Influence de la connectivité multi-échelle via la dispersion larvaire sur la structure des populations et les schémas de biodiversité en mer Méditerranée

Térence Legrand

Aix*Marseille université
Socialement engagée

NNT/NL : 2022AIXM0061/008ED251

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 21 mars 2021 par

Térence Legrand

Influence de la connectivité multi-échelle via la dispersion larvaire sur la structure des populations et les schémas de biodiversité en mer Méditerranée

Discipline
Sciences de l'Environnement

Spécialité
Océanographie

École doctorale
ED 251 SCIENCES DE L'ENVIRONNEMENT

Laboratoire/Partenaires de recherche
Institut Méditerranéen d'Océanologie - MIO

Composition du jury

Dr Frédérique VIARD CNRS - ISEM	Rapporteure
Dr Christophe LETT IRD - MARBEC	Rapporteur
Dr Sakina-Dorothée AYATA Sorbonne Université - LOCEAN	Examinatrice
Pr Eric THIEBAUT Sorbonne Université - SB Roscoff	Examinateur
Pr Jean-Christophe POGGIALE AMU - MIO	Président du jury
Dr Anne CHENUIL CNRS - IMBE	Directrice de thèse
Dr Vincent ROSSI CNRS - MIO	Directeur de thèse

mio Institut Méditerranéen d'Océanologie

Institut Pythéas
Observatoire des Sciences de l'Univers
Aix*Marseille Université

Affidavit

Je soussigné, Térence Legrand, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Anne Chenuil et Vincent Rossi, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille le 13 janvier 2022

Liste de publications et participation aux conférences

Liste des publications réalisées au cours de la période de thèse :

1. A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish. Legrand, T., Di Franco, A., Ser-Giacomi, E., Caló, A., Rossi, V. **2019**. *Marine environmental research*, 151, 104761. https://doi.org/10.1016/j.marenvres.2019.104761
2. Small-scale connectivity of coralligenous habitats : insights from a modelling approach within a semi-opened Mediterranean bay. Rossi, V., Lo, M., Legrand, T., Ser-Giacomi, E., de Jode, A., Thierry De Ville D'avray, L., Pairaud, I., Faure, V., Fraysse, M., Pinazo, C., Chenuil, A. **2020**. *Vie Et Milieu-life And Environment*, 70(3-4), 161-174. https://archimer.ifremer.fr/doc/00680/79225/81737.pdf
3. Explicit and implicit network connectivity : Analytical formulation and application to transport processes. Ser-Giacomi, E., Legrand, T., Hernandez-Carrasco, I., Rossi, V. **2021**. *Physical Review E*, 103(4), 042309. https://doi.org/10.1103/PhysRevE.103.042309
4. Coalescent connectivity through multi-generation dispersal modelling predicts gene flow across marine phyla. Legrand, T., Chenuil, A., Ser-Giacomi, E., Arnaud-Haond, S., Bierne, N., Rossi, V. **In review**. *Nature Communication*.
5. Insights into the spatio-temporal variability of spawning in a territorial coastal fish by combining observations, modelling and literature review. Di Stefano, M., Legrand, T., Di Franco, A., Nerini, D., Rossi, V. **In review**. *Fisheries Oceanography*.

Participation aux conférences au cours de la période de thèse :

1. Integrating Lagrangian modelling with otolith analyses to investigate larval dispersal and fish natal origins. **Poster**. *4th IMarCo conference*, Octobre 2018, Heraklion, Grèce.
2. Rôle de la dispersion larvaire sur la structure des populations et les schémas de biodiversité. **Talk**. *26ième congrès de l'ED 251*, Avril 2019, Marseille, France.
3. A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish. **Talk**. *7th LAPCOD conference.* Juin 2019, Venise, Italie.
4. A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish. **Talk**. *GECOMARS conference*, Fevrier 2020, Marseille, France.
5. Multi-generational connectivity *vs* genetic structures. **Talk**. *5th World Conference on Marine Biodiversity*, Decembre 2020, Auckland, Nouvelle-Zelande.
6. Meta-analysis reveals hidden multi-generational dispersal pathways predicting gene flows across marine taxa. **Talk**. *ASLO Aquatic Science Meeting*, Juin 2021, Palma de Majorque, Espagne.
7. Coalescent connectivity through multi-generation dispersal modelling predicts gene flow across marine phyla. **Talk**. *6th iMarCo Conference*, Decembre 2021, Paris, France.

Résumé

Le monde vivant peut être décrit comme un vaste et complexe réseau de connexions : les organismes, en se déplaçant eux-mêmes dans l'espace et le temps sont les *liens* entre des populations, ou localités. De tels déplacements résultent de la dispersion, soit le mouvement d'un individu depuis son site de naissance jusqu'à son site de reproduction. Les mécanismes de dispersion sont extrêmement variés : ils découlent de la capacité de chaque organisme à se mouvoir par lui-même ou grâce à un tiers. On parle respectivement de dispersion active ou passive. Dans l'océan, la dispersion passive est facilitée par les propriétés physiques intrinsèques de l'environnement marin : de nombreuses espèces benthiques à semi-sédentaires à l'état adulte dispersent ainsi durant leurs premiers stades de vie, en relâchant dans la colonne d'eau des cohortes de propagules (œufs, larves, fruits, etc.) qui sont ensuite transportées par les courants marins. La connectivité entre populations est réalisée lorsque les propagules survivent à une telle phase pélagique puis s'installent sur un habitat favorable pour s'y reproduire. Démographiquement, la connectivité est constitutive de la dynamique spatiale des populations : c'est un processus écologique clef dans la régulation et la persistance des populations. Génétiquement, la connectivité est vectrice de flux génique, qui, avec les autres forces évolutives, module les schémas contemporains de biodiversité. La connectivité via la dispersion est ainsi déterminante dans la résilience des populations face aux pressions anthropiques : sa compréhension est essentielle pour une bonne gestion et sauvegarde des écosystèmes, notamment pour le design d'aires marine protégées (AMPs) ou la gestion des pêches. Dans cette thèse, nous avons défini un *cadre analytique* qui permet de caractériser la connectivité démographique et de localiser les zones de pontes pour toutes espèces présentant une phase pélagique, en combinant l'utilisation d'un modèle biophysique Lagrangien avec des données d'analyses d'otolithes et biogéographiques. Nous avons d'une part montré que les patrons de connectivité sont expliqués par la variabilité spatio-temporelle de la circulation océanique, et d'autre part quantifié le rôle des AMPs dans l'approvisionnement en propagules des zones non protégées. Nous avons ensuite défini analytiquement les probabilités de connexions génétiques résultant d'évènements successifs et cumulatifs de dispersion, cumulant ainsi les différentes possibilités de

flux de gènes entre populations sur plusieurs générations. Pour un nombre de génération fixé, la connectivité filiale, qui quantifie la probabilité qu'une population soit parente à une autre, a été distingué de la connectivité coalescente, qui quantifie la probabilité sous-jacente que deux populations partagent des « populations ancêtres » communes. Nos résultats montrent que les barrières hydrodynamiques, jusqu'ici considérées comme une cause de structuration génétique, sont effectivement perméables à la connectivité coalescente. Ces modèles de connectivité filiale et coalescente ont permis d'estimer le flux de gènes chez 47 espèces à cycle de vie biphasique, compilées dans une méta-analyse couvrant 58 études de génétique des populations en mer Méditerranée. La connectivité coalescente nouvellement définie retourne les meilleures prédictions de flux de gènes pour l'ensemble des espèces et explique environ 50 % de la variabilité des différentiations génétiques observées à l'échelle de la méta-analyse. De plus, nos résultats suggèrent une relation étroite entre les échelles temporelles (i.e. nombre d'évènements de dispersion) et spatiales (i.e. étendu des patrons de diversité génétique) de la connectivité génétique, illustrant les interactions éco-évolutives caractéristiques d'un tel processus. Dans un dernier temps, nous avons étudié l'influence de la considération précise de l'habitat obtenue par modèles de niche, dans la prédiction du flux de gènes. Cette thèse, au-delà d'établir de nouveaux outils méthodologiques pour une meilleure compréhension de la connectivité démographique et génétique, identifie de nouvelles pistes de recherche qui contribueront à mieux évaluer l'impact du flux de gènes et des autres forces évolutives sur la biodiversité marine.

Mots clés : Recherche multidisciplinaire, Modèles biophysiques Lagrangiens, Dynamiques éco-évolutives, Théorie des graphes, Biogéographie, Mer Méditerranée

Abstract

The living world can be viewed as a wide and complex network of connections. By moving over space and time, organisms themselves establish *links* between populations or localities. Movements of an individual from its birthplace to its site of reproduction is called dispersal. This is a multifaceted mechanism: we distinguish active dispersal due to the individual's own locomotion from passive dispersal induced by a third party. In the sea world, where passive dispersal is favoured thanks to the physical properties of the marine environment, most benthic and sedentary species disperse through marine currents during their early-life stages by means of drifting propagules (eggs, larvae, fruits, etc.). Population connectivity is effective when an organism survives its pelagic phase, settles, and eventually reproduces. Demographically, connectivity governs population dynamics and is thus a key ecological process for recruitment and persistence. Genetically, connectivity drives gene flow between populations, which, in interaction with the other evolutionary forces (i.e. drift, mutation, natural selection), shapes the contemporary biodiversity patterns. Hence, connectivity via dispersal processes plays a fundamental role in the resilience of populations to anthropogenic stressors: appraising connectivity is essential for sound management and conservation initiatives, such as the design of Marine Protected Areas (MPAs) and effective fishery management. In this thesis, we build a multidisciplinary *analytical framework* to characterise spatial patterns of demographic connectivity and to pinpoint spawning areas for any species with a pelagic phase, by combining Lagrangian modelling, network theory, otolith analyses and biogeographical information. We show that connectivity patterns are well-explained by the spatio-temporal variability of ocean currents, and we quantify larval export from MPAs to surrounding unprotected areas. We then formulate and model the probability of genetic connections arising from successive dispersal events, depicting gene flow between populations over multiple generations in a cumulative manner. We define filial connectivity, which unveils explicit parents-to-offspring links and, for the first time, coalescent connectivity, which assesses genetic cohesiveness among

populations that share a common ancestor. Our results suggest that the seascape is more open than previously thought since its physical barriers supposedly leading to genetic structuring appear indeed permeable to coalescent connectivity. Filial and coalescent connectivity models allow predicting gene flow over the most extensive meta-analysis of population genetic studies in the Mediterranean Sea, encompassing 58 phylogenetically divergent sedentary species such as corals, endangered territorial fishes, and endemic seagrass. We show that our novel coalescent dispersal model outperforms all previous approaches when simulating gene flow and explains almost 50 % of observed genetic differentiation at the meta-analysis scale. Our results suggest that the temporal (i.e. the number of generations predicting best observed gene flow) and spatial (i.e. extend of observed genetic patterns) scales implied in genetic methods are tightly linked, hence shedding lights on the typical eco-evolutionary scales of genetic connectivity. Finally, we use species' distribution modelling outputs to test the relative importance of habitat heterogeneity in multi-generation dispersal modelling and investigate its impacts on gene flow predictions. Beyond developing new methodological tools to improve our knowledge about demographic and genetic connectivity, this thesis identifies novel research directions that should contribute to better assess the respective influences of gene flow and the other evolutionary forces on biodiversity. For conservation biologists, the results developed in this thesis might help improving the reliability of management plans.

Keywords: Multidisciplinary research, Lagrangian biophysical models, Eco-Evolutionary dynamics, Graph theory, Biogeography, Mediterranean sea

Remerciements

Mes premiers remerciements s'adressent à mes deux encadrants, Anne et Vincent qui, par leur soutien à toute épreuve, leur investissement, leur bienveillance, leur intégrité, ont tout mis en œuvre pour que cette thèse se déroule de la meilleure des manières. Vos conseils et les discussions que nous avons eues m'ont permis d'améliorer mon travail et mes compétences scientifiques, mais également de croire en mes idées et tout simplement de prendre confiance en moi. J'ai beaucoup appris à vos côtés, tant sur le plan scientifique que sur le plan humain. Merci de m'avoir introduit au monde de la recherche de cette manière.

Je remercie également les membres du jury d'avoir accepté d'évaluer mon travail ainsi qu'Audrey Darnaude et Mathias Gauduchon d'avoir accepté de participer à mes comités de thèse.

Merci à Enrico, Antonio, Sophie et Nicolas pour les nombreux échanges stimulants, les discussions et retours pertinents sur mon travail.

Merci aux enseignants-chercheurs de la Licence SVT et du Master OPB de l'université d'Aix-Marseille pour la transmission de savoir durant ces cinq années sur les bancs des amphis.

Merci à tous les jeunes camarades du MIO, arpenteurs eux aussi du grand couloir immaculé dans une recherche existentielle, pour les bons moments passés ensemble à la pause du midi, autour d'un café ou d'une petite bière sous la douce lumière de fin d'après-midi de Luminy. Un merci tout particulier à Claire pour le partage ~~du placard à ba~~ du bureau, à Madiop et Veronica, aka "the elephant team", pour les discussions et les progrès communs faits en anglais, et également à Marine pour toutes les idées qu'on a pu échanger pendant son stage.

Merci à Camille, Auré, Julie, Quentin, Jade, Claire, Alexis, Mathilde et Juliette, soit la belle bande des Océanautes pour tous les moments grandioses qu'on a passé ensemble, depuis le road trip fantastique jusqu'aux skypes confinés et fatigués. Ce n'est que le début des aventures!

Merci à Charley, Pierrot et Jojo d'avoir toujours été là pour moi, depuis le plus fort

de la tempête jusqu'au plus grand des soleils. Merci pour tous ces moments de fête, de rigolade, de partage musical à la *pointe*, que ça soit à Cholière ou aux Fontanettes, mais également à la ville (prononcer avec l'accent savoyard).

Merci aux quatre faces Nord de Pralognan de fournir un si bon refuge contre le monde extérieur : merci pour les couloirs secrets, pour les montées drues dans l'pentu qui tirent sur les mollets, pour les deux heures de soleil en décembre qui réchauffent le corps et pour les chanterelles qui poussent en grand nombre dans la forêt au-dessus de "j[[]hf||bZZ@@JHBL""jhb||eHZV* @[|@]]

Merci à ma famille savoyarde (enfin du Villard) et bretonne (enfin de Bressmeum') de m'avoir soutenu tout au long de mes études, pour leur gentillesse et évidemment les bons repas !

Merci évidemment à mes parents d'être simplement les meilleurs parents, d'avoir toujours été là pour moi et de m'avoir offert la plus grande liberté dans mes choix pendant ces longues années d'étude qui se terminent enfin

Merci à ma sœur de m'avoir supporté jusque-là, pour te dire franchement, pas le temps d'éplucher des kiwis pendant la fin de thèse !

Mes derniers remerciements vont à Juliette pour tous ces moments qu'on partage ensemble depuis un paquet d'année maintenant, pour ton aide, tes relectures, ton soutien moral tellement important (surtout en novembre), mais également pour ton rire, ton enthousiasme, ta panoplie d'animaux, ta haute résistance aux blagues nulles et à la démarche chaloupée de Jean-Luc Smith. Merci de m'avoir fait partager ta passion, découvrir les beautés du monde subaquatique restera une grande aventure pour un petit savoyard comme moi !

Table des matières

1. Introduction

Une chose peut voyager partout
rien qu'en restant immobile
Richard Powers, *l'Arbre-Monde*

1.1. De la connectivité à la dispersion, une histoire de liens et de mouvements

1.1.1. La connectivité : contexte global et définitions

1.1.1.1. Définitions générales

En 2022, et comme depuis un siècle, la science s'écrit en anglais. Il en résulte parfois des difficultés pour certains jeunes produits de l'école républicaine Française à communiquer efficacement au sein de la communauté scientifique. Aussi, du fait de traductions techniques difficiles, de nombreux anglicismes persistent lorsqu'il s'agit d'écrire dans sa langue maternelle, notamment dans le cadre d'une thèse. Ce travail n'échappe pas à la règle : ici, le terme connectivité, mot-clef de cette thèse, est un anglicisme directement tiré de *connectivity* (*the ability of systems, platforms and applications to be connected to each other, Oxford dictionaries* en ligne, s. d.). Bien qu'il soit largement employé et compris instinctivement, le mot « connectivité » n'existe pas, à proprement parler, dans la langue française. Pour bien le définir, il faut chercher des mots étymologiquement proches, partageant la racine latine *connexĭo* (lien, enchainement) : *connexion*, « action de lier par des rapports étroits ; fait d'être lié ; liaison, enchaînement » ; *connexité*, « rapport étroit qui existe entre deux ou plusieurs choses ». Ou encore le verbe *connecter* défini par « unir des choses en les mettant en relation entre elles », (les définitions proviennent du Larousse en ligne, s. d.). Ainsi, ce travail de thèse s'intéresse à la façon dont les connexions s'établissent de manière globale entre deux ou plusieurs entités et aux conséquences de tels liens.

Ça serait enfoncer une porte déjà grande ouverte que de dire que le monde n'a jamais été aussi connecté qu'aujourd'hui : notre époque mondialisée a lié chaque espace du globe par des relations économiques, les réseaux sociaux permettent à toute personne de rentrer en relation avec une autre au moyen d'un « tissu de solitudes reliées » (Damasio, 2019), ou encore l'explosion exponentielle du flux d'informations et de données lie notre expérience du monde avec le monde lui-même comme jamais auparavant. Même si un tel niveau d'interconnexions mondiales a permis d'augmenter notre niveau de vie (surtout pour les habitants des pays industrialisés), cela questionne aussi notre rapport au monde dans un contexte de changement global (e.g. tourisme de masse, échanges de matières premières et de matériaux transformés, coût de stockage des données numériques, etc.). La pandémie actuelle souligne, peut-être, les limites de notre époque interconnectée : il faut justement casser ces liens et s'isoler pour lutter efficacement contre les vagues épidemiques successives de COVID-19. Cela nous oblige à modifier notre rapport direct aux autres, notre *connectivité sociale*, pour éviter de mettre en péril l'économie mondiale. Il en est de même pour le monde numérique, dont paradoxalement, il faut parfois se déconnecter pour mieux se reconnecter (à soi-même ? au vivant ?). Il en reste que la connectivité caractérise et questionne notre rapport au monde, qu'il soit économique, social, numérique ou vivant.

1.1.1.2. Comment représenter la connectivité ?

Les connexions entre deux ou plusieurs entités peuvent être définies de manière quantitative (e.g. nombre de voyageurs entre deux villes) ou qualitative (e.g. relation sociale entre deux personnes). Dans le premier cas, on peut alors quantifier les connexions par un flux ou une probabilité d'échange. Pour le deuxième cas, les connexions sont plutôt binaires : il y a présence ou absence de lien. On peut toutefois y ajouter parfois un degré qualitatif (e.g. ces personnes se connaissent peu et celles-là très bien). Le nombre de connexions possibles est fonction du carré du nombre d'entités considérés : il en résulte des structures complexes, qui peuvent être résumées par des objets mathématiques appelé graphes (Bondy et Murty, 1976). Dans un graphe, les entités sont représentées par des *nœuds* et les connexions par des *liens*. Lorsque les *liens* sont asymétriques, c'est notamment le cas pour des connexions quantitatives, on parle de graphes orientés. De manière opposée, les graphes non orientés sont caractérisés par un unique lien symétrique entre deux nœuds, typique des connexions qualitatives. Lorsque les *nœuds* et *liens* d'un graphe peuvent être appliqués à une

situation réelle, ou du moins lorsqu'on s'affranchit de l'abstraction mathématique, on peut utiliser le terme « réseaux » (Figure 1.1). Ainsi, la science des réseaux applique la théorie des graphes dans diverses disciplines : la physique, l'économie, la climatologie, les sciences sociales, la biologie, etc.

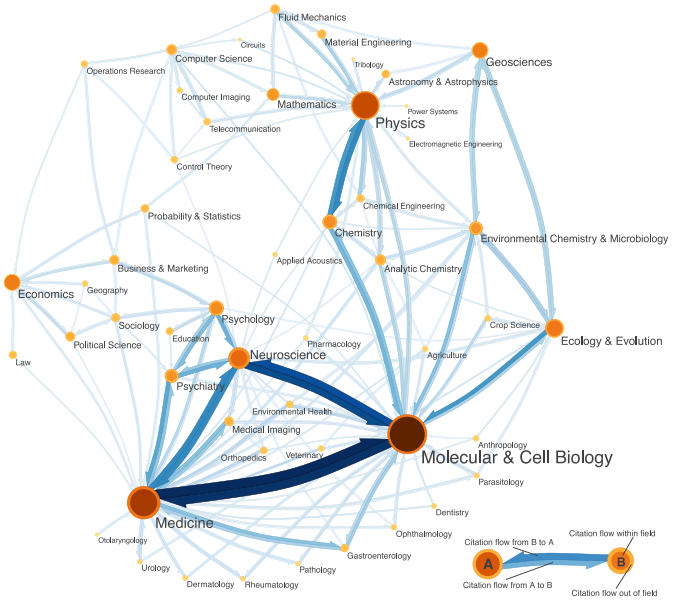


FIGURE 1.1. – Réseau illustrant les patrons de citations entre disciplines scientifiques (Rosvall et Bergstrom, 2008). Ce réseau a été construit en utilisant des informations provenant de 6 128 journaux scientifiques, connectés entre eux par 6 434 916 citations. Cette thèse multidisciplinaire lie au sens large les disciplines d'« Écologie et évolution », de « Biologie moléculaire », de « Physique », de « Géoscience » et, dans une moindre mesure de « Mécanique des fluides ».

1.1.2. La connectivité dans le monde du vivant

Le monde vivant peut être décrit comme un vaste réseau d'interactions entre organismes sur différents niveaux organisationnels (e.g. espèce, population, communauté, écosystème, biomes). En ce qui concerne le vivant « hors-humain », les éthologues ont pu mettre en évidence différentes structures sociales en étudiant les relations entre individus ou catégories d'individus, et ce chez différents phylum (e.g. structure sociale des grands singes ou au sein des colonies de fourmis). De plus, de nombreuses interactions biologiques entres espèces existent au sein d'un même écosystème. Les *liens* d'un réseau peuvent alors définir différents types de symbioses en fonction des effets combinés de telles relations sur les paires de *nœuds*, c.a.d réciproquement entre

individus ou espèces. Par exemple, une relation compétitive va être nuisible pour les deux parties, alors qu'une relation de commensalisme profite à une seule espèce sans nuisance ni bénéfice pour l'autre (Figure 1.2). Dans un écosystème relativement stable, la sélection naturelle favorise le plus souvent les interactions à bénéfice réciproque : les interactions néfastes ont tendance à être réduites ou éliminées car elles entrainent à terme l'extinction des populations concernées (Odum et Barrett, 2005).

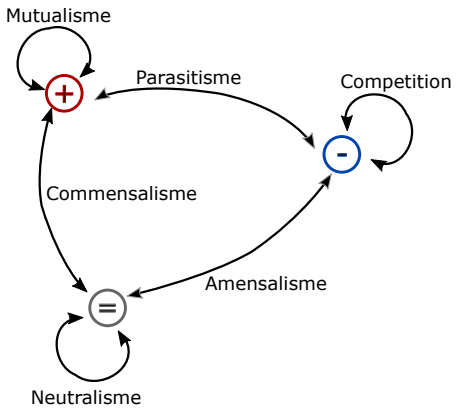


FIGURE 1.2. – Schéma représentant les six interactions biologiques principales. Dans ce réseau, les nœuds représentent l'effet des interactions sur l'individu (le symbole + pour bénéfice, - pour nuisance et = pour sans effet).

A court terme, la prédation a le même effet que le parasitisme : le prédateur bénéficie directement de sa prédation sur la proie, qui comme on peut l'imaginer, ressent comme une nuisance le fait d'être mangée. Au sein d'un écosystème, les relations prédateurs-proies entre espèces peuvent être résumées par un réseau trophique (Pimm et al., 1991 ; Pringle et Hutchinson, 2020). Au sein d'un tel réseau, les espèces représentent les *nœuds*, et les relations de prédation entre espèce les *liens* (Figure 1.3). Ainsi, tous les organismes présents dans un écosystème sont plus ou moins interdépendants : chaque interaction étant caractérisée par sa durée, sa propriété bénéfique ou nuisible ou son degré d'association des organismes impliqués.

La connectivité du vivant s'étend au-delà des relations entre individus. Les organismes, en se déplaçant dans l'espace et le temps, vont être eux même les connecteurs, les *liens*, entre des populations. Le concept de population peut être défini de différentes manières, principalement en fonction du processus écologique ou évolutif auquel on s'intéresse (Waples et Gaggiotti, 2006). Ainsi, aucune définition consensuelle reliant ces deux paradigmes n'émerge dans la littérature. Toutefois, on peut caractériser les populations (i) d'une perspective écologique comme un ensemble d'in-

dividus qui interagissent entre eux et peuvent être identifiés à une aire géographique spécifique et (ii) d'une perspective évolutive comme un ensemble d'individus qui se reproduisent entre eux et qui co-existent dans le temps et l'espace (voir Tableau 1 dans Waples et Gaggiotti, 2006 pour une compilation des différentes définitions dans la littérature). Les déplacements d'individus ont un impact important sur la dynamique spatiale et temporelle des populations et de l'écosystème qu'elles constituent. De tels mouvements induisent des flux d'énergie, de biomasses, de nutriments (Zuercher et Galloway, 2019) ou de carbone (Hyndes et al., 2013, Gounand et al., 2018) entre patchs d'habitat (c.à.d. que les structures des réseaux trophiques vont être dépendantes des mouvements individuels de chaque espèce, donc pour chaque nœud les constituant). Ils induisent également des modifications démographiques et de traits fonctionnels au sein de chaque population, ainsi que des échanges de gènes entre populations (Kool et al., 2013). Le concept de connectivité fonctionnelle intègre toutes ces différentes composantes et pourrait se définir comme le transfert des principales fonctions écologiques (résultant des interactions inter- et intra-espèces) entre écosystèmes ou biomes (e.g. Gerber et al., 2014 ; Selkoe et al., 2016 ; Turgeon et al., 2010).

1.1.3. La dispersion comme vecteur de la connectivité

1.1.3.1. Définition

Le vivant se déplace dans le temps et l'espace au travers différents mécanismes : dispersion, migration, exploration, recherche de nourriture, etc. D'après la définition de Clobert et al., 2012, seule la dispersion permet de maintenir efficacement une ou plusieurs connexions entre des populations. En effet, la dispersion est définie comme le mouvement d'un individu depuis son site de naissance jusqu'à son site de reproduction, soit « tout mouvement amenant à du flux de gènes spatial ». En d'autres termes, tout mouvement amenant au flux de gènes entre populations. Les auteurs ajoutent qu'il n'y a pas besoin que la reproduction s'additionne au mouvement, dans le sens où un individu déplace ses propres gènes dans l'espace. Dans la littérature, la dispersion peut être mentionnée comme migration, bien que la migration fait référence également à des mouvements réguliers et cycliques dans l'espace, n'entrainant pas forcément du flux de gènes (comme les migrations saisonnières ou journalières, Clobert et al., 2012). Par exemple, les sternes arctiques migrent de plusieurs milliers de kilomètres vers le sud pour se nourrir, avant de retourner dans leurs zones de reproduction et nidification (Egevang et al., 2010). En transférant de la matière organique

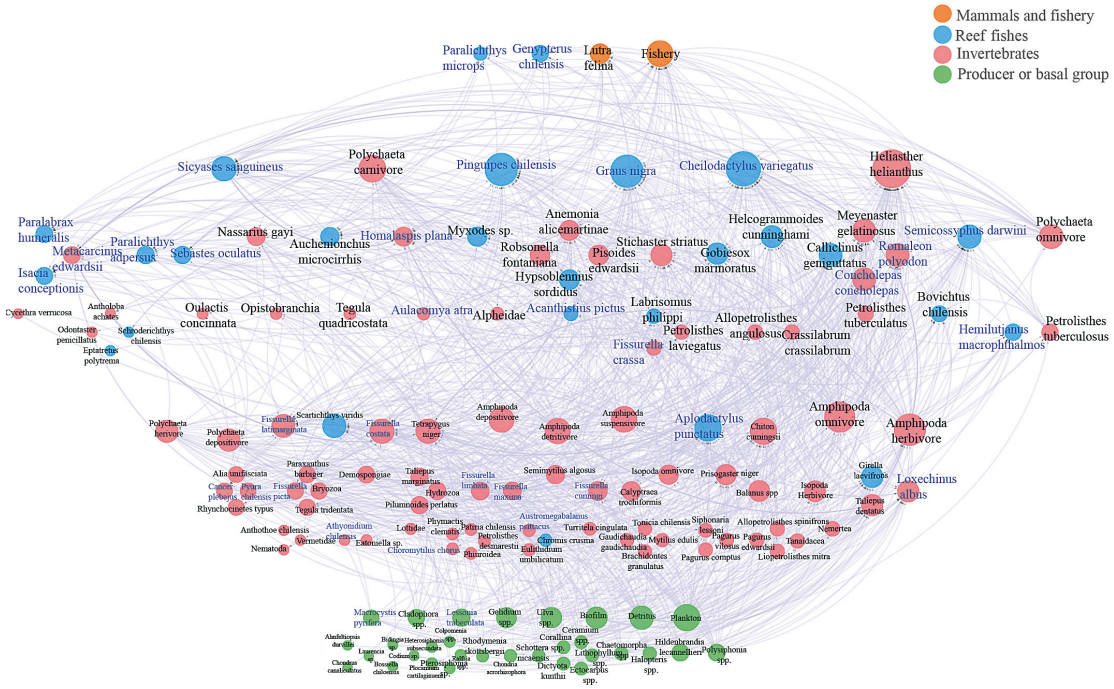


FIGURE 1.3. – Exemple d'un réseau trophique : l'écosystème subtidal des côtes centrales chilienne (Pérez-Matus et al., 2017). Dans ce réseau, 147 taxons sont représentés par des nœuds, et leurs positions verticales dépend du niveau trophique. La taille des nœuds indique le nombre de liens joins à celui-ci : c'est-à-dire le nombre de relation de prédation associé à ce taxon.

(e.g. carbone, azote, etc.) dans l'espace, les sternes arctiques favorisent ainsi la connectivité fonctionnelle (cf section 1.1.2). Les mécanismes amenant à la dispersion sont extrêmement variés. Ils dépendent de la capacité de chaque organisme à se mouvoir par lui-même ou grâce à un tiers (c.f. section 1.1.3.1 et 1.1.3.2, respectivement) et peuvent induire des déplacements directs du site de naissance au site de reproduction jusqu'à des comportements complexes d'exploration et de relations sociales (voir Figure 1.4, Matthysen, 2012). De manière synthétique, le processus de dispersion se découpe en trois étapes successives : émigration depuis le site de naissance, transfert ou déplacement dans l'espace, immigration dans le site de reproduction (Clobert et al., 2012).

La dispersion est donc un mécanisme qui éloigne les individus de leurs sites de naissance. De ce fait, la littérature s'est intéressée à comprendre le réel bénéfice de la dispersion, par rapport au coût potentiel de tels mouvements (Bonte et al., 2012; Burgess et al., 2016; Matthysen, 2012; Pechenik, 1999). Les principaux avantages écologiques et évolutifs à la dispersion sont :

— De réduire les potentielles interactions avec les congénères. Cela permet d'éviter la compétition avec ses parents proches tout en minimisant le risque de dépression de consanguinité.
— D'augmenter la valeur sélective (*fitness*) attendue en confrontant la descendance à un maximum d'environnement variables dans l'espace et le temps. En d'autres termes, d'échelonner les conditions de survie puis de reproduction sur différents contextes environnementaux (stratégie de minimisation des risques ou de « ne pas mettre tous ses œufs dans le même panier »).
— D'échapper à des conditions locales défavorables comme la surpopulation, l'éventuelle absence de partenaires viables ou des pressions externes telles qu'anthropiques.

Parallèlement, en plus du coût stricto sensu sensus-stricto dû au mouvement (dépense énergétique et physiologique, ainsi que le développement de structures de locomotions permettant de se déplacer), les inconvénients de la dispersion sont nombreux : perte de coopération avec les congénères, risque de dépression hybride et possibilité de se déplacer dans un habitat moins favorable, i.e. augmentation de la prédation et réduction des ressources en nourriture (Bonte et al., 2012; Matthysen, 2012). Cependant, ces coûts peuvent être différents entre les milieux terrestre et marin, notamment à cause des propriétés physiques du médium considéré (i.e. air, terre ou eau, voir section 1.1.3.3, Burgess et al., 2016).

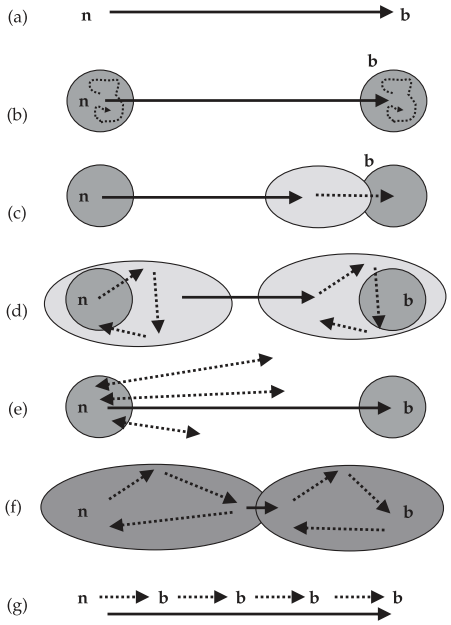


FIGURE 1.4. – Schéma représentant les différents mécanismes amenant à la dispersion des individus, soit au flux de gènes entre des populations distinctes dans l'espace. Ici, n représente le site de naissance (pour *natal site*) et b le site de reproduction (pour *breeding site*). Les flèches pleines représentent le processus de dispersion alors que les flèches en pointillés indiquent d'autre type de mouvements (migration, exploration, etc.). **a** La dispersion est induite par un seul mouvement entre le site de naissance et de reproduction (e.g. dispersion des espèces sessiles). **b** Un individu se déplaçant dans son aire vitale de naissance (représenté par le rond gris foncé) avant de disperser et d'établir son aire vitale de reproduction. **c** L'individu disperse d'abord dans une aire vitale où il ne se reproduit pas (représenté par le rond gris clair), d'où il se déplace pour se reproduire. **d** L'individu se meut dans une zone vitale étendue, qui inclue une zone natale réduite d'où il disperse vers une autre zone vitale étendue, qu'il diminue pour se reproduire. **e** L'individu explore son environnement en dehors de sa zone vitale avant de disperser. **f** L'individu disperse de son groupe social de naissance vers un groupe social voisin. **g** L'individu se déplace tout au long de sa vie et se reproduit de manière intermittente dans des zones de reproduction spécifiques. Schéma et légende tirés de Matthysen, 2012.

1.1.3.2. La dispersion active

Parmi tous les mécanismes de dispersion, la dispersion est dite active lorsque l'individu se déplace dans l'espace grâce à ses propres moyens : il présente alors un comportement orienté de marche, de nage ou de vol. La dispersion active est caractéristique des animaux juvéniles et adultes. Son intensité varie entre les capacités de mouvement de chaque espèce et dépend de plusieurs facteurs, comme la taille de la population locale, la compétition pour les ressources ou la qualité et la taille de l'aire vitale. Cela peut aussi dépendre de la structure sociale spécifique à l'espèce : un système social qui dépend d'un seul mâle ou femelle dominant adulte (i.e. un système de reproduction en harem observable chez certains mammifères, oiseaux ou poissons) force les juvéniles à disperser pour pouvoir se reproduire (Figure 1.4f). Bien que la migration n'implique pas toujours de la dispersion, la méta-analyse de Hays et Scott, 2013 suggère que plus l'individu est lourd, plus il se déplace et peut disperser sur de grandes distances (Figure 1.5). C'est d'autant plus le cas pour les marcheurs. Les espèces qui se déplacent en volant dans l'atmosphère sont celles qui se déplacent le plus loin, suivies des espèces nageant dans l'océan puis celles marchant sur le paysage terrestre. Cela peut s'expliquer par le fait que les barrières physiques au mouvement (e.g. fleuves, montagnes pour le terrestre, détroits et fronts pour le milieu marin) sont moins marquées voire absentes dans l'atmosphère. Les distances peuvent aller de quelques km à près de 10 000 km.

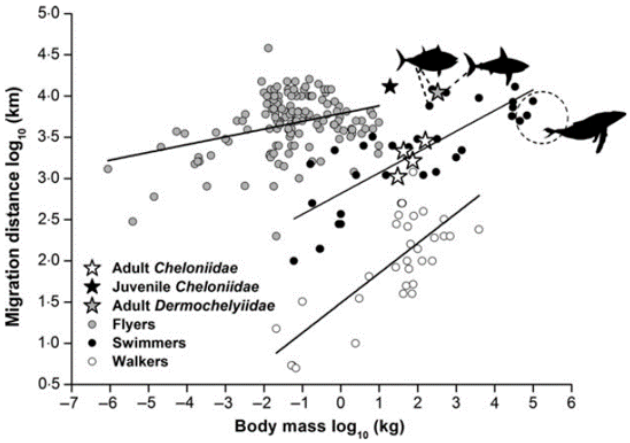


FIGURE 1.5. – Distance de migrations en fonction de la masse corporelle pour les espèces se mouvant en marchant, nageant et volant. Adapté depuis Hays et Scott, 2013, depuis des données de Hein et al., 2012

Il est intéressant de noter que les végétaux peuvent présenter une sorte de dispersion active (i.e. déplacement endogène mais non orienté). Un exemple non exotique est celui du concombre d'âne, ou cornichon sauteur, *Ecballium elaterium*. Cette espèce est une des rares cucurbitacées spontanées en Europe (un individu est présent aux alentours du MIO, Figure 1.6a). Elle disperse en projetant ses graines jusqu'à plusieurs mètres à une vitesse de 10 $m.s^{-1}$ par l'explosion de son fruit préalablement pressurisé à environ 6 bars (Forterre et al., 2016).

1.1.3.3. La dispersion passive

En opposition à la dispersion active, un individu disperse passivement lorsque son déplacement n'est pas induit explicitement par ses capacités de mouvement, réduisant le contrôle de sa trajectoire dans l'espace mais en augmentant possiblement les distances parcourues. Ne dépensant ainsi pas ou peu d'énergie pour se déplacer, le processus de dispersion passive est réalisé pour ces individus, par l'intermédiaire d'un tiers biotique ou abiotique.

Pour les plantes, la dispersion passive biotique est souvent induite par la consommation des fruits par d'autres animaux capables de mouvements, i.e. par *endozoochorie*, et donc dépend d'une stratégie évolutive. Un exemple intéressant parmi beaucoup d'autres est celui du casse-noix moucheté (*Nucifraga caryocatactes*) qui se nourrit des graines de l'arole (*Pinus cembra*), tout en les enfouissant dans le sol pour ses réserves hivernales (Figure 1.6b). En oubliant près de trois quarts de ses cachettes, il assure quasi-exclusivement la dispersion de l'arole, et favorise ainsi sa régénération sur d'anciennes pâtures (Camaret et al., 1998). Cette interaction biologique peut être qualifiée de mutualiste puisqu'elle bénéficie aux deux espèces : la connectivité de l'arole est en effet induite par sa connexion avec le casse-noix moucheté, illustrant l'organisation du vivant en une multitude de réseaux (voir section 1.1.2 et Figure 1.2). Les autres aspects de la dispersion passive biotique relèvent plutôt d'interactions de commensalisme ou d'amensalisme : le processus de dispersion n'a pas d'impact pour l'organisme vecteur, et peut être bénéfique ou non chez l'organisme qui disperse. Pour les plantes, on parle d'*épizoochorie* : le fruit s'attache aux animaux par le biais de mucus collant, de piquants ou de crochets, comme les bardanes par exemple. On retrouve aussi ce système chez certains invertébrés marins comme les bivalves qui se « ferment » sur les pattes des oiseaux afin de disperser sur de longues distances (Figure 1.6c, A. Green et Figuerola, 2005). On peut inclure dans cette catégorie la dispersion d'individus par les eaux de ballast, en considérant les navires comme un vecteur biotique, ou du moins

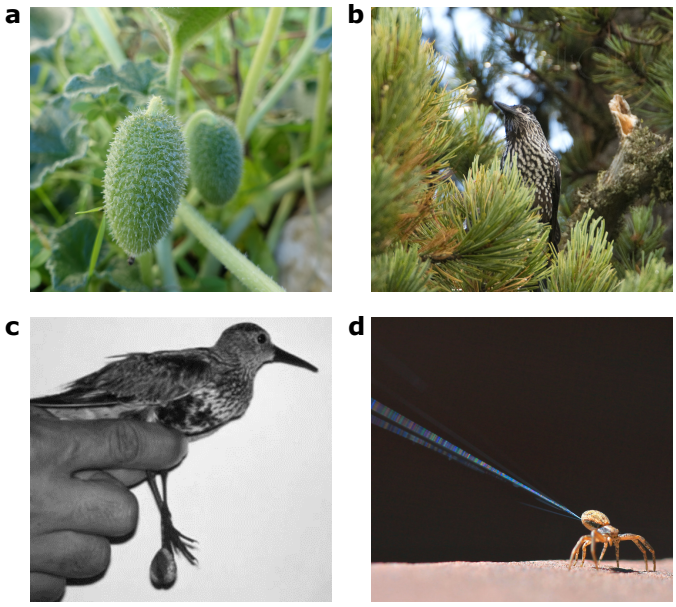


FIGURE 1.6. – Illustrations de quelques exemples de mécanismes de dispersion. **a** Le cornichon sauteur, qui disperse ses graines par l'explosion de son fruit (individu photographié à côté du MIO). **b** L'arole qui disperse grâce au casse noix moucheté (crédit photographique : achrntatrps depuis Flickr). **c** La coque commune qui peut disperser en s'accrochant à la patte d'un Bécasseau variable (A. Green et Figuerola, 2005). **d** Araignée éjectant ses fils de soie, qui vont augmenter sa portance afin d'être dispersée par le vent (crédit photographique : Jeff Mitton depuis Flickr).

anthropogénique. Au-delà d'induire du flux de gènes, le transport d'organismes par les eaux de ballast a causé l'introduction d'espèces non-indigènes (Saebi et al., 2020), amenant de nombreux problèmes écologiques et économiques (Gollasch, 2007).

La dispersion passive dite abiotique concerne toutes les espèces qui sont transportées par le vent ou les courants marins, ce qui conduit à considérer une approche méthodologique convergente pour étudier la dispersion passive aérienne et aquatique (Lett et al., 2020). Dans l'environnement terrestre, les plantes sont majoritairement concernées : le vent permet de disperser leurs pollens et graines, sur parfois de longues distances (Nathan et al., 2002). Il y a également toutes les espèces de petite taille qui composent le plancton aérien dispersant avec le vent : des protistes, champignons, microbes, bactéries, etc. (D. Smith, 2013). On dénote aussi certains arthropodes qui utilisent leurs soies pour avoir assez de portance et ainsi être transportés par le vent (*ballooning dispersal*, Figure 1.6d, Bell et al., 2005). De manière générale, ces organismes terrestres présentent des adaptations à la dispersion passive dans l'atmosphère (Burgess et al., 2016). De façon opposée, les adaptations à la dispersion passive via les courants océaniques sont habituellement absentes chez les organismes marins (Burgess et al., 2016). Cela est notamment dû aux différences de propriétés physiques entre le medium « air » et « eau » : différence de viscosité, de masse volumique, de vitesse terminale, de capacité thermique et d'approvisionnement en oxygène (Burgess et al., 2016). En d'autres termes, la dispersion passive demande des « efforts » dans l'atmosphère, alors qu'elle est relativement « facile » dans l'océan.

1.2. La dispersion dans l'océan, ou comment les courants induisent le déplacement des organismes marins

1.2.1. Des organismes marins à cycle de vie biphasique

Dans l'océan, la dispersion passive est omniprésente chez les espèces dites benthiques (sessiles ou démersales). Le processus de dispersion est réalisé pendant les premiers stades de vie : des cohortes d'individus sont alors relâchées dans la colonne d'eau sous forme de propagules (œufs, larves, spores, fruits, etc.) puis transportées par les courants marins. Durant cette phase pélagique, les individus peuvent échapper à la prédation existante dans le milieu benthique et, pour les espèces ne bénéficiant pas de réserves de nutriment pour se développer, bénéficier de la présence du phyto- et zooplancton dans l'océan ouvert pour se nourrir (Pechenik, 1999). La phase pélagique prend fin lorsque les propagules survivent et s'établissent sur un substrat adéquat (par exemple les zones de nourricerie), grâce, ou non, à des capacités intrinsèques de déplacement et d'orientation. De nouveau fixés ou associés à un habitat, les individus adoptent un mode de vie benthique comme leurs parents, réalisant ainsi un cycle de vie biphasique (Figure 1.7). La durée caractéristique de la phase pélagique, nommée *Pelagic Larval Duration* (PLD), dépend des conditions environnementales *in situ* et des caractéristiques biologiques propres à chaque espèce. Elle peut varier de quelques minutes à plusieurs semaines selon les espèces considérées (voir Tableau 1 dans Shanks, 2009). Parallèlement, les individus peuvent disperser pendant cette phase pélagique sur des distances s'échelonnant de quelques mètres à plusieurs centaines de kilomètres (Shanks, 2009). De manière générale et lorsqu'on considère différents phylums, les variations de la PLD expliquent près de 50 % des variations de distances de dispersion : plus la phase pélagique est longue, plus les organismes tendent à disperser loin (Shanks, 2009). Il est intéressant de noter qu'à l'échelle mondiale, la PLD augmente avec la latitude, illustrant le fait que, comme la plupart des processus physiologiques, des températures plus chaudes raccourcissent la phase pélagique (B. S. Green et Fisher, 2004 ; Munday et al., 2009). Parallèlement, en considérant les vitesses et directions des courants, les espèces de moyennes latitudes (de 20°à 40°) tendent à disperser moins loin que les espèces tropicales et de hautes latitudes (Álvarez-Noriega et al., 2020).

Pour ces espèces à cycle de vie biphasique, la dispersion passive induite par les

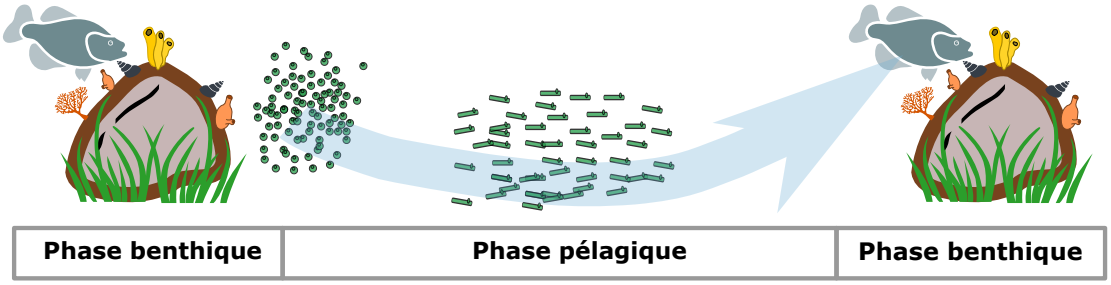


FIGURE 1.7. – Schéma du cycle de vie biphasique caractéristique de nombreux phylum marins. Les organismes dispersent pendant la phase pélagique, lorsque les premiers stades de vie sont advectés par les courant marins. La phase benthique est spécifique de l'état adulte, lorsque les organismes sont fixés au substrat ou sont sédentaires.

courants marins se découpe en trois processus successifs : la ponte ou la remise des propagules dans la colonne d'eau (émigration), le transport pélagique (transfert ou déplacement dans l'espace) et l'installation des propagules dans un habitat favorable (immigration).

1.2.2. Le transport pélagique des propagules par les courants marins

Le transport correspond au déplacement horizontal *sensu-stricto* de propagules entre deux points géographiquement définis dans l'espace (Figure 1.8 encadré bleu, Pineda et al., 2007). Le déplacement dans l'espace est en premier lieu induit par l'advection des propagules par les courants marins, mais il peut également être modulé par les différents comportements larvaires observés chez les espèces à cycle de vie biphasique (Leis, 2006). Même si sa définition se veut simple, l'interaction des différents processus océanographiques associée à la grande variabilité des conditions hydrodynamiques rend le transport larvaire difficile à appréhender et à évaluer.

1.2.2.1. La circulation océanique

Durant la phase pélagique les organismes sont advectés par le déplacement des masses d'eaux, sous l'effet combiné du vent, des forces de marées, des gradients de masse volumique, de la force de Coriolis, des forces de frottement dues à la viscosité et des contraintes géographiques du bassin océanique. Le transport des propagules est ainsi déterminé par les différents processus hydrodynamiques induits par ces forçages

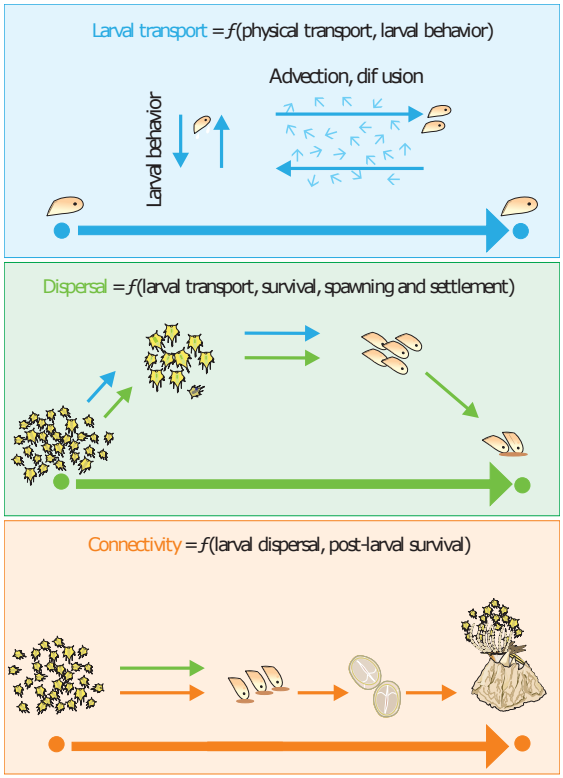


FIGURE 1.8. – Schémas représentant les différents processus imbriqués induisant de la connectivité via la dispersion larvaire dans le milieu marin. Le processus de transport larvaire (encadré bleu) est en premier lieu induit par l'advection (et éventuellement la diffusion, négligée ici) des propagules par les courants marins (section 1.2.2). Au transport larvaire s'additionne la ponte, la survie des propagules pendant le déplacement puis l'installation pour rendre compte du processus de dispersion larvaire (encadré vert, section 1.2.3). La connectivité entre populations à lieu lorsque l'individu, après le processus de dispersion larvaire, survit et s'établit dans la population adulte (encadré orange, section 1.3). D'après Pineda et al., 2007.

(Figure 1.9). Ces processus, constitutifs de la circulation océaniques, sont caractérisés par un continuum d'échelles spatio-temporelles imbriquées, qu'on peut subdiviser par simplification en trois catégories : la sous-méso-échelle (dizaines de m – dizaine de km, horaire à journaliers), la méso-échelle (dizaines de km – centaines de km, dizaines de jours − quelques mois) et la large-échelle (centaines de km – milliers de km, année − plusieurs décennies). Les différentes conditions hydrodynamiques rencontrées par les propagules combinent ainsi des processus turbulents à sous-méso-échelle (e.g. turbulence verticale, ondes internes gravitationnelles, vagues déferlantes, instabilité locale due aux courants de marée et la topographie, etc.) jusqu'aux processus de large échelle (e.g. courants géostrophiques, grandes gyres océaniques stables, ondes de Rossby, oscillations climatiques du type El Niño/Southern Oscillation ou North-Atlantic Oscillations, etc.), en passant par des processus instables à méso-échelle (e.g. fronts séparant des masses d'eau différentes, tourbillons ou *eddies* qui se déplacent, upwelling côtiers, courants de marée, etc.).

Bien que tous les processus cités ci-dessus affectent théoriquement les processus de transport dans l'océan, les échelles spatio-temporelles déterminantes dépendent au premier ordre de la durée de la phase pélagique, mais aussi, dans une moindre mesure, de la zone considérée et des caractéristiques intrinsèques des propagules en question (i.e. taille, densité, etc.). Pour les espèces caractérisées par des PLDs comprises entre quelques minutes et heures (observable chez des invertébrés, comme différents cnidaires, tuniciers ou bryozoaires, Shanks, 2009), le transport est directement induit par les processus turbulents de sous-méso-échelle. Par leurs caractères chaotiques, ces processus sont difficilement mesurables et quantifiables (McWilliams, 2016), et sont souvent mal représentés dans les modèles hydrodynamiques, notamment dans les systèmes côtiers (section 1.2.2.2). Les espèces dispersant pendant des PLDs de plusieurs jours à plusieurs semaines (~ 90 % des espèces pour lesquelles on a des informations biologiques de premiers traits de vie, Shanks, 2009) ce sont les processus de méso− et de large−échelles qui sont déterminant dans le transport des propagules. Les patrons spatiaux de dispersion peuvent ainsi être en premier lieu expliqués par la circulation moyenne induite par les processus océaniques à large-échelle : l'emplacement de la ponte est majoritairement située en amont des courants géostrophiques par rapport au lieu d'installation (voir Chapitre II).

Les courants géostrophiques résultent de l'équilibre entre les forces de pression et la force de Coriolis, et sont spécifiques de la circulation océanique moyenne de surface à large-échelle. Bien qu'ils soient caractérisés par de la variabilité intra-annuelle ou

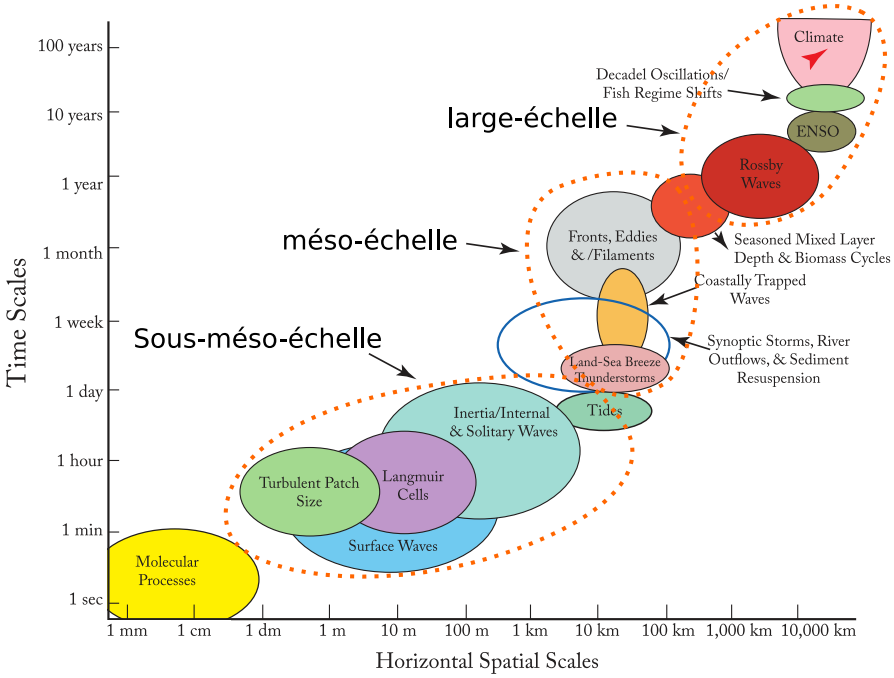


FIGURE 1.9. – Échelles temporelles et spatiales des processus océanographiques. Les processus constituant la circulation océanique peuvent être caractérisés par trois grandess catégories spatio-temporelles : la sous−méso−échelle (dizaines de m – dizaine de km, horaire à journalier), la méso-échelle (dizaines de km – centaines de km, dizaines de jours − quelques mois) et la large-échelle (dizaines de km – centaines de km, dizaines de jours − quelques mois). D'après Dickey et Bidigare, 2005 ; Nichols et Raghukumar, 2020.

saisonnière induite par la variabilité des forçages environnementaux, les courants géostrophiques sont relativement stables dans le temps (large-échelle, Figure 1.9). Cela permet d'identifier et de cartographier les schémas moyens de circulations de surface, grâce notamment à l'utilisation combinée de méthodes d'altimétrie spatiale et de modélisation hydrodynamique, mais aussi de données de flotteurs ARGO (Figure 1.10, Poulain et al., 2012). Ainsi, ce sont ces courants qui sont représentés sur de nombreuses cartes océaniques (e.g. Millot et Taupier-Letage, 2005 ; Pinardi et al., 2015 ; Poulain et al., 2012 pour la Méditerranée) et qui servent bien souvent à interpréter les réseaux de connectivité démographique et génétique (voir section 1.3). La circulation moyenne de surface en Méditerranée est caractérisée par les processus de large échelle, notamment les courants géostrophiques qui présentent des vitesses d'advection importantes jusqu'à 1 $m.s^{-1}$. Par exemple le courant Nord qui s'établit le long de la Méditerranée nord-occidentale ou le courant Algérien depuis le détroit de Gibraltar jusqu'au détroit de Sicile, identifiables également sur la circulation journalière de surface (Figure 1.11). Les structures circulaires relativement stables dans le temps et l'espace sont identifiées comme des gyres à large-échelle (e.g. gyre Nord−Tyrrhénien, gyre Sud−Adriatique ou gyre Est−Alboran, Figure 1.10), induites par des régimes marqués de vent et la topographie (Millot et Taupier-Letage, 2005 ; Poulain et al., 2012). Caractérisés par plus de variabilité, les tourbillons ou *eddies* à méso-échelle sont des structures qui se créent et disparaissent continuellement tout en se déplaçant (Poulain et al., 2012). Elles résultent essentiellement des instabilités des courants géostrophiques (e.g. vortex représentés au nord du courant Algérien, Figure 1.10, identifiables sur la circulation journalière, Figure 1.11). A méso−échelle, la variabilité spatio−temporelle de la circulation océanique rend le transport des propagules difficile à déterminer (e.g $eddies$, front océaniques, courants de marée, upwelling côtiers, etc.). Dans ce cas, les patrons de dispersion peuvent être contraires aux schémas de circulation moyenne, dû à des zones hautement turbulentes (e.g. détroits ou promontoires qui modifient le flux, à l'origine partiellement laminaire des courants moyens) ou des changements rapides dans l'intensité du vent et de son forçage sur la surface (voir Chapitre II).

1.2.2.2. Les modèles hydrodynamiques Eulériens de circulation océanique

Les modèles hydrodynamiques sont largement utilisés pour simuler, comprendre et étudier la circulation océanique et les processus de transport résultant, et ce en englobant, suivant leurs résolutions, les structures à large échelle jusqu'à possiblement

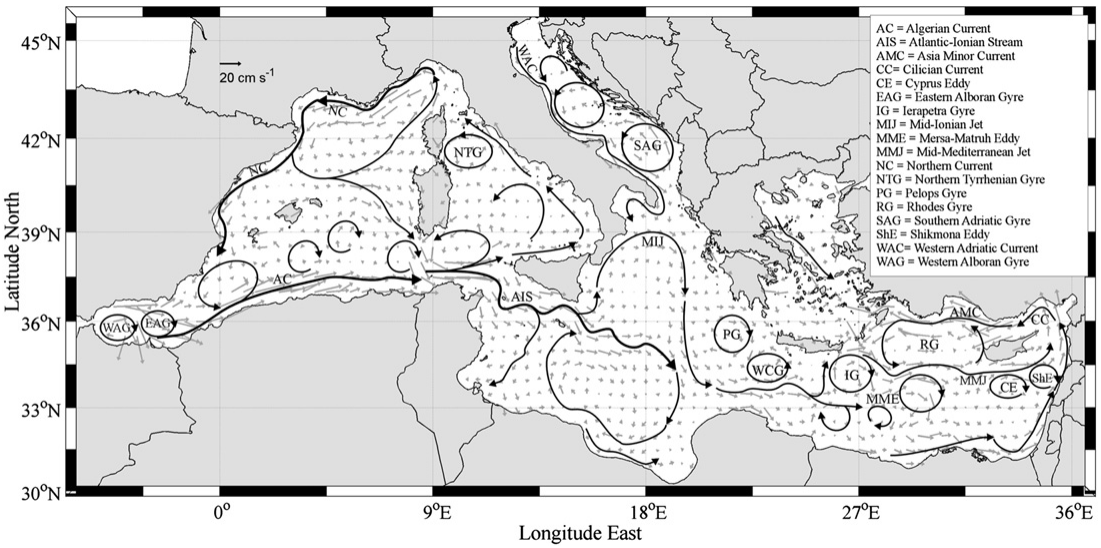


FIGURE 1.10. – Représentation schématique de la circulation moyenne géostrophique de surface de la mer Méditerranée. Cette représentation est obtenue avec des données de flotteurs et de données altimétriques satellitales pour la période 1992-2010 (Poulain et al., 2012, voir également Figure 2 dans Millot et Taupier-Letage, 2005 ou Figure 12 dans Pinardi et al., 2015.

les processus turbulents à sous-méso échelle (un domaine en cours de développement). Ils se basent sur la résolution des équations dites primitives : les équations de Navier-Stokes et leurs équations de continuité, les équations de conservation pour la température et la salinité et l'équation d'état de l'eau de mer de l'UNESCO. Les résolutions des équations sont basées sur plusieurs approximations et hypothèses, comme l'approximation hydrostatique, l'approximation de Boussinesq ou la fermeture Newtonienne de Boussinesq. Les équations primitives sont discrétisées temporellement et spatialement sur des schémas à différences finies, le plus souvent homogènes sur l'horizontale (e.g. grille Arakawa-C représentée sur la Figure 1.12) mais qui peuvent être également hétérogènes et déstructurés (e.g. le modèle SLIM qui utilise les méthodes de Galerkine discontinues, utilisé par exemple dans Dobbelaere et al., 2020). Sur la verticale, on distingue également les discrétisations homogènes de celles hétérogènes. Ces dernières sont qualifiées de modèle à coordonnées sigma. Les modèles sont paramétrés par des conditions aux frontières : (i) en surface, par les tensions dues au vent, les conditions cinématiques, ainsi que les flux de chaleur et d'eau douce, (ii) sur les autres frontières (au bord et au fond du domaine étudié) par les tensions de cisaillement et les conditions cinématiques. Les modèles hydrodynamiques produisent ainsi

des champs de courants grillés et indexés dans le temps (e.g. un champ de courant toutes les heures ou tous les jours pendant une période donnée), indiquant les valeurs de vitesse zonale, méridionale et verticale pour les modèles en trois dimensions, sur chaque point de grille horizontale et ce, sur plusieurs couches verticales (Figure 1.11).

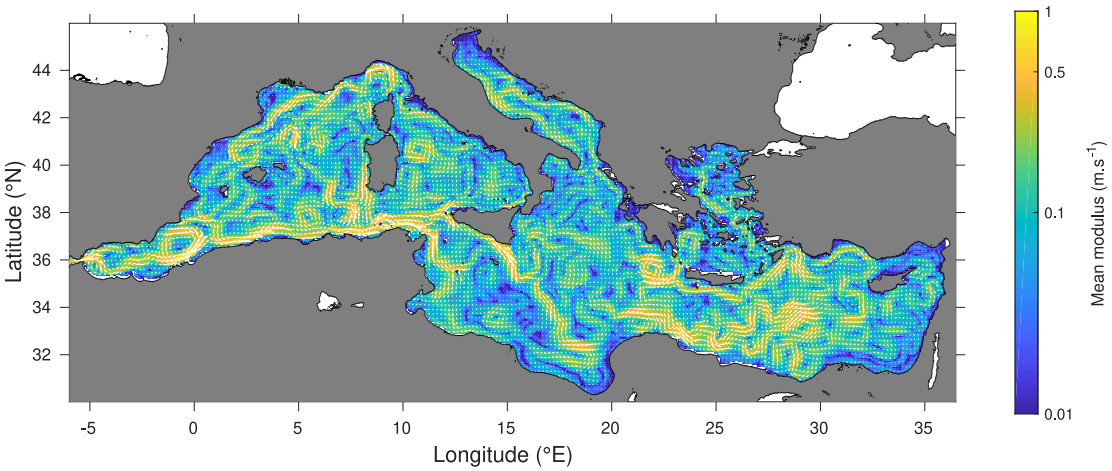


FIGURE 1.11. – Instantané de la circulation de surface journalière en Méditerranée : la direction des courants est indiquée par les flèches blanches, l'intensité du courant (norme du vecteur de courant ou *mean modulus*) par la couleur de fond. Le champ de courants cartographié ici provient du modèle hydrodynamique Mediterranean Forecasting System basé sur NEMO-OPA (Madec et al., 2015). Ce modèle opérationnel est implémenté sur la Méditerranée avec un maillage horizontal régulier de 1/16° (Oddo et al., 2009) et un maillage vertical homogène. Les courants moyens cartographiées sur la Figure 1.10 sont identifiables sur cet instantané (e.g. courant Algérien, courant Nord, gyre de Pelops, etc.).

Lorsque l'on s'intéresse au processus de transport larvaire et que l'on veut étudier la dispersion sous-jacente, deux paramètres sont fondamentaux dans l'implémentation d'un modèle hydrodynamique ou le choix des champs de courants modélisés : (i) la taille du domaine modélisé, pouvant aller d'une baie locale de quelques dizaines de km à un basin océanique de plusieurs milliers de km, (ii) la durée de la simulation, pouvant aller de quelques mois à plusieurs dizaines d'années. Généralement, et en lien avec la puissance de calcul requise, les simulations à l'échelle du bassin sont disponibles pour une période plus longue, avec une itération temporelle journalière ou hebdomadaire, que celles qui rendent compte de la circulation à l'échelle locale. Parallèlement, les modèles *locaux* sont plus précis (maillage horizontal de quelques centaines de mètres, et champs de courant calculés d'heures en heures, Rossi et al., 2020) et rendent mieux comptent des processus physiques à méso- et sous-méso-échelles (Figure 1.9). Pour

les modèles à grande échelle, le coût de calcul potentiel induit un maillage horizontal plus étendu, de l'ordre d'une dizaine de km. Ainsi, les modèles à petites échelles sont adéquats pour des questions de connectivité démographique (cf. section 1.3.1), alors que l'étude de la connectivité génétique demande plutôt des modèles à grande échelle (cf. section 1.3.2).

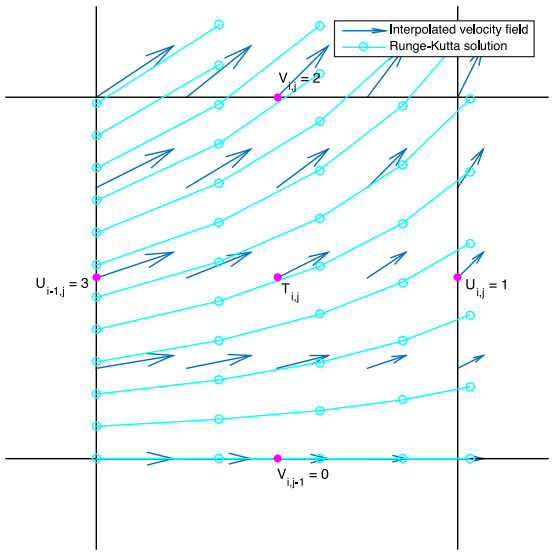


FIGURE 1.12. – Schéma du calcul des trajectoires de particules fluides (en cyan) sur une grille de modèle hydrodynamique (ici, une grille Arakawa-C). Les vitesses méridionales ($U_{i,j}$) et zonales ($V_{i,j}$) des courants sont données sur les points magentas. Les flèches bleues correspondent à l'interpolation linéaire du champ de vitesse. En considérant que les particules sont relâchées sur le bord gauche (i-1), les lignes cyan correspondent à leurs trajectoires. Elles sont obtenues en calculant les positions successives des particules à chaque pas de temps avec une méthode de Runge-Kutta d'ordre 4. Adapté depuis van Sebille et al., 2018.

1.2.2.3. Les modèles biophysiques Lagrangiens pour simuler le transport des propagules par les courants marins

Les champs de courant produits par les modèles de circulation océanique reproduisent les conditions hydrodynamiques observées par les propagules pendant leurs transports et ce pour chaque point de grille, et pour toute la durée de la phase pélagique. Ils décrivent donc, pour un pas de temps donné, l'écoulement hydrodynamique depuis un référentiel fixe : c'est une approche Eulérienne. L'approche Lagrangienne, elle, décrit l'écoulement hydrodynamique depuis un référentiel mobile, une particule

fluide, dans l'espace et le temps. En d'autres termes, l'approche Eulérienne étudiera l'écoulement d'une rivière en mesurant la vitesse d'écoulement à un point fixe, comme depuis un pont par exemple, alors que l'approche Lagrangienne obtiendra une mesure comparable en estimant la vitesse d'un point dérivant avec le fluide, comme depuis un kayak. Les modèles Lagrangiens permettent de simuler le déplacement de particules fluides (i.e. des propagules *numériques*) dans l'espace et le temps (Figure 1.12). Il existe une multitude de modèles, qui utilisent des techniques différentes pour étudier le transport de matière dans l'océan (décrites et détaillées dans van Sebille et al., 2018). Pour cette thèse, on utilise le modèle Lagrangian Flow Network (Ser-Giacomi, Rossi et al., 2015, son principe général étant détaillé dans les chapitres II et III). C'est un modèle offline : les trajectoires de particules fluides sont obtenues par itérations successives de leurs positions sur un champ de vitesse, lui-même simulé par un modèle hydrodynamique Eulérien (Figure 1.11).

Durant le processus de dispersion, trois paramètres biologiques sont primordiaux pour simuler le transport des propagules par les courants marins : le lieu du relargage des propagules dans la colonne d'eau, sa date et la PLD. Ces trois paramètres indiquent respectivement l'emplacement initial des particules fluides, le choix des champs de courants simulés par les modèles hydrodynamiques et le temps d'intégration des particules fluides au sein de ce champs de courants. Les trajectoires des propagules calculées par le modèle biophysique Lagrangien pourront ainsi rendre compte du transport par les courants marins entre deux points géographiquement définis dans l'espace.

1.2.2.4. Influence de la biologie sur le processus de transport

Parmi les différentes types de propagules observée pendant la phase pélagique pour disperser passivement par les courants marins, les larves sont la forme de développement la plus étudiée (Burgess et al., 2016). On distingue habituellement deux types de larves, caractérisées par leurs modes de nutritions :

— Les larves lécithotrophes se développent d'abord dans des œufs relativement grands et peu nombreux (de l'ordre d'un millier par ponte) puis se nourrissent grâce à une vésicule vitelline pendant le transport. La disponibilité en nourriture étant limitée, ces larves ont possiblement une durée de dispersion relativement courte (mais cette hypothèse est rarement testée et peu confirmée, Mercier et al., 2013).

— Les larves planctotrophes se développent d'abord dans des petits œufs très nom-

breux (jusqu'à plusieurs millions par ponte) pendant une durée relativement courte puis se nourrissent de phyto- ou zoo- plancton dans la colonne d'eau. Bien que dépendantes des ressources du milieu, les larves planctotrophes ne sont pas limitées par un stock fixé d'énergie et peuvent disperser sur une période relativement longue.

Il est estimé que 60 à 90 % des invertébrés marins produisent des larves planctotrophes, pour environ 10 % de larves lécithotrophes (Pechenik, 1999). Cependant, les larves lécithotrophes sont peut-être mieux adaptées à l'acidification des océans (Dupont et al., 2010). D'autres formes de développement larvaire existent, sans pour autant être caractérisées par cette dichotomie lécithotrophe/planctotrophe (Allen et Pernet, 2007). Par exemple des larves qui se nourrissent dans la colonne d'eau mais qui possèdent une phase larvaire courte. Cette phase larvaire pourrait avoir été sélectionnée pour permettre une plus grande fécondité, améliorer la croissance et limiter la prédation des larves plutôt que pour permettre l'expansion de l'aire vitale ou répondre aux extinctions locales de population (voir section 1.3.1, Mercier et al., 2013). Bien que les larves soient la forme de développement la plus présente, ou du moins la plus étudiée pour disperser (i.e. dispersion larvaire, Cowen et al., 2007 ; Cowen et Sponaugle, 2009 ; Pineda et al., 2007), on dénombre également d'autres formes pélagiques. Chez les herbiers marins (*Zosteraceae*, *Posidoniaceae*, *Cymodoceae* et *Hydrocharitaceae*), il existe quatre formes de déplacement : les pollens, les propagules sexuées (graines ou fruits), les fragments végétatifs décrochés par fortes contraintes hydrodynamiques et la propagation des individus par clonage (Kendrick et al., 2017 ; Kendrick et al., 2012 ; McMahon et al., 2014). Chacune de ces formes est associée à une période et à un temps particulier de dispersion (Figure 1.13, McMahon et al., 2014). Le même mécanisme s'observe chez les algues, qui peuvent disperser sur de courtes distances via la reproduction et l'envoi de gamètes dans la colonne d'eau, ou sur de plus grandes distances par le détachement de fragments d'individu lors de tempêtes (Buonomo et al., 2017 ; Norton, 1992).

1.2.3. Le processus de dispersion : de la ponte jusqu'à l'installation

La dispersion considère donc, en plus de la survie des propagules pendant le transport, le processus de ponte et d'installation (Figure 1.8 encadré vert, Pineda et al., 2007).

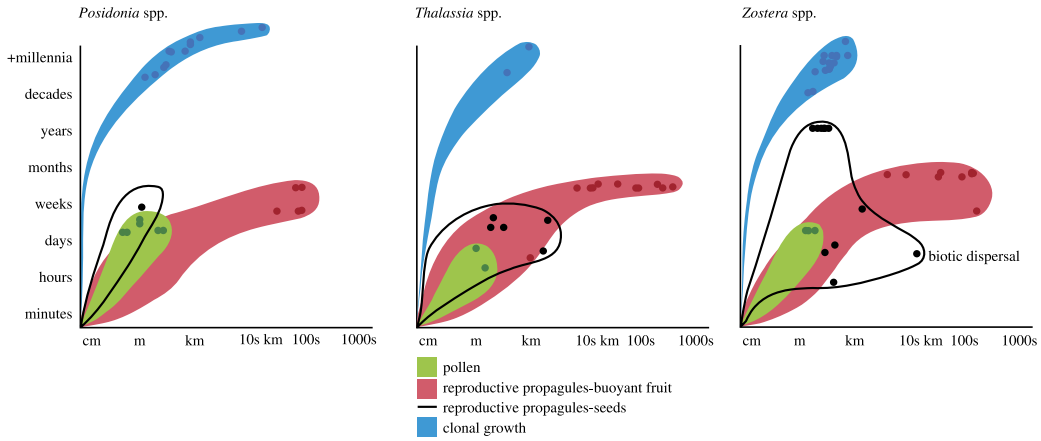


FIGURE 1.13. – Estimation de la distance et du temps de dispersion en fonction du mode de dispersion chez trois groupes d'herbiers marins : *Posidonia spp*, *Thalassia spp* et *Zostera spp*. Il n'y a pas de données pour les mouvements induits par la dispersion de fragments végétatifs, donc la catégorie n'est pas indiquée sur la figure. Tiré de McMahon et al., 2014.

1.2.3.1. La ponte

La ponte, en définissant le moment et le lieu où la phase pélagique débute est déterminante pour l'ensemble du processus de dispersion. En d'autres termes, la ponte fixe le point et la date de départ des possibles routes migratoires décrites par les propagules. Un changement dans l'espace ou le temps du processus de ponte peut potentiellement induire des schémas de dispersion totalement différents du fait de la variabilité spatio-temporelle de la circulation océanique. Plusieurs facteurs peuvent déclencher le mécanisme de ponte, et sont donc décisifs dans le processus global de dispersion. Ils peuvent être abiotiques comme le changement de température de l'eau de mer ou biotiques comme la quantité de réserves lipidiques des géniteurs (c.f. Chapitre II et Tableau 2 dans l'annexe A.2 pour une compilation des différents facteurs influençant le processus de ponte chez les téléostéens, Di Stefano et al., 2022). Au-delà du moment de la ponte, la qualité et la quantité de la descendance, déterminante dans le taux de survie lors du transport puis de l'installation, est fonction de la condition des géniteurs : de leurs âges, de leurs disponibilités en ressources, ainsi que du succès de leurs fécondations (Pineda et al., 2007). L'efficacité de la dispersion larvaire dépend donc de la santé des populations génitrices (voir Chapitre II, Lester et al., 2009 ; Marshall et al., 2019). Dans le cas de certaines espèces de poissons côtiers, plusieurs centaines d'individus se regroupent au même endroit pour se reproduire

et synchroniser leurs pontes (i.e. *spawning aggregation*, Figure 1.14, surtout étudiée dans les récifs coralliens, Domeier et Colin, 1997). Ces agrégations de frai permettent d'augmenter les chances de fécondation des gamètes (i.e. probabilités de rencontre) et la survie des larves (saturation du milieu diminuant la pression de prédation, etc.), mais rendent également les géniteurs vulnérables à la prédation. Ils deviennent en effet une cible facile et lucrative pour les pêcheries lorsque de telles agrégations sont localisées, ce qui rend les populations extrêmement sensibles à la surexploitation (Erisman et al., 2017 ; Mitcheson et al., 2008).

FIGURE 1.14. – Mérous camouflage (*Epinephelus polyphekadion*) pendant une agrégation de frai dans la passe de Fakarava (crédit photographique : Laurent Ballesta – Andromède Océanologie).

Il est important de noter que le passage de la phase benthique à la phase pélagique, donc du début du processus de dispersion, ne fait pas toujours suite à une ponte. Chez les herbiers marins et les algues (c.f. section 1.2.2.4) la séparation des propagules sexuées (graines ou fruits), la libération des gamètes ou encore le détachement de fragment d'organisme lors de fortes contraintes hydrodynamique induisent également le début du processus de dispersion. Dans cette thèse, le terme « ponte » englobe tous ces mécanismes différents.

1.2.3.2. L'installation

Après avoir été transportées par les courants marins, les propagules, en s'installant sur un substrat grâce, ou non, à des capacités intrinsèques de déplacement et d'orientation (Faillettaz, Paris et al., 2018 ; Leis, 2006 ; Paris, Atema et al., 2013), mettent fin

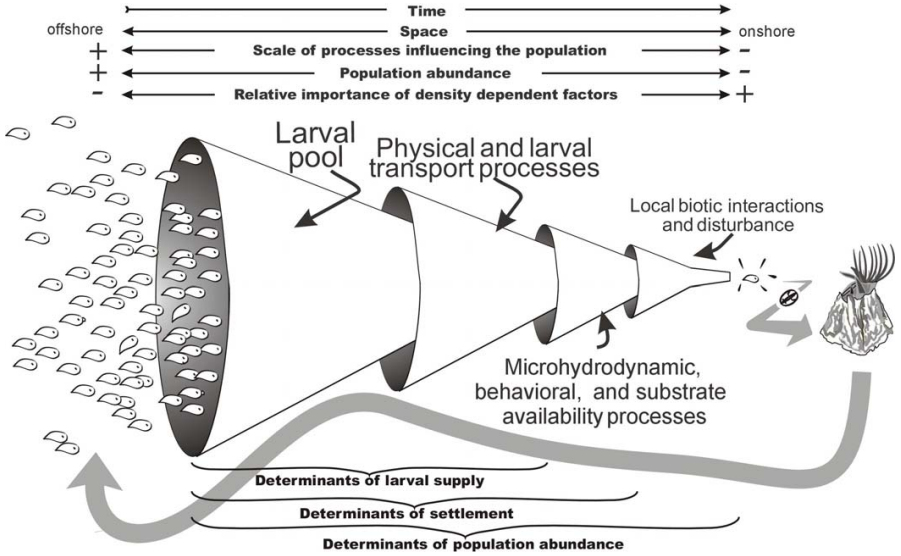


FIGURE 1.15. – Schéma représentant les mécanismes successifs influençant la dispersion depuis la ponte jusqu'à l'installation et la survie sur un substrat ou un habitat. L'abondance relative des individus décroit au fil du temps sur l'effet de la mortalité importante observée durant la phase pélagique. Voir également Figure 2 dans Cowen et Sponaugle, 2009. Adapté depuis Pineda, 2000.

au processus de dispersion. Si certaines espèces s'adaptent à une grande variété de substrats, la plupart doivent parvenir durant leurs phases pélagiques à un habitat spécifique, leur permettant de continuer leur cycle de vie (Pineda et al., 2007). De plus, l'installation est un évènement risqué où le taux de prédation est important. L'efficacité de la dispersion larvaire est alors contrainte par ce processus, où le nombre de larves arrivées sur un lieu d'installation après le transport (*larval supply*) n'est généralement pas corrélé avec le nombre de larves qui s'installent effectivement (Pineda et al., 2010).

Ainsi, depuis la ponte, où entre plusieurs milliers et plusieurs millions de propagules (respectivement pour les larves lécithotrophes et planctotrophes, section 1.2.2.4) sont relâchées dans la colonne d'eau, seule une fraction va survivre jusqu'à l'installation (Figure 1.15, Pineda, 2000). La connectivité entre populations est alors théoriquement réalisée lorsque les propagules survivent au processus complet de dispersion, se développent dans leurs nouveaux habitats, poursuivent leur croissance pour finalement se reproduire entre individus adultes (Figure 1.8 encadré orange, Cowen et Sponaugle, 2009 ; Pineda et al., 2007).

1.3. La connectivité dans le monde marin : de l'échange d'individus au flux de gènes

1.3.1. La connectivité démographique

La connectivité démographique est un des indicateurs de la connectivité entre populations (Kool et al., 2013). Elle se réfère spécifiquement à l'impact du flux d'individus (i.e. immigration et émigration) sur le bilan démographique d'une population, indexée par l'équilibre entre les apports de la natalité et de l'immigration (repeuplement) et les pertes d'effectifs dues à la mortalité et à l'émigration (Lowe et Allendorf, 2010). La connectivité démographique est déterminante dans la dynamique spatio-temporelle des populations en influant ainsi sur leurs croissances, leurs maintiens et leurs extinctions (i.e. bilan démographique respectivement positif, nul et négatif). L'étude de la dynamique des populations est fondamentale dans la gestion des pêcheries pour définir des rendements durables. La connectivité démographique entre populations nécessite une coopération entre des pays possédant des pêcheries limitrophes, et peut entraîner des répercussions importantes sur leur sécurité alimentaire, l'économie et l'emploi (Kough et al., 2013 ; Ramesh et al., 2019).

Pour les espèces à cycle de vie biphasique, l'étude démographique doit être adaptée à la particularité de la phase pélagique et au caractère sédentaire des populations adultes. Lors de la ponte, on peut considérer que toutes les propagules relâchées dans la colonne d'eau sont des émigrants. Seuls les individus, qui, à l'issue du processus de dispersion s'établissent dans la même population émettrice peuvent donc être considérés *stricto-sensu* comme des naissances. Ils sont alors « auto-recrutés » dans leur population d'origine. Le processus d'auto-recrutement est souvent considéré comme prédominant dans le maintien démographique des populations des poissons tropicaux (e.g. Almany et al., 2007 ; Teske et al., 2016), alors que les systèmes tempérées sembleraient dépendre autant des immigrants que des auto-recrutés (Hidalgo et al., 2019). C'est ainsi qu'il a été défini le ratio d'auto-recrutement, soit le rapport $\frac{\text{individus auto-recrutés}}{\text{nombre total d'immigrants}}$ (Botsford et al., 2009 ; Dubois et al., 2016). La dispersion étant réalisée pendant les premiers stades de vie, le reste des émigrants ne doit pas être considéré démographiquement (i.e. temps de résidence nul). On peut tout de même quantifier la rétention locale, un indicateur qui renseigne l'efficacité de l'auto-recrutement par rapport à la production locale de propagule. le ratio de rétention locale est défini par le rapport $\frac{\text{individus auto-recrutés}}{\text{nombre total d'émigrants}}$ (Botsford et al., 2009 ; Dubois et

al., 2016). Contrairement à l'auto-recrutement, la rétention locale est difficilement quantifiable *in situ* du fait de la complexité d'un échantillonnage exhaustif des émigrants. Cependant, ces deux métriques peuvent être calculée grâce à la modélisation biophysique (e.g. Dubois et al., 2016).

L'interdépendance démographique entre deux populations ou entre une population et sa métapopulation (i.e. un assemblage de populations distinctes qui sont connectées entre elles par un flux d'individus, Cowen et Sponaugle, 2009) ne dépend pas seulement du nombre absolu de propagules échangées pendant la dispersion, mais également de la compensation de la mortalité par l'auto-recrutement de chaque population (Figure 1.16, Lowe et Allendorf, 2010). Pour une population caractérisée par une démographie croissante (i.e. auto-recrutement >> mortalité), l'apport de propagules par d'autres populations connectées peut être important mais ne représenter qu'une fraction du nombre total d'immigrant (i.e. ratio d'auto-recrutement ~ 1). Parallèlement, le flux d'individus immigrants peut compenser un auto-recrutement faible ou une mortalité importante dans des populations stables ou en déclins. Dans une situation où une population est proche de l'extinction (i.e. auto-recrutement << mortalité), l'apport d'individus depuis d'autre populations est déterminant dans son maintien potentiel (i.e. *rescue-effect*, Eriksson et al., 2014 ; Sanz-Aguilar et al., 2016) et représente une fraction importante du nombre total d'immigrants (i.e. ratio d'auto-recrutement ~ 0). De manière générale, une population est dite interdépendante démographiquement lorsque son ratio d'auto-recrutement est inférieur ou égale à 0.9 (i.e. au moins 10 % des individus recrutés proviennent d'autres populations, Waples et Gaggiotti, 2006). La dispersion est ainsi déterminante dans la persistance des populations locales (Burgess et al., 2014).

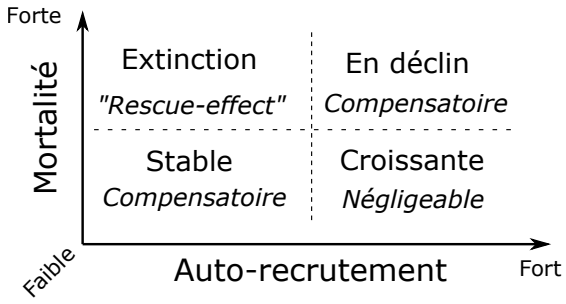


FIGURE 1.16. – Impact de la connectivité (en italique) sur la dynamique démographique des populations, en fonction de l'intensité de l'auto-recrutement et de la mortalité (Lowe et Allendorf, 2010).

La connectivité démographique, permet également la colonisation par la dispersion de patchs d'habitat qui ne contiennent pas de populations adultes. Une telle absence peut être due à la connectivité démographique elle-même et au déficit antérieur d'influx de migrant pour les espèces annuelles (i.e. déclin démographique) ou à l'extinction des populations précédentes (e.g. pression de pêche, pollution, vague de chaleur, etc., Leggat et al., 2019). La colonisation de nouveaux habitats a également lieu lors de l'introduction d'espèces invasives (e.g. Azzurro et al., 2013) ou par des évènements rare de dispersion longue distance en dehors de l'aire de répartition de l'espèce. Dans le dernier cas, les individus qui sont sur le front de l'expansion peuvent se reproduire puis disperser facilement car ils bénéficient d'un environnement vacant, empêchant les seconds colonisateurs de s'établir. Ce processus démographique dépendant de la densité peut avoir des répercussions critiques sur la structure spatiale de la diversité spécifique et génétique (i.e. hypothèse du « *founder takes all* », Waters et al., 2013).

1.3.2. La connectivité génétique

La dispersion, potentiellement suivi de la reproduction, conduit également au flux de gènes, soit à l'échange de matériel génétique d'une population à une autre. Le flux de gènes, en redistribuant continuellement les fréquences alléliques dans l'espace, tend à homogénéiser la variabilité génétique entre populations, contrebalançant l'adaptation locale et la dérive (Hellberg, 2009 ; Slatkin, 1987 ; Slatkin, 1985). En permettant l'introduction d'allèles étrangers dans des populations locales, le flux de gènes minimise l'effet de dépression de consanguinité mais peut contribuer également à réduire l'adaptation des individus de la population aux conditions locales (Hellberg, 2009 ; Lenormand, 2002 ; Lowe et al., 2017).

L'étude de la connectivité génétique consiste à caractériser le flux de gènes entre populations et à évaluer son influence relative par rapport aux autres processus évolutifs (i.e. dérive génétique, mutations et sélection naturelle) sur la distribution spatiale des fréquences alléliques (Lowe et Allendorf, 2010). Le flux de gènes est directement induit par l'échange d'individus entre populations sur une génération. Du fait de la transmission d'allèles d'une génération à l'autre, le flux de gènes peut également être engendré entre des populations éloignées (i.e. non connectées démographiquement) par plusieurs processus de dispersion successifs entre différentes populations intermédiaires. Il en résulte que le flux de gènes impacte les structures génétiques sur des échelles spatio-temporelles caractéristiques des processus écologiques et évolutifs

simultanément. Les méthodes de génétique des populations utilisées dans cette thèse permettent d'étudier les phénomènes évolutifs sur le temps long et ne reflète que l'intégration de la connectivité génétique sur des échelles spatio-temporelles plus grandes que celles de la connectivité démographique.

1.3.3. Différentes méthodes empiriques pour évaluer et quantifier la connectivité multi-échelle

Pour quantifier le processus de dispersion, on utilise des méthodes qui permettent d'évaluer empiriquement les mouvements d'individus dans l'espace et le temps (cf. section 1.1.3). Il existe différentes méthodes pour estimer de manière directe les mouvements individuels. Chacune d'entre elle est caractérisée par une portée et une résolution spatio-temporelle adaptée à l'espèce étudiée (Calò et al., 2013 ; Kool et al., 2013) : observation visuelle, marquage-recapture, télémesure, capteur GPS, radar, surveillance par image satellitaire (e.g. dérive des sargasses dans l'Atlantique, Berline et al., 2020), avion ou récemment drone (Oleksyn et al., 2021).

Même si ces méthodes directes permettent d'obtenir les informations les plus précises sur les déplacements d'individus pour caractériser la connectivité démographique, elles ne sont pas adaptées pour quantifier la dispersion de propagules par les courants marins (e.g. taille des propagules rapportée à l'échelle spatiale de dispersion, marquage à la ponte, etc., Calò et al., 2013). L'emploi de méthodes indirectes est donc nécessaire pour estimer la connectivité. Dans cette thèse, on utilise des données issues de méthodes d'analyse d'otolithes et de méthodes génétiques.

1.3.3.1. La sclérochronologie et l'analyse géochimique des otolithes

Les otolithes sont des concrétions minérales que l'on retrouve dans l'oreille interne des vertébrés, et qui croissent pendant toute la durée de leurs vies (Thorrold et al., 2007). Ce sont des structures uniques, caractérisées par l'ajout successif journalier de concrétions de carbonate de calcium depuis le début de l'ontogenèse, et qui ne sont pas soumises à des modifications ultérieures. L'étude sclérochronologique des otolithes permet ainsi d'estimer la durée des premiers stade de vie, soit la date de ponte et la PLD (Calò et al., 2013 ; Di Franco et al., 2011 ; Di Franco et al., 2013). De plus, les otolithes incorporent dans leurs matrices de carbonate de calcium des traces d'éléments chimiques, reflétant leurs disponibilités dans l'environnement proche des individus (Campana, 1999 ; B. Green et al., 2009 ; Thorrold et al., 2007). Cette propriété

permet d'utiliser les otolithes comme des marqueurs naturels en analysant leurs compositions en élément traces ou en isotopes stables, afin d'identifier des signatures géochimiques particulières pour les différents stades de vie des poissons (Calò et al., 2013 ; Di Franco et al., 2015 ; Di Franco et al., 2012 ; Thorrold et al., 2007). Particulièrement, l'analyse géochimique du cœur de l'otolithe (formé avant l'éclosion des œufs) permet de distinguer chez des larves qui viennent de coloniser des nourriceries, différentes origines natales (Di Franco et al., 2015 ; Di Franco et al., 2012), sans pouvoir toutefois les définir spatialement. Même si l'utilisation des otolithes est adaptée à l'étude indirecte de la connectivité des poissons osseux côtiers (e.g. Ostéichtyens), cette méthode d'analyse sclérochronologique et géochimique peut également être appliquée sur les statolithes (analogue des otolithes pour certains invertébrés), les protoconques (coquilles larvaires de certain mollusques) ou les carapaces des crustacées (Thorrold et al., 2007).

1.3.3.2. Les méthodes moléculaires issues de la génétique des populations

En génétique des populations, différentes méthodes génétiques indirectes sont utilisées pour estimer la connectivité à multi-échelle. L'analyse de parenté utilise un échantillonnage conséquent d'individus (e.g. plusieurs milliers d'adultes et de juvéniles échantillonnées sur quelques dizaines de km), qui sont ensuite génotypés afin d'établir des liens de parenté génétique et d'identifier (au mieux) des couples parents-enfants (on peut parfois simplement estimer des populations d'origine probables pour certains individus sans pouvoir identifier leurs parents précisément). L'analyse de parenté permet ainsi d'estimer la connectivité démographique entre populations pour une génération de dispersion, en particulier son échelle spatiale caractéristique (i.e. noyaux de dispersion, D'Aloia et al., 2015 ; Pinsky et al., 2017), ses différents patrons spatiaux (Harrison et al., 2020) et sa variabilité temporelle (Catalano et al., 2021). Ce type d'analyse est parfois impossible à réaliser pour les espèces peu structurées spatialement, et ce quel que soit le nombre, la qualité et le polymorphisme des marqueurs moléculaires caractérisés.

Les mesures de distribution de fréquence allélique par l'utilisation de différents marqueurs génétiques (e.g. microsatellites, SNPs, RADseq, voir Chapitre IV) permettent d'évaluer les différentiations génétiques entre populations sur une échelle spatiale allant de quelques kilomètres (e.g. Schunter et al., 2019) à plusieurs centaines de kilomètres (e.g. Weber et al., 2015). Dans cette thèse, l'indice de fixation Fst (Wright,

1931), qui est un des indices les plus utilisés dans la génétique de population, permet de quantifier cette différentiation génétique entre paires de populations :

$$Fst = \frac{\sigma^2}{q(1-q)} \tag{1.1}$$

où q et σ^2 sont respectivement la moyenne et la variance de la fréquence allélique entre populations. Cet indice varie entre 0 et 1 et permet de distinguer le cas où les populations ont exactement les mêmes fréquences alléliques (Fst égal à 0) des cas où les populations ont des fréquences plus ou moins différentes entre elles. Quand chaque population a un allèle unique et différent de chaque autre population, leurs Fst égalent 1. Les populations caractérisées par un flux de gènes important voient leurs fréquences alléliques s'homogénéiser et présentent de faibles valeurs de Fst. Dans le cas opposé, un flux de gènes limité ou nul induit des valeurs de Fst élevées : les populations sont alors structurées génétiquement. Les mesures de différentiation génétique estiment ainsi indirectement la connectivité génétique sur différentes échelles spatiales (plus étendues que l'analyse de parenté) et intégrées sur de multiples générations de dispersion (Kool et al., 2013).

1.3.4. Utilisation des modèles biophysiques pour estimer la connectivité démographique et génétique

La connectivité est réalisée lorsque les propagules survivent au processus de dispersion pour potentiellement se reproduire dans une nouvelle population. Les modèles biophysiques permettent de simuler le transport de propagules par les courants marins, et offrent ainsi une approche mécaniste à la quantification spatiale de la connectivité entre population par la dispersion. Les simulations directes de connectivité par les modèles biophysiques sont ainsi complémentaires des analyses empiriques indirectes de la connectivité démographique et génétique.

Cette complémentarité s'illustre spécifiquement avec les méthodes d'otolithométrie pour étudier la connectivité démographique. D'une part, l'analyse sclérochronologique des otolithes, en évaluant la date de ponte et la PLD *in situ*, permet de paramétrer finement les modèles biophysiques (c.f. section 1.2.2.3). D'autre part, les modèles biophysiques, en simulant dans le processus de dispersion, la trajectoire des larves depuis la ponte jusqu'à l'installation, peuvent de manière rétrospective délimiter les différentes origines natales identifiées par les analyses géochimiques (voir Chapitre

II).

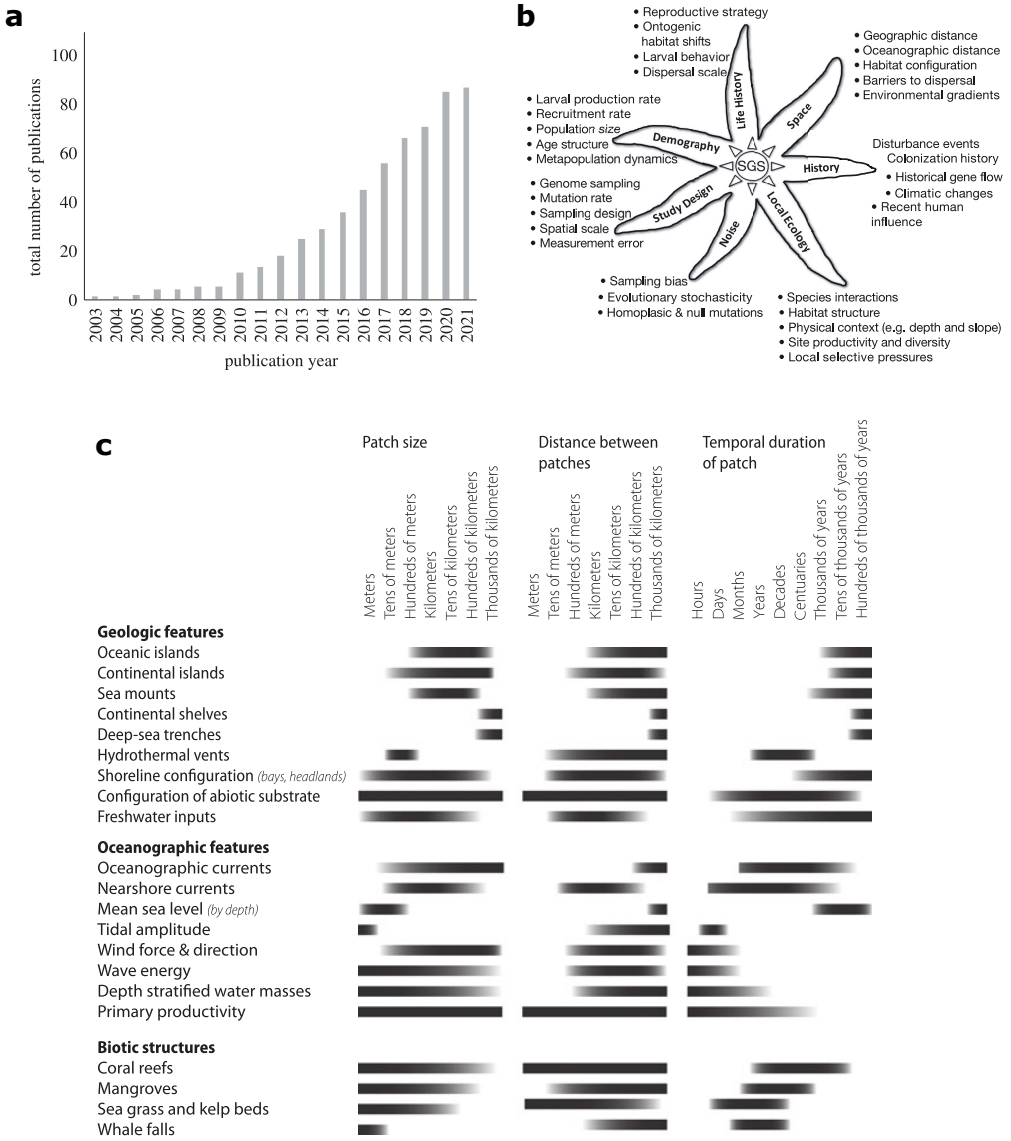


FIGURE 1.17. – L'approche du *seascape genetics* pour comprendre comment les différents processus écologiques et environnementaux influent sur les structures génétiques spatiale. **a** Nombre annuel de publications scientifiques qui combinent l'utilisation de méthodes de génétique des populations et de modèles biophysiques pour étudier l'influence de la connectivité sur les structure génétiques spatiales, d'après Jahnke et Jonsson, 2022. **b** Les forces et contraintes, réparties en sept catégories, qui influent simultanément sur les structures génétiques spatiale (SGS), d'après Selkoe et al., 2016. **c** Echelles spatio-temporelles à lesquelles les différentes caractéristiques de l'environnement marins influent sur les structures spatiales génétiques, d'après Riginos et Liggins, 2013.

Comprendre comment les différents processus écologiques et environnementaux du milieu marin influent sur les structures génétiques spatiales identifiées grâce aux méthodes moléculaires, et pour quelles échelles spatio-temporelles caractéristiques, s'inscrit dans l'approche multidisciplinaire du *seascape genetics* (Figure 1.17, Riginos et Liggins, 2013; Selkoe et al., 2016; Selkoe et al., 2008). Parmi tous les processus environnementaux, l'étude de la relation entre les courants marins et les structures spatiales génétiques est fondatrice dans le *seascape genetics* (Selkoe et al., 2016). Les courants marins influencent les structures génétiques sur des échelles spatiales allant de quelques mètres à plusieurs milliers de kilomètres (i.e. de la sub-méso-échelle à la large-échelle, section 1.2.2.1), et pour des échelles temporelles allant de quelques heures à plusieurs milliers d'années (i.e. échelles écologiques et évolutives, respectivement, Figure 1.15c). Dans ce cadre, les modèles biophysiques ont été largement utilisés pour estimer la connectivité démographique entre paires de population (Figure 1.15a, Jahnke et Jonsson, 2022), et ensuite les comparer avec des mesures de différentiation génétique (Liggins et al., 2013). L'hypothèse sous-jacente est que le nombre d'individus échangés entre paires de population ($N_e m$, voir section 1.3.3.2) est inversement proportionnel au Fst, ou du moins diminue quand le Fst augmente (Lowe et Allendorf, 2010).

Toutefois, les estimations de connectivité démographique sont inadaptées pour évaluer l'influence de la dispersion par les courants marins sur les structures génétiques à des échelles temporelles évolutives de plusieurs centaines de milliers d'années (Riginos et Liggins, 2013). L'utilisation de méthodes issues de la théorie de graphes, elle, permet d'intégrer la dispersion estimée par les modèles biophysiques sur de multiples générations et sur de multiples populations étapes, pour rendre compte de la connectivité génétique à large échelle spatio-temporelle (Kool et al., 2013).

1.4. Problématique générale

Dans cette thèse, nous nous attachons à étudier l'influence de la connectivité démographique et génétique via la dispersion sur la structure des populations et les schémas de biodiversité en mer Méditerranée. Le travail est structuré autour de deux grandes questions :

- **Quel est l'impact de la dispersion sur la structure spatiale des populations ?**
  - Comment délimiter les zones de pontes et caractériser la connectivité démographique chez les poissons côtiers ?
  - Quelles sont les impacts de la variabilité océanique sur la ponte et la connectivité démographique ?
- **Quel est le rôle de la dispersion sur les patrons de diversité génétique ?**
  - Comment définir et quantifier la connectivité génétique à partir d'évènement de dispersion successifs ?
  - Quelles sont les échelles spatio-temporelles qui permettent d'évaluer les structures génétiques entre populations ?

Les quatre prochains chapitres s'attellent à répondre à ces problématique sous la forme d'articles scientifiques publiées, en révision ou qui sera soumis prochainement. Le dernier chapitre consiste à une contextualisation générale du travail de thèse et fournit des perspectives et recommandations de recherche.

2. Délimitation des zones de pontes et quantification de la dispersion larvaire chez des poissons côtiers

2.1. Résumé

Comme vu en introduction, le processus de dispersion larvaire pour les espèces benthiques et sédentaires est réalisé par le transport des premiers stades de vie par les courants marins. Délimitée par la ponte et l'installation, cette phase pélagique dispersive est un vecteur de connectivité démographique entre populations marines.

Dans ce chapitre, nous avons créé un *cadre analytique* qui permet de quantifier les échelles de dispersion, de caractériser les patrons de connectivité et de localiser les zones de pontes. Pour ce faire, nous avons couplé l'utilisation d'un modèle biophysique de dispersion larvaire, de données biologiques sur les premiers stades de vie et d'informations biogéographiques sur l'espèce considéré. Ce *cadre* est pensé de sorte à être appliqué à n'importe quelle espèce à cycle de vie biphasique, et est modulable en fonction des données biologiques à disposition. En effet, les diagnostiques finaux de connectivité sont raffinés au fur et à mesure que sont rajoutés des filtres biogéographiques déduits des informations biologiques sur l'espèce ciblée (i.e. forçages environnementaux sur le comportement de ponte, habitat préférentiel des adultes, ou encore densité et taille d'adultes femelles matures). Nous avons testé et validé cette méthode avec deux études de cas centrées sur les poissons côtiers : le sar commun *Diplodus sargus sargus* (Linnaeus, 1758) et le sar à tête noire *Diplodus vulgaris* (Geoffrey Saint Hilaire, 1817). Pour nos deux études de cas, les informations biologiques sont obtenues grâce à l'étude des otolithes collectés sur des larves installées (*post-settlers* en anglais, i.e. des larves de 1 à 1.5 cm de long qui viennent de s'installer dans des

zones de nurseries avant d'être recrutées dans les populations adultes). Ces larves sont elles-mêmes échantillonnées dans sept localités le long de la côte Apulienne, au sud-ouest de la mer Adriatique. L'étude sclérochronologique des otolithes permet d'obtenir la durée de vie larvaire (PLD) ainsi que la date de ponte pour chaque individu. Ces données vont permettre de paramétrer notre modèle bio-physique afin de contraindre les simulations lagrangiennes (i.e. date de départ et durée d'intégration des particules numériques) aux observations. L'étude géochimique permet quant à elle de distinguer des assemblages différents dans la composition en éléments métalliques du cœur des otolithes, et ainsi de distinguer des origines natales différentes, sans toutefois savoir où elles se situent.

Bien qu'appartenant à la même famille, nos deux sujets d'études sont caractérisées par des traits de premier stade de vie différents : *D. sargus* pond au printemps et a une PLD moyenne de 17 jours, alors que *D. vulgaris* pond en hiver et a une PLD moyenne de 47 jours. Ces différents traits écologiques des premiers stades de vie se répercutent sur les échelles typiques de dispersion : les populations Apuliennes de *D. vulgaris* ont des zones de pontes potentiellement quatre fois plus étendues, et ce sur les deux côtes de l'Adriatique, ainsi qu'une distance maximale de dispersion deux fois supérieure à celle de *D. sargus* (environ 500 km). Grace à l'approche probabiliste du modèle bio-physique, chaque zone de ponte est caractérisée par un potentiel de ponte, c'est-à-dire la probabilité d'être la source larvaire des *post-settlers* échantillonnés le long de la côte Apulienne. Pour les deux espèces, la côte Apulienne affiche un potentiel de ponte important, illustrant la haute proportion d'auto-recrutement chez les espèces à cycle de vie biphasique (e.g. Almany et al., 2007 ; Teske et al., 2016). En prenant en compte le potentiel de ponte, les origines natales identifiées grâce à l'analyse géochimique des otolithes ont pu être délimitées spatialement : on en dénombre trois sur la côte Est et quatre sur la côte Ouest de la mer Adriatique. Les origines natales les mieux représentées pour les deux espèces se trouvent sur la côte Apulienne. Nos résultats montrent que les aires marines protégées (AMPs) représentent une source de larves considérable. Pour respectivement *D. sargus* (2 AMPs) et *D. vulgaris* (9 AMPs), près d'un tiers et un quart du total des larves arrivant sur la côte Apulienne proviennent d'AMPs. Or, il est montré que des AMPs bien gérées ont un effet bénéfique sur la biomasse et la taille des poissions côtiers (Lester et al., 2009), ce qui pourrait augmenter significativement la quantité de larves émises pendant la ponte (jusqu'à cinq fois plus en comparaison d'une zone non protégée, Marshall et al., 2019). La prise en compte, dans notre modèle de connectivité, d'une telle variabilité spatiale de la production larvaire soulignerait

d'autant plus l'effet bénéfique des AMPs dans le réapprovisionnement en migrants des zones non protégées. Nos résultats mettent en évidence qu'une description de la circulation océanique correspondant à la période de dispersion explique les schémas de connectivité démographique pour une période donnée : les plus hauts potentiels de ponte se situent en amont de la circulation moyenne de surface. Parallèlement, la turbulence à fine échelle peut expliquer des connexions non-intuitives comme celles qui vont à l'encontre de la circulation moyenne. Les courants marins présentent par nature une haute variabilité spatiale et temporelle, ce qui induit potentiellement une variabilité interannuelle du processus de dispersion larvaire. Là où notre description des schémas de connectivité représente un instantané pour un événement annuel de ponte, on peut s'interroger sur la possible variabilité interannuelle de ce processus.

Ce dernier point présente une perspective intéressante à ce chapitre. C'est pourquoi, dans la continuité des questions posées par ce travail, nous avons proposé puis encadré un stage de Master ayant pour sujet : « *la variabilité spatio-temporelle de la ponte et de la connectivité démographique chez des poissons côtiers* ». Les résultats de ce stage on fait l'objet d'un article soumis et actuellement en révision dans *Fisheries oceanography*. Le manuscrit est intégré en Annexe A.2, et est résumé dans cette section introductive du Chapitre II.

Dans cet article, nous avons utilisé le *cadre analytique* détaillé précédemment pour identifier les zones de pontes chez *D. sargus* pour deux années successives (2008 et 2009) et à différentes localités (mer Adriatique, Ligurienne et Tyrrhénienne). En d'autres mots, nous avons exploité au mieux l'échantillonnage de post-settlers et l'analyse sclérochronologique sous-jacente des otolithes réalisée dans de précédentes études (Di Franco et al., 2011 ; Di Franco & Guidetti, 2011). Grâce à cette délimitation spatiale des populations sources, nous avons pu identifier une gamme de températures favorable à la ponte pour le nord de la Méditerranée centrale chez *D. sargus*. La température étant considérée dans la littérature comme le principal facteur abiotique environnemental (et donc quantifiable à large échelle) déclencheur de la ponte. En tenant compte de tous les habitats favorables aux populations adultes de *D. sargus* (herbiers de posidonie, substrats durs et rocailleux), Nous avons pu, en utilisant le *cadre analytique* localiser toutes les zones de ponte potentielles et determiner si le processus de ponte était effectif en fonction de la température locale pour une décennie (2005-2014) à l'échelle de la Méditerranée centrale.

Le processus de ponte et la connectivité sous-jacente sont marqués par une haute variabilité intra-annuelle et interannuelle. Les années marquées par des hivers longs

et froids (2012) et parallèlement courts et chauds (2007) sont respectivement délétères et favorables pour l'efficacité du processus de ponte, et apparaissent donc comme des années particulières sur la période étudiée. Cette variabilité interannuelle semble être induite par des oscillations climatiques à large échelle (Adriatic-Ionian Bimodal Oscillating System, BiOS), mais aussi par des conditions océanographiques locales (e.g. événements sporadiques d'upwelling ou downwelling). A l'échelle intra-annuelle, la probabilité pour chaque population d'être une zone de ponte effective est hautement variable au fil de la saison de ponte potentielle (de début janvier à fin juillet). Grâce à une méthode de clusteristation, nous avons pu attribuer à chaque zone de ponte une phénologie caractéristique. Nous avons dénombré quatre phénologies différentes, chacune expliquée par des conditions climatiques et océanographiques particulières. A l'échelle décennale, 50 % des zones de pontes présentent la même phénologie d'une année à l'autre, illustrant à la fois la présence de routes migratoires constantes (notamment en mer Ligure), mais aussi la variabilité interannuelle dans les patrons de connectivité démographique et la structure sous-jacente des populations.

Ainsi, nous avons produit des simulations décennales de connectivité démographique à l'échelle du bassin central Méditerranéen en nous affranchissant des contraintes induites par un échantillonnage souvent limité dans l'espace et le temps. La variabilité temporelle marquée des patrons de connectivité démographique suggère un brassage génétique complexe entre populations à l'échelle de plusieurs générations successives de dispersion. Caractériser ainsi l'ensemble des connexions génétiques nécessite l'emploie de la théorie des graphes, avec comme finalité, la modélisation mécanistique du flux de gènes entre populations.

2.2. Article

Legrand, T., Di Franco, A., Ser-Giacomi, E., Caló, A. & Rossi, V. (2019). A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish. Marine Environmental Research, 151, 104761. https://doi.org/10.1016/j.marenvres.2019.104761

Marine Environmental Research 151 (2019) 104761

Contents lists available at ScienceDirect

Marine Environmental Research

journal homepage: www.elsevier.com/locate/marenvrev

A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish

T. Legrand[a,*], A. Di Franco[b,c], E. Ser-Giacomi[d], A. Caló[c,e], V. Rossi[a]

[a] Mediterranean Institute of Oceanography (UM 110, UMR 7294), CNRS, Aix Marseille Univ., Univ. Toulon, IRD, 13288, Marseille, France
[b] Stazione zoologica Anton Dohrn, Dipartimento Ecologia Marina Integrata, Sede Interdipartimentale della Sicilia, Lungomare Cristoforo Colombo (complesso Roosevelt), 90142 Palermo, Italy
[c] Université Côte d'Azur, CNRS, UMR 7035 ECOSEAS, Parc Valrose 28, Avenue Valrose, 06108, Nice, France
[d] Sorbonne Universités (UPMC, Université Paris 06)-CNRS-IRD-MNHN, LOCEAN, 4 Place JUSSIEU, F-75005, PARIS, France
[e] Dipartimento di Scienze della Terra e del Mare (DiSTeM), Università di Palermo, Via Archirafi 20, 90123 Palermo, Italy

ARTICLE INFO

Keywords:
Marine connectivity
Lagrangian flow network
Conservation
Marine protected area
Fish natal origins
Coastal fishes
Mediterranean sea
Ecosystem management
Population dynamics
Models-hydrodynamics

ABSTRACT

Assessing larval dispersal is essential to understand the structure and dynamics of marine populations. However, knowledge about early-life dispersal is sparse, and so is our understanding of the spawning process, perhaps the most obscure component of biphasic life cycles. Indeed, poorly known species-specific spawning modality and species-specific early-life traits, as well as the high spatio-temporal variability of the oceanic circulation experienced during larval drift, hamper our ability to appraise the realized connectivity of coastal fishes. Here, we propose an analytical framework which combines Lagrangian modelling, network theory, otolith analyses and biogeographical information to pinpoint and characterize larval sources which are then grouped into discrete spawning areas. Such well-delineated larval sources allow improving the quantitative evaluations of both dispersal scales and connectivity patterns. To illustrate its added value, our approach is applied to two case-studies focusing on *Diplodus sargus* and *Diplodus vulgaris* in the Adriatic sea. We evidence robust correlations between otolith geochemistry and modelled spawning areas to assess their relative importance for the larval replenishment of the Apulian coast. Our results show that, contrary to *D. sargus*, *D. vulgaris* larvae originate from both eastern and western Adriatic shorelines. Our findings also suggest that dispersal distances and dispersal surfaces scale differently with the pelagic larval duration. Furthermore, 30.8% of *D. sargus* larvae and 23.6% of *D. vulgaris* larvae of the Apulian populations originate from Marine protected area (MPA), exemplifying larval export from MPAs to surrounding unprotected areas. This flexible multidisciplinary framework, which can be adjusted to any coastal fish and oceanic system, exploits the explanatory power of a dispersal model, fine-tuned and backed-up by observations, to provide more reliable scientific basis for the management and conservation of marine ecosystems.

1. Introduction

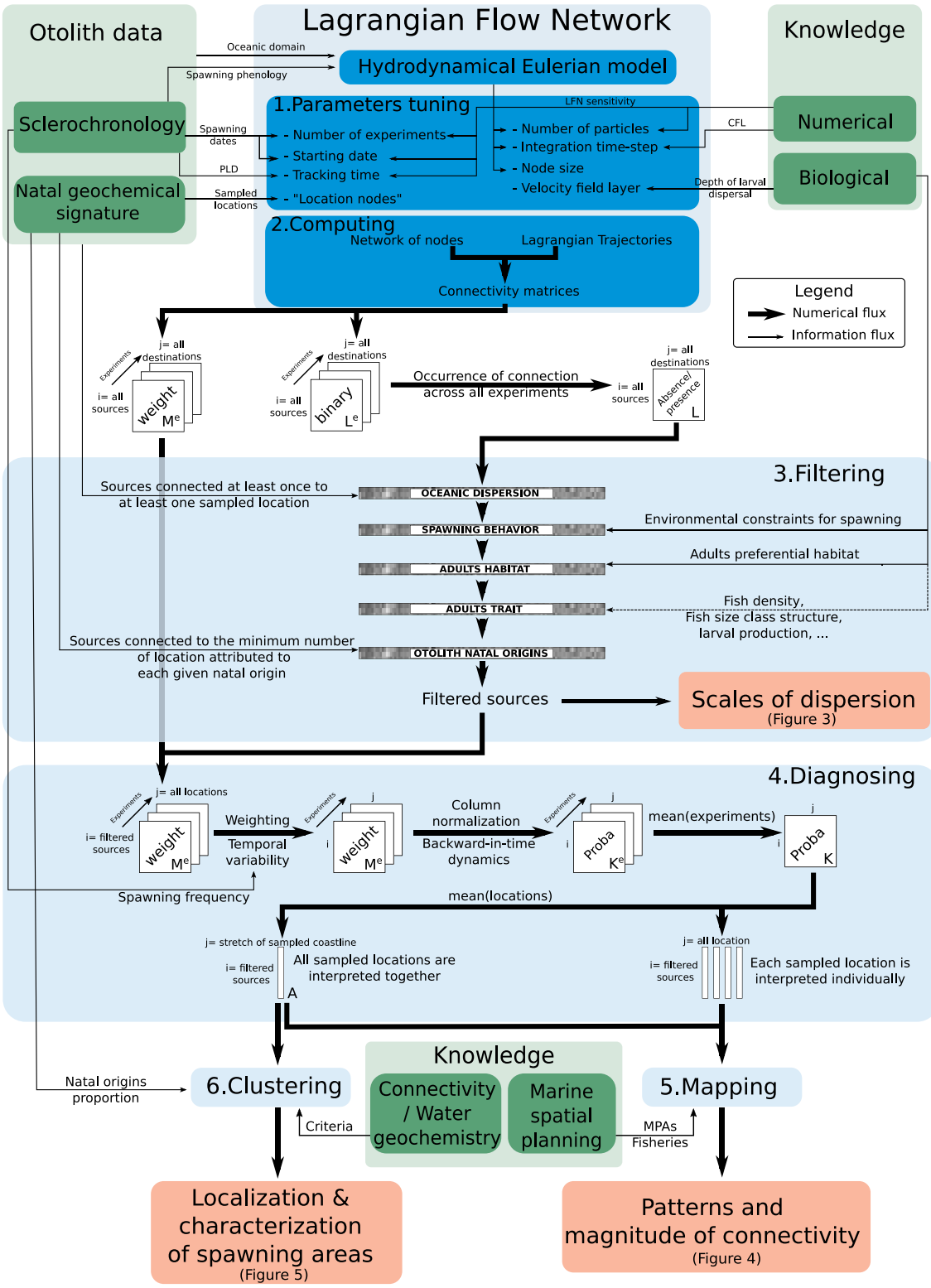
Dispersal has been identified as a crucial ecological and evolutionary process which influences demography, population survival, gene flow and local adaptation (Burgess et al., 2016; Lowe et al., 2017). The persistence and dynamics of marine populations are controlled, in addition to local birth and death rates, by connectivity processes (Kool et al., 2013). Connectivity has been defined as the exchange of individuals among geographically separated populations that comprise a metapopulation (Cowen et al., 2006).

Studying dispersal, and more generally the closely-related concept of connectivity, is thus essential to understand population structuring, a prerequisite for the management and conservation of marine ecosystems. Potential applications of connectivity studies range from informing design of local Marine Protected Areas (MPAs) to replenish fished neighboring areas (Di Franco et al., 2012a; Pelc et al., 2010; Pujolar et al., 2013) and informing maritime spatial planning (Bray et al., 2017; Henry et al., 2018) to improve the design of broad-scale MPAs network and assess their efficiency (Dubois et al., 2016; Rossi et al., 2014).

For coastal fishes, connectivity is assumed to be primarily driven by the dispersion of early-life stages ("propagules") which are transported by ocean currents, a process called "larval dispersal" (Pineda et al., 2007). Indeed, most coastal fishes are characterized by a bipartite life

* Corresponding author.
E-mail address: terence.legrand@mio.osupytheas.fr (T. Legrand).

https://doi.org/10.1016/j.marenvres.2019.104761
Received 5 April 2019; Received in revised form 12 June 2019; Accepted 15 July 2019
Available online 19 July 2019
0141-1136/



(caption on next page)

Fig. 1. Schematic representation of our integrative framework which combines Lagrangian modelling and expert knowledge (including otolith analyses) in order (i) to investigate the scales of dispersion, (ii) to determine the patterns and magnitude of connectivity and (iii) to locate and characterize spawning areas for conservation purposes. Our framework is structured according to the color code: blue annotations refer to the numerical methodologies, green annotations symbolize expert knowledge (from existing bibliography) and red boxes highlight the key results. PLD stands for Pelagic Larval Duration and CFL for Courant-Friedrichs-Lewy (see section 2.2) for a definition of all the terms used). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

cycle, governing how fish populations structure in space and time. A first highly pelagic dispersive phase, encompassing egg and larval stages, is followed by a second relatively sedentary phase as juveniles and adults (Thresher et al., 1989). The transition occurs when individuals take up permanent residence in the demersal habitat, a process named settlement (Levin, 1994), which usually coincides with the metamorphosis of larvae into juveniles (Kingsford, 1988). Then, after a short period, juveniles recruit into the adult fraction of the population (Richards and Lindeman, 1987), which generally constitutes the stages of exploitation and managerial interests.

Even though spawning and subsequent larval dispersal constitute the cornerstones of the fish life cycle, spawning aggregations (areas where fishes gather in high density to reproduce, Domeier and Colin, 1997) can be massively harvested by fisheries, sometimes inducing collapses in aggregating fish stocks (Sadovy and Domeier, 2005) and ultimately leading to the risk of species extinction (Mitcheson et al., 2013). Furthermore, spawning aggregations can be seen as productivity hotspots of ecological significance since they support both coastal and offshore trophic chains (Fuiman et al., 2015; Heithaus et al., 2008).

It has been recently suggested that scientists and managers should focus on spawning areas since their protection could lead to large benefits for fisheries, ecotourism stakeholders and biological conservation (Erisman et al., 2017). Nevertheless, the precise locations where eggs are spawned are actually one of the main unknown of the fish life-cycle since studies are rarely dedicated solely at identifying spawning areas (e.g. Calò et al., 2018). The only few spawning sites that have been well documented are situated mainly in tropical reef systems (Domeier and Colin, 1997; Russell et al., 2014), perhaps where connectivity studies are facilitated by the relatively small-size and closeness of the seascape. It contrasts with the openness and continuous coastlines of temperate systems, such as the Mediterranean Sea, where spawning areas are largely unknown. In addition, while knowledge about fish home range exists (Di Franco et al., 2018), the way it is usually defined disregards the reproductive movements, despite observational evidence that many coastal fishes may move offshore during the spawning season (Aspillaga et al., 2016; Giacalone et al., 2018).

In a context of anthropogenic and environmental perturbations (Ciannelli et al., 2013), a major challenge is thus to provide scientifically-based information about spawning areas in all oceanic systems to achieve a sound spatial management and protection of coastal fishes. However, this objective has been hampered by the complexity of fish life-cycles, the paucity and disparity of observational records and the absence of a unifying methodology to do so. One way to investigate where fishes spawn is to build on available information on larval dispersal. Many methods, each with its own strengths and weaknesses, have been used to investigate the ins and outs of larval dispersal (Calò et al., 2013). They are principally divided into four categories: surveys of marine larvae, artificial tags, natural tags (genetics, otolith sclerochronology, otolith geochemical analyses) and numerical modelling (physical or bio-physical, Nolasco et al., 2018). To increase the accuracy of larval dispersal analyses, it seems fundamental to use a combination of complementary methodologies, maximizing the strengths of each technique. However, while the proportion of works that mixed different methodologies is currently growing, only few studies have used at least two methods to characterize larval dispersal, with a preferential combination of numerical biophysical models with genetic markers and/or otoliths analysis (Nolasco et al., 2018).

Most of these multi-disciplinary studies aimed to locate the potential destinations of propagules by means of forward-in-time advection from pre-determined hypothesized sources, such as MPAs or sampling sites (Bray et al., 2017; Carlson et al., 2016; Di Franco et al., 2012a; Melià et al., 2016; Pujolar et al., 2013). Di Franco et al. (2012b, 2015) have investigated connectivity for two coastal fish species using otolith techniques; their results set the lower limits of dispersal and distinguish several natal origins for both fishes, but there was little information about regional connectivity patterns and no indication at all of the geographical extension and localization of these spawning areas.

In this paper, we propose a flexible analytical framework to delineate and characterize fish spawning areas by integrating model simulations, information from otolith studies and biogeographical knowledge. We test our framework by exploiting two data-rich case-studies focusing on the coastal fishes *Diplodus sargus sargus* (Linnaeus, 1758) and *Diplodus vulgaris* (Geoffrey Saint Hilaire, 1817) in the Adriatic sea.

2. Materials and methods

2.1. General description

Our analytical framework consists of dissecting, backward-in-time, the main steps of dispersal, namely the “origins”, the “transport” process and the “destinations”. “Destinations”, i.e. settlement sites, are predetermined by the locations where juveniles were collected for the otolith studies. A Lagrangian bio-physical model, parameterized with early-life traits of the target species (obtained from otolith schlerochronology data), is used to assess “transport” (“Lagrangian Flow Network” box in Fig. 1). Concerning the “origins”, one of the main novelty is to consider the whole oceanic domain of interest as a mosaic of potential larval sources. These are refined stepwise by biological knowledge (including otolith geochemical data) of our target species (“Filtering” box in Fig. 1). It allows us to assess larval dispersal scales (“Scales of dispersion” box in Fig. 1) in a robust manner thanks to millions of simulated larval trajectories. Normalizations procedure allows computing the proportion of each source to the total larval pool supplied into pre-determined destinations (“Diagnosing” box in Fig. 1). These probabilities are then exploited to quantitatively assess the connectivity induced by larval dispersal (“Patterns and magnitude of connectivity” box in Fig. 1). Then, by incorporating additional ecological and connectivity knowledge, we fit simulated larval sources with otolith geochemistry to better locate and evaluate the relative importance of discrete spawning areas (“Localization & characterization of spawning areas” box in Fig. 1) for the replenishment of the sampled locations.

2.2. Lagrangian Flow Network: tuning parameters and computing

The Lagrangian Flow Network (LFN) methodology combines network theory tools and particle-tracking modelling to investigate transport and dispersal processes in oceanic flows. As most off-line particle tracking model, it can be coupled to any gridded two- or three-dimensional velocity fields available, returning dispersal diagnostics as realistic as the input flow field. Full description can be found in Rossi et al. (2014); Ser-Giacomi et al. (2015a, b) and Dubois et al. (2016). Here, the LFN simulates the dispersal of passively drifting larvae as horizontal Lagrangian trajectories obtained through the integration of a

high-resolution flow field generated by a regional hydrodynamical model.

The ad-hoc LFN configuration is obtained by selecting the most adequate hydrodynamical model and thanks to the fine-tuning of seven LFN parameters ("Parameters tuning" box in Fig. 1) in accord with both biological (including the information derived from the otolith analyses) and numerical knowledge from the scientific literature. Only the most relevant elements are summarized hereafter (see SI A-1.1 for further information). The starting date of each numerical experiment e is simulating a single spawning event while the ensemble of Lagrangian experiments E has to cover the full range of spawning dates obtained from otolith sclerochronology. Tracking time is set to mimic the Pelagic Larval Duration (PLD; i.e the time larvae spend in plankton), which is also estimated from otoliths sclerochronology. The network of nodes (i.e. sub-areas of the discretized oceanic domain) resolution is adjustable, compromising both the level of analyses and the computation time. It must be at least twice larger than the spatial resolution of the velocity field given by the hydrodynamical model. Note that each node has the same area and contains the same initial number of particles. The number of particles per node should be larger or equal to 100 particles, as prescribed by Monroy et al. (2017) and the Runge-Kutta time step should fulfill the Courant-Friedrichs-Lewy (CFL) condition imposed by the velocity field itself (Courant et al., 1928). The vertical layer of the velocity field (for z-coordinates model) must be chosen according to the most probable depth at which larvae of the target species are more likely to be found. Note that our Lagrangian trajectories are currently 2-dimensional, without considering any larval diel migration. We attributed one or more *location nodes* in the close vicinity of each sampled location (i.e. where juveniles were collected for otolith studies).

After computing hundred million Lagrangian trajectories and recording the initial and final positions of each particle, a connectivity matrix of particles is built similarly for each numerical experiment $e \in [1; E]$ ("Computing" box in Fig. 1, see SI A-1.2 for further information). We saved an ensemble of E connectivity matrices, which have as many rows and columns as the total number of nodes N in the network and which contain all information about dispersal. Each matrix element indexedij characterizes the connection between the origin node $i \in [1, N]$ and the destination node $j \in [1, N]$ from a given starting date and during a fixed tracking time.

The connection between any pairs of nodes can be characterized at two levels: the binary link (presence/absence of connection, L^e matrices) and its weight (the number of transported particles associated to the existing link, M^e matrices Ser-Giacomi et al., 2017). The binary links are used to successively filter-out the putative origin nodes in order to locate the effective sources of larvae (see section 2.3). The associated weights are used to compute probabilities of larval emission from the effective sources to the surveyed locations (see section 2.4).

2.3. Filtering

We aim at determining the most relevant sources, among all the putative origins, of the larvae whose trajectories ended in the pre-determined *location nodes* ("Filtering" box in Fig. 1). The effective sources of larvae are investigated by applying successively restrictive filters over the full set of potential origin sources, sequentially refining our characterization. The flexibility of our framework allows us to add or skip any given filter depending on the level of knowledge of the studied species and on the confidence to be attributed to each piece of information. It implies that the more information we have about the biological traits of the studied species ("Biological knowledge" and "otolith data" boxes in Fig. 1), the more precise and realistic are the final larval sources. Different filtering layers are defined hereafter and subsequently applied to the suite of connectivity matrices:

- Among all potential sources (i.e. all nodes), the first filter selects those which are connected, at least once over all experiments (i.e. over all spawning events) to at least one out of all sampled locations (i.e. one of all *location nodes*). Note that this first filter, based only on the cumulative binary matrix L, is a strong constraint imposed by the geographical extent of the sampled locations. When mapping those selected nodes, it gives us the maximal and theoretical extent of all sources (see SI A-2 for further information).
- The second step filters in all putative sources that are favorable for spawning based on the best available knowledge concerning spawning behavior of the studied fish. This could be derived from any environmental criteria triggering spawning (e.g. threshold of temperature, light, etc) or any constraint restricting spawning (e.g. bathymetric limit, see section 2.6.3).
- The third step filters in all putative sources whose environmental characteristics are suitable for adults (e.g. preferential habitats). Indeed, the overall contribution of adults to broad-scale dispersal can be reasonably neglected since these stages are rather territorial, especially littoral fishes that show strong site fidelity (Di Franco et al., 2018).
- One may add as many additional filters as possible to further refine the characterization of the putative larval sources. These extra filters must be spatialized dataset derived from the finest biological knowledge of the target species. For instance, spatial information about fish density, sex-ratio and size class structure, which, in combination with a female size/eggs production relationship (Marshall et al., 2019), could further constrain the initial larval production of each source node.
- The last filter selects those larval sources whose downstream connections (e.g. forward-in-time dispersal) are concordant with the diversity of origins revealed by otolith geochemistry. In other words, larval sources must send larvae to at least the minimal number of sampled location successfully replenished by the less-ubiquitous fish natal origins, as assessed from otolith geochemical analyses.

After the superposition of all filters, the remaining "origin" nodes are the most likely larval sources of the pre-determined sampled locations.

2.4. Diagnosing

The filtered larval sources are analyzed from the weighted connectivity matrices M^e to quantitatively characterize, as explained hereafter, the probabilities of connection with all sampled locations ("Diagnosing" box in Fig. 1, see SI A-3 for further information).

First, to take heed of the temporal heterogeneity of spawning, our ensemble of experiments E must simulate the observed spawning variability. This is achieved by defining a new matrix $\overline{M^e}$ for each experiment which modulates each original connectivity matrix M^e, representing one spawning event, with a weight p^e prescribed by its corresponding spawning dates frequency obtained from the otolith analyses.

To be interpreted as probabilities, connectivity matrices of particles must be normalized. When looking for the probabilities of connection from *location node j* to filtered source node *i*, we approximate backward-in-time dynamics by a column-normalization of each connectivity matrix, even though it has been originally computed forward-in-time (Ser-Giacomi et al., 2015a, b). These backward-in-time "upstream" probabilities are interpreted as follows: if we randomly select a particle settled in any *location node j* after the tracking time (PLD), this particle has a probability $\in [0,1]$ to originate from node i. Therefore, this probability measures the relative contribution of that source node i to the total pool of larvae sent by all sources which successfully settled in *location node j*. The numerical experiments are merged together by averaging all E connectivity matrices into one final probabilistic matrix denoted as K which summarizes all the relevant connectivity information. By selecting in this final matrix K the row i, corresponding to the index of any filtered sources, and the column j, corresponding to the

indices of all *location nodes*, one can investigate the probability of connection for all existing pairs of source node and *location node*. In addition, one can consider all *location nodes* together by averaging the probabilities of their corresponding nodes (column indexed *j*) into one vector called *A*. Here we mainly exploit this configuration as it allows to take a global perspective of the surveyed area, which is compatible with the next clustering step.

2.5. Clustering

Besides the characterization of larval sources described in previous section 2.4, our final objective is to pinpoint spawning areas in our simulations as constrained by the results of otolith geochemistry. As such, once the most relevant sources have been selected through the successive filters and their transport probabilities characterized ("Mapping" box in Fig. 1), we aim here at grouping these sources into several regional clusters whose relative contributions for supplying larvae to the surveyed locations would match the natal origins revealed by otolith geochemistry ("Clustering" box in Fig. 1). Determining pertinent criteria and numerical algorithm for clustering is not a trivial issue (e.g. Fortunato, 2010; Rossi et al., 2014; Ser-Giacomi et al., 2015a). We retain here five criteria which are based on empirical and published knowledge that are used to gather the full set of source nodes into several clusters as follows:

- The number of clusters must be equal to the number of natal origins deduced from otolith geochemical analyses.
- Each cluster must be composed of geographically contiguous or almost contiguous source nodes, fulfilling the assumption that seawater geochemistry is homogeneous at small-scale (e.g. within a given water mass).
- Each cluster should have a maximal extension of about 400 km since it is the typical length-scale at which the geochemical composition of seawater is supposed to vary substantially, hence conferring distinct geochemical signatures in otoliths (Gibb et al., 2017).
- Any relevant connectivity information about the spatial differentiation of fish population must be considered in the cluster delimitation (e.g. through genetics, tagging and tracking techniques; Calò et al., 2013).
- The precise delimitation of each simulated cluster can be further refined by considering the observed proportions of natal origins derived from otolith geochemistry. In other words, the fine-tuning of the clusters' boundaries must be done while maximizing the correlation between natal origin proportions and the aggregated probabilities of larval sources.

2.6. Case-studies

We exploit two data-rich case-studies to test our framework and its effectiveness for locating spawning areas and for evaluating connectivity patterns of both *Diplodus sargus* and *Diplodus vulgaris* in the Adriatic Sea (Fig. 2).

2.6.1. Oceanic domain and hydrodynamical model

The Eulerian gridded velocity field comes from the high-resolution (1/45 °) Adriatic-Ionian REGional configuration (AIREG, Ciliberti et al., 2015; Oddo et al., 2006), which is based on the NEMO kernel (Madec and others, 2015) and has been developed by the CMCC Ocean Lab (Fig. 2a, see SI B-1 for further information).

2.6.2. Otolith data

Post-settlers (i.e. 1–1.5 cm body length) of *D. sargus* and *D. vulgaris* were collected at seven distinct sampled locations separated by 10–30 km, corresponding to about 180 km stretch of the Apulian coast (Fig. 2b; Di Franco and Guidetti, 2011; Di Franco et al., 2013, 2012b, 2015). Data gathered from otolith sclerochronology (Di Franco and Guidetti, 2011; Di Franco et al., 2013) and geochemical analysis (Di Franco et al., 2012b, 2015) are reported in Table 1 (see SI B-3.1).

2.6.3. Biological traits

Knowledge of biological traits is generally sparse and uncertain for most fish species. Concerning our case-studies, quite reliable information is however available for both studied species. Recent reports of the depths at which *D. sargus* and *D. vulgaris* commonly spawn range between 0 and 80 m deep (Aspillaga et al., 2016; Giacalone et al., 2018). Moreover, larvae of *D. sargus* have been mainly observed at around 10 m depth (Olivar and Sabatés, 1997). Both species have similar adult sedentary behavior, with individuals home ranges that are typically smaller or equal to 1 km^2 (Alós et al., 2012; Di Lorenzo et al., 2014; Di Franco et al., 2018). Concerning the preferential habitat of adults, *D. sargus* and *D. vulgaris* inhabit coastal rocky reefs, *Posidonia oceanica* meadows and coralligenous formations (Guidetti, 2000; Harmelin-Vivien et al., 1995; Lenfant and Planes, 1996).

High-resolution maps of these fish preferential habitats over the whole Adriatic were obtained from seabed habitat maps downloaded on the EMODnet portal. They were then hand-corrected using 2-steps procedures (see SI B-3.2). To our knowledge, this improved map represents perhaps the best and most updated geo-referenced seafloor cartography for the Adriatic Sea.

2.6.4. Tuning model parameters from published knowledge

Biological information derived from otolith analyses and literature review on biological traits for *D. sargus* and *D. vulgaris* are used to fine-tune the LFN parameters introduced in section 2.2 (see SI B-4). The parameters retained for this case-study are reported in Table 2. For each connectivity matrix, the LFN model builds a network of around 10000 quasi-rectangular oceanic nodes of 1/16° side (that is about 7 km) and computes around 9 million trajectories. In total, about 380 million of particles trajectories were calculated to produce 43 high-resolution connectivity matrices for both species.

3. Results

3.1. Scales of dispersion

The first filter returns the maximal and theoretical extent of all the larval sources, as constrained by the oceanic circulation experienced during larval drift. For *D. sargus*, filter-1 larval sources spread from 43.2°N to 40.2 °N along the Italian coast (Fig. 3a). Median dispersal distance is 175 km, maximal dispersal distance is 355 km and the "backward" dispersal plume surface is 35662 km^2 (see Tables 3 and SI B-5). For *D. vulgaris*, putative larval sources extend from 44.2 °N to 36 °N in almost all the Adriatic Sea and the northern Ionian Sea (Fig. 3b). The maximal dispersal distance median is 1.7 times higher than *D. sargus*'s one whereas the median distance is quite similar (ratio of 1.0). The dispersal plume surface is 233199 km^2, around 6.5 times larger than *D. sargus*'s one (Table 3).

The second filter selects all larval sources whose environmental characteristics fulfill the favorable conditions for spawning to occur. Here it is a "coarse" bathymetric filter selecting all nodes whose depths range 0–80 m, that is the depths at which *D. sargus* and *D. vulgaris* commonly spawn (see section 2.6.3). For *D. sargus*, filter-2 larval sources extend from 42.5°N to 40.2 along the Italian coastline (Fig. 3a). The dispersal plume surface is 12918 km^2, which corresponds to a reduction of the filter-1 plume by about 64% (Table 3). For *D. vulgaris*, sources spread from 43.8 °N to 37.2 °N (Fig. 3b). The dispersal plume surface is 41606 km^2, indicating a reduction by 82% from filter-1 plume (Table 3). Median and maximal dispersal distances as well as dispersal plume surface are higher for *D. vulgaris* than for *D. sargus*, with a ratio of 1.1, 1.5 and 3.2, respectively (Table 3).

The third filter display all larval sources which house adults preferential habitats, that is rocky reefs, *Posidonia oceanica* meadows and

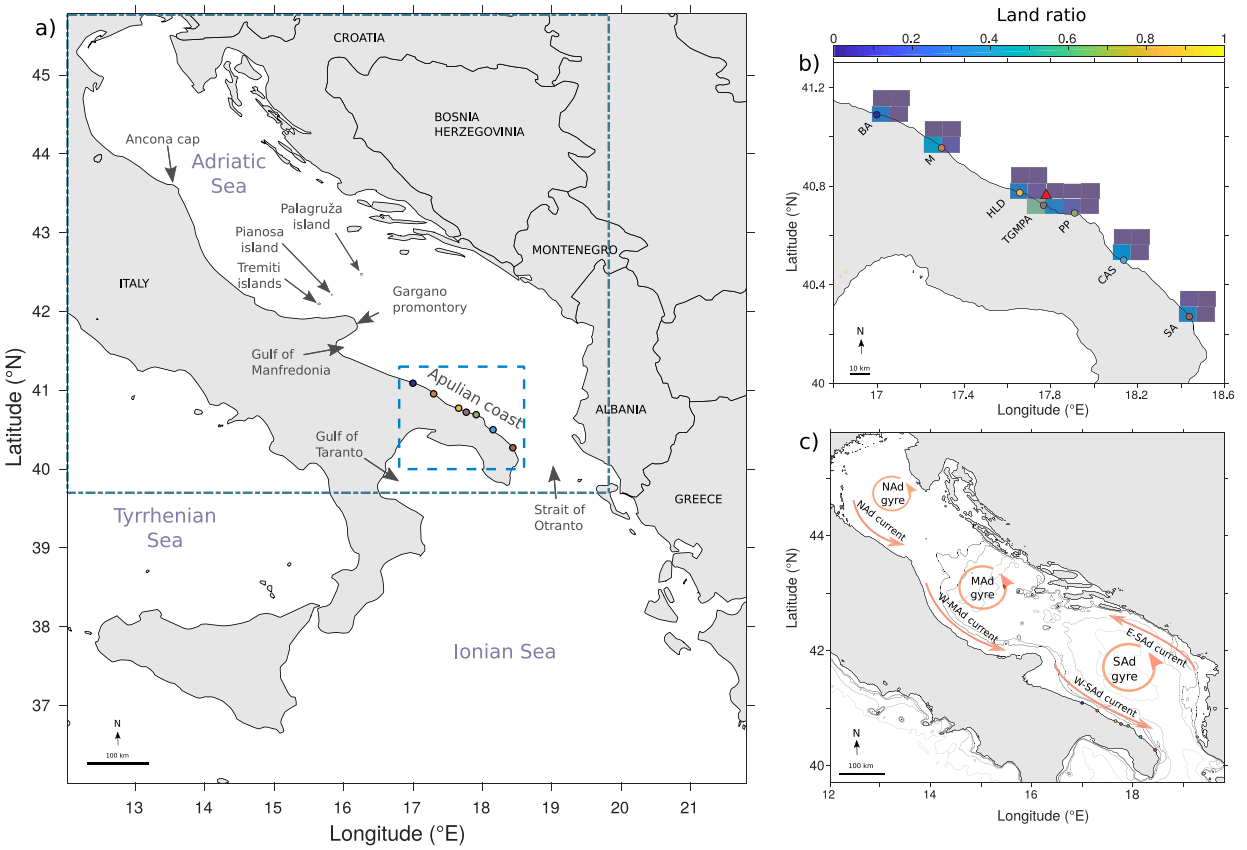


Fig. 2. Geographical location, topography and schematic circulation of the study area. Panel a) displays the network domain (36 °- 46 °N, 12 °E − 22 °E). Panel b) discloses a zoom over the Apulian coast and indicates the seven locations (colored dots) named, from north to south: Bari (BA), Monopoli (M), Hotel La Darsena (HLD), Torre Guaceto Marine Protected Area (TGMPA), Punta Penna (PP), Casalabate (CAS) and San Andrea (SA). *Location nodes*, associated with each sampled location, are colored according to their land ratio (see SI A-1.1). The red triangle corresponds to the barycenter of all *location nodes* centers. Panel c) presents a schematized view of the mean surface circulation of the Adriatic Sea (red arrows). The system is dominated by the Western South Adriatic (W-SAd) current, the Eastern South Adriatic (E-SAd) current, the Western Middle Adriatic (W-MAd) current, the Northern Adriatic (NAd) current, the South Adriatic (SAd) gyre, the Middle Adriatic (MAd) gyre and the North Adriatic (NAd) gyre (adapted from Artegiani et al., 1997b; Millot and Taupier-Letage, 2005). The thin grey lines in panel c) corresponds to isobaths 80 m, 100 m, 200 m and 1000 m, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Summarized information and early-life traits gathered for *D. sargus* (Di Franco and Guidetti, 2011; Di Franco et al., 2012b) and *D. vulgaris* (Di Franco et al., 2013, 2015), as obtained from sclerochronology and geochemical analyses of their otoliths.

Early-life traits	Species	
	D. sargus	*D. vulgaris*
Number of samples	140	160
Period of spawning	04/05/2009–24/05/2009	20/10/2009–14/02/2010
Median PLD (days)	17±1	47±8
Number of natal origins	3	7

Table 2
Summary of species-specific numerical parameters (and their corresponding early-life traits) for *D. sargus* and *D. vulgaris*. The common parameters are a node size of 1/16 °, the selected vertical layer of the flow at 10 m deep, a Runge-Kutta time step of 20 min and an initialization of 900 particles per node.

Numerical parameters	Species	
	D. sargus	*D. vulgaris*
Range of starting dates ~ *spawning duration*	04/05/2009–24/05/2009	05/11/2009–28/01/2010
Starting dates interval (days) ~ *spawning periodicity*	1	4
Tracking time (days) ~ *PLD*	17	47
Number of experiments ~ *spawning events*	21	22

coralligenous formations in our case-studies (see section 2.6.3). For *D. sargus*, filter-3 larval sources extend from from 42.2 °N to 40.2 °N (Fig. 3a). The dispersal plume surface is 3700 km^2, corresponding to a reduction by 71% of the filter-2 surface (Table 3). For *D. vulgaris*, larval sources expand from 43.6 °N to 38.8 °N (Fig. 3b). The dispersal plume surface is 11160 km^2, implying a reduction by 73% of the filter-2 surface (Table 3). The maximal dispersal distance and the dispersal plume of *D. vulgaris* are respectively 1.9 and 3.0 times higher than those of *D. sargus*. The median dispersal distance for *D. sargus* is 1.5 higher than for *D. vulgaris* (Table 3).

The fourth filter selects the larval sources which are connected to the minimum number of sampled locations successfully replenished by the less-ubiquitous fish natal origins, as assessed from otolith

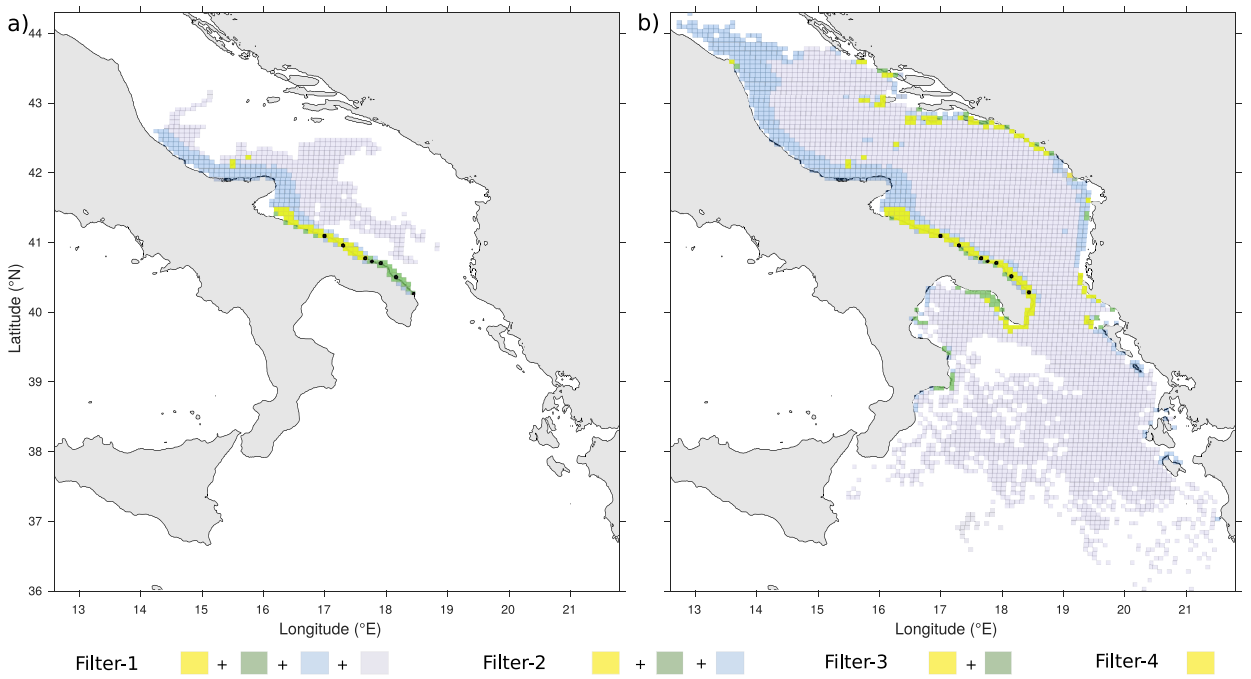


Fig. 3. Spatial scales of dispersion as given by the characterization of larval sources by successive filtering procedure for *D. sargus* (panel a) and *D. vulgaris* (panel b). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

geochemical analyses. In other words, the fourth filter selects the larval sources which send larvae to at least "x" different sampled locations, "x" being determined by the minimum number of sampled locations in which the rarest fish natal origins was found. Here, "x" is four for *D. sargus* and three for *D. vulgaris* (see SI B-3.1). In our case-studies, this last filter provides the most refined sources which take into account all constraints imposed by realistic oceanic dispersal as well as observation-based biological knowledge. For *D. sargus*, filter-4 larval sources expand from 42.2 °N to 40.8 °N (Fig. 3a), located especially along the Apulian coast and around the Tremiti archipelago (composed of the main island and the isolated Pianosa islet; see Fig. 2a). The dispersal plume surface is now 2183 km^2, that is a reduction by 41% of the filter-3 plume (Table 3). For *D. vulgaris* sources spread from 43.6 °N to 39.6 °N on both Adriatic shores. Note that some sources are located within the gulf of Taranto and one noticeable larval source remains in the middle of the Adriatic Sea, next to the Palagruža island at 42.4 °N, 16.3 °E (Fig. 3b). The dispersal plume surface is 7884 km^2, corresponding to a reduction by 29% of the filter-3 plume. The maximal dispersal distance and the dispersal plume of *D. vulgaris* are higher than those of *D. sargus*, with a ratio of 1.9 and 3.6, respectively. The median dispersal distance for *D. sargus* is 1.5 higher than for *D. vulgaris*.

Altogether, the superposition of the four filters has reduced the dispersal plume surface by 94% for *D. sargus* and by 97% for *D. vulgaris*. When comparing three quantitative dispersal diagnostics between both species, the mean ratios of *D. vulgaris* to *D. sargus* are 0.9 ± 0.2 for median distances, 1.7 ± 0.2 for maximum distances and 4.1 ± 1.6 for dispersal surfaces (Table 3).

3.2. Patterns and magnitude of connectivity

The flexibility of our approach allows us to map and analyze the probabilities of all source nodes, i.e. the probability that larvae whose dispersal ended up in one of the seven sampled locations originated from that source node, for each levels of filtering. We chose here to exploit mainly the most constrained larval sources returned by the superposition of the four filters (Fig. 4, see SI B-6 for the other filters).

Table 3
Quantitative evaluation of the dispersal scales of *D. sargus* and *D. vulgaris* for each filter successively applied. Median/maximal dispersal distance corresponds, respectively, to the median/longest distance measured between all source nodes and the barycenter of the seven sampled locations (see SI B-5). The dispersal plume surface corresponds to the total surface of all nodes acting as putative larval sources.

Species	Quantitative diagnostics	Filter-1	Filter-2	Filter-3	Filter-4
D. sargus	Median dispersal distance (km)	175	185	135	135
	Maximal dispersal distance (km)	355	355	245	245
	Dispersal plume surface (km^2)	35662	12918	3700	2183
D. vulgaris	Median dispersal distance (km)	170	210	90	90
	Maximal dispersal distance (km)	590	550	470	470
	Dispersal plume surface (km^2)	233199	41606	11160	7884

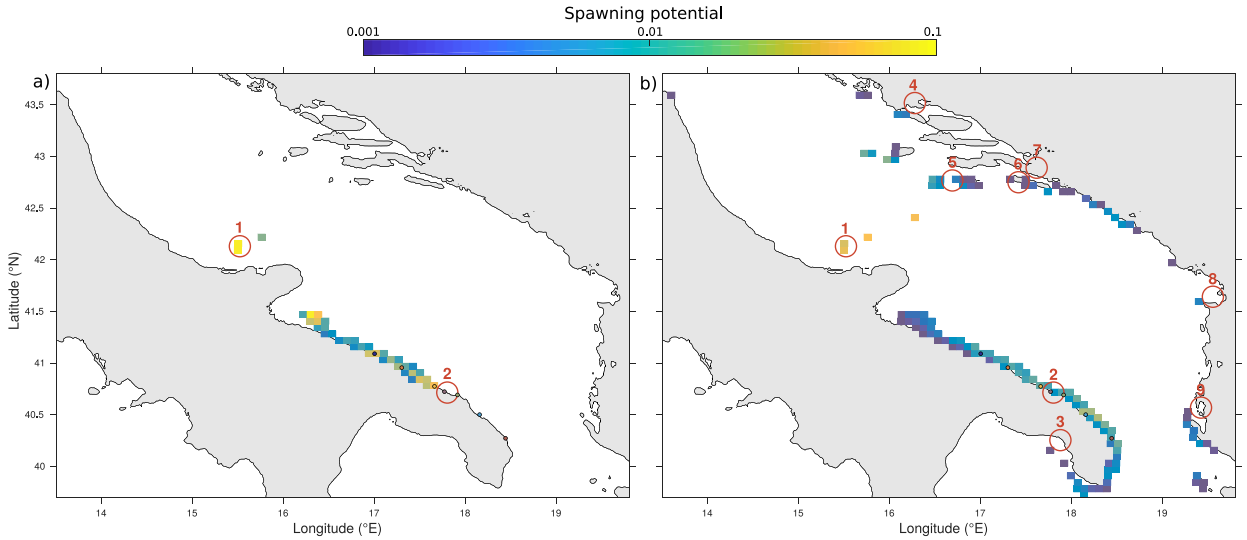


Fig. 4. Quantitative patterns of connectivity for *D. sargus* (panel a) and for *D. vulgaris* (panel b): maps of the most constrained larval sources (after applying the superposition of the four filters), along with their associated probabilities (the probability of that node to supply larvae to the surveyed stretch of Apulian coastline) which are integrated here as the larval source's spawning potential. Red circles represent the current Marine Protected Areas (MPAs) which have connection with the sampled locations: (1) Tremiti islands, (2) Torre Guaceto, (3) Porto Cesaro, (4) Pantana, (5) Lastovo archipelago, (6) Mljet, (7) Mali Ston Bay, (8) Lumi Buna-Velipoje and (9) Karaburun-Sazani Island (see SI B-7). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Note that this larval source probability are integrated here as the larval source's spawning potential. For *D. sargus*, the core larval sources are located along the Apulian coast and around the Tremiti archipelago. Sources situated next to the main islands of the Tremiti archipelago are characterized by high probabilities, i.e. spawning potential (0.1). The other isolated source located next to the isolated Pianosa island has moderate spawning potential ($\sim$ 0.01). Sources along the Apulian coast are characterized by a wide range of spawning potential spanning 0.001–0.1. Three locals maximal, with spawning potential ranging from 0.01 to 0.1, can be distinguished at around 41.5 °N, 41.1 °N and 40.7 °N. The other sources have moderate to low spawning potential of the order of 0.001–0.01. For *D. vulgaris*, the core larval sources are located from 43.6 °N to 39.6 °N on both Adriatic shores (see section 3.1). Most probable sources (spawning potential of 0.1) are around the Tremiti archipelago and the Palagruža island (42.4 °N, 16.3 °E). Sources of intermediate spawning potential ($\sim$ 0.01) are found along the Apulian coast (from 41.2 °N to 40.2 °N), on the eastern part of the gulf of Taranto (around 39.9 °N, 18.1 °E), next to the Croatian islands (from 43 °N to 42.5 °N), along the southern Croatian coastlines (42.3 °N), and along both northern (41.5 °N) and southern Albanian coastlines (40.2 °N). The remaining sources, i.e. near Ancona cape (43.5 °N, 13.8 °E) and along the northern Apulian coast, are associated to low spawning potential ($\sim$ 0.001).

We next analyze, for both species, the strengths of the existing connections between MPAs and the sampled locations. It is equivalent to assess the proportions of larvae supplied by the local MPA network to the sampled locations (Fig. 4, see SI B-7). Our model simulations suggest that two MPAs for *D. sargus* (Fig. 4a) and nine MPAs for *D. vulgaris* (Fig. 4b) are supplying larvae to the sampled locations. The spawning potential associated to each MPA, interpreted here as proportions of the total pool of larvae supplied to the sampled locations, is reported in Table 4. Tremiti MPA emerges as the most important larval supplier, providing 26.4% of all *D. sargus* larvae and 7.5% of all *D. vulgaris* larvae settled along this stretch of Apulian coastline. By summing up probabilities, the total contribution of the local network of MPAs to the total pool of larvae settled in the sampled locations is 30.8% for *D. sargus* and 23.6% for *D. vulgaris*.

Table 4
MPAs associated proportions for *D. sargus* and *D. vulgaris*, indicative of their relative contribution to the total larval pool supplied to the sampled stretch of Apulian coastline. Note that the proportions of larvae supplied by each MPA into each location is given in SI B-7.

MPAs		Species (%)	
Index	Name	*D. sargus*	*D. vulgaris*
1	Tremiti islands	26.4	7.5
2	Torre Guaceto	4.4	7.5
3	Porto Cesaro	–	0.1
4	Pantana	–	0.9
5	Lastovo archipelago	–	5.6
6	Mljet	–	0.5
7	Mali Ston Bay	–	0.4
8	Lumi Buna-Velipoje	–	0.5
9	Karaburun-Sazani Island	–	0.6
–	All	30.8	23.6

3.3. Localization and characterization of spawning areas

To pinpoint the most robust and realistic spawning areas, we exploit the flexibility and statistical power of the LFN model while we constrain the grouping procedure to match the natal origins revealed by otolith geochemistry. Using criteria defined in section 2.5, filter-4 larval sources are grouped into three and seven clusters (e.g. “coherent spawning regions”) for *D. sargus* and for *D. vulgaris*, respectively (Fig. 5a and b). For both species, one cluster embraces the larval sources around the Tremiti archipelago while the sources located along the Apulian coast are separated into two clusters by a boundary that we imposed to both species. Indeed, this boundary was fixed for *D. sargus* by following the last clustering criterion (that is to maximize the fit between model estimations and otolith geochemical analyses, see SI B-8); the same boundary is then used for *D. vulgaris* assuming water geochemistry background is the same for both species. Concerning the four other spawning areas of *D. vulgaris*, one cluster corresponds to the northernmost source located next to Ancona cap and three other clusters

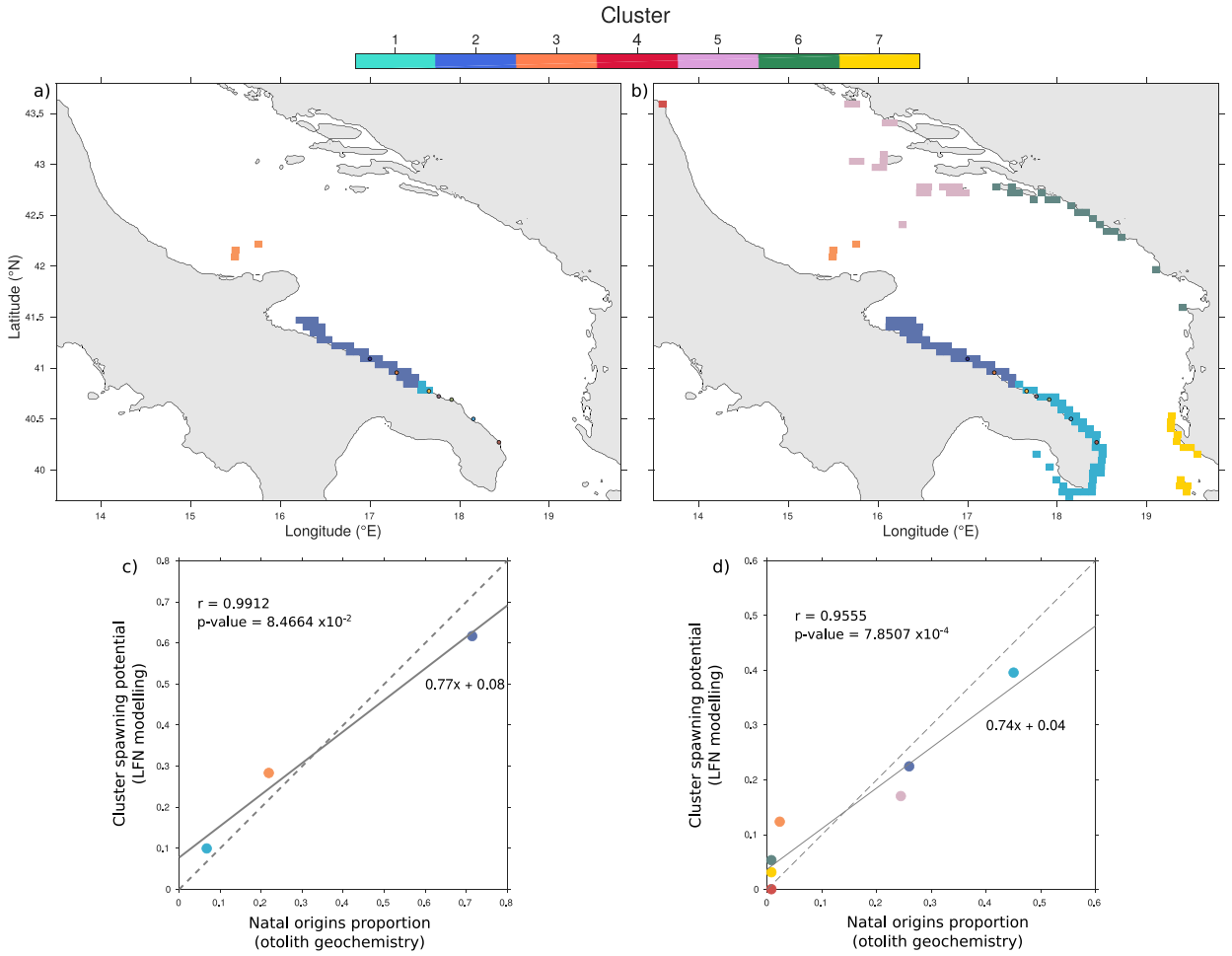


Fig. 5. Spawning areas geographically-delimited thanks to criteria defined in section 2.5 and characterization of their relative contributions to the total pool of larvae settled in the Apulian coast. Clusters of the most constrained larval sources (panels a, b) and correlations between *simulated* larval contributions and *observed* natal origin proportions (panels c, d) for *D. sargus* (panels a, c) and *D. vulgaris* (panels b, d).

represent subgroups of all the source nodes spread along the eastern shores.

We report the relative larval supply proportions attributed to each cluster by summing up the probabilities, i.e. spawning potential, of all the source nodes pertaining to a given cluster. For *D. sargus*, cluster 2 is characterized by the highest spawning potential (0.61), cluster 3 has intermediate spawning potential of 0.29 and cluster 1 exhibits the lowest contribution (0.1; Fig. 5a,c). For *D. vulgaris*, cluster 1 has the highest spawning potential (0.4) whereas clusters 4, 6 and 7 are characterized by the lowest spawning potential (> 0.01, 0.05, 0.03, respectively). Clusters 2, 3 and 5 have moderate contributions of 0.23, 0.12 and 0.17, respectively (Fig. 5b,d).

We analyze the correlations between these cluster spawning potential and the natal origin proportions derived from the otolith geochemical analyses across all sampled locations. Pearson's coefficient *r* is 0.9912 with a *p*-value of 8.4664 x 10^{-2} for *D. sargus* and *r* is 0.9555 with a *p*-value of 7.8507 x 10^{-4} for *D. vulgaris* (Fig. 5c and d). Note that the largest number of natal origins for *D. vulgaris* provides a more robust correlation for *D. vulgaris* (indicated by its smaller *p*-value of 7.8507 x 10^{-4}) than for *D. sargus* (*p*-value of 8.4664 x 10^{-2}). When considering each sampled location independently, it is worth noting that the correlations still hold for both species, despite slightly larger spread (see SI B-8).

4. Discussion

4.1. Fitting modelled larval sources with otolith geochemistry to delineate and evaluate spawning areas

We delineate and evaluate the relative larval supplies of several discrete spawning areas by correlating simulated cluster spawning potential and observed natal origin proportions. By doing so, we confront a probabilistic model using millions of particle trajectories and an observational approach relying only on hundreds of otolith samples. The regression coefficients (i.e. the slope of the regression curve) are around 0,7 for both species, indicating that the model tends to overestimate the poorly represented natal origins and to underestimate the dominant ones (or the other way around, from the observational point-of-view). It suggests that the probabilistic approach is able to capture the larval sources of low probabilities, whereas the observational approach may have missed, or underrepresented, those weak sources since the corresponding post-settlers are rare and scattered in the field. Similarly, the dominant larval sources simulated by our model are one to two orders of magnitude more probable than the rare ones, indicating that the post-settlers sampled in the field have high probability to originate from these prominent sources and that the rare ones are likely to be missed. Ideally, and despite the arduous financial costs and efforts, field

surveys should consider increasing the number of samples to apprehend properly those rare, yet ecologically important, sources. Our integrative approach shows great promises for the "a-priori" testing of different sampling strategies in order to provide scientifically-based guidance for the adequate spacing, sizing and sampling efforts of future field surveys. Despite difficulties owing to the large surfaces involved, further research could consist in surveying the discrete spawning areas revealed here to observe, and eventually subsample, spawning aggregations of mature adults to validate our results.

Notwithstanding the good agreement between cluster spawning potential and natal origin proportions, our empirical clustering procedure, which is based on expert knowledge, could be further generalized by using automatic and objective community detection algorithms (e.g. Rossi et al., 2014; Ser-Giacomi et al., 2015a). However, this is beyond the scope of the present manuscript due to the species- and sites-specific characteristics, the lack of information about seawater geochemistry (Di Franco et al., 2012b, 2015), along with the general opacity and inflexibility of most automatic clustering algorithms. Nonetheless, more precise and realistic characterization of spawning areas could be obtained through our empiric clustering methodology by incorporating additional information such as a geochemical atlas of seawater or supplementary connectivity knowledge inferred from genetic or tracking studies.

Our general approach allows to locate spawning areas and to better quantify larval connectivity patterns from pre-determined origins and destinations (Nolasco et al., 2018). In a context of sparse knowledge about fish spawning aggregations (Erisman et al., 2017), this information is critical for the management and conservation of marine ecosystems (see section 4.4). Furthermore, the existent published knowledge is rather biased toward tropical reef fishes (Russell et al., 2014). In this view, the present study complements recent research efforts (Calò et al., 2018) to improve our ability to locate spawning areas for temperate fish species inhabiting large and opened seascapes.

4.2. Biological controls of connectivity

Our results evidence distinct scales of dispersion for both species, which are primarily due to species-specific biological traits, namely the Pelagic Larval Duration (PLD, i.e. the time that larvae spend drifting in the ocean circulation). Indeed, *D. vulgaris*'s PLD is almost three time longer than *D. sargus*'s one, giving a ratio of ~ 2.8 which resemble most the mean ratio of their dispersal plume which reaches 4.1 ± 1.6. In contrast, the mean ratio of *D. vulgaris* to *D. sargus*'s median dispersal distances is 0.9 ± 0.2, while it is 1.7 ± 0.2 for the maximal distances. It suggests that dispersal distances and surfaces have very different relationship with the PLD, challenging the commonly accepted view that PLD is a good predictor of dispersal (Shanks et al., 2003). We consider here only the median PLDs of both species to design our numerical experiments whereas the range of PLDs derived from otolith sclerochronology is indeed more extended (Di Franco and Guidetti, 2011; Di Franco et al., 2013). Previous findings suggest that considering more appropriately the full ranges of observed PLDs could return slightly different dispersal distances and surfaces. More specifically, the exact dispersal scales would theoretically be individual-specific since each larva (or each meter-scale larval patch) would follow a unique trajectory crossing various environmental conditions which can affect biological traits, including the PLD due to slower or faster development than average (Cowen and Sponaugle, 2009). However, since Monroy et al. (2017) showed that connectivity estimates for long PLDs are more robust against PLD uncertainties than for short PLDs, we expect some substantial influences of the extreme PLDs on the quantitative patterns of connectivity mainly for *D. sargus*.

From spawning to recruitment, the key factors controlling the connectivity of early-life stages are: (i) physically-driven larval transport processes, (ii) larval behavior potentially influencing the drift, (iii) substrate availability for settlement, and (iv) local biotic interactions among the post settlers (Pineda, 2000). Fine-tuning of model parameters by data derived from otoliths of post-settled juveniles, which settled successfully and survived to local biotic interactions, allows to take factors (i), (iii) and (iv) into account in our framework. Factor (ii) has not been explicitly accounted for here, while Sparidae larvae reared in the lab showed some swimming abilities toward the end of their ontogeny (Clark et al., 2005) and possibly orientating themselves via the position of the sun (Faillettaz et al., 2018). The assumption of passive larval drift was retained here because knowledge about species-specific larval behavior in the field are sparse and very uncertain. In addition, Nolasco et al. (2018) recently documented better correlations between connectivity observations and dispersal models simulating passive rather than active drift. Indeed, the diffusion applied to trajectories when adding random components to simulate active movements, as is commonly done, is of the same order of the numerical diffusion induced by the spatial discretization of the oceanic domain, including over the nearshore regions where swimming potential would be more probable (Rossi et al., 2014). Moreover, the sequential filtering reduces successively the dispersal distances that the passive assumption could have overestimated (Shanks, 2009). Although the early-life stages of most marine organisms, including coastal fishes, are thought to be the main drivers of dispersal, the active movements of adults and juveniles may also affect connectivity (Di Franco et al., 2015). Nevertheless, because juveniles have been sampled few days after settlement (Di Franco and Guidetti, 2011; Di Franco et al., 2013), that is before any significant post-settlement movements would occur, these processes do not affect our results. Last but not least, the overall contribution of adults to broad-scale dispersal can be reasonably neglected since these stages are rather territorial, especially littoral fishes that show strong site fidelity (Di Franco et al., 2018).

4.3. Hydrodynamical control of connectivity

Besides the biological control exerted on dispersal distances and surfaces discussed above, finer discrepancies of connectivity pathways and magnitudes are rather well-explained by the spatio-temporal variability of ocean currents. Larval sources found along Italian shorelines for both species are linked to the surface circulation of the Adriatic Sea (Artegiani et al., 1997a; Millot and Taupier-Letage, 2005). Most sources are located north-westward (i.e. upstream) of the sampled locations, which is in agreement with the dominant south-eastward alongshore transport related with the W-NAd, W-MAd and W-SAd currents (Carlson et al., 2016). Along the Apulian coast, source nodes with high probabilities are found close to the sampled locations, in line with Dubois et al. (2016) and Bray et al. (2017) who have already shown that the Apulian coast tends to act as a sink of larvae. Despite the presence of the W-SAd current transporting efficiently larvae south-eastward, prominent nearshore sources can be related to wind-induced coastal downwelling which prevails in this region during winter throughout to early spring (Artegiani et al., 1997b; Bakun and Agostini, 2001). Downwelling, and its associated coastal convergence, pushes larvae toward the coastline and increases retention there, leading to these high probabilities. This is consistent with Dubois et al. (2016) who showed that wind-driven convergent oceanic systems are usually characterized by larval sinks.

The coast-to-coast connections, linking both coastlines of the Adriatic Sea, is related to the sub-gyres cyclonic circulation (Carlson et al., 2016), namely the SAd and MAd gyres. While this finding has ecological and managerial implications (see section 4.4), these inter-coastal connections have a marginal contribution as compared to the sources spread along the Apulian shores, as suggested by Melià et al. (2016). Since the southern branch of the S-Ad gyre induces west-to-east connection (Carlson et al., 2016), the presence of larval sources on both sides of the Otranto strait for *D. vulgaris* would be rather due to the turbulent character of the local circulation, as evidenced by high eddy kinetic energy values (EKE, see SI B-2). Chaotic surface circulation

could create intermittent but repeated pathways which allows inter-coastal connectivity against the main direction of the current. In the same way, filaments of high EKE stretch out until the gulf of Taranto, allowing larval sources to be found there (Fig. 4b). Conversely, the narrower and intensified W-SAd jet-like current associated with low EKE off the Apulian coast in spring imply a directional connectivity with little cross-shore exchanges, explaining why only upstream well-aligned sources were captured for *D. sargus*, contrasting the more diffused and latitudinally-extended sources found for *D. vulgaris* in winter. In fact, while Lagrangian particles are here simulating passive larvae, they can also be seen as a finite fluid parcel. In this view, Ser-Giacomi et al. (2015a) found analytical relationships between diagnostics derived from the LFN connectivity matrices and finite-time Lyapunov exponent, which is a measure of local stretching properties in the ocean, as is EKE (Waugh et al., 2006), hence supporting the link between purely physical variables and our connectivity diagnostics.

While a climatological description of the surface circulation explains well the differential connectivity of both species, ocean currents exhibit in nature high levels of variability, including for longer time-scales (e.g. years/decades). Oceanographic functioning of the Adriatic Sea have been shown to respond to the Bimodal Oscillating system (BiOS) at annual and decadal time-scales inducing substantial variability of the circulation (Mihanović et al., 2015). The dispersal patterns reported here would surely be impacted by such low-frequency variability, similarly to the strong inter-annuality of connectivity demonstrated in the north-west Mediterranean (Hidalgo et al., 2019). Note however that the BiOS has been related to large changes not only of hydrodynamics, which affect directly dispersal (Civitarese et al., 2010), but also of other abiotic properties such as temperature, which indirectly influences larval dispersal by modifying spawning behavior and larval traits (Green and Fisher, 2004). Analyzing inter-annuality of spawning is thus a complex task in which non-linear effects between several counter-acting processes may drive non-predictable results.

4.4. Exploiting connectivity information for conservation purposes

By delineating spawning areas and quantifying connectivity patterns which account for the spatio-temporal variability of ocean currents and the spawning phenology, our framework provides information about when and where spawning aggregations could occur. Considering that many coastal fishes, including those of our case-studies, are exploited by small-scale (professional and recreational) fisheries, this is essential information for managers that is currently lacking (see section 4.1). It could constitute a sound scientific basis to improve fisheries management, for instance thanks to conservation plans that would specially target the discrete spawning areas, that is where putative spawning aggregations take place. The more abilities we have in appraising when and where spawning aggregation events occur, the more appropriate will be the dynamical and adaptative management measures (Heyman, 2014), hence favoring healthier local ecosystem and larval export into neighboring exploited areas (Erisman et al., 2017; Pelc et al., 2010). Indeed, small investments on well-placed and well-timed managerial constraints (e.g. seasonal closures, restricted no-take zone, etc) targeting spawning areas could lead to large benefits for fisheries, ecotourism and biodiversity conservation (Erisman et al., 2017).

In our case-studies, we showed that *D. vulgaris* Apulian subpopulation likely originate from several spawning areas located on both Adriatic shorelines. It suggests the need of a tight international collaboration between adjacent countries, e.g. Italy, Greece, Croatia and Albania, to ensure efficient fishery spatial management in the Adriatic Sea (Hidalgo et al., 2019; Ramesh et al., 2019). Among the total pool of larvae supplied to the Apulian coast, where professional and recreational fishing occurs, 30. 8% of *D. sargus* larvae and 23. 6% of *D. vulgaris* larvae originated from the surrounding MPAs. Tremiti MPA emerges as the most important larval supplier of the studied stretch of coast: 26. 4% of *D. sargus* larvae and 7. 5% of *D. vulgaris* larvae settled in the Apulian coast likely originated from this MPA. Our model results, backed-up by observation, clearly support the fact that an upgrade of Tremiti MPA enforcement level, currently under low enforcement (Guidetti et al., 2008), could induce significant benefit to *D. sargus* and *D. vulgaris* Apulian's populations. Then, using our findings to inform management (Erisman et al., 2017), we would recommend to strengthen conservation policies of Tremiti MPA.

Moreover, properly enforced MPAs have positive effects on biomass and size distribution of larger fish (Lester et al., 2009), leading to an increase of their reproductive outputs. Indeed, MPAs could produce at least five times more offspring than same size unprotected areas (Marshall et al., 2019). Thanks to the flexibility of our model results, MPAs associated proportions (Table 4) could be scale up accordingly. Thus, all connected MPAs (Tremiti MPA alone, respectively) would send almost 70% (60%, respectively) of *D. sargus* larvae and 60% (20%, repectively) of *D. vulgaris* larvae settled in the Apulian coast, dramatically highlighting the benefit of MPAs in the replenishment of Apulian coast. Providing such useful information is well aligned with the necessity to incorporate connectivity knowledge in conservation plans and MPAs design (Dubois et al., 2016; Balbar and Metaxas, 2019).

We also provide spatially-explicit and quantitative information to potentially implement new protection measures for the other non-yet protected regions identified here during the spawning seasons of each species. Indeed, our framework not only appraises the efficiency of the local MPAs network but also inform where and when additional measures could be implemented for the protection of spawning aggregations.

For both target fishes, our results evidence well-defined spawning areas, encompassing several spawning events occurring during a given year (e.g. spring 2009 for *D. sargus* and winter 2009–2010 for *D. vulgaris*), that are directly useable by stakeholders for that given year. What remain to be investigated however is how the geographical delimitation and relative contributions of these spawning regions vary from one year to another. As said in section 4.3, this non-trivial task is kept for future work since investigating properly the inter-annual variability of spawning areas requires the considerations of all biotic and abiotic factors (e.g. in addition of the ocean circulation) displaying inter-annual variations.

5. Conclusion

We proposed an integrative framework which combines Lagrangian particle modelling, network theory and published knowledge encompassing otolith analyses, ecological and biogeographical information. It allows to locate fish spawning areas and assess their respective contributions through an improved quantitative characterization of dispersal scales and connectivity patterns. Our multidisciplinary methodology is flexible as it can be readily adapted to any other case-study, providing a certain degree of knowledge about the target species with biphasic life-cycle. Biological traits such as the PLD and the spawning phenology (derived from otolith analyses), along with larval dispersal pathways imposed by the turbulent oceanic circulation (simulated by state-of-the-art Lagrangian modelling), shape together connectivity patterns and allow us to identify spawning areas, that are critical for adequate spatial conservation planning. We showed that our analytical approach is a powerful tool providing unprecedented information about spawning areas, which can be further refined as ocean circulation models gain in reliability and as connectivity knowledge from complementary observations increases. Furthermore, if the environmental (e.g. temperature) controls of spawning are well-known and considering the current development of forecasts by operational ocean models, our approach could even anticipate some days ahead when and where spawning events are likely to occur, hence providing near-real time information for adaptive management and conservation plans.

Acknowledgment

V.R. thanks Prof. Paolo Guidetti for fruitful discussions. T.L. and V.R. acknowledge financial support from the European project SEAMoBB, funded by ERA-Net Mar-TERA and managed by ANR (number ANR_17_MART-0001_01). T.L. is supported by a PhD grant provided by the French government. A.D.F. and A.C. have been supported by the Safenet (funded by DG Mare) and FishMPABlue2 (co-financed by European Regional Development Fund-ERDF) projects. The project leading to this publication has received funding from European FEDER Fund under project 1166-39417. The authors thank two anonymous reviewers and the editor for their constructive comments that helped improve the original manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marenvres.2019.104761.

References

Alós, J., Cabanellas-Reboredo, M., March, D., 2012. Spatial and temporal patterns in the movement of adult two-banded sea bream Diplodus vulgaris (Saint-Hilaire, 1817). Fish. Res. 115, 82–88.

Artegiani, A., Paschini, E., Russo, A., Bregant, D., Raicich, F., Pinardi, N., 1997a. The Adriatic Sea general circulation. Part I: air–sea interactions and water mass structure. J. Phys. Oceanogr. 27, 1492–1514.

Artegiani, A., Paschini, E., Russo, A., Bregant, D., Raicich, F., Pinardi, N., 1997b. The Adriatic Sea general circulation. Part II: baroclinic circulation structure. J. Phys. Oceanogr. 27, 1515–1532.

Aspillaga, E., Bartumeus, F., Linares, C., Starr, R.M., López-Sanz, À., Díaz, D., Zabala, M., Hereu, B., 2016. Ordinary and extraordinary movement behaviour of small resident fish within a Mediterranean marine protected area. PLoS One 11, e0159813.

Bakun, A., Agostini, V.N., 2001. Seasonal patterns of wind-induced upwelling/downwelling in the Mediterranean Sea. Sci. Mar. 65, 243–257.

Balbar, A.C., Metaxas, A., 2019. The current application of ecological connectivity in the design of marine protected areas. Glob. Ecol. Conserv. 17, e00569. https://doi.org/10.1016/j.gecco.2019.e00569. URL: http://www.sciencedirect.com/science/article/pii/S2351989418304347.

Bray, L., Kassis, D., Hall-Spencer, J., 2017. Assessing larval connectivity for marine spatial planning in the Adriatic. Mar. Environ. Res. 125, 73–81. https://doi.org/10.1016/j.marenvres.2017.01.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0141113617300430.

Burgess, S.C., Baskett, M.L., Grosberg, R.K., Morgan, S.G., Strathmann, R.R., 2016. When is dispersal for dispersal? Unifying marine and terrestrial perspectives. Biol. Rev. 91, 867–882. https://doi.org/10.1111/brv.12198. URL: http://onlinelibrary.wiley.com/doi/abs/10.1111/brv.12198.

Calò, A., Félix-Hackradt, F.C., Garcia, J., Hackradt, C.W., Rocklin, D., Treviño Otón, J., Charton, J.A.G., 2013. A review of methods to assess connectivity and dispersal between fish populations in the Mediterranean Sea. Adv. Oceanogr. Limnol. 4, 150–175.

Calò, A., Lett, C., Mourre, B., Pérez-Ruzafa, Á., García-Charton, J.A., 2018. Use of Lagrangian simulations to hindcast the geographical position of propagule release zones in a Mediterranean coastal fish. Mar. Environ. Res. 134, 16–27. https://doi.org/10.1016/j.marenvres.2017.12.011. URL: https://linkinghub.elsevier.com/retrieve/pii/S0141113617305834.

Carlson, D.F., Griffa, A., Zambianchi, E., Suaria, G., Corgnati, L., Magaldi, M.G., Poulain, P.M., Russo, A., Bellomo, L., Mantovani, C., Celentano, P., Molcard, A., Borghini, M., 2016. Observed and modeled surface Lagrangian transport between coastal regions in the Adriatic Sea with implications for marine protected areas. Cont. Shelf Res. 118, 23–48. https://doi.org/10.1016/j.csr.2016.02.012. URL: https://linkinghub.elsevier.com/retrieve/pii/S0278434316300620.

Ciannelli, L., Fisher, J., Skern-Mauritzen, M., Hunsicker, M., Hidalgo, M., Frank, K., Bailey, K., 2013. Theory, consequences and evidence of eroding population spatial structure in harvested marine fishes: a review. Mar. Ecol. Prog. Ser. 480, 227–243. https://doi.org/10.3354/meps10067. URL: http://www.int-res.com/abstracts/meps/v480/p227-243/.

Ciliberti, S.A., Pinardi, N., Coppini, G., Oddo, P., Vukicevic, T., Lecci, R., Verri, G., Kumkar, Y., Creti, S., 2015. A high resolution Adriatic-Ionian Sea circulation model for operational forecasting. In: EGU General Assembly Conference Abstracts.

Civitarese, G., Gačić, M., Lipizer, M., Eusebi Borzelli, G.L., 2010. On the impact of the bimodal oscillating system (BiOS) on the biogeochemistry and biology of the adriatic and ionian seas (eastern mediterranean). Biogeosciences 7, 3987–3997. URL: https://www.biogeosciences.net/7/3987/2010/bg-7-3987-2010.html. https://doi.org/10.5194/bg-7-3987-2010.

Clark, D.L., Leis, J.M., Hay, A.C., Trnski, T., 2005. Swimming ontogeny of larvae of four temperate marine fishes. Mar. Ecol. Prog. Ser. 292, 287–300.

Courant, R., Friedrichs, K., Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. 100, 32–74. https://doi.org/10.1007/BF01448839. URL: https://doi.org/10.1007/BF01448839.

Cowen, R.K., Paris, C.B., Srinivasan, A., 2006. Scaling of connectivity in marine populations. Science 311, 522–527.

Cowen, R.K., Sponaugle, S., 2009. Larval Dispersal and Marine Population Connectivity.

Di Franco, A., Calò, A., Pennetta, A., De Benedetto, G., Planes, S., Guidetti, P., 2015. Dispersal of larval and juvenile seabream: implications for Mediterranean marine protected areas. Biol. Conserv. 192, 361–368.

Di Franco, A., Coppini, G., Pujolar, J.M., De Leo, G.A., Gatto, M., Lyubartsev, V., Melia, P., Zane, L., Guidetti, P., 2012a. Assessing dispersal patterns of fish propagules from an effective Mediterranean marine protected area. PLoS One 7, e52108.

Di Franco, A., Gillanders, B.M., De Benedetto, G., Pennetta, A., De Leo, G.A., Guidetti, P., 2012b. Dispersal patterns of coastal fish: implications for designing networks of marine protected areas. PLoS ONE 7, e31681. https://doi.org/10.1371/journal.pone.0031681. URL: http://dx.plos.org/10.1371/journal.pone.0031681.

Di Franco, A., Guidetti, P., 2011. Patterns of variability in early-life traits of fishes depend on spatial scale of analysis. Biol. Lett. 7, 454–456.

Di Franco, A., Plass-Johnson, J.G., Di Lorenzo, M., Meola, B., Claudet, J., Gaines, S.D., García-Charton, J.A., Giakoumi, S., Grorud-Colvert, K., Hackradt, C.W., Micheli, F., Guidetti, P., 2018. Linking home ranges to protected area size: the case study of the Mediterranean Sea. Biol. Conserv. 221, 175–181. https://doi.org/10.1016/j.biocon.2018.03.012. URL: http://www.sciencedirect.com/science/article/pii/S0006320717311187.

Di Franco, A., Qian, K., Calò, A., Di Lorenzo, M., Planes, S., Guidetti, P., 2013. Patterns of variability in early life traits of a Mediterranean coastal fish. Mar. Ecol. Prog. Ser. 476, 227–235.

Di Lorenzo, M., D'Anna, G., Badalamenti, F., Giacalone, V.M., Starr, R.M., Guidetti, P., 2014. Fitting the size of no-take zones to species movement patterns: a case study on a Mediterranean seabream. Mar. Ecol. Prog. Ser. 502, 245–255.

Domeier, M.L., Colin, P.L., 1997. Tropical reef fish spawning aggregations: defined and reviewed. Bull. Mar. Sci. 60, 698–726.

Dubois, M., Rossi, V., Ser-Giacomi, E., Arnaud-Haond, S., López, C., Hernández-García, E., 2016. Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems. Glob. Ecol. Biogeogr. 25, 503–515.

Erisman, B., Heyman, W., Kobara, S., Ezer, T., Pittman, S., Aburto-Oropeza, O., Nemeth, R.S., 2017. Fish spawning aggregations: where well-placed management actions can yield big benefits for fisheries and conservation. Fish Fish. 18, 128–144. https://doi.org/10.1111/faf.12132. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/faf.12132.

Faillettaz, R., Durand, E., Paris, C.B., Koubbi, P., Irisson, J.O., 2018. Swimming speeds of Mediterranean settlement-stage fish larvae nuance Hjort's aberrant drift hypothesis. Limnol. Oceanogr. 63, 509–523.

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486, 75–174.

Fuiman, L.A., Connelly, T.L., Lowerre-Barbieri, S.K., McClelland, J.W., 2015. Egg boons: central components of marine fatty acid food webs. Ecology 96, 362–372. https://doi.org/10.1890/14-0571.1. URL: https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/14-0571.1.

Giacalone, V.M., Pipitone, C., Badalamenti, F., Sacco, F., Zenone, A., FERRERI, R., MICALE, V., BASILONE, G., D?ANNA, G., 2018. Home Range, Movements and Daily Activity of the White Seabream Diplodus sargus (Linnaeus, 1758) during the Spawning Season URL. https://doi.org/10.21411/cbm.a.7c19c1b8. http://application.sb-roscoff.fr/cbm/doi/10.21411/CBM.A.7C19C1B8.

Gibb, F.M., Régnier, T., Donald, K., Wright, P.J., 2017. Connectivity in the early life history of sandeel inferred from otolith microchemistry. J. Sea Res. 119, 8–16.

Green, B.S., Fisher, R., 2004. Temperature influences swimming speed, growth and larval duration in coral reef fish larvae. J. Exp. Mar. Biol. Ecol. 299, 115–132. https://doi.org/10.1016/j.jembe.2003.09.001. URL: http://www.sciencedirect.com/science/article/pii/S0022098103004404.

Guidetti, P., 2000. Differences among fish assemblages associated with nearshore Posidonia oceanica seagrass beds, rocky–algal reefs and unvegetated sand habitats in the Adriatic Sea. Estuar. Coast Shelf Sci. 50, 515–529.

Guidetti, P., Milazzo, M., Bussotti, S., Molinari, A., Murenu, M., Pais, A., Spanò, N., Balzano, R., Agardy, T., Boero, F., Carrada, G., Cattaneo-Vietti, R., Cau, A., Chemello, R., Greco, S., Manganaro, A., Notarbartolo di Sciara, G., Russo, G.F., Tunesi, L., 2008. Italian marine reserve effectiveness: does enforcement matter? Biol. Conserv. 141, 699–709. https://doi.org/10.1016/j.biocon.2007.12.013. URL: http://www.sciencedirect.com/science/article/pii/S0006320707004685.

Harmelin-Vivien, M., Harmelin, J., Leboulleux, V., 1995. Microhabitat requirements for settlement of juvenile sparid fishes on Mediterranean rocky shores. Hydrobiologia 300, 309–320.

Heithaus, M.R., Frid, A., Wirsing, A.J., Worm, B., 2008. Predicting ecological consequences of marine top predator declines. Trends Ecol. Evol. 23, 202–210. https://doi.org/10.1016/j.tree.2008.01.003. URL: http://www.sciencedirect.com/science/article/pii/S0169534708000578.

Henry, L.A., Mayorga-Adame, C.G., Fox, A.D., Polton, J.A., Ferris, J.S., McLellan, F., McCabe, C., Kutti, T., Roberts, J.M., 2018. Ocean sprawl facilitates dispersal and connectivity of protected species. Sci. Rep. 8, 11346. https://doi.org/10.1038/s41598-018-29575-4. URL: http://www.nature.com/articles/s41598-018-29575-4.

Heyman, W.D., 2014. Let Them Come to You: Reinventing Management of the Snapper-Grouper Complex in the Western Atlantic: A Contribution to the Data Poor Fisheries Management Symposium, vol. 7.

Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernández-García, E., Guijarro, B., Massutí, E., Alemany, F., Jadaud, A., Perez, J.L., Reglero, P., 2019. Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations. Ecological Applications 0. , e01913. https://doi.org/10.1002/eap.1913. URL: http://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/

eap.1913.
Kingsford, M.J., 1988. The early life history of fish in coastal waters of northern New Zealand: a review. N. Z. J. Mar. Freshw. Res. 22, 463–479. https://doi.org/10.1080/00288330.1988.9516316. URL: https://doi.org/10.1080/00288330.1988.9516316.
Kool, J.T., Moilanen, A., Treml, E.A., 2013. Population connectivity: recent advances and new perspectives. Landsc. Ecol. 28, 165–185.
Lenfant, P., Planes, S., 1996. Genetic differentiation of white sea bream within the Lion's gulf and the Ligurian sea (Mediterranean Sea). J. Fish Biol. 49, 613–621.
Lester, S.E., Halpern, B.S., Grorud-Colvert, K., Lubchenco, J., Ruttenberg, B.I., Gaines, S.D., Airamé, S., Warner, R.R., 2009. Biological effects within no-take marine reserves: a global synthesis. Mar. Ecol. Prog. Ser. 384, 33–46.
Levin, P.S., 1994. Fine-scale temporal variation in recruitment of a temperate demersal fish: the importance of settlement versus post-settlement loss. Oecologia 97, 124–133. https://doi.org/10.1007/BF00317916. URL: https://doi.org/10.1007/BF00317916.
Lowe, W.H., Kovach, R.P., Allendorf, F.W., 2017. Population genetics and demography unite ecology and evolution. Trends Ecol. Evol. 32, 141–152. https://doi.org/10.1016/j.tree.2016.12.002. URL: https://www.cell.com/trends/ecology-evolution/abstract/S0169-5347(16)30235-X.
Madec, G., others, 2015. NEMO Ocean Engine.
Marshall, D.J., Gaines, S., Warner, R., Barneche, D.R., Bode, M., 2019. Underestimating the benefits of marine protected areas for the replenishment of fished populations. Front. Ecol. Environ. https://doi.org/10.1002/fee.2075. URL: http://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/fee.2075.
Melià, P., Schiavina, M., Rossetto, M., Gatto, M., Fraschetti, S., Casagrandi, R., 2016. Looking for hotspots of marine metacommunity connectivity: a methodological framework. Sci. Rep. 6, 23705.
Mihanović, H., Vilibić, I., Dunić, N., Šepić, J., 2015. Mapping of decadal middle Adriatic oceanographic variability and its relation to the BiOS regime: MAPPING OF DECADAL ADRIATIC VARIABILITY. J. Geophys. Res. Oceans 120, 5615–5630. URL: http://doi.org/10.1002/2015JC010725 10.1002/2015JC010725.
Millot, C., Taupier-Letage, I., 2005. Circulation in the Mediterranean Sea. In: The Mediterranean Sea. Springer, pp. 29–66.
Mitcheson, Y.S.d., Craig, M.T., Bertoncini, A.A., Carpenter, K.E., Cheung, W.W.L., Choat, J.H., Cornish, A.S., Fennessy, S.T., Ferreira, B.P., Heemstra, P.C., Liu, M., Myers, R.F., Pollard, D.A., Rhodes, K.L., Rocha, L.A., Russell, B.C., Samoilys, M.A., Sanciangco, J., 2013. Fishing groupers towards extinction: a global assessment of threats and extinction risks in a billion dollar fishery. Fish Fish. 14, 119–136. https://doi.org/10.1111/j.1467-2979.2011.00455.x. URL: http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-2979.2011.00455.x.
Monroy, P., Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., 2017. Sensitivity and robustness of larval connectivity diagnostics obtained from Lagrangian Flow Networks. ICES (Int. Counc. Explor. Sea) J. Mar. Sci. 74, 1763–1779.
Nolasco, R., Gomes, I., Peteiro, L., Albuquerque, R., Luna, T., Dubert, J., Swearer, S.E., Queiroga, H., 2018. Independent estimates of marine population connectivity are more concordant when accounting for uncertainties in larval origins. Sci. Rep. 8, 2641. https://doi.org/10.1038/s41598-018-19833-w. URL: http://www.nature.com/articles/s41598-018-19833-w.
Oddo, P., Pinardi, N., Zavatarelli, M., Coluccelli, A., 2006. The Adriatic basin forecasting system. Acta Adriat.: Int. J. Mar. Sci. 47, 169–184.
Olivar, M., Sabatés, A., 1997. Vertical distribution of fish larvae in the north-west Mediterranean Sea in spring. Mar. Biol. 129, 289–300.
Pelc, R.A., Warner, R.R., Gaines, S.D., Paris, C.B., 2010. Detecting larval export from marine reserves. Proc. Natl. Acad. Sci. 107, 18266–18271. https://doi.org/10.1073/pnas.0907368107. URL: https://www.pnas.org/content/107/43/18266.
Pineda, J., 2000. Linking larval settlement to larval transport: assumptions, potentials, and pitfalls. Oceanogr. East. Pac. 1, 84–105.
Pineda, J., Hare, J.A., Sponaugle, S., 2007. Larval transport and dispersal in the coastal ocean and consequences for population connectivity. Oceanography 20, 22–39.
Pujolar, J.M., Schiavina, M., Di Franco, A., Melià, P., Guidetti, P., Gatto, M., De Leo, G.A., Zane, L., 2013. Understanding the effectiveness of marine protected areas using genetic connectivity patterns and Lagrangian simulations. Divers. Distrib. 19, 1531–1542.
Ramesh, N., Rising, J.A., Oremus, K.L., 2019. The small world of global marine fisheries: the cross-boundary consequences of larval dispersal. Science 364, 1192–1196. https://doi.org/10.1126/science.aav3409. URL: https://science.sciencemag.org/content/364/6446/1192.
Richards, W.J., Lindeman, K.C., 1987. Recruitment dynamics of reef fishes: planktonic processes, settlement and demersal ecologies, and fishery analysis. Bull. Mar. Sci. 41, 392–410.
Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., 2014. Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. Geophys. Res. Lett. 41, 2883–2891.
Russell, M.W., de Mitcheson, Y.S., Erisman, B.E., Hamilton, R.J., Luckhurst, B.E., Nemeth, R.S., 2014. Status Report World's Fish Aggregations 2014. pp. 13.
Sadovy, Y., Domeier, M., 2005. Are aggregation-fisheries sustainable? Reef fish fisheries as a case study. Coral Reefs 24, 254–262. https://doi.org/10.1007/s00338-005-0474-6. URL: http://link.springer.com/10.1007/s00338-005-0474-6.
Ser-Giacomi, E., Rodríguez-Méndez, V., López, C., Hernández-García, E., 2017. Lagrangian Flow Network approach to an open flow model. Eur. Phys. J. Spec. Top. 226, 2057–2068. https://doi.org/10.1140/epjst/e2017-70044-2. URL: https://doi.org/10.1140/epjst/e2017-70044-2.
Ser-Giacomi, E., Rossi, V., López, C., Hernandez-Garcia, E., 2015a. Flow networks: a characterization of geophysical fluid transport. Chaos. Interdiscip. J. Nonlinear Sci. 25, 036404.
Ser-Giacomi, E., Vasile, R., Hernandez-Garcia, E., Lopez, C., 2015b. Most probable paths in temporal weighted networks: an application to ocean transport. Physical Review E 92. https://doi.org/10.1103/PhysRevE.92.012818. arXiv: 1411.6902. URL: http://arxiv.org/abs/1411.6902.
Shanks, A.L., 2009. Pelagic larval duration and dispersal distance revisited. Biol. Bull. 216, 373–385. https://doi.org/10.1086/BBLv216n3p373. URL: https://www.journals.uchicago.edu/doi/10.1086/BBLv216n3p373.
Shanks, A.L., Grantham, B.A., Carr, M.H., 2003. Propagule dispersal distance and the size and spacing of marine reserves. Ecol. Appl. S159–S169.
Thresher, R.E., Colin, P.L., Bell, L.J., 1989. Planktonic duration, distribution and population structure of western and central pacific damselfishes (pomacentridae). Copeia 1989 420–434. https://doi.org/10.2307/1445439. URL: https://www.jstor.org/stable/1445439.
Waugh, D.W., Abraham, E.R., Bowen, M.M., 2006. Spatial variations of stirring in the surface ocean: a case study of the Tasman Sea. J. Phys. Oceanogr. 36, 526–542.

3. De la connectivité démographique à la connectivité génétique : une approche multi-génération cumulative

3.1. Résumé

Dans le chapitre précédent, nous avons détaillé comment la connectivité démographique via la dispersion larvaire structure spatialement les populations, et comment les patrons sous-jacents de connectivité sont influencés par la variabilité environnementale saisonnière et interannuelle. Au sein du cycle de vie biphasique de deux espèces de sparidés, nous avons pu identifier des zones où la phénologie de ponte est similaire à l'échelle décennale, et envisager des trajectoires migratoires stables mais aussi hautement variables au fil des événements successifs de dispersion entre populations. La dispersion larvaire, en permettant le déplacement d'individus et de leurs génomes dans l'espace induit du flux de gènes (Clobert et al., 2012). L'étude de la connectivité génétique consiste à caractériser ces échanges de gènes entre populations et, en raison du caractère héréditaire du matériel génétique, doit considérer les événements successifs de dispersion larvaire. En d'autres termes, la connectivité génétique considère le flux de gènes sur plusieurs générations et entre plusieurs populations successives, là où la connectivité démographique considère la dispersion larvaire d'une génération à l'autre et d'une population à l'autre. Il faut ainsi développer de nouveaux outils théoriques qui s'approchent au plus près des caractéristiques particulières de la connectivité génétique.

Dans ce chapitre, nous avons théorisé et formulé la connectivité de manière à évaluer la probabilité de connexion génétique entre n'importe quelle paire de populations via des événements successifs et cumulatifs de dispersion larvaire. En effet,

cette approche cumule chaque probabilité de connexion associée à un nombre de générations (i.e. un nombre d'événements de dispersion successifs) jusqu'au nombre de générations total. En d'autres termes, elle ne prend pas seulement en compte la probabilité de connexion à un nombre donné de générations comme il était fait jusqu'à présent (Boulanger et al., 2020 ; Buonomo et al., 2017 ; Foster et al., 2012 ; Jahnke et al., 2018 ; Kool et al., 2010 ; White et al., 2010), mais considère également toutes les connexions assurées par les nombres de génération inférieurs. De plus, dans ce chapitre, nous introduisons une nouvelle notion, celle de connexion implicite que nous différencions des connexions que nous définissions comme explicites, les seules qui étaient considérées avant nos travaux. Les connexions explicites sont les plus intuitives lorsqu'on considère la connectivité génétique : elles correspondent au lien filial, c'est à dire la relation parents-enfants entre deux populations. Elles sont utilisées dans toutes les simulations de connectivité génétique (Jahnke & Jonsson, 2022). Les connexions implicites, jusqu'alors inconsidérées dans la littérature, appréhendent le lien entre sœurs et frères, c'est-à-dire lorsque que deux populations peuvent montrer une cohérence génétique car elles accueillent des migrants qui dispersent depuis une même population source.

Les connexions explicites et implicites, nouvellement définies comme cumulatives sur de multiples générations, sont appliquées sur un réseau d'environ 1000 nœuds représentant le transport océanique pour une période de 30 jours à la surface de la mer Méditerranée. En d'autres termes, et en adaptant aux concepts introduits dans cette thèse, ce réseau correspond à l'ensemble des connexions démographiques réalisées par des populations théoriques situées sur l'entière surface du bassin méditerranéen, et pour une espèce qui a une PLD typique de 30 jours. Il apparait que les connexions implicites, contrairement aux connexions explicites, s'affranchissent des barrières de transport, aussi appelé barrières hydrologiques à la dispersion, initialement identifiées sur la période d'intérêt. De plus, la moyenne des probabilités des connexions multi-générations implicites sature plus rapidement (à environ 500 générations), et avec une valeur moyenne de probabilité de connexion cinq fois supérieure que pour les connexions explicites. Les résultats développés dans ce chapitre suggèrent que les processus de connectivité génétique à la surface de l'océan ont pu être largement sous-estimés et que les barrières de transport, jusqu'à présent considérés comme vectrice de structure génétique, sont potentiellement perméables aux connexions implicites.

3.2. Article

Ser-Giacomi, E., Legrand, T., Hernández-Carrasco, I. & Rossi, V. (2021). Explicit and implicit network connectivity : Analytical formulation and application to transport processes [Publisher : American Physical Society]. Phys. Rev. E, 103(4), 042309. https://doi.org/10.1103/PhysRevE.103.042309

PHYSICAL REVIEW E **103**, 042309 (2021)

Explicit and implicit network connectivity: Analytical formulation and application to transport processes

Enrico Ser-Giacomi *
Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 54-1514 MIT, Cambridge, Massachusetts 02139, USA

Térence Legrand
Aix Marseille University, Universite de Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (UMR 7294), Marseille, France

Ismael Hernández-Carrasco
Mediterranean Institute for Advances Studies (IMEDEA, UIB-CSIC), Mallorca (Spain)

Vincent Rossi
Aix Marseille University, University of Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (UMR 7294), Marseille, France

(Received 17 August 2020; revised 7 February 2021; accepted 18 March 2021; published 13 April 2021; corrected 3 May 2021 and 27 January 2022)

Connectivity is a fundamental structural feature of a network that determines the outcome of any dynamics that happens on top of it. However, an analytical approach to obtain connection probabilities between nodes associated with paths of different lengths is still missing. Here, we derive exact expressions for random-walk connectivity probabilities across any range of numbers of steps in a generic temporal, directed, and weighted network. This allows characterizing explicit connectivity realized by causal paths as well as implicit connectivity related to motifs of three nodes and two links called here pitchforks. We directly link such probabilities to the processes of tagging and sampling any quantity exchanged across the network, hence providing a natural framework to assess transport dynamics. Finally, we apply our theoretical framework to study ocean transport features in the Mediterranean Sea. We find that relevant transport structures, such as fluid barriers and corridors, can generate contrasting and counterintuitive connectivity patterns bringing novel insights into how ocean currents drive seascape connectivity.

DOI: 10.1103/PhysRevE.103.042309

I. INTRODUCTION

Connectivity is a key feature of network's structure [1,2] that determines how strongly and quickly different nodes can be linked by consecutive edges [3–5]. Indeed, for any dynamics running over a network, connectivity strongly influences the temporal and spatial evolution of the associated processes and patterns [6]. This has been proven in several contexts, such as epidemic or information spreading [7,8], biological interactions [9], neural networks [10], social systems [11], and fluid transport [12]. Globally, connectivity is determined by topological properties of the network: link density, degree and weight distributions, clustering, modularity, reciprocity, etc. However, these metrics describe statistical features of the network and do not inform about local patterns of connectivity between specific pairs of nodes [3,13].

The conventional approach to characterize pairwise connectivity consists in studying random walks and their trajectories. In fact, random walkers can be seen as agents that navigate through the network drawing paths between pairs of nodes [14–16]. Each of these pathways can be thus defined by the sequence of nodes visited by a random walker and, by multiplying the node-to-node single-step transition probabilities, one can obtain the probability of occurrence of any of them [17–20].

In this way, connectivity can be characterized within a solid probabilistic framework. Moreover, when a given quantity, such as people [21], fluid [12], goods [22], or information [23], is transported across the network, random walker transition probabilities can be related to fractions of exchanged quantities between node pairs. More concretely, this means that it is possible to calculate the probability that an amount of quantity that has been tagged or sampled in a given node will reach another specific destination node forward- or backward-in-time, respectively. As a result, random walks can also mimic transport, dispersion and mixing processes across a network [18,24]. Eventually, this could permit to rigorously establish a quantitative link between the structural features of a network and the dynamics of any transported quantity across it. Such connection would also be relevant in temporal networks, especially under mixing regimes in which network connectivity patterns and random walks unfold on comparable time scales [25,26].

However, to our knowledge, analytical expressions for connectivity probabilities between any pair of nodes that

*enrico.sergiacomi@gmail.com

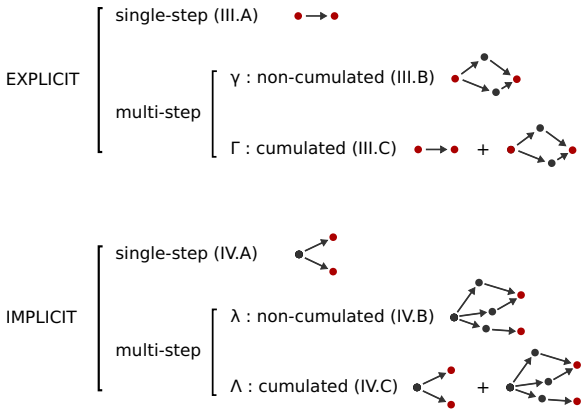


FIG. 1. Sketch of the different connectivity probabilities considered. The sections introducing each quantity are indicated within brackets while greek letters correspond to their mathematical expressions.

take into account connections realized by paths of different lengths (i.e., paths composed of different number of steps) in a "cumulated" manner are, to our knowledge, still lacking. Indeed, while the probabilities of connection realized by paths of the same lengths (i.e., imposing a prescribed number of steps) are readily obtained with simple matrix products, the cumulated probabilities across generic ranges of path lengths (i.e., across different numbers of steps) has not been derived yet. This is mainly due to the fact that connection events between two nodes realized by paths of different lengths are not mutually-exclusive from a probabilistic point of view, making the calculations to obtain them quite convoluted. It is worth noting that this shortcoming holds for both static and temporal network. As such, the current approach to study pairwise connectivity is through Monte Carlo numerical simulations. Specifically, it consists in releasing large numbers of random walkers in a given starting node and in estimating the connection probability with any other destination node from the proportion of walkers that ended up there after a given number of steps.

Moreover, also the concept of connectivity by itself could be extended. Indeed, the *explicit* connectivity probabilities described above are conceptually associated with the pathway of a random walker that joins two nodes, symbolizing a kind of "parent-child" relationship between starting and ending node. Nevertheless, we can also be interested in looking contemporaneously at the entire network in a synoptic fashion. This is the case, for instance, when modeling a transport or spreading phenomena on a network [27] or when tracking differentiation across a phylogenetic tree [28]. In such processes, each pair of nodes could be simultaneously influenced by a third node (or more than one) and such "sibling-sibling" relationships can determine similarities between nodes pairs that we could regard as a form of *implicit* connectivity. These connectivity patterns, at one step, are realized by a particular kind of three-nodes motif, here called *pitchfork*, composed of a node acting as common source (or destination) for two other nodes (see Fig. 1). Examples of such kind of interactions can be found in ecological networks when two species compete for the same resource [29] or in social systems when two agents are both influenced by a third one [30]. Thus, implicit connections associated with pitchforks are in this sense complementary to the aforementioned standard explicit patterns and, despite being mostly overlooked, could play a major role in determining network dynamics.

In this paper, we derive exact analytical expressions for explicit and implicit random-walk connectivity probabilities across any range of numbers of steps in a generic temporal, directed and weighted network. First, in Sec. II we set the theoretical background and delineate the relationships between random walk transition probabilities and the transport dynamics of a given quantity across the network. In Secs. III and IV, we introduce the concept of *cumulated* connectivity that permits to calculate connection probabilities not only for a fixed number of steps but also across an arbitrary range of possible numbers of steps, allowing probability values to eventually saturate toward an asymptotic value. Such approach is adopted to provide exact formulas for: (i) explicit connectivity patterns associated with causal paths among two nodes and (ii) implicit connectivity patterns realized by multistep pitchforks (see a summary of the different connectivity patterns in Fig. 1). Moreover, if a given quantity Q is transported across the network, then we can relate random walk probabilities to processes of tagging and sampling such quantity in specific nodes of the network. This allows linking the probabilistic view of connectivity with an interpretation in terms of transport and diffusion. In Sec. V, we calculate connection probabilities for two simple networks and we numerically confirm our analytical results highlighting significant differences between static and temporal network connectivity. In Sec. VI, we further apply our theoretical approach to characterize connectivity features of a network describing the transport of surface water masses across the Mediterranean Sea [12,31–33]. From probabilistic estimations of connectivity we provide both specific site-to-site and global basin-scale statistics. We find very relevant differences among explicit and implicit probabilities and across different ranges of number of steps. We also show that such probabilities, in average, saturate to different, non-trivial values. Finally, we discuss the implications of such results.

II. RANDOM WALKS AND TRANSPORT PROCESSES

A. Network adjacency matrix and its normalizations

We consider a generic directed, weighted, and temporal network of N nodes. Hence, each of its links is directed and characterized by a positive weight that measures the "intensity" of the connection realized between two nodes. Moreover, due to the temporal character of the network, such weights can change in time. Given a discrete time sequence $\{t_0, t_1, \ldots, t_{M-1}, t_M\}$, the time-dependent structure of the network can be thus described by a set of adjacency matrices in which each element $\mathbf{A}_{ij}^{t_l \to t_{l+1}}$ is the weight of the link from node i to node j during the time interval $[t_l, t_{l+1}]$. For convention, links are hereafter established forward-in-time across different layers representing consecutive discrete times [34].

We define the out-strength and in-strength of node i as

$$S_i^O(t_l \to t_{l+1}) = \sum_j \mathbf{A}_{ij}^{t_l \to t_{l+1}}, \tag{1}$$

$$S_i^I(t_l \to t_{l+1}) = \sum_j \mathbf{A}_{ji}^{t_l \to t_{l+1}}. \tag{2}$$

Assuming that S^O and S^I are always positive, two normalizations for the matrix $\mathbf{A}_{ij}^{t_l \to t_{l+1}}$ are possible:

$$\mathbf{F}_{ij}^{t_l \to t_{l+1}} = \frac{\mathbf{A}_{ij}^{t_l \to t_{l+1}}}{S_i^O(t_l \to t_{l+1})}, \tag{3}$$

$$\mathbf{B}_{ji}^{t_{l+1} \to t_l} = \frac{\mathbf{A}_{ij}^{t_l \to t_{l+1}}}{S_j^I(t_l \to t_{l+1})}, \tag{4}$$

obtaining the following conservation conditions: $\sum_j \mathbf{F}_{ij}^{t_l \to t_{l+1}} = 1$ and $\sum_i \mathbf{B}_{ji}^{t_{l+1} \to t_l} = 1$.

B. Random walk transition probabilities and transport dynamics

Once the adjacency matrices of the network $\mathbf{A}^{t_l \to t_{l+1}}$ are normalized, a random walk can be defined on it. Indeed, in the $[t_l, t_{l+1}]$ time interval, $\mathbf{F}_{ij}^{t_l \to t_{l+1}}$ is the forward-in-time transition probability for a random walker to jump from node i to j while $\mathbf{B}_{ji}^{t_{l+1} \to t_l}$ is the backward-in-time transition probability to go from j to i. Hence, the direction of the links is always associated with the forward-in-time direction but still, for a given link, we are able to define both the forward- and backward-in-time transition probabilities. If we assume a Markovian dynamics, then the probability for a random walker to visit a given sequence of nodes will be given by the product of the associated single-step transition probabilities.

If link weights can be associated with a generic transported quantity Q across the network, random walk transition probabilities can be related to processes of tagging and sampling the transported quantity. Indeed, imagining to tag a portion of Q inside i at t_l, $\mathbf{F}_{ij}^{t_l \to t_{l+1}}$ is the probability that such tagged quantity will arrive to j at t_{l+1}. Consequently, $\mathbf{B}_{ji}^{t_{l+1} \to t_l}$ is the probability of sampling a portion of Q in j at t_{l+1} that was in i at t_l. Pushing forward this analogy, we can quantify the fraction of transported quantity between the pair of nodes i, j in the time interval $[t_l, t_{l+1}]$ by means of transition probabilities [18,24]. Indeed, $\mathbf{F}_{ij}^{t_l \to t_{l+1}}$ is the fraction of Q present in i at t_l that arrives to j at t_{l+1}. Similarly, $\mathbf{B}_{ji}^{t_{l+1} \to t_l}$ is the fraction of Q present in j at t_{l+1} that was in i at t_l.

C. Paths in temporal weighted networks

We denote a path μ of M-steps between nodes i and j as a $(M+1)$-tuple $\{i, k_1, \ldots, k_{M-1}, j\}$ corresponding to the sequence of nodes visited by a random walker at times $\{t_0, t_1, \ldots, t_{M-1}, t_M\}$.

Thus, assuming a Markov process, the forward-in-time probability for a random walker to take the M-steps path μ under the condition of starting in i and ending in j is [18,20,24,35]

$$\mathbf{F}_{ik_1}^{t_0 \to t_1} \mathbf{F}_{k_1k_2}^{t_1 \to t_2} \ldots \mathbf{F}_{k_{M-2}k_{M-1}}^{t_{M-2} \to t_{M-1}} \mathbf{F}_{k_{M-1}j}^{t_{M-1} \to t_M}. \tag{5}$$

Conversely, the backward-in-time probability to take the M-steps path μ under the condition of starting in j and ending in i is

$$\mathbf{B}_{jk_{M-1}}^{t_M \to t_{M-1}} \mathbf{B}_{k_{M-1}k_{M-2}}^{t_{M-1} \to t_{M-2}} \ldots \mathbf{B}_{k_2k_1}^{t_2 \to t_1} \mathbf{B}_{k_1i}^{t_1 \to t_0}. \tag{6}$$

Note that, due to the temporal dependence of the network, the above probabilities depend not only on the number of steps (as in the static case) but also on the specific initial or final time considered.

III. EXPLICIT CONNECTIVITY

In this section we provide exact analytical expressions for random walk probabilities associated with paths. Depending on the number of steps considered, we can define single-step ($M = 1$) or multistep ($M > 1$) connectivity. First, we introduce connectivity for the case of a fixed number of steps M (noncumulated connectivity). Then, we extend this conventional concept by considering connections occurring over a given range of number of steps spanning 1 to M (cumulated connectivity). Hence, the noncumulated connectivity is associated with the probability that a random walker joins two nodes in a specific number of steps. Note that, for the temporal case, since the network is time-dependent, we should also specify the initial time. This probability does not include the possibility of reaching the destination node before or after the exact number of steps chosen. Cumulated connectivity overcomes this limitation by considering the probability that a random walker reaches the destination in an arbitrary number of steps as long as it is comprised within a given range of numbers of steps. When a generic quantity Q is transported across the network, we can also find a relation between the above probabilities and portions of Q (see Sec. II B).

A. Single-step explicit connectivity

Single-step explicit connectivity is associated directly with the elements of the $\mathbf{F}$ and $\mathbf{B}$ matrices (see Sec. II B). Considering the $[t_l, t_{l+1}]$ time interval, we define the *single-step explicit connectivity* calculated forward-in-time from node i to j as

$$\boldsymbol{\gamma}^f(t_l, t_{l+1}) = \mathbf{F}_{ij}^{t_l \to t_{l+1}}. \tag{7}$$

Similarly, we define the backward-in-time *single-step explicit connectivity* as

$$\boldsymbol{\gamma}^b(t_l, t_{l+1}) = \mathbf{B}_{ij}^{t_{l+1} \to t_l}. \tag{8}$$

If some generic quantity Q is transported across the network and one tags an amount of it that is present in node i at time t_l, then the probability that will arrive in node j at time t_{l+1} is exactly $\mathbf{F}_{ij}^{t_l \to t_{l+1}}$. Analogously, if one samples an amount of Q in i at time t_{l+1}, then the probability that was in j at time t_l is $\mathbf{B}_{ij}^{t_{l+1} \to t_l}$.

B. Noncumulated multistep explicit connectivity

To obtain the total probability of connection among any given pair of nodes in exactly M steps, we need to sum the probability of each of the paths that connect that pair. Hence, using the Chapman-Kolmogorov equation, we define, given a fixed number of steps M, the *noncumulated multistep explicit connectivity* calculated forward-in-time as

$$\boldsymbol{\gamma}^f(t_0, t_M) = \mathbf{F}^{t_0 \to t_1} \mathbf{F}^{t_1 \to t_2} \ldots \mathbf{F}^{t_{M-2} \to t_{M-1}} \mathbf{F}^{t_{M-1} \to t_M}. \tag{9}$$

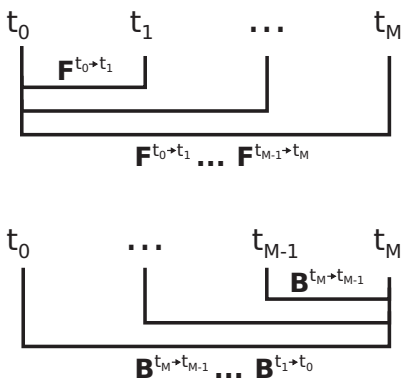


FIG. 2. Several consecutive multistep windows are used together to calculate cumulated connectivity matrices. Each window is defined by an initial time, a final time and a certain number of steps in between. One of the two times is kept fixed (either the initial or the final one) while the other is moving while it draws windows with a progressively larger number of steps. Specifically, in the forward-in-time case, the initial time t_0 is fixed and we increase the number of steps ending at larger t_l's up to t_M (top panel). Going backward-in-time, we instead end always at t_M but starting from decreasing t_l's until reaching t_0 (bottom panel).

Similarly, we define the *noncumulated multistep explicit connectivity* calculated backward-in-time:

$$\boldsymbol{\gamma}^b(t_0, t_M) = \mathbf{B}^{t_M \to t_{M-1}} \mathbf{B}^{t_{M-1} \to t_{M-2}} \dots \mathbf{B}^{t_2 \to t_1} \mathbf{B}^{t_1 \to t_0}. \quad (10)$$

In both definitions above we used the fact that summing probabilities over all the paths corresponds to performing the matrix product of the associated adjacency matrices. Therefore, $\boldsymbol{\gamma}^f$ is a matrix whose element $i - j$ is the probability for a random walker to reach j starting from i after M-steps forward-in-time. Similarly, $\boldsymbol{\gamma}^b$ is a matrix whose element $i - j$ is the probability for a random walker to reach j starting from i after M-steps backward-in-time. It is straightforward to prove that the matrix elements of both $\boldsymbol{\gamma}^f$ and $\boldsymbol{\gamma}^b$ are always bounded in between 0 and 1.

C. Cumulated multistep explicit connectivity

We now consider the case of multiple numbers of steps together to introduce the novel concept of cumulated connectivity (see Fig. 2). We still refer to the discrete time sequence $\{t_0, t_1, \dots, t_{M-1}, t_M\}$ introduced before and we provide the probability for a random walker to connect two nodes in a finite range of possible number of steps. In the forward-in-time case, the initial time t_0 is fixed and the number of steps increases progressively ending up at larger t_l's. Backward-in-time, we instead end always at t_M but starting from decreasing t_l's while increasing the number of steps. Without loss of generality, we consider in the following multistep connectivity realized in a range of number of steps comprised between 1 and M.

1. Deriving up to three-steps cumulated multistep explicit connectivity

Let us start considering the forward-in-time connectivity. Keeping fixed the initial time t_0, we focus on a starting node i and a destination node j and we aim to find an expression for the union of the multistep explicit probabilities of one, two, and three steps. We start defining the following three events:

(1) A: reaching j from i in one step,
(2) B: reaching j from i in two steps,
(3) C: reaching j from i in three steps.

Since the events connecting i to j in different numbers of steps are not mutually exclusive, we cannot obtain the three-events union probability by simply summing their individual probabilities. Such union probability, which will be used to define the cumulated multistep explicit connectivity, can be written as

$$\begin{aligned} P(A \cup B \cup C)_{ij} &= P(A) + P(B) + P(C) + P(A \cap B \cap C) - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &= P(A) + P(B) + P(C) + P(A)P(B|A)P(C|A \cap B) - P(A)P(B|A) - P(A)P(C|A) - P(B)P(C|B). \end{aligned} \quad (11)$$

Following Eq. (9), the multistep connectivity probabilities of the events A, B, and C are

$$P(A) = \mathbf{F}_{ij}^{t_0 \to t_1}, \quad (12)$$

$$P(B) = \sum_k \mathbf{F}_{ik}^{t_0 \to t_1} \mathbf{F}_{kj}^{t_1 \to t_2}, \quad (13)$$

$$P(C) = \sum_{k,l} \mathbf{F}_{ik}^{t_0 \to t_1} \mathbf{F}_{kl}^{t_1 \to t_2} \mathbf{F}_{lj}^{t_2 \to t_3}. \quad (14)$$

Using Eqs. (12), (13), and (14) into Eq. (11) we obtain

$$\begin{aligned} P(A \cup B \cup C)_{ij} = \mathbf{F}_{ij}^{t_0 \to t_1} &+ \sum_k \mathbf{F}_{ik}^{t_0 \to t_1} \mathbf{F}_{kj}^{t_1 \to t_2} + \sum_{k,l} \mathbf{F}_{ik}^{t_0 \to t_1} \mathbf{F}_{kl}^{t_1 \to t_2} \mathbf{F}_{lj}^{t_2 \to t_3} + \mathbf{F}_{ij}^{t_0 \to t_1} \mathbf{F}_{jj}^{t_1 \to t_2} \mathbf{F}_{jj}^{t_2 \to t_3} \\ &- \mathbf{F}_{ij}^{t_0 \to t_1} \mathbf{F}_{jj}^{t_1 \to t_2} - \sum_l \mathbf{F}_{ij}^{t_0 \to t_1} \mathbf{F}_{jl}^{t_1 \to t_2} \mathbf{F}_{lj}^{t_2 \to t_3} - \sum_k \mathbf{F}_{ik}^{t_0 \to t_1} \mathbf{F}_{kj}^{t_1 \to t_2} \mathbf{F}_{jj}^{t_2 \to t_3}. \end{aligned} \quad (15)$$

Developing the second and the third terms in Eq. (15), we find

$$P(A\cup B\cup C)_{ij} = \mathbf{F}_{ij}^{t_0\to t_1} + \sum_{k\neq j}\mathbf{F}_{ik}^{t_0\to t_1}\mathbf{F}_{kj}^{t_1\to t_2} + \sum_{k\neq j}\sum_{l\neq j}\mathbf{F}_{ik}^{t_0\to t_1}\mathbf{F}_{kl}^{t_1\to t_2}\mathbf{F}_{lj}^{t_2\to t_3}$$
$$= \mathbf{F}_{ij}^{t_0\to t_1} + \sum_{k}\mathbf{F}_{ik}^{t_0\to t_1}(1-\delta_{kj})\mathbf{F}_{kj}^{t_1\to t_2} + \sum_{k,l}\mathbf{F}_{ik}^{t_0\to t_1}(1-\delta_{kj})\mathbf{F}_{kl}^{t_1\to t_2}(1-\delta_{lj})\mathbf{F}_{lj}^{t_2\to t_3}. \quad (16)$$

From a geometrical point of view, impeding the indexes k and l from taking the value of j means excluding the contribution to the union probability of all the paths that visit j more than once.

Denoting with a circle the Hadamard (or element-wise) product, we can write Eq. (16) for any pair i-j in a compact form and define the matrix:

$$\mathbf{\Gamma}^{f}(t_0,t_3) = \mathbf{F}^{t_0\to t_1} + \mathbf{F}^{t_0\to t_1}(\mathbb{L}\circ\mathbf{F}^{t_1\to t_2}) + \mathbf{F}^{t_0\to t_1}\{\mathbb{L}\circ[\mathbf{F}^{t_1\to t_2}(\mathbb{L}\circ\mathbf{F}^{t_2\to t_3})]\}, \quad (17)$$

where $\mathbb{L}$ is the all-ones matrix minus the identity matrix, i.e., $\mathbb{L} = \mathbb{J} - \mathbb{I}$ and $\mathbb{L}_{ij} = (1-\delta_{ij})$.

2. Generalizing up to M-steps cumulated multistep explicit connectivity

To generalize the result from the previous section, we consider the probability of the union of M different events $A_1,\ldots,A_M$ and, using the inclusion-exclusion formula, we can write such probability as

$$P\left(\bigcup_{i=1}^{M}A_i\right) = \sum_{i_1=1}^{M}P(A_{i_1}) - \sum_{i_1<i_2}^{M}P(A_{i_1}\cap A_{i_2}) + \cdots + (-1)^{M-1}\sum_{i_1<\cdots<i_M}^{M}P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_M})$$
$$= P(A_1) + P\big(A_1^c\cap A_2\big) + P\big(A_1^c\cap A_2^c\cap A_3\big) + \cdots + P\big(A_1^c\cap\cdots\cap A_{M-1}^c\cap A_M\big). \quad (18)$$

Expanding Eq. (18) we find an expression for the probability union that is a generalization of Eq. (17) to the generic case of M-steps. Keeping fixed the initial time t_0, we define thus the *cumulated multistep explicit connectivity* calculated forward-in-time as

$$\mathbf{\Gamma}^{f}(t_0,t_M) = \mathbf{F}^{t_0\to t_1} + \mathbf{F}^{t_0\to t_1}(\mathbb{L}\circ\mathbf{F}^{t_1\to t_2}) + \mathbf{F}^{t_0\to t_1}\big\{\mathbb{L}\circ\big[\mathbf{F}^{t_1\to t_2}(\mathbb{L}\circ\mathbf{F}^{t_2\to t_3})\big]\big\}$$
$$+\cdots+\mathbf{F}^{t_0\to t_1}\big\{\mathbb{L}\circ\big[\mathbf{F}^{t_1\to t_2}\ldots(\mathbb{L}\circ\mathbf{F}^{t_{M-1}\to t_M})\ldots\big]\big\}. \quad (19)$$

Similarly, keeping fixed instead the final time t_M, we derive the *cumulated multistep explicit connectivity* calculated backward-in-time:

$$\mathbf{\Gamma}^{b}(t_0,t_M) = \mathbf{B}^{t_M\to t_{M-1}} + \mathbf{B}^{t_M\to t_{M-1}}(\mathbb{L}\circ\mathbf{B}^{t_{M-1}\to t_{M-2}}) + \mathbf{B}^{t_M\to t_{M-1}}\big\{\mathbb{L}\circ\big[\mathbf{B}^{t_{M-1}\to t_{M-2}}(\mathbb{L}\circ\mathbf{B}^{t_{M-2}\to t_{M-3}})\big]\big\}$$
$$+\cdots+\mathbf{B}^{t_M\to t_{M-1}}\big\{\mathbb{L}\circ\big[\mathbf{B}^{t_{M-1}\to t_{M-2}}\ldots(\mathbb{L}\circ\mathbf{B}^{t_1\to t_0})\ldots\big]\big\}. \quad (20)$$

Hence, $\mathbf{\Gamma}^{f}$ and $\mathbf{\Gamma}^{b}$ provide the expected probabilities for a random walker to connect pairs of nodes in a range of possible number of steps comprised between 1 and M, forward- and backward-in-time respectively. Consequently, $\mathbf{\Gamma}^{f}$ corresponds also to the probability that a portion of quantity Q tagged in node i arrives into node j, forward-in-time. Similarly, $\mathbf{\Gamma}^{b}$ corresponds to the probability that a portion of sampled quantity Q in node i comes from node j, backward-in-time.

3. Bounding M-steps cumulated multistep explicit connectivity probabilities

Let us consider the forward-in-time dynamics (the same argument holds for the backward-in-time case) and write down Eq. (19) for a specific matrix element associated with the origin node i and destination node j, we have

$$\mathbf{F}_{ij}^{t_0\to t_1} + \sum_{k_1\neq j}\mathbf{F}_{ik_1}^{t_0\to t_1}\mathbf{F}_{k_1j}^{t_1\to t_2} + \sum_{k_1\neq j}\sum_{k_2\neq j}\mathbf{F}_{ik_1}^{t_0\to t_1}\mathbf{F}_{k_1k_2}^{t_1\to t_2}\mathbf{F}_{k_2j}^{t_2\to t_3} + \sum_{k_1\neq j}\cdots\sum_{k_{M-1}\neq j}\mathbf{F}_{ik_1}^{t_0\to t_1}\cdots\mathbf{F}_{k_{M-1}j}^{t_{M-1}\to t_M}$$
$$= \mathbf{F}_{ij}^{t_0\to t_1} + \sum_{k_1\neq j}\mathbf{F}_{ik_1}^{t_0\to t_1}\left\{\mathbf{F}_{k_1j}^{t_1\to t_2} + \sum_{k_2\neq j}\mathbf{F}_{k_1k_2}^{t_1\to t_2}\left[\cdots\left(\mathbf{F}_{k_{M-2}j}^{t_{M-2}\to t_{M-1}} + \sum_{k_{M-1}\neq j}\mathbf{F}_{k_{M-2}k_{M-1}}^{t_{M-2}\to t_{M-1}}\mathbf{F}_{k_{M-1}j}^{t_{M-1}\to t_M}\right)\cdots\right]\right\}. \quad (21)$$

Recalling that $\sum_j \mathbf{F}_{ij} = 1$ and $\mathbf{F}_{ij}\leqslant 1$, we note that the quantity in the inner parenthesis is bounded to 1. This automatically bounds to 1 the quantity in the more external parenthesis. Recursively, we can finally see that all the expression is bounded to 1 too.

IV. IMPLICIT CONNECTIVITY

Paths are not the only connectivity patterns that can be found in a network. In general, one can identify different network motifs composed of an arbitrary number of links and nodes. Such motifs are expected to be associated with different dynamical processes depending on their geometry. In particular, we focus here on the so-called *pitchforks* motifs and their associated random walk connectivity pattern that we call *implicit connectivity*. We define pitchforks as a particular subgroup of motifs composed of three (sometimes two) nodes and two links. We call converging pitchfork a motif of three (or two) nodes and two links pointing to one of them; we call instead diverging pitchfork a motif of three (or two) nodes and two links emanated from one them (see Fig. 3).

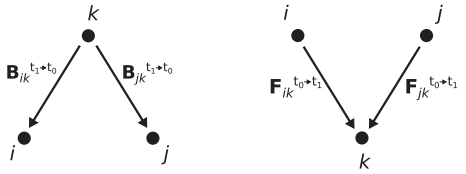


FIG. 3. Schematic representation of diverging (left) and converging (right) pitchforks. Black dots represent network nodes, arrows symbolize directed temporal links.

FIG. 4. Degenerate pitchforks composed of two nodes instead of three, i.e., when $k = i$ or $k = j$ respectively. Black dots represent network nodes, arrows symbolize directed temporal links.

We relate such motifs to an implicit relationship between two nodes i-j that are somehow influenced (or influencing) by a third node k. If "third-party" nodes k's are more than one for a given pair i-j, then we consider them together summing over k. The strength of these implicit relationships can be associated with the probability that two random walkers starting (or arriving) in k end up (or come from) one in i and the other in j. Similarly, the probability can be summed over k to obtain the global implicit connection probability for i-j. As for the explicit case (Sec. III), we can derive (see below) both noncumulated and cumulated implicit connectivity probabilities and, consequently, relate these probabilities with portions of a quantity Q transported across the network (see also Sec. II B).

Note that implicit connections studied here happen "synchronously." For the temporal case, it means that both random walkers ensuring connections start from (or end up in) node k at the same time. It is also the same time at which they reach (or start from) node i and j, respectively. For static networks, it means that we consider for each single noncumulated connection two random walks of the same number of steps. From a physical perspective, this is tantamount to sampling/tagging a transported quantity at the same time. This requirement is consistent with the fact that, for any dynamics running on the network, the states of each node would change in time so that it would be difficult to interpret nonsynchronous relationships. More generally, if we look for a correct synoptic view of a system, then we need to consider comparable snapshots of the associated network, i.e., matching time intervals (temporal case) or the degrees of separation (static case). This would be the case, for instance, when studying indirect interactions between competitors for the same resource in food webs, shared "influencers" of opinions in social systems or common sources of pollutants in fluid transport networks.

A. Pitchfork motifs and the implicit connectivity concept

1. Single pitchfork motifs

From now on, let us focus on diverging pitchforks (an analogous approach can be used for the converging ones) over a time interval $[t_0; t_1]$. Both links composing the pitchfork emanated from the "source" node k and point to nodes i and j. We look for the probability that two random walkers, released simultaneously in i and j at t_1, moving backward-in-time, arrive together into k at t_0. Such probability can be related to a sampling process on the pair $i - j$. Indeed, if we take a sample of the quantity Q in i at time t_1, the probability that such sample was in k at time t_0 is $\mathbf{B}_{ik}^{t_1 \to t_0}$. Similarly, the probability for j would be $\mathbf{B}_{jk}^{t_1 \to t_0}$. Hence, if we sample simultaneously in i and j at t_1 the probability that both samples were in k at t_0 is

$$\mathbf{B}_{ik}^{t_1 \to t_0} \mathbf{B}_{jk}^{t_1 \to t_0}. \tag{22}$$

Note that in the particular case (called here degenerate pitchfork) for which $k = i$ or $k = j$ the formulation is conceptually consistent. For instance, for $k = i$ the probability of Eq. (22) becomes $\mathbf{B}_{ii}^{t_1 \to t_0} \mathbf{B}_{ji}^{t_1 \to t_0}$ and the node i acts as source as well as destination (see Fig. 4).

2. Summing over pitchforks

We now address a more general question: if one samples a quantity in nodes i and j at t_1, then what is the probability that both samples share the same origin at t_0 (regardless of the origin nodes)? This is equivalent to looking for the probability that two random walkers, released simultaneously in i and j at t_1, arrive backward-in-time into the same node at t_0. By generalizing Eq. (22), such probability is the simple sum over all the k nodes that form a pitchfork with i and j. This is because (i) the probability that a sample in i comes from k is independent from the probability that a sample in j comes from k and (ii) sampling quantities coming from different k's inside a single node are mutually-exclusive events. We associate this backward-in-time total probability with what we call as *implicit connectivity* and we define it as:

$$\mathbf{I}_{ij}^{t_1 \to t_0} = \mathbf{I}_{ji}^{t_1 \to t_0} = \sum_k \mathbf{B}_{ik}^{t_1 \to t_0} \mathbf{B}_{jk}^{t_1 \to t_0} = (\mathbf{B}^{t_1 \to t_0}\, {}^{\mathrm{T}}\mathbf{B}^{t_1 \to t_0})_{ij}, \tag{23}$$

where with ${}^{\mathrm{T}}\mathbf{B}^{t_1 \to t_0}$ we denote the transpose of $\mathbf{B}^{t_1 \to t_0}$.

Note that when $i = j$ we have $\mathbf{I}_{ii}^{t_1 \to t_0} = \sum_k (\mathbf{B}_{ik}^{t_1 \to t_0})^2$ that corresponds to the probability that two random samples of the quantity in i came from the same origin (assuming a sampling with replacement). This measure corresponds to the backward-in-time Rényi-entropy for $q = 2$ of the node i defined in Ref. [12]. Interestingly, $\mathbf{I}_{ii}^{t_1 \to t_0}$ is also related to the definition of the well known Simpson index and could be interpreted thus as a measure of diversity of origins of the quantity contained in i.

For the case of converging pitchforks an analogous development can be done. Indeed, when the two links converge to a common "destination" node k, we can calculate the probability that two random walkers, released simultaneously in i and j at t_0, moving forward-in-time arrive together into k at t_1. Such probability corresponds also to the chance that given portions of tagged quantity in i and j at t_0 will reach simultaneously k at t_1.

3. Bounding implicit connectivity probability

We want to prove that $\mathbf{I}_{ij}^{t_1\to t_0} \leqslant 1$ for every i, j. Using that $\mathbf{B}_{jk}^{t_1\to t_0} \leqslant 1$ and $\mathbf{B}_{ik}^{t_1\to t_0} \leqslant 1$ and that $\sum_k \mathbf{B}_{ik}^{t_1\to t_0} = 1$ and $\sum_k \mathbf{B}_{jk}^{t_1\to t_0} = 1$ one can easily find the following relationship:

$$\mathbf{I}_{ij}^{t_1\to t_0} \leqslant \sum_k \mathbf{B}_{ik}^{t_1\to t_0} = 1, \tag{24}$$

$$\mathbf{I}_{ij}^{t_1\to t_0} \leqslant \sum_k \mathbf{B}_{jk}^{t_1\to t_0} = 1. \tag{25}$$

B. Noncumulated multistep implicit connectivity

Here, analogously to what we did for explicit connectivity (Sec. III), we first define the *noncumulated multistep implicit connectivity* by focusing on a fixed number of M steps (instead of single links). We develop only the case of backward implicit connectivity but the same reasoning can be used for forward-in-time dynamics.

For $M = 2$, the multistep implicit connectivity between node i and j is denoted as

$$\sum_k \left(\sum_l \mathbf{B}_{il}^{t_2\to t_1} \mathbf{B}_{lk}^{t_1\to t_0} \sum_m \mathbf{B}_{jm}^{t_2\to t_1} \mathbf{B}_{mk}^{t_1\to t_0} \right) = [(\mathbf{B}^{t_2\to t_1}\mathbf{B}^{t_1\to t_0})^{\mathbf{T}}(\mathbf{B}^{t_2\to t_1}\mathbf{B}^{t_1\to t_0})]_{ij}. \tag{26}$$

We can generalize Eq. (26) formula to M steps to define the *noncumulated multistep implicit connectivity* calculated forward-in-time in matrix form as

$$\boldsymbol{\lambda}^{f}(t_0, t_M) = \left(\mathbf{F}^{t_0\to t_1} \dots \mathbf{F}^{t_{M-1}\to t_M}\right)^{\mathbf{T}} \left(\mathbf{F}^{t_0\to t_1} \dots \mathbf{F}^{t_{M-1}\to t_M}\right) \tag{27}$$

and the matrix form of the *noncumulated multistep implicit connectivity* calculated backward-in-time as

$$\boldsymbol{\lambda}^{b}(t_0, t_M) = \left(\mathbf{B}^{t_M\to t_{M-1}} \dots \mathbf{B}^{t_1\to t_0}\right)^{\mathbf{T}} \left(\mathbf{B}^{t_M\to t_{M-1}} \dots \mathbf{B}^{t_1\to t_0}\right). \tag{28}$$

C. Cumulated multistep implicit connectivity

Similarly to Sec. III C, we now further consider the case of multiple numbers of steps to introduce the *cumulated multistep implicit connectivity* (see Fig. 2). We refer gain to the discrete time sequence $\{t_0, t_1, \dots, t_{M-1}, t_M\}$ introduced before and, without loss of generality, we consider multistep connectivity realized in any number of steps comprised between 1 and M. In other words, we look for a generic analytical expression to obtain the probability of linking two nodes by implicit connections occurring over a range of possible number of steps. For forward-in-time dynamics, the initial time t_0 is fixed while the number of steps considered increase successively up to largest t_l's. Backward-in-time, the final time t_M is fixed while the number of steps considered starts from the lowest t_l's and increases successively.

1. Deriving up to 3-steps cumulated multistep implicit connectivity

We consider in the following the forward-in-time implicit connectivity and, as before, all the derivations are similar for the backward-in-time case. Keeping fixed the initial time t_0, we focus on the nodes i and j and we want to find an expression for the union of the multistep implicit probabilities increasing progressively the number of steps from 1 to M. Since the probabilities at different numbers of steps are not mutually exclusive, we cannot use the simple probability sum. First, we evaluate the probability union from one to three steps and then we generalize it up to a generic M. We define the three events:

(1) A: reaching the same destination from i and j in one step,

(2) B: reaching the same destination from i and j in two steps,

(3) C: reaching the same destination from i and j in three steps.

The union of the probabilities of the above three events, which we call cumulated multistep implicit connectivity, is derived from Eq. (11). Following Eq. (27), we have

$$P(A) = \sum_k \mathbf{F}_{ik}^{t_0\to t_1} \mathbf{F}_{jk}^{t_0\to t_1}, \tag{29}$$

$$P(B) = \sum_k \left(\sum_l \mathbf{F}_{il}^{t_0\to t_1} \mathbf{F}_{lk}^{t_1\to t_2} \sum_m \mathbf{F}_{jm}^{t_0\to t_1} \mathbf{F}_{mk}^{t_1\to t_2} \right), \tag{30}$$

$$P(C) = \sum_k \left(\sum_{l,f} \mathbf{F}_{il}^{t_0\to t_1} \mathbf{F}_{lf}^{t_1\to t_2} \mathbf{F}_{fk}^{t_2\to t_3} \sum_{m,g} \mathbf{F}_{jm}^{t_0\to t_1} \mathbf{F}_{mg}^{t_1\to t_2} \mathbf{F}_{gk}^{t_2\to t_3} \right). \tag{31}$$

Consequently, the remaining terms of Eq. (11) are

$$P(A)P(B|A) = \sum_k \left(\sum_l \mathbf{F}_{il}^{t_0\to t_1} \mathbf{F}_{lk}^{t_1\to t_2} \mathbf{F}_{jl}^{t_0\to t_1} \mathbf{F}_{lk}^{t_1\to t_2} \right), \tag{32}$$

$$P(B)P(C|B) = \sum_k \left(\sum_{l,f} \mathbf{F}_{il}^{t_0\to t_1} \mathbf{F}_{lf}^{t_1\to t_2} \mathbf{F}_{fk}^{t_2\to t_3} \sum_m \mathbf{F}_{jm}^{t_0\to t_1} \mathbf{F}_{mf}^{t_1\to t_2} \mathbf{F}_{fk}^{t_2\to t_3} \right), \tag{33}$$

$$P(A)P(C|A) = \sum_k \left(\sum_{l,f} \mathbf{F}_{il}^{t_0\to t_1} \mathbf{F}_{lf}^{t_1\to t_2} \mathbf{F}_{fk}^{t_2\to t_3} \sum_g \mathbf{F}_{jl}^{t_0\to t_1} \mathbf{F}_{lg}^{t_1\to t_2} \mathbf{F}_{gk}^{t_2\to t_3} \right), \tag{34}$$

$$P(A)P(B|A)P(C|A\cap B)=\sum_k\left(\sum_{l,f}\mathbf{F}^{t_0\to t_1}_{il}\mathbf{F}^{t_1\to t_2}_{lf}\mathbf{F}^{t_2\to t_3}_{fk}\mathbf{F}^{t_0\to t_1}_{jl}\mathbf{F}^{t_1\to t_2}_{lf}\mathbf{F}^{t_2\to t_3}_{fk}\right). \tag{35}$$

Developing properly the sum $\sum_m$ in Eq. (30) and the sums $\sum_{m,g}$ in Eq. (31), the contributions from Eqs. (32)–(35) cancel out inside Eq. (11) and we finally find

$$\begin{aligned}P(A\cup B\cup C)_{ij}=&\sum_k\mathbf{F}^{t_0\to t_1}_{ik}\mathbf{F}^{t_0\to t_1}_{jk}+\sum_k\left(\sum_l\mathbf{F}^{t_0\to t_1}_{il}\mathbf{F}^{t_1\to t_2}_{lk}\sum_{m\neq l}\mathbf{F}^{t_0\to t_1}_{jm}\mathbf{F}^{t_1\to t_2}_{mk}\right)\\&+\sum_k\left(\sum_{l,f}\mathbf{F}^{t_0\to t_1}_{il}\mathbf{F}^{t_1\to t_2}_{lf}\mathbf{F}^{t_2\to t_3}_{fk}\sum_{m\neq l,g\neq f}\mathbf{F}^{t_0\to t_1}_{jm}\mathbf{F}^{t_1\to t_2}_{mg}\mathbf{F}^{t_2\to t_3}_{gk}\right).\end{aligned} \tag{36}$$

From a geometrical point of view, preventing the indexes m and g from taking the values of l and f corresponds to excluding the paths starting from i and j that converge to any common destination node more than once. Recalling the definition of the matrix $\mathbb{L}$ as the all-ones matrix minus the identity matrix i.e., $\mathbb{L}=\mathbb{J}-\mathbb{I}$, we can write Eq. (36) as

$$\begin{aligned}P(A\cup B\cup C)_{ij}=&\sum_k\mathbf{F}^{t_0\to t_1}_{ik}\mathbf{F}^{t_0\to t_1}_{jk}+\sum_k\left(\sum_{l,m}\mathbf{F}^{t_0\to t_1}_{il}\mathbf{F}^{t_1\to t_2}_{lk}\mathbb{L}_{lm}\mathbf{F}^{t_0\to t_1}_{jm}\mathbf{F}^{t_1\to t_2}_{mk}\right)\\&+\sum_k\left(\sum_{l,f,m,g}\mathbf{F}^{t_0\to t_1}_{il}\mathbf{F}^{t_1\to t_2}_{lf}\mathbf{F}^{t_2\to t_3}_{fk}\mathbb{L}_{lm}\mathbb{L}_{fg}\mathbf{F}^{t_0\to t_1}_{jm}\mathbf{F}^{t_1\to t_2}_{mg}\mathbf{F}^{t_2\to t_3}_{gk}\right).\end{aligned} \tag{37}$$

By using the Hadamard product and performing some transpositions, we can finally find an expression for Eq. (37), for every pair $i-j$, in a compact form and define the matrix:

$$\begin{aligned}\mathbf{\Lambda}^{b}(t_0,t_3)=&\,\mathbf{F}^{t_0\to t_1}\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}+\mathbf{F}^{t_0\to t_1}[\mathbb{L}\circ(\mathbf{F}^{t_1\to t_2}\,{}^{\mathbf{T}}\mathbf{F}^{t_1\to t_2})]\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}\\&+\mathbf{F}^{t_0\to t_1}(\mathbb{L}\circ\{\mathbf{F}^{t_1\to t_2}[\mathbb{L}\circ(\mathbf{F}^{t_2\to t_3}\,{}^{\mathbf{T}}\mathbf{F}^{t_2\to t_3})]\,{}^{\mathbf{T}}\mathbf{F}^{t_1\to t_2}\})\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}.\end{aligned} \tag{38}$$

2. Generalizing up to M-steps cumulated multistep implicit connectivity

To generalize the result derived in the previous section, we consider the probability union of M different events $A_1,\ldots,A_M$ using the inclusion-exclusion formula of Eq. (18). Keeping fixed the initial time t_0, we derive thus the *cumulated multistep implicit connectivity* calculated forward-in-time:

$$\begin{aligned}\mathbf{\Lambda}^{f}(t_0,t_M)=&\,\mathbf{F}^{t_0\to t_1}\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}+\mathbf{F}^{t_0\to t_1}[\mathbb{L}\circ(\mathbf{F}^{t_1\to t_2}\,{}^{\mathbf{T}}\mathbf{F}^{t_1\to t_2})]\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}\\&+\mathbf{F}^{t_0\to t_1}(\mathbb{L}\circ\{\mathbf{F}^{t_1\to t_2}[\mathbb{L}\circ(\mathbf{F}^{t_2\to t_3}\,{}^{\mathbf{T}}\mathbf{F}^{t_2\to t_3})]\,{}^{\mathbf{T}}\mathbf{F}^{t_1\to t_2}\})\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}\ldots\\&+\mathbf{F}^{t_0\to t_1}[\mathbb{L}\circ(\mathbf{F}^{t_1\to t_2}\{\mathbb{L}\circ[\ldots(\mathbf{F}^{t_{M-1}\to t_M}\,{}^{\mathbf{T}}\mathbf{F}^{t_{M-1}\to t_M})\ldots]\}\,{}^{\mathbf{T}}\mathbf{F}^{t_1\to t_2})]\,{}^{\mathbf{T}}\mathbf{F}^{t_0\to t_1}.\end{aligned} \tag{39}$$

Similarly, keeping fixed instead the final time t_M, we derive the *cumulated multistep implicit connectivity* calculated backward-in-time:

$$\begin{aligned}\mathbf{\Lambda}^{b}(t_0,t_M)=&\,\mathbf{B}^{t_M\to t_{M-1}}\,{}^{\mathbf{T}}\mathbf{B}^{t_M\to t_{M-1}}+\mathbf{B}^{t_M\to t_{M-1}}[\mathbb{L}\circ(\mathbf{B}^{t_{M-1}\to t_{M-2}}\,{}^{\mathbf{T}}\mathbf{B}^{t_{M-1}\to t_{M-2}})]\,{}^{\mathbf{T}}\mathbf{B}^{t_M\to t_{M-1}}\\&+\mathbf{B}^{t_M\to t_{M-1}}(\mathbb{L}\circ\{\mathbf{B}^{t_{M-1}\to t_{M-2}}[\mathbb{L}\circ(\mathbf{B}^{t_{M-2}\to t_{M-3}}\,{}^{\mathbf{T}}\mathbf{B}^{t_{M-2}\to t_{M-3}})]\,{}^{\mathbf{T}}\mathbf{B}^{t_{M-1}\to t_{M-2}}\})\,{}^{\mathbf{T}}\mathbf{B}^{t_M\to t_{M-1}}\ldots\\&+\mathbf{B}^{t_M\to t_{M-1}}[\mathbb{L}\circ(\mathbf{B}^{t_{M-1}\to t_{M-2}}\{\mathbb{L}\circ[\ldots(\mathbf{B}^{t_1\to t_0}\,{}^{\mathbf{T}}\mathbf{B}^{t_1\to t_0})\ldots]\}\,{}^{\mathbf{T}}\mathbf{B}^{t_{M-1}\to t_{M-2}})]\,{}^{\mathbf{T}}\mathbf{B}^{t_M\to t_{M-1}}.\end{aligned} \tag{40}$$

Hence, $\mathbf{\Lambda}^{f}$ and $\mathbf{\Lambda}^{b}$ provide the expected probabilities for two random walkers released at the same time in two nodes of the network of arriving both into the same node in over a range of steps comprised between 1 and M, forward- and backward-in-time respectively. Consequently, $\mathbf{\Lambda}^{f}$ corresponds also to the probability that two portions of tagged quantity Q from a node pair will arrive to the same node forward-in-time. Similarly, $\mathbf{\Lambda}^{b}$ corresponds to the probability that two portions of sampled quantity Q from a node pair come from the same node backward-in-time.

V. EXAMPLE APPLICATIONS BASED ON SIMPLE NETWORKS

We now study the performance of our novel connectivity metrics when applied on two simple, static networks represented in Fig. 5. To start simple, we compare the analytical and the numerical results for the forward-in-time cumulated multistep explicit and implicit connectivity assuming no time-dependence (Table I). In Table I we report the values of $\mathbf{\Gamma}^{f}$ and $\mathbf{\Lambda}^{f}$ with $M=1,5$ and 100 for every pair of nodes in both networks. We then perform numerical experiments releasing

TABLE I. Tables reporting connectivity values of $\mathbf{\Gamma}^f$ and $\mathbf{\Lambda}^f$ with $M = 1, 5$, and 100 for every pair of nodes of the two example networks shown in Fig. 5. The left table reports values for the network A and the right table reports values for the network B.

	Γ^f			Λ^f		
	$M=1$	$M=5$	$M=100$	$M=1$	$M=5$	$M=100$
$a \to a$	0	0.855	1	0.52	0.84976	1
$a \to b$	0.6	0.9856	1	0.3	0.67607	1
$a \to c$	0.4	0.58	1	0.42	0.755562	1
$a \to d$	0	0.9275	1	0	0.6402272	1
$b \to a$	0	0.9375	1	0.3	0.67607	1
$b \to b$	0.5	0.94	1	0.5	0.73775	1
$b \to c$	0	0.35	1	0.5	0.73775	1
$b \to d$	0.5	0.96875	1	0	0.57058	1
$c \to a$	0	0.9125	1	0.42	0.755562	1
$c \to b$	0.7	0.964	1	0.5	0.73775	1
$c \to c$	0	0.33	1	0.58	0.77971	1
$c \to d$	0.3	0.95625	1	0	0.58398	1
$d \to a$	1	1.0	1	0	0.6402272	1
$d \to b$	0	0.952	1	0	0.57058	1
$d \to c$	0	0.52	1	0	0.58398	1
$d \to d$	0	0.855	1	1	1	1

	Γ^f			Λ^f		
	$M=1$	$M=5$	$M=100$	$M=1$	$M=5$	$M=100$
$a \to a$	0	0.7	0.7	0.58	0.706	1
$a \to b$	0.7	0.7	0.7	0	0.153	1
$a \to c$	0.3	0.51	1	0.3	0.51	1
$a \to d$	0	0.7	0.7	0	0.2601	1
$b \to a$	0	1	1	0	0.153	1
$b \to b$	0	0.7	0.7	1	1	1
$b \to c$	0	0.3	1	0	0.3	1
$b \to d$	1	1	1	0	0.153	1
$c \to a$	0	0	0	0.3	0.51	1
$c \to b$	0	0	0	0	0.3	1
$c \to c$	1	1	1	1	1	1
$c \to d$	0	0	0	0	0.51	1
$d \to a$	1	1.0	1	0	0.2601	1
$d \to b$	0	0.7	0.7	0	0.153	1
$d \to c$	0	0.51	1	0	0.51	1
$d \to d$	0	0.7	0.7	1	1	1

thousands of random walkers across the networks verifying that their encounter probabilities match perfectly the values of $\mathbf{\Gamma}^f$ and $\mathbf{\Lambda}^f$. As highlighted in Fig. 6, we clearly see that explicit and implicit connectivity present marked differences. Moreover, the results show that, while explicit connectivity is not necessarily symmetric with respect to i and j, implicit connectivity is symmetric by definition (i.e., $\mathbf{\Lambda}^f{}_{ij} = \mathbf{\Lambda}^f{}_{ji}$). We note also that, for explicit connectivity with $M = 100$, the probabilities saturate to one only for the network A of Fig. 5. This can be explained by the fact that the network A is strongly connected while the network B is not, for such reason the node c in the network B acts as an absorbing state for random walkers impeding the saturation to one of all probabilities.

To include the effect of temporal dynamics and highlight its relevance for connectivity patterns, we also study the case of a temporal network. To this aim, we consider the network A of Fig. 5 and we cyclically modify some of its weights while keeping the average equal to the original static network. In this way, we can properly assess the differences between a temporal network and its aggregated static counterpart. Specifically, we use the following temporal weights sequences:

$$\begin{aligned} a \to b &: (0.6, 0.5, 0.4, 0.7, 0.8), \\ a \to c &: (0.4, 0.5, 0.6, 0.3, 0.2), \\ c \to b &: (0.7, 0.6, 0.5, 0.8, 0.9), \\ c \to d &: (0.3, 0.4, 0.5, 0.2, 0.1), \end{aligned}$$

where each sequence describes the weights of a link for five time intervals and then is repeated, the other weights are kept constant in time as in network A. In Table II we report the values of $\mathbf{\Gamma}^f$ and $\mathbf{\Lambda}^f$ with $M = 1, 5$, and 100 for every pair of nodes in both networks. We note that, consistently, for $M = 1$

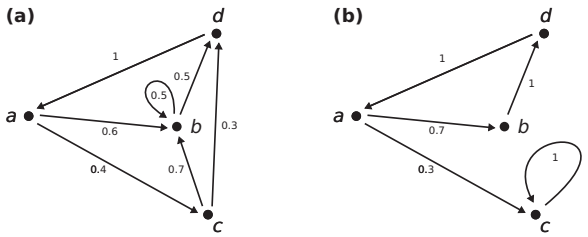


FIG. 5. Examples of two small static networks: a strongly connected network A (left panel) and a weakly connected network B (right panel). Black dots represent network nodes, arrows symbolize directed static links. Small letters label different nodes and numbers are forward-in-time probabilities of transition associated with each link. Note that while the network A is strongly connected, the network B is not.

TABLE II. Table reporting connectivity values of $\mathbf{\Gamma}^f$ and $\mathbf{\Lambda}^f$ with $M = 1, 5$, and 100 for every pair of nodes of the temporal version of the network A shown in Fig. 5.

	Γ^f			Λ^f		
	$M=1$	$M=5$	$M=100$	$M=1$	$M=5$	$M=100$
$a \to a$	0	0.865	1	0.52	0.8705	1
$a \to b$	0.6	0.9952	1	0.3	0.7102	1
$a \to c$	0.4	0.52	1	0.42	0.7829	1
$a \to d$	0	0.9325	1	0	0.6438	1
$b \to a$	0	0.9375	1	0.3	0.7102	1
$b \to b$	0.5	0.94	1	0.5	0.766	1
$b \to c$	0	0.4	1	0.5	0.766	1
$b \to d$	0.5	0.96875	1	0	0.5581	1
$c \to a$	0	0.9125	1	0.42	0.78296	1
$c \to b$	0.7	0.964	1	0.5	0.766	1
$c \to c$	0	0.319	1	0.58	0.80344	1
$c \to d$	0.3	0.956	1	0	0.5869	1
$d \to a$	1	1.0	1	0	0.6438	1
$d \to b$	0	0.95	1	0	0.5581	1
$d \to c$	0	0.55	1	0	0.5869	1
$d \to d$	0	0.87	1	1	1	1

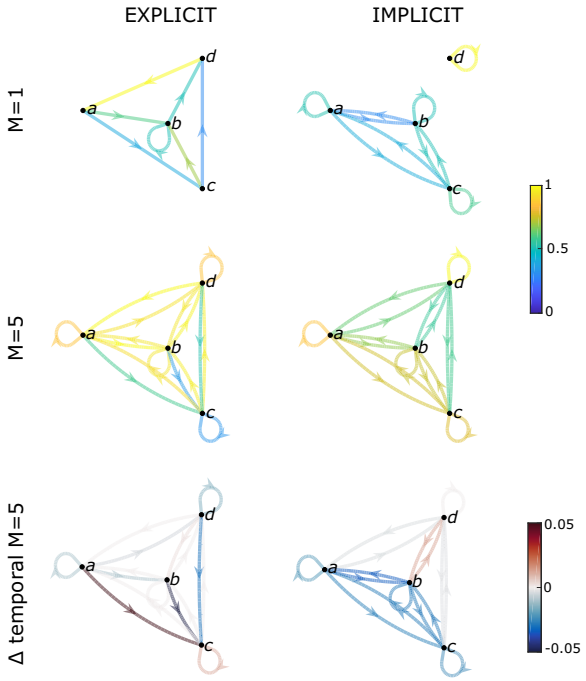


FIG. 6. Schematics highlighting the differences between explicit and implicit connectivity metrics applied to both aggregated and temporal descriptions of the network A of Fig. 5. The four upper panels represent explicit and implicit connections for $M = 1$ and 5 with arrows colored according to their probabilities (see also Table I). The two lower panels show the probability differences (i.e., the aggregated case, Table I, minus the temporal case, Table II) of both explicit and implicit metrics of connectivity for $M = 5$.

and 100 the probabilities coincide with the static case. Indeed, on the one hand, in the first time interval the static network A coincide with its temporal version while, on the other hand, for $M = 100$ probabilities saturate to 1 driven by links geometry rather than weight's values. As shown in Fig. 6, for $M = 5$ we can find instead significant differences between the static and aggregated case that are a clear signature of the temporal dynamics.

All in all, the above results suggest that $\mathbf{\Gamma}^f$ and $\mathbf{\Lambda}^f$ can provide different and complementary information about the connectivity processes occurring across a network. Moreover, as already pointed out by several studies [2], connectivity patterns can change significantly between a full temporal network description and its aggregated counterpart and this is well reflected in our simple examples.

VI. APPLICATION TO OCEAN TRANSPORT

We now apply our theoretical framework on a real-case network representing the dynamics of fluid elements by geophysical transport processes (e.g., oceanic or atmospheric circulation). Network approaches have demonstrated great effectiveness in assessing transport and mixing of fluid parcels in both theoretical and geophysical settings [12,24,36–40]. Studying the connectivity of such networks consists in estimating the probability of exchanging fluid parcels among different geographical locations. Since water (air, respectively) parcels carry numerous particulate and dissolved substances, connectivity is tantamount to evaluating how any almost-passive tracer is transported and dispersed by the oceanic (atmospheric, respectively) circulation. As such, relevant applications of transport networks already include studying the spread of oceanic tracers [41,42], microplastics [43], biological propagules [44–47] and of atmospheric pollutants [48].

Focusing on the ocean, network-based studies recently reported the presence of both preferential corridors and semipermeable barriers of transport within realistic oceanic flows [12,18], as documented also by alternative methods developed from dynamical systems theory [49,50]. These dynamical features, which were associated with relatively persistent fronts [51] (jetlike currents [52], respectively) tend to prevent (facilitate, respectively) the chaotic advection of water parcels across (within, respectively) them. Their existence determines the magnitude of connectivity among distinct oceanic subregions [53] and results in the emergence of broad-scale transport patterns [31,42,44]. However, this view of ocean connectivity has been mainly described by only considering explicit connections associated with a precise transport duration. Hence, the cumulated and implicit approaches introduced in the previous sections can bring new insights into how different places of the ocean can be connected by water parcels dispersal. In the following, we apply our previous analytical results to provide a broader and more general perspective of the connectivity of a realistic transport network in the Mediterranean Sea. In particular, we illustrate how our new metrics allow extracting novel and relevant information (that is well explained by current oceanographic knowledge) from a state-of-the-art oceanic flow field but we by no means intend to assess the reliability of the hydrodynamical model that generated it.

Adopting the Lagrangian flow network approach [12], we define a set of $N = 967$ oceanic nodes representing small, equal-sized subregions of the Mediterranean Sea surface. Links and weights between such set of nodes quantify water parcels exchanges driven by ocean currents over a time-interval of 30 days forward-in-time. To construct the network, we use a reference horizontal flow fields produced by an operational data-assimilating ocean model whose outputs have been validated [54]. More specifically, we exploit realistic daily currents at 10 m depth over a 30-day period spanning 01/06/2012–01/07/2012 [top-left insert in Fig. 7(a)]. The examination of M-steps on this network corresponds to the concatenation of M-times transport events of 30 days under the approximation of negligible diffusion and vertical displacements [18,24]. Explicit forward-in-time connectivity, in this case, is associated with the probability for a fluid parcel of traveling from one node to another and thus, to the probability that tagging a volume of water in one region of the ocean it will arrive to another given destination (after 30 days). Implicit connectivity represents instead the probability for two water parcels, belonging each of them to different nodes, of ending up in a third specific node. Again, this can be seen as the probability that two tagged volumes of water will meet together at a different common place in the ocean afterwards (after 30 days).

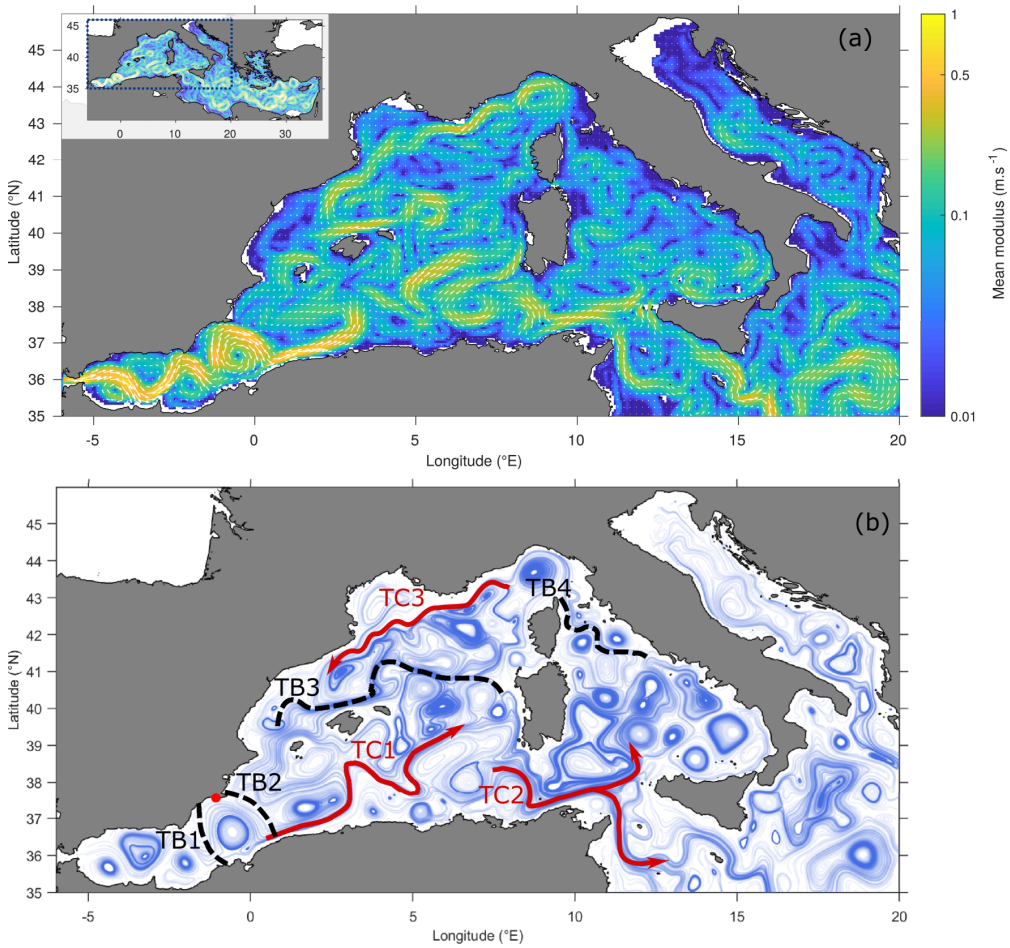


FIG. 7. Maps of the study area covering the western Mediterranean Sea. (a) Horizontal currents direction (thin white arrows) and mean modulus (background colors, in m s^{-1}) of the 10 m flow field averaged over one month (01/06/2012–01/07/2012). The top-left insert displays the whole model domain covering the entire Mediterranean basin. (b) Streamlines of the 10 m flow field averaged over one month (01/06/2012–01/07/2012). The red dot indicates the studied coastal site located south of Cartagena. Other annotations highlight the main transport features, adapted from [57,58]. Transport Barriers (TB) are depicted in black dotted lines with the Almeria-Oran front (TB1), the Cartagena-Tenes front (TB2), the Balearic front (TB3), and the meandering barrier associated with the northern Tyrrhenian gyres (TB4). Mean positions of the preferential Transport Corridors (TC), associated with the main geostrophic jetlike currents, are represented as plain red lines with the Algerian current (TC1), the Atlantic-Ionian jet (TC2), and the Northern current (TC3).

We first investigate how the conventional appraisal of ocean connectivity (i.e., single step explicit) changes when computing our connectivity metrics at a few different time steps. To do so, we arbitrarily select a coastal site located to the south of Cartagena [see the red dot in Fig. 7(b)] in the Alboran Sea and we analyze the evolution of a dispersal plume starting from this reference site using both Γ^f and Λ^f for $M = 1, 2,$ and 5 (Fig. 8). Assigning the index i to the targeted location and by considering all the nonvanishing indexes j, we can map all the nodes, along with their associated probabilities, which are explicitly or implicitly connected with the reference coastal site. Next, we briefly review the main transport barriers and conduits documented by previous research in the study-area and we highlight how explicit connectivity conforms with previous findings while implicit connectivity brings new insights to ocean connectivity.

In the western Mediterranean Sea, previous research highlighted, on the one hand, the presence of several Transport Barriers (TB, black annotations in Fig. 7) associated with major oceanographic fronts: the Oran-Almeria front [44,55], the Carthagena-Tenes front [56] and the North-Balearic front [42,44,57]. On the other hand, preferential Transport Corridors (TC, red annotations in Fig. 7) are associated with the main geostrophic jetlike currents such as the Algerian current, the Atlantic-Ionian jet and the Northern current [18,58,59].

When $M = 1$ in the explicit case (Fig. 8) (e.g., equivalent to single-step explicit estimates), the dispersal plume is spatially inhomogeneous with two cores of medium to high probabilities ($\sim 10^{-1}$ to 10^{-2}) which are well-explained by

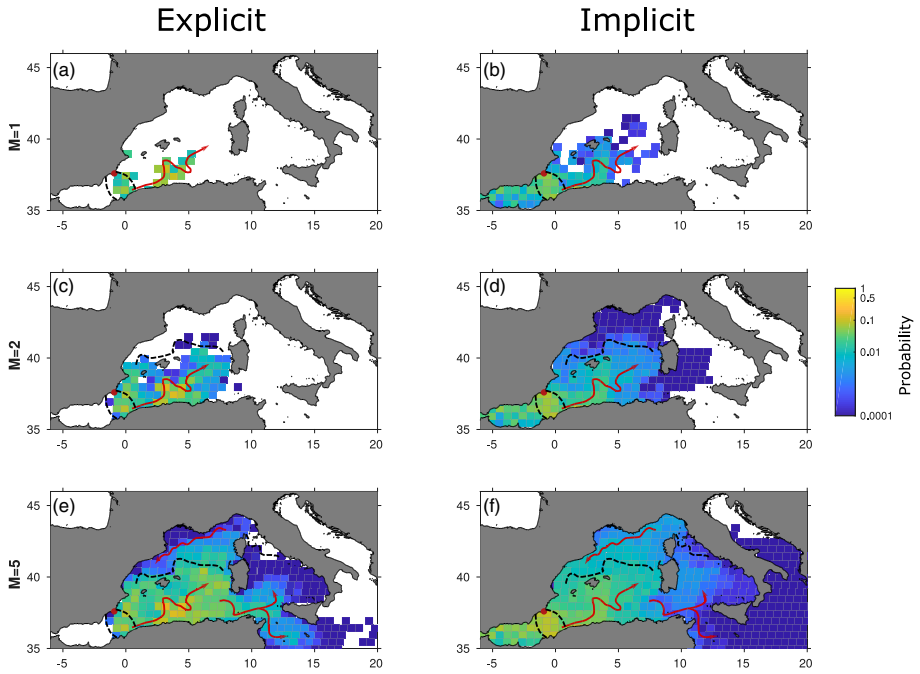


FIG. 8. Forward-in-time dispersal plumes for a tracking-time of 30 days starting from Cartagena, as derived from the explicit (left panels a, c, and e) and implicit (right panels b, d, and f) connectivity metrics using different number of steps (from upper to lower panels: $M = 1$, 2, and 5). Each node is colored according the probability of connection starting from our reference site (red dot) located south of Cartagena. White nodes indicate no connectivity (null probability).

the pre-identified transport features [Figs. 7(a) and 7(b)]. The Almeria-Oran and Cartagena-Tenes fronts, likely associated with an intense quasi-stationary eddy [Fig. 7(b)] trap water parcels south of Cartagena while the nearshore pathway of the Algerian current [Fig. 7(b)] allows some parcels to flow across TB2 and thus to disperse eastward into the Algerian basin (up to 5°E only, Fig. 8). The single-step implicit connectivity plume is much larger, extending from the strait of Gibraltar to about 10°E, and associated with more homogeneous probabilities than in the explicit case. While both cores of high probabilities (ranging from $\sim 10^{-2}$ to 10^{-1}) are similar in both cases, the implicit plume exhibits moderate to low probabilities ($\sim 10^{-3}$) in the western Alboran Sea and in the north-eastern Algerian basin. These implicit patterns are counter-intuitive and more difficult to interpret as they involve indirect connections ensured by "third-party" nodes. The reference site appears to be connected to the western Alboran Sea despite the presence Oran-Almeria front (TB1) and the continuous entrance of Atlantic waters (surface transport is mostly eastward) because they send waters to common downstream locations. Similarly, the low probabilities found in the north-eastern Algerian basin are probably due to recirculation processes and indirect connections ensured by coastal (counter-) currents and the meandering Algerian current. It suggests that, while the Cartagena-Tenes transport barrier constrains strongly the explicit plume [55], it becomes permeable in the case of the implicit plume.

For $M = 2$, the explicit connectivity dispersal plume extends north-eastward, reaching 10° E [Fig. 8(c)]. The cores of high probabilities ($\sim 10^{-1}$) match those revealed by the single-step ($M = 1$) plumes, corroborating the cumulative property of our methodology. Less probable connections ($\sim 10^{-2}$ to 10^{-4}) are found in most of the Algerian basin after approximating 60 days of advection, whereas they were absent for $M = 1$. Acting as a transport barrier, the Balearic front [TB3, Fig. 7(b)] might explain why the plume does not extend further north. As such, explicit two-step connectivity suggests that our reference site remains disconnected from the French and Italian coastlines and from the Alboran Sea. Conversely, the implicit connectivity plume spreads substantially across the western Mediterranean, connecting our reference site to most shorelines until ~10°E and ~45°N [Fig. 8(d)]. Probabilities are larger or equal than $\sim 10^{-2}$ south of TB3 while they drop down to 10^{-4} north of the barrier [Figs. 8(c) and 8(d)]. It indicates that weak implicit connections across the Balearic front occur at $M = 2$ despite the large distances and the supposed transport barrier effect. This can be explained by the fact that, in the vicinity of the front, water parcels coming from the reference site can encounter parcels coming from north of the barrier, realizing thus such implicit connections [Fig. 7(b)].

For $M = 5$ (i.e., surface transport over 150 days), the multistep explicit connectivity plume [Fig. 8(e)] spreads across most of the western Mediterranean basin and penetrates the Ionian Sea. These connectivity patterns are well-explained by the mean circulation highlighted in Figs. 7(a) and 7(b). The northern Tyrrhenian meanders (the Almeria-Oran and Cartagena-Tenes fronts, respectively) prevent effective connections with the northern Tyrrhenian Sea (the Alboran Sea,

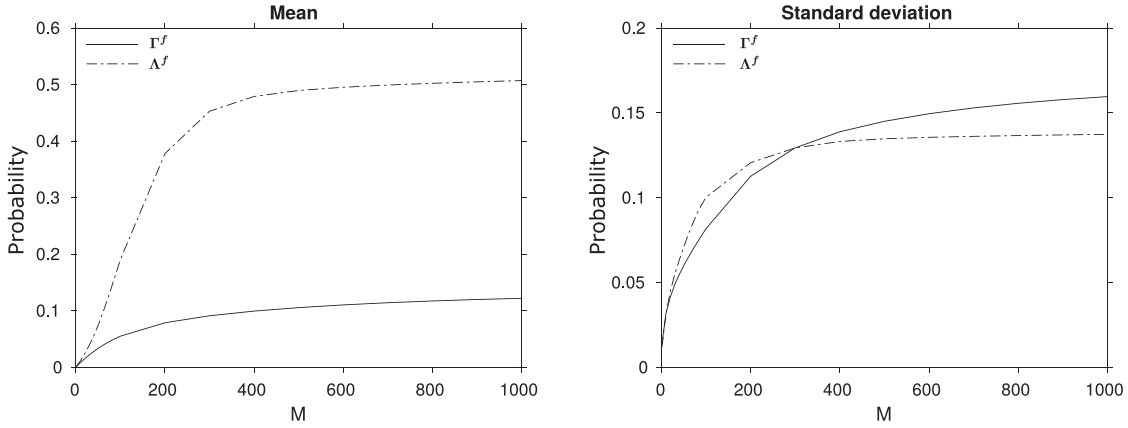


FIG. 9. Mean probabilities of connection (left panel) and the associated standard deviations (right panel) for all possible pairs of nodes in our ocean network for steps ranging $M = 1$ to $M = 1000$. Black plain (dotted, respectively) lines stand for the cumulated multistep explicit (implicit, respectively) connectivity.

respectively). The Northern current (the Atlantic-Ionian jet, respectively) ensure rare connections ($\sim 10^{-4}$; 10^{-3}) with the northern shorelines (with the eastern Ionian Sea, respectively). The multistep implicit connectivity plume [Fig. 8(f)] is larger: it covers the entire western Mediterranean basin and spreads over the Ionian Sea as well as the southern Adriatic Sea, despite the presence of the previously mentioned transport barriers. In comparison with the two-step implicit connectivity, the core of elevated probabilities extends further north, suggesting that the barrier effect of the Balearic front vanishes when longer transport durations are considered. For $M = 5$ and using both explicit and implicit methods, our reference site is weakly but consistently connected to most distant coastlines, except the northern Adriatic shores.

Finally, we analyze the global statistical distribution of our explicit and implicit proxies as a function of the number of steps. To do so, we compare the forward-in-time cumulated multistep explicit (Γ^f) and implicit (Λ^f) connectivity metrics for different M spanning 1–1000 by computing mean probabilities of connection, and their associated standard deviations, for all pairs of nodes (i.e., $N \times N = 935\,089$ pairs) of our flow network (Fig. 9). We find that the mean Γ^f probabilities grow sublinearly with the number of steps until reaching a plateau at around 0.1 after about $M = 800$ steps. Mean Λ^f probabilities grows almost-linearly with the number of steps until reaching a plateau at around 0.5 after approximatively $M = 500$ steps. For both metrics, saturation does not reach 1, as it was shown for one of the theoretical cases (see Sec. V), due to the presence of weakly connected components in our realistic ocean network. At saturation, explicit probabilities are spread-out over a wide range of values since the standard deviation tend to overcome the means. Implicit probabilities are more homogeneous and closer to the mean, even at saturation.

All in all, the newly introduced "implicit connectivity" proxy suggests thus that the connectivity of the surface ocean could have been substantially underestimated by previous methods, providing novel possible directions for the study of dispersion and transport patterns of any tracer across the ocean.

VII. CONCLUSIONS AND PERSPECTIVES

Our theoretical approach can be applied to study any kind of temporal, weighted and directed network in which a random walk can be defined. This should guarantee a broad applicability to various fields such as ecology, epidemics spreading, mobility, genetics and fluid-dynamics. By introducing the *cumulated* connectivity formalism, we provide exact analytical expressions for random walk connection probabilities between any pair of nodes and across arbitrary ranges of number of steps. This framework could constitute a first step for future modeling efforts to characterize network connectivity from a probabilistic perspective. We first focused on *explicit* connectivity patterns realized by paths and then for a novel *implicit* connectivity concept associated with network pitchforks. Such implicit view of connectivity highlighted network topological features overlooked until now. Future studies could indeed investigate how different network topologies, such as random, small-world or scale free, would be reflected in implicit connectivity patterns and how the latter would be related to different network dynamical regimes. Moreover, when random walk single-step transition probabilities can be mapped to fractions of a given quantity exchanged across the network, it is possible to link the probabilistic connectivity interpretation to transport dynamics. Indeed, we showed that explicit connection probabilities correspond to probabilities related to processes of tagging or sampling the transported quantity in a node forward- or backward-in-time, respectively. Analogously, implicit connection probabilities are also related to tagging or sampling processes but in two nodes simultaneously. These relationships can be further developed both theoretically and for practical applications, such as tagging and sampling experiments on spatial systems, discovering indirect interactions in complex ecological networks or further characterize diseases spreading and opinion dynamics in social systems. Possible extensions of our approach can also include nonconservative dynamics such as production, consumption and transformation of a transported quantity by modulating the probabilities at node scale. We finally illustrated an example of how our results can be applied to characterize fluid

transport driven by ocean currents in the Mediterranean Sea. We showed that our approach extends and generalizes the way physical connectivity in the ocean was understood until now and unveils hidden connections between different regions of the Mediterranean Sea. Consequently, this changes also our understanding of the role of some oceanographic features, such as transport barriers and transport corridors, in controlling fluid connections across the seascape. Applications of this methodology to geophysical flows could provide novel insights on the spreading of drifting organisms, pollutants and, more generally, any tracer that is transported by the flow.

The Python codes used to compute connectivity probabilities are available online [60].

ACKNOWLEDGMENTS

E.S.-G. is grateful to C. Payrató Borrás for discussions on the analytical derivations. T.L. is funded by a Doctoral fellowship obtained through Aix-Marseille University. V.R. and T.L. acknowledge financial support from the European project SEAMoBB (Solutions for sEmi-Automated Monitoring of Benthic Biodiversity), funded by ERA-Net Mar-TERA (id. 145) and managed by the ANR (Grant No. ANR-17-MART-0001-02, P.I.: A. Chenuil). I. H.-C acknowledges the Vicenç Mut contract funded by the Government of the Balearic Island and the European Social Fund (ESF) Operational Programme. This study has been conducted using E.U. Copernicus Marine Service Information.

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[2] P. Holme and J. Saramäki, Phys. Rep. **519**, 97 (2012).
[3] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).
[4] E. Estrada and N. Hatano, Phys. Rev. E **77**, 036111 (2008).
[5] H. H. K. Lentz, T. Selhorst, and I. M. Sokolov, Phys. Rev. Lett. **110**, 118701 (2013).
[6] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, UK, 2008).
[7] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[8] R. E. ULANOWlCZ and J. S. Norden, Int. J. Syst. Sci. **21**, 429 (1990).
[9] D. Garlaschelli, G. Caldarelli, and L. Pietronero, Nature **423**, 165 (2003).
[10] M. Rubinov and O. Sporns, Neuroimage **52**, 1059 (2010).
[11] D. Kempe, J. Kleinberg, and É. Tardos, in *International Colloquium on Automata, Languages, and Programming* (Springer, Berlin, 2005), pp. 1127–1138.
[12] E. Ser-Giacomi, V. Rossi, C. López, and E. Hernández-García, Chaos **25**, 036404 (2015).
[13] M. E. Newman, Nat. Phys. **8**, 25 (2012).
[14] L. Lovász *et al.*, Combinatorics **2**, 1 (1993).
[15] M. E. Newman, Soc. Netw. **27**, 39 (2005).
[16] N. Masuda, M. A. Porter, and R. Lambiotte, Phys. Rep. **716**, 1 (2017).
[17] D. Brockmann and D. Helbing, Science **342**, 1337 (2013).
[18] E. Ser-Giacomi, R. Vasile, E. Hernández-García, and C. López, Phys. Rev. E **92**, 012818 (2015).
[19] A. Gautreau, A. Barrat, and M. Barthélemy, J. Stat. Mech.: Theory Exp. (2007) L09001.
[20] S. Huntsman, in *International Conference on Complex Networks and their Applications* (Springer, Berlin, 2018), pp. 433–444.
[21] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci. USA **101**, 3747 (2004).
[22] M. A. Serrano and M. Boguná, Phys. Rev. E **68**, 015101(R) (2003).
[23] R. Pastor-Satorras and A. Vespignani, *Evolution and structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, UK, 2007).
[24] E. Ser-Giacomi, R. Vasile, I. Recuerda, E. Hernández-García, and C. López, Chaos **25**, 087413 (2015).
[25] N. Perra, A. Baronchelli, D. Mocanu, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **109**, 238701 (2012).
[26] M. Starnini, A. Baronchelli, A. Barrat, and R. Pastor-Satorras, Phys. Rev. E **85**, 056115 (2012).
[27] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein, Rep. Prog. Phys. **80**, 036603 (2017).
[28] C. R. Woese, Proc. Natl. Acad. Sci. USA **97**, 8392 (2000).
[29] C. Gracia-Lázaro, L. Hernández, J. Borge-Holthoefer, and Y. Moreno, Sci. Rep. **8**, 1 (2018).
[30] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, Physica A **391**, 1777 (2012).
[31] P. Miron, F. J. Beron-Vera, M. J. Olascoaga, J. Sheinbaum, P. Pérez-Brunius, and G. Froyland, Sci. Rep. **7**, 7021 (2017).
[32] R. V. Donner, M. Lindner, L. Tupikina, and N. Molkenthin, in *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems* (Springer, Berlin, 2019), pp. 197–226.
[33] R. Banisch, P. Koltai, and K. Padberg-Gehle, Chaos **29**, 063125 (2019).
[34] H. Kim and R. Anderson, Phys. Rev. E **85**, 026107 (2012).
[35] P. Koltai and D. M. Renger, J. Nonlin. Sci. **28**, 1915 (2018).
[36] E. Ser-Giacomi, V. Rodríguez-Méndez, C. López, and E. Hernández-García, Eur. Phys. J.: Spec. Top. **226**, 2057 (2017).
[37] V. Rodríguez-Méndez, E. Ser-Giacomi, and E. Hernández-García, Chaos **27**, 035803 (2017).
[38] M. Lindner and R. V. Donner, Chaos **27**, 035806 (2017).
[39] K. Padberg-Gehle and C. Schneide, Nonlin. Process. Geophys. **24**, 661 (2017).
[40] D. Wichmann, C. Kehl, H. A. Dijkstra, and E. van Sebille, Nonlin. Process. Geophys. **27**, 501 (2020).
[41] A. Baudena, E. Ser-Giacomi, C. López, E. Hernandez-Garcia, and F. d'Ovidio, J. Marine Syst. **192**, 1 (2019).
[42] E. Ser-Giacomi, G. Jordá-Sánchez, J. Soto-Navarro, S. Thomsen, J. Mignot, F. Sevault, and V. Rossi, Geophys. Res. Lett. **47**, e2020GL089941 (2020).
[43] D. Wichmann, P. Delandmeter, and E. van Sebille, J. Geophys. Res. Oceans **124**, 6086 (2019).
[44] V. Rossi, E. Ser-Giacomi, C. López, and E. Hernández-García, Geophys. Res. Lett. **41**, 2883 (2014).

[45] C. J. Thomas, J. Lambrechts, E. Wolanski, V. A. Traag, V. D. Blondel, E. Deleersnijder, and E. Hanert, Ecol. Modell. **272**, 160 (2014).
[46] M. Dubois, V. Rossi, E. Ser-Giacomi, S. Arnaud-Haond, C. López, and E. Hernández-García, Global Ecol. Biogeogr. **25**, 503 (2016).
[47] N. Ramesh, J. A. Rising, and K. L. Oremus, Science **364**, 1192 (2019).
[48] S. Fellini, P. Salizzoni, L. Soulhac, and L. Ridolfi, Atmos. Environ. **198**, 291 (2019).
[49] S. Wiggins, Annu. Rev. Fluid Mech. **37**, 295 (2005).
[50] G. Haller and G. Yuan, Physica D **147**, 352 (2000).
[51] M. J. Olascoaga, I. Rypina, M. G. Brown, F. J. Beron-Vera, H. Koçak, L. E. Brand, G. Halliwell, and L. K. Shay, Geophys. Res. Lett. **33**, L22603 (2006).
[52] M. V. Budyansky, M. Y. Uleysky, and S. V. Prants, Phys. Rev. E **79**, 056215 (2009).
[53] P. Celentano, P. Falco, and E. Zambianchi, Deep Sea Res. **166**, 103431 (2020).
[54] S. Simoncelli, C. Fratianni, N. Pinardi, A. Grandi, M. Drudi, P. Oddo, and S. Dobricic, *Mediterranean Sea Physical Reanalysis (CMEMS MED-Physics)* (Version 1) [Data set], Copernicus Monitoring Environment Marine Service (CMEMS) (2019), https://doi.org/10.25423/MEDSEA_REANALYSIS_PHYS_006_004.
[55] J. Tintore, P. La Violette, I. Blade, and A. Cruzado, J. Phys. Oceanogr. **18**, 1384 (1988).
[56] I. Hernández-Carrasco and A. Orfila, J. Geophys. Res.: Oceans **123**, 4398 (2018).
[57] A. M. Mancho, E. Hernandez-Garcia, D. Small, S. Wiggins, and V. Fernandez, J. Phys. Oceanogr. **38**, 1222 (2006).
[58] C. Millot and I. Taupier-Letage, in *The Mediterranean Sea* (Springer, Berlin, 2005), pp. 29–66.
[59] P.-M. Poulain, M. Menna, and E. Mauri, J. Phys. Oceanogr. **42**, 973 (2012).
[60] https://github.com/serjaaa/cumulated-net-conn.

Correction: The original Ref. [54] has been deleted. The source information for the original Ref. [55], now set as [54], has been modified. Subsequent references have been renumbered. The omission of a support statement has been fixed.

Second Correction: A wording error in the abstract and several wording errors in text have been fixed.

4. La connectivité coalescente sur de multiples générations prédit le flux de gènes pour divers phylums marins

4.1. Résumé

Dans le chapitre précédent, nous avons défini la connectivité génétique à partir d'évènements successifs et cumulatifs de dispersion larvaire, représentant les différentes possibilités de flux de gènes sur plusieurs générations entre populations. Les connexions que nous appelons explicites appréhendent le lien filial : la transmission sur une ou plusieurs générations de matériel génétique entre individus. Pour une population cible, nous allons pouvoir quantifier la probabilité qu'une autre population soit parente à un nombre de générations fixé. Les connexions que nous appelons implicites considèrent quant à elles la similitude génétique entre individus provenant de mêmes populations ancêtres : elles permettent de quantifier la probabilité sous-jacente que deux populations partagent les mêmes populations sources. Le parallèle est net avec la théorie coalescente qui, en génétique des populations est un modèle rétrospectif qui estime le temps écoulé depuis un possible ancêtre commun (Sigwart, 2009). Ainsi, toutes les populations considérées dans la connectivité coalescente sont contemporaines (i.e. on estime la probabilité qu'elles partagent une population ancêtre commune). Ce n'est pas le cas pour la connectivité filiale, où les paires de populations sont temporellement séparées par le nombre de génération considéré.

Ici, nos modèles mécanistiques de connectivité filiale et coalescente (i.e. utilisant respectivement les connexions explicite et implicite) vont estimer le flux de gènes potentiel entre paires de populations afin de les comparer aux patrons de structures génétiques observés. Ces structures sont elles-mêmes évaluées à partir de mesures

de différentiation génétique (c'est-à-dire de variation de fréquence allélique) entre paires de populations. Chaque paire de population est ainsi définie par un indice de différentiation génétique, i.e. *Fst* (c.f. section 1.3.3.2). Même si le flux de gène est le pendant évolutif de la dispersion larvaire, de nombreuses incertitudes empêchent de relier précisément la probabilité d'échange d'individus entre populations et les *Fst* mesurés (Whitlock & Mccauley, 1999). Ces incertitudes concernent d'une part la modélisation de la connectivité génétique à partir d'événements successifs de dispersion larvaire (voir Chapitre III), et d'autre part les restrictions quantitatives et spatiales de l'échantillonnage de chaque étude individuelle de génétique des populations. Il faut également mentionner les phénomènes de sélection naturelle et les phénomènes aléatoires agissant sur les fréquences alléliques des gènes marqueurs utilisés pour calculer les *Fst* entre des populations.

Dans ce chapitre, nous avons modélisé la connectivité filiale et coalescente pour estimer le flux de gène chez 47 espèces à cycle de vie biphasique, compilées depuis 58 études de génétique des populations. Les espèces considérées dans cette méta-analyse appartiennent à divers représentants marins du vivant (e.g. anthozoaires, crustacés, mollusques, téléostéens, ou encore des phanérogames) avec des traits de premier stade de vie et des stratégies de dispersion différentes. Les probabilités de connexions explicites et implicites calculées sur de multiples générations successives de dispersion sont comparées avec les mesures de *Fst* entre chaque paire de population.

Le modèle de connectivité coalescente qui considère des événements de dispersion sur de multiples générations montre les meilleures prédictions de flux de gènes pour l'ensemble des taxons, et explique en moyenne ~ 50 % de la variabilité des différentiations génétiques observées à l'échelle de la méta-analyse. De plus, le modèle de connectivité filiale prédit le flux de gène avec plus de précision que les modèles classiques d'isolement par la distance (i.e. *IBD*). L'échange de gènes sur une seule génération, est quant à lui le modèle le moins adéquat. Nos résultats mettent également en évidence le caractère déterminant de l'échantillonnage dans la compréhension des patrons de diversité génétique : il faut un échantillonnage conséquent et réparti de manière homogène à l'échelle du bassin étudié pour pouvoir prédire efficacement le flux de gène. D'autre part, pour une capacité de dispersion constante (i.e. une PLD fixée), plus l'échantillonnage est étendu spatialement, plus le modèle qui rend compte du flux de gène de façon optimale correspond à un nombre de générations élevé. Parallèlement, pour un échantillonnage fixé, plus la capacité de dispersion est

grande, plus ce nombre de générations est petit. Cette relation entre le nombre de générations adéquates pour comprendre les patrons de flux de gène et l'étendue de l'échantillonnage modulé par la capacité de dispersion suggère une relation étroite entre les échelles temporelles et spatiales de la connectivité génétique. De manière générale, le nombre moyen de générations adéquates à l'échelle de la méta-analyse (20 générations) est assez nettement inférieur à la valeur de saturation (celle pour laquelle la probabilité de connexion atteint un plateau) identifiée dans le chapitre précédant (500 générations, pour une PLD de 30 jours). Cela suggère que le flux de gène et la dérive génétique ne sont pas à l'équilibre, potentiellement à cause de la variabilité des forçages environnementaux, elle-même modifiée par les activités anthropiques à l'origine du changement climatique.

Modéliser la connectivité génétique coalescente a permis d'estimer avec une meilleure précision le flux de gène chez des organismes marins multicellulaires couvrant un large éventail phylogénétique. Cette méthode pourrait également être appliquée pour les espèces terrestres dispersant par le vent (Lett et al., 2020). Nos résultats ouvrent de nouvelles perspectives quant à la compréhension de l'impact relatif des différentes forces évolutives sur la structure génétique et à l'évaluation des échelles spatiales caractéristiques sous-jacentes.

4.2. Article

Coalescent connectivity through multi-generation dispersal modelling predicts gene flow across marine phyla

Térence Legrand[1], Anne Chenuil[2], Enrico Ser-Giacomi[3], Sophie Arnaud-Haond[4], Nicolas Bierne[5] and Vincent Rossi [1]

[1]Aix Marseille University, Universite de Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (UMR 7294), Marseille, France.
[2]Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, CNRS (UMR 7263), Station Marine d'Endoume, Marseille, France
[3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 54514 MIT, Cambridge, Massachusetts, USA
[4]Marine Biodiversity, Exploitation and Conservation, UMR 9190 IRD – IFREMER – UM – CNRS, Sète, France.
[5]Institut des Sciences de l'Evolution Montpellier, UMR 5554 UM – CNRS – IRD – EPHE, Station Marine OREME, Sète, France.

En révision dans *Nature Communication*, soumis en Septembre 2021

4.2.1. Abstract

Gene flow governs the contemporary spatial structure and dynamic of populations as well as their long-term evolution. For species whom migration is realized through propagule dispersal driven by atmospheric or oceanic flows, biophysical models allow predicting the migratory component of gene flow, which facilitates the interpretation of broad-scale spatial structure inferred from observed allele frequencies among populations. However, frequent mismatches between dispersal estimates and genetic diversity prevent an operational synthesis for eco-evolutionary projections. Here we use an extensive compilation of 58 population genetic studies of 47 phylogenetically divergent marine sessile species over the Mediterranean basin to compare genetic differentiation observations against gene-flow predictions obtained with Isolation-By-Distance, single-generation dispersal and novel multi-generation dispersal models. Unlike previous approaches, they unveil explicit parents-to-offspring links (filial connectivity) and, for the first time, implicit links among siblings from a common ancestor (coalescent connectivity). We find that genetic sampling strategy is determinant to predict accurately genetic structure and that 50 % of observed genetic differentiation variance is explained by coalescent connectivity over multiple generations, significantly outperforming other models. Our results offer great promises to untangle the eco-evolutionary forces that shape sessile population structure and to anticipate climate-driven redistributions, altogether improving spatial conservation planning.

Keywords: Connectivity, Population genetics, Meta-analysis, Seascape genetics, Stepping-stone dispersal, Network structure, Eco-Evolutionary dynamics, Isolation-by-distance, Conservation biology, Lagrangian biophysical models.

Statement of authorship: T.L and V.R. planned and designed research. T.L., A.C. and S.A-H. contributed substantially to the compilation of genetic population studies. E.S-G. developed the theoretical formulations of the multi-generation dispersal models. T.L. performed data analysis with important contributions from V.R., E.S-G. and A.C. The manuscript was written by T.L. and V.R., and all co-authors provided important contributions and critical revisions. All authors approved the final version of this manuscript.

Data accessibility statement: Population genetic data used in this study were compiled from the published literature; the full dataset is provided as a supplementary file and will be uploaded on a public repository upon acceptation, along with correspond-

ing bibliographic references in the supplementary material. The Python codes used to compute multi-generation explicit and implicit dispersal probabilities are available online here

4.2.2. Introduction

Gene flow counterbalances natural selection and genetic drift, reshuffles mutations among spatial locations and contributes to shaping the contemporary spatial patterns of biodiversity (Hellberg, 2009; Lenormand, 2002; Lowe et al., 2017; Slatkin, 1987; Slatkin, 1985). By introducing foreign alleles to local population, gene flow spreads adaptative changes and tends to alleviate the effect of inbreeding depression (Hellberg, 2009; Lowe et al., 2017). Simultaneously, gene flow homogenises allele frequency among populations, which counterbalances local adaptation, hence reducing the mean fitness of populations (Lenormand, 2002). Fundamentally, gene flow is ensured when dispersal is followed by reproduction and subsequent offspring survival (Duputié and Massol, 2013). A common confusion prevails between demographic connectivity (i.e. the number of migrants exchanged among populations), which is usually assessed by direct detection of individuals (field observations and parentage analyses of genetic data), and genetic connectivity (i.e. the efficient transfer of genetic material between distant populations), which is indirectly estimated thanks to population genetics (Hellberg, 2009; Lowe and Allendorf, 2010; Lowe et al., 2017; Selkoe et al., 2016; Slatkin, 1987; Weersing and Toonen, 2009; Whitlock and Mccauley, 1999). In this way, demographic and genetic connectivity seem to interact on specific -yet poorly appreciated- temporal and spatial scales (Benestan et al., 2021; Bode et al., 2019; Gagnaire, 2020; Lowe et al., 2017; Pinsky et al., 2017). This may explain the numerous mismatches between demographic connectivity and gene flow estimates (Lowe and Allendorf, 2010; Selkoe et al., 2016; Weersing and Toonen, 2009), all of which question the relative roles of eco-evolutionary forces in shaping population spatial structure.

This paradox could stem from the fact that dispersal is a complex and multi-aspect process involving interlocked ecological and evolutionary mechanisms (Clobert et al., 2012; Cowen and Sponaugle, 2009; Duputié and Massol, 2013). Dispersal results from movements of individuals themselves or those induced by third-parties categorized as biotic, (e.g. thanks to other moving organisms, Tomback et al., 2001; Viana et al., 2016), or abiotic, that is driven by winds and ocean currents (Cowen and Sponaugle, 2009; Nathan et al., 2002; Selkoe et al., 2016). This study focuses on sessile population which adults have no or little displacement abilities, so that their connectivity is mostly ensured by the abiotic dispersal of propagules, like numerous marine and terrestrial taxa. This alleviates the difficulties in appraising the movements of wild populations and biotic third parties. In marine sessile populations, early-life non-

motile stages (e.g. seeds, eggs, larvae) are regularly released in the water column and are then passively transported across the seascape by anisotropic currents over various spatial scales (Cowen et al., 2006; Shanks, 2009) ensuring the replenishment of both local and distant populations (Hidalgo et al., 2019; Legrand et al., 2019). As such, the proper evaluation of current-driven dispersal should help us disentangling the evolutionary forces (gene flow, natural selection, genetic drift or long term mutation) shaping marine biodiversity and its climate-driven redistributions (Blowes et al., 2019). While a small proportion of migrants could be sufficient to ensure gene flow between distant populations (Lowe and Allendorf, 2010), the inherent spatial scales of genetic structures are generally a few orders of magnitude higher than potential dispersal distances over a single generation (Marshall et al., 2010), even for species exhibiting extremely rare long-distance dispersal (Crandall et al., 2012; T. M. Smith et al., 2018). Likewise, a single generation dispersal event does not allow to evaluate the evolutionary timescale over which gene flow shapes genetic diversity (Duputié and Massol, 2013). Theory predicts instead that consecutive dispersal events of numerous propagules, acting in synergy with other evolutionary forces, shape together the genetic diversity observed at broad-scale (Crandall et al., 2012; Saura et al., 2014). Consequently, modelling genetic diversity from the unique perspective of water-borne dispersal should enlighten the typical scales and relative importance of evolutionary forces that shape the spatial structure of marine sessile populations.

Modelling water-borne dispersal is a multidisciplinary challenge sharing tight commonalities with air-borne dispersal (Lett et al., 2020). First, it requires to jointly account for the spatio-temporal variability of currents, the species-specific early-life traits, the habitat patchiness (D'Aloia et al., 2015; Weersing and Toonen, 2009) as well as to consider all possible connectivity pathways (Kool et al., 2013). Second, it must simulate consecutive dispersal events by considering multiple generations of migrants where each intermediary connected population acts as a steppingstone. Bio-physical models, which simulate explicitly the dispersal of propagules by oceanic chaotic flows, have been widely used in the last few decades to derive physical connectivity metrics such as dispersal kernels (Cowen et al., 2006; Legrand et al., 2019; Mari et al., 2020). Simulations of single-generation dispersal commonly provide quantitative estimates of how distant populations are connected with each other. However they rarely match observed gene flows (Selkoe et al., 2016), possibly due to intrinsic flaws such as disregarding the multi-generational character of successive dispersal events (Boulanger et al., 2020; Jahnke et al., 2018) while overlooking intermediate

stepping-stone connexions.

In the marine realm, current models considering multi-generational dispersal are seldom, concern only a few specific species or taxa, and still inadequately explain genetic differentiation measures. They rely on the computation of connectivity matrices, which are mathematical objects describing dispersal of propagule exchanged between discrete populations. Such matrices can be interpreted as adjacency matrices of directed and weighted network (or graph). Thus, an approach consists in considering network theory tools such as shortest paths analysis to estimate the strength of connections among two distinct populations over multiple dispersal events (Boulanger et al., 2020; Buonomo et al., 2017). As shortest and most-probable paths of such networks differ (Ser-Giacomi, Vasile, et al., 2015), these methods neglect all other possible pathways that may drastically change the resulting connectivity (McRae and Beier, 2007). Other approach uses Markov chains and matrix multiplications to estimate the probability of connection between population-pairs over a given number of generations (Foster et al., 2012; Jahnke et al., 2018; Kool et al., 2010; White et al., 2010). Studies using this method did not consider all the inherent dispersal pathways as they only assessed the connection probabilities occurring at a given number of generations (equivalent to the exact number of multiplication) while neglecting all intermediate connections associated to any number of generations lower or equal to the prescribed number of dispersal events (Ser-Giacomi et al., 2021). Moreover, and to our knowledge, all present modelling approaches simulate stepping-stone dispersal of single lineages; in other words, they estimate the connectivity resulting from all explicit parents-to-offspring connections, i.e. filial connectivity. However, it is conspicuous that two fully disconnected populations which are both replenished by migrants originating from the same source population should share common alleles, thus displaying similar allele frequencies. The consideration in dispersal models of this new conceptual view of connectivity (Ser-Giacomi et al., 2021) that highlights links among siblings through common ancestors, i.e. coalescent connectivity, has been overlooked to-date, although it could largely alter gene flow predictions and contribute to the aforementioned discrepancies between predicted dispersal and realized gene flow assessments (Hellberg, 2009).

This paper presents an exhaustive comparison between demographic and genetic basin-scale connectivity based on classical and novel dispersal metrics across a meta-analysis of several marine taxa. While our results mainly apply to marine sessile populations whose dispersal is mediated by ocean currents, our new models and con-

clusions have the potential to transform how water- as well as air-borne dispersal of sessile terrestrial populations are evaluated. Here, we introduce state-of-the-art multi-generation dispersal models that evaluate all connections among population-pairs for a fixed number of generation while simultaneously cumulating those ensured by previous generations (Ser-Giacomi et al., 2021). Our models provide not only a precise estimation of the explicit links (filial connectivity) but also allow computing, for the first time, the implicit links existing between any population-pair having common source populations through multi-generational dispersal (coalescent connectivity). After parameterizing our model with the main dispersal traits of various taxa encompassing seagrasses, algae and metazoans, we test modelled gene flow predictions against an extensive compilation of observed genetic structures (i.e. genetic differentiation estimates given by *Fst* between population-pairs) over the whole Mediterranean Sea. We find that coalescent connectivity through multi-generation dispersal explains 50 % of the observed variance of genetic structures, substantially improving gene flow predictions with respect to previous approaches. Furthermore, the optimal number of generations to best predict gene flow significantly correlates with the sampling coverage scaled by the species-specific dispersal abilities, enlightening the typical scales of eco-evolutionary processes. It suggests that our model could be used to infer population genetic structures, a key pre-requisite for management and protection.

4.2.3. Results

We test the predictions of our multi-generation explicit and implicit dispersal models (Ser-Giacomi et al., 2021), simulating filial and coalescent connectivity respectively, against an extensive compilation of 58 genetic structures observed in the Mediterranean Sea. The dataset contains 3821 *Fst* measures between population-pairs for phylogenetically divergent 47 marine species (Figure 4.1a) which are characterized by a biphasic life-cycle, i.e. early-life free-swimming dispersing propagules and full to semi-sedentary adult (Figure 4.2a). We model the full range of variability of current-driven dispersal over the whole Mediterranean basin for each species using a fine-tuned particle-tracking model (Dubois et al., 2016; Monroy et al., 2017; Rossi et al., 2014; Ser-Giacomi, Rossi, et al., 2015) fed by the horizontal multi-year velocity field from an operational data-assimilative ocean model (Oddo et al., 2009). Each species is characterized by three main dispersal traits: Pelagic Larval Duration (PLD, i.e. the time propagules spend drifting with ocean currents, Fig. 2a), spawning seasons and adult habitats (Figure 4.1b,c). These biological factors were identified as the major ones

governing how the variability of ocean currents reflect on our probabilistic connectivity metrics (Monroy et al., 2017). Based on Wright's island model that expresses *Fst* as a function of the effective population size and the migration rates (Wright, 1931), *modelled Fst* are computed using the connection probabilities obtained from five dispersal models as an equivalent of the migration rates. It allows directly comparing *modelled Fst* with *observed Fst* values through Mantel tests.

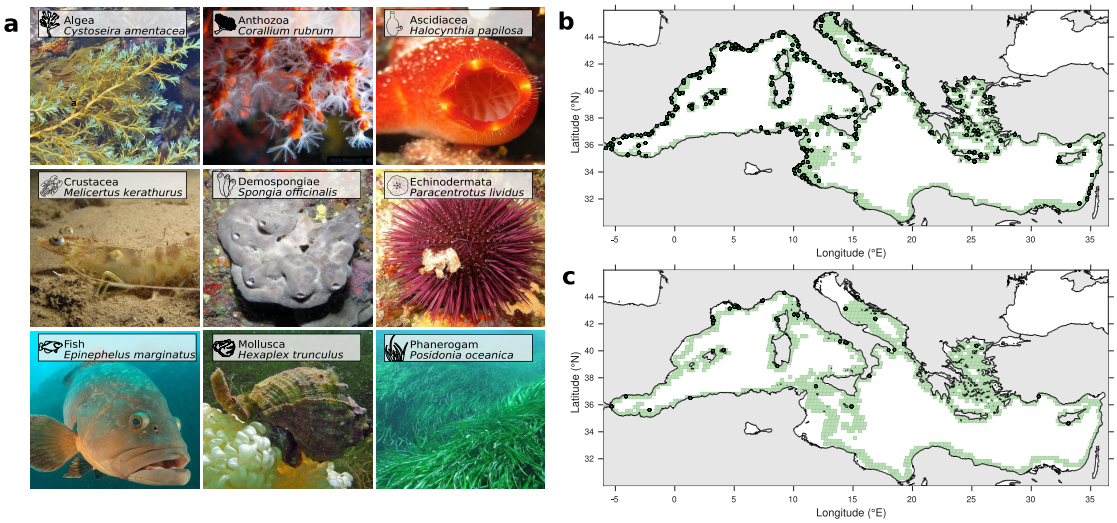


Figure 4.1. – **Meta-analysis summary, estimated habitats and geographical locations of the sampled populations**. **a** Exemplary pictures of marine species belonging to the nine taxonomic groups comprised in the compilation of 58 population genetics studies. **b** Basin-scale view of all sampling locations compiled in the meta-analysis (dark green dots) and putative populations (light green squares) that act as steppingstones in our multi-generation dispersal model for the shallow coastal habitat. **c**, same as **b** but for the neritic shelf habitat

Among the 58 compiled basin-scale population genetic studies (Figure 4.1), we search for Mantel correlations between *observed Fst* and *modelled Fst* obtained with our novel cumulated explicit and implicit dispersal models (Figure 4.2b,c) considering both single- and multi-generation estimates. We also test for conventional Isolation-By-Distance (IBD) models, using either Euclidian or sea-least cost distances. The number of significant Mantel tests is 10 for single-generation explicit, 19 for Euclidian IBD, 21 for multi-generation explicit, 22 for sea least-cost IBD and 23 for multi-generation implicit models respectively (Figure 4.3). Comparing the set of 58 p-values of Mantel tests for each study, Fisher's combined probability tests are significant for all methods (p-values***). Among the significant Mantel tests, the lowest

mean Mantel R^2 is found for the single-generation explicit dispersal model (0.19 ± 0.05, i.e. 95% confidence intervals), followed by both sea-least cost (0.31 ± 0.07) and Euclidian (0.35 ± 0.11) IBD models. The highest mean Mantel R^2 stem from both multi-generation models, with 0.45 ± 0.13 for explicit dispersal (Figure 4.3d) and 0.46 ± 0.13 for implicit dispersal (Figure 4.3e) models.

Next, we compare the five different models based on the correlations with *observed Fst* by computing pairwise relative differences of mean Mantel R^2 (i.e ΔR^2). The multi-generation implicit dispersal model displays mean mantel R^2 higher than any other method: mean ΔR^2 is 0.20 ± 0.12 with Euclidian IBD, 0.16 ± 0.10 with sea-least cost IBD, 0.32 ± 0.11 with single-generation explicit and 0.05 ± 0.05 with multi-generation explicit dispersal models (Figure 4.3e). Furthermore, the *modelled* and *observed Fst* correlations obtained with a given model are tested for being greater than those obtained with other models (Hendrickson et al., 1970; Silver et al., 2004), applying Fisher's combined probability tests on p-values. It shows that multi-generation implicit dispersal model is greatly (p-values**) to highly (p-values***) significantly better than any other method to predict gene flow at the meta-analysis scale (Figure 4.3e). Note that single-generation explicit dispersal model displays the lowest Mantel R^2 (Figure 4.3c). Hence, our novel multi-generation implicit dispersal model provides the best correlations with *observed Fst*.

When inspecting the study-specific accuracy of the best multi-generation implicit dispersal model, Mantel R^2 values range from 0.07 for the European hake *Merluccius merluccius* (Milano et al., 2014) to 0.97 for green crab *Carcinus aestuarii* (Schiavina et al., 2014, Figure 4.4a). For studies that include abundant genetic markers, it is possible to identify markers with particularly high *Fst* values (i.e. outlier; based on appropriate models), suggesting that natural selection filtered alleles differentially among populations. For a sea urchin *Paracentrotus lividus* (Carreras et al., 2020), *Fst*-outlier loci returns a Mantel R^2 of 0.36*, which is much higher than the R^2 of 0.23^{ns} obtained considering all the loci. Note that two studies focusing on the same species using both microsatellite markers, can display contradictory results: *Corallium rubrum* is characterized by a highly significant and tight correlation (Mantel R^2 = 0.506***, Aurelle et al., 2011) as well as a non-significant loose correlation (Mantel R^2 = 0.206^{ns},Costantini et al., 2013, Figure 4.4a), exemplifying inter-study variability. For the flathead grey mullet (Durand et al., 2013), which have been sampled homogenously across the Mediterranean basin (i.e. the "Spatial Sampling Representativeness" SSRep, which evaluates the spatial coverage of the sampling by computing the mean distance

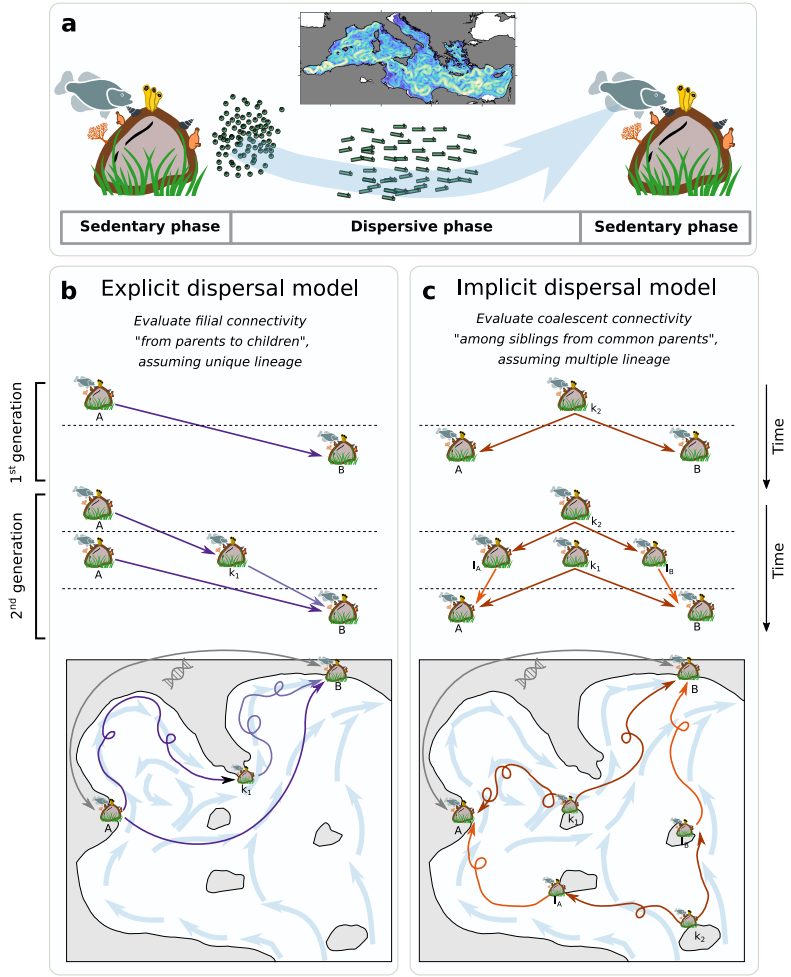


Figure 4.2. – **Modelling multi-generational water-borne dispersal**. **a** Schematic representation of a biphasic life cycle composed of sessile or sedentary adults and dispersive early-life stages, that is a distinctive feature of all species included in the meta-analysis. During the dispersive phase, numerous individuals are dispersed across the seascape by turbulent currents (represented in the Mediterranean miniature). **b** Explicit dispersal model evaluates filial connections between population-pairs. **c** Implicit dispersal model estimates coalescent connections between population-pairs. Schematics highlight the simulated genealogy over two generations and illustrate one of many multi-generational dispersal pathways that are considered by our models when estimating the connectivity between distant populations A and B.

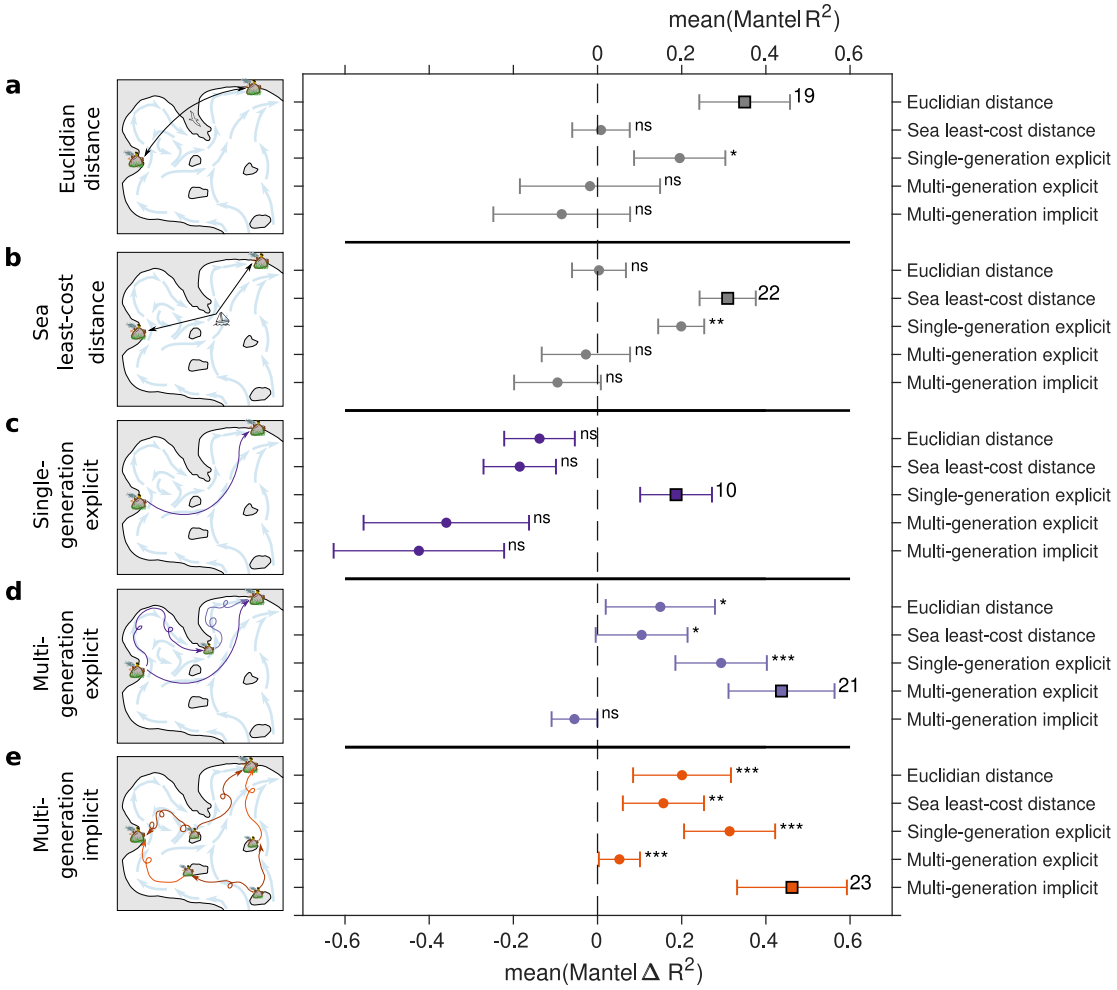


Figure 4.3. – **Cross-comparison of gene flow predictive models using pairwise unilateral tests**. Squares indicate mean Mantel R^2 and dots indicate mean ΔMantel R^2 computed as R^2 of each reference model (left) minus R^2 of all models (right) for **a** IBD (Euclidian) model, **b** IBD (sea least-cost) model, **c** single-generation explicit dispersal connectivity, **d** multi-generation explicit dispersal model and **e** multi-generation implicit dispersal model. Only the significant predictions (p-value*) of each reference model (left) are considered in these comparative analyses. The number of significant predictions per model (over the total of 58 studies) is reported on the right of each square. Error-bars represent the 95 % confidence intervals. Asterisks indicate if the reference model (left) is significantly better than the others (right) to predict observed gene flows at the meta-analysis scale. Note that asterisks inform us about the significant of the statistical tests as follows: * ≤ 0.05; ** ≤ 0.005; *** ≤ 0.0005 and "*ns*" stands for not significant.

between sampling sites and their barycenter, is 1,200 km for this study, Table SI), the network representation of *modelled Fst* mimics well the one of observed Fst (Figure 4.4b). *Observed Fst* are low between population-pairs located in the western basin but relatively high between western and eastern Mediterranean populations, suggesting spatial genetic structuring that is well predicted by coalescent connectivity (see the scatter plot of Figure 4.4b, Mantel $R^2 = 0.42^{**}$). Similar results are obtained for instance for a seagrass species (Alberto et al., 2008, Figure 4.4c). Although in this case spatial sampling is spatially more restricted (SSRep = 830 km), the genetic structure (high observed *Fst*) between the Adriatic and Spanish populations is well reproduced by *modelled Fst* (Mantel $R^2 = 0.62^{**}$).

We then test the robustness of the multi-generation implicit dispersal model with respect to the species and studies attributes. None of these factors (taxa, PLD, spawning season, genetic marker, and SSRep) has a significant effect on Mantel correlation results (R^2, p-value, Table SI-2 in annexe B). Yet, we find a significant linear negative correlation between the logarithm of Mantel p-values and the number of sampled population ($R^2 = 0.15^{**}$) as well as with the range of *observed Fst* ($R^2 = 0.25^{***}$, Table SI2). Furthermore, the probability to obtain a significant Mantel correlation between *modelled Fst* and *observed Fst* (i.e. successful gene flow predictions) as a function of the number of sampled populations is well predicted by a logit model using a binomial distribution ($R^2 = 0.50^{***}$, Figure 4.5a). It suggests that ≥ 11 populations must be sampled across the Mediterranean basin to achieve more than 50 % of successful model predictions, while it reaches 90 % with more than 23 sampled populations (Figure 4.5a). Finally, we find a positive linear relationship between the base 10 logarithm of the optimal number of modelled generations and the base 10 logarithm of SSRep normalized by the PLD ($R^2 = 0.59^{***}$, Figure 4.5b), considering only the studies whose genetic observations are significantly correlated with our model predictions. This surprisingly tight relationship has several interesting implications. First, if the spatial structures of two species are evaluated through the same sampling design, our dispersal model needs more generations for short PLD than for long PLD species to reproduce well the genetic observations. Consequently, our model conforms to the intuitive view that a species needs more successive events of dispersion across generations to disperse widely across the seascape. Moreover, if two species have similar PLDs, our model requires a higher number of generations to simulate well the genetic structure of the one whose sampling is wider and more comprehensive. Last, when parametrized correctly with the dispersal ability (i.e. PLD) of the target

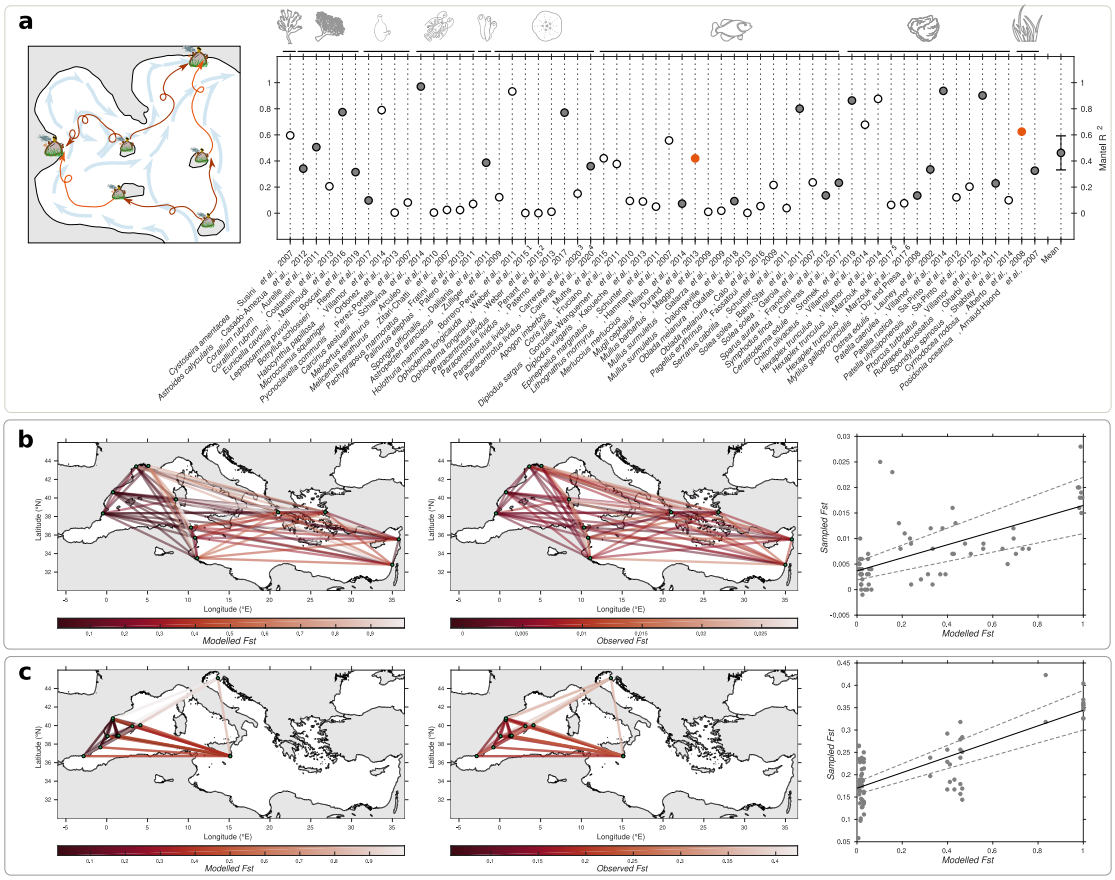


Figure 4.4. – **Accuracy of the multi-generation implicit dispersal model in explaining compiled genetic structures**. **a** Mantel R^2 between *modelled Fst* and *observed Fst*. Filled dots highlight the 23 significant correlations (p-value*). Note that some results reported by a given study are analysed separately: (i) Weber et al., 2015 used SNPs marker (1) and mtDNA marker (2); (ii) Carreras et al., 2020 considered all the loci together (3) and then only the Mediterranean outliers loci (4); (iii) Marzouk et al., 2017 analysed SNPs marker (5) and mtDNA marker (6). **b** network representation of *modelled Fst* (left) and *observed Fst* (right) and their corresponding scatterplot for the flathead grey mullet (*Mugil cephalus*; Durand et al., 2013; red dot in **a**). **c** same as **b** but for a seagrass (*Cymodocea nodosa*; Alberto et al., 2008; red dot in **a**).

species, the predictive ability of the implicit dispersal model scales with the SSRep. Altogether and assuming that the species-specific dispersal traits have been accurately parametrized, it suggests that the non-significant gene flow predictions (e.g. 35 studies among the meta-analysis) could be attributed to too scarcely (Figure 4.5a) and too spatially-restricted sampling (Figure 4.5b) rather than to the model abilities themselves.

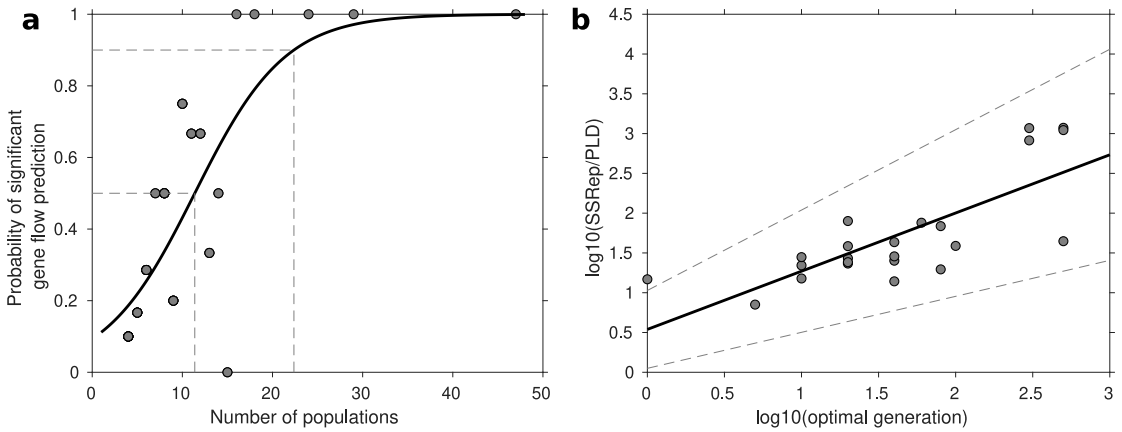


Figure 4.5. – **Sensitivity of the multi-generation implicit dispersal model**. **a** Probability of significant gene flow predictions (i.e. probability to obtain significant Mantel correlations between *modelled Fst* and *observed Fst*) as a function of the number of sampled populations. The thick black line represents the logit model (R^2 = 0.50 and p-value***). **b** Correlation between the 10 base logarithm of the optimal number of generation and the 10 base logarithm of the spatial sampling representativeness normalized by the species-specific PLDs (R^2 = 0.59 and p-value***). The thick black line represents the regression curve and the dotted grey ones the 95 % interval of confidence.

4.2.4. Discussion

For gene flow predictions at the meta-analysis scale, the cumulated multi-generation implicit dispersal model (Ser-Giacomi et al., 2021), which evaluates coalescent connectivity, significantly outperformed explicit dispersal models, which assess filial connectivity, as well as IBD models. Nearly half of the compiled studies, which spanned a wide phylogenetic range of mostly sessile taxa with contrasted dispersal traits, displayed significant Mantel correlations between modelled and observed genetic differentiation estimates. It is more than twice the proportion displayed by the explicit single-generation dispersal model, which emerges as the worst model in our

meta-analysis. Overall, the best models are the multi-generation implicit and explicit dispersal, suggesting unambiguously that modelling multiple generations to predict accurately genetic connectivity is crucial. To our knowledge, the only model that considers both coalescent and filial connectivity uses circuit theory to approximate how barriers and corridors of habitat affect genetic connectivity through a process called Isolation-By-Resistance (McRae, 2006). While this empirical model helped interpreting gene flows for self-dispersing organisms across well-known and relatively stable landscape (McRae and Beier, 2007), it has not yet been applied to the marine realm probably because the seascape is highly variable and in perpetual movements. Contrarily, our new dispersal models are mechanistic and plainly consider the dynamical properties of ocean currents that drives water-borne dispersal, so that it can be readily apply to air-borne dispersal (Lett et al., 2020).

We find that multi-generation dispersal models performed significantly better than IBD models. Similar results are found when using explicit multi-generation models for the seagrass *Zostera marina* in the North Sea (Jahnke et al., 2018) and the mollusc *Kelletia kelletii* along the Californian shores (White et al., 2010). Yet, Boulanger et al., 2020 found a tighter and more significant correlation of observed genetic structure with sea least-cost IBD model than with their explicit multi-generation dispersal model for the fish *Mullus surmuletus* (Dalongeville et al., 2018). When using the same data of *observed Fst*, our multi-generation implicit dispersal model returns a better correlation than IBD models while explicit dispersal models are not significant.

Our results also show that IBD models (Euclidian and sea-least cost distance) better explain observed genetic differentiation than the single-generation dispersal model, in accord with previous studies (Boulanger et al., 2020; Jahnke et al., 2018). About one third of the compiled studies displayed significant IBD predictions with a mean Mantel R^2, which is comparable to a previous meta-analysis (Jenkins et al., 2010). Note however that our results highlight no significant improvement when using sea least-cost distance rather than straight line distance, even though it may provide more realistic estimation of spatial isolation (Pante and Simon-Bouhet, 2013). Still, for some studies which do not consider multi-generation, single-generation dispersal models were reported to improve gene flow prediction as compared to IBD models (e.g. Alberto et al., 2011; Selkoe et al., 2010; Xuereb et al., 2018). The apparent contradiction with the present results may be due to a publication bias: single-generation dispersal predictions that were worse than IBD's ones could have been withheld by authors. Single-generation dispersal models are worse than IBD models to predict genetic

connectivity probably because IBD is supported by robust theory (Rousset, 1997). Indeed, IBD models have been proven useful for various terrestrial and marine taxa (Jenkins et al., 2010; Selkoe and Toonen, 2011). Second, broad-scale single-generation dispersal modelling studies often reported that most distant populations are not connected, suggesting genetic isolation that is often explained by dispersal barriers due to major oceanographic features such as fronts and jet-like currents (Pascual et al., 2017). Our results contradict this view: multi-generation dispersal models suggest that these population-pairs are connected through stepping-stone dispersal despite the supposed physical barriers (Ser-Giacomi et al., 2021). Notwithstanding, our models showed that physical barriers hamper the levels of connectivity (see Figure 8 of Ser-Giacomi et al., 2021 [1]) while exhibiting different levels of connectivity, as reflected in observed genetic differentiation and as predicted by theoretical gene flow magnitudes (Lowe and Allendorf, 2010; Lowe et al., 2017). Since IBD is an analytical model explaining genetic differentiation only by geographical distances, it does not allow disentangling the relative importance of evolutionary processes that control gene flow. In line with the distinction between IBD and Isolation-by-Environment (Wang et al., 2013), and since our mechanistic multi-generation dispersal models realistically simulate stepping-stone dispersal, one can tease apart the respective role of evolutionary forces in driving gene flow. Altogether, our results suggest that the seascape and its supposed physical barriers are more permeable to dispersal than previously thought (Pascual et al., 2017) and that genetic isolation in the marine realm could be rather due to environmental selection acting on drifting propagules and settled adults as well as intrinsic reproductive isolation (Bierne et al., 2011). Since both ocean currents (Sen Gupta et al., 2021), transport and mixing (Ser-Giacomi et al., 2020) and temperatures (Jorda et al., 2020) are already changing fast, the structure of marine populations is expected to fluctuate accordingly, consistently with the recent evidence of spatial reorganization of marine biodiversity (Blowes et al., 2019).

Our models perform better than previous ones probably also because they consider properly the mesoscale variability of ocean currents, they are parametrized with species-specific dispersal traits, and they allow testing explicitly what number of generations maximize correlations with observed data. Since there is no consensus on the adequate number of generations required to comprehend gene flow, other multi-generational approaches used shortest path algorithms (minimum number of steps to connect sampled population-pairs) with 25 or less intermediate steps (Boulanger et al.,

1. Chapitre III

2020; Buonomo et al., 2017) or set arbitrarily the number of generations from dozens (as it was considered sufficient to span the studied domain, Jahnke et al., 2018) to thousands (i.e. the number of Markov chain iterations needed to reach convergence) of generations (White et al., 2010). It illustrates that the typical time-scales over which demographic connectivity interplays with genetic connectivity are difficult to infer (Hellberg, 2009; Lowe et al., 2017). The relationship between the optimal number of generations and the SSRep normalized by the main dispersal trait (PLD) implies that the temporal scales estimated with our model (from 1 to several tens of generations, e.g. ecological time) and spatial scales derived from genetic methods (over typical evolutionary scales, from a hundred to a few thousands of kilometres) are tightly linked. This is aligned with estimates of dispersal kernels that were found congruent over ecological and evolutionary time (Pinsky et al., 2017). Moreover, *Fst* theory assumes that the whole population has reached an equilibrium between gene flow and genetic drift, suggesting that the coalescent connectivity model should predict best *observed Fst* for long-term multi-generation dispersal, that is when dispersal probabilities reach convergence (White et al., 2010), i.e. after about 500 generations in our case (see Figure 9 of Ser-Giacomi et al., 2021). The relatively low optimal number of generations disclosed here (~ 20) further suggests that gene flow and drift have insufficient time to reach equilibrium due to environmental stochasticity and rapidly changing ecological forces (Whitlock and Mccauley, 1999). Moreover, the substantial impacts of ecological processes on genetic structures shown here could explain why chaotic genetic patchiness has been recently documented at small space and time scales (Benestan et al., 2021; Eldon et al., 2016; Jackson et al., 2018; Pascual et al., 2016; Pérez-Portela et al., 2019; Schunter et al., 2019), which are also the scales over which dispersal and environment co-vary. As such, dispersal could be characterized as one of the evolutionary force shaping the contemporary spatial patterns of biodiversity, along with natural selection, providing evolutionary changes occurring over ecological timescales of few generations (Carroll et al., 2007; Schunter et al., 2019).

Last but not least, our multi-generation mechanistic dispersal models, which allow assessing both filial and coalescent connectivity and are applicable to other taxonomic groups and ecosystems, could serve as future guidelines to optimize sampling design for population genetic studies and anticipate the structure of wild sessile populations. In the context of biodiversity loss (Butchart et al., 2010) and spatial reorganization (Blowes et al., 2019), it is urgent to better understand the eco-evolutionary dynamics that continuously shape population structures to improve protection strategies (Sala

et al., 2021) and natural resources management natural resources such as fisheries (Hauser and Carvalho, 2008).

4.2.5. Methods

4.2.5.1. Studies characteristics

We screen the published literature of the last two decades to gather the population genetic studies focussing on marine species at basin-scale in the Mediterranean Sea. While our meta-analysis intends to be the most comprehensive possible in terms of collected data and taxa covered, pre-selected studies are filtered out based on two criteria: the biological traits of the species and the sampling design. In this way, we exclude datasets that are not appropriate to address our research question while optimizing statistical discriminatory power. More specifically, we select studies (i) whose target species is mostly sessile, i.e. that is characterized by a biphasic life cycle with early-life free-swimming dispersing propagules (e.g. seeds, eggs, larvae or body fragments) and full to semi-sedentary adult behaviour (Figure 4.1a), and (ii) which must present at least four distinct sampled populations with at least 15 individuals in each one of them. As a result, this meta-analysis compiles 58 population genetic studies published between 2002 and 2020, encompassing 47 different marine species distributed in nine taxonomic groups: Algae, Anthozoa, Ascidiacea, Crustacea, Demospongiae, Echinodermata, Fish, Mollusca and Phanerogam (Figure 4.1a). In total, 559 populations were sampled across the basin (dark green dots in Figure 4.1b,c), representing 3821 population-pairs. *Fst* fixation index allows evaluating the genetic differentiation between population-pairs using six types of genetic markers (allozymes, microsatellites, mitochondrial DNA sequences, RADseq/GBS and SNPs). Note that, when applicable, we analyse separately different genetic markers extracted from the same study, i.e. mitochondrial DNA and SNPs for *Ophioderma longicauda* (Weber et al., 2015) and *Hexaplex trunculus* (Marzouk et al., 2017).

To gauge the sampling strategy of each selected study, we compute the mean straight-line geographical distances (in km) from all sampled populations to their barycentre. This geometrical metric, that we called “Spatial Sampling Representativeness” (SSRep), quantitatively evaluates the spatial coverage of the sampling strategy carried out in each study. The larger the study-specific SSRep, the more scattered and spatially comprehensive is the sampling (see Annexe B, Table SI-1).

4.2.5.2. Species characteristics

Based on literature review (see Annexe B, section I), all species are classified according to their main dispersal traits. Reproductive phenology comprises five groups to reflect seasonal (spring, summer, fall, winter) and annual spawning strategies. Pelagic Larval Durations (PLD) are categorized in five groups: very-low (1 day), low (10 days), low-to-medium (20 days), medium-to-high (30 days) and high (45 days) dispersal abilities. Finally, two broad-scale classes of habitats are distinguished (see Annexe B, section II): the shallow coastal habitat (inner continental shelf whose depths span 0-50 m; Figure 4.1b) and the neritic shelf habitat (mid to outer continental shelf whose depths range is 50-200 m; Figure 4.1c).

4.2.5.3. Bio-physical modelling

Tracking passive Lagrangian particles is a common approach to characterize flow-driven dispersal of propagules (Cowen et al., 2000; Cowen et al., 2006; Dubois et al., 2016; Legrand et al., 2019; Lett et al., 2020; Rossi et al., 2014; Figure 4.2a). To provide synthetic -yet realistic- views of basin-scale propagules dispersal while encompassing the full variability of ocean currents and for various dispersal abilities, we use the Lagrangian Flow Network framework (LFN, Dubois et al., 2016; Rossi et al., 2014; Ser-Giacomi, Rossi, et al., 2015). It combines a particle-tracking model with graph theory tools to generate and analyse connectivity matrices (see Annexe B, section III), allowing us to investigate oceanic dispersal in a robust and efficient manner (Monroy et al., 2017).

The Mediterranean Sea surface characterized by favourable habitats is subdivided into several $\frac{1}{4}$ °sub-areas that represent theoretically isolated marine populations, resulting in $n = 1170$ populations in the shallow coastal habitat (Figure 4.1b) while the neritic shelf habitat is composed of $n = 1163$ populations (Figure 4.1c; see Annexe B section II). For each LFN experiment, we track 100 propagules per population by integrating daily gridded velocity fields generated by a data-assimilative operational ocean model implemented in the Mediterranean Sea at a $\frac{1}{16}$ °horizontal resolution (Oddo et al., 2009). We use the horizontal flow field at 10 m and 100 m for species inhabiting shallow coastal and neritic shelf habitats, respectively. Overall, virtual propagules trajectories are modelled at two specific depths during the five PLDs simulating consecutive propagule release events with a 10-day periodicity over 2000 to 2010. Assuming that the long-term (e.g. decadal, centennial and longer) variability of ocean currents is negligible as compared to their inter- and intra-annual variations,

our approximation using 10 recent years allows comprehending the full variability of both contemporary and past oceanic flows. By computing billions of Lagrangian trajectories and recording their initial and final positions, the LFN construct 402 matrices for each of the 10 habitat/PLD combinations, resulting in 4020 matrices in total. The elements m_{ij} of each raw n x n connectivity matrix encode the number of propagules advected between all population-pairs; they are converted into backward-in-time dispersal probabilities thanks to a column-normalization $m_{ij} = \frac{m_{ij}}{\sum_{i=1}^{n} m_{ij}}$. Then connectivity matrices M are aggregated (i.e. element-by-element averaged) according to their starting dates to match each species-specific spawning phenology (i.e. averaging 402 matrices for year-round release and about 100 matrices for seasonal release, see Annexe B Table SI-4). In other words, we compute for each species a composite matrix P which fits best with its dispersal traits, averaging ten years of realistic current-driven dispersal in the Mediterranean Sea. Single-generation dispersal estimates are directly extracted from one of these composite matrices. As explained next, multi-generational dispersal estimates are finally obtained by applying additional computations on these composite matrices.

4.2.5.4. Cumulating implicit and explicit links in multi-generation dispersal models

Explicit pathways of connectivity correspond to direct filial links from parents to children, assuming unique lineage (i.e filial connectivity, Figure 4.2b). It is the usual proxy of connectivity used in other multi-generation models to assess gene flow between populations (Boulanger et al., 2020; Buonomo et al., 2017; Foster et al., 2012; Jahnke et al., 2018; Kool et al., 2010; White et al., 2010). Implicit pathways of connectivity (Ser-Giacomi et al., 2021) evaluates indirect links among siblings with common parents, considering multiple lineages (i.e. coalescent connectivity, Figure 4.2c). To estimate multi-generation dispersal probabilities between all population-pairs, considering explicit or implicit links, we apply the theoretical formulations described in (Ser-Giacomi et al., 2021) on composite matrices. The main novelties are the adequate consideration of putative intermediate stepping-stones as any of the non-sampled populations of both habitats (Figure 4.1b,c) and the fact that it allows cumulating connectivity pathways over each consecutive generation, i.e. from generation 1 to a fixed number of generations. Analytical formulations for both cumulated explicit and implicit probabilities for any number of generation are established in (Ser-Giacomi et al., 2021) and are theoretically bounded to 1 for an infinite number of generations.

— The general expression of explicit dispersal probabilities over $G = 2$ generations based on the composite matrix P is:

$$P^{G=2} = P + P(L \circ P) \tag{4.1}$$

The circle denotes the Hadamard product, and L is the all-ones matrix minus the identity matrix. When applying equation (4.1) for two populations A and B (example illustrated in Figure 4.2b), explicit link cumulates: (i) the sampled population-pair explicit probability (P_{AB}) and (ii) the products of probabilities between sampled populations and their second generation intermediate population ($P_{Ak_1}P_{k_1B}$, Figure 4.2b), that is :

$$P_{AB}^{G=2} = P_{AB} + P_{Ak_1}P_{k_1B} \tag{4.2}$$

— The general expression of implicit dispersal probabilities over $G = 2$ generations based on the composite matrix P is:

$$P^{G=2} = P^tP + P[L \circ (P^tP)]^tP \tag{4.3}$$

As before, the circle denotes the Hadamard product, and L is the all-ones matrix minus the identity matrix. When applying equation (4.3) for two populations A and B (example illustrated in Figure 4.2c) implicit link cumulates: (i) the product of probability between sampled populations and their common source (i.e. parent) populations ($P_{Ak_1}P_{Bk_1}$); and (ii) the product of probability between sampled populations and their common source populations through two generations ($P_{AI_A}P_{I_Ak_2}P_{BI_B}P_{I_Bk_2}$, Figure 4.2c), that is:

$$P_{AB}^{G=2} = P_{Ak_1}P_{Bk_1} + P_{AI_A}P_{I_Ak_2}P_{BI_B}P_{I_Bk_2} \tag{4.4}$$

The Hadamard product vanished in equation (4.2) and equation (4.4) because there is no self-loop (i.e. self-recruitment) in none of our simplified exemplary populations. If self-recruitment exists, e.g. if $k_1 = A$ in Figure 4.2c), siblings are found in both origin and destination populations implying that implicit links also encompass explicit links. Since *Fst* are theoretically symmetrical (i.e. Fst_{AB} equals Fst_{BA}), explicit dispersal probabilities have been transformed following $1-(1-P_{AB})*(1-P_{BA})$ to be symmetrical. Note that implicit dispersal probabilities between population-pairs are symmetrical by construction. Both

multi-generation explicit and implicit dispersal models are computed for 1, 5, 10, 20, 40, 60, 80, 100, 150, 200, 300, 400 and 500 generations using species-specific composite matrices.

4.2.5.5. Translating dispersal probabilities into proxies of genetic differentiation

We use the theoretical formulation of the Island model (Wright, 1931), $Fst = 1/(4N_em+1)$, which assumes no selection nor mutation, and equilibrium between migration and drift. Two parameters are used in this model: Ne the effective population size (i.e. the number of individuals that reproduce) and m the migration rate (ranging between 0 and 1). N_em is thus the number of migrants successfully entering a population. Considering the obvious analogy between the migration rate m and our simulated dispersal probabilities, we compute *modelled Fst* for each sample pair of each study by applying the Island model formulae on our dispersal probabilities. As such, a *modelled Fst* can be interpreted as a probability of gene exchange through dispersal, and it can be readily compared against *observed Fst* for all population-pairs compiled in our meta-analysis. Moreover, the Island model permits to handle null connection probabilities (which are often found between distant population-pairs for small generations) without artificially adding very low probabilities due to the $\log^{10}$ transformation done in Crandall et al., 2012 and Jahnke et al., 2018. Note that Ne is only a non-linear scaling factor that do not affect our results since they consist exclusively in Mantel correlations (R^2 primarily depends on m which is given by multi-generation dispersal probabilities) rather than the slope of the *observed vs modelled Fst* relationship. Since estimating Ne requires complex analyses (reviewed in Hare et al., 2011), sensitivity tests on the values of Ne have been performed and revealed little dependence of our results for any Ne spanning 100 to 10^4 (see Annexe B Figure SI-4). We choose then to set $Ne = 100$ because it matches the number of artificial propagules released per population. Moreover, we neglect local demography dynamics for all populations, as assumed in the Island model, and we assume constant larval mortality and recruitment over space and time. We then translate single- and multi-generation explicit and implicit estimates into *modelled Fst* for all the species types described above.

4.2.5.6. Mantel tests

We investigate correlations between the *modelled Fst* derived from multi-generation dispersal models for 1 to 500 generations and *observed Fst* for all studies of the meta-

analysis with Mantel tests. For each study, we parametrize our models by selecting the optimal number of generations as the one for which the Mantel R^2 display the highest value. When imposing the same number of generations for all studies, the one maximising the significant Mantel correlations appears to be 40 generations for both explicit and implicit models (17 and 18 significant studies respectively, see B section I and IV). In general, for medium to high dispersal species (i.e. PLD ≥ 20 days), the optimal number of generations is between 20 and 40 generations for explicit model and 20 generations for implicit model. The optimal highest number of generations (≥ 100) are only obtained for very-low dispersal species characterized by 1-day PLDs (see B section I and IV). We use the optimal generations of each study for all subsequent analyses, including the *modelled Fst* derived from all dispersal probabilities (single- and multi-generation, considering both explicit and implicit models). In addition, we perform Isolation-By-Distance (IBD) analyses of all compiled genetic structures with two proxies of distance between population-pairs: the Euclidian distance (i.e. straight-line geographical distances) and the sea least-cost distance, which corresponds to the length of the shortest path considering only maritime areas. Sea least-cost distances are calculated thanks to the Marmap package (Pante & Simon-Bouhet, 2013, version 1.0.4) on R (version 4.0.2). For the IBD analyses, we consider a two-dimensional dispersal model and thus compared $\log_e$(distance) with $1/(1-Fst)$ (Rousset, 1997). Mantel tests are performed with the scikit-bio package (version 0.5.6) with Python (version 3.7.3) using Pearson correlations and 10000 randomizations (see Annexe B section V).

4.2.5.7. Statistics at the meta-analysis scale

We apply Fisher's combined probability tests (Fisher, 1934) to examine all study-specific Mantel *p-values* associated to the single-generation dispersal models and to the optimal generation estimates using the same null hypothesis (H_o: *modelled Fst* and *observed Fst* are not correlated). If H_o is rejected, the Fisher's combined probability test gives the significance of the correlation at the meta-analysis scale. We applied the same method to test the IBD models. For each study, we compare Mantel correlations among them to determine the best method to appraise gene flow. More specifically, we cross-compare model predictions across the entire meta-analysis: (i) single-generation explicit dispersal model against both multi-generation explicit and implicit dispersal models (Hendrickson et al., 1970), (ii) multi-generation explicit dispersal model against multi-generation implicit dispersal model (Hendrickson et

al., 1970), (iii) IBD model against single-generation explicit dispersal model (Silver et al., 2004) and (iv) IBD model against both multi-generation explicit and implicit dispersal models (Silver et al., 2004) using the R-package COCOR (Diedenhofen & Musch, 2015, version 1.1-3). For all tests, we set the one-sided "model A correlations are greater than model B correlations" alternate hypothesis at an 0.05 alpha level (see B section VI). We finally use Fisher's combined probability test on the p-values for the meta-analysis interpretation. We test the sensitivity of the Mantel correlations results (R^2 and p-values) against species-specific (taxonomic group, PLD, spawning season) and studies-specific (marker, number of sampled populations, SSRep and *Fst* range) characteristics with ANOVA and linear regressions (see Annexe B section VII). We use a logistical regression with a binomial distribution to model the probability of obtaining significant gene flow predictions as a function of the number of samples. Sensitivity tests are performed using the Matlab software (version 9.4). Throughout the entire manuscript, asterisks inform about the significance of all statistical tests, as follows: * ≤ 0.05; ** ≤ 0.005 ; *** ≤ 0.0005 and "^{ns}" stands for not-significant.

4.2.6. Acknowledgments

T.L. is funded by a Doctoral fellowship obtained through Aix-Marseille University. T.L., A.C. and V.R acknowledge financial support from the European project SEAMoBB (Solutions for sEmi-Automated Monitoring of Benthic Biodiversity), funded by ERA-Net Mar-TERA (id. 145) and managed by the ANR (Grant No. ANR7-MART0001-02, P.I.: A. Chenuil). V.R., E.S.-G., N.B. and S. A.-H. acknowledge financial support obtained through the HYDROGENCONNECT project (P.I. V. Rossi) funded by the French program MISTRALS ENVI-Med. T.L. and V.R. warmly thank Madiop Lo for the technical improvements implemented in the LFN model. V.R., A.C., S.A.-H. and N.B. thank Barbara Porro for her 2015/2016 master thesis that allow initiating the data compilation.

5. La considération de l'habitat dans la modélisation de la connectivité génétique

5.1. Avant-propos

Avant de conclure, nous allons, dans ce dernier chapitre, utiliser des estimations de distribution spatiale d'espèce ou d'écosystème afin de mieux comprendre comment la fragmentation des habitats affecte le calcul de la dispersion multi-génération. En d'autres termes, ce travail met l'accent sur la prise en compte de l'hétérogénéité spatiale de la démographie (et implicitement sur l'efficacité du processus de ponte et d'installation) dans la modélisation de la connectivité génétique, et questionne l'impact relatif de son intégration dans la prédiction du flux de gènes. Ce dernier chapitre est particulier : plus court que les précédents, il n'est pas sous la forme d'un article scientifique publié ou soumis, et est rédigé en français. Il se veut comme un approfondissement du travail réalisé dans le chapitre précédent et pourra être transformé en article ultérieurement.

5.2. Introduction

Pour définir les limites spatiales des deux habitats présentés dans le Chapitre IV (i.e. *shallow coastal* and *neritic shelf*), nous avons couplé un seuil bathymétrique avec des données spatiales de type de substrat (SIG, issue du portail européen EMODnet Seabed Habitats [1]). Précisément, nous avons rendu « *actifs* ou *inactifs* » les nœuds des réseaux de connectivité représentant le transport de propagule pour une PLD fixée en

1. information contained here has been derived from data that is made available under the European Marine Observation Data Network (EMODnet) Seabed Habitats initiative (http://www.emodnet-seabedhabitats.eu/), financed by the European Union under Regulation (EU) No 508/2014 of the European Parliament and of the Council of 15 May 2014 on the European Maritime and Fisheries Fund.

mer Méditerranée, en fonction de leurs positions. Si ces nœuds étaient situés entre deux isobathes préalablement définis (i.e. 0-50 m et 50-200 m) ou qu'ils contenaient un substrat particulier (i.e. associé à l'infralittoral ou le circalittoral), ils étaient considérés comme un habitat possible dans le calcul de la dispersion multi-génération. Dans le cas contraire, c'est-à-dire pour des nœuds non caractérisées par la présence d'habitat favorable, le flux de gènes est rendu impossible. Il était difficile d'obtenir un raffinement supplémentaire à l'échelle de la méta-analyse en considérant la grande variété de substrats (e.g. sables, vase, rochers, etc.) car la disponibilité de telles données spatiales est hétéroclite en Méditerranée. Les côtes nord occidentales sont en général bien représentées, alors que pour les côtes sud ou orientales, les données sont souvent éparses et imprécises. Cette disparité représente une des difficultés principales pour une délimitation d'habitat adéquate : les nœuds localisés le long des côtes sud et orientales de la Méditerranée seraient le plus souvent considérés inactifs, ce qui induirait une évaluation biaisée du flux de gènes à l'échelle du bassin. Considérant l'ensemble des espèces compilées, une détermination relativement « rudimentaire » de ces deux types d'habitats à partir de la bathymétrie et des données géographiques de substrats était une approximation nécessaire à la généralisation et la cohérence de l'approche méta-analyse, qui a permis d'avoir une vision globale des processus de connectivité génétique.

En se focalisant sur une seule espèce et sur un domaine moins étendu, d'autres études de connectivité génétique couplée à de la dispersion multi-génération simulée ont utilisé (i) des données précises de substrat cartographié par des experts (i.e. habitat rocheux dans le sud de l'Italie et le détroit de Sicile afin de modéliser l'habitat de l'algue *Cystoseira amentacea*, Buonomo et al., 2017, de la même manière que réalisé dans le Chapitre II pour la définition des habitats des sars) ou (ii) des données récentes et historiques de distribution de l'espèce considérée (l'herbier *Zostera marina* en mer Baltique, Jahnke et al., 2018). Toutefois, à l'échelle du bassin méditerranéen, Boulanger et al., 2020 n'ont pas pris pas en compte le substrat préférentiel du rouget-barbet *Mullus surmuletus*, et ont considéré tous les nœuds côtiers de leur réseau de connectivité comme habitat favorable. Il en est de même à une échelle beaucoup plus locale pour White et al., 2010 avec l'étude du gastéropode *Kelletia kelletii* dans le canal de Santa Barbara au sein de l'upwelling de Californie (Pacifique Nord-Est).

L'utilisation de modèles de niche (*species distribution models*, SDMs) permet de mieux définir les habitats dans la compréhension de la connectivité à large échelle (Mari et al., 2020), et pourrait apporter une réelle plus-value à la modélisation de

la connectivité génétique sur de multiples générations de dispersion. Ces modèles permettent de prédire la répartition spatiale d'une espèce ou d'un écosystème à partir d'observations de présence (et plus rarement d'absence) qui sont corrélées à des variables environnementales afin d'en déterminer des enveloppes favorables autorisant ensuite une extrapolation spatiale (Guisan & Thuiller, 2005). Les prédictions quantitatives de distribution ainsi obtenues sont continues et homogènes dans l'espace, et ce potentiellement à large échelle. Il faut toutefois considérer les possibles incertitudes liées aux données d'occurrences d'espèces et à la précision des données environnementales. En estimant la probabilité d'occurrence d'une espèce ou d'un écosystème, les SDMs permettent d'aller plus loin que la délimitation binaire d'habitats favorables : cette probabilité peut servir à moduler les probabilités de connexion obtenues par la modélisation du transport de propagules par les courants marins. Cette modulation peut se faire en amont du transport, en émettant l'hypothèse que plus la probabilité de présence d'un organisme est grande, plus l'émission de propagules est importante, ou en d'autres termes, plus le processus de ponte est efficace. De même en aval : plus la probabilité de présence est grande, plus le processus d'installation est efficace (notamment pour les espèces non limitées par une taille maximale de leurs populations). Cette méthodologie de modulation de la connectivité entre population par le taux de propagules émises et l'efficacité de l'installation a été appliquée sur les récifs coralliens caribéens (Foster et al., 2012 ; Kool et al., 2010) et les herbiers de posidonie en Méditerranée (Mari et al., 2020), en bénéficiant respectivement des données cartographiques précise de présence d'habitat (Andréfouët & Bionaz, 2021) et des données de probabilité d'occurrence d'espèce issues de SDMs (Giannoulaki et al., 2013).

Ainsi, la question scientifique abordée ici est la suivante : quel est l'impact de la considération de la démographie et de sa variabilité spatiale sur la prédiction de flux de gènes en utilisant le modèle de connectivité coalescente (Chapitre III et IV) ? Une première hypothèse serait que le nombre de générations optimales pour représenter la connectivité génétique augmenterait, les probabilités de connexion sous-jacentes étant inférieures à cause de la modulation par les probabilités d'occurrences à la ponte et à l'installation. Une deuxième hypothèse serait qu'en définissant mieux l'habitat et en considérant dans les simulations la variabilité spatiale du processus de ponte et d'installation de l'espèce considérée, les prédictions du flux de gènes seraient améliorées. Pour vérifier ces hypothèses, nous avons utilisé des données de probabilité de présence d'herbiers de posidonie (Giannoulaki et al., 2013) et de coralligène (Martin

TABLEAU 5.1. – Traits biologiques des premiers stades de vie des quatre espèces modèles.

Espèce	Habitat spécifique	Habitat global	PLD (j)	Saison de ponte	Référence
P. oceanica	Posidonie	Shallow coastal	30	Toute l'année	Melià et al., 2016; Serra et al., 2010
P. lividus	Posidonie	Shallow coastal	30	Printemps	Pedrotti, 1993
P. clavata	Coralligène surface	shallow coastal	1	Printemps	Guizien et al., 2020
C. rubrum	Coralligène profond	Neritic shelf	10	Eté	Coelho et Lasker, 2016; Costantini et al., 2013; Teixidó et al., 2011

et al., 2014) en Méditerranée, pour mieux définir l'habitat de la posidonie, de l'oursin violet, du corail rouge et de la gorgone pourpre dans la modélisation de la connectivité coalescente via la dispersion multi-génération.

5.3. Méthode

Espèces modèles. Nous avons utilisé quatre espèces modèles et leurs études phylogéographiques correspondantes : la posidonie (*Posidonia oceanica*, Arnaud-Haond et al., 2007) et l'oursin violet (*Paracentrotus lividus*, Paterno et al., 2017) respectivement constitutives et habitant l'écosystème de l'herbier de posidonie; le corail rouge (*Corallium rubrum*, Aurelle et al., 2011) et la gorgone pourpre (*Paramuricea clavata*, Mokhtar-Jamaï et al., 2011) habitant l'écosystème à coralligène. Leurs traits biologiques des premiers stades de vie, ainsi que les caractéristique des études de génétique des populations associées sont référencées respectivement dans le Tableau 5.1 et 5.2.

Modèle de distribution spatiale d'habitat. Nous avons utilisé les données de distribution spatiales modélisées d'écosystèmes spécifiques sur l'ensemble de côtes Méditerranéennes. Il s'agit de l'herbier de posidonie (Giannoulaki et al., 2013) et du coralligène (Martin et al., 2014), tous deux disponibles depuis le portail EMODnet *Seabed Habitat*[1]. La modélisation de la distribution de ces deux écosystèmes clefs de Méditerranée, ici considérés comme des habitats, s'inscrit dans le projet Européen MEDISEH (Giannoulaki et al., 2013) qui avait pour but de mieux définir et caractériser

TABLEAU 5.2. – Caractéristique des études de génétique de population utilisées SSRep signifie Spatial Sampling Representativenes soit la distance moyenne entre les sites d'échantillonnage et leurs barycentre (voir Chapitre IV).

Espèce	Marqueur génétique	Nombre de pop	SSRep (km)	Reference
P. oceanica	Microsatellites	29	859	Arnaud-Haond et al., 2007
P. lividus	SNPs	10	758	Paterno et al., 2017
P. clavata	Microsatellites	24	687	Mokhtar-Jamaï et al., 2011
C. rubrum	Microsatellites	39	842	Aurelle et al., 2011

les habitats sensibles de Méditerranée. Dans ce travail, on émet l'hypothèse que la probabilité d'occurrence de ces deux habitats est corrélée avec la probabilité d'occurrence d'espèces constitutives (i.e. la posidonie) ou inféodées (i.e. l'oursin violet, la gorgone pourpre et le corail rouge). La prédiction de la distribution spatiale des herbiers de posidonie est basée sur des informations de présence-absence de l'espèce (voir Figure 1 de Telesca et al., 2015) et l'utilisation de 36 variables environnementales (e.g. bathymétrie, caractéristiques physico-chimiques, concentration de polluants et indicateurs d'impacts anthropiques). L'utilisation d'un modèle de niche corrélatif permet ainsi d'obtenir une cartographie de la probabilité d'occurrence des herbiers (comprise entre 0 et 1) à haute résolution 1/240°, soit environ un point tous les 500 m sur l'ensemble du bassin méditerranéen. De la même manière, la modélisation de la distribution du coralligène à haute résolution (environ un point tous les 400 m) fut établit grâce à la compilation d'observations de présence d'espèces le composant et de 17 variables environnementales prédictives. Il en résulte une cartographie continue de la probabilité d'occurrence (comprise entre 0 et 1) du coralligène le long des côtes méditerranéennes (voir Figure 4a de Martin et al., 2014).

Interpolation des données de distribution spatiale sur le réseau de connectivité. En tirant profit des outils de la théorie des graphes, le modèle Lagrangian Flow Network (LFN ; Dubois et al., 2016 ; Rossi et al., 2014 ; Ser-Giacomi, Rossi et al., 2015) utilisé pour simuler le transport larvaire, subdivise premièrement le domaine océanique en un ensemble de nœuds d'une résolution de 1/4°. De la même manière que dans le Chapitre IV, nous avons choisi d'implémenter cette subdivision sur les champs de courant du modèle hydrodynamique opérationnel Mediterranean Forecasting System (Oddo et al., 2009, voir Annexe B section III) aux profondeurs verticales de 10 m (3516 nœuds) et 100 m (3137 nœuds). Dans ce travail, nous avons ainsi modélisé la dispersion uniquement sur la dimension horizontale en retenant la profondeur

moyenne de présence de l'espèce considérée. Les données de distribution spatiale des herbiers de posidonie et du coralligène sont ensuite interpolées (« *coarse-graining* ») sur la grille du LFN en calculant la moyenne pondérée des probabilités d'occurrences contenues dans chaque nœud de 1/4°. Tous les nœuds caractérisés par une probabilité d'occurrence non nulle représentent alors des populations distinctes établies sur des patchs d'habitat de posidonie ou de coralligène et considérées comme isolées (i.e. reliées uniquement par la dispersion de propagules). Nous avons ainsi caractérisé pour chaque population les probabilités d'occurrence d'herbiers de posidonie et de l'oursin violet (Figure 5.1a, 1106 populations) de la gorgone pourpre en surface (Figure 5.1b, 1404 populations) ainsi que du corail rouge en profondeur (Figure 5.1c, 1085 populations). L'interpolation des données de distribution spatiale sur le réseau de connectivité représentent ainsi les trois habitats spécifiques utilisés pour modéliser la connectivité coalescente de nos quatre espèces cibles. Les habitats « globaux » sont les mêmes que décrits dans le Chapitre V (i.e. *shallow coastal* ou *neritic shelf*), et servent de références dans cette étude (Figure 5.1).

Modélisation Lagrangienne. La même méthodologie que celle du Chapitre IV est utilisée pour obtenir, pour chaque espèce modèle, deux matrices de connectivité composites synthétisant 10 années de dispersion en mer Méditerranée et s'ajustant aux traits de premier stade de vie de l'espèce étudié. La matrice P_G représente les habitats globaux du Chapitre IV (de dimension $1170 * 1170$ pour l'habitat *shallow coastal* et $1163 * 1163$ pour l'habitat *neritic shelf*) et la matrice P_S représente l'habitat spécifique défini précédemment (de dimension $1106 * 1106$ pour l'habitat Posidonie, de dimension $1404 * 1404$ pour l'habitat Coralligène surface et de dimension $1085 * 1085$ pour l'habitat Coralligène profond). Ces matrices indiquent ainsi la probabilité moyenne pour une population de recevoir des propagules d'une autre population en considérant de 100 à 400 évènements de ponte (évènement de ponte saisonnier ou annuel respectivement, voir Figure 5.1) pendant 10 ans : elles quantifient de manière synthétique des processus de transport complexes dus à la circulation marine (Figure 5.2a).

Prise en compte de la probabilité d'occurrence dans la modélisation de la dispersion. En utilisant la probabilité d'occurrence des deux habitats extrapolé à chaque population définie précédemment, nous avons construit une matrice H de même dimension que P_s qui associe à chaque paire de population la probabilité de ponte de la particule envoyée (Figure 5.2b). De manière similaire, sa transposé ${}^tH$ associe à chaque paire de population la probabilité d'installation d'une particule reçue. Ainsi,

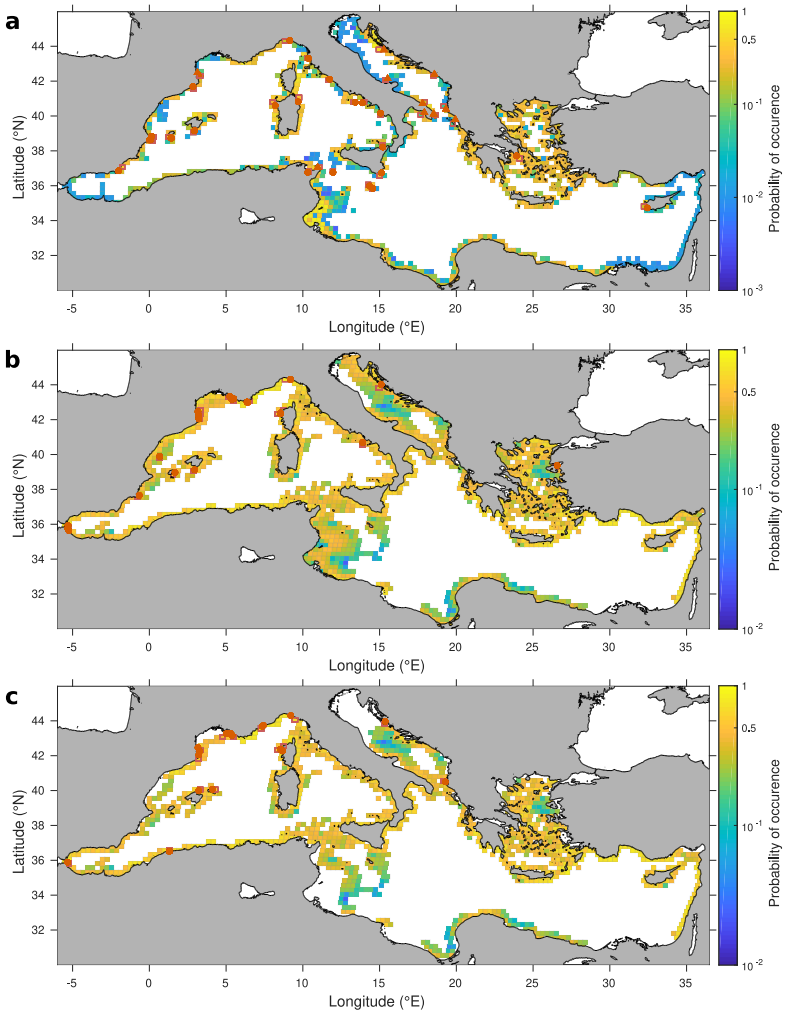


FIGURE 5.1. – Cartographie des différents habitats spécifiques et des probabilités d'occurrence associées. Les sites d'échantillonnage des quatre espèces modèles sont indiqués par les symboles rouges. Les nœuds associés aux populations échantillonnées sont encadrés en rouge. **a** Habitat Posidonie, interpolé sur la couche du modèle hydrodynamique à 10 m de profondeur. Les sites d'échantillonnage de la posidonie (Arnaud-Haond et al., 2007) et de l'oursin violet (Paterno et al., 2017) sont représentés respectivement par les cercles et les triangles rouges. **b** Habitat Coralligène de surface, interpolé sur la couche du modèle hydrodynamique à 10 m de profondeur. Les sites d'échantillonnage de la gorgone pourpre (Mokhtar-Jamaï et al., 2011) sont représentés par les cercles rouges. **c** Habitat Coralligène de fond, interpolé sur la couche du modèle hydrodynamique à 100 m de profondeur. Les sites d'échantillonnage du corail rouge (Aurelle et al., 2011) sont représentés par les cercles rouges.

nous avons défini une matrice finale qui considère successivement la probabilité de ponte, la probabilité de transport et la probabilité d'installation (Figure 5.2c ; Foster et al., 2012 ; Kool et al., 2010 ; Mari et al., 2020) :

$$P_S^H = H \circ P_S \circ {}^tH \tag{5.1}$$

Le signe $\circ$ correspond au produit matriciel de Hadamard. La matrice P_S^H est ensuite normalisée en divisant chaque terme par la somme de sa colonne, afin que pour chaque population, la probabilité totale de recevoir des particules soit égale à 1.

Calcul de la connectivité coalescente. Pour chaque étude, on applique le modèle de dispersion multi-génération identifiant les liens implicites (Chapitre III, Ser-Giacomi et al., 2021) sur les matrices P_G et P_S^H. Nous considèrons ainsi la connectivité génétique pour des habitats globaux (de la même manière que dans le chapitre IV) et pour des habitats spécifiques (i.e. herbiers de posidonie et coralligène). Le calcul des probabilités de connexion coalescente est réalisé pour un nombre de génération allant de 1 à 250 (jusqu'à 500 génération pour *Paramuricea clavata* car elle est caractérisée par une PLD de un jour, voir Figure 4.5b du Chapitre IV).

Prédiction du flux de gènes. Enfin, les indices de fixation génétique (Fst) et les probabilités de connexion multi-génération associées aux mêmes paires de populations sont comparés pour chaque étude avec des tests de Mantel réalisés sur R (version 4.0.5) via le package $vegan$ (version 2.5-7). De la même manière que dans le chapitre IV (section 4.2.5.5), les probabilités de connexion, étant fonctions inverses des Fst, sont d'abords transformées en suivant le modèle d'île de Wright, 1931 : $Fst = \frac{1}{4N_e m+1}$. N_e est fixée à 100 (équivalent au nombre de propagules modélisées pour chaque populations) et m est donné par les probabilités de connexion coalescente calculées entre paires de populations. Notons que toutes les corrélations présentées en résultat sont significatives au seuil de 5 %.

5.4. Résultats & Discussion

Dans ce chapitre, nous avons utilisé des données de différentiation génétique de quatre espèces constitutives ou inféodées dans deux habitats écologiquement clefs de Méditerranée. L'objectif est ici de tester l'effet de la considération de la distribution spatiale de ces habitats dans le modèle de dispersion multi-génération sur la précision des prédictions de flux de gènes via la connectivité coalescente. Les deux hypothèses

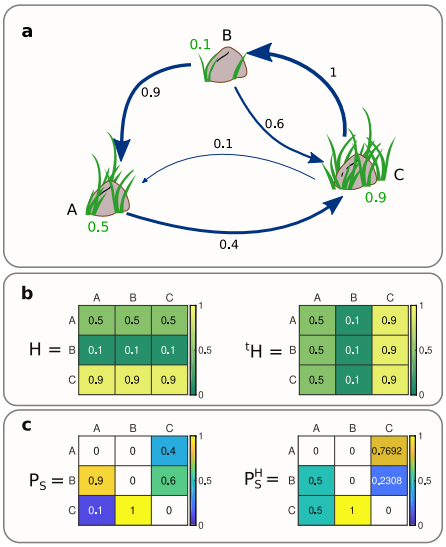


FIGURE 5.2. – Schéma illustrant la prise en compte de la probabilité d'occurrence des deux habitats considérés dans la modélisation de la dispersion. **a** Probabilité de connexions entre trois populations caractérisées par des probabilités d'occurrence d'herbiers de posidonie différentes. **b** Matrices associant à chaque paire de population la probabilité de ponte (H) ou d'installation (^tH). **c** Matrices composites synthétisant les probabilités de connexions entre paires de populations (Ps), et en prenant en compte les probabilités d'occurrence d'herbiers de posidonie (P_s^H).

de recherche sous-jacentes à cette considération des probabilité d'occurrence sont : (i) que le nombre de générations optimales augmenterait, (ii) que les prédictions du flux de gènes seraient améliorées.

Parmi les quatre espèces considérées, seule la prédiction de flux de gènes de la posidonie (Arnaud-Haond et al., 2007) répond positivement à ces deux hypothèses. Le nombre de générations optimal avec l'habitat spécifique est de 72 générations, soit 24 générations de plus qu'avec l'habitat global. Le coefficient de corrélation de Mantel est également faiblement supérieur : $R^2_{spécifique} = 0.35$ contre $R^2_{global} = 0.33$ (Figure 5.3a). Toutefois, la corrélation avec l'habitat spécifique est significativement supérieure à celle obtenue avec l'habitat global (p-value = 0.0185, Hendrickson et al., 1970). Pour l'oursin violet (Paterno et al., 2017), le nombre de générations optimal avec l'habitat spécifique est également supérieur, 74 générations contre 40 générations pour l'habitat global. Cependant, le coefficient de corrélation est similaire entre les deux habitats ($R^2_{spécifique} = R^2_{global} = 0.77$). Cela peut être dû au fait que l'oursin violet n'est pas uniquement présent sur les herbiers de posidonie et peut également s'établir sur d'autres substrats, comme les fonds rocheux ou sableux. La distribution spatiale de la posidonie serait ainsi trop restrictive par rapport à un habitat global (i.e. qui prend en compte tous les substrats possibles d'un étage littoral, voir Chapitre IV) pour modéliser plus efficacement son flux de gènes. Le caractère restrictif des prédictions de distribution spatiale par les SDMs pourrait ainsi induire un biais potentiel dans la modélisation de la connectivité génétique. Il faut également considérer les inconsistances potentielles dans les données spatiales empiriques de présence (et possiblement d'absence pour les espèces sessiles) des espèces, utilisées pour paramétrer les SDMs : erreurs d'identification, approximations géographiques, étendue des données d'occurrence non représentative de l'aire de répartition, biais d'échantillonnage (e.g. entre côte nord et sud de la mer Méditerranée), etc. Pour ces deux espèces constitutives ou présentes dans les herbiers de posidonie, le rapport entre le nombre de générations optimal pour l'habitat global et l'habitat spécifique est du même ordre de grandeur, soit 1/3 pour la posidonie et 1/2 pour l'oursin violet. Du fait que ces deux espèces sont caractérisées par la même PLD (i.e. 30 jours, Tableau 1) et par une SSRep quasi-similaire (i.e. 859 km et 758 km, respectivement), cela indiquerait que la relation entre les échelles spatiales et temporelles montrée dans le chapitre précédant (i.e. Figure 4.5b du Chapitre V) serait conservative.

Pour les deux espèces présentes sur le coralligène, aucune des deux hypothèses de recherche n'est vérifiée par les résultats. Le nombre de générations optimales

avec l'habitat spécifique est deux fois inférieur à celui avec l'habitat global pour la gorgone pourpre (11 vs 22), et le coefficient de corrélation est également légèrement inférieur (R^2_{global} = 0.22 vs $R^2_{spécifique}$ = 0.21, Figure 5.3c). Pour le corail rouge, et contrairement aux trois autres études, les courbes décrivant la valeur du coefficient de Mantel par rapport au nombre de générations sont quasi-similaires pour les deux habitats : elles augmentent linéairement jusqu'à atteindre à 36 générations un seuil à R^2 = 0.5. Ces résultats contradictoires aux hypothèses peuvent être dus à l'effet de l'échantillonnage centré sur la Méditerranée nord occidentale pour ces deux espèces (87 % pour la gorgone pourpre et 86 % pour le corail rouge), qui coïncide avec une zone d'homogénéité relative de la distribution spatiale du coralligène qui y apparait peu fragmenté (Figure 5.1b,c).

Afin d'étudier l'adaptation locale, Mokhtar-Jamaï et al., 2011 ont pour moitié échantillonné la gorgone pourpre de manière très localisée et sur plusieurs profondeurs (e.g. 7 sites dans la région catalane, 13 sites dans la région marseillaise, 5 sites au nord de la Corse), tout en prélevant d'autres populations plus distantes (e.g. un unique site d'échantillonnage en mer Adriatique et Egée). Appliqué à nos problématiques et avec notre méthodologie, un tel échantillonnage combinant fine et large échelle est moins adapté. En effet, en étudiant la connectivité génétique et modélisant le flux de gènes à l'échelle du bassin Méditerranéen, notre modèle ne peut pas résoudre les différences génétiques à l'échelle de quelques dizaine de km ni sur différentes profondeurs. Une telle étude de la connectivité génétique à fine échelle nécessiterait un modèle hydrodynamique tridimensionnel local à haute résolution, comme celui utilisé dans Rossi et al., 2020 pour modéliser la connectivité démographique du coralligène (à partir d'une cartographie haute résolution de l'écosystème) dans la baie de Marseille. Il faudrait également un modèle bio-physique qui simule la trajectoire des propagules sur les trois dimensions de l'espace. Ici, les sites d'échantillonnage qui sont seulement distants d'une dizaine de km sont pour la plupart caractérisés par un même nœud, ou par un nœud adjacent, voir Figure 5.1b. La même situation s'observe pour le corail rouge (Figure 5.1c). Pour une espèce dispersant peu comme la gorgone pourpre (i.e. PLD de 1 jour), il en résulte que toutes ces populations très proches sont connectées entre elles pour un faible nombre de générations alors que les probabilités de connexions sont nulles entre les autres paires de populations. Sur la courbe représentant le coefficient de corrélation en fonction du nombre de générations, le pic observé à $\sim 10-20$ générations (Figure 5.3c) est induit par la situation « artificielle » où la proportion de paires de populations connectées par rapport au

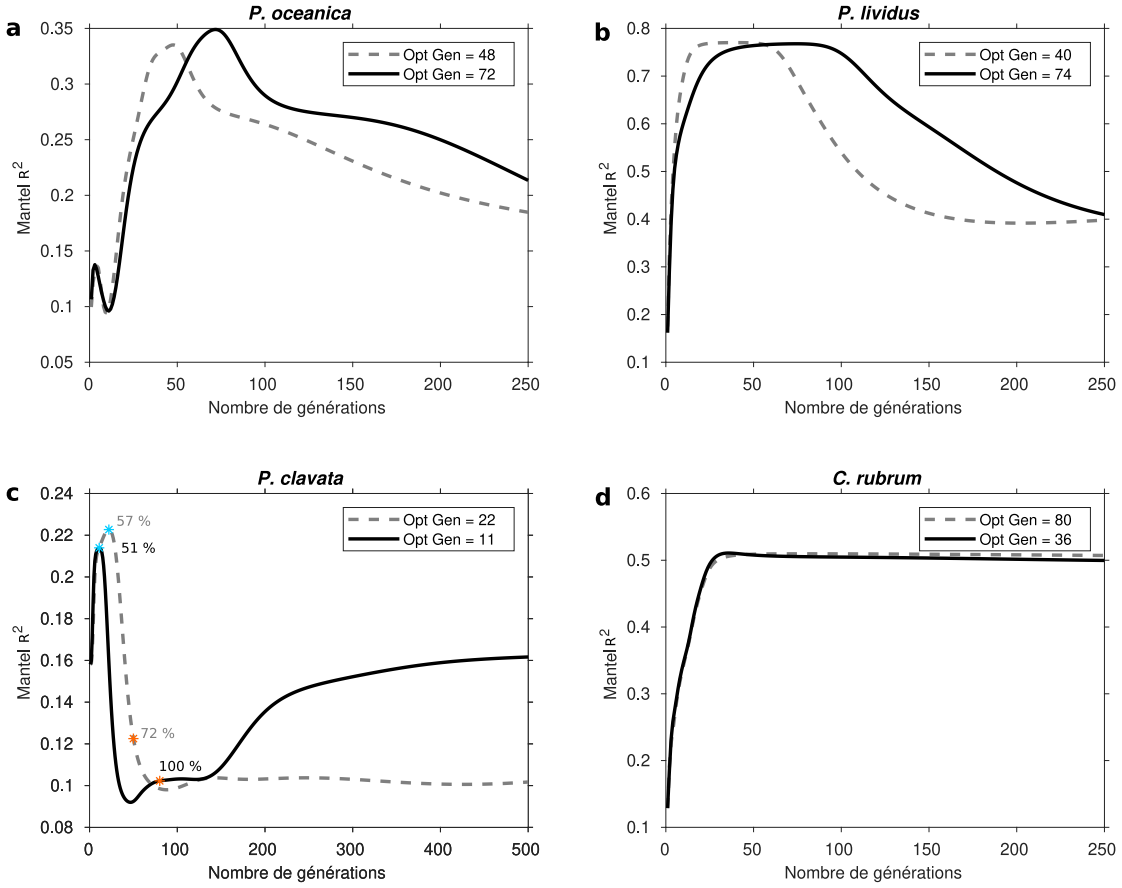


FIGURE 5.3. – Coefficient de corrélation de Mantel entre les Fst et les probabilités de connexion en fonction du nombre de générations considérées dans le modèle de dispersion multi-génération. Les courbes pleines et noires correspondent à l'habitat spécifique représentant l'ecosystème à herbier de posidonie (Giannoulaki et al., 2013) ou à coralligène (Martin et al., 2014). Les courbes pointillées et grises correspondent à l'habitat global. **a** La posidonie *Posidonia oceanica* (Arnaud-Haond et al., 2007). **b** L'oursin violet *Paracentrotus lividus* (Paterno et al., 2017). **c** La gorgone pourpre *Paramuricea clavata* (Mokhtar-Jamaï et al., 2011). Les pourcentages représentent la proportion de paires de populations connectées pour un nombre de générations fixé par rapport au nombre maximal potentiellement possible (ici 561 paires de populations). Les étoiles bleus correspondent au nombre de génération pour lequel le coefficient de corrélation de Mantel est maximal. Les étoiles orange indiquent le nombre de génération pour lequel le nombre de paires de population connectées est maximal (i.e. égale au nombre à 500 générations). **d** Le corail rouge *Corallium rubrum* (Aurelle et al., 2011).

nombre maximum potentiel est d'environ 50 %. Ainsi, les paires de populations voisines sont caractérisées par des probabilités de connexions proches de 1 et des Fst relativement faibles et de manière contraire, les paires de populations plus éloignées sont caractérisées par des probabilités de connexions nulles et des Fst relativement plus élevées. Toutefois, les prédictions de flux de gènes sont meilleures avec l'habitat spécifique qu'avec l'habitat global lorsque toutes les populations sont connectées entre elles pour l'habitat spécifique (pour environ 70 générations, étoile orange sur la Figure 5.3c) alors que ~ 30 % des paires de populations restent déconnectées pour l'habitat global, et ce jusqu'à 500 générations. Le modèle de niche, en définissant de manière plus fine les habitats de coralligène, notamment au niveau de la mer Adriatique, du canal de Sicile et des îles Baléares (voir Figure 4.1b du Chapitre IV par rapport à la Figure 5.1b), a permis de considérer des populations clefs dans le calcul de la connectivité génétique multi-génération. Sans la prise en compte de ces populations, notre modèle ne peut estimer le flux de gènes entre un tiers des paires de population échantillonnées (Mokhtar-Jamaï et al., 2011). La considération de la distribution spatiale est ainsi déterminante pour étudier les processus de connectivité génétique à large échelle. Pour les espèces caractérisées par une PLD très courte, le nombre de générations optimal est obtenu pour le plus grand nombre de générations modélisé (Figure 4.5b du Chapitre IV), similairement à la saturation des probabilités de dispersion multi-génération (Figure 9 du Chapitre III).

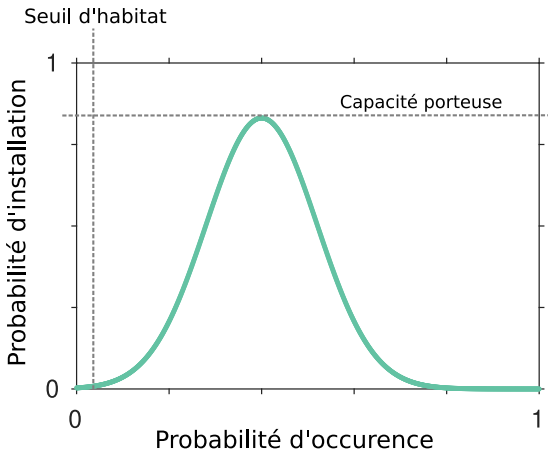


FIGURE 5.4. – Exemple d'intégration de la capacité porteuse dans la transformation des probabilité d'occurrence d'écosystème en probabilité d'installation

Une des raisons pour lesquelles les prédictions de flux de gènes ne sont pas significativement améliorées avec l'utilisation des habitats spécifiques réside dans la

prise en compte des probabilités d'occurrence et leur interprétation biologique dans le modèle de dispersion multi-génération. Nous avons suggéré que plus la probabilité d'occurrence de l'habitat considéré est grande, plus le processus de ponte et d'installation est efficace. En utilisant les mêmes données de distribution spatiale des herbiers de posidonie, Mari et al., 2020 appliquent cette même méthodologie pour caractériser la connectivité démographique de la posidonie sur l'ensemble des côtes méditerranéennes. De plus, ils définissent un seuil de probabilité d'occurrence pour définir un habitat viable ou non (Figure 5.4). Kool et al., 2010 ont également estimés la probabilité de ponte et d'installation en fonction du nombre de patchs d'habitats favorables (i.e. les récifs coralliens de la mer des Caraïbes) contenus dans les limites géographiques des nœuds de leur réseau de connectivité. Cette relation linéaire entre probabilité d'occurrence d'un habitat favorable et probabilité d'installation est adaptée pour des espèces à court cycle de vie, ou connues pour ne pas saturer leurs habitats (i.e. populations caractérisées par une forte mortalité, donc avec une dynamique démographique potentiellement dépendante du nombre de migrant, voir section 1.3.1), mais ne semble pas justifiée lorsque les populations ou patchs d'habitats sont limités par le nombre d'individus qu'ils peuvent accueillir (e.g. les herbiers marins). En d'autres termes, la prise en compte de la capacité porteuse (Del Monte-Luna et al., 2004) induirait qu'une population ne peut plus supporter d'autres immigrants lorsque un seuil de probabilité d'occurrence d'un habitat est franchi. Dans ce cas, la probabilité d'installation serait alors fortement réduite passé ce seuil défini par la capacité porteuse (Figure 5.4). A la génération suivante, les propagules des quelques migrants qui ont pu s'installer vont être diluées dans le nombre considérable de propagules relâchée par les populations autochtones, et vont avoir une très faible chance de contribuer au flux de gènes (i.e. *High-density-blocking*, voir *Box 1* dans Waters et al., 2013). De telles transformations de probabilités d'occurrence pourraient, en fonction de l'espèce considérée, apporter des solutions simples et modulables à l'implémentation de la dynamique spatiale des populations dans l'étude de la connectivité génétique.

6. Conclusions générales

6.1. Résumé et contextualisation

L'objectif de ce travail de thèse a été d'étudier l'influence de la connectivité multi-échelle via la dispersion sur la structure des populations et les schémas de biodiversité en mer Méditerranée. La modélisation des mécanismes de dispersion passive par les courant marins et la compréhension de la connectivité multi-échelle attenante ont nécessité l'utilisation couplée de modèles biophysiques paramétrés par les premiers traits de vie des espèces cibles, de données biogéographiques et de développements théoriques s'appuyant sur la théorie des graphes.

En premier lieu, l'étude des paramètres biologiques des premiers traits de vie, c'est-à-dire de la période de ponte et de la durée de la phase de dispersion pélagique, a rendu possible la paramétrisation du modèle biophysique. Dans cette thèse, ce sont des données *in situ* d'analyse d'otolithes de deux espèces de poissons côtiers pour un évènement de dispersion en mer Adriatique, et des compilations d'information bibliographiques sur près de cinquante espèces à cycle de vie biphasique à l'échelle de la mer Méditerranée. Ensuite, l'analyse biogéographique des espèces étudiées a permis de définir les lieux d'habitats des populations adultes via l'intégration d'informations spatiales sur les différents types de substrats, la bathymétrie, ou les prédictions d'occurrence d'espèce via des modèles de niches. Finalement, en modélisant le transport larvaire spécifique entre des sous-divisions spatiales des habitats favorables (i.e. des populations) et en utilisant les outils de la théorie des graphes, le modèle biophysique quantifie des $liens$ dans un réseau de $n\text{œ}uds$. Ces liens représentent alors des flux d'individus ou de gènes entre des populations, et caractérisent respectivement la connectivité démographique ou génétique.

Au-delà de la nécessité d'une recherche multidisciplinaire pour étudier ces deux aspects de la connectivité, on a, dans cette thèse, considéré de manière simultanée les échelles écologiques et évolutives spécifiques induites par le processus de dispersion. Ce processus influe, par définition (Clobert et al., 2012), simultanément sur

l'échange d'individus et de gènes dans l'espace. La dispersion impacte la dynamique spatio-temporelle démographique des populations, et dans le même temps, leurs structures génétiques. Cependant, il est difficile de faire un lien méthodologique entre ces deux mécanismes écologiques et évolutifs théoriquement distincts, notamment du fait des échelles temporelles typiques supposées différentes : que ce soit pour quantifier les échanges d'individus à partir de mesures de différentiation génétique entre populations ou de façon opposée, d'estimer le flux de gènes à partir de modèles biophysiques d'échanges d'individus. En définissant théoriquement les liens explicites (i.e enfants-parents) et implicites (i.e. partage d'un parent commun) entre populations sur des générations successives d'événements de dispersion cumulatifs, on a pu convertir la quantification de flux d'individus en estimation de flux de gènes (Chapitre III). Les liens implicites, qui n'étaient jusqu'à présent pas considérés dans l'étude de la connectivité via la dispersion, retournent les meilleures prédictions de flux de gènes lorsqu'elles sont comparées aux observations de structure génétique chez 47 espèces différentes sur l'ensemble du bassin méditerranéen (Chapitre IV). Ces meilleures prédictions de flux de gènes sont obtenues à un nombre de générations optimal qui est fonction de l'échelle spatiale considérée (i.e. étendu des patrons génétiques échantillonnés), liant ainsi les échelles temporelles et spatiales pertinentes pour évaluer la structure génétique. De plus, pour la méta-analyse, le nombre moyen de générations optimal est relativement faible (~ 20 générations) par rapport au nombre de générations nécessaires pour que les probabilités de connexions implicites saturent (~ 500 générations, Chapitre III), ce qui questionne le postulat d'un temps évolutif « long » nécessaire à l'établissement des structures génétiques observées. Cela suggère que le flux de gènes n'est pas à l'équilibre avec la dérive génétique à cause du caractère stochastique des forçages environnementaux et des rapides fluctuations écologiques biotiques. La variabilité à méso-échelle du processus de dispersion, en impactant directement la dynamique des populations, peut être un moteur de ces changements écologiques. Cette variabilité est due à des conditions hydrodynamiques locales changeantes (e.g. cycle saisonnier dans le transport des courants, fronts, *eddies*, événements sporadiques d'upwelling, etc.) mais également à des oscillations climatiques à large-échelle (oscillations décennales, exemple du BiOS, Di Stefano et al., 2022, en Annexe A.2). Pourtant, pour une génération de dispersion, la circulation océanique in situ, par les courants géostrophiques dominants, les mécanismes de convergence côtière amenant à de la rétention (e.g. downwelling) ou la turbulence, explique relativement bien les patrons de connectivité démographique (Chapitre II).

Dans le cas de l'estimation du flux de gènes, la prise en compte des liens implicites entre populations sur plusieurs générations montre que la connectivité génétique coalescente sous-jacente s'affranchit des barrières hydrodynamiques induites par les courants géostrophiques ou les gyres identifiés à large-échelle, contrairement à la connectivité filiale (Chapitre III). Il en résulte que les explications des structures génétiques observées par la circulation océanique moyenne doivent être revisitées. Même si la dispersion influe simultanément sur la démographie et la structure génétique des populations pour une échelle spatio-temporelle qui semble congruente, la connectivité démographique ne caractérise pas les même liens entre populations que la connectivité génétique. En effet, la connectivité génétique coalescente, en considérant des liens implicites que deux populations contemporaines partagent une population « ancêtre » commune, ne quantifie pas le flux d'individus mais le flux de gènes potentiel, et permet de mieux interpréter les patrons de diversité génétique observées. Finalement, la prise en compte, dans la modélisation de la connectivité coalescente, de la distribution spatiale d'une espèce (Chapitre V) permet de considérer de la variabilité spatiale de la taille de la population et potentiellement des phénomènes densité-dépendants, qui sont également directement influencé par la dispersion.

Pour conclure, cette thèse, en utilisant des outils et données complémentaires issus de différentes disciplines, a permis de cerner les lacunes et d'améliorer nos connaissances sur la relation entre connectivité démographique et génétique. Par son caractère éco-évolutif, le processus de dispersion influe simultanément sur la structure spatiale des populations et les patrons de diversité génétiques à des échelles spatio-temporelles qui semblent similaires, mais via des connexions intrinsèquement différentes. La définition de la connectivité coalescente ouvre de nombreuses perspectives quant à l'évaluation de l'impact relatif du flux de gènes par rapport aux autres forces évolutives sur la diversité génétique, et notamment du rôle potentiellement dominant de l'adaptation aux fluctuations environnementales, mais questionne également l'influence de la dynamique temporelle de la démographie locale sur de telles structures génétiques.

6.2. Quelles sont les perspectives apportées par cette thèse ?

6.2.1. De nouveaux outils méthodologiques

6.2.1.1. Le *cadre analytique*

Le *cadre analytique*, détaillé dans le chapitre II (c.f Figure 1 du Chapitre II), permet de quantifier les échelles de dispersion larvaire, de caractériser les patrons de connectivité démographique et de localiser les zones de pontes. L'intérêt premier de ce *cadre* est sa modularité en fonction de la précision des données biologiques à disposition sur l'espèce considérée. A partir d'une ou plusieurs simulations Lagrangiennes de dispersion larvaire sur l'intégralité de la zone d'étude, l'ajout de filtres successifs représentant des contraintes environnementales pour la ponte, les habitats préférentiels des adultes puis des données démographiques (densité et taille d'adultes matures pour chaque population ou taille de la population effective si l'on s'intéresse à la génétique des populations) permet théoriquement d'affiner, étape par étape, la précision et la pertinence des diagnostiques de connectivité. Techniquement parlant, les caractéristiques biologiques et démographiques sont considérées en modifiant certains éléments d'une matrice de connectivité *brute*, c'est-à-dire en supprimant des lignes ou colonnes (avant normalisation) et en modulant les probabilités de connexions (comme procédé par exemple dans Foster et al., 2012 ; Kool et al., 2010 ; Mari et al., 2020, Figure 6.1). Cette méthodologie probabiliste résumant l'information de millions de trajectoires apparait plus flexible qu'une approche centrée sur l'individu (i.e. se réfère ici à l'approche proposée par les modèles biophysiques comme Ichthyop, Lett et al., 2008 ou Connectivity Modeling System,Paris, Helgers et al., 2013, c.f. section 6.2.2.2), dans laquelle les caractéristiques biologiques et démographiques paramètrent directement l'implémentation des particules numériques. En effet, on peut étudier la variabilité spatio-temporelle des conditions biologiques et démographiques en comparant le résultat de plusieurs modulations conditionnelles appliquées sur les mêmes matrices $brutes$ (Figure 1.1, comme il a été fait dans le chapitre V). Une perspective intéressante serait de créer des banques de matrices de connectivité, représentant un événement précis de dispersion (Chapitre II et III) ou composites sur plusieurs années (Chapitre IV et V), pour pouvoir ensuite les moduler selon les questions scientifiques posées. Ce $cadre$ est utilisé dans le chapitre II pour étudier la variabilité spatio-temporelle du processus de ponte à l'échelle interannuelle,

dans le chapitre IV pour considérer les deux catégories d'habitats, les cinq périodes de pontes et les cinq PLDs différentes pour les 47 espèces de la méta-analyse et dans le chapitre V pour moduler les probabilités de connexion en fonction de la probabilité d'occurrence d'herbiers de posidonie et de coralligène dans chaque population.

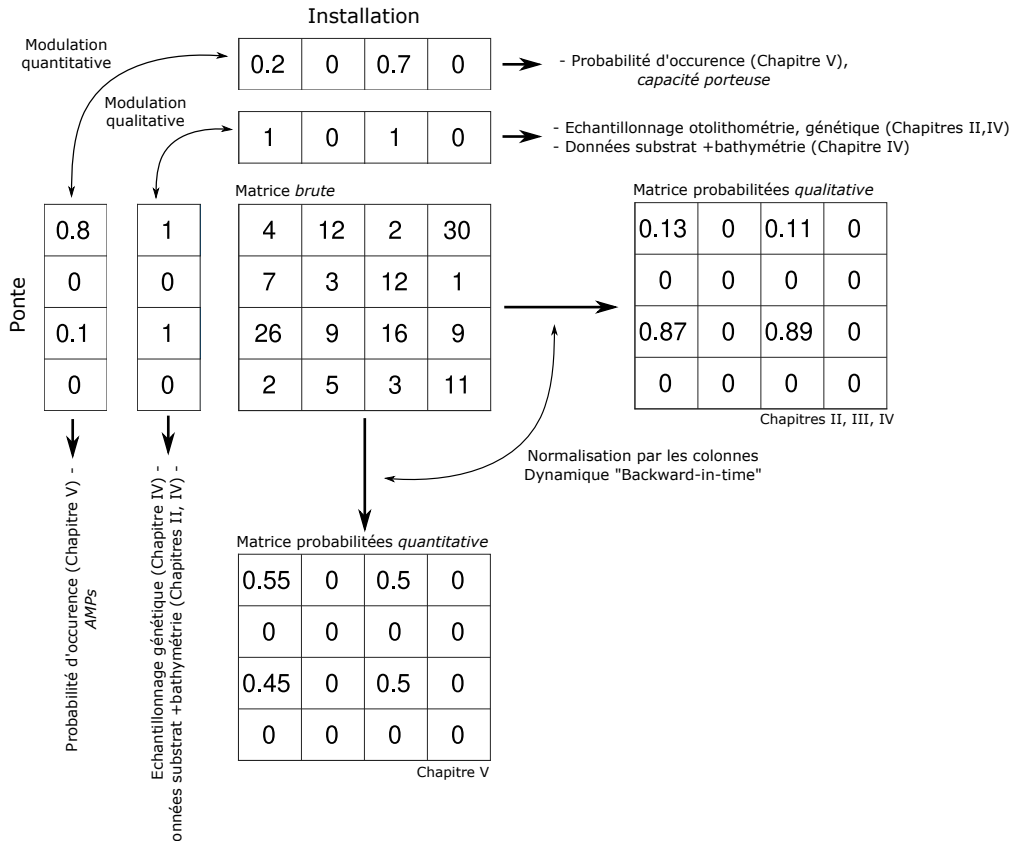


FIGURE 6.1. – Schéma simplifié du *cadre analytique* général et de son utilisation dans cette thèse. La matrice brute est le produit du modèle bio-physique LFN, où chaque élément m_{ij} correspond au nombre de particules échangées entre les populations i et j. Les caractéristiques biologiques et démographiques (habitat préférentiel et site d'échantillonnage) sont considérées en modulant la matrice $brute$ de manière qualitative (Chapitre II, III, et IV) ou quantitative (V). Les perspectives de modulation quantitatives sont en italiques, et concernent le processus de ponte (considération de la possible production plus importante de larves dans les AMPs) et d'installation (considération de la capacité d'accueil).

6.2.1.2. La connectivité filiale et coalescente via la dispersion multi-génération pour caractériser le flux de gènes

La formulation théorique des connexions explicites et implicites multi-générations cumulatives décrites dans le Chapitre III est un outil méthodologique novateur pouvant être appliqué à n'importe quel réseau temporel et orienté. Les connexions explicites et implicites sont appliquées dans le Chapitre IV pour modéliser respectivement la connectivité filiale et coalescente induite par la dispersion sur de multiples générations chez des espèces marines de différents phylums. Grace à cette méta-analyse, nous avons montré que la connectivité coalescente est le meilleur modèle pour estimer de manière mécanistique le flux de gènes inhérent au processus de dispersion par les courants marins. Les connexions implicites multi-générations cumulatives représentent ainsi une contribution importante à la discipline récente du *seascape genetics*, soit l'étude conjointe des patrons génétiques identifiés par la génétique des populations, de l'environnement marin et des conditions écologiques et démographiques (c.f. section 1.3.4, Jahnke et Jonsson, 2022 ; Riginos et Liggins, 2013 ; Selkoe et al., 2016 ; Selkoe et al., 2008). L'estimation du flux de gènes par un modèle mécanistique ouvre de nouvelles perspectives quant à la dissociation des influences relatives de la sélection environnementale et du flux de gènes sur les structures génétiques observées (e.g.Wang et al., 2013). La connectivité filiale pourraient être appliquées pour étudier les processus de colonisation sur plusieurs générations de dispersion successives. Théoriquement, les connexions explicites permettent d'évaluer comment une espèce étend son aire de répartition par évènement de dispersion successif. D'un point de vue démographique, les probabilités de connexion filiale sur plusieurs générations peuvent être comparées à l'historique des observations de présence d'une espèce invasive. Par exemple, cela peut être appliquée chez les espèce invasives en Méditerranée après un évènement de migration lessepsienne depuis la mer Rouge, et qui colonisent de génération en génération les côtes orientales puis occidentales (e.g. *Fistularia commersonii*, Azzurro et al., 2013).

6.2.2. La modélisation Lagrangienne et le comportement larvaire

6.2.2.1. Le comportement larvaire

Certaines espèces, notamment chez les poissons côtiers (Leis, 2006) mais également les invertébrés marins (Chia et al., 2011), présentent des comportements larvaires

actifs : cela correspond à la capacité des larves à se mouvoir en dehors de la dérive induite par les courants océaniques. De telles capacités natatoires se développent au cours de l'ontogénèse : pour les poissons, il y a d'abord le stade de flexion (courbure de la colonne vertébrale), puis le développement complet des vertèbres, des nageoires caudales et enfin des nageoires dorsales (Clark et al., 2005). Le potentiel maximal de mouvement endogène est ainsi obtenue à la fin de la phase pélagique. Pour quantifier le comportement larvaire, la majorité des études sont menées en laboratoire dans des aquariums simulant les contraintes des conditions pélagiques. La vitesse critique des larves (i.e. un proxy de la vitesse maximale atteinte sous incrémentation successives des courants) chez des poissons tempérés peut aller de 10 $cm.s^{-1}$ à 30 $cm.s^{-1}$ à la fin de leur développement (Figure 6.2a, Baptista et al., 2019 ; Clark et al., 2005 ; Faillettaz, Durand et al., 2018 ; Leis, 2006). Sous la contrainte de courants de 10 $cm.s^{-1}$ (i.e. caractéristique d'un courant moyen), les larves peuvent nager à la fin de la phase pélagique pendant une dizaine de km environ (Figure 6.2b, Baptista et al., 2019 ; Clark et al., 2005). Dans le cas d'un cisaillement vertical de la circulation océanique, cette capacité de mouvement accroit la rétention locale en permettant aux larves de ne pas être exporté au large en nageant verticalement entre les couches de surface (Ayata et al., 2010 ; Leis, 2006 ; Sponaugle et al., 2002). Cela s'observe notamment pendant les cycles de marées, dans les estuaires ou les upwellings (Menge et al., 2011 ; Queiroga et Blanton, 2005 ; Thiébaut et al., 1992). Il a été montré *in situ* que les larves répondent à des signaux olfactifs envoyés par leurs substrats préférentiels (ici les récifs de coraux) en nageant plus rapidement vers la direction voulue (Paris, Atema et al., 2013). A la fin de la phase pélagique, les larves de plusieurs espèces de poissons côtiers de Méditerranée peuvent également tenir un cap dans leurs mouvements, cap qui peut être indexé par la position du soleil (Berenshtein et al., 2014 ; Faillettaz et al., 2015) ou du champ magnétique (Bottesch et al., 2016). Ces comportements natatoires et d'orientation permettent aux larves d'augmenter leurs chances de s'installer sur un substrat adéquat en fin de phase pélagique (Faillettaz, Paris et al., 2018).

6.2.2.2. L'approche de modélisation Lagrangienne centrée sur l'individu

Plusieurs modèles Lagrangiens (e.g. Ichthyop, Lett et al., 2008, Connectivity Modeling System, Paris, Helgers et al., 2013) adoptent une approche centrée sur l'individu (*individual-based model*) pour rendre compte du comportement larvaire pendant le processus de transport. Ils considèrent à chaque pas de temps de l'intégration des trajectoires des particules, la vitesse horizontale (e.g. Faillettaz, Durand et al., 2018) ou

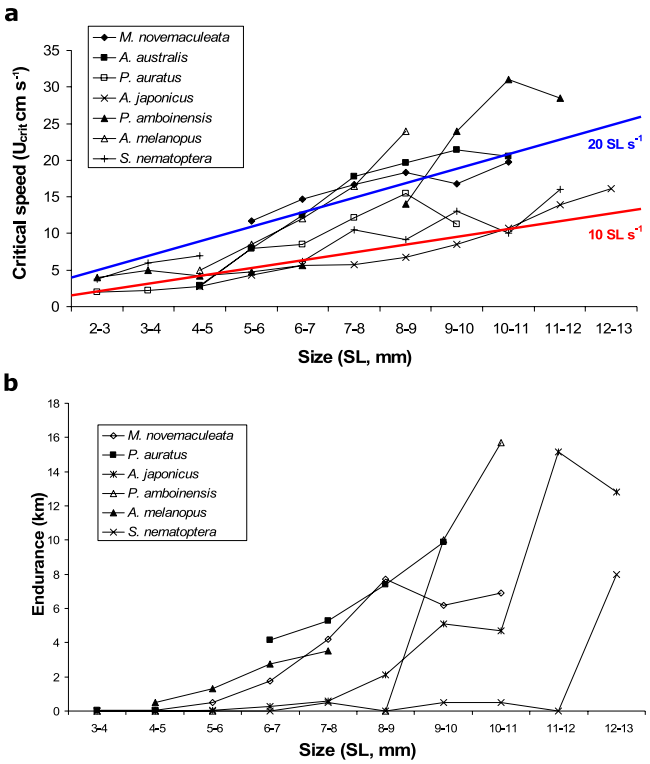


FIGURE 6.2. – Evolution des capacités natatoires durant l'ontogenèse pour trois espèces tropicales (*A. melanopus, P. amboinensis, S. nematoptera*) et trois espèces tempérées (*M. novemaculeata, P. auratus, A. japonicus*) évaluée en laboratoire. **a** Vitesse critique de nage (cm.s^{-1}), soit la vitesse maximale atteinte sous incrémentation successives des courants. Les courbes rouges et vertes indiquent les vitesses critiques de nage par unité de taille par seconde (SL, *standard length*), soit respectivement 10 SL.s^{-1} et 20 SL.s^{-1}. A la taille d'installation (10-12 mm SL), la vitesse critique moyenne est d'environ 15 cm.s^{-1}. **b** Endurance de nage sous la contrainte de courant de 10 cm.s^{-1} . A la taille d'installation, l'endurance et de l'ordre d'une à deux dizaines de km. D'après Clark et al., 2005 ; Leis, 2006

.

verticale (e.g. Ayata et al., 2010 ; Guizien et al., 2020) de nage potentielle de l'espèce étudiée. La température observée pendant la trajectoire de chaque propagule numérique peut être considéré dans le taux de croissance supposé des larves (revue dans Lett et al., 2010), et peut également impacter la mortalité sous-jacente (Kool et al., 2010). De tels modèles peuvent possiblement mettre fin à la simulation du processus de transport lorsque la propagule se situe sur un habitat favorable après une période fixée de dispersion, qui correspond au temps nécessaire à la larve de se développer suffisamment pour pouvoir s'installer (Faillettaz, Paris et al., 2018, voir Ayata et al., 2009 avec l'utilisation d'un modèle Eulérien). En suivant individuellement chaque propagule numérique, cette approche centrée sur l'individu est évidemment plus couteuse en temps et puissance de calcul. Elle nécessite des connaissances biologiques, le plus souvent inexistantes, sur le comportement larvaire pendant la dispersion : taux de croissance en fonction de l'environnement, développement des capacité de mouvement pendant l'ontogénie, capacité d'orientation, schéma de migration verticale (Ayata et al., 2010 ; Guizien et al., 2020), etc. De plus, conditionner *in silico* la croissance et la mortalité des propagules par des facteurs abiotiques demande de modéliser avec précision les variations environnementales à sous-méso-échelle, ce qui n'est pas encore le cas de nos jours.

6.2.2.3. Perspectives d'amélioration du modèle

Dans cette thèse, nous avons utilisé le modèle Lagrangian Flow Network (Dubois et al., 2016 ; Rossi et al., 2014 ; Ser-Giacomi, Rossi et al., 2015) pour modéliser le transport de propagules par les courants. Nous avons choisi de ne pas intégrer de comportement larvaire dans le calcul des trajectoires de chaque particule. L'hypothèse sous-jacente est que la simplification faite en négligeant le comportement larvaire est du même ordre de grandeur que la diffusion numérique induite par la définition spatiale large des nœuds du modèle (i.e. l'emplacement final des propagules numériques au sein d'un nœud n'a pas d'importance). En effet, la taille de coté de ces nœuds, qui est de 1/16° (soit ~ 7 km) dans le chapitre II et de 1/4° (soit ~ 25 km) dans le chapitre IV et V, et les capacité moyennes d'endurance observées en laboratoire sont d'une à deux dizaines de km (Figure 6.2b, voir également discussion dans la section 4.2 du Chapitre II).

Comme proposé par les modèles centrés sur l'individu, de nombreuses perspectives existent quant à une meilleure intégration du comportement larvaire dans la modélisation Lagrangienne proposée dans cette thèse. Une solution idéale serait d'im-

plémenter les considérations biologiques de comportement des modèles centrés sur l'individu (Lett et al., 2008 ; Paris, Helgers et al., 2013) sur la routine de modélisation Lagrangienne du modèle LFN. On pourrait ainsi bénéficier des avantages de ce modèle, soit l'approche probabilistique et les pre- et post-traitement des matrices de connectivité facilitant la modulation et l'utilisation des outils de la théorie des graphes, tout en considérant les comportements larvaires. Toutefois, lorsqu'il est question d'étudier la connectivité génétique, il faut prendre en considération les processus physiques résolus par les modèles Eulériens à l'échelle du bassin (voir discussion Pineda et al., 2007). Pour les modèles à grande échelle, avec des simulations disponibles sur plusieurs décennies, les incertitudes sur les processus à méso-échelle (e.g. les ondes de gravité, les ondes de surface et interne de marée, les vagues, le cisaillement aux frontières, les vitesses et la stratification verticales, etc.) sont comparables voire supérieures aux effets potentiels des capacités natatoires des larves (le modèle Eulérien utilisé dans le Chapitre III, IV et V produit des champs de courants horizontaux à une précision d'1/16°).

En plus de la diffusion numérique induite par la taille des nœuds et si l'on considère la possible variabilité temporelle induite pour modéliser la connectivité génétique, implémenter un sous-modèle de marche aléatoire pendant l'intégration des trajectoires permettrait de prendre en compte les processus stochastiques de turbulence (e.g. Ayata et al., 2010) et de nage (e.g. Paris, Atema et al., 2013) à la fois sur l'horizontale et la verticale.

Il serait intéressant d'implémenter dans notre modèle biophysique la considération des conditions environnementales (notamment la température) observées par les particules pendant l'intégration de leurs trajectoires. De ce fait, en plus de modéliser les probabilités de ponte et d'installation en fonction de l'habitat, on pourrait également prendre en compte la mortalité des propagules, ainsi que leurs pré-compétence pour s'installer dans un habitat favorable pendant le processus de transport en fonction des conditions environnementales (Ayata et al., 2009 ; Faillettaz, Paris et al., 2018). La prise en compte de la variabilité spatio-temporelle de l'environnement pourrait également être indexée pour chaque population (Di Stefano et al., 2022, voir annexe A.2), afin d'étudier l'influence relative de la sélection environnementale sur les structures génétiques.

6.2.3. Délimitation spatiale de communautés dans les réseaux de connectivité génétique

Pour étudier l'influence de l'environnement sur la diversité génétique, les méthodes employées se basent généralement sur la comparaison de mesures de différentiations génétiques (e.g. Fst) avec diverses mesures de distances ou différences entre paires de populations : par exemple l'isolement par la distance (Rousset, 1997), l'isolement par l'environnement (Wang et Bradburd, 2014), l'isolement par la résistance (McRae, 2006), ou l'isolement par l'océanographie (terme qui regroupe l'évaluation par des modèles biophysiques des connexions entre populations via la dispersion passive des propagules par les courants, Jahnke et Jonsson, 2022). Pour le dernier cas, on a effectivement comparé dans le Chapitre IV des matrices de Fst avec des matrices de distances entre paires de populations (isolement par la distance et par l'océanographie). Cependant, la considération de quelques paires de populations directement indexées par l'échantillonnage, peut être restrictif dans la compréhension des structures spatiales à l'échelle globale du bassin étudié. En effet, en utilisant l'exemple de la posidonie, échantillonnée de manière conséquente sur les côtes Méditerranéennes (29 sites d'échantillonnages, Arnaud-Haond et al., 2007), on a testé dans le Chapitre IV et V la corrélation entre Fst et probabilité de connexion pour 406 paires de populations, soit moins de 1% du nombre total de connexions du réseau de connectivité coalescent sous-jacent (i.e. 683 865 et 611 065 paires de populations pour respectivement l'habitat global et l'habitat spécifique, Chapitre V). Au-delà de l'indexation des paires de populations contrainte par l'échantillonnage, chaque population peut être caractérisée par sa propre importance au sein du réseau de connectivité (quantifiée par des métriques issues de la théorie des graphes, e.g. *local retention*, *self recruitment*, *source-sink degree*, *betwenness centrality* et *eigenvector centrality*, Andrello et al., 2017 ; Dubois et al., 2016 ; Ser-Giacomi, Rossi et al., 2015 ; Treml et al., 2008). Parallèlement, l'étude des caractéristiques des réseaux de connectivité permet d'identifier des communautés, rassemblant les populations qui sont bien connectées entre elles et peu avec leurs voisines (e.g. Jacobi et al., 2012 ; Rossi et al., 2020 ; Rossi et al., 2014 ; Rosvall et Bergstrom, 2008 ; Ser-Giacomi, Rossi et al., 2015). De telles communautés peuvent correspondre à l'identification de sous-populations (Jacobi et al., 2012) et permettent de localiser les possibles barrières à la dispersion (De Wit et al., 2020 ; Jahnke et al., 2018), indépendamment de la contrainte de l'échantillonnage. Dans le cas de la connectivité coalescente, identifier des communautés dans un tel réseau de liens implicites reviendrait (i) à grouper ensemble les populations qui partagent les

mêmes populations ancêtres et (ii) à délimiter des régions à l'intérieur desquelles le flux de gènes induit par la dispersion peut être considéré comme relativement homogène mais distinct du flux de la région voisine. Les frontières entre communautés représentent ainsi des barrières au flux de gènes inhérentes au réseau de connectivité coalescent, c'est-à-dire qui représentent à la fois l'entièreté des liens implicites entre chaque population et l'hétérogénéité spatiale de l'habitat. La considération de telles « méta-populations coalescentes » (voir Figure 6.3 pour l'exemple de la posidonie) permettrait de s'émanciper de l'analyse comparative entre populations et d'optimiser la stratégie d'échantillonnage en répartissant les sites d'échantillonnage au sein des différentes régions. Ces méta-populations coalescentes peuvent faciliter l'interprétation des clusters génétiques, et de leurs interpolations spatiales (e.g. De Wit et al., 2020 ; Jahnke et al., 2018). Elles permettent également, en estimant la structure génétique spatiale et en définissant des unités de conservation à l'échelle d'un bassin, de fournir des informations cruciales pour la gestion des populations marines, alors que le génotypage peut être couteux pour couvrir de telles échelles spatiales.

6.2.4. Direction future et recommandations

Les résultats dévelopés dans cette thèse ont été obtenus en réutilisant des données biologiques et génétiques déjà publiées, et en employant des données biogéographiques et hydrodynamiques disponibles sur les portails européens (EMODnet et Copernicus respectivement). D'une part, la réutilisation de données à de nouvelles hypothèses de travail à de nombreux avantages : réduction des coûts de recherche, approfondissement des questionnements scientifiques, vision d'ensemble grâce aux méta-analyse, etc. D'autre part, la stratégie d'échantillonnage proposée pour répondre aux problématiques initiales n'est pas forcément adaptée à d'autres hypothèses de recherche : nos résultats ont été en quelque sorte contraints par le choix antérieur des espèces cibles et des stratégies d'échantillonnage. Inversement, nos résultats permettent d'offrir des pistes de réflexions pour mieux étudier les processus de connectivité multi-échelle :

— Selon le Chapitre IV, un échantillonnage de plus de 11 populations est nécessaire pour avoir au moins 50 % de prédiction significative de flux de gènes par le modèle de connectivité coalescente. De plus, un échantillonnage combinant fine et grande échelle est inadapté pour l'évaluation de la connectivité génétique (Chapitre V). Pour de prochaines études de connectivité génétique se plaçant à l'échelle d'un bassin, nos résultats suggèrent un échantillonnage homogène et

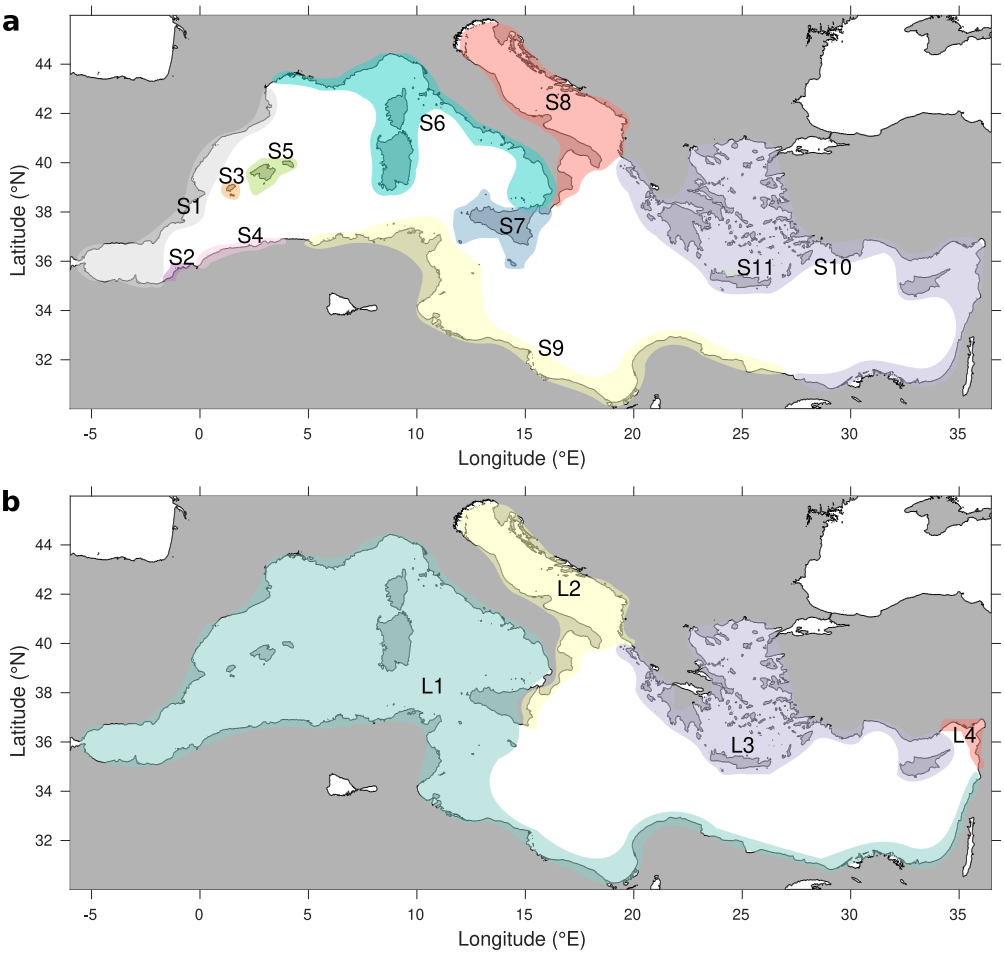


FIGURE 6.3. – Délimitation de « méta-populations » dans le réseau de connectivité coalescente de la posidonie par l'algorithme InfoMap (Rosvall et Bergstrom, 2008). Le réseau de connectivité coalescente a été modélisé en considérant les probabilités d'occurrence des herbiers de posidonie détaillée dans le Chapitre V. Au sein de chaque région, le flux de gènes modélisé peut être considéré comme homogène. **a** 11 communautés sont détectées pour une courte période de dispersion (i.e. 1 jour, connectivité coalescente calculée pour 600 générations), caractéristique du transport des pollens et des graines par les courant (McMahon et al., 2014, c.f. Figure 1.13). **b** 4 communautés sont détectées pour une longue période de dispersion (i.e. 30 jours, connectivité coalescente calculée pour 72 générations), caractéristique du transport des fruits par les courant (McMahon et al., 2014).

conséquent (plus de 20 populations échantillonnées) le long des côtes du domaine étudié. La répartition spatiale des sites d'échantillonnage peut être orientée par la délimitation préliminaire des méta-populations coalescentes donnée par nos modélisations biophysiques. Les sites d'échantillonnages doivent être espacés à minima d'un ordre de grandeur supérieur à la précision des modèles hydrodynamiques grande échelle, soit plusieurs dizaines à une centaine de km. En effet, la précision horizontale des modèles hydrodynamiques opérationnels, fournissant des champs de courants journaliers à l'échelle d'un bassin pour une période d'une trentaine d'années est actuellement d'environ 4 km pour la Baltique-Mer du Nord [1] et la mer Noire [2] , 7 km pour la mer Méditerranée [3] et 10 km pour les côtes Atlantiques Nord [4] et Sud [5]. Les processus physiques côtiers caractérisés par quelques dizaines de km et moins ne sont ainsi pas résolus, et ne permettent pas de différencier des données, par exemple de différentiation génétique, échantillonnées à cette échelle. Pour les études de génétique des population combinant un échantillonnage à sous-méso-échelle (sites séparées de quelques km) et à large-échelle (sites séparées de plusieurs centaines de km, e.g. Aurelle et al., 2011 ; Mokhtar-Jamaï et al., 2011), il serait intéressant de pouvoir regrouper les populations proches (i.e. distantes de moins d'une dizaine km) en une unique population lors de l'application de modèle biophysiques à large-échelle. A noter que les champs de courants issus de données altimétriques possèdent une précision moindre (~ 1/4° , soit environ 25 km) et résolvent mal les processus physiques côtiers. L'utilisation de produits altimétriques est ainsi réservée à l'étude du transport « transcontinental » par les courant marins (e.g. sargasse, Beron-Vera et al., 2015).

— L'étude de la connectivité génétique à sous-méso-échelle serait idéale dans des milieux fermés ou semi-fermés, comme les étangs ou les lagons En d'autres termes, un domaine où les populations présentes sont constitutives d'un unique réseau de connectivité, et ne sont pas seulement un sous échantillonnage d'un réseau global. Par exemple, pour l'étude précise de la baie de Marseille (Rossi et al., 2020, où la résolution horizontale du modèle hydrodynamique est de 400 m), les populations échantillonnées dans cette zone (~ 10 km, e.g. Aurelle et al.,

1. https://doi.org/10.48670/moi-00013
2. https://doi.org/10.25423/CMCC/BLKSEA$_{M}ULTIYEAR_{P}HY_{0}07_{0}04$
3. https://doi.org/10.25423/CMCC/MEDSEA$_{M}ULTIYEAR_{P}HY_{0}06_{0}04_{E}3R1$
4. https://doi.org/10.48670/moi-00059
5. https://doi.org/10.48670/moi-00029

2011 ; Mokhtar-Jamaï et al., 2011) sont aussi connectées avec des populations qui ne sont pas prises en compte au sein du petit domaine modélisé, ce qui peut nuire à la compréhension des patrons spatiaux génétiques identifiés. De plus, la modélisation de la circulation hydrodynamique à haute-résolution est facilitée dans des milieux fermés (e.g. absence de considération de la circulation globale, simplification des hypothèses hydrodynamiques, données plus précises des forçage environnementaux). Même si c'est une technologie couteuse en argent et entretien, il serait envisageable d'obtenir des observations des courants de surface à haute résolution spatio-temporelle grâce aux radars hautes fréquences combiné à des méthodologies de réseaux de neurones (Hernández-Carrasco et al., 2018).

— Pour étudier la variabilité temporelle des patrons de différentiation génétique, il serait intéressant de procéder à un échantillonnage à l'échelle du bassin considéré (voir premier point) sur plusieurs années (e.g Jackson et al., 2018 ; Pascual et al., 2016 ; Pérez-Portela et al., 2019). On pourrait ainsi comparer l'échelle temporelle induite par notre modèle de connectivité coalescente (nombre de génération optimale) et celle induite par l'observation de patrons de structure génétique entre des populations échantillonnée à un intervalle de quelques années (Pascual et al., 2017 ; Pérez-Portela et al., 2019).

— Pour une approche de *seascape genetics* utilisant un modèle biophysique, le choix d'une espèce cible doit être conditionnée, à un degré moindre que son statut de conservation, par rapport (i) à l'état de ses connaissances biologiques, notamment sur les premiers traits de vie (e.g. période de ponte, facteurs biotiques ou abiotiques déclenchant la ponte, comportement larvaire pendant la dispersion) mais également à l'état adulte (e.g. habitat préférentiel, déplacement des juvéniles, e.g. Di Franco et al., 2015, déplacement vers des zones de ponte) et (ii) aux données disponibles sur sa distribution spatiale (i.e. cartographie par observation à petite échelle, e.g. Rossi et al., 2020 ou par modèle de niche corrélatif à grande échelle, Chapitre V).

— En plus de l'utilisation de données de distribution spatiale spécifique, la considération de la capacité porteuse de chaque population, ou de la saturation des patchs d'habitats, dans les modèles biophysiques de connectivité génétiques semble primordiale pour l'amélioration des estimations de flux de gènes. De plus, cela permettrait d'étudier théoriquement l'impact des phénomènes densité-dépendants sur le flux de gènes et les structures génétiques sous-jacentes avec

un modèle mécanistique de dispersion (e.g. *high-density blocking*, *gene surfing*, etc., Waters et al., 2013).

En prenant tous ces points en considération, les espèces constituants les herbiers marins seraient des espèces cibles adéquates dans la mise en place d'un projet futur, qui aurait pour objectif de comprendre l'impact relatif du flux de gènes par rapport aux autres forces évolutives. Elles ne présentent pas de comportement larvaire (transport passif des fruits, des graines ou des rameaux décrochés par les courants de surface), et leurs distributions spatiales sont bien cartographiées (e.g. Jahnke et al., 2018 ; Telesca et al., 2015), ce qui permet de développer des modèles de niches corrélatifs pour définir des probabilités d'occurrence sur l'ensemble du domaine considéré, quand ce n'est pas déjà fait (Giannoulaki et al., 2013). L'étude de la connectivité génétique à grande échelle pourrait se concentrer sur la zostère *Zostera marina* en mer Baltique où des produits réanalysés de champs de courant horaires à la surface sont disponibles à une résolution d'environ 4 km. De plus, mis à part la Russie qui occupe un espace côtier restreint, l'ensemble des pays riverains de la mer Baltique appartiennent à l'Union Européenne, ce qui faciliterait les démarches pour un échantillonnage homogène et pluriannuel, contrairement à la Méditerranée, où un tel échantillonnage de la posidonie, du bassin occidental à l'oriental et du Nord au Sud, est irréalisable à ce jour et au vu des tensions géopolitiques.

Bibliographie

Alberto, F., Massa, S., Manent, P., Diaz-Almela, E., Arnaud-Haond, S., Duarte, C. M. & Serrão, E. A. (2008). Genetic differentiation and secondary contact zone in the seagrass Cymodocea nodosa across the Mediterranean–Atlantic transition region. Journal of Biogeography, 35(7), 1279-1294. https://doi.org/10.1111/j.1365-2699.2007.01876.x

Alberto, F., Raimondi, P. T., Reed, D. C., Watson, J. R., Siegel, D. A., Mitarai, S., Coelho, N. & Serrão, E. A. (2011). Isolation by oceanographic distance explains genetic structure for Macrocystis pyrifera in the Santa Barbara Channel. Molecular Ecology, 20(12), 2543-2554. https://doi.org/10.1111/j.1365-294X.2011.05117.x

Allen, J. D. & Pernet, B. (2007). Intermediate modes of larval development : bridging the gap between planktotrophy and lecithotrophy. Evolution & Development, 9(6), 643-653. https://doi.org/10.1111/j.1525-142X.2007.00202.x

Almany, G. R., Berumen, M. L., Thorrold, S. R., Planes, S. & Jones, G. P. (2007). Local Replenishment of Coral Reef Fish Populations in a Marine Reserve [Publisher : American Association for the Advancement of Science]. Science. https://doi.org/10.1126/science.1140597

Álvarez-Noriega, M., Burgess, S. C., Byers, J. E., Pringle, J. M., Wares, J. P. & Marshall, D. J. (2020). Global biogeography of marine dispersal potential. Nat Ecol Evol, 4(9), 1196-1203. https://doi.org/10.1038/s41559-020-1238-y

Andréfouët, S. & Bionaz, O. (2021). Lessons from a global remote sensing mapping project. A review of the impact of the Millennium Coral Reef Mapping Project for science and management. Science of The Total Environment, 776, 145987. https://doi.org/10.1016/j.scitotenv.2021.145987

Andrello, M., Guilhaumon, F., Albouy, C., Parravicini, V., Scholtens, J., Verley, P., Barange, M., Sumaila, U. R., Manel, S. & Mouillot, D. (2017). Global mismatch between fishing dependency and larval supply from marine reserves. Nature Communications, 8, 16039. https://doi.org/10.1038/ncomms16039

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van De Vliet, M. S., Alberto, F., Procaccini, G., Duarte, C. M. & Serrão, E. A. (2007). Vicariance patterns in the Mediterranean Sea : east–west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. Journal of Biogeography, 34(6), 963-976. https://doi.org/10.1111/j.1365-2699.2006.01671.x

Aurelle, D., Ledoux, J.-B., Rocher, C., Borsa, P., Chenuil, A. & Féral, J.-P. (2011). Phylogeography of the red coral (Corallium rubrum) : inferences on the evolutionary history of a temperate gorgonian. Genetica, 139(7), 855-869. https://doi.org/10.1007/s10709-011-9589-6

Ayata, S.-D., Ellien, C., Dumas, F., Dubois, S. & Thiébaut, É. (2009). Modelling larval dispersal and settlement of the reef-building polychaete Sabellaria alveolata : Role of hydroclimatic processes on the sustainability of biogenic reefs. Continental Shelf Research, 29(13), 1605-1623. https://doi.org/10.1016/j.csr.2009.05.002

Ayata, S.-D., Lazure, P. & Thiébaut, É. (2010). How does the connectivity between populations mediate range limits of marine invertebrates? A case study of larval dispersal between the Bay of Biscay and the English Channel (North-East Atlantic). Progress in Oceanography, 87(1), 18-36. https://doi.org/10.1016/j.pocean.2010.09.022

Azzurro, E., Soto, S., Garofalo, G. & Maynou, F. (2013). Fistularia commersonii in the Mediterranean Sea : invasion history and distribution modeling based on presence-only records. Biol Invasions, 15(5), 977-990. https://doi.org/10.1007/s10530-012-0344-4

Baptista, V., Morais, P., Cruz, J., Castanho, S., Ribeiro, L., Pousão-Ferreira, P., Leitão, F., Wolanski, E. & Teodósio, M. A. (2019). Swimming Abilities of Temperate Pelagic Fish Larvae Prove that They May Control Their Dispersion in Coastal Areas. Diversity, 11(10), 185. https://doi.org/10.3390/d11100185

Bell, J. R., Bohan, D. A., Shaw, E. M. & Weyman, G. S. (2005). Ballooning dispersal using silk : world fauna, phylogenies, genetics and models [Publisher : Cambridge University Press]. Bulletin of Entomological Research, 95(2), 69-114. https://doi.org/10.1079/BER2004350

Benestan, L., Fietz, K., Loiseau, N., Guerin, P. E., Trofimenko, E., Rühs, S., Schmidt, C., Rath, W., Biastoch, A., Pérez-Ruzafa, A., Baixauli, P., Forcada, A., Arcas, E., Lenfant, P., Mallol, S., Goñi, R., Velez, L., Höppner, M., Kininmonth, S., … Manel, S. (2021). Restricted dispersal in a sea of gene flow [Publisher : Royal Society].

Proceedings of the Royal Society B : Biological Sciences, 288(1951), 20210458. https://doi.org/10.1098/rspb.2021.0458

Berenshtein, I., Kiflawi, M., Shashar, N., Wieler, U., Agiv, H. & Paris, C. B. (2014). Polarized Light Sensitivity and Orientation in Coral Reef Fish Post-Larvae [Publisher : Public Library of Science]. PLOS ONE, 9(2), e88468. https://doi.org/10.1371/journal.pone.0088468

Berline, L., Ody, A., Jouanno, J., Chevalier, C., André, J.-M., Thibaut, T. & Ménard, F. (2020). Hindcasting the 2017 dispersal of Sargassum algae in the Tropical North Atlantic. Marine Pollution Bulletin, 158, 111431. https://doi.org/10.1016/j.marpolbul.2020.111431

Beron-Vera, F. J., Olascoaga, M. J., Haller, G., Farazmand, M., Triñanes, J. & Wang, Y. (2015). Dissipative inertial transport patterns near coherent Lagrangian eddies in the ocean. Chaos, 25(8), 087412. https://doi.org/10.1063/1.4928693

Bierne, N., Welch, J., Loire, E., Bonhomme, F. & David, P. (2011). The coupling hypothesis : why genome scans may fail to map local adaptation genes. Molecular Ecology, 20(10), 2044-2072. https://doi.org/10.1111/j.1365-294X.2011.05080.x

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., Moyes, F., Magurran, A., McGill, B., Myers-Smith, I. H., Winter, M., Bjorkman, A. D., Bowler, D. E., Byrnes, J. E. K., Gonzalez, A., Hines, J., Isbell, F., Jones, H. P., Navarro, L. M., ... Dornelas, M. (2019). The geography of biodiversity change in marine and terrestrial assemblages [Publisher : American Association for the Advancement of Science Section : Research Article]. Science, 366(6463), 339-345. https://doi.org/10.1126/science.aaw1620

Bode, M., Leis, J. M., Mason, L. B., Williamson, D. H., Harrison, H. B., Choukroun, S. & Jones, G. P. (2019). Successful validation of a larval dispersal model using genetic parentage data. PLOS Biology, 17(7), e3000380. https://doi.org/10.1371/journal.pbio.3000380

Bondy, A. & Murty, U. S. R. (1976). Graph theory with applications. The Macmillan Press Ltd.

Bonte, D., Dyck, H. V., Bullock, J. M., Coulon, A., Delgado, M., Gibbs, M., Lehouck, V., Matthysen, E., Mustin, K., Saastamoinen, M., Schtickzelle, N., Stevens, V. M., Vandewoestijne, S., Baguette, M., Barton, K., Benton, T. G., Chaput-Bardy, A., Clobert, J., Dytham, C., ... Travis, J. M. J. (2012). Costs of dispersal. Biological Reviews, 87(2), 290-312. https://doi.org/10.1111/j.1469-185X.2011.00201.x

Botsford, L. W., White, J. W., Coffroth, M. A., Paris, C. B., Planes, S., Shearer, T. L., Thorrold, S. R. & Jones, G. P. (2009). Connectivity and resilience of coral reef metapopulations in marine protected areas : matching empirical efforts to predictive needs. Coral Reefs, 28(2), 327-337. https://doi.org/10.1007/s00338-009-0466-z

Bottesch, M., Gerlach, G., Halbach, M., Bally, A., Kingsford, M. J. & Mouritsen, H. (2016). A magnetic compass that might help coral reef fish larvae return to their natal reef. Current Biology, 26(24), R1266-R1267. https://doi.org/10.1016/j.cub.2016.10.051

Boulanger, E., Dalongeville, A., Andrello, M., Mouillot, D. & Manel, S. (2020). Spatial graphs highlight how multi-generational dispersal shapes landscape genetic patterns. Ecography, 43(8), 1167-1179. https://doi.org/https://doi.org/10.1111/ecog.05024

Buonomo, R., Assis, J., Fernandes, F., Engelen, A. H., Airoldi, L. & Serrão, E. A. (2017). Habitat continuity and stepping-stone oceanographic distances explain population genetic connectivity of the brown alga Cystoseira amentacea. Molecular Ecology, 26(3), 766-780. https://doi.org/10.1111/mec.13960

Burgess, S. C., Baskett, M. L., Grosberg, R. K., Morgan, S. G. & Strathmann, R. R. (2016). When is dispersal for dispersal? Unifying marine and terrestrial perspectives : When is dispersal for dispersal? Biological Reviews, 91(3), 867-882. https://doi.org/10.1111/brv.12198

Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A. K., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W. & Botsford, L. W. (2014). Beyond connectivity : how empirical methods can quantify population persistence to improve marine protected-area design. Ecological Applications, 24(2), 257-270. https://doi.org/10.1890/13-0710.1

Butchart, S. H. M., Walpole, M., Collen, B., Strien, A. v., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., ... Watson, R. (2010). Global Biodiversity : Indicators of Recent Declines [Publisher : American Association for the Advancement of Science Section : Report]. Science, 328(5982), 1164-1168. https://doi.org/10.1126/science.1187512

Calò, A., Félix-Hackradt, F. C., Garcia, J., Hackradt, C. W., Rocklin, D., Treviño Otón, J. & Charton, J. A. G. (2013). A review of methods to assess connectivity and disper-

sal between fish populations in the Mediterranean Sea. Advances in Oceanography and Limnolog 4(2), 150-175.

Camaret, S., Guerin, B. & Leclerc, D. (1998). Impact of nutcracker (Nucifraga caryocatactes L.) on the spatial distribution of Swiss stone pine (Pinus cembra L.) regeneration. Bulletin de la Societe Zoologique de France (France).

Campana, S. E. (1999). Chemistry and composition of fish otoliths : pathways, mechanisms and applications. Marine Ecology Progress Series, 263-297.

Carreras, C., García-Cisneros, A., Wangensteen, O. S., Ordóñez, V., Palacín, C., Pascual, M. & Turon, X. (2020). East is East and West is West : Population genomics and hierarchical analyses reveal genetic structure and adaptation footprints in the keystone species *Paracentrotus lividus* (Echinoidea) (A. Zhan, Éd.). Divers Distrib, 26(3), 382-398. https://doi.org/10.1111/ddi.13016

Carroll, S. P., Hendry, A. P., Reznick, D. N. & Fox, C. W. (2007). Evolution on ecological time-scales. Functional Ecology, 21(3), 387-393. https://doi.org/10.1111/j.1365-2435.2007.01289.x

Catalano, K. A., Dedrick, A. G., Stuart, M. R., Puritz, J. B., Montes, H. R. & Pinsky, M. L. (2021). Quantifying dispersal variability among nearshore marine populations. Molecular Ecology, 30(10), 2366-2377. https://doi.org/10.1111/mec.15732

Chia, F.-S., Buckland-Nicks, J. & Young, C. M. (2011). Locomotion of marine invertebrate larvae : a review [Publisher : NRC Research Press Ottawa, Canada]. Canadian Journal of Zoology. https://doi.org/10.1139/z84-176

Clark, D. L., Leis, J. M., Hay, A. C. & Trnski, T. (2005). Swimming ontogeny of larvae of four temperate marine fishes. Marine Ecology Progress Series, 292, 287-300.

Clobert, J., Baguette, M., Benton, T. G. & Bullock, J. M. (2012). Dispersal Ecology and Evolution. Oxford University Press.

Coelho, M. A. G. & Lasker, H. R. (2016). Larval Dispersal and Population Connectivity in Anthozoans. In S. Goffredo & Z. Dubinsky (Éd.), The Cnidaria, Past, Present and Future : The worl (p. 291-315). Springer International Publishing. https://doi.org/10.1007/978-3-319-31305-4_19

Costantini, F., Carlesi, L. & Abbiati, M. (2013). Quantifying Spatial Genetic Structuring in Mesophotic Populations of the Precious Coral Corallium rubrum (J. M. Roberts, Éd.). PLoS ONE, 8(4), e61546. https://doi.org/10.1371/journal.pone.0061546

Cowen, R. K., GAWARKIEWICZ, G., PINEDA, J., THORROLD, S. R. & WERNER, F. E. (2007). Population Connectivity in Marine Systems An Overview. Oceanography,

20(3), 14-21. Récupérée 19 février 2019, à partir de https://www.jstor.org/stable/24860093

Cowen, R. K., Lwiza, K. M., Sponaugle, S., Paris, C. B. & Olson, D. B. (2000). Connectivity of marine populations : open or closed? Science, 287(5454), 857-859.

Cowen, R. K., Paris, C. B. & Srinivasan, A. (2006). Scaling of connectivity in marine populations. Science, 311(5760), 522-527.

Cowen, R. K. & Sponaugle, S. (2009). Larval dispersal and marine population connectivity.

Crandall, E. D., Treml, E. A. & Barber, P. H. (2012). Coalescent and biophysical models of stepping-stone gene flow in neritid snails. Molecular Ecology, 21(22), 5579-5598. https://doi.org/https://doi.org/10.1111/mec.12031

D'Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G. & Buston, P. M. (2015). Patterns, causes, and consequences of marine larval dispersal. PNAS, 112(45), 13940-13945. https://doi.org/10.1073/pnas.1513754112

Dalongeville, A., Andrello, M., Mouillot, D., Lobreaux, S., Fortin, M.-J., Lasram, F., Belmaker, J., Rocklin, D. & Manel, S. (2018). Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. Evolutionary Applications, 11(8), 1437-1447. https://doi.org/10.1111/eva.12638

Damasio, A. (2019). Les furtifs. La Volte. Récupérée 29 janvier 2022, à partir de https://lavolte.net/livres/les-furtifs-alain-damasio/

De Wit, P., Jonsson, P. R., Pereyra, R. T., Panova, M., André, C. & Johannesson, K. (2020). Spatial genetic structure in a crustacean herbivore highlights the need for local considerations in Baltic Sea biodiversity management. Evolutionary Applications, 13(5), 974-990. https://doi.org/10.1111/eva.12914

Del Monte-Luna, P., Brook, B. W., Zetina-Rejón, M. J. & Cruz-Escalona, V. H. (2004). The carrying capacity of ecosystems. Global Ecology and Biogeography, 13(6), 485-495. https://doi.org/10.1111/j.1466-822X.2004.00131.x

Di Franco, A., De Benedetto, G., De Rinaldis, G., Raventos, N., Sahyoun, R. & Guidetti, P. (2011). Large scale-variability in otolith microstructure and microchemistry : the case study of Diplodus sargus sargus (Pisces : Sparidae) in the Mediterranean Sea. Italian Journal of Zoology, 78(2), 182-192.

Di Franco, A., Calò, A., Pennetta, A., De Benedetto, G., Planes, S. & Guidetti, P. (2015). Dispersal of larval and juvenile seabream : Implications for Mediterranean marine protected areas. Biological Conservation, 192, 361-368.

Di Franco, A., Gillanders, B. M., De Benedetto, G., Pennetta, A., De Leo, G. A. & Guidetti, P. (2012). Dispersal Patterns of Coastal Fish : Implications for Designing Networks of Marine Protected Areas (R. K. F. Unsworth, Éd.). PLoS ONE, 7(2), e31681. https://doi.org/10.1371/journal.pone.0031681

Di Franco, A. & Guidetti, P. (2011). Patterns of variability in early-life traits of fishes depend on spatial scale of analysis. Biology Letters, 7(3), 454-456.

Di Franco, A., Qian, K., Calò, A., Di Lorenzo, M., Planes, S. & Guidetti, P. (2013). Patterns of variability in early life traits of a Mediterranean coastal fish. Marine Ecology Progress Series, 476, 227-235.

Dickey, T. D. & Bidigare, R. R. (2005). Interdisciplinary oceanographic observations : the wave of the future [Number : S1]. Scientia Marina, 69(S1), 23-42. https://doi.org/10.3989/scimar.2005.69s123

Diedenhofen, B. & Musch, J. (2015). cocor : A Comprehensive Solution for the Statistical Comparison of Correlations [Publisher : Public Library of Science]. PLOS ONE, 10(4), e0121945. https://doi.org/10.1371/journal.pone.0121945

Dobbelaere, T., Muller, E. M., Gramer, L. J., Holstein, D. M. & Hanert, E. (2020). Coupled Epidemio-Hydrodynamic Modeling to Understand the Spread of a Deadly Coral Disease in Florida. Frontiers in Marine Science, 7, 1016. https://doi.org/10.3389/fmars.2020.591881

Domeier, M. L. & Colin, P. L. (1997). Tropical reef fish spawning aggregations : defined and reviewed. Bulletin of Marine Science, 60(3), 698-726.

Dubois, M., Rossi, V., Ser-Giacomi, E., Arnaud-Haond, S., López, C. & Hernández-García, E. (2016). Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems. Global ecology and biogeography, 25(5), 503-515.

Dupont, S., Lundve, B. & Thorndyke, M. (2010). Near future ocean acidification increases growth rate of the lecithotrophic larvae and juveniles of the sea star Crossaster papposus. Journal of Experimental Zoology Part B : Molecular and Developmental Ev 314B(5), 382-389. https://doi.org/10.1002/jez.b.21342

Duputié, A. & Massol, F. (2013). An empiricist's guide to theoretical predictions on the evolution of dispersal [Publisher : The Royal Society]. Interface focus, 3(6), 20130028.

Durand, J., Blel, H., Shen, K., Koutrakis, E. & Guinand, B. (2013). Population genetic structure of Mugil cephalus in the Mediterranean and Black Seas : a single

mitochondrial clade and many nuclear barriers. Mar. Ecol. Prog. Ser., 474, 243-261. https://doi.org/10.3354/meps10080

Egevang, C., Stenhouse, I. J., Phillips, R. A., Petersen, A., Fox, J. W. & Silk, J. R. D. (2010). Tracking of Arctic terns Sterna paradisaea reveals longest animal migration [Publisher : National Academy of Sciences Section : Biological Sciences]. PNAS, 107(5), 2078-2081. https://doi.org/10.1073/pnas.0909493107

Eldon, B., Riquet, F., Yearsley, J., Jollivet, D. & Broquet, T. (2016). Current hypotheses to explain genetic chaos under the sea. Current Zoology, 62(6), 551-566. https://doi.org/10.1093/cz/zow094

Eriksson, A., Elías-Wolff, F., Mehlig, B. & Manica, A. (2014). The emergence of the rescue effect from explicit within- and between-patch dynamics in a metapopulation [Publisher : Royal Society]. Proceedings of the Royal Society B : Biological Sciences, 281(1780), 20133127. https://doi.org/10.1098/rspb.2013.3127

Erisman, B., Heyman, W., Kobara, S., Ezer, T., Pittman, S., Aburto-Oropeza, O. & Nemeth, R. S. (2017). Fish spawning aggregations : where well-placed management actions can yield big benefits for fisheries and conservation. Fish and Fisheries, 18(1), 128-144. https://doi.org/10.1111/faf.12132

Faillettaz, R., Durand, E., Paris, C. B., Koubbi, P. & Irisson, J.-O. (2018). Swimming speeds of Mediterranean settlement-stage fish larvae nuance Hjort's aberrant drift hypothesis. Limnology and Oceanography, 63(2), 509-523.

Faillettaz, R., Paris, C. B. & Irisson, J.-O. (2018). Larval Fish Swimming Behavior Alters Dispersal Patterns From Marine Protected Areas in the North-Western Mediterranean Sea. Frontiers in Marine Science, 5, 97.

Faillettaz, R., Blandin, A., Paris, C. B., Koubbi, P. & Irisson, J.-O. (2015). Sun-compass orientation in Mediterranean fish larvae. PloS one, 10(8), e0135213.

Fisher, R. A. (1934). Statistical methods for research workers. [Publisher : Oliver and Boyd, Edinburgh and London]. Statistical methods for research workers., (5th Ed). Récupérée 15 janvier 2021, à partir de https://www.cabdirect.org/cabdirect/abstract/19351601205

Forterre, Y., Marmottant, P., Quilliet, C. & Noblin, X. (2016). Physics of rapid movements in plants. Europhysics News, 47(1), 27-30. https://doi.org/10.1051/epn/2016104

Foster, N. L., Paris, C. B., Kool, J. T., Baums, I. B., Stevens, J. R., Sanchez, J. A., Bastidas, C., Agudelo, C., Bush, P., Day, O., Ferrari, R., Gonzalez, P., Gore, S., Guppy, R., McCARTNEY, M. A., McCOY, C., Mendes, J., Srinivasan, A., Steiner, S., ... Mumby,

P. J. (2012). Connectivity of Caribbean coral populations : complementary insights from empirical and modelled gene flow. Molecular Ecology, 21(5), 1143-1157. https://doi.org/10.1111/j.1365-294X.2012.05455.x

Gagnaire, P.-A. (2020). Comparative genomics approach to evolutionary process connectivity. Evolutionary Applications, 13(6), 1320-1334. https://doi.org/10.1111/eva.12978

Gerber, L. R., Mancha-Cisneros, M. D. M., O'Connor, M. I. & Selig, E. R. (2014). Climate change impacts on connectivity in the ocean : Implications for conservation. Ecosphere, 5(3), art33. https://doi.org/10.1890/ES13-00336.1

Giannoulaki, M., Belluscio, A., Colloca, F., Fraschetti, S., Scardi, M., Smith, C., Panayotidis, P., Valavanis, V. & Spedicato, M. (2013). Mediterranean Sensitive Habitats (MEDISEH), final project report. Récupérée 23 juillet 2020, à partir de https://gis.ices.dk/geonetwork/srv/api/records/c690ac1c-015b-43e2-8aa1-0aff72eec50e

Gollasch, S. (2007). Is Ballast Water a Major Dispersal Mechanism for Marine Organisms? In W. Nentwig (Éd.), Biological Invasions (p. 49-57). Springer. https://doi.org/10.1007/978-3-540-36920-2_4

Green, A. & Figuerola, J. (2005). Recent advances in the study of long-distance dispersal of aquatic invertebrates via birds. Diversity and Distributions, 11(2), 149-156. https://doi.org/10.1111/j.1366-9516.2005.00147.x

Green, B., Mapstone, B. D., Carlos, G. & Begg, G. A. (2009). Tropical otoliths–where to next? Tropical Fish Otoliths : Information for Assessment, Management and Ecology (p. 296-301). Springer.

Green, B. S. & Fisher, R. (2004). Temperature influences swimming speed, growth and larval duration in coral reef fish larvae. Journal of Experimental Marine Biology and Ecology, 299(1), 115-132. https://doi.org/10.1016/j.jembe.2003.09.001

Guisan, A. & Thuiller, W. (2005). Predicting species distribution : offering more than simple habitat models. Ecology Letters, 8(9), 993-1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x

Guizien, K., Viladrich, N., Martínez-Quintana, Á. & Bramanti, L. (2020). Survive or swim : different relationships between migration potential and larval size in three sympatric Mediterranean octocorals [Number : 1 Publisher : Nature Publishing Group]. Sci Rep, 10(1), 18096. https://doi.org/10.1038/s41598-020-75099-1

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg, K. & Palstra, F. (2011). Understanding and Estimating Effective Population Size for Practical Application in Marine Species Management. Conservation Biology, 25(3), 438-449. https://doi.org/10.1111/j.1523-1739.2010.01637.x

Harrison, H. B., Bode, M., Williamson, D. H., Berumen, M. L. & Jones, G. P. (2020). A connectivity portfolio effect stabilizes marine reserve performance [Publisher : National Academy of Sciences Section : Biological Sciences]. PNAS, 117(41), 25595-25600. https://doi.org/10.1073/pnas.1920580117

Hauser, L. & Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics : ugly hypotheses slain by beautiful facts. Fish and Fisheries, 9(4), 333-362. https://doi.org/10.1111/j.1467-2979.2008.00299.x

Hays, G. C. & Scott, R. (2013). Global patterns for upper ceilings on migration distance in sea turtles and comparisons with fish, birds and mammals. Functional Ecology, 27(3), 748-756. https://doi.org/10.1111/1365-2435.12073

Hein, A. M., Hou, C. & Gillooly, J. F. (2012). Energetic and biomechanical constraints on animal migration distance. Ecology Letters, 15(2), 104-110. https://doi.org/10.1111/j.1461-0248.2011.01714.x

Hellberg, M. E. (2009). Gene Flow and Isolation among Populations of Marine Animals. Annual Review of Ecology, Evolution, and Systematics, 40, 291-310. Récupérée 20 février 2020, à partir de https://www.jstor.org/stable/20744042

Hendrickson, G. F., Stanley, J. C. & Hills, J. R. (1970). Olkin's New Formula for Significance of r13 vs. r23 Compared with Hotelling's Method [Publisher : American Educational Research Association]. American Educational Research Journal, 7(2), 189-195. https://doi.org/10.3102/00028312007002189

Hernández-Carrasco, I., Solabarrieta, L., Rubio, A., Esnaola, G., Reyes, E. & Orfila, A. (2018). Impact of HF radar current gap-filling methodologies on the Lagrangian assessment of coastal dynamics [Publisher : Copernicus GmbH]. Ocean Science, 14(4), 827-847. https://doi.org/10.5194/os-14-827-2018

Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernández-García, E., Guijarro, B., Massutí, E., Alemany, F., Jadaud, A., Perez, J. L. & Reglero, P. (2019). Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations. Ecological Applications, 0(0), e01913. https://doi.org/10.1002/eap.1913

Jackson, T. M., Roegner, G. C. & O'Malley, K. G. (2018). Evidence for interannual variation in genetic structure of Dungeness crab (Cancer magister) along the

California Current System. Molecular Ecology, 27(2), 352-368. https://doi.org/10.1111/mec.14443

Jacobi, M. N., André, C., Döös, K. & Jonsson, P. R. (2012). Identification of subpopulations from connectivity matrices. Ecography, 35(11), 1004-1016. https://doi.org/10.1111/j.1600-0587.2012.07281.x

Jahnke, M. & Jonsson, P. R. (2022). Biophysical models of dispersal contribute to seascape genetic analyses. Phil. Trans. R. Soc. B, 377(1846), 20210024. https://doi.org/10.1098/rstb.2021.0024

Jahnke, M., Jonsson, P. R., Moksnes, P.-O., Loo, L.-O., Jacobi, M. N. & Olsen, J. L. (2018). Seascape genetics and biophysical connectivity modelling support conservation of the seagrass Zostera marina in the Skagerrak–Kattegat region of the eastern North Sea [Publisher : John Wiley & Sons, Ltd]. Evolutionary Applications, 11(5), 645-661. https://doi.org/10.1111/eva.12589

Jenkins, D. G., Carey, M., Czerniewska, J., Fletcher, J., Hether, T., Jones, A., Knight, S., Knox, J., Long, T., Mannino, M., McGuire, M., Riffle, A., Segelsky, S., Shappell, L., Sterner, A., Strickler, T. & Tursi, R. (2010). A meta-analysis of isolation by distance : relic or reference standard for landscape genetics? [Publisher : John Wiley & Sons, Ltd]. Ecography, 33(2), 315-320. https://doi.org/10.1111/j.1600-0587.2010.06285.x

Jorda, G., Marbà, N., Bennett, S., Santana-Garcon, J., Agusti, S. & Duarte, C. M. (2020). Ocean warming compresses the three-dimensional habitat of marine life [Bandiera_abtest : a Cg_type : Nature Research Journals Number : 1 Primary_atype : Research Publisher : Nature Publishing Group Subject_term : Climate-change ecology;Marine biology;Physical oceanography Subject_term_id : climate-change-ecology;marine-biology;physical-oceanography]. Nat Ecol Evol, 4(1), 109-114. https://doi.org/10.1038/s41559-019-1058-0

Kendrick, G. A., Orth, R. J., Statton, J., Hovey, R., Montoya, L. R., Lowe, R. J., Krauss, S. L. & Sinclair, E. A. (2017). Demographic and genetic connectivity : the role and consequences of reproduction, dispersal and recruitment in seagrasses. Biological Reviews, 92(2), 921-938. https://doi.org/10.1111/brv.12261

Kendrick, G. A., Waycott, M., Carruthers, T. J. B., Cambridge, M. L., Hovey, R., Krauss, S. L., Lavery, P. S., Les, D. H., Lowe, R. J., Vidal, O. M. i., Ooi, J. L. S., Orth, R. J., Rivers, D. O., Ruiz-Montoya, L., Sinclair, E. A., Statton, J., van Dijk, J. K. & Verduin, J. J. (2012). The Central Role of Dispersal in the Maintenance and

Persistence of Seagrass Populations. BioScience, 62(1), 56-65. https://doi.org/10.1525/bio.2012.62.1.10

Kool, J. T., Moilanen, A. & Treml, E. A. (2013). Population connectivity : recent advances and new perspectives. Landscape Ecology, 28(2), 165-185.

Kool, J. T., Paris, C. B., Andréfouët, S. & Cowen, R. K. (2010). Complex migration and the development of genetic structure in subdivided populations : an example from Caribbean coral reef ecosystems. Ecography, 33(3), 597-606. https://doi.org/10.1111/j.1600-0587.2009.06012.x

Kough, A. S., Paris, C. B. & Iv, M. J. B. (2013). Larval Connectivity and the International Management of Fisheries. PLOS ONE, 8(6), e64970. https://doi.org/10.1371/journal.pone.0064970

Leggat, W. P., Camp, E. F., Suggett, D. J., Heron, S. F., Fordyce, A. J., Gardner, S., Deakin, L., Turner, M., Beeching, L. J., Kuzhiumparambil, U., Eakin, C. M. & Ainsworth, T. D. (2019). Rapid Coral Decay Is Associated with Marine Heatwave Mortality Events on Reefs. Current Biology, 29(16), 2723-2730.e4. https://doi.org/10.1016/j.cub.2019.06.077

Legrand, T., Di Franco, A., Ser-Giacomi, E., Caló, A. & Rossi, V. (2019). A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish. Marine Environmental Research, 151, 104761. https://doi.org/10.1016/j.marenvres.2019.104761

Leis, J. M. (2006). Are larvae of demersal fishes plankton or nekton ? Advances in marine biology, 51, 57-141.

Lenormand, T. (2002). Gene flow and the limits to natural selection. Trends in Ecology & Evolution, 17(4), 183-189. https://doi.org/10.1016/S0169-5347(02)02497-7

Lester, S. E., Halpern, B. S., Grorud-Colvert, K., Lubchenco, J., Ruttenberg, B. I., Gaines, S. D., Airamé, S. & Warner, R. R. (2009). Biological effects within no-take marine reserves : a global synthesis. Marine Ecology Progress Series, 384, 33-46.

Lett, C., Ayata, S.-D., Huret, M. & Irisson, J.-O. (2010). Biophysical modelling to investigate the effects of climate change on marine population dispersal and connectivity. Progress in Oceanography, 87(1), 106-113. https://doi.org/10.1016/j.pocean.2010.09.005

Lett, C., Barrier, N. & Bahlali, M. (2020). Converging approaches for modeling the dispersal of propagules in air and sea. Ecological Modelling, 415, 108858. https://doi.org/10.1016/j.ecolmodel.2019.108858

Lett, C., Verley, P., Mullon, C., Parada, C., Brochier, T., Penven, P. & Blanke, B. (2008). A Lagrangian tool for modelling ichthyoplankton dynamics. Environmental Modelling & Software, 23(9), 1210-1214. https://doi.org/10.1016/j.envsoft.2008.02.005

Liggins, L., Treml, E. A. & Riginos, C. (2013). Taking the Plunge : An Introduction to Undertaking Seascape Genetic Studies and using Biophysical Models [_eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/gec3.12031]. Geography Compass, 7(3), 173-196. https://doi.org/10.1111/gec3.12031

Lowe, W. H. & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? Molecular Ecology, 19(15), 3038-3051. https://doi.org/10.1111/j.1365-294X.2010.04688.x

Lowe, W. H., Kovach, R. P. & Allendorf, F. W. (2017). Population Genetics and Demography Unite Ecology and Evolution. Trends in Ecology & Evolution, 32(2), 141-152. https://doi.org/10.1016/j.tree.2016.12.002

Madec, G. et al. (2015). NEMO ocean engine.

Mari, L., Melià, P., Fraschetti, S., Gatto, M. & Casagrandi, R. (2020). Spatial patterns and temporal variability of seagrass connectivity in the Mediterranean Sea. Diversity and Distributions, 26(2), 169-182. https://doi.org/10.1111/ddi.12998

Marshall, D. J., Monro, K., Bode, M., Keough, M. J. & Swearer, S. (2010). Phenotype–environment mismatches reduce connectivity in the sea. Ecology Letters, 13(1), 128-140. https://doi.org/https://doi.org/10.1111/j.1461-0248.2009.01408.x

Marshall, D. J., Gaines, S., Warner, R., Barneche, D. R. & Bode, M. (2019). Underestimating the benefits of marine protected areas for the replenishment of fished populations. Frontiers in Ecology and the Environment, 0(0). https://doi.org/10.1002/fee.2075

Martin, C. S., Giannoulaki, M., De Leo, F., Scardi, M., Salomidi, M., Knittweis, L., Pace, M. L., Garofalo, G., Gristina, M., Ballesteros, E., Bavestrello, G., Belluscio, A., Cebrian, E., Gerakaris, V., Pergent, G., Pergent-Martini, C., Schembri, P. J., Terribile, K., Rizzo, L., … Fraschetti, S. (2014). Coralligenous and maërl habitats : predictive modelling to identify their spatial distributions across the Mediterranean Sea. Scientific Reports, 4, 5073. https://doi.org/10.1038/srep05073

Marzouk, Z., Aurelle, D., Said, K. & Chenuil, A. (2017). Cryptic lineages and high population genetic structure in the exploited marine snail Hexaplex trunculus (Gastropoda : Muricidae). Biological Journal of the Linnean Society, 122(2), 411-428. https://doi.org/10.1093/biolinnean/blx070

Matthysen, E. (2012). Multicausality of dispersal : a review. Dispersal ecology and evolution (p. 3-18.). Oxford University Press.

McMahon, K., van Dijk, K.-j., Ruiz-Montoya, L., Kendrick, G. A., Krauss, S. L., Waycott, M., Verduin, J., Lowe, R., Statton, J., Brown, E. & Duarte, C. (2014). The movement ecology of seagrasses [Publisher : Royal Society]. Proceedings of the Royal Society B : Biologi 281(1795), 20140878. https://doi.org/10.1098/rspb.2014.0878

McRae, B. H. & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. Proceedings of the National Academy of Sciences, 104(50), 19885-19890. https://doi.org/10.1073/pnas.0706568104

McRae, B. H. (2006). Isolation by Resistance. Evolution, 60(8), 1551-1561. https://doi.org/10.1111/j.0014-3820.2006.tb00500.x

McWilliams, J. C. (2016). Submesoscale currents in the ocean [Publisher : Royal Society]. Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences, 472(2189), 20160117. https://doi.org/10.1098/rspa.2016.0117

Melià, P., Schiavina, M., Rossetto, M., Gatto, M., Fraschetti, S. & Casagrandi, R. (2016). Looking for hotspots of marine metacommunity connectivity : a methodological framework. Scientific Reports, 6, 23705.

Menge, B. A., Gouhier, T. C., Freidenburg, T. & Lubchenco, J. (2011). Linking long-term, large-scale climatic and environmental variability to patterns of marine invertebrate recruitment : Toward explaining “unexplained” variation. Journal of Experimental Marine Biology and Ecology, 400(1), 236-249. https://doi.org/10.1016/j.jembe.2011.02.003

Mercier, A., Sewell, M. A. & Hamel, J.-F. (2013). Pelagic propagule duration and developmental mode : reassessment of a fading link [_eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/geb.12018]. Global Ecology and Biogeography, 22(5), 517-530. https://doi.org/10.1111/geb.12018

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., Espiñeira, M., Fiorentino, F., Garofalo, G., Geffen, A. J., Hansen, J. H., Helyar, S. J., Nielsen, E. E., Ogden, R., Patarnello, T., Stagioni, M., FishPopTrace Consortium, Tinti, F. & Bargelloni, L. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). Mol Ecol, 23(1), 118-135. https://doi.org/10.1111/mec.12568

Millot, C. & Taupier-Letage, I. (2005). Circulation in the Mediterranean sea. The Mediterranean Sea (p. 29-66). Springer.

Mitcheson, Y. S. D., Cornish, A., Domeier, M., Colin, P. L., Russell, M. & Lindeman, K. C. (2008). A Global Baseline for Spawning Aggregations of Reef Fishes. Conservation Biology, 22(5), 1233-1244. https://doi.org/10.1111/j.1523-1739.2008.01020.x

Mokhtar-Jamaï, K., Pascual, M., Ledoux, J.-B., Coma, R., Féral, J.-P., Garrabou, J. & Aurelle, D. (2011). From global to local genetic structuring in the red gorgonian Paramuricea clavata : the interplay between oceanographic conditions and limited larval dispersal. Molecular Ecology, 20(16), 3291-3305. https://doi.org/10.1111/j.1365-294X.2011.05176.x

Monroy, P., Rossi, V., Ser-Giacomi, E., López, C. & Hernández-García, E. (2017). Sensitivity and robustness of larval connectivity diagnostics obtained from Lagrangian Flow Networks. ICES Journal of Marine Science, 74(6), 1763-1779.

Munday, P. L., Leis, J. M., Lough, J. M., Paris, C. B., Kingsford, M. J., Berumen, M. L. & Lambrechts, J. (2009). Climate change and coral reef connectivity. Coral Reefs, 28(2), 379-395. https://doi.org/10.1007/s00338-008-0461-9

Nathan, R., Katul, G. G., Horn, H. S., Thomas, S. M., Oren, R., Avissar, R., Pacala, S. W. & Levin, S. A. (2002). Mechanisms of long-distance dispersal of seeds by wind [Bandiera_abtest : a Cg_type : Nature Research Journals Number : 6896 Primary_atype : Research Publisher : Nature Publishing Group]. Nature, 418(6896), 409-413. https://doi.org/10.1038/nature00844

Nichols, C. R. & Raghukumar, K. (2020). Marine Environmental Characterization [Publisher : Morgan & Claypool Publishers]. Synthesis Lectures on Ocean Systems Engineering, 1(1), 1-103. https://doi.org/10.2200/S01006ED1V01Y202004OSE002

Norton, T. (1992). Dispersal by macroalgae [Publisher : Taylor & Francis _eprint : https://doi.org/10.1080/00071619200650271]. British Phycological Journal, 27(3), 293-301. https://doi.org/10.1080/00071619200650271

Oddo, P., Adani, M., Pinardi, N., Fratianni, C., Tonani, M., Pettenuzzo, D. et al. (2009). A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting. Ocean science.

Odum, E. P. & Barrett, G. W. (2005). Fundamentals of ecology [OCLC : 56476957]. Thomson Brooks/Cole.

Oleksyn, S., Tosetto, L., Raoult, V. & Williamson, J. E. (2021). Drone-Based Tracking of the Fine-Scale Movement of a Coastal Stingray (Bathytoshia brevicaudata) [Number : 1 Publisher : Multidisciplinary Digital Publishing Institute]. Remote Sensing, 13(1), 40. https://doi.org/10.3390/rs13010040

Pante, E. & Simon-Bouhet, B. (2013). marmap : A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R [Publisher : Public Library of Science]. PLOS ONE, 8(9), e73051. https://doi.org/10.1371/journal.pone.0073051

Paris, C. B., Atema, J., Irisson, J.-O., Kingsford, M., Gerlach, G. & Guigand, C. M. (2013). Reef odor : a wake up call for navigation in reef fish larvae. PloS one, 8(8), e72808.

Paris, C. B., Helgers, J., van Sebille, E. & Srinivasan, A. (2013). Connectivity Modeling System : A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. Environmental Modelling & Software, 42, 47-54. https://doi.org/10.1016/j.envsoft.2012.12.006

Pascual, M., Palero, F., García-Merchán, V. H., Macpherson, E., Robainas-Barcia, A., Mestres, F., Roda, T. & Abelló, P. (2016). Temporal and spatial genetic differentiation in the crab Liocarcinus depurator across the Atlantic-Mediterranean transition. Sci Rep, 6(1), 29892. https://doi.org/10.1038/srep29892

Pascual, M., Rives, B., Schunter, C. & Macpherson, E. (2017). Impact of life history traits on gene flow : A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. PLOS ONE, 12(5), e0176419. https://doi.org/10.1371/journal.pone.0176419

Paterno, M., Schiavina, M., Aglieri, G., Souissi, J. B., Boscari, E., Casagrandi, R., Chassanite, A., Chiantore, M., Congiu, L., Guarnieri, G., Kruschel, C., Macic, V., Marino, I. A. M., Papetti, C., Patarnello, T., Zane, L. & Melià, P. (2017). Population genomics meet Lagrangian simulations : Oceanographic patterns and long larval duration ensure connectivity among Paracentrotus lividus populations in the Adriatic and Ionian seas. Ecology and Evolution, 7(8), 2463-2479. https://doi.org/10.1002/ece3.2844

Pechenik, J. (1999). On the advantages and disadvantages of larval stages in benthic marine invertebrate life cycles. Mar. Ecol. Prog. Ser., 177, 269-297. https://doi.org/10.3354/meps177269

Pedrotti, M. L. (1993). Spatial and temporal distribution and recruitment of echinoderm larvae in the Ligurian Sea. J. Mar. Biol. Ass., 73(3), 513-530. https://doi.org/10.1017/S0025315400033075

Pérez-Matus, A., Ospina-Alvarez, A., Camus, P. A., Carrasco, S. A., Fernandez, M., Gelcich, S., Godoy, N., Ojeda, F. P., Pardo, L. M., Rozbaczylo, N., Subida, M. D., Thiel, M., Wieters, E. A. & Navarrete, S. A. (2017). Temperate rocky subtidal reef com-

munity reveals human impacts across the entire food web. Marine Ecology Progress Series, 567, 1-16. https://doi.org/10.3354/meps12057

Pérez-Portela, R., Wangensteen, O. S., Garcia-Cisneros, A., Valero-Jiménez, C., Palacín, C. & Turon, X. (2019). Spatio-temporal patterns of genetic variation in Arbacia lixula, a thermophilous sea urchin in expansion in the Mediterranean [Number : 2 Publisher : Nature Publishing Group]. Heredity, 122(2), 244-259. https://doi.org/10.1038/s41437-018-0098-6

Pimm, S. L., Lawton, J. H. & Cohen, J. E. (1991). Food web patterns and their consequences [Bandiera_abtest : a Cg_type : Nature Research Journals Number : 6320 Primary_atype : Reviews Publisher : Nature Publishing Group]. Nature, 350(6320), 669-674. https://doi.org/10.1038/350669a0

Pinardi, N., Zavatarelli, M., Adani, M., Coppini, G., Fratianni, C., Oddo, P., Simoncelli, S., Tonani, M., Lyubartsev, V., Dobricic, S. & Bonaduce, A. (2015). Mediterranean Sea large-scale low-frequency ocean variability and water mass formation rates from 1987 to 2007 : A retrospective analysis. Progress in Oceanography, 132, 318-332. https://doi.org/10.1016/j.pocean.2013.11.003

Pineda, J. (2000). Linking larval settlement to larval transport : assumptions, potentials, and pitfalls. Oceanography of the eastern Pacific, 1(2000), 84-105.

Pineda, J., Hare, J. A. & Sponaugle, S. (2007). Larval transport and dispersal in the coastal ocean and consequences for population connectivity. Oceanography, 20(3), 22-39.

Pineda, J., Porri, F., Starczak, V. & Blythe, J. (2010). Causes of decoupling between larval supply and settlement and consequences for understanding recruitment and population connectivity. Journal of Experimental Marine Biology and Ecology, 392(1-2), 9-21. https://doi.org/10.1016/j.jembe.2010.04.008

Pinsky, M. L., Saenz-Agudelo, P., Salles, O. C., Almany, G. R., Bode, M., Berumen, M. L., Andréfouët, S., Thorrold, S. R., Jones, G. P. & Planes, S. (2017). Marine Dispersal Scales Are Congruent over Evolutionary and Ecological Time. Current Biology, 27(1), 149-154. https://doi.org/10.1016/j.cub.2016.10.053

Poulain, P.-M., Menna, M. & Mauri, E. (2012). Surface geostrophic circulation of the Mediterranean Sea derived from drifter and satellite altimeter data. Journal of Physical Oceanograph 42(6), 973-990.

Pringle, R. M. & Hutchinson, M. C. (2020). Resolving Food-Web Structure. Annu. Rev. Ecol. Evol. Syst., 51(1), 55-80. https://doi.org/10.1146/annurev-ecolsys-110218-024908

Queiroga, H. & Blanton, J. (2005). Interactions between behaviour and physical forcing in the control of horizontal transport of decapod crustacean larvae. Adv Mar Biol, 47, 107-214. https://doi.org/10.1016/s0065-2881(04)47002-3

Ramesh, N., Rising, J. A. & Oremus, K. L. (2019). The small world of global marine fisheries : The cross-boundary consequences of larval dispersal. Science, 364(6446), 1192-1196. https://doi.org/10.1126/science.aav3409

Riginos, C. & Liggins, L. (2013). Seascape Genetics : Populations, Individuals, and Genes Marooned and Adrift. Geography Compass, 7(3), 197-216. https://doi.org/10.1111/gec3.12032

Rossi, V., Lo, M., Legrand, T., Ser-Giacomi, E., De Jode, A., Thierry De Ville D'avray, L., Pairaud, I., Faure, V., Fraysse, M., Pinazo, C. & Chenuil, A. (2020). Small-scale connectivity of coralligenous habitats : insights from a modelling approach within a semi-opened Mediterranean bay. Vie Et Milieu-life And Environment, 70(3-4).

Rossi, V., Ser-Giacomi, E., López, C. & Hernández-García, E. (2014). Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. Geophysical Research Letters, 41(8), 2883-2891.

Rosvall, M. & Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences, 105(4), 1118-1123.

Rousset, F. (1997). Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance [Publisher : Genetics Section : Investigations]. Genetics, 145(4), 1219-1228. Récupérée 7 avril 2020, à partir de https://www.genetics.org/content/145/4/1219

Saebi, M., Xu, J., Curasi, S. R., Grey, E. K., Chawla, N. V. & Lodge, D. M. (2020). Network analysis of ballast-mediated species transfer reveals important introduction and dispersal patterns in the Arctic. Sci Rep, 10(1), 19558. https://doi.org/10.1038/s41598-020-76602-4

Sala, E., Mayorga, J., Bradley, D., Cabral, R. B., Atwood, T. B., Auber, A., Cheung, W., Costello, C., Ferretti, F., Friedlander, A. M., Gaines, S. D., Garilao, C., Goodell, W., Halpern, B. S., Hinson, A., Kaschner, K., Kesner-Reyes, K., Leprieur, F., McGowan, J., ... Lubchenco, J. (2021). Protecting the global ocean for biodiversity, food and climate [Number : 7854 Publisher : Nature Publishing Group]. Nature, 592(7854), 397-402. https://doi.org/10.1038/s41586-021-03371-z

Sanz-Aguilar, A., Igual, J. M., Tavecchia, G., Genovart, M. & Oro, D. (2016). When immigration mask threats : The rescue effect of a Scopoli's shearwater colony in the Western Mediterranean as a case study. Biological Conservation, 198, 33-36. https://doi.org/10.1016/j.biocon.2016.03.034

Saura, S., Bodin, Ö. & Fortin, M.-J. (2014). EDITOR'S CHOICE : Stepping stones are crucial for species' long-distance dispersal and range expansion through habitat networks [Publisher : John Wiley & Sons, Ltd]. Journal of Applied Ecology, 51(1), 171-182. https://doi.org/10.1111/1365-2664.12179

Schiavina, M., Marino, I. a. M., Zane, L. & Melià, P. (2014). Matching oceanography and genetics at the basin scale. Seascape connectivity of the Mediterranean shore crab in the Adriatic Sea. Molecular Ecology, 23(22), 5496-5507. https://doi.org/10.1111/mec.12956

Schunter, C., Pascual, M., Raventos, N., Garriga, J., Garza, J. C., Bartumeus, F. & Macpherson, E. (2019). A novel integrative approach elucidates fine-scale dispersal patchiness in marine populations. Sci Rep, 9(1), 1-10. https://doi.org/10.1038/s41598-019-47200-w

Selkoe, K. A. & Toonen, R. J. (2011). Marine connectivity : a new look at pelagic larval duration and genetic metrics of dispersal. Marine Ecology Progress Series, 436, 291-305. https://doi.org/10.3354/meps09238

Selkoe, K. A., Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., von der Heyden, S. & Toonen, R. J. (2016). A decade of seascape genetics : contributions to basic and applied marine connectivity. Marine Ecology Progress Series, 554, 1-19.

Selkoe, K. A., Henzler, C. M. & Gaines, S. D. (2008). Seascape genetics and the spatial ecology of marine populations. Fish and Fisheries, 9(4), 363-377. https://doi.org/10.1111/j.1467-2979.2008.00300.x

Selkoe, K. A., Watson, J. R., White, C., Horin, T. B., Iacchei, M., Mitarai, S., Siegel, D. A., Gaines, S. D. & Toonen, R. J. (2010). Taking the chaos out of genetic patchiness : seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species [Publisher : John Wiley & Sons, Ltd]. Molecular Ecology, 19(17), 3708-3726. https://doi.org/10.1111/j.1365-294X.2010.04658.x

Sen Gupta, A., Stellema, A., Pontes, G. M., Taschetto, A. S., Vergés, A. & Rossi, V. (2021). Future changes to the upper ocean Western Boundary Currents across two

generations of climate models. Sci Rep, 11(1), 9538. https://doi.org/10.1038/s41598-021-88934-w

Ser-Giacomi, E., Rossi, V., Lopez, C. & Hernandez-Garcia, E. (2015). Flow networks : A characterization of geophysical fluid transport [arXiv : 1409.4171]. Chaos : An Interdisciplinary Jou 25(3), 036404. https://doi.org/10.1063/1.4908231

Ser-Giacomi, E., Jordá-Sánchez, G., Soto-Navarro, J., Thomsen, S., Mignot, J., Sevault, F. & Rossi, V. (2020). Impact of Climate Change on Surface Stirring and Transport in the Mediterranean Sea. Geophysical Research Letters, 47(22), e2020GL089941. https://doi.org/10.1029/2020GL089941

Ser-Giacomi, E., Legrand, T., Hernández-Carrasco, I. & Rossi, V. (2021). Explicit and implicit network connectivity : Analytical formulation and application to transport processes [Publisher : American Physical Society]. Phys. Rev. E, 103(4), 042309. https://doi.org/10.1103/PhysRevE.103.042309

Ser-Giacomi, E., Vasile, R., Hernández-García, E. & López, C. (2015). Most probable paths in temporal weighted networks : An application to ocean transport. Phys. Rev. E, 92(1), 012818. https://doi.org/10.1103/PhysRevE.92.012818

Serra, I. A., Innocenti, A. M., Maida, G. D., Calvo, S., Migliaccio, M., Zambianchi, E., Pizzigalli, C., Arnaud-Haond, S., Duarte, C. M., Serrao, E. A. & Procaccini, G. (2010). Genetic structure in the Mediterranean seagrass Posidonia oceanica : disentangling past vicariance events from contemporary patterns of gene flow. Molecular Ecology, 19(3), 557-568. https://doi.org/10.1111/j.1365-294X.2009.04462.x

Shanks, A. L. (2009). Pelagic Larval Duration and Dispersal Distance Revisited. The Biological Bulletin, 216(3), 373-385. https://doi.org/10.1086/BBLv216n3p373

Sigwart, J. (2009). Coalescent Theory : An Introduction. Systematic Biology, 58(1), 162-165. https://doi.org/10.1093/schbul/syp004

Silver, N. C., Hittner, J. B. & May, K. (2004). Testing Dependent Correlations With Nonoverlapping Variables : A Monte Carlo Simulation [Publisher : Routledge _eprint : https ://doi.org/10.3200/JEXE.71.1.53-70]. The Journal of Experimental Education, 73(1), 53-69. https://doi.org/10.3200/JEXE.71.1.53-70

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. Science, 236(4803), 787-792. https://doi.org/10.1126/science.3576198

Slatkin, M. (1985). Gene Flow in Natural Populations. Annual Review of Ecology and Systematics, 1, 393-430. http://www.jstor.org/stable/2097054

Smith, D. (2013). Aeroplankton and the Need for a Global Monitoring Network. BioScience, 63(7), 515-516. https://doi.org/10.1525/bio.2013.63.7.3

Smith, T. M., York, P. H., Broitman, B. R., Thiel, M., Hays, G. C., Sebille, E. v., Putman, N. F., Macreadie, P. I. & Sherman, C. D. H. (2018). Rare long-distance dispersal of a marine angiosperm across the Pacific Ocean. Global Ecology and Biogeography, 27(4), 487-496. https://doi.org/https://doi.org/10.1111/geb.12713

Sponaugle, S., Cowen, R. K., Shanks, A., Morgan, S. G. & Leis, J. M. (2002). Predicting self-recruitment in marine populations : Biophysical correlates and mechanisms. BULLETIN OF MARINE SCIENCE, 70(1), 35.

Teixidó, N., Garrabou, J. & Harmelin, J.-G. (2011). Low Dynamics, High Longevity and Persistence of Sessile Structural Species Dwelling on Mediterranean Coralligenous Outcrops (S. Thrush, Éd.). PLoS ONE, 6(8), e23744. https://doi.org/10.1371/journal.pone.0023744

Telesca, L., Belluscio, A., Criscoli, A., Ardizzone, G., Apostolaki, E. T., Fraschetti, S., Gristina, M., Knittweis, L., Martin, C. S., Pergent, G., Alagna, A., Badalamenti, F., Garofalo, G., Gerakaris, V., Louise Pace, M., Pergent-Martini, C. & Salomidi, M. (2015). Seagrass meadows (Posidonia oceanica) distribution and trajectories of change [Number : 1 Publisher : Nature Publishing Group]. Scientific Reports, 5(1), 12505. https://doi.org/10.1038/srep12505

Teske, P. R., Sandoval-Castillo, J., van Sebille, E., Waters, J. & Beheregaray, L. B. (2016). Oceanography promotes self-recruitment in a planktonic larval disperser. Sci Rep, 6(1), 34205. https://doi.org/10.1038/srep34205

Thiébaut, E., Dauvin, J.-C. & Lagadeuc, Y. (1992). Transport of Owenia fusiformis larvae (Annelida : Polychaeta) in the Bay of Seine. I. Vertical distribution in relation to water column stratification and ontogenic vertical migration [Publisher : Inter-Research Science Center]. Marine Ecology Progress Series, 80(1), 29-39. Récupérée 30 janvier 2022, à partir de https://www.jstor.org/stable/24826542

Thorrold, S., Zacherl, D. & Levin, L. (2007). Population Connectivity and Larval Dispersal Using Geochemical Signatures in Calcified Structures. Oceanography, 20(3), 80-89. https://doi.org/10.5670/oceanog.2007.31

Tomback, D. F., Anderies, A. J., Carsey, K. S., Powell, M. L. & Mellmann-Brown, S. (2001). Delayed Seed Germination in Whitebark Pine and Regeneration Patterns Following the Yellowstone Fires. Ecology, 82(9), 2587-2600. https://doi.org/https://doi.org/10.1890/0012-9658(2001)082[2587:DSGIWP]2.0.CO;2

Treml, E. A., Halpin, P. N., Urban, D. L. & Pratson, L. F. (2008). Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. Landscape Ecol, 23(1), 19-36. https://doi.org/10.1007/s10980-007-9138-y

Turgeon, K., Robillard, A., Grégoire, J., Duclos, V. & Kramer, D. L. (2010). Functional connectivity from a reef fish perspective : behavioral tactics for moving in a fragmented landscape. Ecology, 91(11), 3332-3342. https://doi.org/10.1890/09-2015.1

van Sebille, E., Griffies, S. M., Abernathey, R., Adams, T. P., Berloff, P., Biastoch, A., Blanke, B., Chassignet, E. P., Cheng, Y., Cotter, C. J., Deleersnijder, E., Döös, K., Drake, H. F., Drijfhout, S., Gary, S. F., Heemink, A. W., Kjellsson, J., Koszalka, I. M., Lange, M., ... Zika, J. D. (2018). Lagrangian ocean analysis : Fundamentals and practices. Ocean Modelling, 121, 49-75. https://doi.org/10.1016/j.ocemod.2017.11.008

Viana, D. S., Santamaría, L. & Figuerola, J. (2016). Migratory Birds as Global Dispersal Vectors. Trends in Ecology & Evolution, 31(10), 763-775. https://doi.org/10.1016/j.tree.2016.07.005

Wang, I. J. & Bradburd, G. S. (2014). Isolation by environment. Molecular Ecology, 23(23), 5649-5662. https://doi.org/10.1111/mec.12938

Wang, I. J., Glor, R. E. & Losos, J. B. (2013). Quantifying the roles of ecology and geography in spatial genetic divergence. Ecology Letters, 16(2), 175-182. https://doi.org/10.1111/ele.12025

Waples, R. S. & Gaggiotti, O. (2006). INVITED REVIEW : What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology, 15(6), 1419-1439. https://doi.org/10.1111/j.1365-294X.2006.02890.x

Waters, J. M., Fraser, C. I. & Hewitt, G. M. (2013). Founder takes all : density-dependent processes structure biodiversity. Trends in Ecology & Evolution, 28(2), 78-85. https://doi.org/10.1016/j.tree.2012.08.024

Weber, A. a.-T., Mérigot, B., Valière, S. & Chenuil, A. (2015). Influence of the larval phase on connectivity : strong differences in the genetic structure of brooders and broadcasters in the Ophioderma longicauda species complex. Molecular Ecology, 24(24), 6080-6094. https://doi.org/https://doi.org/10.1111/mec.13456

Weersing, K. & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. Marine Ecology Progress Series, 393, 1-12. https://doi.org/10.3354/meps08287

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C. & Toonen, R. J. (2010). Ocean currents help explain population genetic structure [Publisher : Royal Society]. Proceedings of the Royal Society B : Biological Sciences, 277(1688), 1685-1694. https://doi.org/10.1098/rspb.2009.2214

Whitlock, M. C. & Mccauley, D. E. (1999). Indirect measures of gene flow and migration : FST1/(4Nm+1). Heredity, 82(2), 117-125. https://doi.org/10.1046/j.1365-2540.1999.00496.x

Wright, S. (1931). Evolution in Mendelian Populations. Genetics, 16(2), 97-159. Récupérée 30 octobre 2020, à partir de https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1201091/

Xuereb, A., Benestan, L., Normandeau, É., Daigle, R. M., Curtis, J. M. R., Bernatchez, L. & Fortin, M.-J. (2018). Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (Parastichopus californicus). Molecular Ecology, 27(10), 2347-2364. https://doi.org/10.1111/mec.14589

Zuercher, R. & Galloway, A. W. E. (2019). Coastal marine ecosystem connectivity : pelagic ocean to kelp forest subsidies. Ecosphere, 10(2), e02602. https://doi.org/10.1002/ecs2.2602

A. ANNEXES

A. Chapitre II

A.1. SI - A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish

SI - A multidisciplinary analytical framework to locate spawning areas and quantify larval dispersal in coastal fish

T. Legrand[a,*], A. Di Franco[b,c], E. Ser Giacomi[d], A. Caló[e,c], V. Rossi[a]

[a] *Mediterranean Institute of Oceanography (UM 110, UMR 7294), CNRS, Aix Marseille Univ., Univ. Toulon, IRD, 13288, Marseille, France*
[b] *Stazione zoologica Anton Dohrn, Dipartimento Ecologia Marina Integrata, Sede Interdipartimentale della Sicilia, Lungomare Cristoforo Colombo (complesso Roosevelt), 90142 Palermo, Italy*
[c] *Université Côte d'Azur, CNRS, UMR 7035 ECOSEAS, Parc Valrose 28, Avenue Valrose, 06108 Nice, France*
[d] *Sorbonne Universités (UPMC, Université Paris 06)-CNRS-IRD-MNHN, LOCEAN, 4 place JUSSIEU, F-75005 PARIS, France*
[e] *Dipartimento di Scienze della Terra e del Mare (DiSTeM), Università di Palermo, Via Archirafi 20, 90123 Palermo, Italy*

Contents

*Corresponding author

Part A

Methodological details of our multidisciplinary analytical framework

1. Lagrangian Flow Network

The Lagrangian Flow Network (LFN) is a methodology which combines network theory tools and Lagrangian trajectories to investigate transport and dispersal processes in oceanic flows. Full description can be found in Rossi et al. (2014); Ser-Giacomi et al. (2015a,b) and Dubois et al. (2016).

1.1. Parameters tuning

The ad-hoc LFN configuration is obtained through the selection of the most adequate hydrodynamical model and the fine-tuning of seven LFN parameters in accord with otolith analyses and both biological and numerical information.

The domain of the circulation model generating the required velocity field should be larger or equal than the oceanic region of interest. Its time range must be in accord with the spawning phenology determined by otolith sclerochronology analyses. The vertical layer (for z-coordinates model) must be chosen according to the most probable depth at which larvae of the target species are more likely to be found.

The starting date of each numerical experiment is simulating a single spawning event while the ensemble of Lagrangian experiments has to cover the full range of spawning dates derived from otolith sclerochronology. Moreover, if a broadcast spawner species is known to spawn repeatedly over a few months, a starting date periodicity of 5-10 days still provides robust results while minimizing computation needs (Monroy et al., 2017). When the spawning occurs

repeatedly but over shorter periods (e.g. few weeks to a month), a daily periodicity should be implemented to properly consider connectivity fluctuations due to the variability of ocean currents (Monroy et al., 2017).

Tracking time must be chosen to mimic the Pelagic Larval Duration (PLD), estimated from otoliths sclerochronology. According to the results of Monroy et al. (2017), LFN connectivity diagnostics are robust to small uncertainties on the PLD when it is greater than 15 days (which is true for many Mediterranean fishes, Macpherson and Raventos, 2006; Dubois et al., 2016). In this case, a single tracking time can be retained, corresponding to the median PLD.

The discretization of the domain of interest into a network of nodes provides a coarse-grained version of the information contained within the time- and space-dependent velocity field (Ser-Giacomi et al., 2015a,b). The spatial discretization can be of any size, compromising both the level of analyses and the computation time, given that it must be at least twice larger than the spatial resolution of the velocity field itself (so that each node contains at least 4 grid points of the flow field). To circumvent the typical issue of poorly resolved nearshore areas, each node is characterized by a variable called "land ratio" which estimates the proportion of land area within them (Ser-Giacomi et al., 2015a).

Land ratio is calculated as the ratio of grid points per node characterized by missing values ("land") to those for which the Eulerian model simulates realistic velocities ("ocean"). A node with a land ratio of 1 is considered fully covered by land and there is no current data in this node, hence no particles are seeded there. Contrarily, a node with a land ratio of 0 is considered to be entirely covered by the ocean. In this case, the entire node is covered by particles and all velocity field grid points comprised within this node have reliable current speeds and directions to integrate trajectories for all particles. For intermediate situations (e.g. node half covered by land, half by ocean), the seeding of particle is adjusted to the coverage of the flow field, initializing a particle only if it has four neighboring grid points with realistic currents to compute trajectories. In this way, all particles move around during the simulations and the slightly different number of initial particles per node is accounted for when normalizing

the matrices, ensuring its stochasticity.

Furthermore, we attribute for each sampling location $s \in [1; S]$ one or more *location nodes*, e.g. those few nodes which are located in the close vicinity of each sampled location (see section 4). Thus, for each sampling location $s \in [1; S]$ with, $x_\eta^s \in \mathbb{N}$ is the η^{th} index i or j assigned to the *location node* s, with $\eta \in [1; H]$. We may consider x_η as the η^{th} index i or j of all the combined *location nodes*.

The specification of purely numerical parameters, such as the Runge-Kutta integration time step and the number of particles per node, controls the robustness of the Lagrangian trajectories and corresponding connectivity matrix. The number of particles per node should be larger or equal to 100 particles, as prescribed by Monroy et al. (2017) and the Runge-Kutta time step should follows the Courant-Friedrichs-Lewy (CFL) condition imposed by the properties of the velocity field (Courant et al., 1928). Even though, due to numerical limitations when integrating trajectories, some particles may be confronted to the "beaching" effect. Indeed, when a particle is advected within the current field in the vicinity of the shoreline at a time t, its next position at time $t + \Delta t$, Δt being the Runge-Kutta time-step, might end up on land (i.e. without any velocity). Because this particle is no longer in the flow field, its next position cannot be integrated, and the particle is considered as lost. However, beaching issues are minimal in this study with the retained Runge-Kutta time-step, i.e. 20 min, satisfying the Courant-Friedrichs-Lewy (CLF) criteria. It concerns around less than 5 % of all particles and it occurs mainly in nodes with very high land ratio (Ser-Giacomi et al., 2015a, 2017).

1.2. Computing

Given E experiences (i.e. a set of simulations with different starting times but with similar tracking duration, depth of dispersion and network grid) indexed by $e \in [1; E]$, the associated Lagrangian Flow Networks (LFNs) are characterized by a set of weighted, directed connectivity matrices (i.e. adjacency

matrices) M^e (Figure SI-1b) defined as:

$$M^e_{ij} = \text{number of particles from node } i \text{ to node } j \text{ during experience } e \quad (1)$$

Another set of unweighted, directed matrices L^e represent the binary connectivity matrices (Figure SI-1c) defined as:

$$L^e_{ij} = \begin{cases} 1 & \text{if } M^e_{ij} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Both matrices describe, with different levels of details, the connections between any pair of nodes of the entire network (Ser-Giacomi et al., 2015a, 2017).

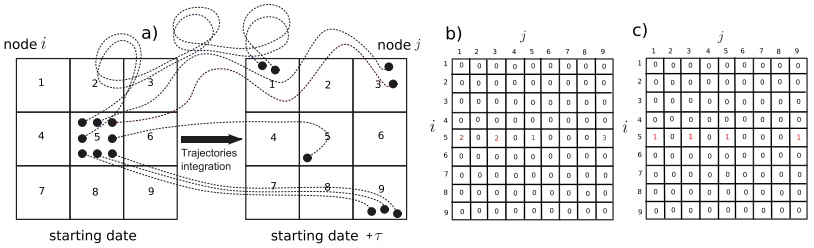


Figure SI-1: A schematic representation of the transport network (panel a) which permits the construction of a connectivity matrix M^e (panel b) and a binary connectivity matrix L^e (panel c). For a fixed tracking time τ, particles display trajectories from the origin node $i \in [1, N]$ to the destination node $j \in [1, N]$. Adapted from Ser-Giacomi et al. (2015a).

2. Filtering

The first filter is obtained by computing first the cumulative binary matrix L through the sum of all the binary connectivity matrices for all experiments $e \in [1; E]$ as follows: $\sum_{e=1}^{E} L^e$. Then, we select in L the corresponding columns j to all the x_η indices of location nodes. Finally, we sum these selected columns into one vector: $\sum_{\eta=1}^{H*S} L_{ix^s_\eta}$. By displaying the non-zero elements of that vector on a map, the first filter gives all the putative sources which are connected at least once to at least one of the seven locations.

3. Diagnosing

To take heed of the temporal variability of spawning in our ensemble of experiments, each connectivity matrix M^e must be weighted in accord with the observed spawning distribution. Thereby, we define a vector of multiplicating factors p^e (ranging from 0 to 1 and whose total sum is equal to 1) to weight each connectivity matrix according the starting date of that experiment. The factors $p^e \in [0, 1]$ are deduced by fitting a normal distribution on the observed distribution of spawning dates and by selecting the "normalized frequency" p^e for the retained spawning dates of our numerical experiments. We define thus a set of matrix $\bar{M^e}_{ij} = p^e M^e_{ij}$ that are a modulation of the original connectivity matrix M^e with the weight p^e. Finally, we compute the cumulative weighted matrix $\bar{M}$ as:

$$\bar{M} = \sum_{e=1}^{E} \bar{M^e} \tag{3}$$

To be interpreted as probabilities, the particle matrix $\bar{M^e}$ must be row- or column-normalized. To do so, each element is divided by the sum along each row ($\sum_{j=1}^{N} \bar{M^e_{ij}}$) or along each column ($\sum_{i=1}^{N} \bar{M^e_{ij}}$), respectively. It returns a normalized weighted matrix that we denote as K^e. By doing such normalization, it ensures the row- or column-stochasticity of the matrix, i.e all elements of each row or column sums to 1 (Ser-Giacomi et al., 2015a,b). In the case of a row-normalized matrix, the element K^e_{ij} gives the probability for a particle randomly chosen within departure node i to end up into node j. In the case of a column-normalized matrix the matrix element K^e_{ij} returns instead the probability for a particle randomly chosen in arrival node j to originate from node i.

Hence, the column-normalization of a particle matrix computed forward-in-time provides an approximation of the backward-in-time dynamics. In this study, we build and exploit column-normalized probabilistic matrices to identify spawning areas from pre-determined settlement sites. We denote f^α_n the index of a node selected by filter (α), $\alpha \in [1, 2, 3, 4]$ with $n \in [1, F^\alpha]$ (being F^α the total number of filtered-in nodes). Consequently, we retain in $\bar{M^e}$ the rows i

corresponding to all the f_i^α indices of filtered-in nodes and the columns j corresponding to all the x_k location nodes. For each filtered-in node f_n^α, transport probability is obtained as follows:

(i) column-normalize each matrix $\bar{M}^e$ to obtain the corresponding normalized matrix K^e:

$$K^e_{f_n^\alpha x_\eta} = \frac{\bar{M}^e{}_{f_n^\alpha x_\eta}}{\sum_{n=1}^{F^\alpha} \bar{M}^e{}_{f_n^\alpha x_\eta}} \tag{4}$$

(ii) sum all the column-normalized matrix K^e into one matrix:

$$K = \sum_{e=1}^{E} K^e_{f_n^\alpha x_\eta} \tag{5}$$

(iii) construct an array $A_{f_n^\alpha}$ which averages the H columns corresponding to all the location nodes indices x_η into one column:

$$A_{f_n^\alpha} = \frac{\sum_{\eta=1}^{H*S} K_{f_n^\alpha x_\eta}}{\sum_{n=1}^{F^\alpha} \sum_{\eta=1}^{H*S} K_{f_n^\alpha x_\eta}} \tag{6}$$

Thus, $A_{f_n^\alpha}$ indicates for each filter (α), the probability for each putative node source f_n^α to be connected to any of the S sampled locations.

Part B

Case-studies: *Diplodus sargus* and *Diplodus vulgaris* in the Adriatic Sea

1. Hydrodynamical Eulerian model

The velocity field comes from the Adriatic-Ionian REGional model (AIREG) developed by the CMCC Ocean Lab (Ciliberti et al., 2015; Oddo et al., 2006). This configuration is forced by momentum, water and heat fluxes using the horizontal-resolution operational atmospheric data provided by the European Centre for Medium-Range Weather Forecast. The initial and lateral boundary conditions originated from the Mediterranean Forecasting System (Oddo et al., 2009; Tonani et al., 2008). It covers the Adriatic and the Ionian Sea, from 31 °N to 46 °N and 6 °E to 22 °E over the years 2009-2010. The AIREG is based on the Nucleus for European Modelling of the Ocean model (Madec and others, 2015). The primitive equations are discretized horizontally on a 1/45°grid, with a final resolution of approximatively 2.2 km. It uses 121 vertical levels with high resolution in the top 100 m and stretched at the subsurface.

2. Overview of the regional circulation

To have a first glimpse at the circulation in the study area, as simulated by AIFS, two simple physical diagnostics were calculated over specific temporal periods: the mean modulus and the mean Eddy Kinetic Energy (EKE; Figure SI-2). The time ranges are chosen according to the total period of dispersion of both species, that is from the first spawning date to the last day of pelagic larvae drifting (i.e. the last starting date plus the tracking time). Thus, for *D.*

sargus the time range is from 04/05/2009 to 10/06/2009 and for *D. vulgaris* the time range is from 05/11/2009 to 16/03/2010.

The mean modulus is a time-averaged modulus (e.g. the current vector norm, $(u^2 + v^2)^{\frac{1}{2}}$) at each grid point of the flow field.

The EKE is first calculated by decomposing the full velocity in an average term, $< u >$ and a turbulent (or eddy) term, u' with $u =< u > +u'$. The averaged term corresponds to the time average for each grid point of the model. Then, u' is calculated by subtracting $(< u >)$ to the total velocity at each grid point for each day of the time range. Second, EKE is calculated according to the formula, $\frac{1}{2} * (u'^2 + v'^2)$. EKE is a physical variable commonly used to diagnose mesoscale activity (filament, eddies, frontal structures with spatial scales spanning 10-100 km and temporal scales ranging from of a few weeks to several months) and its variability (Waugh et al., 2006).

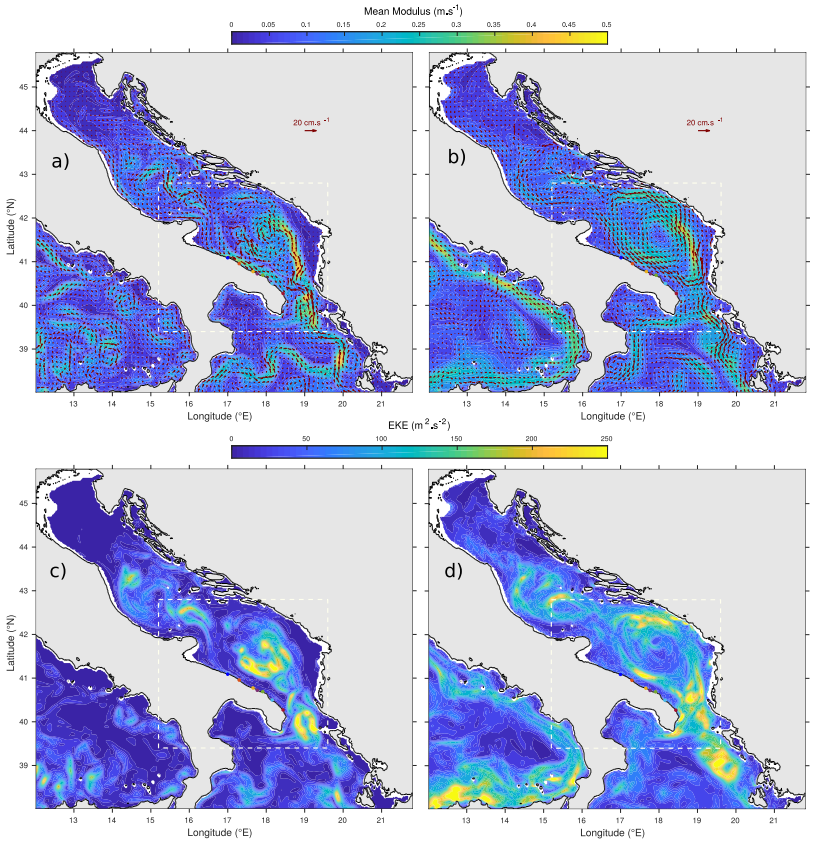


Figure SI-2: Averaged velocity fields (red arrows) superimposed on the mean modulus (colored background) over the periods of larval dispersal for a) *D. sargus* and b) *D. vulgaris*. Maps of mean EKE over the periods of larval dispersal for c) *D. sargus* and c) *D. vulgaris*.

3. Biological and biogeographical knowledge

3.1. Ototith analyses

Post-settlers (i.e. 1-1.5 cm body length) of *D. sargus* and *D. vulgaris* were collected at $S = 7$ distinct locations $s \in [1; S] \in \mathbb{N}$ separated by about 10-30 km, corresponding to about 180 km stretch of the Apulian coast. Within each location, two sites separated by 2 to 6 km were randomly selected for sampling. In this study, we decided to restrict our analyses at the scale of the locations by

considering both replicate sites together. Preliminary tests (not shown) revealed that the distances between sites are too small to be well discriminated in our transport network; moreover, by summing up the natal origins of both sites for each location, the number of representative samples per location increase, hence favoring the statistical power of our discriminating analyses.

Note that the central location is situated within the Torre Guaceto Marine Protected Area (TGMPA), while the other locations are not under any type of protection. The TGMPA was founded in 1991 but enforcement became effective in 2000-2001. The whole MPA, stretching along 8 km of coastline and covering about 22.27 km^2 in total, is divided into three zones.

Repeated visual surveys by divers allowed estimating the settlement peak; all specimens were then collected within 10 days to reduce temporal bias. After the samples collection, otoliths were removed from each specimen. First, sclerochronology techniques applied on otolith growth-increment allow estimating larval biological traits of interests (Di Franco and Guidetti, 2011; Di Franco et al., 2013). Second, the geochemical compositions of otolith cores were analyzed to gain insights into the fish natal origins (Di Franco et al., 2012, 2015). Three and seven different natal origins emerged for *D. sargus* and *D. vulgaris.*, respectively. The proportions of each natal origin for each location are reported in Table SI-1. Note that the minimum number of sampled location in which the rarest natal origin was found is four for *D. sargus* (Table SI-1). For *D. vulgaris*, it is three hen removing those three outliers natal origins named A, B and D which appeared only once at one location (Table SI-1).

Table SI-1: The proportion (%) of each natal origin for *D. sargus* (Di Franco et al., 2012) and *D. vulgaris* (Di Franco et al., 2015) at each sampling location and at all locations together as obtained from the otolith analyses. Here we are not considering otoliths with a "NA" origin.

Species		BA	M	HLD	TGMPA	PP	CAS	SA	ALL
	Number of samples	14	17	16	18	16	19	19	119
	Natal origin (%)								
	A	0	6	0	6	19	0	16	7
D. sargus	B	14	12	6	50	0	42	21	22
	C	86	82	94	44	81	58	63	71
	Number of samples	14	21	22	21	17	16	20	131
	Natal origin (%)								
	A	0	0	5	0	0	0	0	1
	B	0	5	0	0	0	0	0	1
	C	21	38	32	24	6	31	25	26
D. vulgaris	D	0	0	0	0	0	0	5	1
	E	36	38	55	48	53	31	50	45
	F	43	19	9	24	41	31	15	24
	G	0	0	0	5	0	6	5	2

3.2. Habitat mapping

Fish preferential habitats mapping was produced in GIS using QGIS software (QGIS Development Team, 2009). We used as base-layer the seabed habitat map available on the EMODnet portal (European Marine Observation and Data Network, http://www.emodnet.eu/seabed-habitats) that represents the most updated georeferenced seafloor map for the Mediterranean Sea. Habitat information were extracted from the EUSeaMap2016, a broad-scale habitat map with at 250m resolution, aggregated into the Benthic Broad Habitat Types of the Marine Strategy Framework Directive (MSFD) habitat classification. We retained only substrate types that represent preferential habitats for adults and sub-adults of *D. sargus* and *D. vulgaris*, i.e. the types '*infra-littoral and circa-littoral rock and other biogenic reefs*'. This category accounts for coastal rocky reefs, Posidonia oceanica meadows and coralligenous formations, representing

the habitats inhabited by the selected species (Harmelin-Vivien et al., 1995; Lenfant and Planes, 1996). The EUNIS habitat map is associated with a confidence layer that indicates the level of reliability (from 1=low to 3=high) of the source data (Figure SI-3). Some areas of the map included in the model domain are associated to a low level of confidence (i.e. eastern part of the Adriatic Sea). For areas with low confidence we carried out a two-steps approach aimed at refining the information and improve substrate map accuracy. Specifically, 1) we crossed-checked information on the original map with the distribution of coastline substrate type reported in Furlani et al. (2014), and for all the areas were information did not match, 2) we analyzed high-resolution satellite images (Google-Earth) to ascertain substrate type. In case of mismatch, that is when Furlani et al. (2014) reported rocky areas where EUNIS did not, our analysis of images confirmed the presence of rocky substrates. In these cases, we hand-corrected the map in QGIS by adding a buffer of rocky substrate along the coast with its extent being proportional to the sea bottom slope (the higher the slope, the shorter the width of the buffer). The same correction was applied for the portion of EUNIS original map where habitat information was absent (e.g. the north-eastern Ionian Sea, Figure SI-3).

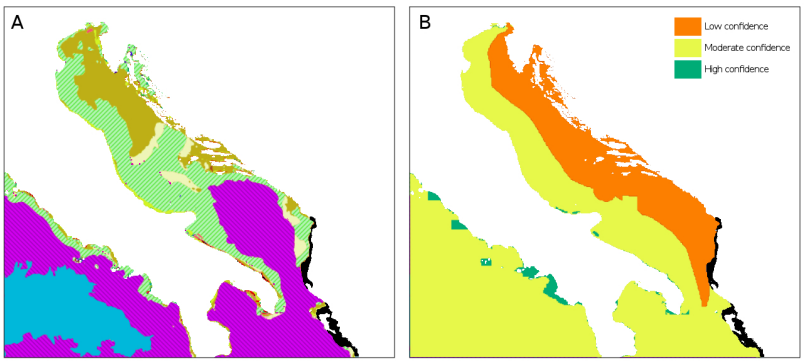


Figure SI-3: EUNIS original habitat map for the sector considered (A) and Confidence map for the habitat layer (B). Areas in black indicate portions of the map where habitat information was absent.

4. Parameters tuning: specific parametrization of the framework

Parameterizing spawning temporal variability

The spawning phenology and duration is specific for each species. *D. sargus* was found to spawn over a relatively short period (20 days) in spring (Figure SI-4a). In contrast, *D. vulgaris* spawns repeatedly over a longer period (around three months) in winter SI-4b. Following the prescriptions of Monroy et al. (2017), we use a spawning periodicity of one day for *D. sargus* while a spawning periodicity of four days is retained for *D. vulgaris*. Note that the full distribution of spawning dates for *D. vulgaris* revealed four outliers, which would significantly increase the range of spawning period but that are disregarded here due to their rareness (see Figure SI-4). Overall, we generate and analyze 21 numerical experiments for *D. sargus* and 22 experiments for *D. vulgaris*.

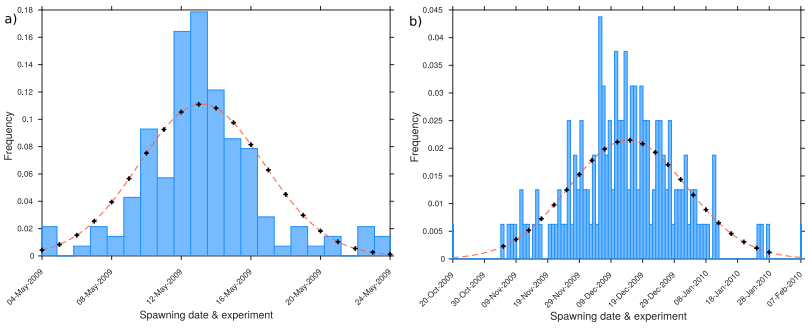


Figure SI-4: Frequency of occurrence (histogram) of the otolith-derived spawning dates for *D. sargus* (panel a) and for *D. vulgaris* (panel b). The fitted normal distributions (red lines) are constructed using the mean and standard deviation of the spawning dates of each species. The starting dates of our numerical experiments are indicated by black crosses.

Location node: dealing with nearshore model limitations

To account for the poorly resolved small-scale nearshore circulation, we attributed $H = 4$ location nodes over and in the vicinity of each sampling location while minimizing the land ratio of the selected nodes (Figure 2b). Preliminary tests (not shown) were performed using different number of location nodes and different geometrical shapes of the retained area. This "square form" with $H = 4$ location nodes was retained to fit the profile of the Apulian coastline and to

obtain a minimized averaged land ratio that is comparable among the seven locations (from north to south, the mean land ratio of location nodes is 0.08, 0.16, 0.1, 0.26, 0.05, 0.12, 0.11).

5. Quantitative dispersal diagnosis

The dispersal distances are calculated as the euclidean distances between the center of any source node and the barycenter of all location node centers. Then, all dispersal distances were classified into categorical bins. The probability associated to each bin corresponds to the summed probabilities of all source nodes whose dispersal distances fall within the bin edges. The median dispersal distance is given by the center of the bin (i.e. the mean of its two edges) associated to the first cumulative probability equal to or greater than 0.5. The maximal dispersal distance is given by the center of the last bin which contains the maximal euclidean distance (given by the source node the furthest away from the locations).

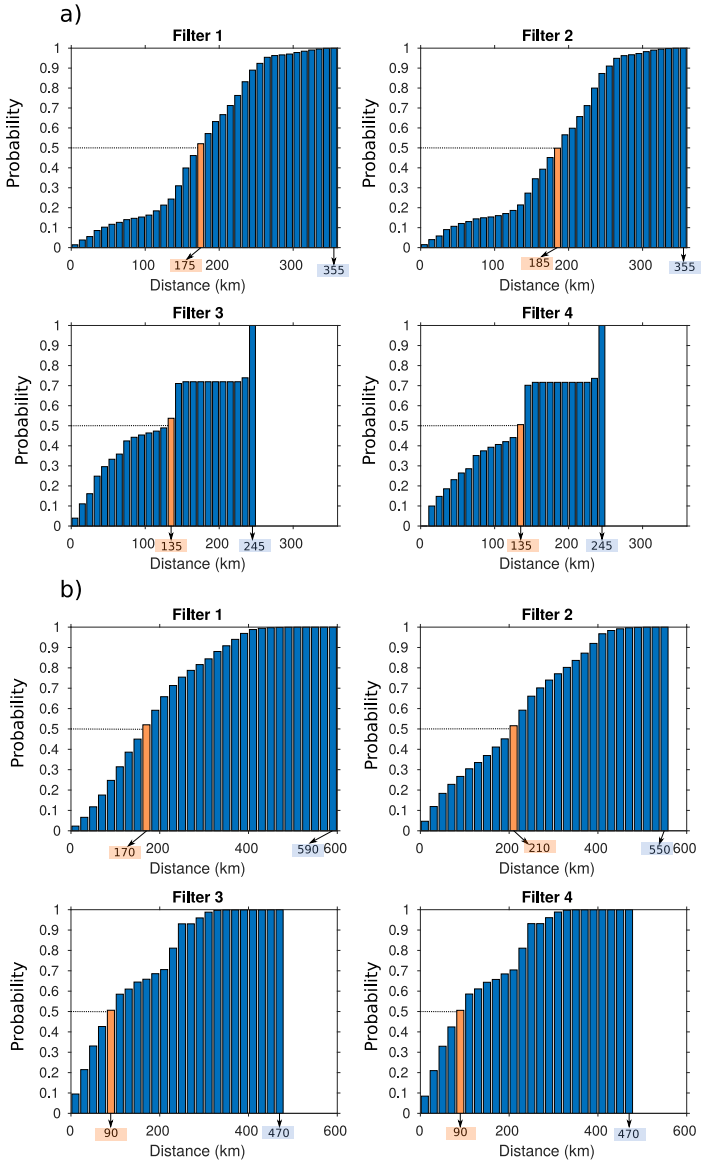


Figure SI-5: Cumulative histograms of dispersal distances obtained for the four consecutive filters (as indicated in the subtitle) for *D. sargus* (panel a) and *D. vulgaris* (panel b). The orange bars indicate the median distance bin.

6. Patterns and magnitude of connectivity as derived from the filter-1, filter-2 and filter-3

Figure SI-6 displays probabilities for each source node selected by the filter-1: it indicates the probability for that source node to be a supplier of larvae whose dispersal ends up in one of the seven sampled locations. Another interpretation is as follows: it gives the relative proportion of larvae supplied by that node source, among all selected sources (e.g. summing to 1), to the total pool of larvae received by sampled locations. For *D. sargus*, sources characterized by high probabilities, i.e. a high spawning potential (close to 0.01), are located along the Gargano promontory, next to the Apulian coast ($\sim$ 40.8 °N and 40.4 °N) and offshore ($\sim$ 42 °N, 17 °E; Figure SI-6a). Sources situated within the gulf of Manfredonia are characterized by low spawning potential (0.0001). For *D. vulgaris*, sources characterized by high spawning potential (0.001), are situated in the western middle Adriatic sub-basin and across the western and eastern south Adriatic sub-basin. Note that the core of the south Adriatic sub-basin is characterized by medium spawning potential ($\sim$ 0.0001). The northern Ionian Sea is characterized by medium to spawning potential, spanning 0.0001 to 0.00001.

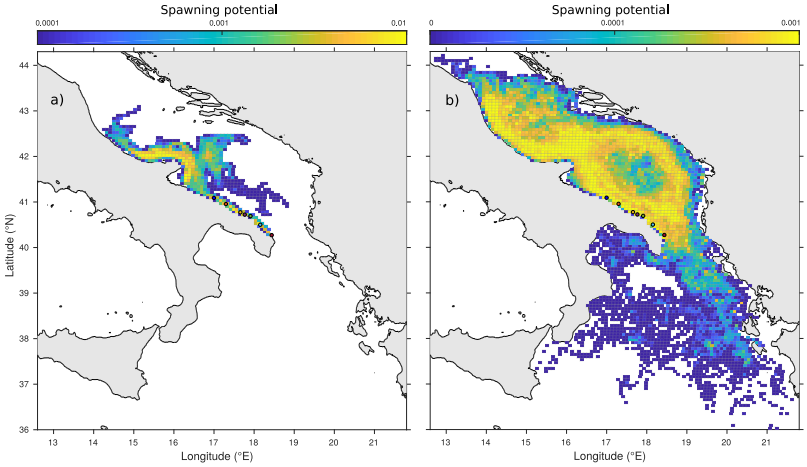


Figure SI-6: Quantitative patterns of connectivity for *D. sargus* (panel a) and for *D. vulgaris* (panel b): maps of the larval sources (after applying filter-1), along with their associated probabilities (the probability that larvae whose dispersal ended up in one of the seven sampled locations originated form that larval source) which are integrated here as the larval source's spawning potential.

Figure SI-7 displays spawning potential for each source node selected by the filter-2, i.e all sources whose environmental characteristics fulfill the favorable conditions for spawning to occur. Here it is a bathymetric filter selecting all nodes whose depths range 0-80 m (see section 2.6.3). For *D. sargus*, sources characterized by high spawning potential (close to 0.05) are located along the Gargano promontory (Figure SI-7a). Note that sources characterized by spawning potential (0.005) are situated next to the Apulian coast ($\sim$ 40.8 °N and 40.4 °N). For *D. vulgaris*, sources characterized by high spawning potential (close to 0.01) are located all along the Italian coastline, especially from $\sim$ 43.2 °N to 42 °N and along the Apulian coast, and around the Tremiti archipelago and the Palagruža island (42 °N, 17 °N, Figure SI-7b). Medium spawning potential are found next to the Croatian islands (from 43 °N to 42.5 °N).

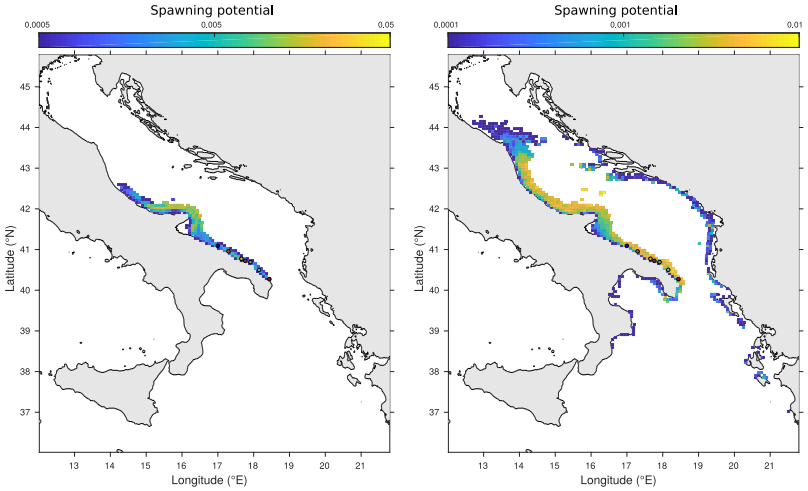


Figure SI-7: Quantitative patterns of connectivity for *D. sargus* (panel a) and for *D. vulgaris* (panel b): maps of the larval sources (after applying filter-2), along with their associated probabilities (the probability that larvae whose dispersal ended up in one of the seven sampled locations originated form that larval source) which are integrated here as the larval source's spawning potential.

Figure SI-8 displays spawning potential for each source node selected by the filter-3, i.e all sources which house adults preferential habitats (rocky reefs, Posidonia oceanica meadows and coralligenous formations in our case-studies, see section 2.6.3). For *D. sargus*, Sources situated next to the main islands of the Tremiti archipelago are characterized by high spawning potential (0.1, Figure SI-8a). Along the Apulian coast, two locals maximal (with probabilities ranging from 0.01 to 0.1) can be found at around 41.5 °N, 41.1 °N and 40.7 °N. The other sources have moderate to low spawning potential of the order of 0.001 to 0.01. For *D. vulgaris*, sources characterized by high spawning potential (0.1) are around the Tremiti archipelago and the Palagruža island (Figure SI-8a). Sources of intermediate spawning potential ($\sim$ 0.01) are found along the Apulian coast (from 41.2 °N to 40.2 °N) and next to the Croatian islands (from 43 °N to 42.5 °N).

When comparing the extents and magnitudes of larval sources over the gulf of Manfredonia for filter-1 and filter-2 (Figure SI-6,SI-7), *D. vulgaris* sources are

found until the inner gulf with a quite wide maxima of probabilities offshore, contrasting those of *D. sargus* displaying a thin local maxima that does not prevail inside the gulf. Indeed, during spring, the W-SAd current flows as a narrow intensified jet following closely the 100 m isobath (Figure SI-3a). It tends to create a transport barrier preventing exchange across the main current axis, thus isolating the inner gulf (Figure SI-6a,SI-7a). In contrast, a wider and more unstable W-SAd current during winter provides pathways for the sources located inside the gulf to connect with our locations (Figure SI-6b,SI-7b).

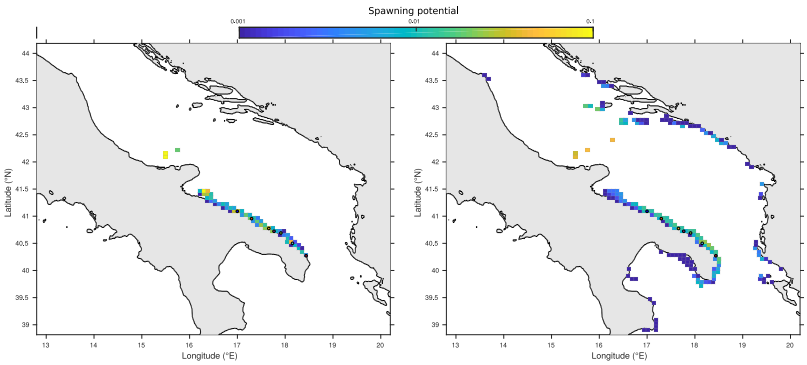


Figure SI-8: Quantitative patterns of connectivity for *D. sargus* (panel a) and for *D. vulgaris* (panel b): maps of the larval sources (after applying filter-3), along with their associated probabilities (the probability that larvae whose dispersal ended up in one of the seven sampled locations originated form that larval source) which are integrated here as the larval source's spawning potential.

7. Contributions of the current Marine Protected Areas network

Connected Marine Protected Areas (MPAs) are indexed following the MAPAMED database (`http://www.medpan.org/main_activities/mapamed/`) and without considering any specific legal status (i.e. marine national park, site of community importance or protected landscape). Note that Natura 2000 areas are not considered since they do not have fish protection as objective and have been showed to not deliver any direct benefit to fish populations (Guidetti et al., 2019).

Similarly to the association of four "location nodes" to each sampled location, we attribute to each MPA a number of representative nodes to account for the poorly resolved small-scale nearshore circulation. As such, a MPA is considered connected if a node source is located within a 20 km radius from the exact location of the MPA.

The proportion of larvae supplied by each connected MPA into each location is given in the Figure SI-9. These results revealed that TGMPA is both an important larval supplier of the sampled stretch of the Apulian coastline, itself included (Figure SI-9) due to elevated self-recruitment (Pujolar et al., 2013). Considering that MPA implementation should favor both network and self-persistence to sustain marine populations (Planes et al., 2009), our results conform with the view that TGMPA is well placed. This finding is even more interesting considering that TGMPA has been showed to be an highly effective MPAs in protecting fish populations (Giakoumi et al., 2017; Guidetti et al., 2014).

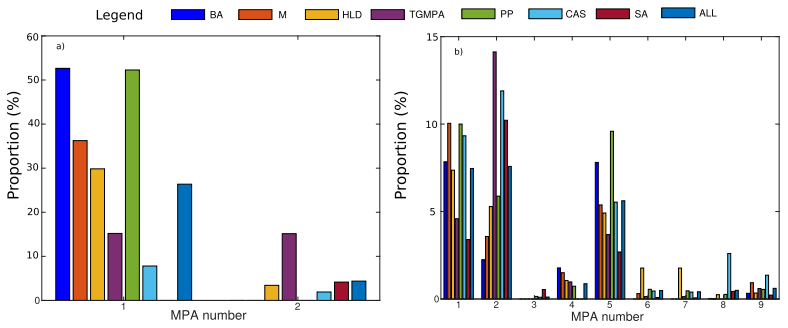


Figure SI-9: Proportions of larvae supplied by each connected MPA, as given by the last filter 4, into each location for a) *D. sargus* and b) *D. vulgaris*. Note that TGMPA is located within the Torre Guaceto MPA. Connected MPAs are indexed as follow: (1) Tremiti islands, (2) Torre Guaceto, (3) Porto Cesaro, (4) Pantana, (5) Lastovo archipelago, (6) Mljet, (7) Mali Ston Bay, (8) Lumi Buna-Velipoje and (9) Karaburun-Sazani Island.

8. Full description of the clustering procedure

Otolith geochemistry revealed that the number of clusters to be determined for *D. sargus* and *D. vulgaris* is respectively three and seven. In spite of the blind use of an automatic clustering algorithm, we apply an empiric grouping based on the criteria described in the section 2.5 of the manuscript.

The most constrained larval sources (i.e. obtained after applying our four consecutive filters) have a bimodal spatial distribution (see section 3.1 of the manuscript) with a consistent group of source nodes located along the Apulian coast whereas all the remaining sources are rather patchy. These isolated sources must be assigned to different clusters.

For both species, one cluster embraces the sources located around the Tremiti archipelago. For *D. vulgaris*, one cluster corresponds to the northernmost node source located near Ancona cap and three clusters are obtained by grouping all the node sources spread along the eastern shores. Consequently, the two remaining clusters to be defined for both species enclose all sources located along the Apulian coast. We must set the optimal latitudinal limits between these two Apulian clusters.

Starting with the simpler case of *D. sargus*, we indexed all the Apulian node sources by $D \in [1, 36]$, with $D = 1$ the southernmost source and $D = 36$ the northernmost source. Considering the cluster criteria, optimally setting the delimitation between these two clusters corresponds to finding the "boundary" node D which maximizes the fit between clusters probabilities and natal origin proportions derived from otoliths. We tested all the possible clustering options by shifting the boundary node D along the Apulian shores, one node after another while evaluating the fit between cluster probabilities and natal origin proportions.

To do so, we computed the Pearson correlation r between cluster probabilities and natal origin proportions for all locations (i.e. the seven locations are aggregated into one) and for each location. We also computed the corresponding regression coefficients β_1 and the y-intercepts β_2 (Figure SI-10).

The most adequate boundary node should maximize the correlation, that is the Pearson correlation coefficients r closest to 1 for all locations and for each location. It should also have a regression coefficient β_1 closes to 1 and a y-intercept β_2 close to 0. Those conditions are best fulfilled for node $D = 3$, and the 2 nodes surrounding it, while the correlations decrease tremendously when selecting other boundary nodes (Figure SI-10). We finally retain the delimitation between the two Apulian clusters at node $D = 3$, returning the best fit against otolith data.

We then apply the same boundary for *D. vulgaris* since the geochemical imprints of ambient waters left in otoliths is assumed comparable between both species.

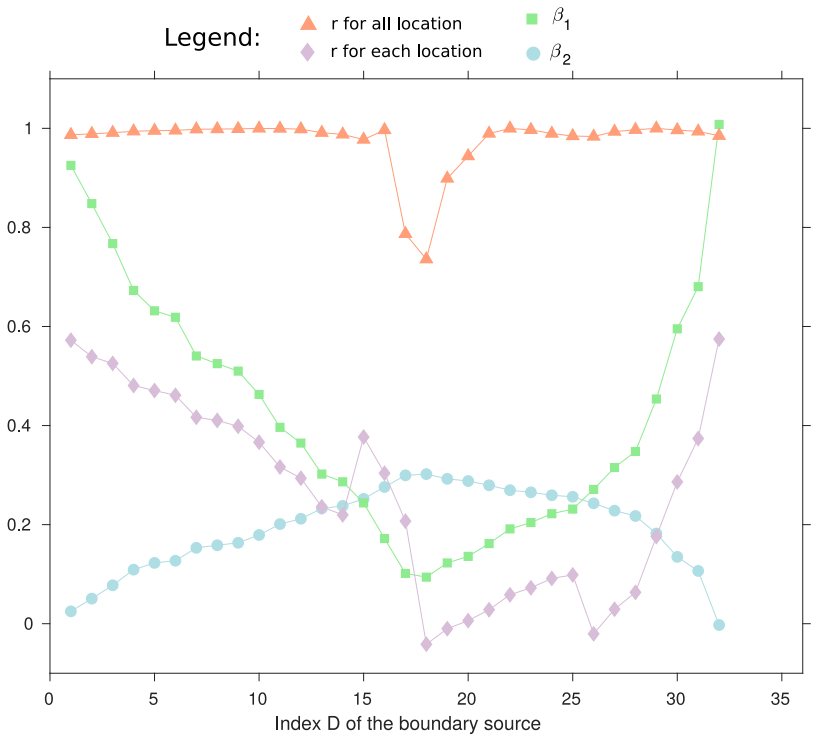


Figure SI-10: Pearson correlation r for all locations (triangle) and for each location (diamond); regression coefficient β_1 (square) and the y-intercept β_2 (circle) between cluster probabilities and natal origin proportions for all the possible clustering combinations (the value of D indicates the "boundary" node).

When considering each sampling location independently, it is worth noting that the correlations still hold for both species, despite slightly larger spread for *D. sargus*, as reported in Figure SI-11. Then, these reliable fits found for *D. vulgaris* when aggregating all locations (see Figure 5b), as well as for both species when treating each location (see Figure SI-11), suggest that the optimization procedure of the correlation between dispersal model and otoliths geochemistry, performed on *D. sargus*'s case only, is pertinent.

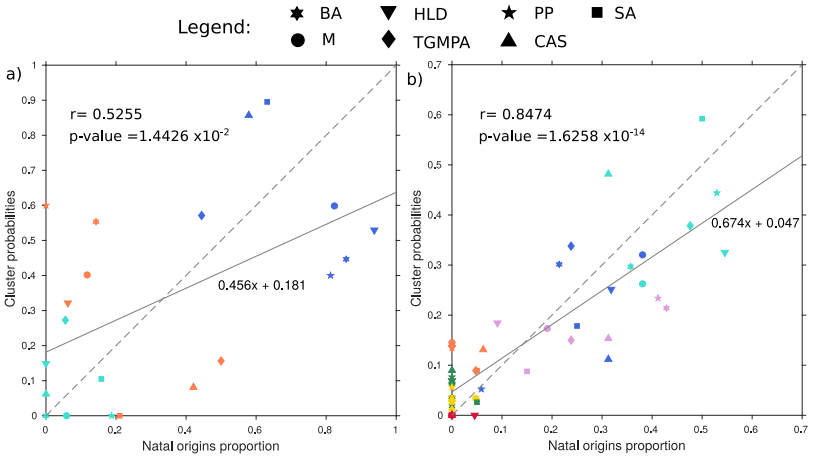


Figure SI-11: Correlations between natal origin proportions and cluster probabilities, considering each location independently, for a) *D. sargus* and b) *D. vulgaris*.

References

Ciliberti, S.A., Pinardi, N., Coppini, G., Oddo, P., Vukicevic, T., Lecci, R., Verri, G., Kumkar, Y., Creti, S., 2015. A high resolution Adriatic-Ionian Sea circulation model for operational forecasting, in: EGU General Assembly Conference Abstracts.

Courant, R., Friedrichs, K., Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. 100, 32–74. URL: https://doi.org/10.1007/BF01448839, doi:10.1007/BF01448839.

Di Franco, A., Calò, A., Pennetta, A., De Benedetto, G., Planes, S., Guidetti, P., 2015. Dispersal of larval and juvenile seabream: Implications for Mediterranean marine protected areas. Biological Conservation 192, 361–368.

Di Franco, A., Gillanders, B.M., De Benedetto, G., Pennetta, A., De Leo, G.A., Guidetti, P., 2012. Dispersal Patterns of Coastal Fish: Implications for Designing Networks of Marine Protected Areas. PLoS ONE 7, e31681. URL: http://dx.plos.org/10.1371/journal.pone.0031681, doi:10.1371/journal.pone.0031681.

Di Franco, A., Guidetti, P., 2011. Patterns of variability in early-life traits of fishes depend on spatial scale of analysis. Biology Letters 7, 454–456.

Di Franco, A., Qian, K., Calò, A., Di Lorenzo, M., Planes, S., Guidetti, P., 2013. Patterns of variability in early life traits of a Mediterranean coastal fish. Marine Ecology Progress Series 476, 227–235.

Dubois, M., Rossi, V., Ser-Giacomi, E., Arnaud-Haond, S., López, C., Hernández-García, E., 2016. Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems. Global ecology and biogeography 25, 503–515.

Furlani, S., Pappalardo, M., Gómez-Pujol, L., Chelli, A., 2014. The rock coast of the Mediterranean and Black seas. Geological Society, London, Memoirs 40, 89–123.

Giakoumi, S., Scianna, C., Plass-Johnson, J., Micheli, F., Grorud-Colvert, K., Thiriet, P., Claudet, J., Carlo, G.D., Franco, A.D., Gaines, S.D., García-Charton, J.A., Lubchenco, J., Reimer, J., Sala, E., Guidetti, P., 2017. Ecological effects of full and partial protection in the crowded Mediterranean Sea: a regional meta-analysis. Scientific Reports 7, 8940. URL: https://www.nature.com/articles/s41598-017-08850-w, doi:10.1038/s41598-017-08850-w.

Guidetti, P., Addis, P., Atzori, F., Bussotti, S., Calò, A., Cau, A., Culioli, J.M., Lucia, G.D., Franco, A.D., Lorenzo, M.D., Follesa, M.C., Gazale, V., Massaro, G., Mura, F., Navone, A., Pala, D., Panzalis, P.A., Pusceddu, A., Ruiu, A., Cau, A., 2019. Assessing the potential of marine Natura 2000 sites to produce ecosystem-wide effects in rocky reefs: A case study from Sardinia Island (Italy). Aquatic Conservation: Marine and Freshwater Ecosystems 29, 537–545. URL: http://onlinelibrary.wiley.com/doi/abs/10.1002/aqc.3026, doi:10.1002/aqc.3026.

Guidetti, P., Baiata, P., Ballesteros, E., Franco, A.D., Hereu, B., Macpherson, E., Micheli, F., Pais, A., Panzalis, P., Rosenberg, A.A., Zabala, M., Sala, E., 2014. Large-Scale Assessment of Mediterranean Marine Protected Areas Effects on Fish Assemblages. PLOS ONE 9, e91841. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0091841, doi:10.1371/journal.pone.0091841.

Harmelin-Vivien, M., Harmelin, J., Leboulleux, V., 1995. Microhabitat requirements for settlement of juvenile sparid fishes on Mediterranean rocky shores. Hydrobiologia 300, 309–320.

Lenfant, P., Planes, S., 1996. Genetic differentiation of white sea bream within the Lion's Gulf and the Ligurian Sea (Mediterranean Sea). Journal of Fish Biology 49, 613–621.

Macpherson, E., Raventos, N., 2006. Relationship between pelagic larval du-

ration and geographic distribution of Mediterranean littoral fishes. Marine Ecology Progress Series 327, 257–265.

Madec, G., others, 2015. NEMO ocean engine .

Monroy, P., Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., 2017. Sensitivity and robustness of larval connectivity diagnostics obtained from Lagrangian Flow Networks. ICES Journal of Marine Science 74, 1763–1779.

Oddo, P., Adani, M., Pinardi, N., Fratianni, C., Tonani, M., Pettenuzzo, D., others, 2009. A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting. Ocean science .

Oddo, P., Pinardi, N., Zavatarelli, M., Coluccelli, A., 2006. The Adriatic basin forecasting system. Acta Adriatica: international journal of Marine Sciences 47, 169–184.

Planes, S., Jones, G.P., Thorrold, S.R., 2009. Larval dispersal connects fish populations in a network of marine protected areas. PNAS , pnas.0808007106URL: https://www.pnas.org/content/early/2009/03/20/0808007106, doi:10.1073/pnas.0808007106.

Pujolar, J.M., Schiavina, M., Di Franco, A., Melià, P., Guidetti, P., Gatto, M., De Leo, G.A., Zane, L., 2013. Understanding the effectiveness of marine protected areas using genetic connectivity patterns and Lagrangian simulations. Diversity and Distributions 19, 1531–1542.

Rossi, V., Ser-Giacomi, E., López, C., Hernández-García, E., 2014. Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. Geophysical Research Letters 41, 2883–2891.

Ser-Giacomi, E., Rodríguez-Méndez, V., López, C., Hernández-García, E., 2017. Lagrangian Flow Network approach to an open flow model. Eur. Phys. J. Spec. Top. 226, 2057–2068. URL: https://doi.org/10.1140/epjst/e2017-70044-2, doi:10.1140/epjst/e2017-70044-2.

Ser-Giacomi, E., Rossi, V., López, C., Hernandez-Garcia, E., 2015a. Flow networks: A characterization of geophysical fluid transport. Chaos: An Interdisciplinary Journal of Nonlinear Science 25, 036404.

Ser-Giacomi, E., Vasile, R., Hernandez-Garcia, E., Lopez, C., 2015b. Most probable paths in temporal weighted networks: An application to ocean transport. Physical Review E 92. URL: http://arxiv.org/abs/1411.6902, doi:10.1103/PhysRevE.92.012818. arXiv: 1411.6902.

Tonani, M., Pinardi, N., Dobricic, S., Pujol, I., Fratianni, C., 2008. A high-resolution free-surface model of the Mediterranean Sea. Ocean Science 4, 1–14.

Waugh, D.W., Abraham, E.R., Bowen, M.M., 2006. Spatial variations of stirring in the surface ocean: A case study of the Tasman Sea. Journal of Physical Oceanography 36, 526–542.

A.2. Insights into the spatio-temporal variability of spawning in a territorial coastal fish by combining observations, modelling and literature review

ORIGINAL ARTICLE

Insights into the spatio-temporal variability of spawning in a territorial coastal fish by combining observations, modelling and literature review

Marine DI STEFANO[1] | Térence LEGRAND[1] | Antonio DI FRANCO[2] | David NERINI[1] | Vincent ROSSI[1]

[1]Mediterranean Institute of Oceanography (UM 110, UMR 7294), CNRS, Aix Marseille Univ., Univ. Toulon, IRD, 13288, Marseille, France

[2]Stazione Zoologica Anton Dohrn, Dipartimento Ecologia Marina Integrata, Sede Interdipartimentale della Sicilia, Lungomare Cristoforo Colombo (complesso Roosevelt), 90149 Palermo, Italy

Correspondence
Marine DI STEFANO, Mediterranean Institute of Oceanography (UM 110, UMR 7294), CNRS, Aix Marseille Univ., Univ. Toulon, IRD, 13288, Marseille, France
Email: marine.distefano@mio.osupytheas.fr & vincent.rossi@mio.osupytheas.fr

In bipartite life-cycle fishes, spawning represents the onset of propagules dispersal, with eggs and larvae experiencing anisotropic transport and high mortality rates, before eventually metamorphosing and settling. Hence, early-life stages operate as bottlenecks for population demography by strongly constraining recruitment. Despite its significance, spawning is rarely explicitly considered in ecosystem management due to a lack of knowledge, for many species, about where and when spawning occurs. Previous evidences suggest that temperature is among the main drivers of spawning in Teleosts. Using the ecologically and economically relevant white seabream *Diplodus sargus* in the central Mediterranean Sea as a case study, we assess the abiotic factors that regulate the onset and duration of spawning and subsequent dispersal. Lagrangian backtracking simulations fed with early-life observations allow locating 11 spawning events, which are then associated to simulated temperatures ranging from 14.8 to 20.6°C, in close agreement with previous estimates. Based on this range of suitable temperature, we model the spatio-temporal variability of spawning success at broad-scale over 10 years (2005-2014) following the backtracking approach with hypothetical constant settlement areas. It highlights a prominent inter-annual variability in the Adriatic and Siculo-Tunisian strait driven by oceanographic processes. Moreover, a powerful clustering method uncovers relatively stable spawning areas in the Ligurian and Tyrrhenian seas with both early (January to Mid-February) and late (April to June) spawning peaks. Our methodology can be applied to other species and oceanic systems to investigate how oceanic processes impact spawning success, enabling the design of sound management strategies.

KEYWORDS

Spawning event, larval dispersal, marine connectivity, Mediterranean sea, Diplodus sargus, Lagrangian Flow Network, ecosystem management, lagrangian backtracking model, oceanic variability

1 | INTRODUCTION

The Mediterranean is one of the most over-exploited sea in the world with about 75 percent of the assessed fish stocks being currently overfished, mostly by large scale fisheries (Tsikliras et al., 2015; FAO, 2018). Moreover, among coastal species, which are essentially targeted by small scale and recreational fisheries, many are currently not assessed despite declining catches (Lloret & Planes, 2003). While local conservation initiatives (e.g. Marine Protected Areas, MPAs) have been shown to have positive effects on such coastal species locally (Sala et al., 2012; Abecasis et al., 2015; Di Franco et al., 2016; Giakoumi et al., 2017), their large-scale protection and management remain challenging because of complex life cycles and high environmental variability (Cadrin & Secor, 2009). Due to the current lack of information to design sound management strategies and fisheries restrictions for coastal fish at basin-scale, there is an urgent need to develop research and management for a more sustainable exploitation of fish stocks in this semi-enclosed and densely populated sea.

In this context, determining the location and onset of spawning now represents a key element for fishery science (Sadovy de Mitcheson, 2016). Indeed, spawning aggregations are often targeted by fisheries as it represents relatively easy and lucrative catches in a first place, but ultimately lead to a rapid decrease in stocks and to long-term reduction of fishery activities (Sadovy & Domeier, 2005). Moreover, spotting spawning areas is crucial to appraise the dispersal of propagules (i.e. eggs and larvae) and hence understand population connectivity (Bauer et al., 2014; Dubois et al., 2016; Calò et al., 2018; Legrand et al., 2019), especially for bipartite life-cycle for which the spatial scales of ecological connectivity are mainly determined by the dispersal of early-life stages by currents across the seascape (Nathan et al., 2003; Gaines et al., 2007; Burgess et al., 2016). These early-life stages start with the spawning process, when fish aggregates and gametes are released into the water column, then eggs hatch into larvae which, after a first phase mostly passive, will acquire the capability to actively swim, just before metamorphosing in juveniles and settling in coastal habitats (Thresher et al., 1989). Enormous amounts of propagules are generally assumed to have the potential to disperse over large distances, spanning tens to hundreds of kilometres (Cowen et al., 2006) while exhibiting high mortality (Pineda et al., 2007). Nevertheless, by introducing migrants into a population, dispersal can impact growth and mortality rates, a process called demographic connectivity (Cowen & Sponaugle, 2009). Demographic connectivity is fundamental in fishery and conservation science as it has been shown to influence dynamics and persistence of spatially structured populations (Burgess et al., 2014; Hidalgo et al., 2019). Few management tools have already been considering demographic connectivity explicitly, thus limiting their efficiency (Fogarty & Botsford, 2007) and pointing out the necessity of a change in the management approach to a more sustainable exploitation of marine living resources (Carr et al., 2019).

However, the location and timing of spawning events and their eventual stability/variability in space and time remain major unknown of the fish life cycle (Calò et al., 2018). Since field observations of spawning events are very

sparse and fragmentary, they do not provide sufficient quantitative information about when and where spawning occurs. Moreover, monitoring the spawning activity of a given fish species over broad-scale would be overly costly, if not impossible. As such, indirect methodologies must be considered. One approach, which we follow here, consists in using field observations of just-settled juveniles (i.e. post-settlers) as indirect evidence of the successful occurrence of one or multiple spawning events. Then backtracking larvae trajectories from settlement sites to putative spawning areas and identifying the abiotic conditions that prevail during these spawning events may help gain insights into which abiotic conditions favour spawning. By extrapolating these relationships and thanks to constant monitoring of abiotic conditions, one could then predict future events and use this information to design sound management strategies enabling the protection of these propagules' sources, contributing to the replenishment of both local and distant sub-populations (Pelc et al., 2010; Di Franco et al., 2012; Erisman et al., 2015; Pittman & Heyman, 2020). Development of hydrodynamical models backtracking larvae from settlement areas to potential spawning areas is actively trying to fill this gap (Bauer et al., 2014; Calò et al., 2018; Legrand et al., 2019). Combining these models with the variability of environmental factors, known to control the onset of fish spawning in favourable areas, could improve the simulation of the spawning process and thus enhance scientific-based knowledge of this event (Werner et al., 2007; Huret et al., 2010).

Among all potentially relevant abiotic factors, previous findings suggest seawater temperature playing a key role as a trigger of fish spawning, both in freshwater (King et al., 2016) and in marine species (Gill et al., 1996; Winters & Wheeler, 1996; Hereu et al., 2006). In the Mediterranean sea, temperature as a spawning cue has been already reported for Sparids: changing temperatures seem to trigger the onset of spawning for this family (Gonçalves, 2000; Pajuelo et al., 2003; Pajuelo et al., 2006; Pajuelo et al., 2008; Mouine et al., 2011). Here, we focus on a demersal fish of the Sparid family, the white sea bream *Diplodus sargus* (Linnaeus, 1758) as it is a relatively data-rich case study in the Mediterranean. *D. sargus*, is known to spawn once a year (Mouine et al., 2007). More precisely, the rise of seawater temperature from the winter minimum throughout spring has been related with the onset and the duration of the spawning period for this species (Morato et al., 2003; Mouine et al., 2007, 2012; Di Lorenzo et al., 2014; Potts et al., 2014; Aspillaga et al., 2016). The white seabream is a keystone species as it feeds on sea urchins *Paracentrotus lividus*, and regulate their grazing pressure on algae, therefore potentially controlling regime shifts (from vegetated to bare rocky coastal habitats) in the Mediterranean ecosystem (Hereu et al., 2005). Moreover, this species is economically relevant in the Mediterranean Sea as targeted by both recreational and professional fisheries (Sala et al., 1998), despite the lack of stock assessment for this species (FAO, 2006).

Our objectives here are to understand how abiotic factors influence the timing and location of spawning events for *D. sargus* in order to appraise the locations and periods that are potentially favourable for spawning, hence delivering useful inputs to design sound management and conservation strategies. We first determine a range of temperatures that trigger spawning by building a backward-in-time particle tracking model fed with early life traits and settlement locations for the investigated species. Based on this range, we use multi-year outputs of a regional ocean model to simulate and analyse the inter-annual variability / stability of dispersal and spawning over ten recent years at large scale. We finally discuss our results against the published literature and interpret them in the context of dynamic spatio-temporal management.

2 | MATERIAL AND METHODS

2.1 | Modelling the dispersive phases using Lagrangian Flow Networks

The Lagrangian Flow Network (LFN) model, which combines network theory tools and a particle-tracking model, has been successfully applied to study marine connectivity (Rossi et al., 2014; Dubois et al., 2016; Hidalgo et al., 2019; Legrand et al., 2019; Rossi et al., 2020) thanks to robust theoretical foundations (Ser-Giacomi et al., 2015) and comprehensive sensitivity studies (Monroy et al., 2017). It simulates the dispersal of fish early-life stages (eggs, larvae) by calculating passive horizontal trajectories at a given depth for predetermined starting times and durations.

Here the LFN is coupled offline with an hydrodynamical model, the Adriatic-Ionian REGional configuration (AIREG; Oddo et al., 2006; Ciliberti et al., 2015, Figure 1), which is based on the NEMO kernel (Madec & the NEMO team, 2008). We use the high-resolution (1/45°) AIREG outputs that consist in daily three-dimensional velocity and temperature fields on 121 depths for years ranging from 2005 to 2014. The LFN first discretizes the AIREG domain in a non-regular mesh of equal-area boxes, called nodes, and computes a land ratio for each of them (e.g. a node with a land ratio of 0/1 is considered fully covered by sea/land, respectively). The node size is set to a horizontal resolution of 1/16° to properly consider the typical intricate coastlines and small islands found in the full AIREG domain. In total, the LFN builds 48 839 nodes, which are quasi-squared boxes of about 6 km width and a surface of around 40 km². Just-released eggs and subsequent larvae are represented as passive particles drifting at a fixed depth of about 10 metres, following the few available evidence for Sparid larvae dispersal (Olivar & Sabatés, 1997). While changing in buoyancy or ontogeny of larvae could potentially modify vertical positioning of larvae of some species, hence modifying trajectories (Pineda et al., 2007), no study has documented this effect for *D. sargus* yet. Moreover, small vertical displacements (e.g. tens of meters) within the surface mixed layer (which ranges from 30 to 100 m or more in the Mediterranean Sea) would return similar trajectories as modelled near-surface currents are mostly barotropic. Another potentially relevant biological behaviour occurs at the end of the larval phase, when competent larvae gain swimming abilities to move toward more suitable settlement sites located 1 or 2 km away (Clark et al., 2005; Staaterman et al., 2012; Faillettaz et al., 2018). Our model indirectly includes the effect of such behaviour since such small nearshore movements are clearly imperceptible in nodes of 40 km². In addition, note that this coarse-graining procedure allows the consideration of adult movements in our simulations since the home range of adult individuals is, on average, smaller than 1 km² (Di Lorenzo et al., 2014; Aspillaga et al., 2016) while individuals can display longer distance movement (for 1 or 2 km) during spawning. For each numerical experiment, about 100 Lagrangian particles (for full-sea node; otherwise slightly less, proportionally to the land ratio) are homogeneously seeded in each node, representing around 2.5 million particles for the entire AIREG domain. The Runge-Kutta time step used to integrate Lagrangian trajectories is 30 min, fulfilling the Courant-Friedrichs-Lewy condition (Courant et al., 1928; Legrand et al., 2019). Each experiment is characterized by specific starting dates and particle-tracking durations, which are chosen according to *D. sargus* spawning dates and Pelagic Larval Durations (PLD, here assumed to be the time comprised between spawning and settlement) determined through otoliths' reading (Di Franco et al., 2011; Di Franco and Guidetti, 2011; see Table S1).

Finally, the LFN builds a connectivity matrix that encodes the particles exchanged between any pairs of nodes and retained within them. Such a connectivity matrix contains all the dispersal information: each element of the raw matrix is the number of particles exchanged between the departure node, corresponding to the row, and the arrival node, corresponding to the column. Therefore, the diagonal element of the matrix represents the retention of particles

in each node.

2.2 | Identification of successful spawning areas and determination of favourable temperatures using observation-constrained simulations

The first part of the study consists in spatially delimiting successful spawning areas associated with settlement areas previously identified and for which specific information on early life traits is available (Di Franco et al., 2011; Di Franco & Guidetti, 2011). To do so, we track early-life stages from their settlement sites back to their spawning areas (Legrand et al., 2019) using our dispersal model and fine-tuned parameters following the biological observations collected at each site (see Table S1 and Figure S2). Then, we determine and compile all the modelled temperatures found within these discrete spawning areas over well-defined temporal periods. This range of spawning suitable temperatures is finally used as a parameter of the second modelling exercise (see following section 2.3).

Along the Italian coasts, 10 *D. sargus* juveniles per site were caught at 4 sites in June 2008 and 20 juveniles per site at 7 sites in June 2009 (Di Franco et al., 2011; Di Franco and Guidetti, 2011; Figure 1). Since the timing of the sampling have been chosen in accord with the supposed settlement peak of each region (Di Franco et al., 2011; Di Franco & Guidetti, 2011), these observations should provide reliable information about the maximum spawning activities associated with each site and year. For each site, we derive a range of spawning dates (minimum and maximum) obtained by keeping 95 percent of a normal probability distribution adjusted on the spawning dates obtained thanks to otoliths' reading (see Table S1). We consider this range of spawning dates as the temporal duration of that distinct spawning event: as such, we run as many daily LFN experiments (characterized by a connectivity matrix attributed to each spawning date) as the number of days comprised within this period. For the LFN tracking-time, we retain the median of all observed PLDs (see Table S1) that have been also determined by otolithometry (Di Franco et al., 2011; Di Franco & Guidetti, 2011).

Based on the analytical framework developed by Legrand et al., 2019, the full dispersal information contained in the raw connectivity matrices must be pre-processed for our purpose. The first filter is applied on arrival nodes (note that in backward-in-time dynamics, they act as departure nodes) to consider only the surveyed settlement sites. To account for the coastal processes non-resolved by the AIREG model and to limit the impacts of "beaching" artefacts, we retain here four nodes encompassing each sampled site and its vicinity (note that the mean land ratio of the retained four nodes is low and similar between sampling sites). The second filter, applied on departure nodes (becoming the arrival nodes in backward-in-time), simulates the suitable habitat of low-moving adult stage of *D. sargus*. Adults are generally associated to rocky substrate and *Posidonia oceanica* meadows (Guidetti, 2000) and have been observed at least until 80 metres depth during the spawning period (Aspillaga et al., 2016; Giacalone et al., 2018). The filter is thus constructed by keeping all nodes that encompass those substrates and a conservative bathymetric limit that is shallower than 100 metres depth (see Figure 1). To do so, we exploit the hand-made corrected habitat maps published in Legrand et al., 2019 and the ETOPO1 Global Relief Model (National Geophysical Data Center, 2021). After applying both filters, each element of the filtered matrices represents the number of particles sent by the effective spawning habitat toward the surveyed settlement site.

The elements of such matrices can be readily transformed into dispersal probabilities thanks to a column-normalization of the number of particles, hence approximating backward-in-time dynamics (Rossi et al., 2014; Ser-Giacomi et al., 2015; Legrand et al., 2019). After the spatial filtering and the column-normalization, all daily matrices (i.e. 6 to 31 ma-

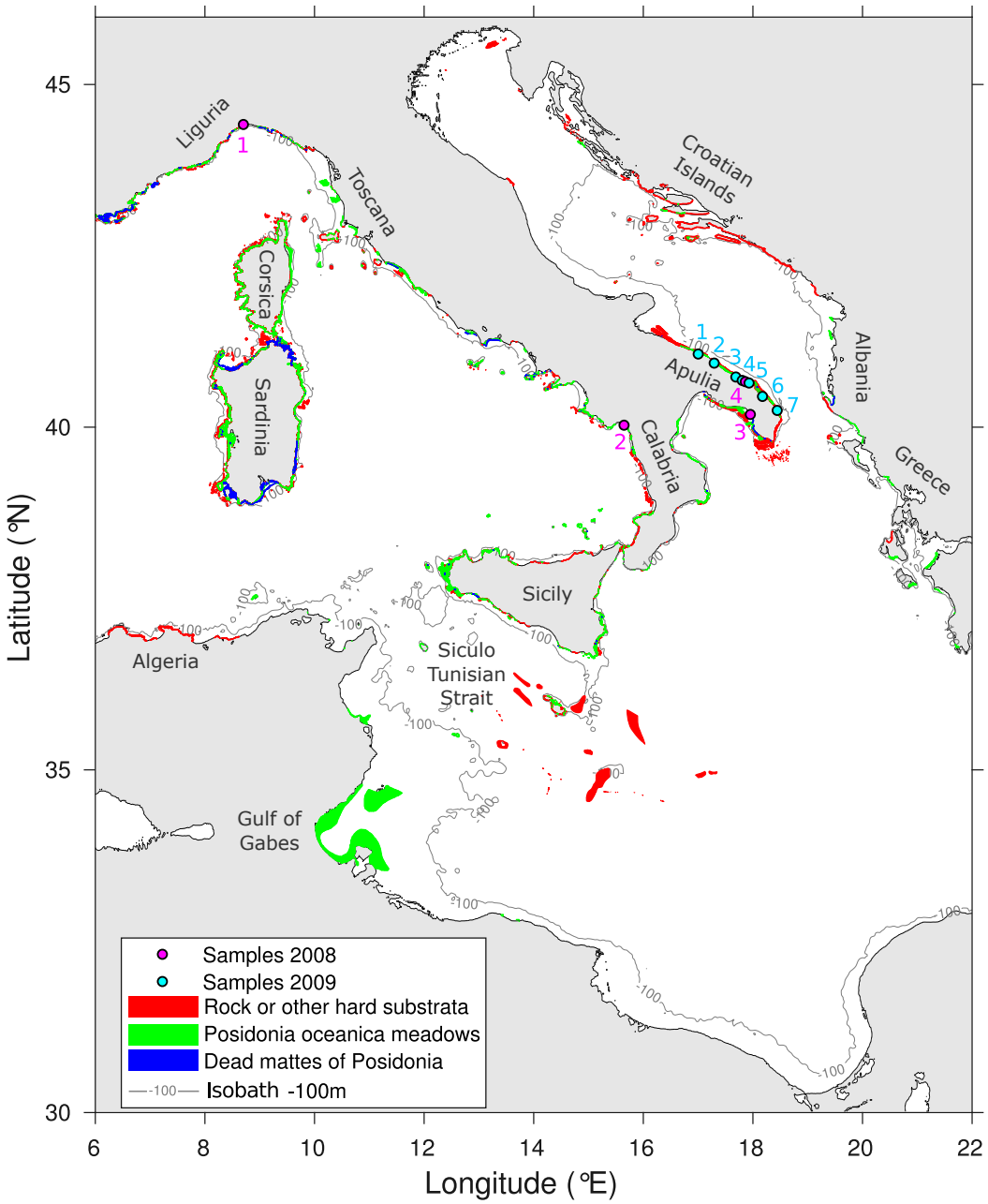


FIGURE 1 Seabed substrates downloaded from the EMODnet portal (European Marine Observation and Data Network, EMODnet, 2013) and 100 meters isobath (dark thin line) used to determine the spawning habitat in section 2.2 and the spawning and settlement habitat in section 2.3, in accord with the literature concerning *D. sargus* (Guidetti, 2000; Di Lorenzo et al., 2014; Aspillaga et al., 2016; Giacalone et al., 2018). The sites of 2008 are in pink colour: 1. Genova, 2. Maratea, 3. San Isidoro, 4. Torre Guaceto. The sites of 2009 are light blue: 1. Bari, 2. Monopoli, 3. Hotel La Darsena, 4. Torre Guaceto MPA, 5. Punta Penne, 6. Casalabate, 7. San Andrea. Those sites represent the settlement habitat in section 2.2.

trices, depending on the site considered) representative of a given surveyed site are weighted averaged into a unique composite matrix to obtain the total probability of all spawning nodes to send larvae into the surveyed settlement site during the entire spawning period (spanning 6 to 31 days). More specifically, the weight attributed to each daily matrix is determined according to the position of the corresponding spawning date in the Gaussian distribution fitted on the total spawning period of each site (see Figure S1 and Table S1). Only non-zero elements of the composite matrix are kept, meaning that the retained spawning areas effectively deliver larvae to the settlement site. Overall, we delineate 11 spawning events composed of several nodes representing the spawning areas during their corresponding spawning period for each of the 11 surveyed settlement sites.

Finally, we gather all the temperature values simulated by the AIREG model (Oddo et al., 2006; Ciliberti et al., 2015) for depths ranging from 0 to 50 metres within each previously identified spawning area for the entire effective spawning period. All those temperatures values are compiled into a single dataset while taking into account the larval contributions of each spawning node (e.g. dispersal probabilities are interpreted as proportion of larvae) to the settlement sites they replenished. Quantile values are then estimated using a kernel density estimation method (Simonoff, 1996). It allows retaining a 95% temperature interval that controls the spawning outbreak of *D. sargus* in the north western Mediterranean Sea. These favourable seawater temperatures are then used as a threshold for the broad-scale modelling exercise of spawning success (Figure 2; see following section 2.3).

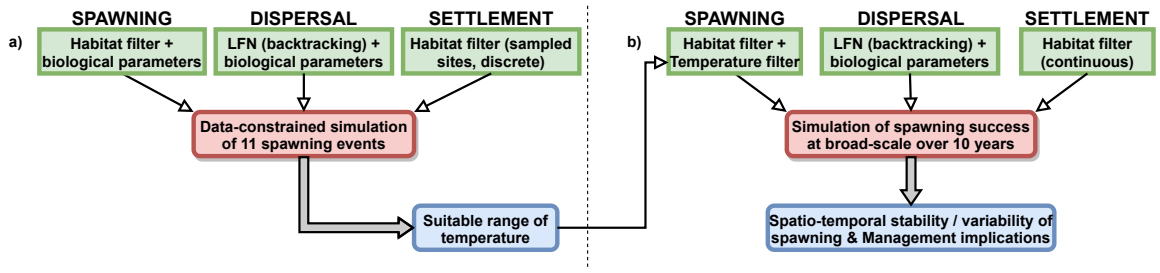


FIGURE 2 Schematic representation of the modelling approach for a) the data-constrained simulation and b) the inverse simulations of spawning success at broad-scale. Green boxes identify the inputs of the numerical experiments represented by the red boxes. Blue boxes represent the analysed outputs of the study. A more detailed version of this figure is available in Supplementary Material (see Figure S2).

2.3 | Inverse simulations of spawning success at broad-scale

Building on the range of temperature suitable for spawning, determined through a combination of modelling and field observations (see previous section 2.2), we now investigate the spatio-temporal variability of spawning over the entire model domain over 10 years (2005 to 2014). In our inverse approach that considers non-varying habitats, the sources of variability affecting early-life stages are ocean currents, which control the backward-in-time dispersal from settlement to spawning areas, and seawater temperature, which govern where and when spawning occur.

We performed LFN experiments every five days (Monroy et al., 2017) from the 5th of January to the 29th of June each year from 2005 to 2014 (available data of AIREG model; Oddo et al., 2006; Ciliberti et al., 2015). Consequently, we compute 36 connectivity matrices per year. The conservative choice of the spawning period relies on the information gathered from the literature (Table 1), revealing that *D. sargus* spawning may occur as early as January and until

June. Since Monroy et al., 2017 showed little to none effects of small PLD changes, we fix the PLD to 16 days, equal to the median PLD from individuals collected at all settlement sites (see Table S1; Di Franco et al., 2011; Di Franco and Guidetti, 2011). Note that some output files of the AIREG model are corrupted; as such, a few connectivity matrices (10/01, 15/01, 20/01, 25/01 and 30/01/2012) cannot be modelled so that year 2012 has only 31 matrices.

For each connectivity matrix, we apply the same habitat filter (e.g. *Posidonia oceanica*, rocks and maximum 100 metres depth, see section 2.2 and Figure 1) on both departure (i.e. spawning areas) and arrival (i.e. settlement areas) nodes to consider all possible spawning and settlement connections across the full domain (Figure 2). The habitat filter used for settlement is the same as the spawning one whereas, in nature, literature specifies that juveniles recruit between 0 and 2 metres deep in sandy bays and rocky shores (Harmelin-Vivien et al., 1995; Lenfant & Planes, 1996). However, there is no continuous and reliable information about marine micro-habitats at basin-scale. Moreover, the spatial resolution of our approach is too coarse and the node surface too wide to allow distinguishing settlement sites at this scale. Therefore, based on data constraints we operate this choice to perform a filter of the potential settlement sites based on depth and habitat. Then, we sum all the columns of each matrix into a vector (i.e. all the settlement nodes are now represented as one global settlement area). Each element of this vector gives the total number of particles sent by a given spawning node to the total settlement area over one spawning event. This vector is column-normalized to transform its elements into backward-in-time probabilities of "physical connectivity", co-varying only with the oceanic circulation. They are interpreted as the probabilities for any node pertaining to the spawning habitat to send larvae into any of the favourable settlement sites for a given spawning event.

Then, for each non-zero element of a vector (meaning a physical connection exists between the corresponding pair of spawning and settlement nodes), we evaluate if surface and subsurface temperatures (from 0 to 50 metres) encompassed in a given spawning node fall into the reference range of spawning favourable temperatures. More specifically, we determine the "thermal niche" probability as follows: a kernel density estimation has been fitted over the temperature distribution of each node and integrated between the bounds of the reference thermal range (i.e. we calculate the surface under the estimator, bounded by the temperature range, see Figures S2 and S3).

Finally, the "successful spawning probability" for each node and for the 2005-2014 period/each spawning event is obtained by multiplying the "physical connectivity" probability vector and the "thermal niche" probability vector (Figure 2), considered independent (not shown). For each year over 2005-2014, we thus compute 36 vectors of probability of successful spawning (31 vectors for 2012), each of them representing a spawning event. Note that the range of successful spawning probabilities depends on the total number of nodes of the studied domain. Hence, our subsequent analyses disregard absolute probabilities but rather focus on highlighting and interpreting the relative differences among node probabilities. Note also that our "successful spawning probabilities" comprehend how ocean temperatures control spawning and how ocean currents influence dispersal while it neglects the time and space variability of early-life stages natural mortality.

These vectors are further exploited by first averaging all events and all years and computing associated standard deviations to highlight spatial differences between regions potentially due to interannual or seasonal variability (see Figure 6). To dig into the origins of this variability, the "successful spawning probability" vectors are also exploited by computing seasonally averages and associated standard deviations to observe if the variability is seasonal or if there are interannual differences in the spawning process according to the season and the region (see Figures S4 and S5). Moreover, the "successful spawning probability" vectors are yearly averaged and their standard deviations computed to highlight spatio-temporal patterns between regions, according to the year without considering the seasonal aspect

(see Figure S6). With all the "successful spawning probability" vectors of all years, we run a non-normalized Principal Component Analysis on a data table composed with nodes as observations (1500 nodes $\times$ 10 years) and 36 dates as variables (FactoMineR package/library on R version 3.6.0 (Lê et al., 2008)). The 2D mapping of the principal components (coordinates of each node) allows to display the different years (yearly mean of the node coordinates). Note that the unfavourable nodes with probabilities equal to zero for all years and events were removed from the Principal Component Analysis.

Finally, a Model Based Clustering (R package mclust (Scrucca et al., 2016)) is implemented using the 10 first principal component coordinates accounting for about 80% of the entire variability. The selection of a small number of node clusters in which spawning is expected to have the same pattern over time is operated using a BIC criterion. These clusters highlight the dynamical behaviour of spawning probabilities through years and thus show potential interannual stability of some areas (Figure 2). Our objective is to highlight and categorize the spatio-temporal windows suitable for spawning. Consequently, the unfavourable nodes with probabilities equal to zero for all years and events were removed from the analysis.

3 | RESULTS

3.1 | Data-constrained simulation

Discrete spawning areas are simulated, along with the respective larval contribution of each node to the total larval pool, through back-tracking experiments using the LFN model for each of all settlement sites sampled in 2008 (Figure 3) and 2009 (Figure 4). In the Adriatic Sea, our model suggests that Apulian juveniles settled in 2009 (light blue dots for sites 1 to 7 in Figure 1 and red-edged nodes in Figure 4) originated mostly (90% to 100% of the larval contribution) from nearby spawning areas located along the Apulian shores, more or less north-westward of each settlement site. The same pattern occurred in 2008 for Torre Guaceto, with few additional larvae (less than 10%) coming from the south (Figure 3d). In the Ligurian sea, our model simulates that the spawning areas situated south-eastward of Genova are the main contributors ($\simeq$ 75%) of this settlement site in 2008, while a relatively small proportion ($\simeq$ 25%) of propagules were released westward of Genova (Figure 3a). In the Tyrrhenian sea, the Maratea settlement site was replenished equally by the northern ($\simeq$ 50%) and southern ($\simeq$ 50%) spawning areas surrounding the settlement site (Figure 3b). For San Isidoro, the larvae would have been released from Southern Apulia following the coastline (Figure 3c).

Concerning the node-specific probabilities of larval replenishment, which can also be interpreted as larval contribution, the minimum is 10^{-4} and the maximum is around 0.5 for both years while probabilities are not equally distributed over space. For most sites, a few nearby nodes constitute the main suppliers (probabilities spanning 0.1 to 0.5, yellowish colours, Figures 3 and 4) while a large proportion of widespread nodes act as minor but significant contributors (probabilities lower than 0.1). Note that some rather distant nodes constitute non negligible suppliers despite their weak contribution (from $\simeq 10^{-3}$ to 3.10^{-1}) as in Tremiti Islands (42.1°N, 15.6°E) and inside the Gulf of Manfredonia (41.5°N, 16.3°E) for Apulian sites (Figures 3d and 4) or such as contributors of San Isidoro localised on the Eastern side of Apulia (Figure 3c). Furthermore, the locations of the main spawning areas (nodes with higher probabilities) replenishing Torre Guaceto vary among years, as this site was sampled twice (Torre Guaceto, Figure 3d for 2008 and Figure 4d for 2009). Moreover, it shows that the spatial extent and larval replenishment probabilities of spawning areas vary among eco-regions, exemplifying a substantial spatio-temporal variability of the spawning

process.

When compiling all the modelled temperatures extracted from the specific dates and locations defined by the 11 backtracked spawning areas for depth spanning 0 to 50 m, we find a range of temperature with a minimum around 13°C and a maximum of 24°C (Figure 5a). 95% of temperature values that are suitable for spawning ranged between 14.8°C and 20.6°C (Figure 5b). This data-constrained range of seawater temperatures is then used as a threshold triggering spawning in the following section.

3.2 | Simulated spawning success at broad-scale

3.2.1 | Inter-annual variability

Adopting an inverse modelling approach, we now identify spawning areas and their associated larval contributions at broad-scale over a 10-year period for all potential *D. sargus* populations inhabiting our domain of interest (see Figure 1). The 10-year averages of spawning success (Figure 6a) are analysed along with the standard deviation (Figure 6b) to highlight spatial patterns and assess their inter-annual variability.

High averaged probabilities of larval replenishment (around 10^{-3}) are modelled in the Croatian Islands and on

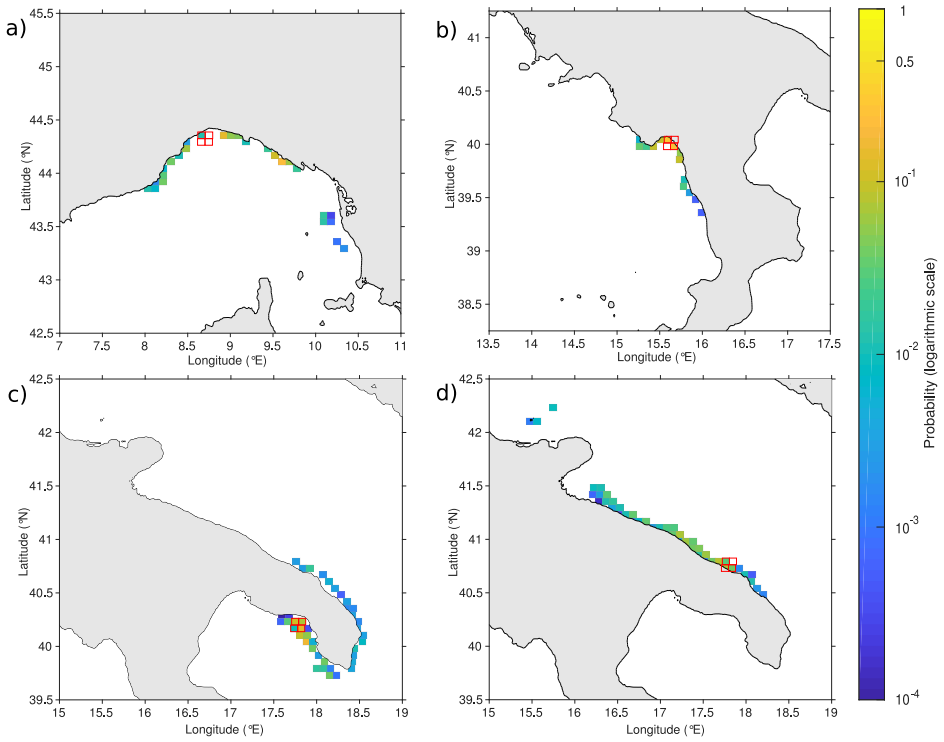


FIGURE 3 Backtracked spawning areas and associated probabilities of larval replenishment for four spawning events inferred from the post-settlers sampled in 2008 around a) Genova; b) Maratea; c) San Isidoro; d) Torre Guaceto (see also Figure 1). The red-edged nodes represent the settlement sites;

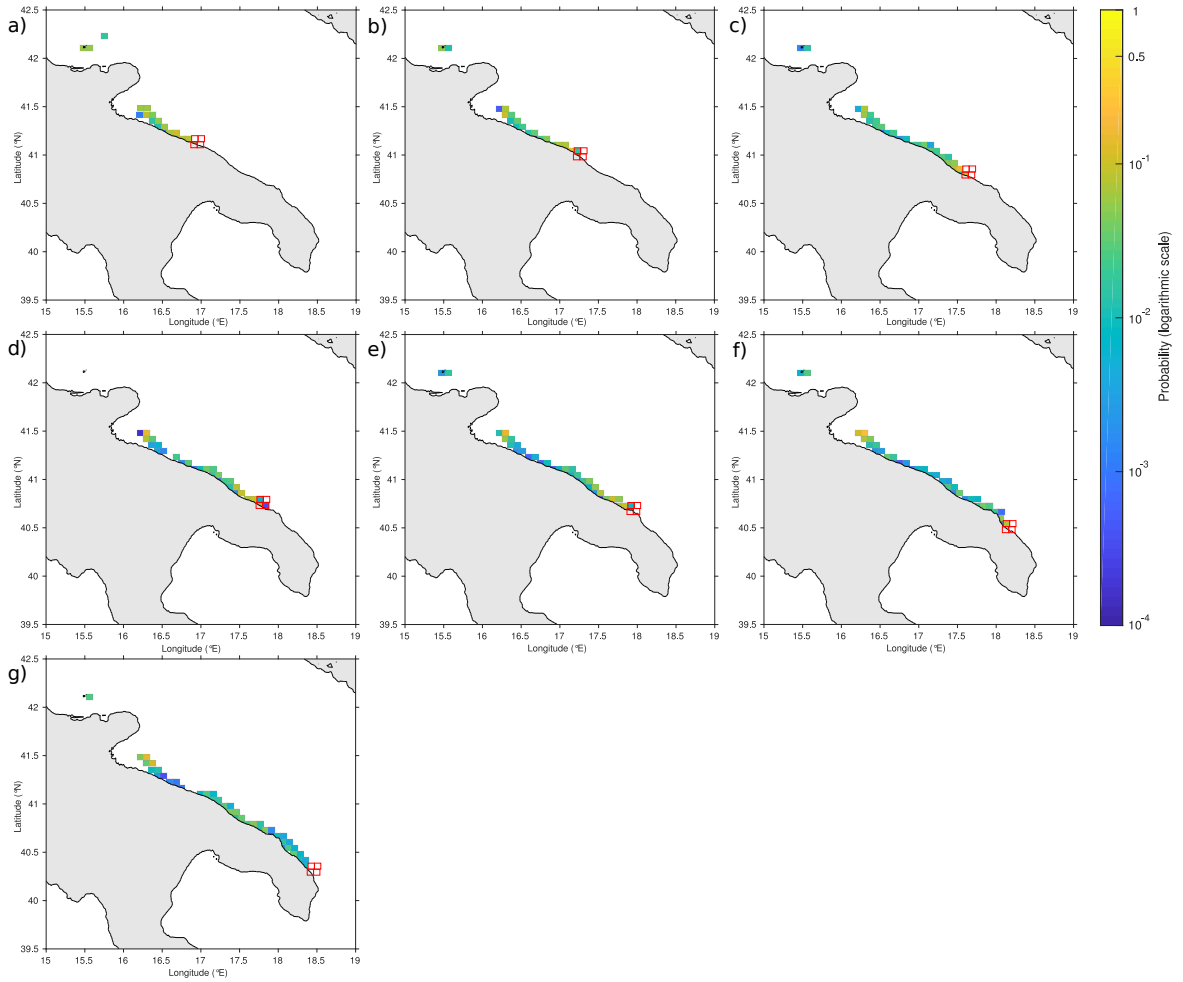


FIGURE 4 Backtracked spawning areas and associated probabilities of larval replenishment for four spawning events inferred from the post-settlers sampled in 2008 around a) Bari; b) Monopoli; c) Hotel La Darsena; d) Torre Guaceto; e) Punta Penne; f) Casalabate; g) San Andrea (see also Figure 1). The red-edged nodes represent the settlement sites;

both sides of Calabria (Figure 6a) with moderate variability (around 5.10^{-4}, Figure 6b). Similar mean values are visible on both sides of Apulia, in Eastern Corsica and Eastern Sardinia, combined with high variability (standard deviation around $2{,}5.10^{-3}$, Figure 6b). Low probabilities (from 10^{-5} to 10^{-4}) appear around the islands of the Siculo-Tunisian Strait (Figure 6a), associated with very little variability around 10^{-4} (Figure 6b). Thus, the situation is rather contrasted with the emergence of high (low, respectively) means that are not necessarily associated with high (low, respectively) variability. This conclusion highlights the necessity of disentangling further the temporal variability into seasonal and interannual signals to better evaluate how they manifest themselves over space.

The Principal Component Analysis run on the spawning success probabilities of 36 yearly events over 10 years reveal that both first axes explain 54% of the inter-annual variability of spawning success (Figure 7a). 2007 and 2012 emerge as the most contrasted years while a sub-group encompasses 2008/2014 and another larger one concerns 2005/2006/2009/2010/2013. 2007 appears as a very particular year and seems to represent, along with 2012 that

is closer to the other sub-groups, a great part of the variability of both first and second dimensions. The main reason behind this interannual structuring could be the magnitudes of probabilities, 2007 being a year with particularly high mean probabilities across the entire domain and the whole studied period (around 10^{-3}, Figure 7b) as compared with other years such as 2012 (Figures 7c and S6).

3.2.2 | Intra-annual variability

Intra-annual variability of spawning success is analysed using climatological averages (see Figure S4) and standard deviation (see Figure S5) by considering the 10 yearly replicates for each of the 36 events spanning January-June. Overall, the probability associated to an area is strongly dependent on the date (see Figure S4). In fact, average

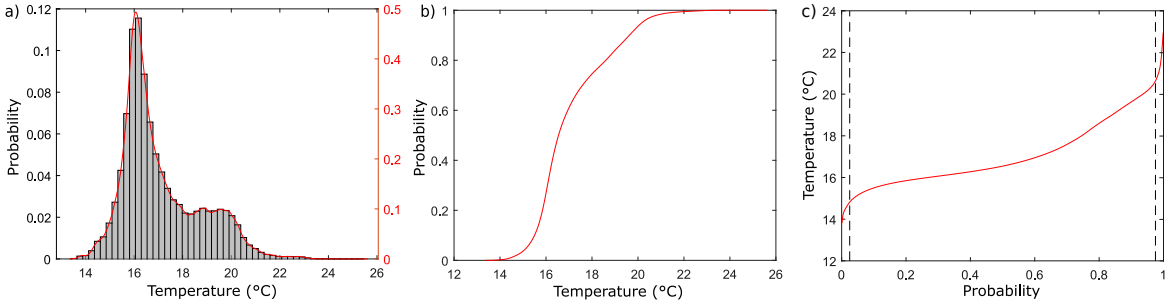


FIGURE 5 a) Kernel density function (red line) fitted on the histogram (grey bars) of modelled temperatures extracted from the 11 discrete spawning areas. b) Corresponding cumulated probability distribution function (red line). c) Corresponding quantile function (red line) with the 2.5^{th} and 97.5^{th} quantiles (dotted dark lines) used to determine the temperature range triggering spawning.

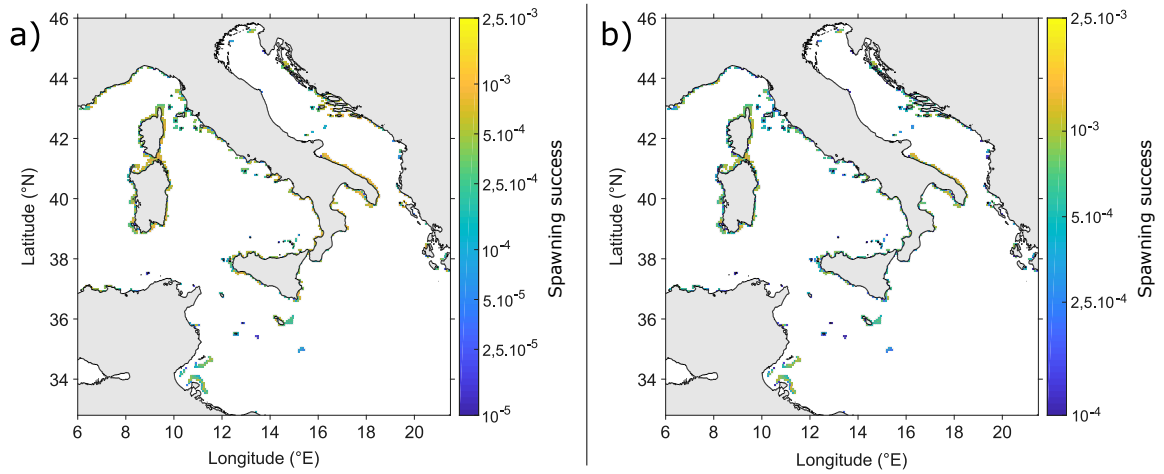


FIGURE 6 Total (10-year) a) average and b) standard deviation of successful spawning probabilities (logarithmic scale) over 2005-2014 considering all 36 spawning events simulated each year from January to June.

probabilities over most of the domain are quite high in January ($\simeq 5.10^{-4}$) and then decrease to the lowest values around February and March ($\simeq 2{,}5.10^{-5}$), except in Apulia where it remains elevated in late winter. Then, spawning success probabilities rise nearly everywhere from April until mid-June (up to $2{,}5.10^{-3}$) and then decrease rapidly, especially in Apulia, Gulf of Gabes and Toscana. Apulia, Eastern Calabria, Corsica and Sardinia are the sub-regions where intra-annual variability is the highest ($\simeq 10^{-3}$, see Figure S5). March appears as a stable month with the lowest variability on the whole domain (less than $2{,}5.10^{-4}$). On the contrary, there is a high variability from April to mid-June (around 10^{-3}) and then this variability decreases.

3.2.3 | Clustering of spawning areas

The Model-Based Clustering methodology defines in an objective manner 4 clusters, i.e. sub-groups of nodes sharing similar magnitude and intra-annual variability of the spawning success probabilities. Each cluster is described by a specific intra-annual cycle of the spawning success on the studied period (Figure 8). For each year, all the nodes with the cluster they belong to are represented in the Figure S7. Clusters 1 and 4 characterize areas with elevated spawning potential from late March to the end of June, peaking at mean values of around 10^{-3} from mid-April to the

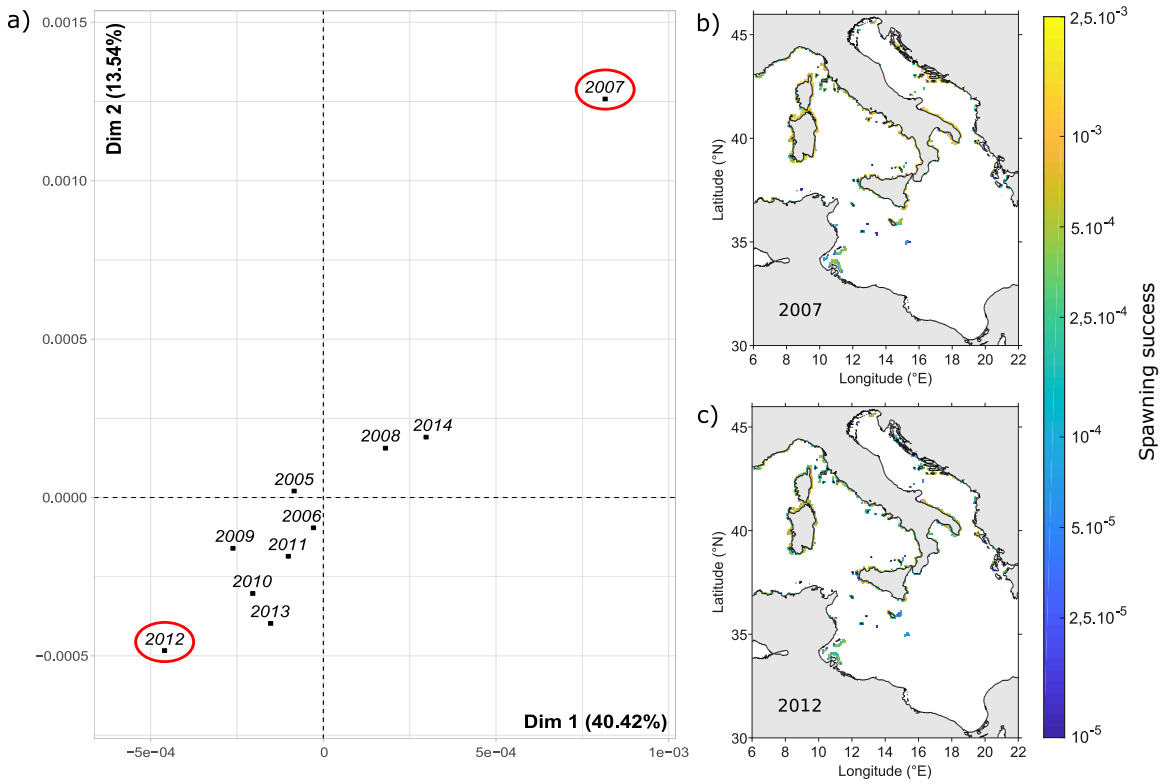


FIGURE 7 a) 2D-mapping of the Principal Component Analysis of the spawning success probabilities. Displayed years correspond to the mean coordinates of all nodes belonging to a given year. Spatial representation of the yearly mean spawning success probabilities (logarithmic scale) for the most contrasted years: b) 2007 and c) 2012. Yearly mean maps for the remaining years are reported in Figure S6.

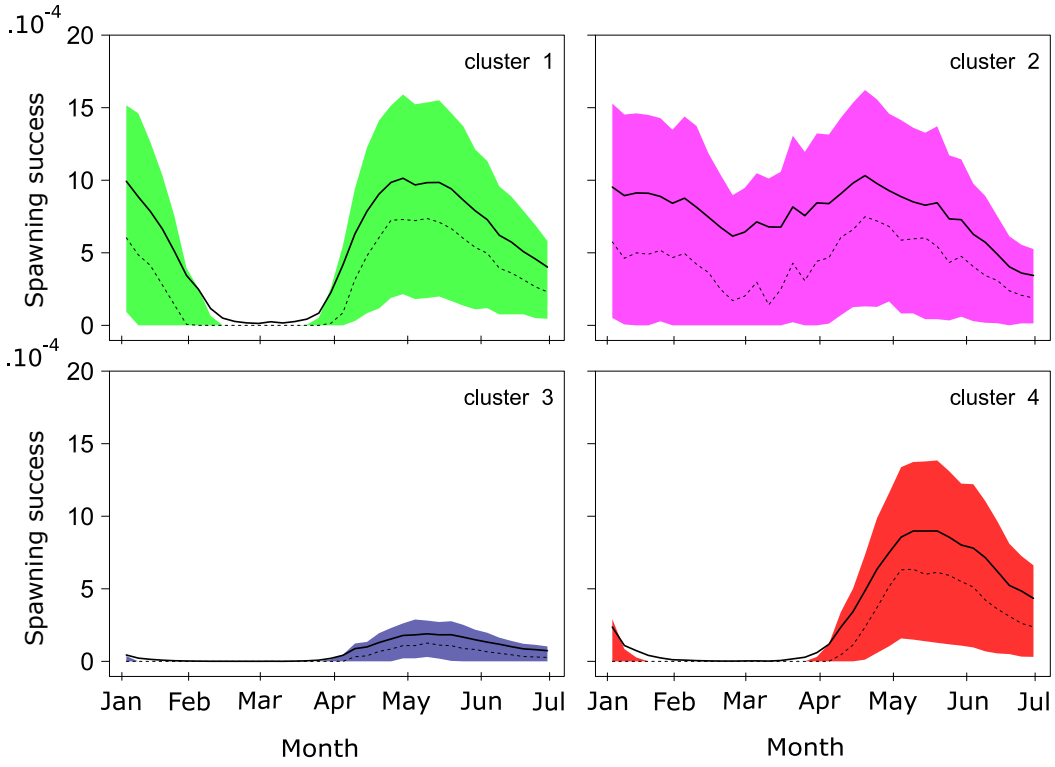


FIGURE 8 Intra-annual cycles of the spawning success probabilities for the four clusters obtained thanks to the Model-Based Clustering. The coloured shading represents the 0.25 and 0.75 quantiles, the black dot line is the median and the black bold line is the mean. The colours are used in the following to identify the clusters.

end of May. Cluster 4 is uni-modal with a unique spawning peak in late season, contrarily to the bi-modality of cluster 1 that also presents an early spawning peak in winter (January to mid-February). As both previous clusters, cluster 3 characterises areas with a late spawning success, peaking from the end of March to the end of July but with weaker probabilities (mean of about $2{,}5.10^{-4}$). Moreover, it is worth noting that the seasonal cycles reported in Figure S4 are consistent with those revealed by clusters 1, 3 and 4 with very weak to null spawning success around February to April and very low variability (see Figure S5), followed by high spawning success, with high variability, until the end of June. Cluster 2 represents areas characterized by elevated spawning potential pretty much over the whole studied period (January to June average is around 10^{-3}, Figure 8) but with high variability throughout the spawning season. The probabilities and the variability are decreasing from about early May to the end of June, but the exact duration of the spawning period is too difficult to forecast for this cluster due to its high variability over the studied period.

We now analyse the spatial distribution of the four clusters, representative of four seasonal modalities of spawning success. Each node belonging to a given cluster, one can investigate at which frequency a node is objectively characterized by the same cluster across years. We set an arbitrary threshold to distinguish the discrete spawning areas whose seasonal behaviour is consistent from one year to another, from those whose seasonality varies at inter-annual time scales. In other words, the nodes whose cluster frequency is lower than 0.7 are subjected to prominent inter- and intra-annual variability; conversely, nodes characterized by cluster frequency greater or equal to 0.7 can be considered with relatively stable seasonal behaviour across years. The coloured nodes displayed in Figure 9 represent

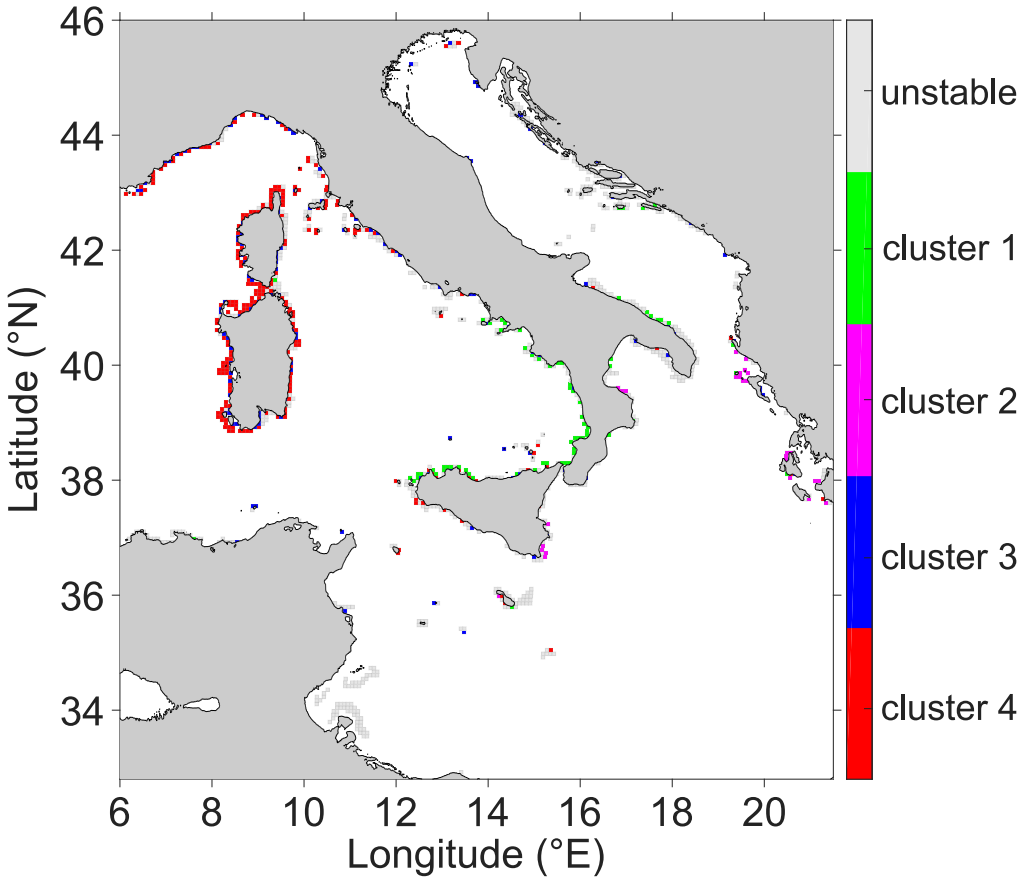


FIGURE 9 Spawning areas whose intra-annual variability (see also Figure 8) is stable across years (i.e. whose cluster frequency is greater or equal to 0.7) are coloured according to the cluster they belong to. Light grey nodes represent the overly variable spawning areas (i.e. whose cluster frequency is lower than 0.7).

the spawning areas whose simulated success is highly stable across years, representing about 50

4 | DISCUSSION

4.1 | Abiotic and biotic factors affecting spawning

Using the white seabream *Diplodus sargus* in the central Mediterranean Sea as a case study, we investigate the role of seawater temperature in controlling broad-scale spawning onset and duration of coastal fish by identifying the suitable temperature range for spawning while ensuring that spawning and settlement areas are effectively connected by oceanic dispersal. The key role of temperature for triggering spawning has been indeed already documented in previous research (Table 2, see “Temperature" row). Literature also suggests that spawning can be affected by multiple other abiotic or biotic factors (Table 2). Some papers focus primarily on abiotic factors (Table 2), such as oceanic

TABLE 1 Literature review reporting the minimum and maximum values of temperature (in °C) suitable for spawning and the favourable period (in month) for spawning documented for *Diplodus sargus* in the Mediterranean Sea. The last line reports the values derived from the present study.

Article	Minimum temperature	Maximum temperature	Minimum spawning date	Maximum spawning date
Aspillaga et al., 2016	13	14.5	March	April
Di Lorenzo et al., 2014	13	18	April	May
Mouine et al., 2007	15	18	March	May
Mouine et al., 2012	14.8	15.6	March	June
El Maghraby et al., 1982	17	20	January	April
Benchalel and Kara, 2010	15	18	February	April
Man-Wai, 1985	14	16	April	May
Present study	**14.8**	**20.6**	**January**	**June**

temperature, upwelling, photoperiod, solar irradiance, lunar cycle (also linked to tides) and daylight/dark diurnal alternation. These studies often discuss the physiological internal processes (feeding status, gonad maturation) which would explain how environmental forcing somehow indirectly control fish spawning behaviour. Other researches focus on biotic factors (Table 2) at both population, such as demography and inter-specific competition, and individual levels, such as feeding status through lipid storage.

Here we specifically focused on temperature as this abiotic factor has been predominantly reported as crucial in the compiled literature, comparatively with other factors (see Table 2). While this overrepresentation might originate from a publication bias, it is easily ruled out by some multi-factorial studies that reported major influences of seawater temperature on spawning as compared to other factors (such as photoperiod or lunar cycle, especially in the Mediterranean (De Vlaming, 1972; Hereu et al., 2006; Sheaves, 2006; Potts et al., 2014)). More generally, this is consistent with the key role of seawater temperatures in driving major ecological patterns in the global ocean, such as fish biodiversity and dynamics (Tittensor et al., 2010; Burrows et al., 2019). The current warming trend of the coastal ocean due to climate change (Lima & Wethey, 2012; Liao et al., 2015) thus raise concerns about possible temporal and spatial shifts of spawning for coastal fishes, as already observed in demersal fish such as cod *Gadus morhua* and sole *Solea solea* (Fincham et al., 2013; McQueen & Marshall, 2017; Auth et al., 2017). Our modelling framework, which has been especially designed to consider the full environmental variability of both ocean temperature and current in evaluating spawning success, could be used to predict future changes as well as to guide future sampling efforts.

Our findings highlighted a range of seawater temperatures suitable for spawning (14.8°C - 20.6°C, see section 3.1) that is consistent with the ranges compiled from the literature, especially for the minimum temperature (Table 1). The small mismatches between our range derived from modelled temperatures over multiple depths and those already published (Table 1) may result from the fact that they exclusively used Sea Surface Temperature, despite the fact that the vertical distribution of coastal fishes when living and/or spawning is much higher than just the sea surface (Aspillaga et al., 2016; Giacalone et al., 2018). Moreover, the surprisingly good agreement for the lower temperature threshold that triggers spawning may be partly explained by the fact that the spawning onset and associated temperatures are often derived from observational studies of juveniles targeting especially the supposed settlement peak (Table 2 and Di Franco and Guidetti, 2011). On the other hand, the upper thermal range is slightly higher than previous estimates, suggesting that our approach may slighty overestimate the spawning duration. While this flaw would

especially concern the warmer eco-regions (e.g. Gulf of Gabes) during the late spawning season (May/June), observations showed that some spawning events have indeed occurred in these space/time windows (Morato et al., 2003; Benchalel & Kara, 2010). It is worth noting that the cessation of spawning could also be related to a combination of both external (environmental) and internal (biological) factors. In fact, some fishes could have stopped spawning after a certain duration of active spawning not necessarily due to non-favourable abiotic factor but rather due to biological limitations such as spent gonads (Mouine et al., 2012) or behavioural cessation of spawning aggregation (Colin, 2010; Aspillaga et al., 2016). Overall, it suggests that the control exerted by oceanic temperature on the spawning process may be more relevant to simulate the spatio-temporal variations of its onset rather than its termination.

Furthermore, more detailed analyses of temperatures for 2008 and 2009 spawning events point out different suitable temperatures: mean optimal temperatures are around 18°C for all the sites of 2008 and around 16°C for all those of 2009. This could be explained by the different sampling design between 2008 (four sites relatively well scattered across the studied domain) and 2009 (seven sites, all located in Eastern Apulia). These regional differences of triggering temperatures could suggest local adaptation (Takahashi et al., 2012) but the small statistical samples prevent any further analyses. It is worth noting that other factors neglected here, such as saturation of the settlement habitat (Di Franco et al., 2013) or local hydrodynamics (Cuadros et al., 2018), could also be responsible for the variations of PLDs and spawning dates. Moreover, when analysing early-life traits of *D. sargus* settlers in another area (Marseilles, France) during years 1994 and 1995, Vigliola, 1998 found a mean PLD that is considerably higher than the one used in the present study (i.e. 28 vs 16 days). Hence, expanding our study to other areas and other years with more numerous observations could help studying the spatio-temporal variability of these early-life traits.

4.2 | Disentangling the spatial and temporal variability of spawning

4.2.1 | Spatial variability of backtracked spawning areas

Backtracking model experiments constrained by settlers observations (Calò et al., 2018; Legrand et al., 2019) allow us to characterize several discrete spawning events by well-delineated areas associated with their respective contribution to the larval replenishment of 11 settlement areas, considered here (Di Franco et al., 2011; Di Franco & Guidetti, 2011). Their spatial locations and relative larval contributions can be reasonably well explained by both local and regional hydrodynamics and topography.

Our model identifies the spatial locations of the major spawning areas (i.e. highest probabilities) within the pathways of surface currents prevailing in the region (Millot & Taupier-Letage, 2005; Poulain et al., 2013), generally upstream from their settlement locations. Once released off Livorno and La Spezia, fertilized eggs and larvae tend to follow the Liguro-Provencal current toward the Genova settlement site (Figure 3a). In the Tyrrhenian Sea, the influx of Atlantic Water entering from the Sardinia Channel (Iacono et al., 2021) carries propagules in a counter-clockwise direction following the shorelines of western Calabria up to Maratea (Figure 3b). In the Gulf of Taranto, the San Isidoro settlement site is replenished thanks to a coastal current that originates from the southern Adriatic Sea, then circles around the Salanto peninsula while being pushed westward by the North Ionian Gyre (Gačić et al., 2010; Crisciani and Mosetti, 2016; Figure 3c). Finally, the south-westward coastal current in the Adriatic Sea is the main contributor to the Apulian settlement sites sampled in 2008 and 2009 (Figures 3d and 4), which is in good agreement with Pujolar et al., 2013, Bray et al., 2017 and Legrand et al., 2019.

TABLE 2 List of factors influencing the spawning process of Teleosts (in blue: of Sparids only) according to a literature review.

Factor	Species	Area	Articles
Temperature	Teleosts	Victoria, Australia	King et al., 2016
	Clupea harengus	Newfoundland, Canada	Winters and Wheeler, 1996
	Pseudogobius olorum	Western Australia, Australia	Gill et al., 1996
	Teleosts	Oregon, USA	Auth et al., 2017
	Solea solea	Irish sea, North Sea, Channel	Fincham et al., 2013
	Epinephelus marginatus	Medes Islands MPA, Catalonia, Spain	Hereu et al., 2006
	Gadus morhua	North Sea, Scotland	McQueen and Marshall, 2017
	Thunnus thynnus, Auxis rochei, T. alalunga	Balearic archipelago, Spain	Reglero et al., 2012
	Sparidae	Meta-analysis	Sheaves, 2006
	D. sargus subspecies	Meta-analysis	Potts et al., 2014
	D. sargus, D. annularis, D. puntazzo, D. vulgaris	Gulf of Tunis, Tunisia	Mouine et al., 2011
	D. cervinus cervinus	Canarian archipelago, Spain	Pajuelo et al., 2003
	D. vulgaris	Canarian archipelago, Spain	Pajuelo et al., 2006
	D. puntazzo	Canarian archipelago, Spain	Pajuelo et al., 2008
	D. vulgaris	South Western Portugal	Gonçalves, 2000
	D. sargus	Azores archipelago, Portugal	Morato et al., 2003
	D. sargus sargus	Gulf of Tunis, Tunisia	Mouine et al., 2007
	D. sargus sargus	Torre Guaceto MPA, Puglia, Italia	Di Lorenzo et al., 2014
	D. sargus	Medes Islands MPA, Catalonia, Spain	Aspillaga et al., 2016
Photoperiod	Teleosts	Meta-analysis	De Vlaming, 1972
	Lepomis macrochirus	Story County, Iowa, USA	Mischke and Morris, 1997
	Dicentrarcus labrax	West Coast of Portugal	Vinagre et al., 2009
	D. vulgaris	South West Coast of Portugal	Gonçalves, 2000
	Sparids	Meta-analysis	Sheaves, 2006
Irradiance	*D. vulgaris*	South West Coast of Portugal	Gonçalves, 2000
Moon / Tides	*Cheilinus undulatus*	Republic of Palau	Colin, 2010
	Paracirrhites forsteri	Kuchierabu-jima Island, Japan	Kadota et al., 2010
	Takifugu niphobles	Kyushu Island, Japan	Yamahira, 1997
	Epinephelus marginatus	Medes Islands MPA, Catalonia, Spain	Hereu et al., 2006
	Thalassoma bifasciatum	Puerto Rico	Appeldoorn et al., 1993
	Acanthopagrus pacificus	Queensland, Australia	Sheaves and Molony, 2013
	Acanthopagrus berda	Kosi estuary, South Africa	Garratt, 1993
Dark / Night	*Osmerus mordax*	Newfoundland, Canada	Bradbury et al., 2004
	Acanthopagrus berda	Kosi estuary, South Africa	Garratt, 1993
Upwelling	*Centroberyx gerrardi*	South-Western Australia	Coulson et al., 2019
	Sardinops, Engraulis	California, USA	Lluch-Belda et al., 1991
Demography	Teleosts	Meta-analysis	Wright and Trippel, 2009
Lipid storage	*Gambusia affinis*	Illinois, North Carolina, New Jersey, USA	Reznick and Braun, 1987
Competition	*D. annularis, D. sargus, D. puntazzo, D. vulgaris, Oblada melanura, Sarpa salpa*	Marseille, France	Harmelin-Vivien et al., 1995
	D. puntazzo, D. sargus, D. vulgaris	North-West Mediterranean Sea	Planes et al., 1999

Counter-intuitively, some minor spawning areas (i.e. lowest probabilities) are found downstream for Genova, Maratea and Torre Guaceto in 2008. These more complex patterns can be caused by seasonal slowdowns of the main currents, by the presence of mesoscale eddies creating turbulent and anisotropic dispersal or by coastal counter-currents creating small-scale recirculation patterns (Millot & Taupier-Letage, 2005; Poulain et al., 2013). Moreover, the settlement site of Torre Guaceto was sampled in both 2008 and 2009 (Figures 3d and 4d), returning different spawning dates and spatial distribution of backtracked spawning areas. It exemplifies the substantial inter-annual variability of hydrodynamics that could be linked to the current reversal of the North Ionian Gyre occurring between these years (Gačić et al., 2010; Crisciani & Mosetti, 2016).

Concerning the levels of contribution, a general pattern is that the closest to the settlement area, the highest the larval contribution for Maratea and San Isidoro in 2008 (Figures 3b and 3c). It can be related to the relatively short PLDs of *D. sargus* and the sluggish circulation in these regions, resulting in limited dispersal. Note that the probabilities reported here must be interpreted with caution as we assumed similar larval production in each spawning area. In other words, the spatial heterogeneity of larval contributions originates essentially from the effective connectivity (controlled by ocean currents) between spawning and settlement areas. Since the mean number of fish eggs produced by mature adults is proportional to the size of the individuals (Lester et al., 2009), spatially explicit quantitative information about adult biomasses, or directly about eggs production (Blanco et al., 2017), could help improving the accuracy of our simulations in the future. Even without this valuable information, a more affordable improvement could be to scale the larval production of each spawning node according to its protection status, as done in Legrand et al., 2019. Indeed, current MPAs are expected, when effectively managed and enforced, to host bigger individuals producing large number of eggs (Di Franco et al., 2012; Marshall et al., 2019) that could act as effective source of propagules for surrounding unprotected populations. Some distant sources highlighting weak but non negligible larval supply to the settlement site, such as Tremiti Islands (42.1°N, 15.6°E; Figures 3d and 4), could thus contribute to a higher level of larval replenishment than what is initially expected because they are protected by an effective MPA.

4.2.2 | Spatio-temporal variability of simulated spawning success

One of the overarching goals of this study is to evince how the variability of oceanographic processes, namely hydrodynamics and seawater temperatures, affects the simulated spawning success in order to draw more general conclusions applicable to other oceanic systems. Globally, our model suggests that there are two peaks of favourable conditions for spawning to occur: January to early/mid-February and from April to June for all clusters, primarily due to suitable temperatures for spawning during these periods. Beyond too simplistic latitudinal patterns, note that the early occurrence of suitable conditions for spawning concerns not only the most southern locations (e.g. Gulf of Gabes, also backed-up by observations of mature spawners during winter by Mouine et al., 2007 or Benchalel and Kara, 2010 in Algeria) but also more septentrional locations (such as Sicily and Southern Italy), suggesting the need to collect field observations earlier in the season in these regions. For clusters 1, 3 and 4 (as well as for cluster 2, to a lesser extent), our results suggest weak to null spawning success from early February through to March, which seems to be due to seawater temperature lower than the suitable range (i.e. < 14.8°C) in these areas. This bi-modality of the simulated spawning success could be explained by the general seasonal cycle of oceanic temperatures that is slighty delayed as compared to air temperatures: coldest sea temperatures occur in late winter/early spring while air temperature are already warming up. The non-favourable conditions simulated from February to April are probably linked to deep convection and dense water formation events that have been documented in the southern Adriatic

and the Ligurian sea (Millot & Taupier-Letage, 2005).

More specifically, south-eastern Sicily, Greece and Albania are represented by cluster 2. For this cluster, the period between February and April exhibits lower spawning success (but still positive) than the early and late peaks. It suggests that temperature rarely goes below the prescribed threshold in these regions (far from the deep water formation areas) and that the small variations could be due to the variability of ocean currents. These areas are indeed characterized by highly turbulent mesoscale circulation such as the North Ionian Gyre and the Mid-Ionian Jet (Millot & Taupier-Letage, 2005; Poulain et al., 2012).

The cluster 4 represents areas in Ligurian Sea, Corsica, Sardinia, Toscana and French Riviera, which seem more impacted by temperature due to higher latitudes. The Ligurian sea is characterized by the Northern Current and is also known for deep-water formations (Poulain et al., 2012) which, along with the latitude, could explain why cluster 4 prevails in the Ligurian sea whereas cluster 1 predominates in the Tyrrhenian Sea (e.g. North Sicily, Calabria and Campania). In fact, this relates to the influx of warmer Atlantic waters coming from the Algerian Current: a major part, known as the Atlantic-Ionian stream, goes through the Siculo-Tunisian Strait to warm-up the southernmost locations (e.g. Gulf of Gabes) while another smaller contribution invades the Northern Sicily coastlines and the rest of the Tyrrhenian Sea (Millot & Taupier-Letage, 2005; Poulain et al., 2012; Iacono et al., 2021). Note that the large variability of the Atlantic-Ionian stream may also explain why the spawning success is characterized by unstable clusters in the Siculo-Tunisian Strait.

In fact, the Siculo-Tunisian Strait region is known to present particularly high eddy kinetic energy due to the presence of the particularly unstable Atlantic-Ionian stream (D'Ovidio et al., 2004), explaining why the spawning success probability is very low due to non-favourable dispersal routes that poorly connect spawning and settlement areas (Cuttitta et al., 2016; Figure 6a). Moreover, the regular coastal upwelling occurring on the western coast of the Sicily (Piccioni et al., 1988) could play a role too but its effect on the spawning success probability is difficult to disentangle from the influence of the Atlantic-Ionian stream in this region. More local analyses on spawning and larval dispersal, as realised by Cuttitta et al., 2016, could help to further understand the coupled bio-physical processes occurring in the region. On the other hand, the high variability of the spawning success in the Gulf of Gabes (unstable interannual clusters, see Figures 9 and S7) could be explained by a quick rise of the temperature and the large air-sea interactions (e.g. evaporation and heat fluxes) enhanced by the very shallow bathymetry in this region.

Finally, cluster 3 is found around some islands of the Siculo-Tunisian Strait and in Aeolian Islands which are highly turbulent areas. This cluster is represented by a very weak peak from April to June, however temperature is in the good range in the concerned regions (Figure 8), but the high variability of currents in these areas could be responsible of the limited connectivity among spawning and settlement areas. On the other hand, clusters 3 and 4 are interlocked in the same areas (Ligurian Sea, Corsica, Sardinia, French Riviera, Toscana; see Figure 9) with cluster 3 appearing regularly in similar nearshore areas. These patterns could be a methodological artefact and must be interpreted with caution as these nodes present a mean land ratio of 0.75 (i.e. meaning that three quarter of the node is represented by land) due to the poor resolution of the ocean model around these complex coastlines.

Despite intricate spatio-temporal variations, we also highlight high probabilities and high variability of spawning success in the Adriatic, especially in Apulia and in Croatian Islands (Figure 6), supported by the unstable interannual clustering (Figures 9 and S7). For the Croatian Islands, this pattern can be explained by a shallow bathymetry (< 100 metres depth) and sluggish circulation probably implying a quick rise of temperature in these areas. Concerning Apulia, Gačić et al., 2014 studied the BiOS (Adriatic-Ionian Bimodal Oscillating System), which is a regime-shift of the

North Ionian Gyre occurring at inter-annual/decadal time scales, modifying the regional-scale currents, including at the entrance of the Adriatic Sea (Gačić et al., 2010; Crisciani & Mosetti, 2016). Those circulation changes could partly explain the inter-annual variability of spawning highlighted here. In fact, the path of the waters entering the Adriatic through the Strait of Otranto can considerably vary from one year to another as a function of the BiOS regimes, leading to large variations in temperature and salinity (Gačić et al., 2013). When the gyre is anticyclonic (e.g. 2007-2008), the waters come from the Atlantic-Ionian stream and the Levantine waters originate from the west. When the regime is cyclonic (e.g. 2012-2013), the waters are directly coming from the Levantine Sea located to the South-East of the Adriatic Sea (Gačić et al., 2010). Consequently, it could be more or less favourable for spawning, depending on the BiOS regime.

Moreover, Gačić et al., 2014 observed a transition from cyclonic to anticyclonic circulation during 2005-2006, which is consistent with the central positions of both years in Figure 7a. In accord with the prominence of anticyclonic circulation during 2007-2008, years 2007 (and 2008, to a lesser extent) appears in the upper right corner of Figure 7a. It is characterized by extreme positive values for both dimensions and by the highest probabilities over the entire domain (Figure 7b). Thereafter, a second reversal of the basin-wide circulation occurred in 2009/2011, potentially explaining why these years are close to the centre of Figure 7a. Years 2012/2013 are characterized by negative values for both dimensions (lower left corner, Figure 7a), which is coherent with the prevailing cyclonic circulation during this period. Moreover, 2012 has been described as a very harsh winter by Gačić et al., 2014, probably explaining the lowest negative values for both dimensions (Figure 7a) and the very low probabilities of spawning success (Figure 7c). While the very specific position of year 2012 could have been slightly affected by the lack of five dates in January, it is likely that the main conclusions still hold since these dates are characterized by very low probabilities during the other years (see Figure S4).

Overall, our results suggest that colder and longer winters than average, as monitored by the BiOS index, are deleterious for the spawning success in our studied domain while warmer and milder winters than average are favouring spawning success at regional scale. In fact, cluster 2, characterized by favourable spawning success over the entire season, is predominant over year 2007 (see Figure S7) which seems to be a warmer year than average due to a particularly short winter in comparison with the extreme 2012 winter (Gačić et al., 2014). The opposite behaviour of both years (Figure 7a) confirm the particularly high spawning success probability over the entire domain during 2007 (see Figure S6). Hence, 2014 must be a mild year too, mainly represented by the cluster 2 (see Figures S6 and S7) and by positive values for both dimensions on Figure 7, such as 2007.

Note that the variability of spawning success in the Adriatic Sea could also reflect local rather than regional oceanographic dynamics. It is indeed a semi-enclosed sea with its own intrinsic variability associated to local oceanographic processes such as coastal cooling/warming due to sporadic upwelling/downwelling events forced by Bora winds and/or fine-scale circulation changes (Magaldi et al., 2010). However, additional temperature analyses (not shown) highlight a change in the general dynamic of the temperature at regional scale, consistent with the BiOS regime. In fact, from 2005 to 2008, the warming of the waters is realised from the North-East to the South-West (i.e. Adriatic first, Liguria second, Tunisia last) contrary to the period 2009-2014 for which the warming of the waters occurs from the South-East to the North-West (from the Eastern Mediterranean sub-basin to the Adriatic and the Western Mediterranean sub basin).

Overall, seawater temperature seems to be a preponderant factor of the triggering and intensity of spawning success, whereas currents act like a “background noise” of the spawning success, except in places where highly turbulent

currents have a major impact in the spawning success due to limited connectivity.

4.3 | Implications for management

Despite the fact that stocks of *Diplodus sargus* are not currently assessed by the General Fisheries Commission for the Mediterranean (FAO, 2006), coastal fishes like Sparids are targeted by small scale and recreational fisheries with reports of declining catches (FAO, 2004). In this context, the specific variability of the spawning success demonstrated by our results could help selecting the most adapted management tools and best spatially-allocate restrictions, as advised by Erisman et al., 2015; Sadovy de Mitcheson, 2016. Note that while Hussein et al., 2011 reported that permanent fishing closures appeared more efficient on juveniles habitat (e.g. settlement) than on spawning zones, these conclusions need revisions as the ISIS-Fish model neglects the dispersal of early-life stages.

By assuming that any management initiative should be consistent over several years to be worth implementing, all nodes whose cluster frequency is lower than 0.7 (Figure 9) could be excluded since they exhibit too much inter-annual variability. In contrast, all nodes whose simulated spawning success and its intra-annual evolution are relatively stable across years are good candidates following the current trend of spatially fixed management measures such as the Fisheries Restricted Areas (FAO, 2021). Areas characterized by clusters 1 and 4 consistently across years with favourable spawning success in late season could be managed by restricting fishing when spawning potential is the highest (e.g. through April, May, June, see Figure 8). These measures represent useful management tools to protect and rebuild populations of exploited as well as patrimonial species (Di Franco et al., 2012; Guidetti et al., 2014; Di Franco et al., 2018). Note that the relevance of considering the early spawning peak (early January to mid-February) of cluster 1 for management could be questioned as Mouine et al., 2007 and Mouine et al., 2012 documented in the Gulf of Tunis that gonads of *D. sargus* were not mature for spawning during this period. Using a gonado-somatic index, they stated that spawning onset is linked to temperature increase, whereas in January it generally decreases or stabilizes in Western Mediterranean, except in places where winter is briefer like Tunisia or Algeria (see spawning period of Benchalel and Kara, 2010 in Table 1). Thus, the regions of North Sicily, Western Calabria and Apulia represented by cluster 1 could probably be protected during the late season only, such as cluster 4. In this study, we only look at external factors triggering spawning but they could be favourable in periods where the gonado-somatic index is not, so it would be important to consider biotic processes in future studies. Such as clusters 1 and 4, cluster 3 presents a spawning potential late in the season but the weakness of the spawning success peak questions the relevance of a protection for the concerned areas, that require more information to be correctly managed. Finally, the spawning success of the areas typified by cluster 2 (south-eastern Sicily, Albania and Greece in Figure 9) is very stable and elevated over time, which require long-term measures. In fact, these areas could be protected all year long with "static" measures like MPA allowing the resilience of neighbouring exploited areas thanks to larval export (Fogarty & Botsford, 2007; Jones et al., 2007; Gaines et al., 2010).

Furthermore, dynamic measures could be required in order to adapt management with the oceanic variability. Some regions, highly dependent on oceanic temperature (areas represented by clusters 1 and 4 such as Corsica, Sardinia or North Sicily), could be ideal places to test implementing flexible protected areas and adaptive measures of protection or regulation, based on the variation of this environmental trigger of spawning (Maxwell et al., 2015). It is worth mentioning that, in the coming years, fishes are more likely to move to higher latitudes due to climate change and ocean warming, especially exacerbated near the coast (Perry et al., 2005; Lima & Wethey, 2012; Liao et al., 2015),

and dynamic management measures could thus be a standard in the future.

However, planning management measures at this scale could be challenged by the political context of the Mediterranean Sea, with potential straddling fish stocks encompassed by multiple jurisdictions or multiple countries (Hidalgo et al., 2019), for instance between Greece and Albania, making these measures difficult to apply without the concertation and collaboration of the various stakeholders.

Nevertheless, the methodology of this study could be used for other species with a bipartite life cycle in any other oceanic system for which operational ocean model exist. Our results could help considering the interannual variability of spawning in the management of resources, and more especially in a context of adaptive and dynamic management (Maxwell et al., 2015). As future perspective, even more effective recommendations could probably be achieved by considering an ecosystem-based approach that would study together multiple key species of an ecosystem to enhance their sustainability and reduce the negative impacts of human activities on these resources (Levin & Lubchenco, 2008). While our modelling approach could be well-adapted to guide these efforts, the accuracy and reliability of its outcome still rely on the performance levels of operational ocean models (Reglero et al., 2018). Finally yet importantly, it is worth noting that the forecast lead-time of operational ocean models does not currently exceed a few days. It implies that historical analyses such as this one are promising while future projections at these scales are still out-of-reach.

5 | CONCLUSION

In this study, we use backward-in-time particle tracking model from juveniles' settlement observations to delineate 11 spawning events and to determine a range of suitable temperatures for spawning. With this range that is consistent with the literature, we model and analyse the intra- and inter-annual variability of spawning locations and duration over ten years, from 2005 to 2014. Thanks to a Model Based Clustering method, we disentangle the different levels and sources of variability in the simulated spawning success, putting the emphasis on the role of oceanographic processes. The approach and tools developed in this study could guide future efforts that aims at evaluating when and where spawning occur for other territorial fish species to help managers and conservationists.

Acknowledgments

The authors acknowledge support from Région Sud for M.D.S's PhD grant and co-funding by Port-Cros National Park through a project supported by the Prince Albert II of Monaco Foundation. T.L. is funded by a Doctoral fellowship obtained through Aix-Marseille University.

Data Availability

Sharing observed data is not applicable to this article as no new data were created or analyzed in this study (original published in Di Franco et al., 2011 and Di Franco and Guidetti, 2011). Simulated results are available upon reasonable request to the authors.

References

Abecasis, D., Horta e Costa, B., Afonso, P., Gonçalves, E. J., and Erzini, K. (2015). "Early reserve effects linked to small home ranges of a commercial fish, Diplodus sargus, Sparidae". In: *Marine Ecology Progress Series* 518, pp. 255–266. issn: 0171-8630, 1616-1599. doi: 10.3354/meps11054.

Appeldoorn, R. S., Hensley, D. A., Shapiro, D. Y., Kioroglou, S., and Sanderson, B. G. (1993). "Egg Dispersal in a Caribbean Coral Reef Fish, *Thalassoma bifasciatum*. II. Dispersal off the Reef Platform". In: *Bulletin of Marine Science* 54.1, pp. 271–280.

Aspillaga, E., Bartumeus, F., Linares, C., Starr, R. M., López-Sanz, À., Díaz, D., Zabala, M., and Hereu, B. (2016). "Ordinary and Extraordinary Movement Behaviour of Small Resident Fish within a Mediterranean Marine Protected Area". In: *PLOS ONE* 11.7. Ed. by H. M. Patterson, e0159813. issn: 1932-6203. doi: 10.1371/journal.pone.0159813.

Auth, T. D., Daly, E. A., Brodeur, R. D., and Fisher, J. L. (2017). "Phenological and distributional shifts in ichthyoplankton associated with recent warming in the northeast Pacific Ocean". In: *Global Change Biology* 24, pp. 259–272. issn: 13541013. doi: 10.1111/gcb.13872.

Bauer, R. K., Gräwe, U., Stepputtis, D., Zimmermann, C., and Hammer, C. (2014). "Identifying the location and importance of spawning sites of Western Baltic herring using a particle backtracking model". In: *ICES Journal of Marine Science* 71.3, pp. 499–509. issn: 1095-9289, 1054-3139. doi: 10.1093/icesjms/fst163.

Benchalel, W. and Kara, M. H. (2010). "Age, croissance et reproduction du sar commun *Diplodus sargus sargus* (Sparidae) des côtes de l'Est algérien". In: *Rapport de la Commission Internationale pour l'Exploration scientifique de la mer Méditerranée* 39, p. 451.

Blanco, M., Ospina-Álvarez, A., González, C., and Fernández, M. (2017). "Egg production patterns of two invertebrate species in rocky subtidal areas under different fishing regimes along the coast of central Chile". In: *PLOS ONE* 12.5. Ed. by J. P. Meador, e0176758. issn: 1932-6203. doi: 10.1371/journal.pone.0176758.

Bradbury, I. R., Campana, S. E., Bentzen, P., and Snelgrove, P. V. R. (2004). "Synchronized hatch and its ecological significance in rainbow smelt *Osmerus mordax* in St. Mary's Bay, Newfoundland". In: *Limnology and Oceanography* 49.6, pp. 2310–2315. issn: 0024-3590. doi: 10.4319/lo.2004.49.6.2310.

Bray, L., Kassis, D., and Hall-Spencer, J. M. (2017). "Assessing larval connectivity for marine spatial planning in the Adriatic". In: *Marine Environmental Research* 125, pp. 73–81. issn: 0141-1136. doi: 10.1016/j.marenvres.2017.01.006.

Burgess, S. C., Baskett, M. L., Grosberg, R. K., Morgan, S. G., and Strathmann, R. R. (2016). "When is dispersal for dispersal? Unifying marine and terrestrial perspectives". In: *Biological Reviews* 91.3, pp. 867–882. issn: 14647931. doi: 10.1111/brv.12198.

Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A. K., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W., and Botsford, L. W. (2014). "Beyond connectivity: how empirical methods can quantify population persistence to improve marine protected-area design". In: *Ecological Applications* 24.2, pp. 257–270. issn: 1051-0761. doi: 10.1890/13-0710.1.

Burrows, M. T., Bates, A. E., Costello, M. J., Edwards, M., Edgar, G. J., Fox, C. J., Halpern, B. S., Hiddink, J. G., Pinsky, M. L., Batt, R. D., García Molinos, J., Payne, B. L., Schoeman, D. S., Stuart-Smith, R. D., and Poloczanska, E. S. (2019). "Ocean community warming responses explained by thermal affinities and temperature gradients". In: *Nature Climate Change* 9, pp. 959–963. issn: 1758-6798. doi: 10.1038/s41558-019-0631-5.

Cadrin, S. X. and Secor, D. H. (2009). "Accounting for Spatial Population Structure in Stock Assessment: Past, Present and Future." In: *The future of fisheries science in North America*. Ed. by R. J. Beamish, B. J. Rothschild, and American

Institute of Fishery Research Biologists. Fish & fisheries series v. 31. Dordrecht: Springer. isbn: 978-1-4020-9209-1.

Calò, A., Lett, C., Mourre, B., Pérez-Ruzafa, Á., and García-Charton, J. A. (2018). "Use of Lagrangian simulations to hindcast the geographical position of propagule release zones in a Mediterranean coastal fish". In: *Marine Environmental Research* 134, pp. 16–27. issn: 01411136. doi: 10.1016/j.marenvres.2017.12.011.

Carr, M., White, J. W., Saarman, E., Lubchenco, J., and Milligan, K. (2019). "Marine Protected Areas Exemplify the Evolution of Science and Policy". In: *Oceanography* 32.3, pp. 94–103. issn: 10428275. doi: 10.5670/oceanog.2019.315.

Ciliberti, S. A., Pinardi, N., Coppini, G., Oddo, P., Vukicevic, T., Lecci, R., Verri, G., Kumkar, Y., and Creti', S. (2015). "A high resolution Adriatic-Ionian Sea circulation model for operational forecasting". In: *Geophysical Research Abstracts* 17.

Clark, D. L., Leis, J. M., Hay, A. C., and Trnski, T. (2005). "Swimming ontogeny of larvae of four temperate marine fishes". In: *Marine Ecology Progress Series* 292, pp. 287–300. issn: 0171-8630, 1616-1599. doi: 10.3354/meps292287.

Colin, P. L. (2010). "Aggregation and spawning of the humphead wrasse *Cheilinus undulatus* (Pisces: Labridae): general aspects of spawning behaviour". In: *Journal of Fish Biology* 76.4, pp. 987–1007. issn: 00221112, 10958649. doi: 10.1111/j.1095-8649.2010.02553.x.

Coulson, P. G., Norriss, J. V., Jackson, G., and Fairclough, D. V. (2019). "Reproductive characteristics of the fishery important temperate demersal berycid *Centroberyx gerrardi* indicate greater reproductive output in regions of upwelling". In: *Fisheries Management and Ecology* 26, pp. 236–248. issn: 0969-997X, 1365-2400. doi: 10.1111/fme.12343.

Courant, R., Friedrichs, K., and Lewy, H. (1928). "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Mathematische Annalen* 100, pp. 32–74. doi: https://doi.org/10.1007/BF01448839.

Cowen, R. K., Paris, C. B., and Srinivasan, A. (2006). "Scaling of Connectivity in Marine Populations". In: *Science* 311.5760, pp. 522–527. issn: 0036-8075, 1095-9203. doi: 10.1126/science.1122039.

Cowen, R. K. and Sponaugle, S. (2009). "Larval Dispersal and Marine Population Connectivity". In: *Annual Review of Marine Science* 1.1, pp. 443–466. issn: 1941-1405, 1941-0611. doi: 10.1146/annurev.marine.010908.163757.

Crisciani, F. and Mosetti, R. (2016). "Is the Bimodal Oscillating Adriatic-Ionian Circulation a Stochastic Resonance?" In: *Bollettino di Geofisica Teorica ed Applicata* 57.3, pp. 275–285. issn: 0006-6729. doi: 10.4430/bgta0176.

Cuadros, A., Basterretxea, G., Cardona, L., Cheminée, A., Hidalgo, M., and Moranta, J. (2018). "Settlement and post-settlement survival rates of the white seabream (*Diplodus sargus*) in the western Mediterranean Sea". In: *PLOS ONE* 13.1. Ed. by L. Zane, e0190278. issn: 1932-6203. doi: 10.1371/journal.pone.0190278.

Cuttitta, A., Quinci, E. M., Patti, B., Bonomo, S., Bonanno, A., Musco, M., Torri, M., Placenti, F., Basilone, G., Genovese, S., Armeri, G. M., Spanò, A., Arculeo, M., Mazzola, A., and Mazzola, S. (2016). "Different key roles of mesoscale oceanographic structures and ocean bathymetry in shaping larval fish distribution pattern: A case study in Sicilian waters in summer 2009". In: *Journal of Sea Research* 115, pp. 6–17. issn: 1385-1101. doi: 10.1016/j.seares.2016.04.005.

D'Ovidio, F., Fernandez, V., Hernandez-Garcia, E., and Lopez, C. (2004). "Mixing structures in the Mediterranean Sea from finite-size Lyapunov exponents". In: *Geophysical Research Letters* 31, p. L17203. doi: doi:10.1029/2004GL020328.

De Vlaming, V. L. (1972). "Environmental control of teleost reproductive cycles: a brief review". In: *Journal of Fish Biology* 4.1, pp. 131–140. issn: 0022-1112, 1095-8649. doi: 10.1111/j.1095-8649.1972.tb05661.x.

Di Franco, A., Coppini, G., Pujolar, J. M., De Leo, G. A., Gatto, M., Lyubartsev, V., Melià, P., Zane, L., and Guidetti, P. (2012). "Assessing Dispersal Patterns of Fish Propagules from an Effective Mediterranean Marine Protected Area". In: *PLOS ONE* 7.12. Ed. by S. C. A. Ferse, e52108. issn: 1932-6203. doi: 10.1371/journal.pone.0052108.

Di Franco, A., De Benedetto, G., De Rinaldis, G., Raventos, N., Sahyoun, R., and Guidetti, P. (2011). "Large scale-variability in otolith microstructure and microchemistry: The case study of *Diplodus sargus sargus* (Pisces: Sparidae) in the Mediterranean Sea". In: *Italian Journal of Zoology* 78.2, pp. 182–192. issn: 1125-0003, 1748-5851. doi: 10.1080/11250003.2011.566227.

Di Franco, A. and Guidetti, P. (2011). "Patterns of variability in early-life traits of fishes depend on spatial scale of analysis". In: *Biology Letters* 7.3, pp. 454–456. issn: 1744-9561, 1744-957X. doi: 10.1098/rsbl.2010.1149.

Di Franco, A., Plass-Johnson, J. G., Di Lorenzo, M., Meola, B., Claudet, J., Gaines, S. D., García-Charton, J. A., Giakoumi, S., Grorud-Colvert, K., Hackradt, C. W., Micheli, F., and Guidetti, P. (2018). "Linking home ranges to protected area size: The case study of the Mediterranean Sea". In: *Biological Conservation* 221, pp. 175–181. issn: 00063207. doi: 10.1016/j.biocon.2018.03.012.

Di Franco, A., Qian, K., Calò, A., Di Lorenzo, M., Planes, S., and Guidetti, P. (2013). "Patterns of variability in early life traits of a Mediterranean coastal fish". In: *Marine Ecology Progress Series* 476, pp. 227–235. issn: 0171-8630, 1616-1599. doi: 10.3354/meps10117.

Di Franco, A., Thiriet, P., Di Carlo, G., Dimitriadis, C., Francour, P., Gutiérrez, N. L., Jeudy de Grissac, A., Koutsoubas, D., Milazzo, M., Otero, M. d. M., Piante, C., Plass-Johnson, J., Sainz-Trapaga, S., Santarossa, L., Tudela, S., and Guidetti, P. (2016). "Five key attributes can increase marine protected areas performance for small-scale fisheries management". In: *Scientific Reports* 6.1, p. 38135. issn: 2045-2322. doi: 10.1038/srep38135.

Di Lorenzo, M., D'Anna, G., Badalamenti, F., Giacalone, V. M., Starr, R. M., and Guidetti, P. (2014). "Fitting the size of no-take zones to species movement patterns: a case study on a Mediterranean seabream". In: *Marine Ecology Progress Series* 502, pp. 245–255. issn: 0171-8630, 1616-1599. doi: 10.3354/meps10723.

Dubois, M., Rossi, V., Ser-Giacomi, E., Arnaud-Haond, S., López, C., and Hernández-García, E. (2016). "Linking basin-scale connectivity, oceanography and population dynamics for the conservation and management of marine ecosystems: Large-scale connectivity and management of marine ecosystems". In: *Global Ecology and Biogeography* 25.5, pp. 503–515. issn: 1466822X. doi: 10.1111/geb.12431.

El Maghraby, A. M., Hashem, M. T., Botros, G. A., and Wassef, E. A. (1982). "Maturation, spawning and fecundity of two Sparid fish *Diplodus sargus*, L. and *Diplodus vulgaris*, Geoff. in the Egyptian Mediterranean waters". In: *Bulletin of the Institute of Oceanography and Fisheries* 8.2, pp. 51–67.

EMODnet (2013). *Seabed Habitats*.

Erisman, B., Heyman, W., Kobara, S., Ezer, T., Pittman, S., Aburto-Oropeza, O., and Nemeth, R. S. (2015). "Fish spawning aggregations: where well-placed management actions can yield big benefits for fisheries and conservation". In: *Fish and Fisheries* 18.1, pp. 128–144. issn: 14672960. doi: 10.1111/faf.12132.

Faillettaz, R., Durand, E., Paris, C. B., Koubbi, P., and Irisson, J.-O. (2018). "Swimming speeds of Mediterranean settlement-stage fish larvae nuance Hjort's aberrant drift hypothesis: Larval swimming nuances aberrant drift hypothesis". In: *Limnology and Oceanography* 63.2, pp. 509–523. issn: 00243590. doi: 10.1002/lno.10643.

FAO (2004). *FAO yearbook. 94,1: Fishery statistics, Capture production 2002*. FAO fisheries series 66. isbn: 978-92-5-005139-0.

FAO (2006). *Report of the ninth session of the Scientific Advisory Committee / General Fisheries Commission for the Mediterranean: Rome, 24-27 October 2006 = Rapport de la neuvième session du Comité Scientifique Consultatif*. FAO Fisheries Report 814. Rome: Food and Agriculture Organization of the United Nations. isbn: 978-92-5-005637-1.

FAO (2018). *The State of Mediterranean and Black Sea fisheries / General Fisheries Commission for the Mediterranean: Rome.* isbn: 978-92-5-131152-3.

FAO (2021). *Fisheries Restricted Areas | General Fisheries Commission for the Mediterranean - GFCM | Food and Agriculture Organization of the United Nations*.

Fincham, J. I., Rijnsdorp, A. D., and Engelhard, G. H. (2013). "Shifts in the timing of spawning in sole linked to warming sea temperatures". In: *Journal of Sea Research* 75, pp. 69–76. issn: 13851101. doi: 10.1016/j.seares.2012.07.004.

Fogarty, M. J. and Botsford, L. W. (2007). "Population Connectivity and Spatial Management of Marine Fisheries". In: *Oceanography* 20.3, pp. 112–123. issn: 10428275. doi: 10.5670/oceanog.2007.34.

Gačić, M., Civitarese, G., Kovačević, V., Ursella, L., Bensi, M., Menna, M., Cardin, V., Poulain, P.-M., Cosoli, S., Notarstefano, G., and Pizzi, C. (2014). "Extreme winter 2012 in the Adriatic: an example of climatic effect on the BiOS rhythm". In: *Ocean Science* 10.3, pp. 513–522. issn: 1812-0792. doi: 10.5194/os-10-513-2014.

Gačić, M., Eusebi Borzelli, G. L., Civitarese, G., Cardin, V., and Yari, S. (2010). "Can internal processes sustain reversals of the ocean upper circulation? The Ionian Sea example". In: *Geophysical Research Letters* 37.9, p. L09608. issn: 00948276. doi: 10.1029/2010GL043216.

Gačić, M., Schroeder, K., Civitarese, G., Cosoli, S., Vetrano, A., and Eusebi Borzelli, G. L. (2013). "Salinity in the Sicily Channel corroborates the role of the Adriatic–Ionian Bimodal Oscillating System (BiOS) in shaping the decadal variability of the Mediterranean overturning circulation". In: *Ocean Science* 9.1, pp. 83–90. issn: 1812-0792. doi: 10.5194/os-9-83-2013.

Gaines, S. D., Gaylord, B., Gerber, L. R., Hastings, A., and Kinlan, B. P. (2007). "Connecting Places: The Ecological Consequences of Dispersal in the Sea". In: *Oceanography* 20.3, pp. 90–99.

Gaines, S. D., White, C., Carr, M. H., and Palumbi, S. R. (2010). "Designing marine reserve networks for both conservation and fisheries management". In: *Proceedings of the National Academy of Sciences* 107.43, pp. 18286–18293. issn: 0027-8424, 1091-6490. doi: 10.1073/pnas.0906473107.

Garratt, P. A. (1993). "Spawning of riverbream, *Acanthopagrus berda*, in Kosi estuary". In: *South African Journal of Zoology* 28.1, pp. 26–31. issn: 0254-1858. doi: 10.1080/02541858.1993.11448284.

Giacalone, V. M., Pipitone, C., Badalamenti, F., Sacco, F., Zenone, A., Ferreri, R., Micale, V., Basilone, G., and D'Anna, G. (2018). "Home range, movements and daily activity of the white seabream Diplodus sargus (Linnaeus, 1758) during the spawning season". In: *Cahiers de Biologie Marine* 59, pp. 421–429. issn: 2262-3094. doi: 10.21411/CBM.A.7C19C1B8.

Giakoumi, S., Scianna, C., Plass-Johnson, J., Micheli, F., Grorud-Colvert, K., Thiriet, P., Claudet, J., Di Carlo, G., Di Franco, A., Gaines, S. D., García-Charton, J. A., Lubchenco, J., Reimer, J., Sala, E., and Guidetti, P. (2017). "Ecological effects of full and partial protection in the crowded Mediterranean Sea: a regional meta-analysis". In: *Scientific Reports* 7.1, p. 8940. issn: 2045-2322. doi: 10.1038/s41598-017-08850-w.

Gill, H. S., Wise, B. S., Potter, I. C., and Chaplin, J. A. (1996). "Biannual spawning periods and resultant divergent patterns of growth in the estuarine goby *Pseudogobius olorum*: temperature-induced?" In: *Marine Biology* 125, pp. 453–466.

Gonçalves, J. M. D. S. (2000). "Fisheries Biology and Population Dynamics of *Diplodus vulgaris* (Geoffr.) and *Spondyliosoma cantharus* (L.) (Pisces, Sparidae) from the southwest coast of Portugal." PhD thesis. Unidade de Ciências e Tecnologias dos recursos Aquaticos: Universidade do Algarve.

Guidetti, P. (2000). "Differences Among Fish Assemblages Associated with Nearshore *Posidonia oceanica* Seagrass Beds, Rocky–algal Reefs and Unvegetated Sand Habitats in the Adriatic Sea". In: *Estuarine, Coastal and Shelf Science* 50.4, pp. 515–529. issn: 02727714. doi: 10.1006/ecss.1999.0584.

Guidetti, P., Baiata, P., Ballesteros, E., Di Franco, A., Hereu, B., Macpherson, E., Micheli, F., Pais, A., Panzalis, P., Rosenberg, A. A., Zabala, M., and Sala, E. (2014). "Large-Scale Assessment of Mediterranean Marine Protected Areas Effects on Fish Assemblages". In: *PLoS ONE* 9.4. Ed. by B. R. MacKenzie, e91841. issn: 1932-6203. doi: 10.1371/journal.pone.0091841.

Harmelin-Vivien, M. L., Harmelin, J.-G., and Leboulleux, V. (1995). "Microhabitat requirements for settlement of juvenile sparid fishes on Mediterranean rocky shores". In: *Hydrobiologia* 300/301, pp. 309–320.

Hereu, B., Diaz, D., Pasqual, J., Zabala, M., and Sala, E. (2006). "Temporal patterns of spawning of the dusky grouper *Epinephelus marginatus* in relation to environmental factors". In: *Marine Ecology Progress Series* 325, pp. 187–194. issn: 0171-8630, 1616-1599. doi: 10.3354/meps325187.

Hereu, B., Zabala, M., Linares, C., and Sala, E. (2005). "The effects of predator abundance and habitat structural complexity on survival of juvenile sea urchins". In: *Marine Biology* 146.2, pp. 293–299. issn: 0025-3162, 1432-1793. doi: 10.1007/s00227-004-1439-y.

Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernández-García, E., Guijarro, B., Massutí, E., Alemany, F., Jadaud, A., Perez, J. L., and Reglero, P. (2019). "Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations". In: *Ecological Applications* 29.5, e01913. issn: 1051-0761, 1939-5582. doi: 10.1002/eap.1913.

Huret, M., Petitgas, P., and Woillez, M. (2010). "Dispersal kernels and their drivers captured with a hydrodynamic model and spatial indices: A case study on anchovy (Engraulis encrasicolus) early life stages in the Bay of Biscay". In: *Progress in Oceanography* 87.1-4, pp. 6–17. issn: 00796611. doi: 10.1016/j.pocean.2010.09.023.

Hussein, C., Verdoit-Jarraya, M., Pastor, J., Ibrahim, A., Saragoni, G., Pelletier, D., Mahévas, S., and Lenfant, P. (2011). "Assessing the impact of artisanal and recreational fishing and protection on a white seabream (Diplodus sargus sargus) population in the north-western Mediterranean Sea, using a simulation model. Part 2: Sensitivity analysis and management measures". In: *Fisheries Research* 108.1, pp. 174–183. issn: 01657836. doi: 10.1016/j.fishres.2010.12.018.

Iacono, R., Napolitano, E., Palma, M., and Sannino, G. (2021). "The Tyrrhenian Sea Circulation: A Review of Recent Work". In: *Sustainability* 13.11, p. 6371. issn: 2071-1050. doi: 10.3390/su13116371.

Jones, G. P., Srinivasan, M., and Almany, G. R. (2007). "Population Connectivity and Conservation of Marine Biodiversity". In: *Oceanography* 20.3, pp. 100–111. issn: 10428275. doi: 10.5670/oceanog.2007.33.

Kadota, T., Sakai, Y., Hashimoto, H., and Gushima, K. (2010). "Diel and lunar spawning periodicity of the hawkfish *Paracirrhites forsteri* (Cirrhitidae) on the reefs of Kuchierabu-jima Island, southern Japan". In: *Ichthyological Research* 57.1, pp. 102–106. issn: 1341-8998, 1616-3915. doi: 10.1007/s10228-009-0124-z.

King, A. J., Gwinn, D. C., Tonkin, Z., Mahoney, J., Raymond, S., and Beesley, L. (2016). "Using abiotic drivers of fish spawning to inform environmental flow management". In: *Journal of Applied Ecology* 53.1. Ed. by J. Heino, pp. 34–43. issn: 00218901. doi: 10.1111/1365-2664.12542.

Lê, S., Josse, J., and Husson, F. (2008). "**FactoMineR** : An *R* Package for Multivariate Analysis". In: *Journal of Statistical Software* 25.1, pp. 1–18. issn: 1548-7660. doi: 10.18637/jss.v025.i01.

Legrand, T., Di Franco, A., Ser-Giacomi, E., Caló, A., and Rossi, V. (2019). "A multidisciplinary analytical framework to delineate spawning areas and quantify larval dispersal in coastal fish". In: *Marine Environmental Research* 151, pp. 1–13. issn: 0141-1136. doi: 10.1016/j.marenvres.2019.104761.

Lenfant, P. and Planes, S. (1996). "Genetic differentiation of white sea bream within the Lion's Gulf and the Ligurian Sea (Mediterranean Sea)". In: *Journal of Fish Biology* 49.4, pp. 613–621. issn: 0022-1112, 1095-8649. doi: 10.1111/j.1095-8649.1996.tb00058.x.

Lester, S. E., Halpern, B. S., Grorud-Colvert, K., Lubchenco, J., Ruttenberg, B. I., Gaines, S. D., Airamé, S., and Warner, R. R. (2009). "Biological effects within no-take marine reserves: a global synthesis". In: *Marine Ecology Progress Series* 384, pp. 33–46. doi: 10.3354/meps08029.

Levin, S. A. and Lubchenco, J. (2008). "Resilience, Robustness, and Marine Ecosystem-based Management". In: *BioScience* 58.1, pp. 27–32. issn: 1525-3244, 0006-3568. doi: 10.1641/B580107.

Liao, E., Lu, W., Yan, X.-H., Jiang, Y., and Kidwell, A. (2015). "The coastal ocean response to the global warming acceleration and hiatus". In: *Scientific Reports* 5.1, p. 16630. issn: 2045-2322. doi: 10.1038/srep16630.

Lima, F. P. and Wethey, D. S. (2012). "Three decades of high-resolution coastal sea surface temperatures reveal more than warming". In: *Nature Communications* 3.1, p. 704. issn: 2041-1723. doi: 10.1038/ncomms1713.

Lloret, J. and Planes, S. (2003). "Condition, feeding and reproductive potential of white seabream *Diplodus sargus* as indicators of habitat quality and the effect of reserve protection in the northwestern Mediterranean". In: *Marine Ecology Progress Series* 248, pp. 197–208. issn: 0171-8630, 1616-1599. doi: 10.3354/meps248197.

Lluch-Belda, D., Lluch-Cota, D. B., Hernandez-Vazquez, S., Salinas-Zavala, C. A., and Schwartzlose, R. A. (1991). "Sardine and Anchovy Spawning as related to Temperature and Upwelling in the California Current System". In: 32, pp. 105–111.

Madec, G. and the NEMO team (2008). "NEMO ocean engine". In: *Institut Pierre-Simon Laplace*. Note du Pôle de modélisation 27, p. 412. issn: 1288-1619.

Magaldi, M. G., Özgökmen, T. M., Griffa, A., and Rixen, M. (2010). "On the response of a turbulent coastal buoyant current to wind events: the case of the Western Adriatic Current". In: *Ocean Dynamics* 60.1, pp. 93–122. issn: 1616-7341, 1616-7228. doi: 10.1007/s10236-009-0247-9.

Man-Wai, R. (1985). "Les sars du Golfe du Lion: *Diplodus sargus*, *D. vulgaris*, *D. annularis* (Pisces, Sparidae)." Ecobiologie-Pêche. Université des sciences et Techniques du Languedoc.

Marshall, D. J., Gaines, S. D., Warner, R. R., Barneche, D. R., and Bode, M. (2019). "Underestimating the benefits of marine protected areas for the replenishment of fished populations". In: *Frontiers in Ecology and the Environment* 17.7, pp. 407–413. issn: 1540-9295, 1540-9309. doi: 10.1002/fee.2075.

Maxwell, S. M., Hazen, E. L., Lewison, R. L., Dunn, D. C., Bailey, H., Bograd, S. J., Briscoe, D. K., Fossette, S., Hobday, A. J., Bennett, M., Benson, S., Caldwell, M. R., Costa, D. P., Dewar, H., Eguchi, T., Hazen, L., Kohin, S., Sippel, T., and Crowder, L. B. (2015). "Dynamic ocean management: Defining and conceptualizing real-time management of the ocean". In: *Marine Policy* 58, pp. 42–50. issn: 0308597X. doi: 10.1016/j.marpol.2015.03.014.

McQueen, K. and Marshall, C. T. (2017). "Shifts in spawning phenology of cod linked to rising sea temperatures". In: *ICES Journal of Marine Science* 74.6. Ed. by V. Bartolino, pp. 1561–1573. issn: 1054-3139, 1095-9289. doi: 10.1093/icesjms/fsx025.

Millot, C. and Taupier-Letage, I. (2005). "Circulation in the Mediterranean Sea". In: *The Mediterranean Sea*. Ed. by A. Saliot. Vol. 5-K. The Mediterranean Sea. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 29–66. isbn: 978-3-540-25018-0. doi: 10.1007/b107143.

Mischke, C. C. and Morris, J. E. (1997). "Out-of-Season Spawning of Sunfish *Lepomis* spp. in the Laboratory". In: *The Progressive-Fish Culturist* 59.4, pp. 297–302. doi: 10.1577/1548-8640(1997)059<0297:OOSSOS>2.3.CO;2.

Monroy, P., Rossi, V., Ser-Giacomi, E., López, C., and Hernández-García, E. (2017). "Sensitivity and robustness of larval connectivity diagnostics obtained from Lagrangian Flow Networks". In: *ICES Journal of Marine Science* 74.6. Ed. by C. Paris, pp. 1763–1779. issn: 1054-3139, 1095-9289. doi: 10.1093/icesjms/fsw235.

Morato, T., Afonso, P., Lourinho, P., Nash, R. D. M., and Santos, R. S. (2003). "Reproductive biology and recruitment of the white sea bream in the Azores". In: *Journal of Fish Biology* 63.1, pp. 59–72. issn: 0022-1112, 1095-8649. doi: 10.1046/j.1095-8649.2003.00129.x.

Mouine, N., Francour, P., Ktari, M.-H., and Chakroun-Marzouk, N. (2007). "The reproductive biology of *Diplodus sargus sargus* in the Gulf of Tunis (central Mediterranean)". In: *Scientia Marina* 71.3, pp. 461–469. issn: 0214-8358. doi: 10.3989/scimar.2007.71n3461.

Mouine, N., Francour, P., Ktari, M.-H., and Chakroun-Marzouk, N. (2012). "Reproductive biology of four *Diplodus* species *Diplodus vulgaris*, *D. annularis*, *D. sargus sargus* and *D. puntazzo* (Sparidae) in the Gulf of Tunis (central

Mediterranean)". In: *Journal of the Marine Biological Association of the United Kingdom* 92.3, pp. 623–631. issn: 0025-3154, 1469-7769. doi: 10.1017/S0025315411000798.

Mouine, N., Ktari, M.-H., and Chakroun-Marzouk, N. (2011). "Reproductive characteristics of *Spondyliosoma cantharus* (Linnaeus, 1758) in the Gulf of Tunis: Reproductive characteristics of *S. cantharus*". In: *Journal of Applied Ichthyology* 27.3, pp. 827–831. issn: 01758659. doi: 10.1111/j.1439-0426.2010.01518.x.

Nathan, R., Perry, G., Cronin, J. T., Strand, A. E., and Cain, M. L. (2003). "Methods for estimating long-distance dispersal". In: *Oikos* 103.2, pp. 261–273. issn: 00301299. doi: 10.1034/j.1600-0706.2003.12146.x.

National Geophysical Data Center (2021). *ETOPO1 Global Relief Model*.

Oddo, P., Pinardi, N., Zavatarelli, M., and Coluccelli, A. (2006). "The Adriatic Basin forecasting system". In: *ACTA ADRIATICA* 47.Suppl. Pp. 169–184. issn: 0001-5113.

Olivar, M. P. and Sabatés, A. (1997). "Vertical distribution of fish larvae in the north-west Mediterranean Sea in spring". In: *Marine Biology* 129, pp. 289–300.

Pajuelo, J. G., Lorenzo, J. M., Bilbao, A., Ayza, O., and Ramos, A. G. (2006). "Reproductive characteristics of the benthic coastal fish *Diplodus vulgaris* (Teleostei: Sparidae) in the Canarian archipelago, northwest Africa". In: *Journal of Applied Ichthyology* 22.5, pp. 414–418. issn: 0175-8659, 1439-0426. doi: 10.1111/j.1439-0426.2006.00766.x.

Pajuelo, J. G., Lorenzo, J. M., and Domínguez-Seoane, R. (2008). "Gonadal development and spawning cycle in the digynic hermaphrodite sharpsnout seabream *Diplodus puntazzo* (Sparidae) off the Canary Islands, northwest of Africa". In: *Journal of Applied Ichthyology* 24, pp. 68–76. issn: 0175-8659, 1439-0426. doi: 10.1111/j.1439-0426.2007.01010.x.

Pajuelo, J. G., Lorenzo, J. M., Domınguez, R., Ramos, A., and Gregoire, M. (2003). "On the population ecology of the zebra seabream *Diplodus cervinus cervinus* (Lowe 1838) from the coasts of the Canarian archipelago, North West Africa". In: *Environmental Biology of Fishes* 67, pp. 407–416.

Pelc, R. A., Warner, R. R., Gaines, S. D., and Paris, C. B. (2010). "Detecting larval export from marine reserves". In: *Proceedings of the National Academy of Sciences* 107.43, pp. 18266–18271. issn: 0027-8424, 1091-6490. doi: 10.1073/pnas.0907368107.

Perry, A. L., Low, P. J., Ellis, J. R., and Reynolds, J. D. (2005). "Climate Change and Distribution Shifts in Marine Fishes". In: *Science* 308.5730, pp. 1912–1915. doi: 10.1126/science.1111322.

Piccioni, A., Gabriele, M., Salusti, E., and Zambianchi, E. (1988). "Wind-induced upwellings off the southern coast of Sicily". In: *Oceanologica Acta* 11.4, pp. 309–314.

Pineda, J., Hare, J. A., and Sponaugle, S. (2007). "Larval Transport and Dispersal in the Coastal Ocean and Consequences for Population Connectivity". In: *Oceanography* 20.3, pp. 22–39.

Pittman, S. J. and Heyman, W. D. (2020). "Life below water: Fish spawning aggregations as bright spots for a sustainable ocean". In: *Conservation Letters* 13.5. issn: 1755-263X, 1755-263X. doi: 10.1111/conl.12722.

Planes, S., Macpherson, E., Biagi, F., Garcia-Rubies, A., Harmelin, J.-G., Harmelin-Vivien, M. L., Jouvenel, J.-Y., Tunesi, L., Vigliola, L., and Galzin, R. (1999). "Spatio-temporal variability in growth of juvenile sparid fishes from the Mediterranean littoral zone". In: *Journal of the Marine Biological Association of the United Kingdom* 79.1, pp. 137–143. issn: 0025-3154, 1469-7769. doi: 10.1017/S0025315498000150.

Potts, W. M., Booth, A. J., Richardson, T. J., and Sauer, W. H. H. (2014). "Ocean warming affects the distribution and abundance of resident fishes by changing their reproductive scope". In: *Reviews in Fish Biology and Fisheries* 24.2, pp. 493–504. issn: 0960-3166, 1573-5184. doi: 10.1007/s11160-013-9329-3.

Poulain, P.-M., Bussani, A., Gerin, R., Jungwirth, R., Mauri, E., Menna, M., and Notarstefano, G. (2013). "Mediterranean Surface Currents Measured with Drifters: From Basin to Subinertial Scales". In: *Oceanography* 26.1, pp. 38–47. issn: 10428275. doi: 10.5670/oceanog.2013.03.

Poulain, P.-M., Menna, M., and Mauri, E. (2012). "Surface Geostrophic Circulation of the Mediterranean Sea Derived from Drifter and Satellite Altimeter Data". In: *Journal of Physical Oceanography* 42.6, pp. 973–990. issn: 0022-3670, 1520-0485. doi: 10.1175/JPO-D-11-0159.1.

Pujolar, J. M., Schiavina, M., Di Franco, A., Melià, P., Guidetti, P., Gatto, M., De Leo, G. A., and Zane, L. (2013). "Understanding the effectiveness of marine protected areas using genetic connectivity patterns and Lagrangian simulations". In: *Diversity and Distributions* 19.12. Ed. by D. Richardson, pp. 1531–1542. issn: 13669516. doi: 10.1111/ddi.12114.

Reglero, P., Ciannelli, L., Alvarez-Berastegui, D., Balbín, R., López-Jurado, J. L., and Alemany, F. (2012). "Geographically and environmentally driven spawning distributions of tuna species in the western Mediterranean Sea". In: *Marine Ecology Progress Series* 463, pp. 273–284. issn: 0171-8630, 1616-1599. doi: 10.3354/meps09800.

Reglero, P., Alvarez-Berastegui, D., Alemany, F., Rossi, V., Torres, A. P., Balbin, R., and Hidalgo, M. (2018). "Operational Oceanography and the Management of Pelagic Resources: The Mediterranean Sea as a Case-Study". In: *New Frontiers in Operational Oceanography*. Ed. by E. P. Chassignet, A. Pascual, J. Tintoré, and J. Verron. GODAE OceanView, pp. 713–728. isbn: 978-1-72054-997-0. doi: 10.17125/gov2018.ch26.

Reznick, D. N. and Braun, B. (1987). "Fat cycling in the mosquitofish (*Gambusia affinis*): fat storage as a reproductive adaptation". In: *Oecologia* 73.3, pp. 401–413. issn: 0029-8549, 1432-1939. doi: 10.1007/BF00385257.

Rossi, V., Lo, M., Legrand, T., Ser-Giacomi, E., De Jode, A., Thierry De Ville D'avray, L., Pairaud, I., Faure, V., Fraysse, M., Pinazo, C., and Chenuil, A. (2020). "Small-scale connectivity of coralligenous habitats: insights from a modeling approach within a semi-opened Mediterranean bay". In: *Life and Environment*, p. 14.

Rossi, V., Ser-Giacomi, E., López, C., and Hernández-García, E. (2014). "Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves". In: *Geophysical Research Letters* 41.8, pp. 2883–2891. issn: 00948276. doi: 10.1002/2014GL059540.

Sadovy, Y. J. and Domeier, M. (2005). "Are aggregation-fisheries sustainable? Reef fish fisheries as a case study". In: *Coral Reefs* 24.2, pp. 254–262. issn: 0722-4028, 1432-0975. doi: 10.1007/s00338-005-0474-6.

Sadovy de Mitcheson, Y. J. (2016). "Mainstreaming Fish Spawning Aggregations into Fishery Management Calls for a Precautionary Approach". In: *BioScience* 66.4, pp. 295–306. issn: 1525-3244, 0006-3568. doi: 10.1093/biosci/biw013.

Sala, E., Ballesteros, E., Dendrinos, P., Di Franco, A., Ferretti, F., Foley, D., Fraschetti, S., Friedlander, A., Garrabou, J., Güçlüsoy, H., Guidetti, P., Halpern, B. S., Hereu, B., Karamanlidis, A. A., Kizilkaya, Z., Macpherson, E., Mangialajo, L., Mariani, S., Micheli, F., Pais, A., Riser, K., Rosenberg, A. A., Sales, M., Selkoe, K. A., Starr, R., Tomas, F., and Zabala, M. (2012). "The Structure of Mediterranean Rocky Reef Ecosystems across Environmental and Human Gradients, and Conservation Implications". In: *PLoS ONE* 7.2. Ed. by T. N. Romanuk, e32742. issn: 1932-6203. doi: 10.1371/journal.pone.0032742.

Sala, E., Boudouresque, C.-F., and Harmelin-Vivien, M. L. (1998). "Fishing, Trophic Cascades, and the Structure of Algal Assemblages: Evaluation of an Old but Untested Paradigm". In: *Oikos* 82.3, pp. 425–439. issn: 0030-1299. doi: 10.2307/3546364.

Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E. (2016). "mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models". In: *The R Journal* 8.1, pp. 289–317. issn: 2073-4859. doi: 10.32614/RJ-2016-021.

Ser-Giacomi, E., Rossi, V., Lopez, C., and Hernandez-Garcia, E. (2015). "Flow networks: A characterization of geophysical fluid transport". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.3, pp. 1–15. issn: 1054-1500, 1089-7682. doi: 10.1063/1.4908231.

Sheaves, M. (2006). "Is the timing of spawning in sparid fishes a response to sea temperature regimes?" In: *Coral Reefs* 25.4, pp. 655–669. issn: 0722-4028, 1432-0975. doi: 10.1007/s00338-006-0150-5.

Sheaves, M. and Molony, B. (2013). "Reproductive periodicity of the sparid, *Acanthopagrus pacificus*, on a hierarchy of temporal scales". In: *Journal of Fish Biology* 82.2, pp. 538–554. issn: 00221112. doi: 10.1111/jfb.12010.

Simonoff, J. S. (1996). *Smoothing methods in statistics*. Springer series in statistics. Springer. isbn: 0-387-94716-7.

Staaterman, E., Paris, C. B., and Helgers, J. (2012). "Orientation behavior in fish larvae: A missing piece to Hjort's critical period hypothesis". In: *Journal of Theoretical Biology* 304, pp. 188–196. issn: 0022-5193. doi: 10.1016/j.jtbi.2012.03.016.

Takahashi, M., McCormick, M. I., Munday, P. L., and Jones, G. P. (2012). "Influence of seasonal and latitudinal temperature variation on early life-history traits of a coral reef fish". In: *Marine and Freshwater Research* 63.10, p. 856. issn: 1323-1650. doi: 10.1071/MF11278.

Thresher, R. E., Colin, P. L., and Bell, L. J. (1989). "Planktonic Duration, Distribution and Population Structure of Western and Central Pacific Damselfishes (Pomacentridae)". In: *Copeia* 1989.2, p. 420. issn: 00458511. doi: 10.2307/1445439.

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Vanden Berghe, E., and Worm, B. (2010). "Global patterns and predictors of marine biodiversity across taxa". In: *Nature* 466.7310, pp. 1098–1101. issn: 0028-0836, 1476-4687. doi: 10.1038/nature09329.

Tsikliras, A. C., Dinouli, A., Tsiros, V.-Z., and Tsalkou, E. (2015). "The Mediterranean and Black Sea Fisheries at Risk from Overexploitation". In: *PLOS ONE* 10.3. Ed. by C. Durif, e0121188. issn: 1932-6203. doi: 10.1371/journal.pone.0121188.

Vigliola, L. (1998). "Contrôle et régulation du recrutement des Sparidae (Poissons, Téléostéens) en Méditerranée : importance des processus pré- et post-installation benthique". PhD thesis. Centre d'Océanologie de Marseille: Université de la Méditerranée (Aix Marseille II).

Vinagre, C., Ferreira, T., Matos, L., Costa, M. J., and Cabral, H. N. (2009). "Latitudinal gradients in growth and spawning of sea bass, *Dicentrarchus labrax*, and their relationship with temperature and photoperiod". In: *Estuarine, Coastal and Shelf Science* 81.3, pp. 375–380. issn: 02727714. doi: 10.1016/j.ecss.2008.11.015.

Werner, F. E., Cowen, R. K., and Paris, C. B. (2007). "Coupled Biological and Physical Models: Present Capabilities and Necessary Developments for Future Studies of Population Connectivity". In: *Oceanography* 20.3, pp. 54–69.

Winters, G. H. and Wheeler, J. P. (1996). "Environmental and phenotypic factors affecting the reproductive cycle of Atlantic herring". In: *ICES Journal of Marine Science* 53.1, pp. 73–88. issn: 10543139. doi: 10.1006/jmsc.1996.0007.

Wright, P. J. and Trippel, E. A. (2009). "Fishery-induced demographic changes in the timing of spawning: consequences for reproductive success". In: *Fish and Fisheries* 10.3, pp. 283–304. issn: 14672960, 14672979. doi: 10.1111/j.1467-2979.2008.00322.x.

Yamahira, K. (1997). "Hatching success affects the timing of spawning by the intertidally spawning puffer *Takifugu niphobles*". In: *Marine Ecology Progress Series* 155, pp. 239–248. issn: 0171-8630, 1616-1599. doi: 10.3354/meps155239.

B. Chapitre IV

B.1. SI - Coalescent connectivity through multi-generation dispersal modelling predicts gene flow across marine phyla

Supplementary Information:

Coalescent connectivity through multi-generation dispersal modelling predicts gene flow across marine phyla

Térence Legrand[1]; Anne Chenuil[2]; Enrico Ser-Giacomi[3]; Sophie Arnaud-Haond[4]; Nicolas Bierne[5]; Vincent Rossi[6]

[1]Aix Marseille University, Universite de Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (UMR 7294), Marseille, France. (terence.legrand@mio.osupytheas.fr). Corresponding author.
[2] Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, CNRS (UMR 7263), Station Marine d'Endoume, Marseille, France. (anne.chenuil@imbe.fr).
[3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, 54-1514 MIT, Cambridge, Massachusetts, USA. (enrico.sergiacomi@gmail.com).
[4] Marine Biodiversity, Exploitation and Conservation, UMR 9190 IRD – IFREMER – UM – CNRS, Sète, France. (sophie.arnaud-haond@umontpellier.fr).
[5] Institut des Sciences de l'Evolution UMR 5554, Université de Montpellier, CNRS-IRD-EPHE-UM, Montpellier, France. (nicolas.bierne@umontpellier.fr).
[6]Aix Marseille University, Universite de Toulon, CNRS, IRD, Mediterranean Institute of Oceanography (UMR 7294), Marseille, France. (vincent.rossi@mio.osupytheas.fr).

I. Mediterranean meta-analysis: selected references, model parameterization and literature review

Table SI-1: ***Summarized description of the 58 population genetic studies included in the meta-analysis****. Note that some results reported by a given study are analysed separately: (i) Weber et al., 2015 used SNPs marker ([1]) and mtDNA marker ([2]); (ii) Carrera et al., 2020 considered all the loci together ([3]) and then only the Mediterranean outliers loci ([4]); (iii) Marzouk et al., 2017 analyzed SNPs marker ([5]) and mtDNA marker ([6]). Asterisks indicate the significance of the Mantel tests performed between explicit/implicit modelled Fst and observed Fst for each study: * ≤ 0.05; ** ≤ 0.005 and *** ≤ 0.0005 and "ns" stands for not-significant. SSRep stands for spatial sampling representativeness, i.e. evaluating the effective spatial scattering and seascape coverage by the sampling carried out in each study.*

Species characteristics					Study characteristics					
Taxa	Species	Habitat	PLD	Season	Study	Marker	Nbr of population	SSRep	Optimal M explicit	Optimal M implicit
Algea	*Cystoseira amentacea*	shallow coastal	1	all	Susini et al., 2007	RAPD	4	161	5 ns	1 ns
Anthozoa	*Astroides calycularis*	shallow coastal	1	summer	Casado-Amezua et al., 2012	microsat	16	1172	5 ***	300 *
Anthozoa	*Corallium rubrum*	neritic shelf	10	summer	Aurelle et al., 2011	microsat	24	687	40 ***	80 **
Anthozoa	*Corallium rubrum*	neritic shelf	10	summer	Costantini et al., 2013	microsat	5	234	1 ns	1 ns
Anthozoa	*Eunicella cavolinii*	shallow coastal	1	all	Masmoudi et al., 2016	microsat	18	821	300 ***	300 **
Anthozoa	*Leptopsammia pruvoti*	shallow coastal	10	spring	Boscari et al., 2019	RADseq/GBS	10	385	10 ***	20 **
Ascidiacea	*Botryllus schlosseri*	shallow coastal	1	all	Reem et al., 2017	microsat	11	1182	500 *	500 *
Ascidiacea	*Halocynthia papillosa*	neritic shelf	20	summer	Villamor et al., 2014	mtDNA	4	444	5 ns	10 ns
Ascidiacea	*Microcosmus squamiger*	shallow coastal	1	summer	Ordóñez et al., 2013	microsat	6	18	1 ns	1 ns
Ascidiacea	*Pycnoclavella communis*	shallow coastal	1	all	Pérez-Portela et al., 2007	microsat	4	459	500 ns	500 ns
Crustacea	*Carcinus aestuarii*	shallow coastal	30	fall	Schiavina et al., 2014	microsat	8	417	60 *	40 *
Crustacea	*Melicertus kerathurus*	shallow coastal	20	all	Arculeo et al., 2010	microsat	9	614	1 ns	10 ns
Crustacea	*Melicertus kerathurus*	shallow coastal	20	all	Zitari-Chatti et al., 2007	zymes	9	173	1 ns	5 ns
Crustacea	*Pachygrapsus marmoratus*	shallow coastal	30	all	Fratini et al., 2013	microsat	8	57	500 ns	5 ns
Crustacea	*Palinurus elephas*	shallow coastal	45	spring	Palero et al., 2011	microsat	5	894	40 ns	20 ns
Demospongiae	*Spongia officinalis*	shallow coastal	1	summer	Dailianis et al., 2011	microsat	9	1106	500 *	500 *
Echinodermata	*Astropecten aranciacus*	shallow coastal	45	all	Zulliger et al., 2009	microsat	7	1001	10 ns	1 ns
Echinodermata	*Holothuria mammata*	shallow coastal	20	spring	Borrero-Pérez et al., 2011	mtDNA	4	1188	150 *	60 ns
Echinodermata	*Ophioderma longicauda*	shallow coastal	10	spring	Weber et al., 2015[1]	mtDNA	13	1142	500 ns	500 ns
Echinodermata	*Ophioderma longicauda*	shallow coastal	10	spring	Weber et al., 2015[2]	SNPs	11	1324	200 ns	500 ns
Echinodermata	*Paracentrotus lividus*	shallow coastal	30	spring	Penant et al., 2013	mtDNA	12	1019	500 ns	40 ns
Echinodermata	*Paracentrotus lividus*	shallow coastal	30	spring	Paterno et al., 2017	SNPs	10	758	40 **	40 **
Echinodermata	*Paracentrotus lividus*	shallow coastal	30	spring	Carreras et al., 2020[3]	RADseq/GBS	8	1296	40 ns	40 ns
Echinodermata	*Paracentrotus lividus*	shallow coastal	30	spring	Carreras et al., 2020[4]	RADseq/GBS	8	1296	40 *	40 *
Fish	*Apogon imberbis*	shallow coastal	20	summer	Muths et al., 2015	microsat	4	17	5 ns	1 ns
Fish	*Coris julis*	shallow coastal	30	summer	Fruciano et al., 2011	mtDNA	10	819	80 ns	300 ns
Fish	*Diplodus sargus*	shallow coastal	20	spring	González-Wangüemert et al., 2010	microsat	5	590	150 ns	10 ns
Fish	*Diplodus vulgaris*	shallow coastal	45	winter	Kaouèche et al., 2013	zymes	6	179	5 ns	1 ns
Fish	*Epinephelus marginatus*	shallow coastal	30	summer	Schunter et al., 2011	microsat	9	669	10 ns	20 ns
Fish	*Lithognathus mormyrus*	shallow coastal	30	all	Hammami et al., 2007	zymes	4	180	5 ns	500 ns
Fish	*Merluccius merluccius*	neritic shelf	45	all	Milano et al., 2014	SNPs	14	992	10 *	10 *
Fish	*Mugil cephalus*	shallow coastal	45	all	Durand et al., 2013	microsat	12	1221	20 **	20 **
Fish	*Mullus barbartus*	shallow coastal	30	spring	Maggio et al., 2009	microsat	14	513	1 ns	1 ns
Fish	*Mullus surmuletus*	shallow coastal	30	spring	Galarza et al., 2009	microsat	6	1066	1 ns	60 ns
Fish	*Mullus surmuletus*	shallow coastal	30	spring	Dalongeville et al., 2018	SNPs	47	1163	60 ns	100 *
Fish	*Oblada melanura*	shallow coastal	20	spring	Gkafas et al., 2013	microsat	8	198	10 ns	500 ns
Fish	*Oblada melanura*	shallow coastal	20	spring	Calò et al., 2016	microsat	9	376	1 ns	1 ns
Fish	*Pagellus erythrinus*	shallow coastal	45	spring	Fassatoui et al., 2009	zymes	6	243	1 ns	20 ns
Fish	*Serranus cabrilla*	shallow coastal	30	spring	Schunter et al., 2011	microsat	13	987	500 ns	1 ns
Fish	*Solea solea*	shallow coastal	30	winter	Bahri-Sfar et al., 2011	zymes	10	698	60 **	20 **
Fish	*Solea solea*	shallow coastal	30	winter	Garoia et al., 2007	microsat	4	1047	500 ns	5 ns
Fish	*Sparus aurata*	shallow coastal	45	fall	Franchini et al., 2012	microsat	12	319	10 *	5 *
Fish	*Symphodus tinca*	shallow coastal	10	spring	Carreras et al., 2017	SNPs	6	242	5 ns	20 *
Mollusca	*Cerastoderma edule*	shallow coastal	10	all	Sromek et al., 2019	RADseq/GBS	7	796	10 *	20 *
Mollusca	*Chiton olivaceus*	shallow coastal	10	spring	Villamor et al., 2014	mtDNA	4	444	500 ns	80 ns
Mollusca	*Hexaplex trunculus*	shallow coastal	1	all	Villamor et al., 2014	mtDNA	4	398	500 ns	150 ns
Mollusca	*Hexaplex trunculus*	shallow coastal	1	all	Marzouk et al., 2017[5]	SNPs	15	973	60 ns	400 ns
Mollusca	*Hexaplex trunculus*	shallow coastal	1	all	Marzouk et al., 2017[6]	mtDNA	15	973	80 ns	200 ns
Mollusca	*Mytilus galloprovincialis*	shallow coastal	20	spring	Diz and Presa, 2008	microsat	8	295	1 *	1 *
Mollusca	*Ostrea edulis*	shallow coastal	10	summer	Launey et al., 2002	microsat	5	758	20 *	60 *
Mollusca	*Patella caerulea*	shallow coastal	20	fall	Villamor et al., 2014	mtDNA	6	393	80 *	80 *
Mollusca	*Patella rustica*	shallow coastal	20	fall	Sá-Pinto et al., 2012	zymes	6	1102	1 ns	1 ns
Mollusca	*Patella ulyssiponensis*	shallow coastal	20	summer	Sá-Pinto et al., 2012	zymes	5	1104	150 ns	150 ns
Mollusca	*Phorcus turbinatus*	shallow coastal	10	all	Villamor et al., 2014	mtDNA	4	444	40 ns	500 *
Mollusca	*Ruditapes decussatus*	shallow coastal	10	summer	Gharbi et al., 2011	zymes	11	151	10 *	10 *
Mollusca	*Spondylus spinosus*	shallow coastal	20	all	Shabtay et al., 2014	mtDNA	5	38	5 ns	500 ns
Phanerogam	*Cymodocea nodosa*	shallow coastal	30	all	Alberto et al., 2008	microsat	13	840	80 *	10 **
Phanerogam	*Posidonia oceanica*	shallow coastal	30	all	Arnaud-Haond et al., 2007	microsat	29	859	60 ***	40 **

Bibliography Table SI-1:

Alberto, F., Massa, S., Manent, P., Diaz-Almela, E., Arnaud-Haond, S., Duarte, C.M., Serrão, E.A., 2008. Genetic differentiation and secondary contact zone in the seagrass Cymodocea nodosa across the Mediterranean–Atlantic transition region. J. Biogeogr. 35, 1279–1294. https://doi.org/10.1111/j.1365-2699.2007.01876.x

Arculeo, M., Pellerito, R., Bonhomme, F., 2010. Isolation and use of microsatellite loci in *Melicertus kerathurus* (Crustacea, Penaeidae). Aquat. Living Resour. 23, 103–107. https://doi.org/10.1051/alr/2010008

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van De Vliet, M.S., Alberto, F., Procaccini, G., Duarte, C.M., Serrão, E.A., 2007. Vicariance patterns in the Mediterranean Sea: east–west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. J. Biogeogr. 34, 963–976. https://doi.org/10.1111/j.1365-2699.2006.01671.x

Aurelle, D., Ledoux, J.-B., Rocher, C., Borsa, P., Chenuil, A., Féral, J.-P., 2011. Phylogeography of the red coral (Corallium rubrum): inferences on the evolutionary history of a temperate gorgonian. Genetica 139, 855–869. https://doi.org/10.1007/s10709-011-9589-6

Bahri-Sfar, L., Kaouèche, M., Haffani, M., Ouanes, K., Ben Hassine, O.K., 2011. Genetic population structure of the common sole, *Solea solea* Linnaeus, 1758 (Pisces, Pleuronectiformes) along the southern shores of the Mediterranean Sea (Tunisian coasts). Ital. J. Zool. 78, 157–167. https://doi.org/10.1080/11250003.2010.532513

Borrero-Pérez, G.H., González-Wangüemert, M., Marcos, C., Pérez-Ruzafa, A., 2011. Phylogeography of the Atlanto-Mediterranean sea cucumber Holothuria (Holothuria) mammata: the combined effects of historical processes and current oceanographical pattern: PHYLOGEOGRAPHY OF HOLOTHURIA MAMMATA. Mol. Ecol. 20, 1964–1975. https://doi.org/10.1111/j.1365-294X.2011.05068.x

Boscari, E., Abbiati, M., Badalamenti, F., Bavestrello, G., Benedetti-Cecchi, L., Cannas, R., Cau, A., Cerrano, C., Chimienti, G., Costantini, F., Fraschetti, S., Ingrosso, G., Marino, I.A.M., Mastrototaro, F., Papetti, C., Paterno, M., Ponti, M., Zane, L., Congiu, L., 2019. A population genomics insight by 2b-RAD reveals populations' uniqueness along the Italian coastline in *Leptopsammia pruvoti* (Scleractinia, Dendrophylliidae). Divers. Distrib. 25, 1101–1117. https://doi.org/10.1111/ddi.12918

Calò, A., Muñoz, I., Pérez-Ruzafa, Á., Vergara-Chen, C., García-Charton, J.A., 2016. Spatial genetic structure in the saddled sea bream (Oblada melanura [Linnaeus, 1758]) suggests multi-scaled patterns of connectivity between protected and unprotected areas in the Western Mediterranean Sea. Fish. Res. 176, 30–38. https://doi.org/10.1016/j.fishres.2015.12.001

Carreras, C., García-Cisneros, A., Wangensteen, O.S., Ordóñez, V., Palacín, C., Pascual, M., Turon, X., 2020. East is East and West is West: Population genomics and hierarchical analyses reveal genetic structure and adaptation footprints in the keystone species *Paracentrotus lividus* (Echinoidea). Divers. Distrib. 26, 382–398. https://doi.org/10.1111/ddi.13016

Carreras, C., Ordóñez, V., Zane, L., Kruschel, C., Nasto, I., Macpherson, E., Pascual, M., 2017. Population genomics of an endemic Mediterranean fish: differentiation by fine scale dispersal and adaptation. Sci. Rep. 7, 43417. https://doi.org/10.1038/srep43417

Casado-Amezúa, P., Goffredo, S., Templado, J., Machordom, A., 2012. Genetic assessment of population structure and connectivity in the threatened Mediterranean coral *Astroides calycularis* (Scleractinia, Dendrophylliidae) at different spatial scales: GENETIC STRUCTURE IN *ASTROIDES CALYCULARIS*. Mol. Ecol. 21, 3671–3685. https://doi.org/10.1111/j.1365-294X.2012.05655.x

Costantini, F., Carlesi, L., Abbiati, M., 2013. Quantifying Spatial Genetic Structuring in Mesophotic Populations of the Precious Coral Corallium rubrum. PLoS ONE 8, e61546. https://doi.org/10.1371/journal.pone.0061546

Dailianis, T., Tsigenopoulos, C.S., Dounas, C., Voultsiadou, E., 2011. Genetic diversity of the imperilled bath sponge Spongia officinalis Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation: GENETIC DIVERSITY OF SPONGIA OFFICINALIS. Mol. Ecol. 20, 3757–3772. https://doi.org/10.1111/j.1365-294X.2011.05222.x

Dalongeville, A., Benestan, L., Mouillot, D., Lobreaux, S., Manel, S., 2018. Combining six genome scan methods to detect candidate genes to salinity in the Mediterranean striped red mullet (Mullus surmuletus). BMC Genomics 19, 217. https://doi.org/10.1186/s12864-018-4579-z

Diz, A.P., Presa, P., 2008. Regional patterns of microsatellite variation in Mytilus galloprovincialis from the Iberian Peninsula. Mar. Biol. 154, 277–286. https://doi.org/10.1007/s00227-008-0921-3

Durand, J., Blel, H., Shen, K., Koutrakis, E., Guinand, B., 2013. Population genetic structure of Mugil cephalus in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers. Mar. Ecol. Prog. Ser. 474, 243–261. https://doi.org/10.3354/meps10080

Fassatoui, C., Mdelgi, E., Romdhane, M.S., 2009. A preliminary investigation of allozyme genetic variation and population structure in common pandora (Pagellus erythrinus, Sparidae) from Tunisian and Libyan coasts. Ichthyol. Res. 56, 301–307. https://doi.org/10.1007/s10228-008-0094-6

Franchini, P., Sola, L., Crosetti, D., Milana, V., Rossi, A.R., 2012. Low levels of population genetic structure in the gilthead sea bream, Sparus aurata, along the coast of Italy. ICES J. Mar. Sci. 69, 41–50. https://doi.org/10.1093/icesjms/fsr175

Fratini, S., Ragionieri, L., Cutuli, G., Vannini, M., Cannicci, S., 2013. Pattern of genetic isolation in the crab Pachygrapsus marmoratus within the Tuscan Archipelago (Mediterranean Sea). Mar. Ecol. Prog. Ser. 478, 173–183. https://doi.org/10.3354/meps10247

Fruciano, C., Tigano, C., Ferrito, V., 2011. Geographical and morphological variation within and between colour phases in *Coris julis* (L. 1758), a protogynous marine fish: GEOGRAPHICAL VARIATION IN A PROTOGYNOUS FISH. Biol. J. Linn. Soc. 104, 148–162. https://doi.org/10.1111/j.1095-8312.2011.01700.x

Galarza, J.A., Turner, G.F., Macpherson, E., Rico, C., 2009. Patterns of genetic differentiation between two co-occurring demersal species: the red mullet (Mullus barbatus) and the striped red mullet (Mullus surmuletus). Can. J. Fish. Aquat. Sci. 66, 1478–1490. https://doi.org/10.1139/F09-098

Garoia, F., Guarniero, I., Grifoni, D., Marzola, S., Tinti, F., 2007. Comparative analysis of AFLPs and SSRs efficiency in resolving population genetic structure of Mediterranean Solea vulgaris: EFFICIENCY OF SSRs AND AFLPs IN FISH POPULATION ANALYSIS. Mol. Ecol. 16, 1377–1387. https://doi.org/10.1111/j.1365-294X.2007.03247.x

Gharbi, A., Zitari-Chatti, R., Van Wormhoudt, A., Dhraief, M.N., Denis, F., Said, K., Chatti, N., 2011. Allozyme Variation and Population Genetic Structure in the Carpet Shell Clam Ruditapes decussatus Across the Siculo-Tunisian Strait. Biochem. Genet. 49, 788–805. https://doi.org/10.1007/s10528-011-9450-8

Gkafas, G., Tsigenopoulos, C., Magoulas, A., Panagiotaki, P., Vafidis, D., Mamuris, Z., Exadactylos, A., 2013. Population subdivision of saddled seabream Oblada melanura in the Aegean Sea revealed by genetic and morphometric analyses. Aquat. Biol. 18, 69–80. https://doi.org/10.3354/ab00490

González-Wangüemert, M., Cánovas, F., Pérez-Ruzafa, A., Marcos, C., Alexandrino, P., 2010. Connectivity patterns inferred from the genetic structure of white seabream (Diplodus sargus L.). J. Exp. Mar. Biol. Ecol. 383, 23–31. https://doi.org/10.1016/j.jembe.2009.10.010

Hammami, I., Bahri-Sfar, L., Kaoueche, M., Hassine, O.K.B., 2007. Genetic characterization of striped sea bream (Lithognathus mormyrus) populations on both sides of a boundary area between eastern and western Mediterranean basins 5.

Kaouèche, M., Bahri-Sfar, L., Hammami, I., Hassine, O.K.B., 2013. Morphological and genetic variations of Diplodus vulgaris along the Tunisian coasts 10.

Launey, S., 2002. Geographic Structure in the European Flat Oyster (Ostrea edulis L.) as Revealed by Microsatellite Polymorphism. J. Hered. 93, 331–351. https://doi.org/10.1093/jhered/93.5.331

Maggio, T., Lo Brutto, S., Garoia, F., Tinti, F., Arculeo, M., 2009. Microsatellite analysis of red mullet Mullus barbatus (Perciformes, Mullidae) reveals the isolation of the Adriatic Basin in the Mediterranean Sea. ICES J. Mar. Sci. 66, 1883–1891. https://doi.org/10.1093/icesjms/fsp160

Marzouk, Z., Aurelle, D., Said, K., Chenuil, A., 2017. Cryptic lineages and high population genetic structure in the exploited marine snail Hexaplex trunculus (Gastropoda: Muricidae). Biol. J. Linn. Soc. 122, 411–428. https://doi.org/10.1093/biolinnean/blx070

Masmoudi, M.B., Chaoui, L., Topçu, N.E., Hammami, P., Kara, M.H., Aurelle, D., 2016. Contrasted levels of genetic diversity in a benthic Mediterranean octocoral: Consequences of different demographic histories? Ecol. Evol. 6, 8665–8678. https://doi.org/10.1002/ece3.2490

Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G.R., Espiñeira, M., Fiorentino, F., Garofalo, G., Geffen, A.J., Hansen, Jakob.H., Helyar, S.J., Nielsen, E.E., Ogden, R., Patarnello, T., Stagioni, M., FishPopTrace Consortium, Tinti, F., Bargelloni, L., 2014. Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). Mol. Ecol. 23, 118–135. https://doi.org/10.1111/mec.12568

Muths, D., Rastorgueff, P.-A., Selva, M., Chevaldonné, P., 2015. Local scale connectivity in the cave-dwelling brooding fish Apogon imberbis. J. Sea Res. 95, 70–74. https://doi.org/10.1016/j.seares.2014.10.009

Ordóñez, V., Pascual, M., Rius, M., Turon, X., 2013. Mixed but not admixed: a spatial analysis of genetic variation of an invasive ascidian on natural and artificial substrates. Mar. Biol. 160, 1645–1660. https://doi.org/10.1007/s00227-013-2217-5

Palero, F., Abelló, P., Macpherson, E., Beaumont, M., Pascual, M., 2011. Effect of oceanographic barriers and overfishing on the population genetic structure of the European spiny lobster (*Palinurus elephas*): POPULATION GENETICS OF PALINURUS. Biol. J. Linn. Soc. 104, 407–418. https://doi.org/10.1111/j.1095-8312.2011.01728.x

Paterno, M., Schiavina, M., Aglieri, G., Souissi, J.B., Boscari, E., Casagrandi, R., Chassanite, A., Chiantore, M., Congiu, L., Guarnieri, G., Kruschel, C., Macic, V., Marino, I.A.M., Papetti, C., Patarnello, T., Zane, L., Melià, P., 2017. Population genomics meet Lagrangian simulations: Oceanographic patterns and long larval duration ensure connectivity among Paracentrotus lividus populations in the Adriatic and Ionian seas. Ecol. Evol. 7, 2463–2479. https://doi.org/10.1002/ece3.2844

Penant, G., Aurelle, D., Feral, J., Chenuil, A., 2013. Planktonic larvae do not ensure gene flow in the edible sea urchin Paracentrotus lividus. Mar. Ecol. Prog. Ser. 480, 155–170. https://doi.org/10.3354/meps10194

Pérez-Portela, R., Palacín, C., Duran, S., Turon, X., 2007. Biological traits of three closely related species of Pycnoclavella (Ascidiacea) in the Western Mediterranean. Mar. Biol. 152, 1031–1038. https://doi.org/10.1007/s00227-007-0750-9

Reem, E., Douek, J., Paz, G., Katzir, G., Rinkevich, B., 2017. Phylogenetics, biogeography and population genetics of the ascidian Botryllus schlosseri in the Mediterranean Sea and beyond. Mol. Phylogenet. Evol. 107, 221–231. https://doi.org/10.1016/j.ympev.2016.10.005

Sá-Pinto, A., Branco, M.S., Alexandrino, P.B., Fontaine, M.C., Baird, S.J.E., 2012. Barriers to Gene Flow in the Marine Environment: Insights from Two Common Intertidal Limpet Species of the Atlantic and Mediterranean. PLoS ONE 7, e50330. https://doi.org/10.1371/journal.pone.0050330

Schiavina, M., Marino, I. a. M., Zane, L., Melià, P., 2014. Matching oceanography and genetics at the basin scale. Seascape connectivity of the Mediterranean shore crab in the Adriatic Sea. Mol. Ecol. 23, 5496–5507. https://doi.org/10.1111/mec.12956

Schunter, C., Carreras-Carbonell, J., Macpherson, E., Tintoré, J., Vidal-Vijande, E., Pascual, A., Guidetti, P., Pascual, M., 2011. Matching genetics with oceanography: directional gene flow in a Mediterranean fish species. Mol. Ecol. 20, 5167–5181. https://doi.org/10.1111/j.1365-294X.2011.05355.x

Shabtay, A., Tikochinski, Y., Benayahu, Y., Rilov, G., 2014. Preliminary data on the genetic structure of a highly successful invading population of oyster suggesting its establishment dynamics in the Levant. Mar. Biol. Res. 10, 407–415. https://doi.org/10.1080/17451000.2013.814790

Sromek, L., Forcioli, D., Lasota, R., Furla, P., Wolowicz, M., 2019. Next-generation phylogeography of the cockle *Cerastoderma glaucum* : Highly heterogeneous genetic differentiation in a lagoon species. Ecol. Evol. 9, 4667–4682. https://doi.org/10.1002/ece3.5070

Susini, M.-L., Thibaut, T., Meinesz, A., Forcioli, D., 2007. A preliminary study of genetic diversity in Cystoseira amentacea (C. Agardh) Bory var. stricta Montagne (Fucales, Phaeophyceae) using random amplified polymorphic DNA. Phycologia 46, 605–611. https://doi.org/10.2216/06-100.1

Villamor, A., Costantini, F., Abbiati, M., 2014. Genetic Structuring across Marine Biogeographic Boundaries in Rocky Shore Invertebrates. PLOS ONE 9, e101135. https://doi.org/10.1371/journal.pone.0101135

Weber, A.A.-T., Mérigot, B., Valière, S., Chenuil, A., 2015. Influence of the larval phase on connectivity: strong differences in the genetic structure of brooders and broadcasters in the *Ophioderma longicauda* species complex. Mol. Ecol. 24, 6080–6094. https://doi.org/10.1111/mec.13456

Zitari-Chatti, R., Chatti, N., Elouaer, A., Said, K., 2007. Genetic variation and population structure of the caramote prawn Penaeus kerathurus (Forskäl) from the eastern and western Mediterranean coasts in Tunisia: Genetic variation in Penaeus kerathurus. Aquac. Res. 39, 70–76. https://doi.org/10.1111/j.1365-2109.2007.01874.x

Zulliger, D.E., Tanner, S., Ruch, M., Ribi, G., 2009. Genetic structure of the high dispersal Atlanto-Mediterreanean sea star Astropecten aranciacus revealed by mitochondrial DNA sequences and microsatellite loci. Mar. Biol. 156, 597–610. https://doi.org/10.1007/s00227-008-1111-z

Table SI-2 Literature reference used to configure species characteristics (i.e. habitat, PLD, spawning season). We also use FishBase (Froese and Pauly 2000, https://www.fishbase.se/search.php) and Doris (Willis et al., 2016, https://doris.ffessm.fr/) webpages for global information about the species of interest.

Species	References
Apogon imberbis	(Macpherson and Raventos, 2006; Raventos, 2007)
Astroides calycularis	(Casado-Amezúa et al., 2012; Goffredo et al., 2010)
Astropecten aranciacus	(Baeta et al., 2016; Zulliger et al., 2009)
Balanophyllia europaea	(Goffredo et al., 2004)
Botryllus schlosseri	(Reem et al., 2017)
Perforatus perforatus	(Villamor et al., 2014)
Carcinus aestuarii	(Carlton and Cohen, 2003; Schiavina et al., 2014)
Cerastoderma edule	(Boyden and Russell, 1972)
Chiton olivaceus	(Villamor et al., 2014; Wanninger and Haszprunar, 2002)
Chondrosia reniformis	(Lazoski et al., 2001; Villamor et al., 2014)
Cladocora caespitosa	(Casado-Amezúa et al., 2012; Kersting et al., 2013; Kružić et al., 2008)
Corallium rubrum	(Coelho and Lasker, 2016; Costantini et al., 2013; Teixidó et al., 2011)
Coris julis	(Fruciano et al., 2011; Macpherson and Raventos, 2006)
Crassostrea gigas	(Ernande et al., 2003)
Cymodocea nodosa	(Alberto et al., 2008; Orth et al., 2006)
Cystoseira amentacea	(Susini et al., 2007; Thibaut et al., 2016)
Diplodus puntazzo	(Di Franco and Guidetti, 2011)
Diplodus sargus	(Di Franco et al., 2013)
Diplodus vulgaris	(Di Franco et al., 2013; Macpherson and Raventos, 2006)
Eunicella cavolinii	(Cánovas-Molina et al., 2018; Masmoudi et al., 2016)
Fistularia commersonii	(Bernardi et al., 2016)
Epinephelus marginatus	(Macpherson and Raventos, 2006)
Halocynthia papillosa	(Villamor et al., 2014)
Hexaplex trunculus	(Vasconcelos et al., 2004)
Holothuria mammata	(Borrero-Pérez et al., 2011; Santos et al., n.d.)
Leptopsammia pruvoti	(Boscari et al., 2019; Goffredo et al., 2006)
Lithognathus mormyrus	Fishbase & Doris
Melicertus kerathurus	(Arculeo et al., 2010; Roberts et al., 2012; Zitari-Chatti et al., 2007)
Merluccius merluccius	(Hidalgo et al., 2019; Morales-Nin and Moranta, 2004)
Microcosmus squamiger	(Rius et al., 2010, 2009)
Mugil cephalus	(Kuo et al., 1973)
Mullus barbartus	(Félix-Hackradt et al., 2013; Maggio et al., 2009)
Mullus surmuletus	(Félix-Hackradt et al., 2013; Macpherson and Raventos, 2006)
Mytilus galloprovincialis	(Caceres-Martinez et al., 1993)
Oblada melanura	(Macpherson and Raventos, 2006)
Phorcus turbinatus	(Villamor et al., 2014)
Ophioderma longicauda	(Weber et al., 2014, 2015)
Ostrea edulis	(Bierne et al., 1998)
Pachygrapsus marmoratus	(Cuesta and Rodríguez, 2000)
Pagellus erythrinus	(Macpherson and Raventos, 2006)
Palinurus elephas	(Hunter, 1999)
Paracentrotus lividus	(Pedrotti, 1993)
Patella caerulea	(Villamor et al., 2014)

Patella rustica	(Sá-Pinto et al., 2012)
Patella ulyssiponensis	(Sá-Pinto et al., 2012)
Posidonia oceanica	(Melià et al., 2016; Serra et al., 2010)
Pycnoclavella communis	(Pérez-Portela et al., 2007)
Ruditapes decussatus	(Gharbi et al., 2011)
Scopalina lophyropoda	(Garoia et al., 2004)
Serranus cabrilla	(Macpherson and Raventos, 2006)
Solea solea	(Bahri-Sfar et al., 2011; Vaz et al., 2019)
Sparus aurata	(Franchini et al., 2012)
Spondylus spinosus	(Soria et al., 2010)
Spongia officinalis	(Baldacconi et al., 2007; Gaino et al., 2007)
Symphodus tinca	(Macpherson and Raventos, 2006; Pallaoro and Jardas, 2003)

Bibliography Table SI-2:

Alberto, F., Massa, S., Manent, P., Diaz-Almela, E., Arnaud-Haond, S., Duarte, C.M., Serrão, E.A., 2008. Genetic differentiation and secondary contact zone in the seagrass Cymodocea nodosa across the Mediterranean–Atlantic transition region. J. Biogeogr. 35, 1279–1294. https://doi.org/10.1111/j.1365-2699.2007.01876.x

Arculeo, M., Pellerito, R., Bonhomme, F., 2010. Isolation and use of microsatellite loci in *Melicertus kerathurus* (Crustacea, Penaeidae). Aquat. Living Resour. 23, 103–107. https://doi.org/10.1051/alr/2010008

Baeta, M., Galimany, E., Ramón, M., 2016. Growth and reproductive biology of the sea star Astropecten aranciacus (Echinodermata, Asteroidea) on the continental shelf of the Catalan Sea (northwestern Mediterranean). Helgol. Mar. Res. 70, 1. https://doi.org/10.1186/s10152-016-0453-z

Bahri-Sfar, L., Kaouèche, M., Haffani, M., Ouanes, K., Ben Hassine, O.K., 2011. Genetic population structure of the common sole, *Solea solea* Linnaeus, 1758 (Pisces, Pleuronectiformes) along the southern shores of the Mediterranean Sea (Tunisian coasts). Ital. J. Zool. 78, 157–167. https://doi.org/10.1080/11250003.2010.532513

Baldacconi, R., Nonnis-Marzano, C., Gaino, E., Corriero, G., 2007. Sexual reproduction, larval development and release in Spongia oYcinalis L. (Porifera, Demospongiae) from the Apulian coast. Mar Biol 11.

Bernardi, G., Azzurro, E., Golani, D., Miller, M.R., 2016. Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader. Mol. Ecol. 25, 3384–3396. https://doi.org/10.1111/mec.13682

Bierne, N., Launey, S., Naciri-Graven, Y., Bonhomme, F., 1998. Early Effect of Inbreeding as Revealed by Microsatellite Analyses on Ostrea edulis Larvae. Genetics 148, 1893–1906.

Borrero-Pérez, G.H., González-Wangüemert, M., Marcos, C., Pérez-Ruzafa, A., 2011. Phylogeography of the Atlanto-Mediterranean sea cucumber Holothuria (Holothuria) mammata: the combined effects of historical processes and current oceanographical pattern: PHYLOGEOGRAPHY OF HOLOTHURIA MAMMATA. Mol. Ecol. 20, 1964–1975. https://doi.org/10.1111/j.1365-294X.2011.05068.x

Boscari, E., Abbiati, M., Badalamenti, F., Bavestrello, G., Benedetti-Cecchi, L., Cannas, R., Cau, A., Cerrano, C., Chimienti, G., Costantini, F., Fraschetti, S., Ingrosso, G., Marino, I.A.M., Mastrototaro, F., Papetti, C., Paterno, M., Ponti, M., Zane, L., Congiu, L., 2019. A population genomics insight by 2b-RAD reveals populations' uniqueness along the Italian coastline in *Leptopsammia pruvoti* (Scleractinia, Dendrophylliidae). Divers. Distrib. 25, 1101–1117. https://doi.org/10.1111/ddi.12918

Boyden, C.R., Russell, P.J.C., 1972. The Distribution and Habitat Range of the Brackish Water Cockle (Cardium (Cerastoderma) glaucum) in the British Isles. J. Anim. Ecol. 41, 719. https://doi.org/10.2307/3205

Caceres-Martinez, J., Robledo, J., Figueras, A., 1993. Settlement of mussels Mytilus galloprovincialis on an exposed rocky shore in Ria de Vigo, NW Spain. Mar. Ecol. Prog. Ser. 93, 195–198. https://doi.org/10.3354/meps093195

Cánovas-Molina, A., Montefalcone, M., Bavestrello, G., Masmoudi, M.B., Haguenauer, A., Hammami, P., Chaoui, L., Kara, M.H., Aurelle, D., 2018. From depth to regional spatial genetic differentiation of Eunicella cavolini in the NW Mediterranean. C. R. Biol. 341, 421–432. https://doi.org/10.1016/j.crvi.2018.09.002

Carlton, J.T., Cohen, A.N., 2003. Episodic global dispersal in shallow water marine organisms: the case history of the European shore crabs *Carcinus maenas* and *C. aestuarii*: Episodic global invasion patterns in shore crabs. J. Biogeogr. 30, 1809–1820. https://doi.org/10.1111/j.1365-2699.2003.00962.x

Casado-Amezúa, P., Goffredo, S., Templado, J., Machordom, A., 2012. Genetic assessment of population structure and connectivity in the threatened Mediterranean coral *Astroides calycularis* (Scleractinia, Dendrophylliidae) at different spatial scales: GENETIC STRUCTURE IN *ASTROIDES CALYCULARIS*. Mol. Ecol. 21, 3671–3685. https://doi.org/10.1111/j.1365-294X.2012.05655.x

Coelho, M.A.G., Lasker, H.R., 2016. Larval Dispersal and Population Connectivity in Anthozoans, in: Goffredo, S., Dubinsky, Z. (Eds.), The Cnidaria, Past, Present and Future: The World of Medusa and Her Sisters. Springer International Publishing, Cham, pp. 291–315. https://doi.org/10.1007/978-3-319-31305-4_19

Costantini, F., Carlesi, L., Abbiati, M., 2013. Quantifying Spatial Genetic Structuring in Mesophotic Populations of the Precious Coral Corallium rubrum. PLoS ONE 8, e61546. https://doi.org/10.1371/journal.pone.0061546

Cuesta, J.A., Rodríguez, A., 2000. Zoeal stages of the intertidal crab Pachygrapsus marmoratus (Fabricius, 1787) (Brachyura, Grapsidae) reared in the laboratory. Hydrobiologia 436, 119–130. https://doi.org/10.1023/A:1026576614590

Di Franco, A., Guidetti, P., 2011. Patterns of variability in early-life traits of fishes depend on spatial scale of analysis. Biol. Lett. 7, 454–456. https://doi.org/10.1098/rsbl.2010.1149

Di Franco, A., Qian, K., Calò, A., Di Lorenzo, M., Planes, S., Guidetti, P., 2013. Patterns of variability in early life traits of a Mediterranean coastal fish. Mar. Ecol. Prog. Ser. 476, 227–235.

Ernande, B., Clobert, J., McCombie, H., Boudry, P., 2003. Genetic polymorphism and trade-offs in the early life-history strategy of the Pacific oyster, Crassostrea gigas (Thunberg, 1795): a quantitative genetic study: Early life-history strategy in a marine bivalve. J. Evol. Biol. 16, 399–414. https://doi.org/10.1046/j.1420-9101.2003.00543.x

Félix-Hackradt, F.C., Hackradt, C.W., Pérez-Ruzafa, Á., García-Charton, J.A., 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. https://doi.org/10.1016/j.marenvres.2013.08.008

Franchini, P., Sola, L., Crosetti, D., Milana, V., Rossi, A.R., 2012. Low levels of population genetic structure in the gilthead sea bream, Sparus aurata, along the coast of Italy. ICES J. Mar. Sci. 69, 41–50. https://doi.org/10.1093/icesjms/fsr175

Fruciano, C., Tigano, C., Ferrito, V., 2011. Geographical and morphological variation within and between colour phases in *Coris julis* (L. 1758), a protogynous marine fish: GEOGRAPHICAL VARIATION IN A PROTOGYNOUS FISH. Biol. J. Linn. Soc. 104, 148–162. https://doi.org/10.1111/j.1095-8312.2011.01700.x

Gaino, E., Baldacconi, R., Corriero, G., 2007. Post-larval development of the commercial sponge Spongia officinalis L. (Porifera, Demospongiae). Tissue Cell 39, 325–334. https://doi.org/10.1016/j.tice.2007.06.006

Garoia, F., Guarniero, I., Ramšak, A., Ungaro, N., Landi, M., Piccinetti, C., Mannini, P., Tinti, F., 2004. Microsatellite DNA variation reveals high gene flow and panmictic populations in the Adriatic shared stocks of the European squid and cuttlefish (Cephalopoda). Heredity 93, 166–174. https://doi.org/10.1038/sj.hdy.6800489

Gharbi, A., Zitari-Chatti, R., Van Wormhoudt, A., Dhraief, M.N., Denis, F., Said, K., Chatti, N., 2011. Allozyme Variation and Population Genetic Structure in the Carpet Shell Clam Ruditapes decussatus Across the Siculo-Tunisian Strait. Biochem. Genet. 49, 788–805. https://doi.org/10.1007/s10528-011-9450-8

Goffredo, S., Airi, V., Radetić, J., Zaccanti, F., 2006. Sexual reproduction of the solitary sunset cup coral Leptopsammia pruvoti (Scleractinia, Dendrophylliidae) in the Mediterranean. 2. Quantitative aspects of the annual reproductive cycle. Mar. Biol. 148, 923–931. https://doi.org/10.1007/s00227-005-0137-8

Goffredo, S., Gasparini, G., Marconi, G., Putignano, M.T., Pazzini, C., Zaccanti, F., 2010. Gonochorism and planula brooding in the Mediterranean endemic orange coral Astroides calycularis (Scleractinia: Dendrophylliidae). Morphological aspects of gametogenesis and ontogenesis. Mar. Biol. Res. 6, 421–436. https://doi.org/10.1080/17451000903428488

Goffredo, S., Mezzomonaco, L., Zaccanti, F., 2004. Genetic differentiation among populations of the Mediterranean hermaphroditic brooding coral Balanophyllia europaea (Scleractinia: Dendrophylliidae). Mar. Biol. 145, 1075–1083. https://doi.org/10.1007/s00227-004-1403-x

Hidalgo, M., Rossi, V., Monroy, P., Ser-Giacomi, E., Hernández-García, E., Guijarro, B., Massutí, E., Alemany, F., Jadaud, A., Perez, J.L., Reglero, P., 2019. Accounting for ocean connectivity and hydroclimate in fish recruitment fluctuations within transboundary metapopulations. Ecol. Appl. 0, e01913. https://doi.org/10.1002/eap.1913

Hunter, 1999. BIOLOGY OF THE EUROPEAN SPINY LOBSTER, PALINURUS ELEPHAS (FABRICIUS, 1787) (DECAPODA, PALINURIDEA). Crustaceana 72, 545–565. https://doi.org/10.1163/156854099503609

Kersting, D., Casado, C., López-Legentil, S., Linares, C., 2013. Unexpected patterns in the sexual reproduction of the Mediterranean scleractinian coral Cladocora caespitosa. Mar. Ecol. Prog. Ser. 486, 165–171. https://doi.org/10.3354/meps10356

Kružić, P., Žuljević, A., Nikolić, V., 2008. Spawning of the colonial coral Cladocora caespitosa (Anthozoa, Scleractinia) in the Southern Adriatic Sea. Coral Reefs 27, 337–341. https://doi.org/10.1007/s00338-007-0334-7

Kuo, C.-M., Shehadeh, Z.H., Milken, K.K., 1973. A preliminary report on the development, growth and survival of laboratory reared larvae of the grey mullet, Mugil cephalus L. J. Fish Biol. 5, 459–470. https://doi.org/10.1111/j.1095-8649.1973.tb04475.x

Lazoski, C., Solé-Cava, A., Boury-Esnault, N., Klautau, M., Russo, C., 2001. Cryptic speciation in a high gene flow scenario in the oviparous marine sponge Chondrosia reniformis. Mar. Biol. 139, 421–429. https://doi.org/10.1007/s002270100542

Macpherson, E., Raventos, N., 2006. Relationship between pelagic larval duration and geographic distribution of Mediterranean littoral fishes. Mar. Ecol. Prog. Ser. 327, 257–265.

Maggio, T., Lo Brutto, S., Garoia, F., Tinti, F., Arculeo, M., 2009. Microsatellite analysis of red mullet Mullus barbatus (Perciformes, Mullidae) reveals the isolation of the Adriatic Basin in the Mediterranean Sea. ICES J. Mar. Sci. 66, 1883–1891. https://doi.org/10.1093/icesjms/fsp160

Masmoudi, M.B., Chaoui, L., Topçu, N.E., Hammami, P., Kara, M.H., Aurelle, D., 2016. Contrasted levels of genetic diversity in a benthic Mediterranean octocoral: Consequences of different demographic histories? Ecol. Evol. 6, 8665–8678. https://doi.org/10.1002/ece3.2490

Melià, P., Schiavina, M., Rossetto, M., Gatto, M., Fraschetti, S., Casagrandi, R., 2016. Looking for hotspots of marine metacommunity connectivity: a methodological framework. Sci. Rep. 6, 23705.

Morales-Nin, B., Moranta, J., 2004. Recruitment and post-settlement growth of juvenile *Merluccius merluccius* on the western Mediterranean shelf. Sci. Mar. 68, 399–409. https://doi.org/10.3989/scimar.2004.68n3399

Orth, R.J., Harwell, M.C., Inglis, G.J., 2006. Ecology of Seagrass Seeds and Seagrass Dispersal Processes, in: LARKUM, A.W.D., ORTH, R.J., DUARTE, C.M. (Eds.), SEAGRASSES: BIOLOGY, ECOLOGYAND CONSERVATION. Springer Netherlands, Dordrecht, pp. 111–133. https://doi.org/10.1007/978-1-4020-2983-7_5

Pallaoro, A., Jardas, I., 2003. Some biological parameters of the peacock wrasse, *Symphodus (Crenilabrus) tinca* (L. 1758) (Pisces: Labridae) from the middle eastern Adriatic (Croatian coast). Sci. Mar. 67, 33–41. https://doi.org/10.3989/scimar.2003.67n133

Pedrotti, M.L., 1993. Spatial and temporal distribution and recruitment of echinoderm larvae in the Ligurian Sea. J. Mar. Biol. Assoc. U. K. 73, 513–530. https://doi.org/10.1017/S0025315400033075

Pérez-Portela, R., Palacín, C., Duran, S., Turon, X., 2007. Biological traits of three closely related species of Pycnoclavella (Ascidiacea) in the Western Mediterranean. Mar. Biol. 152, 1031–1038. https://doi.org/10.1007/s00227-007-0750-9

Raventos, N., 2007. Age, growth and reproductive parameters of the Mediterranean cardinal fish, Apogon imberbis. J. Appl. Ichthyol. 23, 675–678. https://doi.org/10.1111/j.1439-0426.2007.00847.x

Reem, E., Douek, J., Paz, G., Katzir, G., Rinkevich, B., 2017. Phylogenetics, biogeography and population genetics of the ascidian Botryllus schlosseri in the Mediterranean Sea and beyond. Mol. Phylogenet. Evol. 107, 221–231. https://doi.org/10.1016/j.ympev.2016.10.005

Rius, M., Pineda, M.C., Turon, X., 2009. Population dynamics and life cycle of the introduced ascidian Microcosmus squamiger in the Mediterranean Sea. Biol. Invasions 11, 2181–2194. https://doi.org/10.1007/s10530-008-9375-2

Rius, M., Turon, X., Dias, G.M., Marshall, D.J., 2010. Propagule size effects across multiple life-history stages in a marine invertebrate: *Propagule size across life-history stages*. Funct. Ecol. 24, 685–693. https://doi.org/10.1111/j.1365-2435.2009.01668.x

Roberts, S.D., Dixon, C.D., Andreacchio, L., 2012. Temperature dependent larval duration and survival of the western king prawn, Penaeus (Melicertus) latisulcatus Kishinouye, from Spencer Gulf, South Australia. J. Exp. Mar. Biol. Ecol. 411, 14–22. https://doi.org/10.1016/j.jembe.2011.10.022

Santos, R., Dias, S., Tecelão, C., Pedrosa, R., Pombo, A., n.d. Reproductive biological characteristics and fatty acid profile of Holothuria mammata (Grube, 1840) 8.

Sá-Pinto, A., Branco, M.S., Alexandrino, P.B., Fontaine, M.C., Baird, S.J.E., 2012. Barriers to Gene Flow in the Marine Environment: Insights from Two Common Intertidal Limpet Species of the Atlantic and Mediterranean. PLoS ONE 7, e50330. https://doi.org/10.1371/journal.pone.0050330

Schiavina, M., Marino, I. a. M., Zane, L., Melià, P., 2014. Matching oceanography and genetics at the basin scale. Seascape connectivity of the Mediterranean shore crab in the Adriatic Sea. Mol. Ecol. 23, 5496–5507. https://doi.org/10.1111/mec.12956

Serra, I.A., Innocenti, A.M., Maida, G.D., Calvo, S., Migliaccio, M., Zambianchi, E., Pizzigalli, C., Arnaud-Haond, S., Duarte, C.M., Serrao, E.A., Procaccini, G., 2010. Genetic structure in the Mediterranean seagrass Posidonia oceanica: disentangling past vicariance events from contemporary patterns of gene flow. Mol. Ecol. 19, 557–568. https://doi.org/10.1111/j.1365-294X.2009.04462.x

Soria, G., Tordecillas-Guillen, J., Cudney-Bueno, R., Shaw, W., 2010. Spawning Induction, Fecundity Estimation, and Larval Culture of *Spondylus calcifer* (Carpenter, 1857) (Bivalvia: Spondylidae). J. Shellfish Res. 29, 143–149. https://doi.org/10.2983/035.029.0108

Susini, M.-L., Thibaut, T., Meinesz, A., Forcioli, D., 2007. A preliminary study of genetic diversity in Cystoseira amentacea (C. Agardh) Bory var. stricta Montagne (Fucales, Phaeophyceae) using random amplified polymorphic DNA. Phycologia 46, 605–611. https://doi.org/10.2216/06-100.1

Teixidó, N., Garrabou, J., Harmelin, J.-G., 2011. Low Dynamics, High Longevity and Persistence of Sessile Structural Species Dwelling on Mediterranean Coralligenous Outcrops. PLoS ONE 6, e23744. https://doi.org/10.1371/journal.pone.0023744

Thibaut, T., Bottin, L., Aurelle, D., Boudouresque, C.-F., Blanfuné, A., Verlaque, M., Pairaud, I., Millet, B., 2016. Connectivity of Populations of the Seaweed Cystoseira amentacea within the Bay of Marseille (Mediterranean Sea): Genetic Structure and Hydrodynamic Connections. Cryptogam. Algol. 37, 233–255. https://doi.org/10.7872/crya/v37.iss4.2016.233

Vasconcelos, P., Gaspar, M.B., Joaquim, S., Matias, D., Castro, M., 2004. Spawning of *Hexaplex (Trunculariopsis) trunculus* (Gastropoda: Muricidae) in the laboratory: description of spawning behaviour, egg masses, embryonic development, hatchling and juvenile growth rates. Invertebr. Reprod. Dev. 46, 125–138. https://doi.org/10.1080/07924259.2004.9652616

Vaz, A.C., Scarcella, G., Pardal, M.A., Martinho, F., 2019. Water temperature gradients drive early life-history patterns of the common sole (Solea solea L.) in the Northeast Atlantic and Mediterranean. Aquat. Ecol. 53, 281–294. https://doi.org/10.1007/s10452-019-09688-2

Villamor, A., Costantini, F., Abbiati, M., 2014. Genetic Structuring across Marine Biogeographic Boundaries in Rocky Shore Invertebrates. PLOS ONE 9, e101135. https://doi.org/10.1371/journal.pone.0101135

Wanninger, A., Haszprunar, G., 2002. Chiton myogenesis: Perspectives for the development and evolution of larval and adult muscle systems in molluscs. J. Morphol. 251, 103–113. https://doi.org/10.1002/jmor.1077

Weber, A.A.-T., Stöhr, S., Chenuil, A., 2014. Genetic data, reproduction season and reproductive strategy data support the existence of biological species in Ophioderma longicauda. C. R. Biol. 337, 553–560. https://doi.org/10.1016/j.crvi.2014.07.007

Weber, A. a.-T., Mérigot, B., Valière, S., Chenuil, A., 2015. Influence of the larval phase on connectivity: strong differences in the genetic structure of brooders and broadcasters in the Ophioderma longicauda species complex. Mol. Ecol. 24, 6080–6094. https://doi.org/10.1111/mec.13456

Zitari-Chatti, R., Chatti, N., Elouaer, A., Said, K., 2007. Genetic variation and population structure of the caramote prawn Penaeus kerathurus (Forskäl) from the eastern and western Mediterranean coasts in Tunisia: Genetic variation in Penaeus kerathurus. Aquac. Res. 39, 70–76. https://doi.org/10.1111/j.1365-2109.2007.01874.x

Zulliger, D.E., Tanner, S., Ruch, M., Ribi, G., 2009. Genetic structure of the high dispersal Atlanto-Mediterreanean sea star Astropecten aranciacus revealed by mitochondrial DNA sequences and microsatellite loci. Mar. Biol. 156, 597–610. https://doi.org/10.1007/s00227-008-1111-z

II. Shallow coastal and neritic shelf habitat description

Species-specific habitat

Based on the compiled literature of Table SI-2, we attribute to each species compiled in our meta-analysis (Table SI-1) one habitat among two broad types: shallow coastal habitat or neritic shelf habitat. If a species could fall into both habitat types, we select only one by (i) for fishes, retaining the habitat that contains the depth where the post-settlers/juveniles inhabit; (ii) for other organisms, retaining the habitat that comprises the depth of the sampling.

Network habitat filter

i. Bathymetric filter

Gridded bathymetry data provided by ETOPO1 1 Arc-Minute Global Relief Model (doi:10.7289/V5C8276M) are co-located with on our network grid to select all the topographic values encompassed in each node. Nodes whose shallowest depth is comprised between 0 m and 50 m are kept in the shallow coastal bathymetric filter and those whose shallowest depth is comprised between 50 m and 200 m are filtered in our neritic shelf bathymetric filter; the remaining deeper nodes are excluded.

ii. Substrate filter

We use EMODnet Seabed Habitats data[1] to define our substrate filter combining both the EUNIS classification and the MSFD Benthic Broad Habitat typology. We distinguish *infralittoral* and *circalittoral* substrate categories, while considering all substrates together including rocks, fine muds, coarse sediments, etc. (Table SI-3). By co-locating substrate data on our network grid, nodes which encompass *infralittoral* substrate are filtered in our shallow coastal substrate filter and nodes which encompass *circalittoral* substrate are filtered in our neritic shelf substrate filter.

*Table SI-3: **Substrate typology used to define the two composite substrates**, called shallow coastal and neritic shelf substrates, following EUNIS and MFSD classifications referenced in EMODnet Seabed Habitats data.*

Classifications	Shallow coastal substrate	Neritic shelf substrate
EUNIS	-'A3: Infralittoral rock and other hard substrata' -'A4.26 or A4.32: Mediterranean coralligenous communities moderately exposed to hydrodynamic action or Mediterranean coralligenous communities sheltered from hydrodynamic action' -'A5.13: Infralittoral coarse sediment' -'A5.23 or A5.33 or A5.34: Infralittoral fine sands or Infralittoral sandy mud or Infralittoral fine mud' -'A5.23: Infralittoral fine sands' -'A5.33: Infralittoral sandy mud'	-'A4.26 or A4.32: Mediterranean coralligenous communities moderately exposed to hydrodynamic action or Mediterranean coralligenous communities sheltered from hydrodynamic action' -'A4.27: Faunal communities on deep moderate energy circalittoral rock' -'A4: Circalittoral rock and other hard substrata -'A5.14: Circalittoral coarse sediment' -'A5.25: Circalittoral fine sand' -'A5.26: Circalittoral muddy sand' -'A5.35: Circalittoral sandy mud'

[1] *Information contained here has been derived from data that is made available under the European Marine Observation Data Network (EMODnet) Seabed Habitats initiative (http://www.emodnet-seabedhabitats.eu/), financed by the European Union under Regulation (EU) No 508/2014 of the European Parliament and of the Council of 15 May 2014 on the European Maritime and Fisheries Fund.*

	-'A5.34: Infralittoral fine mud' -'A5.5353: Facies of dead "mattes" of [Posidonia oceanica]' -'A5.535: [Posidonia] beds'	-'A5.36: Circalittoral fine mud'
MFSD	-'Infralittoral coarse sediment' -'Infralittoral mixed sediment' -'Infralittoral mud' -'Infralittoral rock and biogenic reef' -'Infralittoral sand'	-'Circalittoral coarse sediment' -'Circalittoral mixed sediment' -'Circalittoral mud' -'Circalittoral mud or Offshore circalittoral mud' -'Circalittoral rock and biogenic reef' -'Circalittoral sand' -'Offshore circalittoral coarse sediment' -'Offshore circalittoral mixed sediment' -'Offshore circalittoral mud' -'Offshore circalittoral rock and biogenic reef' -'Offshore circalittoral sand'

iii. Habitat filter

The final habitat filter consists in the superposition of both bathymetric and substrate filters: coloured nodes characterize the shallow coastal habitat (Fig. SI-1 and Fig. 1b main manuscript) and the neritic shelf habitat (Fig. SI-2, Fig. 1c main manuscript). All nodes selected in the bathymetric filter are coloured in transparent light grey, those belonging to the substrate filter are displayed in red, while pink nodes represent the superposition of both filters.

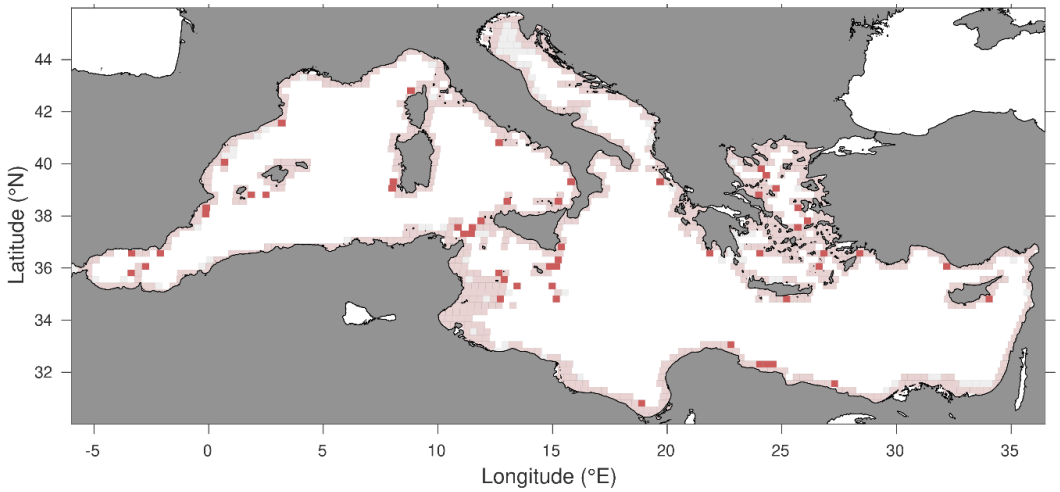


*Figure SI-1: **Shallow coastal habitat** represented by a bathymetric filter (light grey nodes) and a substrate filter (red nodes). Pink nodes characterize the superposition of both bathymetric and substrate filters.*

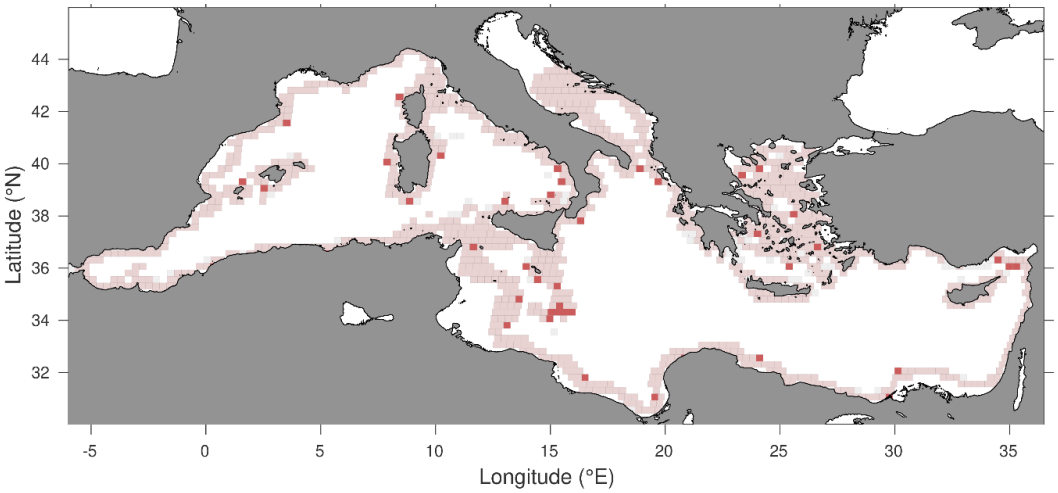


Figure SI-2: ***Neritic shelf habitat*** *represented by a bathymetric filter (light grey nodes) and a substrate filter (red nodes). Pink nodes characterize the superposition of both bathymetric and substrate filters.*

III. Bio-physical modelling

Hydrodynamical model

We use the hydrodynamic model Mediterranean Forecasting System (MFS) based on NEMO-OPA (Nucleus for European Modelling of the Ocean-PArellelis, version 3.2; Madec, 2008). This data-assimilative operational model is implemented in the Mediterranean at 1/16° degree horizontal regular resolution and 72 unevenly spaced verticals levels (Oddo et al., 2009). We use the physics reanalysis products for years spanning 2000-2010 downloaded from Marine Copernicus website (https://marine.copernicus.eu/).

Lagrangian modelling

Following the procedure described in Ser-Giacomi et al., (2015), we discretise the Mediterranean basin on a network of 8196 node of 1/4° degree horizontal resolution. About 100 numerical particles are evenly initialised in each node. Horizontal trajectories are simulated by integrating the velocity field, bilinearly interpolated using a *Runge-Kutta* 4 algorithm with a time step of 0.3 h, fulfilling the Courant-Friedrichs-Lewy condition (CFL, Courant et al., (1928)) over the period of interest (Fig. SI-3). We use the MFS horizontal velocity fields of the 3rd and 17th vertical levels, which correspond to about 12 m for the shallow coastal habitat and 102 m for the neritic shelf habitat, respectively. Numerical propagules are tracked over five different drifting times (simulating different Pelagic Larval Durations): 1, 10, 20, 30 and 45 days considering successive starting times with a 10-day periodicity over years 2000 to 2010. Altogether, it represents 402 Lagrangian experiments per PLD and per vertical layer, that is 4020 numerical experiments in total (Table SI-4).

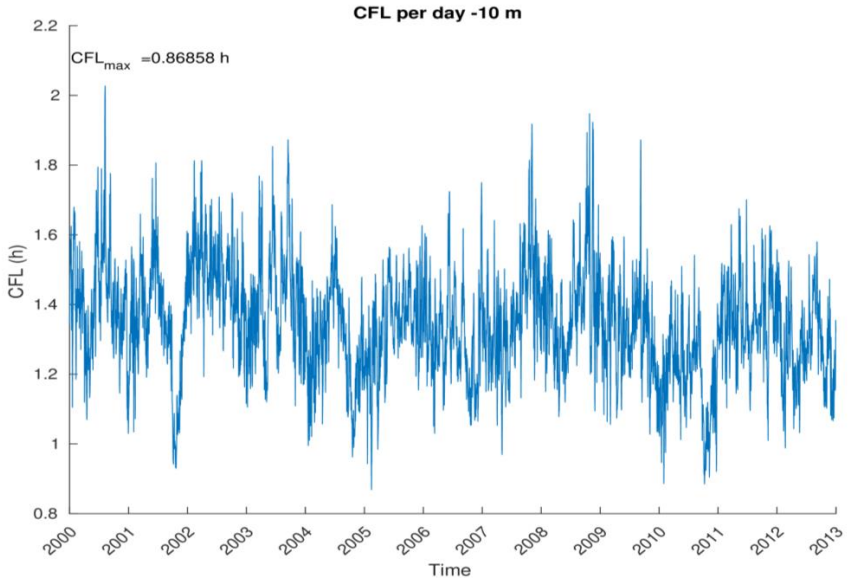


Figure SI-3: ***CFL condition indexed for each MFS daily velocity field at 12 m deep between 2000 and 2010****. The maximum time-step value to fulfil CFL condition is 0.87 h, we use a time-step of 0.3 h in our Lagrangian experiments.*

Table SI-4: ***Summary of the numerical experiments performed*** *using a 10-day periodicity over 2000-2010; each number refer to the total number of connectivity matrices that are averaged for each PLD (spanning 1, 10, 20, 30 and 45 days) for each vertical layer (12 m and 102 m).*

Season	Winter	Spring	Summer	Fall	All year
Start date	01/01	01/04	01/07	01/10	01/01
End date	31/03	30/06	30/09	31/12	31/12
Nbr of Lagrangian experiment	100	99	100	103	402

Bibliography SI-III

Courant, R., Friedrichs, K., Lewy, H., 1928. Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. 100, 32–74. https://doi.org/10.1007/BF01448839

Madec, G.: NEMO ocean engine, Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27 ISSN No1288-1619, 2008.

Oddo, P., Adani, M., Pinardi, N., Fratianni, C., Tonani, M., Pettenuzzo, D., others, 2009. A nested Atlantic-Mediterranean Sea general circulation model for operational forecasting. Ocean Sci.

Ser-Giacomi, E., Rossi, V., López, C., Hernandez-Garcia, E., 2015. Flow networks: A characterization of geophysical fluid transport. Chaos Interdiscip. J. Nonlinear Sci. 25, 036404.

IV. Parametrization of our multi-generation dispersal model

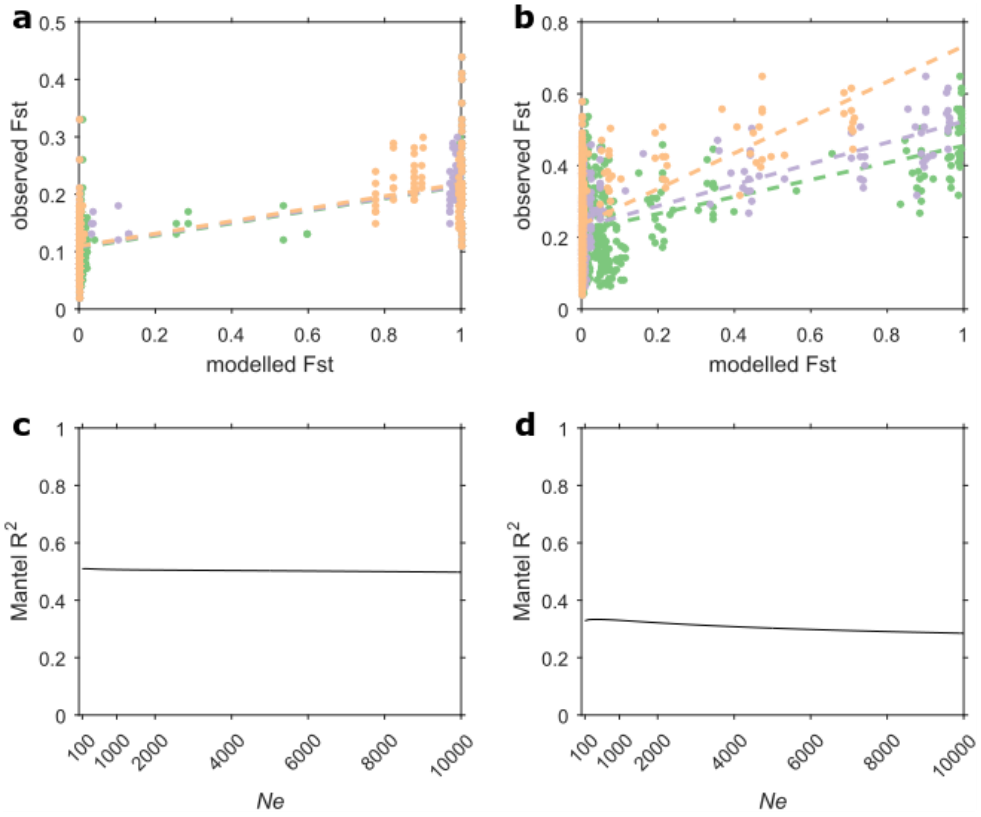


Figure SI-4. ***Sensitivity tests on the values of Ne. a*** *Correlation between* modelled Fst *and* observed Fst *with Ne = 100 (green dots and linear regression curve), Ne = 1000 (purple dots and linear regression curve) and Ne = 10^4 (orange dots and linear regression curve) for* Corallium rubrum*, Aurelle et al., 2011.* ***b****, same as a for* Posidonia oceanica*, Arnaud-Haond et al., 2007.* ***c****, Mantel R^2 between* modelled Fst *and* observed Fst *with Ne spanning 100 to 10^4 for* Corallium rubrum*, Aurelle et al., 2011.* ***d****, same as* ***c*** *for* Posidonia oceanica*, Arnaud-Huond et al., 2007.*

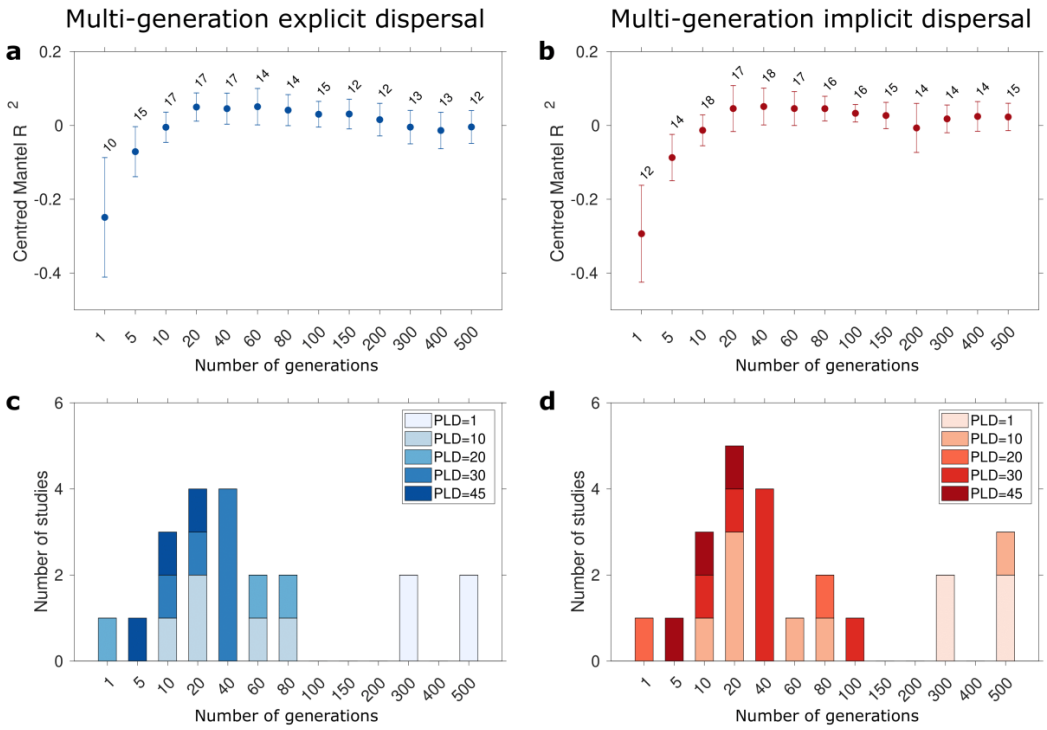


Fig SI-5: ***Correlations between our multi-generation dispersal models and observed genetic structures. a*** *mean centred Mantel R^2 with 95 % IC and the number of significant Mantel correlation (p-value ≤ 0.05) between* modelled Fst *calculated with the multi-generation explicit dispersal model and* observed Fst *per number of generations.* ***b*** *same as* ***a*** *but with our multi-generation implicit dispersal model.* ***c*** *Histograms of the optimal number of generations for all significant Mantel correlations between explicit* modelled Fst *and* observed Fst*; the blueish colours indicate the PLD of the studied species.* ***d*** *same as* ***c*** *but with implicit modelled Fst.*

V. Correlation between isolation-by-distance models or dispersal models and observed genetic structure

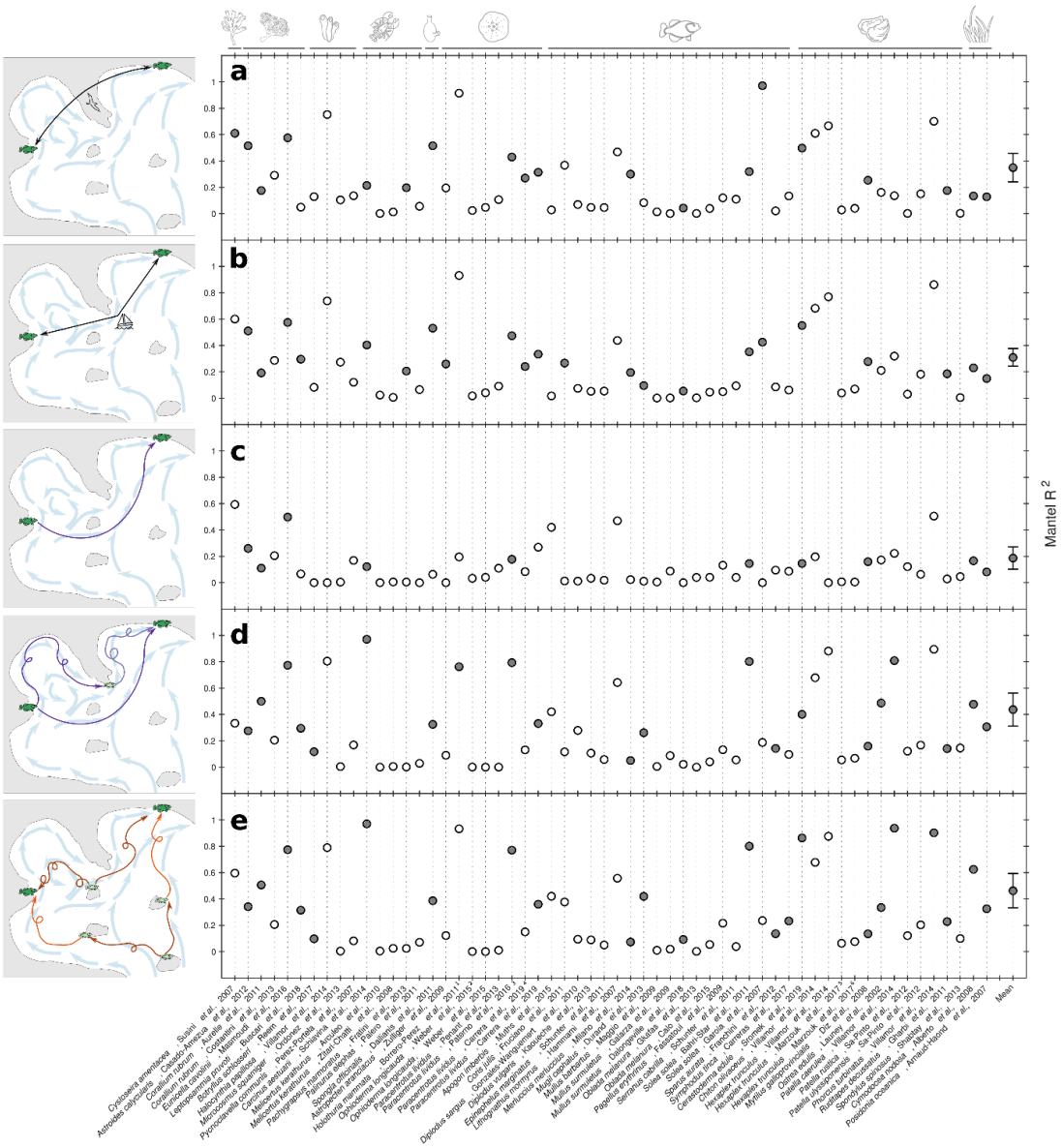


Figure SI-6: ***Models' accuracy in explaining observed genetic structures across the meta-analysis.*** *Mantel R² are referenced for* ***a*** *IBD (Euclidian) model,* ***b*** *IBD (sea least-cost) model,* ***c*** *single-generation explicit dispersal model,* ***d*** *multi-generation explicit dispersal model and* ***e*** *multi-generation implicit dispersal model. Filled dots correspond to significant correlations (p-value*). Note that (i) Weber et al., 2015 use SNPs marker ([1]) and mtDNA marker ([2]), (ii) Carrera et al., 2019 considers all the loci ([3]) and only the Mediterranean outliers loci ([4]), and (iii) Marzouk et al., 2017 use SNPs marker ([5]) and mtDNA marker ([6]).*

Table SI-5: ***Accuracy of the Isolation-by-distance (Euclidian and Sea least-cost distance) models in explaining observed genetic structures across the meta-analysis.*** *Mantel R^2 between log_e(distance) and Fst(1-Fst) and associated p-values are indexed for each study. Note that (i) Weber et al., 2015 use SNPs marker ([1]) and mtDNA marker ([2]), (ii) Carrera et al., 2019 considers all the loci ([3]) and only the Mediterranean outliers loci ([4]), and (iii) Marzouk et al., 2017 use SNPs marker ([5]) and mtDNA marker ([6]). Significant studies (p-values*) are highlighted in grey.*

		IBD Euclidian distance		IBD Sea least-cost distance	
Species	**Study**	**Mantel R^2**	**p-value**	**Mantel R^2**	**p-value**
Cystoseira amentacea	Susini et al., 2007	0,61	0,0397	0,60	0,0860
Astroides calycularis	Casado-Amezua et al., 2012	0,52	0,0001	0,51	0,0001
Corallium rubrum	Aurelle et al., 2011	0,18	0,0001	0,19	0,0001
Corallium rubrum	Costantini et al., 2013	0,29	0,0955	0,29	0,1040
Eunicella cavolinii	Masmoudi et al., 2016	0,58	0,0001	0,57	0,0001
Leptopsammia pruvoti	Boscari et al., 2019	0,05	0,2406	0,30	0,0002
Botryllus schlosseri	Reem et al., 2017	0,13	0,0980	0,08	0,0899
Halocynthia papillosa	Villamor et al., 2014	0,75	0,1301	0,74	0,1999
Microcosmus squamiger	Ordóñez et al., 2013	0,10	0,2113	0,27	0,1452
Pycnoclavella communis	Pérez-Portela et al., 2007	0,14	0,5847	0,12	0,7084
Carcinus aestuarii	Schiavina et al., 2014	0,21	0,0050	0,40	0,0036
Melicertus kerathurus	Arculeo et al., 2010	0,00	0,8539	0,02	0,4842
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,01	0,5010	0,01	0,6405
Pachygrapsus marmoratus	Fratini et al., 2013	0,20	0,0137	0,21	0,0142
Palinurus elephas	Palero et al., 2011	0,06	0,6515	0,07	0,6075
Spongia officinalis	Dailianis et al., 2011	0,52	0,0033	0,53	0,0044
Astropecten aranciacus	Zulliger et al., 2009	0,20	0,0517	0,26	0,0231
Holothuria mammata	Borrero-Pérez et al., 2011	0,91	0,1225	0,93	0,1232
Ophioderma longicauda	Weber et al., 2015[1]	0,02	0,4746	0,02	0,5433
Ophioderma longicauda	Weber et al., 2015[2]	0,05	0,2219	0,04	0,2707
Paracentrotus lividus	Penant et al., 2013	0,11	0,1262	0,09	0,1691
Paracentrotus lividus	Paterno et al., 2017	0,43	0,0194	0,47	0,0176
Paracentrotus lividus	Carreras et al., 2020[3]	0,27	0,0049	0,24	0,0109
Paracentrotus lividus	Carreras et al., 2020[4]	0,31	0,0155	0,33	0,0102
Apogon imberbis	Muths et al., 2015	0,03	0,7487	0,02	0,8322
Coris julis	Fruciano et al., 2011	0,37	0,0513	0,27	0,0236
Diplodus sargus	González-Wangüemert et al., 2010	0,07	0,3555	0,07	0,3626
Diplodus vulgaris	Kaouèche et al., 2013	0,05	0,3759	0,05	0,3601
Epinephelus marginatus	Schunter et al., 2011	0,05	0,3970	0,05	0,3724
Lithognathus mormyrus	Hammami et al., 2007	0,47	0,1699	0,44	0,1637
Merluccius merluccius	Milano et al., 2014	0,30	0,0001	0,19	0,0007
Mugil cephalus	Durand et al., 2013	0,08	0,0535	0,10	0,0312
Mullus barbartus	Maggio et al., 2009	0,01	0,4538	0,00	0,8701
Mullus surmuletus	Galarza et al., 2009	0,00	0,8976	0,00	0,9014
Mullus surmuletus	Dalongeville et al., 2018	0,04	0,0002	0,05	0,0001
Oblada melanura	Gkafas et al., 2013	0,00	0,7958	0,00	0,7772
Oblada melanura	Calò et al., 2016	0,04	0,2154	0,05	0,1673
Pagellus erythrinus	Fassatoui et al., 2009	0,12	0,2832	0,05	0,4901
Serranus cabrilla	Schunter et al., 2011	0,11	0,0806	0,09	0,1003
Solea solea	Bahri-Sfar et al., 2011	0,32	0,0031	0,35	0,0039
Solea solea	Garoia et al., 2007	0,97	0,0422	0,43	0,0418
Sparus aurata	Franchini et al., 2012	0,02	0,4093	0,09	0,1486
Symphodus tinca	Carreras et al., 2017	0,13	0,3353	0,06	0,5232
Cerastoderma edule	Sromek et al., 2019	0,50	0,0037	0,55	0,0014
Chiton olivaceus	Villamor et al., 2014	0,61	0,2132	0,68	0,1278
Hexaplex trunculus	Villamor et al., 2014	0,67	0,1253	0,77	0,1225
Hexaplex trunculus	Marzouk et al., 2017[5]	0,03	0,3212	0,04	0,2472
Hexaplex trunculus	Marzouk et al., 2017[6]	0,04	0,2163	0,07	0,1010
Mytilus galloprovincialis	Diz and Presa, 2008	0,25	0,0100	0,28	0,0078
Ostrea edulis	Launey et al., 2002	0,16	0,3905	0,21	0,1445
Patella caerulea	Villamor et al., 2014	0,14	0,1818	0,32	0,0546
Patella rustica	Sá-Pinto et al., 2012	0,00	0,9113	0,03	0,5970
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,15	0,2457	0,18	0,2450
Phorcus turbinatus	Villamor et al., 2014	0,70	0,2995	0,86	0,2909
Ruditapes decussatus	Gharbi et al., 2011	0,18	0,0028	0,19	0,0045
Spondylus spinosus	Shabtay et al., 2014	0,00	0,8636	0,00	0,8536
Cymodocea nodosa	Alberto et al., 2008	0,13	0,0201	0,23	0,0165
Posidonia oceanica	Arnaud-Haond et al., 2007	0,13	0,0001	0,15	0,0001

Table SI-6: ***Accuracy of our dispersal models (i.e. single-generation explicit, multi-generation explicit, and multi-generation implicit) in explaining observed genetic structures across the meta-analysis.*** *Mantel R² between* modelled *and* observed Fst *and associated p-values are indexed for each study. Note that (i) Weber et al., 2015 use SNPs marker ([1]) and mtDNA marker ([2]), (ii) Carrera et al., 2019 considers all the loci ([3]) and only the Mediterranean outliers loci ([4]), and (iii) Marzouk et al., 2017 use SNPs marker ([5]) and mtDNA marker ([6]). Significant studies (p-values*) are highlighted in grey.*

		Single-generation explicit		Multi-generation explicit			Multi-generation implicit		
Species	**Study**	**Mantel R²**	**p-value**	**Mantel R²**	**p-value**	**Optimal M**	**Mantel R²**	**p-value**	**Optimal M**
Cystoseira amentacea	Susini et al., 2007	0,59	0,1631	0,33	0,0849	5	0,60	0,0819	1
Astroides calycularis	Casado-Amezua et al., 2012	0,26	0,0001	0,28	0,0001	5	0,34	0,0060	300
Corallium rubrum	Aurelle et al., 2011	0,11	0,0001	0,50	0,0001	40	0,51	0,0010	80
Corallium rubrum	Costantini et al., 2013	0,21	0,0705	0,21	0,0705	1	0,21	0,0819	1
Eunicella cavolinii	Masmoudi et al., 2016	0,50	0,0001	0,77	0,0001	300	0,77	0,0010	300
Leptopsammia pruvoti	Boscari et al., 2019	0,07	0,0584	0,29	0,0001	10	0,31	0,0020	20
Botryllus schlosseri	Reem et al., 2017	0,00	1,0000	0,12	0,0147	500	0,10	0,0070	500
Halocynthia papillosa	Villamor et al., 2014	0,00	1,0000	0,80	0,1668	5	0,79	0,1768	10
Microcosmus squamiger	Ordóñez et al., 2013	0,00	1,0000	0,00	1,0000	1	0,00	1,0000	1
Pycnoclavella communis	Pérez-Portela et al., 2007	0,17	0,5038	0,17	0,4636	500	0,08	0,8062	500
Carcinus aestuarii	Schiavina et al., 2014	0,12	0,0237	0,97	0,0178	60	0,97	0,0150	40
Melicertus kerathurus	Arculeo et al., 2010	0,00	0,9541	0,00	0,9541	1	0,01	0,7772	10
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,01	0,6600	0,01	0,6600	1	0,03	0,4795	5
Pachygrapsus marmoratus	Fratini et al., 2013	0,01	0,6792	0,00	0,7097	500	0,02	0,2637	5
Palinurus elephas	Palero et al., 2011	0,00	1,0000	0,03	0,6986	40	0,07	0,8232	20
Spongia officinalis	Dailianis et al., 2011	0,06	0,1457	0,32	0,0057	500	0,39	0,0080	500
Astropecten aranciacus	Zulliger et al., 2009	0,00	1,0000	0,09	0,2136	10	0,12	0,0939	1
Holothuria mammata	Borrero-Pérez et al., 2011	0,20	0,4996	0,76	0,0430	150	0,93	0,0759	60
Ophioderma longicauda	Weber et al., 2015[1]	0,03	0,4368	0,00	0,8801	500	0,00	0,8302	500
Ophioderma longicauda	Weber et al., 2015[2]	0,04	0,2902	0,00	0,9989	200	0,00	0,9201	500
Paracentrotus lividus	Penant et al., 2013	0,11	0,1598	0,00	0,9836	500	0,01	0,7143	40
Paracentrotus lividus	Paterno et al., 2017	0,18	0,0348	0,79	0,0045	40	0,77	0,0030	40
Paracentrotus lividus	Carreras et al., 2020[3]	0,08	0,2851	0,13	0,0810	40	0,15	0,0739	40
Paracentrotus lividus	Carreras et al., 2020[4]	0,27	0,0693	0,33	0,0085	40	0,36	0,0060	40
Apogon imberbis	Muths et al., 2015	0,42	0,5047	0,42	0,4975	5	0,42	0,4925	1
Coris julis	Fruciano et al., 2011	0,01	0,4102	0,12	0,0633	80	0,38	0,0559	300
Diplodus sargus	González-Wangüemert et al., 2010	0,01	1,0000	0,28	0,0696	150	0,09	0,3586	10
Diplodus vulgaris	Kaouèche et al., 2013	0,03	0,5417	0,11	0,2378	5	0,09	0,2877	1
Epinephelus marginatus	Schunter et al., 2011	0,02	0,4578	0,06	0,2208	10	0,05	0,3606	20
Lithognathus mormyrus	Hammami et al., 2007	0,47	0,2027	0,64	0,0813	5	0,56	0,0839	500
Merluccius merluccius	Milano et al., 2014	0,02	0,1058	0,05	0,0457	10	0,07	0,0210	10
Mugil cephalus	Durand et al., 2013	0,01	0,3320	0,26	0,0007	20	0,42	0,0010	20
Mullus barbartus	Maggio et al., 2009	0,01	0,6078	0,01	0,6078	1	0,01	0,3846	1
Mullus surmuletus	Galarza et al., 2009	0,09	0,1508	0,09	0,1508	1	0,02	0,8641	60
Mullus surmuletus	Dalongeville et al., 2018	0,00	0,9769	0,02	0,0782	60	0,09	0,0250	100
Oblada melanura	Gkafas et al., 2013	0,04	0,3266	0,00	0,9552	10	0,00	0,7742	500
Oblada melanura	Calò et al., 2016	0,04	0,2626	0,04	0,2626	1	0,05	0,1528	1
Pagellus erythrinus	Fassatoui et al., 2009	0,13	0,1276	0,13	0,1276	1	0,22	0,0609	20
Serranus cabrilla	Schunter et al., 2011	0,04	0,1515	0,05	0,0994	500	0,04	0,1389	1
Solea solea	Bahri-Sfar et al., 2011	0,15	0,0047	0,80	0,0031	60	0,80	0,0030	20
Solea solea	Garoia et al., 2007	0,00	1,0000	0,19	0,0883	500	0,24	0,1299	5
Sparus aurata	Franchini et al., 2012	0,10	0,0883	0,14	0,0483	10	0,14	0,0480	5
Symphodus tinca	Carreras et al., 2017	0,09	0,3748	0,10	0,4430	5	0,23	0,0260	20
Cerastoderma edule	Sromek et al., 2019	0,15	0,0482	0,40	0,0165	10	0,86	0,0080	20
Chiton olivaceus	Villamor et al., 2014	0,20	0,6666	0,68	0,1708	500	0,68	0,0919	80
Hexaplex trunculus	Villamor et al., 2014	0,00	1,0000	0,88	0,0806	500	0,88	0,0759	150
Hexaplex trunculus	Marzouk et al., 2017[5]	0,01	0,4010	0,05	0,1873	60	0,06	0,1399	400
Hexaplex trunculus	Marzouk et al., 2017[6]	0,01	0,4732	0,07	0,1506	80	0,08	0,0939	200
Mytilus galloprovincialis	Diz and Presa, 2008	0,16	0,0122	0,16	0,0122	1	0,14	0,0120	1
Ostrea edulis	Launey et al., 2002	0,17	0,3095	0,49	0,0170	20	0,34	0,0410	60
Patella caerulea	Villamor et al., 2014	0,22	0,0900	0,81	0,0190	80	0,94	0,0220	80
Patella rustica	Sá-Pinto et al., 2012	0,12	0,1996	0,12	0,1996	1	0,12	0,2168	1
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,06	0,3954	0,17	0,2061	150	0,20	0,1089	150
Phorcus turbinatus	Villamor et al., 2014	0,51	0,3364	0,89	0,2887	40	0,90	0,0470	500
Ruditapes decussatus	Gharbi et al., 2011	0,03	0,2616	0,14	0,0107	10	0,23	0,0070	10
Spondylus spinosus	Shabtay et al., 2014	0,05	0,7992	0,15	0,3700	5	0,10	0,6334	500
Cymodocea nodosa	Alberto et al., 2008	0,17	0,0049	0,48	0,0062	80	0,62	0,0010	10
Posidonia oceanica	Arnaud-Haond et al., 2007	0,08	0,0001	0,31	0,0001	60	0,33	0,0010	40

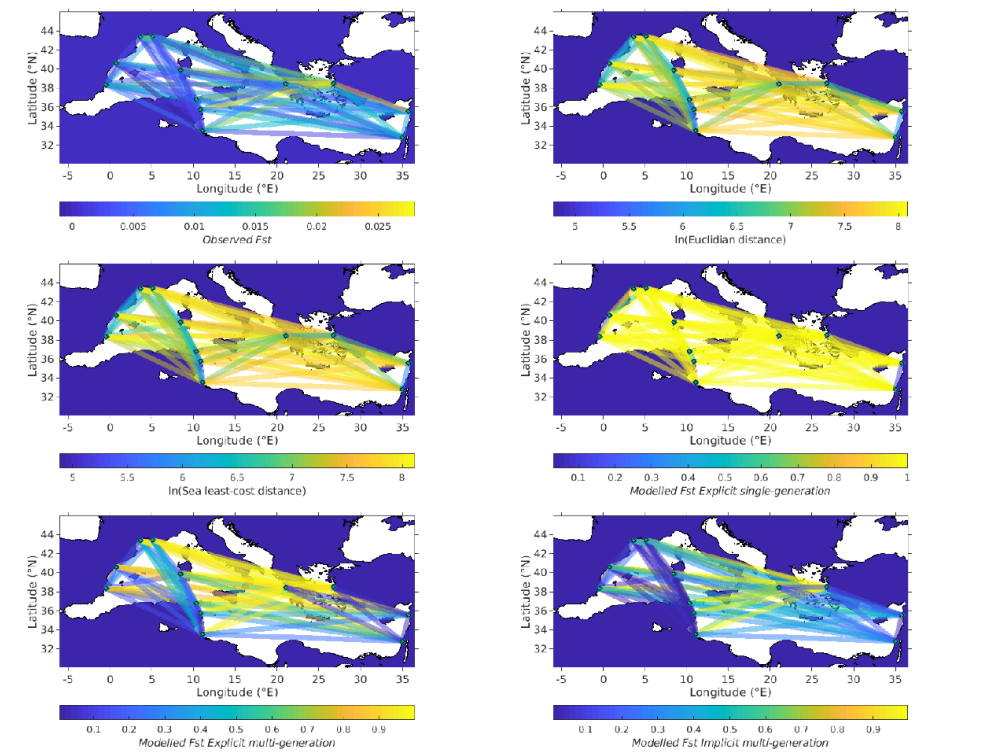


Figure SI-7: ***Network representation of*** **observed** ***and*** **modelled** ***genetic structures*** *for* Mugil cephalus*; Durand et al., (2013).* ***a*** Observed Fst *values,* ***b*** log_e*(Euclidian distance),* ***c*** log_e*(Sea least-cost distance),* ***d*** *single-generation explicit* modelled Fst *,* ***e*** *multi-generation explicit* modelled Fst *and* ***f*** *multi-generation implicit* modelled Fst*. Similar plots made for each study of our meta-analysis can be viewed at: https://nuage.osupytheas.fr/s/wajsAZrRCRrwHHk*

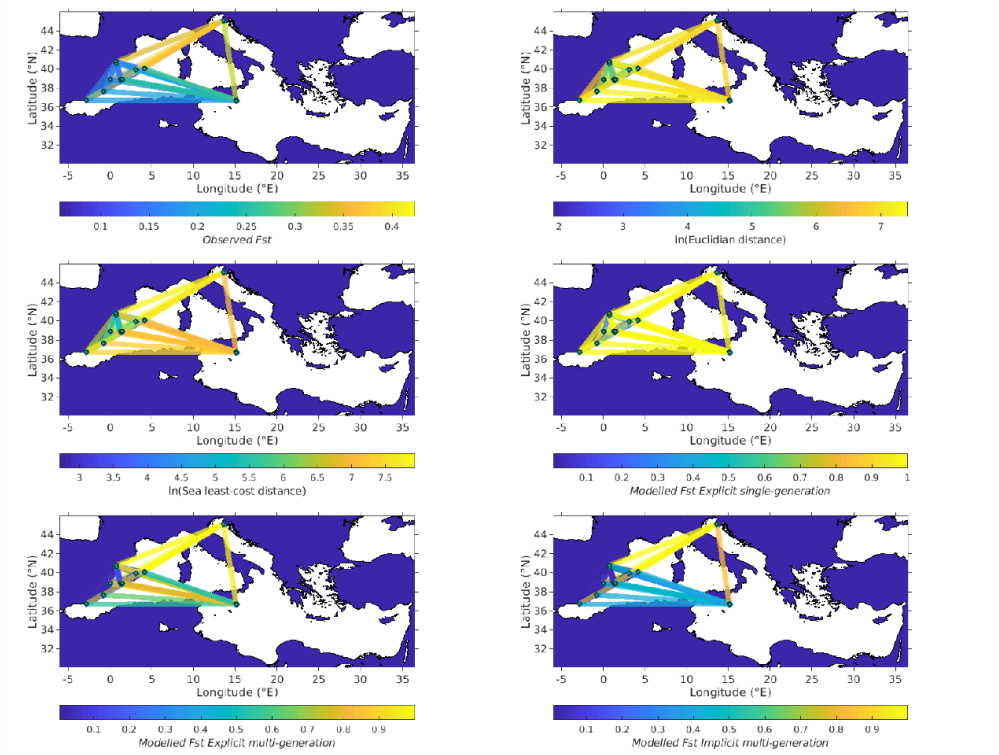


Figure SI-8: ***Network representation of*** **observed** ***and*** **modelled** ***genetic structures*** *for* Cymodocea nodosa*; Alberto et al., (2008).* ***a*** Observed Fst *values,* ***b*** *$\log_e$(Euclidian distance),* ***c*** *$\log_e$(Sea least-cost distance),* ***d*** *single-generation explicit* modelled Fst *,* ***e*** *multi-generation explicit* modelled Fst *and* ***f*** *multi-generation implicit* modelled Fst*. Similar plots made for each study of our meta-analysis can be viewed at: https://nuage.osupytheas.fr/s/wajsAZrRCRrwHHk*

VI. Cross-comparison of Mantel correlations

Table SI-7: ***Significance (p-values) of Mantel correlations between IBD Euclidian distance and Fst/(1-Fst) being greater*** *than (i) correlations between IBD sea least-cost distance and Fst/(1-Fst) (Hendrickson et al., 1970), (ii) correlations between single-generation explicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004), (iii) correlations between multi-generation explicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004) and (iv) correlations between multi-generation implicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004).*

		IBD Euclidian distance *vs*			
Species	**Study**	**IBD sea least-cost distance**	**Single-generation explicit**	**Multi-generation explicit**	**Multi-generation implicit**
Cystoseira amentacea	Susini et al., 2007	0,3816	0,4759	0,3546	0,4789
Astroides calycularis	Casado-Amezua et al., 2012	0,1330	0,0355	0,0586	0,1896
Corallium rubrum	Aurelle et al., 2011	0,7612	0,2935	0,9573	0,9608
Corallium rubrum	Costantini et al., 2013	0,4081	0,4149	0,4149	0,4153
Eunicella cavolinii	Masmoudi et al., 2016	0,4832	0,2806	0,9282	0,9303
Leptopsammia pruvoti	Boscari et al., 2019	0,9731	0,5565	0,8393	0,8197
Botryllus schlosseri	Reem et al., 2017	0,3644	NaN	0,4776	0,4284
Halocynthia papillosa	Villamor et al., 2014	0,4739	NaN	0,7675	0,5911
Microcosmus squamiger	Ordóñez et al., 2013	0,2844	0,6454	0,6454	0,6454
Pycnoclavella communis	Pérez-Portela et al., 2007	0,5560	0,4641	0,4641	0,6090
Carcinus aestuarii	Schiavina et al., 2014	0,9303	0,3160	0,9945	0,9943
Melicertus kerathurus	Arculeo et al., 2010	0,3491	0,5459	0,5459	0,6219
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,3261	0,4360	0,4360	0,5693
Pachygrapsus marmoratus	Fratini et al., 2013	0,5502	0,1061	0,1770	0,1709
Palinurus elephas	Palero et al., 2011	0,8660	NaN	0,4534	0,5217
Spongia officinalis	Dailianis et al., 2011	0,6902	0,0498	0,1982	0,2678
Astropecten aranciacus	Zulliger et al., 2009	0,5890	NaN	0,3810	0,4220
Holothuria mammata	Borrero-Pérez et al., 2011	0,8032	0,1684	0,2616	0,6123
Ophioderma longicauda	Weber et al., 2015[1]	0,6810	0,4487	0,8007	0,6803
Ophioderma longicauda	Weber et al., 2015[2]	0,6872	0,5289	0,8535	0,8522
Paracentrotus lividus	Penant et al., 2013	0,6195	0,4902	0,7951	0,8849
Paracentrotus lividus	Paterno et al., 2017	0,6573	0,1304	0,9655	0,9419
Paracentrotus lividus	Carreras et al., 2020[3]	0,3744	0,1467	0,2128	0,2867
Paracentrotus lividus	Carreras et al., 2020[4]	0,5758	0,4164	0,5333	0,5737
Apogon imberbis	Muths et al., 2015	0,7376	0,3360	0,7146	0,3360
Coris julis	Fruciano et al., 2011	0,2252	0,0795	0,1748	0,5139
Diplodus sargus	González-Wangüemert et al., 2010	0,5894	0,4210	0,1581	0,5293
Diplodus vulgaris	Kaouèche et al., 2013	0,5372	0,4641	0,5513	0,6277
Epinephelus marginatus	Schunter et al., 2011	0,7109	0,3945	0,5310	0,5106
Lithognathus mormyrus	Hammami et al., 2007	0,4210	0,5023	0,6643	0,7487
Merluccius merluccius	Milano et al., 2014	0,1846	0,0867	0,0968	0,1240
Mugil cephalus	Durand et al., 2013	0,7225	0,2650	0,8066	0,9095
Mullus barbartus	Maggio et al., 2009	0,2813	0,4245	0,4245	0,4717
Mullus surmuletus	Galarza et al., 2009	0,3970	0,7817	0,7817	0,6586
Mullus surmuletus	Dalongeville et al., 2018	0,7947	0,0848	0,3613	0,6927
Oblada melanura	Gkafas et al., 2013	0,5238	0,2391	0,4268	0,5056
Oblada melanura	Calò et al., 2016	0,6479	0,5072	0,5072	0,5819
Pagellus erythrinus	Fassatoui et al., 2009	0,1920	0,5245	0,5245	0,5996
Serranus cabrilla	Schunter et al., 2011	0,1811	0,2823	0,0629	0,2651
Solea solea	Bahri-Sfar et al., 2011	0,7911	0,1653	0,9679	0,9580
Solea solea	Garoia et al., 2007	0,1281	NaN	0,1299	0,1233
Sparus aurata	Franchini et al., 2012	0,7538	0,7069	0,7355	0,7368
Symphodus tinca	Carreras et al., 2017	0,3153	0,4508	0,4394	0,5902
Cerastoderma edule	Sromek et al., 2019	0,7104	0,2006	0,3908	0,9388
Chiton olivaceus	Villamor et al., 2014	0,6202	0,2362	0,6188	0,6142
Hexaplex trunculus	Villamor et al., 2014	0,6510	NaN	0,7503	0,7390
Hexaplex trunculus	Marzouk et al., 2017[5]	0,7044	0,2268	0,6849	0,7402
Hexaplex trunculus	Marzouk et al., 2017[6]	0,8758	0,2074	0,6752	0,7239
Mytilus galloprovincialis	Diz and Presa, 2008	0,6658	0,2656	0,2656	0,2466
Ostrea edulis	Launey et al., 2002	0,5692	0,5174	0,7991	0,6294
Patella caerulea	Villamor et al., 2014	0,8472	0,5932	0,9069	0,9647
Patella rustica	Sá-Pinto et al., 2012	0,7776	0,6831	0,6831	0,6831
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,6712	0,3788	0,5176	0,5596
Phorcus turbinatus	Villamor et al., 2014	0,9038	0,1251	0,9623	0,9063
Ruditapes decussatus	Gharbi et al., 2011	0,5587	0,1609	0,4106	0,6166
Spondylus spinosus	Shabtay et al., 2014	0,3749	0,6732	0,6407	0,7205
Cymodocea nodosa	Alberto et al., 2008	0,9817	0,5829	0,9046	0,9797
Posidonia oceanica	Arnaud-Haond et al., 2007	0,8878	0,2993	0,8578	0,8610

Table SI-8: ***Significance (p-values) of Mantel correlations between IBD sea least-cost distance and Fst/(1-Fst) being greater*** *than (i) correlations between IBD Euclidian distance and Fst/(1-Fst) (Hendrickson et al., 1970), (ii) correlations between single-generation explicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004), (iii) correlations between multi-generation explicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004) and (iv) correlations between multi-generation implicit dispersal* modelled Fst *and* observed Fst *(Silver et al., 2004).*

		IBD sea least-cost distance *vs*			
Species	**Study**	**IBD Euclidian distance**	**Single-generation explicit**	**Multi-generation explicit**	**Multi-generation implicit**
Cystoseira amentacea	Susini et al., 2007	0,6184	0,4893	0,3629	0,4928
Astroides calycularis	Casado-Amezua et al., 2012	0,8670	0,0377	0,0622	0,1966
Corallium rubrum	Aurelle et al., 2011	0,2388	0,2613	0,9522	0,9562
Corallium rubrum	Costantini et al., 2013	0,5919	0,4224	0,4224	0,4228
Eunicella cavolinii	Masmoudi et al., 2016	0,5168	0,2868	0,9273	0,9292
Leptopsammia pruvoti	Boscari et al., 2019	0,0269	0,1487	0,4986	0,5267
Botryllus schlosseri	Reem et al., 2017	0,6356	NaN	0,5865	0,5513
Halocynthia papillosa	Villamor et al., 2014	0,5261	NaN	0,7439	0,9634
Microcosmus squamiger	Ordóñez et al., 2013	0,7156	0,7656	0,7656	0,7656
Pycnoclavella communis	Pérez-Portela et al., 2007	0,4440	0,4566	0,4566	0,6301
Carcinus aestuarii	Schiavina et al., 2014	0,0697	0,0830	0,9948	0,9946
Melicertus kerathurus	Arculeo et al., 2010	0,6509	0,6572	0,6572	0,7996
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,6739	0,4939	0,4939	0,6142
Pachygrapsus marmoratus	Fratini et al., 2013	0,4498	0,1105	0,1848	0,1764
Palinurus elephas	Palero et al., 2011	0,1340	NaN	0,4411	0,5083
Spongia officinalis	Dailianis et al., 2011	0,3098	0,0502	0,1789	0,2386
Astropecten aranciacus	Zulliger et al., 2009	0,4110	NaN	0,2500	0,3453
Holothuria mammata	Borrero-Pérez et al., 2011	0,1968	0,1529	0,2229	0,5102
Ophioderma longicauda	Weber et al., 2015[1]	0,3190	0,4050	0,7846	0,6701
Ophioderma longicauda	Weber et al., 2015[2]	0,3128	0,5016	0,8424	0,8389
Paracentrotus lividus	Penant et al., 2013	0,3805	0,4317	0,8073	0,8995
Paracentrotus lividus	Paterno et al., 2017	0,3427	0,1213	0,9743	0,9486
Paracentrotus lividus	Carreras et al., 2020[3]	0,6256	0,1787	0,2195	0,3028
Paracentrotus lividus	Carreras et al., 2020[4]	0,4242	0,3827	0,4946	0,5500
Apogon imberbis	Muths et al., 2015	0,2624	0,2942	0,6753	0,2942
Coris julis	Fruciano et al., 2011	0,7748	0,1277	0,2408	0,6432
Diplodus sargus	González-Wangüemert et al., 2010	0,4106	0,4141	0,1558	0,5236
Diplodus vulgaris	Kaouèche et al., 2013	0,4628	0,4490	0,5491	0,6205
Epinephelus marginatus	Schunter et al., 2011	0,2891	0,3808	0,5117	0,4908
Lithognathus mormyrus	Hammami et al., 2007	0,5790	0,6005	0,7085	0,7903
Merluccius merluccius	Milano et al., 2014	0,8154	0,1448	0,1309	0,1761
Mugil cephalus	Durand et al., 2013	0,2775	0,2414	0,7884	0,9005
Mullus barbartus	Maggio et al., 2009	0,7187	0,5667	0,5667	0,6212
Mullus surmuletus	Galarza et al., 2009	0,6030	0,7852	0,7852	0,6623
Mullus surmuletus	Dalongeville et al., 2018	0,2053	0,0569	0,2863	0,6555
Oblada melanura	Gkafas et al., 2013	0,4762	0,2430	0,4216	0,4986
Oblada melanura	Calò et al., 2016	0,3521	0,4737	0,4737	0,5445
Pagellus erythrinus	Fassatoui et al., 2009	0,8080	0,6996	0,6996	0,7077
Serranus cabrilla	Schunter et al., 2011	0,8189	0,3198	0,0662	0,3046
Solea solea	Bahri-Sfar et al., 2011	0,2089	0,1163	0,9619	0,9509
Solea solea	Garoia et al., 2007	0,8719	NaN	0,3427	0,1952
Sparus aurata	Franchini et al., 2012	0,2462	0,5285	0,6648	0,6604
Symphodus tinca	Carreras et al., 2017	0,6847	0,5344	0,5639	0,6786
Cerastoderma edule	Sromek et al., 2019	0,2896	0,1737	0,3357	0,9316
Chiton olivaceus	Villamor et al., 2014	0,3798	0,2534	0,4667	0,4694
Hexaplex trunculus	Villamor et al., 2014	0,3490	NaN	0,7647	0,7383
Hexaplex trunculus	Marzouk et al., 2017[5]	0,2956	0,2063	0,6034	0,6560
Hexaplex trunculus	Marzouk et al., 2017[6]	0,1242	0,1637	0,4874	0,5347
Mytilus galloprovincialis	Diz and Presa, 2008	0,3342	0,2375	0,2375	0,2260
Ostrea edulis	Launey et al., 2002	0,4308	0,4489	0,8337	0,6289
Patella caerulea	Villamor et al., 2014	0,1528	0,3979	0,8914	0,9683
Patella rustica	Sá-Pinto et al., 2012	0,2224	0,6313	0,6313	0,6313
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,3288	0,3251	0,4841	0,5244
Phorcus turbinatus	Villamor et al., 2014	0,0962	0,0908	0,8806	0,6953
Ruditapes decussatus	Gharbi et al., 2011	0,4413	0,1562	0,3715	0,6095
Spondylus spinosus	Shabtay et al., 2014	0,6251	0,6853	0,6438	0,7288
Cymodocea nodosa	Alberto et al., 2008	0,0183	0,3522	0,8515	0,9718
Posidonia oceanica	Arnaud-Haond et al., 2007	0,1122	0,2313	0,8296	0,8358

Table SI-9: ***Significance (p-values) of Mantel correlations between single-generation explicit dispersal*** **modelled Fst *and* observed Fst *being greater*** *than (i) correlations between IBD Euclidian distance and Fst/(1-Fst) (Silver et al., 2004), (ii) correlations between IBD sea least-cost distance and Fst/(1-Fst) (Silver et al., 2004), (iii) correlations between multi-generation explicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970) and (iv) correlations between multi-generation implicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970).*

		Single-generation explicit *vs*			
Species	**Study**	**IBD Euclidian distance**	**IBD sea least-cost distance**	**Multi-generation explicit**	**Multi-generation implicit**
Cystoseira amentacea	Susini et al., 2007	0,5241	0,5107	0,4028	0,6320
Astroides calycularis	Casado-Amezua et al., 2012	0,9645	0,9623	0,5787	0,6371
Corallium rubrum	Aurelle et al., 2011	0,7065	0,7387	0,9748	0,9764
Corallium rubrum	Costantini et al., 2013	0,5851	0,5776	NaN	0,6225
Eunicella cavolinii	Masmoudi et al., 2016	0,7194	0,7132	0,9695	0,9708
Leptopsammia pruvoti	Boscari et al., 2019	0,4435	0,8513	0,7761	0,7766
Botryllus schlosseri	Reem et al., 2017	NaN	NaN	0,7568	0,7355
Halocynthia papillosa	Villamor et al., 2014	NaN	NaN	0,7943	0,7881
Microcosmus squamiger	Ordóñez et al., 2013	0,3546	0,2344	0,5000	0,5000
Pycnoclavella communis	Pérez-Portela et al., 2007	0,5359	0,5434	0,5010	0,5542
Carcinus aestuarii	Schiavina et al., 2014	0,6840	0,9170	0,9995	0,9995
Melicertus kerathurus	Arculeo et al., 2010	0,4541	0,3428	NaN	0,5526
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,5640	0,5061	0,5000	0,5907
Pachygrapsus marmoratus	Fratini et al., 2013	0,8939	0,8895	0,6102	0,7101
Palinurus elephas	Palero et al., 2011	NaN	NaN	0,5604	0,5960
Spongia officinalis	Dailianis et al., 2011	0,9502	0,9498	0,8538	0,8654
Astropecten aranciacus	Zulliger et al., 2009	NaN	NaN	0,6603	0,6872
Holothuria mammata	Borrero-Pérez et al., 2011	0,8316	0,8471	0,7600	0,8744
Ophioderma longicauda	Weber et al., 2015[1]	0,5513	0,5950	0,7592	0,7043
Ophioderma longicauda	Weber et al., 2015[2]	0,4711	0,4984	0,7212	0,7478
Paracentrotus lividus	Penant et al., 2013	0,5098	0,5683	0,8164	0,8981
Paracentrotus lividus	Paterno et al., 2017	0,8696	0,8787	0,9831	0,9762
Paracentrotus lividus	Carreras et al., 2020[3]	0,8533	0,8213	0,5907	0,6104
Paracentrotus lividus	Carreras et al., 2020[4]	0,5836	0,6173	0,5815	0,6072
Apogon imberbis	Muths et al., 2015	0,6640	0,7058	0,5000	NaN
Coris julis	Fruciano et al., 2011	0,9205	0,8723	0,6932	0,8814
Diplodus sargus	González-Wangüemert et al., 2010	0,5790	0,5859	0,2098	0,5963
Diplodus vulgaris	Kaouèche et al., 2013	0,5359	0,5510	0,5835	0,6614
Epinephelus marginatus	Schunter et al., 2011	0,6055	0,6192	0,5989	0,5702
Lithognathus mormyrus	Hammami et al., 2007	0,4977	0,3995	0,6458	0,6994
Merluccius merluccius	Milano et al., 2014	0,9133	0,8552	0,5924	0,6483
Mugil cephalus	Durand et al., 2013	0,7350	0,7586	0,8802	0,9448
Mullus barbartus	Maggio et al., 2009	0,5755	0,4333	NaN	0,6217
Mullus surmuletus	Galarza et al., 2009	0,2183	0,2148	NaN	0,4100
Mullus surmuletus	Dalongeville et al., 2018	0,9152	0,9431	0,7874	0,9425
Oblada melanura	Gkafas et al., 2013	0,7609	0,7570	0,6478	0,6831
Oblada melanura	Calò et al., 2016	0,4928	0,5263	NaN	0,5727
Pagellus erythrinus	Fassatoui et al., 2009	0,4755	0,3004	NaN	0,5867
Serranus cabrilla	Schunter et al., 2011	0,7177	0,6802	0,1334	0,4868
Solea solea	Bahri-Sfar et al., 2011	0,8347	0,8837	0,9854	0,9841
Solea solea	Garoia et al., 2007	NaN	NaN	0,6038	0,6185
Sparus aurata	Franchini et al., 2012	0,2931	0,4715	0,5881	0,5796
Symphodus tinca	Carreras et al., 2017	0,5492	0,4656	0,5117	0,6136
Cerastoderma edule	Sromek et al., 2019	0,7994	0,8263	0,7413	0,9714
Chiton olivaceus	Villamor et al., 2014	0,7638	0,7466	0,7137	0,7130
Hexaplex trunculus	Villamor et al., 2014	NaN	NaN	0,8293	0,8266
Hexaplex trunculus	Marzouk et al., 2017[5]	0,7732	0,7937	0,8155	0,8239
Hexaplex trunculus	Marzouk et al., 2017[6]	0,7926	0,8363	0,8244	0,8332
Mytilus galloprovincialis	Diz and Presa, 2008	0,7344	0,7625	NaN	0,4153
Ostrea edulis	Launey et al., 2002	0,4826	0,5511	0,7903	0,5943
Patella caerulea	Villamor et al., 2014	0,4068	0,6021	0,9022	0,9765
Patella rustica	Sá-Pinto et al., 2012	0,3169	0,3687	0,5000	0,5000
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,6212	0,6749	0,5850	0,6075
Phorcus turbinatus	Villamor et al., 2014	0,8749	0,9092	0,9380	0,8972
Ruditapes decussatus	Gharbi et al., 2011	0,8391	0,8438	0,7337	0,8199
Spondylus spinosus	Shabtay et al., 2014	0,3268	0,3147	0,5593	0,7133
Cymodocea nodosa	Alberto et al., 2008	0,4171	0,6478	0,8742	0,9700
Posidonia oceanica	Arnaud-Haond et al., 2007	0,7007	0,7687	0,9143	0,9177

Table SI-10: ***Significance (p-values) of Mantel correlations between multi-generation explicit dispersal*** **modelled Fst *and* observed Fst *being greater*** *than (i) correlations between IBD Euclidian distance and Fst/(1-Fst) (Silver et al., 2004), (ii) correlations between IBD sea least-cost distance and Fst/(1-Fst) (Silver et al., 2004), (iii) correlations between single-generation explicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970) and (iv) correlations between multi-generation implicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970).*

		Multi-generation explicit *vs*			
Species	**Study**	**IBD Euclidian distance**	**IBD sea least-cost distance**	**Single-generation explicit**	**Multi-generation implicit**
Cystoseira amentacea	Susini et al., 2007	0,6454	0,6371	0,5972	0,6907
Astroides calycularis	Casado-Amezua et al., 2012	0,9414	0,9378	0,4213	0,8243
Corallium rubrum	Aurelle et al., 2011	0,0427	0,0478	0,0252	0,6750
Corallium rubrum	Costantini et al., 2013	0,5851	0,5776	NaN	0,7374
Eunicella cavolinii	Masmoudi et al., 2016	0,0718	0,0727	0,0305	0,6217
Leptopsammia pruvoti	Boscari et al., 2019	0,1607	0,5014	0,2239	0,5813
Botryllus schlosseri	Reem et al., 2017	0,5224	0,4135	0,2432	0,3104
Halocynthia papillosa	Villamor et al., 2014	0,2325	0,2561	0,2057	0,3885
Microcosmus squamiger	Ordóñez et al., 2013	0,3546	0,2344	0,5000	0,5000
Pycnoclavella communis	Pérez-Portela et al., 2007	0,5359	0,5434	0,4990	0,6073
Carcinus aestuarii	Schiavina et al., 2014	0,0055	0,0052	0,0005	0,3955
Melicertus kerathurus	Arculeo et al., 2010	0,4541	0,3428	NaN	0,6258
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,5640	0,5061	0,5000	0,7109
Pachygrapsus marmoratus	Fratini et al., 2013	0,8230	0,8152	0,3898	0,7893
Palinurus elephas	Palero et al., 2011	0,5466	0,5589	0,4396	0,6560
Spongia officinalis	Dailianis et al., 2011	0,8018	0,8211	0,1462	0,8192
Astropecten aranciacus	Zulliger et al., 2009	0,6190	0,7500	0,3397	0,5959
Holothuria mammata	Borrero-Pérez et al., 2011	0,7384	0,7771	0,2400	0,9302
Ophioderma longicauda	Weber et al., 2015[1]	0,1993	0,2154	0,2408	0,5031
Ophioderma longicauda	Weber et al., 2015[2]	0,1465	0,1576	0,2788	0,7134
Paracentrotus lividus	Penant et al., 2013	0,2049	0,1927	0,1836	0,8203
Paracentrotus lividus	Paterno et al., 2017	0,0345	0,0257	0,0169	0,0994
Paracentrotus lividus	Carreras et al., 2020[3]	0,7872	0,7805	0,4093	0,6979
Paracentrotus lividus	Carreras et al., 2020[4]	0,4667	0,5054	0,4185	0,7200
Apogon imberbis	Muths et al., 2015	0,2854	0,3247	0,5000	0,5000
Coris julis	Fruciano et al., 2011	0,8252	0,7592	0,3068	0,9976
Diplodus sargus	González-Wangüemert et al., 2010	0,8419	0,8442	0,7902	0,9894
Diplodus vulgaris	Kaouèche et al., 2013	0,4487	0,4509	0,4165	0,4705
Epinephelus marginatus	Schunter et al., 2011	0,4690	0,4883	0,4011	0,4525
Lithognathus mormyrus	Hammami et al., 2007	0,3357	0,2915	0,3542	0,3404
Merluccius merluccius	Milano et al., 2014	0,9032	0,8691	0,4076	0,8964
Mugil cephalus	Durand et al., 2013	0,1934	0,2116	0,1198	0,9947
Mullus barbartus	Maggio et al., 2009	0,5755	0,4333	NaN	0,8145
Mullus surmuletus	Galarza et al., 2009	0,2183	0,2148	NaN	0,3144
Mullus surmuletus	Dalongeville et al., 2018	0,6387	0,7137	0,2126	1,0000
Oblada melanura	Gkafas et al., 2013	0,5732	0,5784	0,3522	0,8326
Oblada melanura	Calò et al., 2016	0,4928	0,5263	NaN	0,6717
Pagellus erythrinus	Fassatoui et al., 2009	0,4755	0,3004	NaN	0,6792
Serranus cabrilla	Schunter et al., 2011	0,9371	0,9338	0,8666	0,9991
Solea solea	Bahri-Sfar et al., 2011	0,0321	0,0381	0,0146	0,4832
Solea solea	Garoia et al., 2007	0,8701	0,6573	0,3962	0,5658
Sparus aurata	Franchini et al., 2012	0,2645	0,3352	0,4119	0,3629
Symphodus tinca	Carreras et al., 2017	0,5606	0,4361	0,4883	0,7637
Cerastoderma edule	Sromek et al., 2019	0,6092	0,6643	0,2587	1,0000
Chiton olivaceus	Villamor et al., 2014	0,3812	0,5333	0,2863	0,4864
Hexaplex trunculus	Villamor et al., 2014	0,2497	0,2353	0,1707	0,3785
Hexaplex trunculus	Marzouk et al., 2017[5]	0,3151	0,3966	0,1845	0,6891
Hexaplex trunculus	Marzouk et al., 2017[6]	0,3248	0,5126	0,1756	0,7427
Mytilus galloprovincialis	Diz and Presa, 2008	0,7344	0,7625	NaN	0,3094
Ostrea edulis	Launey et al., 2002	0,2009	0,1663	0,2097	0,2691
Patella caerulea	Villamor et al., 2014	0,0931	0,1086	0,0978	0,9963
Patella rustica	Sá-Pinto et al., 2012	0,3169	0,3687	0,5000	0,5000
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,4824	0,5159	0,4150	0,8239
Phorcus turbinatus	Villamor et al., 2014	0,0377	0,1194	0,0620	0,5812
Ruditapes decussatus	Gharbi et al., 2011	0,5894	0,6285	0,2663	0,9823
Spondylus spinosus	Shabtay et al., 2014	0,3593	0,3562	0,4407	0,4481
Cymodocea nodosa	Alberto et al., 2008	0,0954	0,1485	0,1258	0,9989
Posidonia oceanica	Arnaud-Haond et al., 2007	0,1422	0,1704	0,0857	0,7641

Table SI-11: ***Significance (p-values) of Mantel correlations between multi-generation implicit dispersal*** **modelled Fst *and* observed Fst *being greater*** *than (i) correlation between IBD Euclidian distance and Fst/(1-Fst) (Silver et al., 2004), (ii) correlations between IBD sea least-cost distance and Fst/(1-Fst) (Silver et al., 2004), (iii) correlation between single-generation explicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970) and (iv) correlations between multi-generation explicit dispersal* modelled Fst *and* observed Fst *(Hendrickson et al., 1970).*

		Multi-generation implicit *vs*			
Species	**Study**	**IBD Euclidian distance**	**IBD sea least-cost distance**	**Single-generation explicit**	**Multi-generation explicit**
Cystoseira amentacea	Susini et al., 2007	0,5211	0,5072	0,3680	0,3093
Astroides calycularis	Casado-Amezua et al., 2012	0,8104	0,8034	0,3629	0,1757
Corallium rubrum	Aurelle et al., 2011	0,0392	0,0438	0,0236	0,3250
Corallium rubrum	Costantini et al., 2013	0,5847	0,5772	0,3775	0,2626
Eunicella cavolinii	Masmoudi et al., 2016	0,0697	0,0708	0,0292	0,3783
Leptopsammia pruvoti	Boscari et al., 2019	0,1803	0,4733	0,2234	0,4187
Botryllus schlosseri	Reem et al., 2017	0,5716	0,4487	0,2645	0,6896
Halocynthia papillosa	Villamor et al., 2014	0,4089	0,0366	0,2119	0,6115
Microcosmus squamiger	Ordóñez et al., 2013	0,3546	0,2344	0,5000	0,5000
Pycnoclavella communis	Pérez-Portela et al., 2007	0,3910	0,3699	0,4458	0,3927
Carcinus aestuarii	Schiavina et al., 2014	0,0057	0,0054	0,0005	0,6045
Melicertus kerathurus	Arculeo et al., 2010	0,3781	0,2004	0,4474	0,3742
Melicertus kerathurus	Zitari-Chatti et al., 2007	0,4307	0,3858	0,4093	0,2891
Pachygrapsus marmoratus	Fratini et al., 2013	0,8291	0,8236	0,2899	0,2107
Palinurus elephas	Palero et al., 2011	0,4783	0,4917	0,4040	0,3440
Spongia officinalis	Dailianis et al., 2011	0,7322	0,7614	0,1346	0,1808
Astropecten aranciacus	Zulliger et al., 2009	0,5780	0,6547	0,3128	0,4041
Holothuria mammata	Borrero-Pérez et al., 2011	0,3877	0,4898	0,1256	0,0698
Ophioderma longicauda	Weber et al., 2015[1]	0,3197	0,3299	0,2957	0,4969
Ophioderma longicauda	Weber et al., 2015[2]	0,1478	0,1611	0,2522	0,2866
Paracentrotus lividus	Penant et al., 2013	0,1151	0,1005	0,1019	0,1797
Paracentrotus lividus	Paterno et al., 2017	0,0581	0,0514	0,0238	0,9006
Paracentrotus lividus	Carreras et al., 2020[3]	0,7133	0,6972	0,3896	0,3021
Paracentrotus lividus	Carreras et al., 2020[4]	0,4263	0,4500	0,3928	0,2800
Apogon imberbis	Muths et al., 2015	0,6640	0,7058	NaN	0,5000
Coris julis	Fruciano et al., 2011	0,4861	0,3568	0,1186	0,0024
Diplodus sargus	González-Wangüemert et al., 2010	0,4707	0,4764	0,4037	0,0106
Diplodus vulgaris	Kaouèche et al., 2013	0,3723	0,3795	0,3386	0,5295
Epinephelus marginatus	Schunter et al., 2011	0,4894	0,5092	0,4298	0,5475
Lithognathus mormyrus	Hammami et al., 2007	0,2513	0,2097	0,3006	0,6596
Merluccius merluccius	Milano et al., 2014	0,8760	0,8239	0,3517	0,1036
Mugil cephalus	Durand et al., 2013	0,0905	0,0995	0,0552	0,0053
Mullus barbartus	Maggio et al., 2009	0,5283	0,3788	0,3783	0,1855
Mullus surmuletus	Galarza et al., 2009	0,3414	0,3377	0,5900	0,6856
Mullus surmuletus	Dalongeville et al., 2018	0,3073	0,3445	0,0575	0,0000
Oblada melanura	Gkafas et al., 2013	0,4944	0,5014	0,3169	0,1674
Oblada melanura	Calò et al., 2016	0,4181	0,4555	0,4273	0,3283
Pagellus erythrinus	Fassatoui et al., 2009	0,4004	0,2923	0,4133	0,3208
Serranus cabrilla	Schunter et al., 2011	0,7349	0,6954	0,5132	0,0009
Solea solea	Bahri-Sfar et al., 2011	0,0420	0,0491	0,0159	0,5168
Solea solea	Garoia et al., 2007	0,8767	0,8048	0,3815	0,4342
Sparus aurata	Franchini et al., 2012	0,2632	0,3396	0,4204	0,6371
Symphodus tinca	Carreras et al., 2017	0,4098	0,3214	0,3864	0,2363
Cerastoderma edule	Sromek et al., 2019	0,0612	0,0684	0,0286	0,0000
Chiton olivaceus	Villamor et al., 2014	0,3858	0,5306	0,2870	0,5136
Hexaplex trunculus	Villamor et al., 2014	0,2610	0,2617	0,1734	0,6215
Hexaplex trunculus	Marzouk et al., 2017[5]	0,2598	0,3440	0,1761	0,3109
Hexaplex trunculus	Marzouk et al., 2017[6]	0,2761	0,4653	0,1668	0,2573
Mytilus galloprovincialis	Diz and Presa, 2008	0,7534	0,7740	0,5847	0,6906
Ostrea edulis	Launey et al., 2002	0,3706	0,3711	0,4057	0,7309
Patella caerulea	Villamor et al., 2014	0,0353	0,0317	0,0235	0,0037
Patella rustica	Sá-Pinto et al., 2012	0,3169	0,3687	0,5000	0,5000
Patella ulyssiponensis	Sá-Pinto et al., 2012	0,4404	0,4756	0,3925	0,1761
Phorcus turbinatus	Villamor et al., 2014	0,0937	0,3047	0,1028	0,4188
Ruditapes decussatus	Gharbi et al., 2011	0,3834	0,3905	0,1801	0,0177
Spondylus spinosus	Shabtay et al., 2014	0,2795	0,2712	0,2867	0,5519
Cymodocea nodosa	Alberto et al., 2008	0,0203	0,0282	0,0300	0,0011
Posidonia oceanica	Arnaud-Haond et al., 2007	0,1390	0,1642	0,0823	0,2359

VII. Statistical analyses on the impact of species and study characteristics on Mantel correlations

Table SI-12: ***Sensitivity of the Mantel correlations between multi-generation implicit dispersal*** **modelled Fst** ***and*** **observed Fst** ***for the 58 population genetic studies included in the meta-analysis.*** *We test the sensitivity of the Species characteristics (taxa, PLD and spawning season) and study characteristics (Marker, Fst range, Number of populations sampled and Spatial Sampling Representativeness) on* ***a*** *R^2 and* ***b*** *p-values either with ANOVA or a linear regression (after a log10 transformation of Mantel R^2 and p-values).*

a		**Mantel R^2**			
	Characteristics	*Statistic method*	*p-value*	*R^2*	β *slope*
Species	Taxa	ANOVA	0.4397	NaN	NaN
	PLD	ANOVA	0.559	NaN	NaN
	Spawning season	ANOVA	0.2998	NaN	NaN
Study	Marker	ANOVA	0.1773	NaN	NaN
	Fst range	log regression	0.883	0.0011	-0.0516
	Nbr of populations	log regression	0.128	0.1068	-0.0104
	SSR	log regression	0.425	0.0306	-0.0001

b		**Mantel p-value**			
	Characteristics	Statistic method	p-value	R^2	β slope
Species	Taxa	ANOVA	0.0681	NaN	NaN
	PLD	ANOVA	0.7162	NaN	NaN
	Spawning season	ANOVA	0.4876	NaN	NaN
Study	Marker	ANOVA	0.6516	NaN	NaN
	Fst range	log regression	0.0001	0.2518	-7.2038
	Nbr of populations	log regression	0.0029	0.1476	-0.0495
	SSR	log regression	0.1181	0.0431	-0.0005