Isolation-by-Distance and Genetic Parentage Analysis Provide Strikingly Similar Dispersal Estimates in the Coral Reef Fish, *Elacatinus lori*

by

Taylor M. Naaykens

BSc. Hons, University of Manitoba, 2016

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in the Graduate Academic Unit of Biology

Supervisors: Cassidy D'Aloia, PhD, Department of Biological Sciences, UNB Saint John and Department of Biology, University of Toronto Mississauga

Examining Board: Scott Pavey, PhD, Department of Biological Sciences, UNB Saint John
Ben Speers-Roesch, PhD, Department of Biological Sciences, UNB Saint John
Oscar Puebla, PhD, Department of Ecology, Leibniz Centre for Tropical Marine Research & University of Oldenburg

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

April 2022

ABSTRACT

Larval exchange among marine populations is a vital driver of population dynamics and has the potential to inform conservation actions, but accurately measuring dispersal remains challenging. Here, I test whether accurate dispersal estimates can be obtained indirectly from an isolation-by-distance (IBD) model in the coral reef fish *Elacatinus lori* by comparing indirect estimates to direct measurements from genetic parentage analyses. Using the IBD approach, the spread of the dispersal kernel, σ, was estimated to be 2.9 – 4.1 km, remarkably similar to σ measured directly through genetic parentage analyses (σ = 3.9 km). Additionally, sensitivity analyses revealed that the IBD dispersal estimates were robust to genetic marker type and uncertainty in effective population size. Taken together, these findings suggest that accurate dispersal estimates can be produced by indirect IBD methods. The results indicate that this indirect—and more feasible—approach may be broadly applicable to the study of marine larval dispersal.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Cassidy D'Aloia for her mentorship and guidance throughout the past two years. I would also like to thank Dr. Scott Pavey and Dr. Ben Speers-Roesch as my supervisory committee for their advice and guidance in planning this project. I thank Diana Acosta, Jose Andrés, Steve Bogdanowicz, Eliana Bondra, Pete Buston, Alben David, Kevin David, Udell Foreman, Robin Francis, Rick Harrison, John Majoris, Alissa Rickborn, and Derek Scolaro for contributions to the data sets analyzed in this study. Pete Buston, Steve Bogdanowicz, Cynthia Riginos and Eric Crandall provided valuable feedback that improved this thesis. This research was supported by a NSERC Discovery Grant and NSF Grant #OCE-1260424, as well as Vaughan Graduate Fellowships.

Finally, I want to express my deep and sincere gratitude to my family and friends for their continuous love, help and support in encouraging me to strive for my goals. Thank you to Mario Solis for inspiring me, encouraging me, and for always being there to celebrate my successes and comfort me through my setbacks. You help me to recognize and appreciate my own personal growth and to see every challenge as an adventure. Thank you to Mom and Dad for supporting me every step of the way, reminding me always that I am loved, and giving me strength and courage to pursue my dreams. Thank you to Gord and Dici, for making me feel like part of the family and for all that you've done to make me truly feel at home. I treasure our friendship and am immeasurably grateful for your wonderful company and support throughout my graduate experience.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

AL: Adult lifespan.

CI: Confidence interval.

De: Effective population density

Fst: Fixation index

HWE: Hardy-Weinberg equilibrium

IBD: Isolation-by-distance

m: Slope of IBD relationship.

N_b: Effective number of breeders

N_e: Effective population size

P_{crit}: Critical value. The value below which rare alleles are excluded from the calculation of N_e.

S.D: Standard deviation

SNP: Single nucleotide polymorphism

α: Age of maturity.

σ: Sigma or standard deviation of dispersal

1. General Introduction

1.1: The evolutionary and ecological importance of dispersal

Natal dispersal is the movement of individuals away from their place of birth to another location for reproduction. From an evolutionary and ecological standpoint, the process of dispersal has several major consequences (Bonte et al., 2012; Matthysen, 2012). First, it reduces the likelihood of interacting with kin and, by extension, inbreeding (Szulkin & Sheldon, 2008). Second, it can increase the variance in fitness among offspring from the same parents if different offspring disperse to habitats of varying quality (Roff, 1974; Shaw et al., 2019). Third, dispersal can allow individual organisms to escape from unfavorable conditions by moving to higher quality areas to breed, or to take advantage of previously unexploited habitat, thereby increasing a species' range (Kokko & Lopez-Sepulcre, 2006). Finally, dispersal itself can impose significant costs to the individual (Bonte et al., 2012; Burgess et al., 2012). These can be the energetic costs associated with movement and the development and maintenance of locomotory structures, as well as mortality during or post dispersal.

Dispersal also drives gene flow among populations and therefore determines the degree of genetic and demographic connectivity between spatially distinct sites. Demographic connectivity refers to the cumulative effect of dispersing individuals on population growth, extinction, or recolonization, whereas genetic connectivity is concerned with the exchange of alleles, rather than individuals, and therefore limits consideration to only those individuals that successfully reproduce after dispersing. In turn, the degree of connectivity between populations influences population dynamics,

spatial patterns of genetic diversity, the rate of microevolution, and the resulting resilience of populations to changing environmental conditions (e.g.: disease, habitat fragmentation, climate change) (Huang et al., 2020; Kokko & Lopez-Sepulcre, 2006; Martensen et al., 2017).

1.2: Measuring dispersal

Measuring dispersal is a fundamental goal in ecology and evolutionary biology given the importance of this process in determining gene flow. Demographic tools for estimating dispersal include capture-recapture techniques and, to a lesser extent, telemetry tracking of individual movements (Cayuela et al., 2018; Hooten et al., 2017; Hussey et al., 2015; Royle et al., 2018; Shafer et al., 2016). These methods either directly track or infer the movements of individual organisms and are best suited for species that exhibit active dispersal (Cayuela et al., 2018). Genetic approaches to quantifying dispersal represent an alternative and have become increasingly accessible in recent decades (Broquet & Petit, 2009). These methods utilize tools such as summary statistics of allele frequencies, genetic assignment tests, and Bayesian inferences based on coalescent theory to estimate dispersal (Broquet & Petit, 2009). These suites of tools can be used separately or in combination to answer different evolutionary and ecological questions (Cayuela et al., 2018).

Once an approach is chosen, dispersal distances must be measured. For an individual dispersal event, the simplest approach is to measure the straight-line Euclidean distance (d) between the origin and destination locations (Figure 1A). At a fine spatial scale, this could represent the distance between the locations of a parent and offspring

pair, or, at a coarser spatial scale, the distance between an individual's birth and settlement populations or geographic neighborhoods. When dispersal is measured directly using demographic level techniques like mark-recapture, empiricists most often measure Euclidean distances. A Euclidean dispersal distance (*d*) can also be plotted on a Cartesian coordinate system (Figure 1B), where the two axial components of dispersal (*x*, *y*) can be visualized. The difference between axial and Euclidean dispersal distance is subtle but important for this thesis which draws on dispersal theory that uses axial distances (Rousset 1997; 2000; 2004). The directionality of the dispersal event relative to the origin location can also be accounted for, and this represents *signed* dispersal distances. For example, in Figure 1B, dispersal is shown in a positive direction, relative to the origin, whereas *unsigned* dispersal distance (Figure 1A) simply considers the absolute value of the dispersal distance.

Expanding beyond individual observations, if we take many measurements of distances between parent-offspring pairs or natal versus settlement locations, we can begin to characterize dispersal patterns at the population level. The simplest option is to plot a distribution of dispersal frequencies at varying distances (Figure 1C). Histograms constructed from these data provide useful qualitative depictions of dispersal patterns. Fitting a probability density function to the dispersal distance data produces a *dispersal kernel* (Figure 1D). Put simply, a dispersal kernel is a statistical distribution describing the probability that an individual born at a given location will successfully disperse to other locations at varying distances (Nathan et al., 2012). Dispersal kernels are important for informing population dynamics models. This information can also be used to test hypotheses about the likelihood of population expansion/contraction or

colonization/extinction events, and to evaluate the potential benefits of implementing marine protected area networks. Despite the utility and need for dispersal measurements, few empirical dispersal kernels have, to date, been measured for marine species.

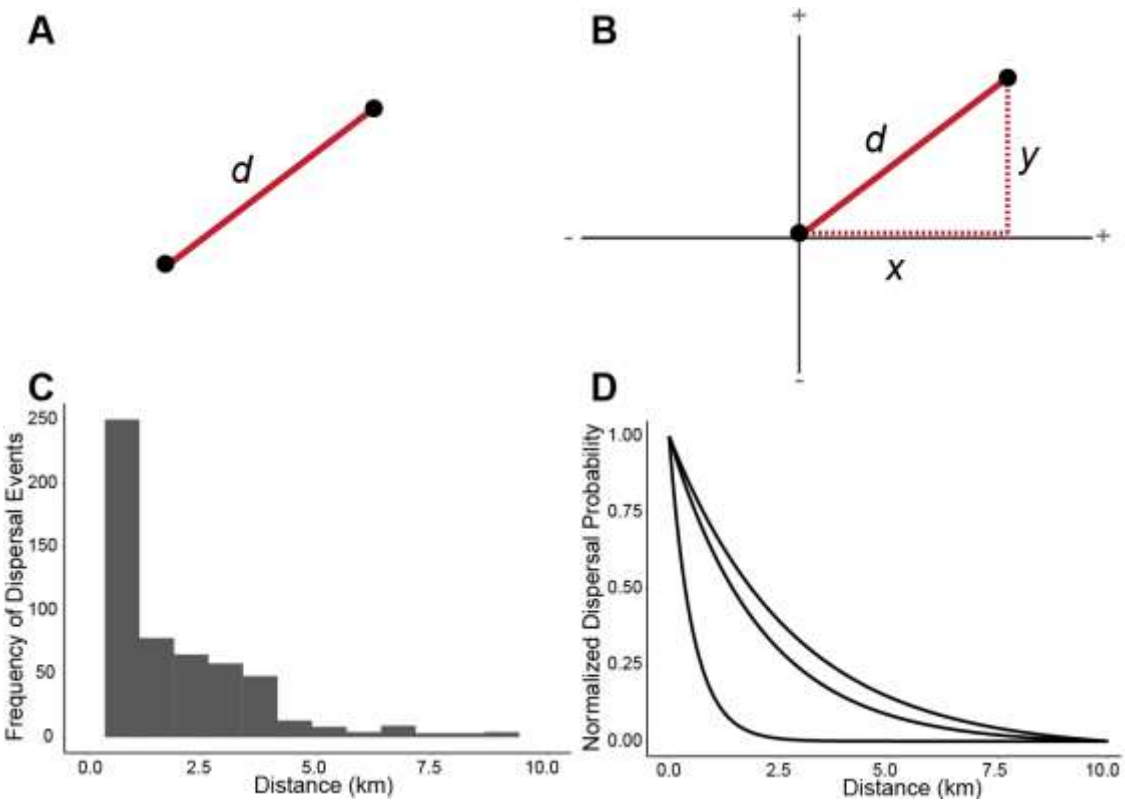

Figure 1. Measurements of dispersal. (A) Euclidean dispersal distance between two points; (B) That same dispersal event, with its signed axial (x, y) components; (C) Histogram of many dispersal events in a population; (D) Examples of alternative dispersal kernels that could represent the dispersal pattern in the population (here, with dispersal probability normalized from 0-1).

1.3: The challenge of quantifying marine larval dispersal

In the marine environment, the biology of organisms and the fluid medium through which they disperse pose significant challenges to quantifying dispersal patterns (Chan et al., 2018; Cowen & Sponaugle, 2009; Jones et al., 2009). Most benthic marine species exhibit some form of pelagic larval dispersal, whereby individuals develop and disperse in the open ocean before metamorphosing and settling back onto benthic habitat. These microscopic larvae are notoriously difficult to observe directly, making genetic methods invaluable tools for measuring or estimating dispersal in the ocean. It was previously thought that marine larvae were passively transported by ocean currents, which led to the expectation of ubiquitous long-distance dispersal through the pelagic environment and open marine populations (Caley et al., 1996; Roughgarden et al., 1988). Additionally, given that larvae can remain in the pelagic zone for several weeks, their capacity for long distance dispersal was generally predicted to increase with their pelagic larval duration (Scheltema, 1975; Thorson, 1950). This paradigm has since shifted to recognize that dispersal can be limited in the ocean, resulting in local retention and spatial genetic structuring in many marine populations (Almany et al., 2017; Hameed et al., 2016; Jones et al., 2005; Pinsky et al., 2017). Indeed, despite the fluid environment through which marine populations could theoretically be connected, both intrinsic and extrinsic factors represent barriers to long distance dispersal in marine species. Stretches of open ocean (or otherwise unsuitable habitat) and variable depths between habitat patches such as coral reefs or seagrass beds present significant barriers to dispersal (D'Aloia et al., 2014; Shlesinger & Loya, 2021). Additionally, evidence has accumulated to suggest that larvae of some species actively orient towards suitable habitat and can

maintain proximity to their birth site despite the presence of hydrodynamic forces that could otherwise transport them orders of magnitude further than observed (Faillettaz et al., 2018; Leis et al., 2014; Majoris et al., 2021).

Demographic tools that are commonly employed to directly measure dispersal in terrestrial species cannot be applied to dispersing marine larvae (Cayuela et al., 2018; Lebreton et al., 2003; Serrano et al., 2005). Therefore, genetic methods have emerged as valuable tools for estimating larval dispersal. Marine larval dispersal can reliably be measured using genetic parentage analysis. This method serves as the gold standard for directly measuring larval dispersal, as it captures actual dispersal events within a single generation. This method genotypes individuals at polymorphic loci and assigns parents to putative offspring from a pool of candidate parents. The geographic distance between assigned parent-offspring pairs is the net dispersal distance for the subset of correctly assigned offspring at the time and place of sampling. Repeating this for a large number of individuals creates the distributions over which dispersal kernels can then be applied (e.g Figure 1C-D). While this method can produce robust dispersal estimates (D'Aloia et al., 2015; Almany et al., 2017), it necessitates very large sample sizes which pose significant logistical limitations to its widespread application. Hence, there have been few directly measured marine larval dispersal kernels to date and the dispersal patterns of most marine species remain unknown.

Indirect estimates of dispersal provide a promising solution to address this knowledge gap and include inferences based off population genetics principals, as well as biophysical models (Pinsky et al., 2010; Puebla et al., 2012; Swearer et al., 2019). The latter method simulates dispersal of larvae based on their transport on ocean currents

(Counsell et al., 2022; Hawkins et al., 2019). Some biophysical models have been shown to overestimate dispersal distances, however with the incorporation of larval behavior data, these models could be a useful tool (Bode et al., 2019).

Alternatively, measures of spatial genetic structure in marine populations have been used to infer levels of connectivity among populations (De Wit et al., 2020; Fisher et al., 2022). Isolation-by-distance is one such pattern of spatial genetic structure, and it occurs when the genetic similarity of individuals decreases over geographic distance (Malécot, 1948; Wright, 1943). This pattern is generated when organisms tend to disperse only short distances, creating increased levels of genetic isolation between populations. This is a commonly observed pattern in natural populations (Benestan et al., 2021; Drinan et al., 2018; Hedgecock et al., 2007; Wright et al., 2015). When combined with estimates of effective population density, a metric that describes the effect of genetic drift acting on a given population, the standard deviation (or spread) of a dispersal kernel can be estimated (Rousset, 1997; Vekemans & Hardy, 2004). This method requires fewer samples than genetic parentage analysis, making it a promising tool for assessing dispersal patterns in marine species. However, due to the limited application of this method thus far, there is a need to assess its accuracy.

1.4: *Elacatinus lori* as a model species for studying marine dispersal

The genus *Elacatinus*, of the family Gobiidae, contains over 25 species and is the one of the most speciose genera of fish on Caribbean coral reefs (Nelson, 2006). *Elacatinus* spp. are small, microhabitat-specialist fishes that exhibit a classic biphasic life cycle. As adults, they live on species-specific corals or sponges and lay demersal eggs

within their respective microhabitats. Upon hatching, larvae enter a pelagic larval stage that lasts approximately three to four weeks (Colin, 1975, Majoris et al., 2018; Taylor & Hellberg, 2003), after which time they return to the reef and settle in association with their particular microhabitat. Species of *Elacatinus* can be subdivided into five distinct eco-morphological suites, two of which associate with stony corals and typically exhibit cleaning behaviors (differentiated by inferior vs sub-terminal mouth location), two associate with sponges at varying depths and do not exhibit cleaning behaviors, and one distinguishes its member species as hovering planktivores (Colin, 2010). Members of each suite occur throughout most of the tropical western North Atlantic, however individual species are distributed allopatrically and only a single member of each suite occurs in any area (Colin, 2010). Limited dispersal capacity has been suggested to contribute to the high degree of speciation observed within the genus (Taylor & Hellberg, 2005).

Elacatinus lori is an obligate sponge-dweller (Figure 2), found only in the Gulf of Honduras, occurring along the Mesoamerican barrier reef and the Bay Islands of Honduras (Colin, 2002). This species is characterized by a white lateral stripe and a thin white bar on its snout (Colin, 2002; Figure 2). *E. lori's* distribution is tightly linked with that of its preferred microhabitat sponge host, the yellow tube sponge *Aplysina fistularis* (D'Aloia et al., 2011). *E. lori* have evolved behavioral preferences for this species of sponge, as it maximizes their post-settlement persistence, relative to other microhabitat choices (Majoris et al., 2018). When settling on the benthos, they rely largely on visual cues to orient towards their preferred host sponges (Majoris et al., 2018). *E. lori* are equipped with all sensory organs upon hatching, which increase in size and complexity

during the larval phase (Majoris et al., 2021). Their sensory and locomotor capabilities as larvae are sufficient to allow *E. lori* to actively orient shortly after hatching (Majoris et al., 2021).

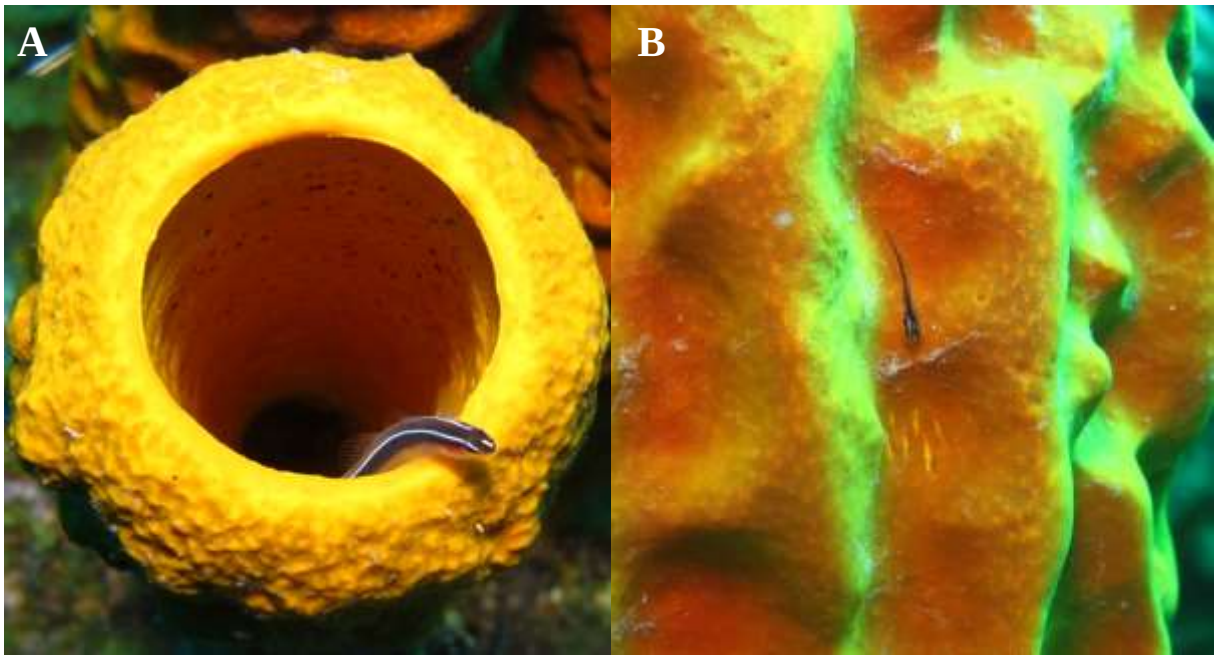

Figure 2. (A) Adult female *E. lori* inside its host sponge. (B) Post settlement *E. lori* on outside of host sponge. (Photo credit: Pete Buston).

Due to its association with a sessile and easily observable invertebrate host, *Elacatinus lori* has emerged as a tractable species to study marine larval dispersal patterns. A complete empirical dispersal kernel has been constructed for this species using direct parentage analysis based on 20 microsatellite loci (D'Aloia et al., 2015). The parentage kernel found strongly limited dispersal over a 41 km stretch of reef, with an average unsigned dispersal distance of just 2.8 km. Subsequently, at a larger spatial scale of 190 km, D'Aloia et al. (2020) uncovered an isolation-by-distance pattern along the

Belize barrier reef using three nuclear genetic marker panels. In this thesis, I will attempt to indirectly estimate a dispersal kernel for *E. lori* and reconcile this indirect estimate with the estimate generated from genetic parentage analysis. Specifically, I will use the slope of the isolation-by-distance relationship, along with an estimate of effective population size to indirectly estimate dispersal for *E. lori.* Few studies have applied the indirect isolation-by-distance method to explicitly estimate larval dispersal. While this method presents a promising approach to estimate dispersal for marine species, there is a need to further validate its accuracy with known dispersal measurements. This study will be the second known validation of an indirect IBD dispersal estimate with direct parentage dispersal data.

Bibliography

Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., Bonin, M.C., Frisch, A.J., Harrison, H.B., Messmer, V., Nanninga, G.B., Priest, M.A., Srinivasan, M., Sinclair-Taylor, T., Williamson, D.H., & Jones, G. P. (2017). Larval fish dispersal in a coral-reef seascape. *Nature Ecology & Evolution, 1(6)*, 1-7.

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., Mouillot, D., Puebla, O., & Manel, S. (2021). Restricted dispersal in a sea of gene flow. *Proceedings of the Royal Society B*, *288*(1951), 20210458.

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.

Bonte, D., Van Dyck, H., 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., Hovestadt, T., Meier, C.M., Palmer, S.C.F., Turlure, C., & Travis, J. M. (2012). Costs of dispersal. *Biological Reviews, 87*(2), 290-312.

Broquet, T., & Petit, E. J. (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics, 40*, 193-216.

Burgess, S. C., Treml, E. A., & Marshall, D. J. (2012). How do dispersal costs and habitat

selection influence realized population connectivity? *Ecology*, *93*(6), 1378-1387.

Caley, M.J., M.H. Carr, M.A. Hixon, T.P. Hughes, G.P. Jones, and B.A. Menge. (1996). Recruitment and the local dynamics of open marine populations. *Annual Review of Ecology and Systematics 27*, 477–500.

Cayuela, H., Rougemont, Q., Prunier, J. G., Moore, J. S., Clobert, J., Besnard, A., & Bernatchez, L. (2018). Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology*, *27*(20), 3976-4010.

Chan, K. Y. K., Sewell, M. A., & Byrne, M. (2018). Revisiting the larval dispersal black box in the Anthropocene. *ICES Journal of Marine Science*, *75*(6), 1841-1848.

Colin, P. L. (1975). *The neon gobies: the comparative biology of the gobies of the genus Gobiosoma, subgenus Elacatinus,(Pisces: Gobiidae) in the tropical western North Atlantic Ocean.* Hong Kong: TFH Publications.

Colin, P. L. (2002). A new species of sponge-dwelling Elacatinus (Pisces: Gobiidae) from the western Caribbean. *Zootaxa*, *106*(1), 1-7.

Colin, P. L. (2010). Fishes as living tracers of connectivity in the tropical western North Atlantic: I. Distribution of the neon gobies, genus Elacatinus (Pisces: Gobiidae). *Zootaxa*, *2370*(1), 36-52.

Counsell, C. W. W., Coleman, R. R., Lal, S. S., Bowen, B. W., Franklin, E. C., Neuheimer, A. B., Powell, B.S., Toonen, R. J., Donahue, M. J., Hixon, M. A., & McManus, M. A. (2022). Interdisciplinary analysis of larval dispersal for a coral reef fish: opening the black box. *Marine Ecology Progress Series, 684*, 117-132.

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

connectivity. *Annual Review of Marine Science, 1,* 443-466.

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. *Proceedings of the National Academy of Sciences, 112*(45), 13940-13945.

D'Aloia, C. C., Majoris, J. E., & Buston, P. M. (2011). Predictors of the distribution and abundance of a tube sponge and its resident goby. *Coral Reefs, 30*(3), 777-786.

D'Aloia, C. C., Bogdanowicz, S. M., Harrison, R. G., & Buston, P. M. (2014). Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish. *Molecular Ecology, 23*(12), 2902-2913.

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.

Drinan, D. P., Gruenthal, K. M., Canino, M. F., Lowry, D., Fisher, M. C., & Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (*Gadus macrocephalus*). *Evolutionary Applications*, *11*(8), 1448-1464.

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.

Fisher, M. C., Helser, T. E., Kang, S., Gwak, W., Canino, M. F., & Hauser, L. (2022). Genetic structure and dispersal in peripheral populations of a marine fish (Pacific

cod, Gadus macrocephalus) and their importance for adaptation to climate change. *Ecology and Evolution, 12*(1), e8474.

Hameed, S. O., White, J. W., Miller, S. H., Nickols, K. J., & Morgan, S. G. (2016). Inverse approach to estimating larval dispersal reveals limited population connectivity along 700 km of wave-swept open coast. *Proceedings of the Royal Society B: Biological Sciences, 283*(1833), 20160370.

Hawkins, S. J., Allcock, A. L., Bates, A. E., Firth, L. B., Smith, I. P., Swearer, S. E., & Todd, P. A. (2019). A review of biophysical models of marine larval dispersal. *Annual Reviews of Oceanography and Marine Biology, 57*, 325-356.

Hedgecock, D., Barber, P. H., & Edmands, S. (2007). Genetic approaches to measuring connectivity. *Oceanography, 20*(3), 70-79.

Hooten, M. B., Johnson, D. S., McClintock, B. T., & Morales, J. M. (2017*). Animal movement: statistical models for telemetry data.* Boca Raton: CRC press.

Huang, J. L., Andrello, M., Martensen, A. C., Saura, S., Liu, D. F., He, J. H., & Fortin, M. J. (2020). Importance of spatio–temporal connectivity to maintain species experiencing range shifts. *Ecography, 43*(4), 591-603.

Hussey, N. E., Kessel, S. T., Aarestrup, K., Cooke, S. J., Cowley, P. D., Fisk, A. T., Harcourt, R. G., Holland, K. N., Iverson, S. J., Kocik, J. F., Mills Flemming, J. E., & Whoriskey, F. G. (2015). Aquatic animal telemetry: a panoramic window into the underwater world. *Science, 348*(6240), 1255642.

Jones, G. P., Almany, G. R., Russ, G. R., Sale, P. F., Steneck, R. S., Van Oppen, M. J. H., & Willis, B. L. (2009). Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. *Coral Reefs, 28*(2),

307-325.

Jones, G. P., Planes, S., & Thorrold, S. R. (2005). Coral reef fish larvae settle close to home. *Current Biology, 15*(14), 1314-1318.

Kokko, H., & López-Sepulcre, A. (2006). From individual dispersal to species ranges: perspectives for a changing world. *Science, 313*(5788), 789-791.

Lebreton, J. D., Hines, J. E., Pradel, R., Nichols, J. D., & Spendelow, J. A. (2003). Estimation by capture-recapture of recruitment and dispersal over several sites. *Oikos, 101*(2), 253-264.

Leis, J. M., Paris, C. B., Irisson, J. O., Yerman, M. N., & Siebeck, U. E. (2014). Orientation of fish larvae in situ is consistent among locations, years and methods, but varies with time of day. *Marine Ecology Progress Series, 505*, 193-208.

Majoris, J. E., Foretich, M. A., Hu, Y., Nickles, K. R., Di Persia, C. L., Chaput, R., Schlatter, E., Webb, J.F., Paris, C.B., & Buston, P. M. (2021). An integrative investigation of sensory organ development and orientation behavior throughout the larval phase of a coral reef fish. *Scientific Reports, 11*(1), 1-13.

Majoris, J. E., Francisco, F. A., Atema, J., & Buston, P. M. (2018). Reproduction, early development, and larval rearing strategies for two sponge-dwelling neon gobies, *Elacatinus lori* and *E. colini*. *Aquaculture, 483*, 286-295.

Malécot, G. (1948). *Mathématiques de l'hérédité*. Paris: Masson et Cie.

Martensen, A. C., Saura, S., & Fortin, M. J. (2017). Spatio-temporal connectivity: assessing the amount of reachable habitat in dynamic landscapes. *Methods in Ecology and Evolution, 8*(10), 1253-1264.

Matthysen, E. (2012). Multicausality of dispersal: a review. In J. Clobert, M. Baguette,

T.G. Benton, & J.M. Bullock (Eds.), *Dispersal ecology and evolution* (pp. 3-18). Oxford, UK: Oxford University Press.

Nathan, R., Klein, E. K., Robledo-Arnuncio, J. J., & Revilla, E. (2012). Dispersal kernels: review. J. Clobert, M. Baguette, T.G. Benton, & J.M. Bullock (Eds.), *Dispersal ecology and evolution* (pp. 187-210). Oxford, UK: Oxford University Press.

Nelson, J. S. (2006). *Fishes of the World*. Hoboken, NJ: John Wiley & Sons, Inc.

Pinsky, M. L., Montes, Jr, H. R., & Palumbi, S. R. (2010). Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution, 64*(9), 2688-2700.

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.

Puebla, O., Bermingham, E., & McMillan, W. O. (2012). On the spatial scale of dispersal in coral reef fishes. *Molecular Ecology*, *21*(23), 5675-5688.

Roff, D. A. (1974). The analysis of a population model demonstrating the importance of dispersal in a heterogeneous environment. *Oecologia, 15*(3), 259-275.

Roughgarden, J., Gaines, S., & Possingham, H. (1988). Recruitment dynamics in complex life cycles. *Science, 241(4872)*, 1460-1466.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219-1228.

Rousset, F. (2000). Genetic differentiation between individuals. *Journal of Evolutionary*

Biology, 13(1), 58-62.

Rousset, F. (2004). *Genetic structure and selection in subdivided populations (Vol. 40).* Princeton, New Jersey: Princeton University Press.

Royle, J. A., Fuller, A. K., & Sutherland, C. (2018). Unifying population and landscape ecology with spatial capture–recapture. *Ecography, 41*(3), 444-456.

Scheltema, R. S. (1975). Relationship of larval dispersal, gene-flow and natural selection to geographic variation of benthic invertebrates in estuaries and along coastal regions. *Estuarine research, 1,* 372-391.

Serrano, D., Oro, D., Ursua, E., & Tella, J. L. (2005). Colony size selection determines adult survival and dispersal preferences: Allee effects in a colonial bird. *The American Naturalist, 166*(2), E22-E31.

Shafer, A. B., Northrup, J. M., Wikelski, M., Wittemyer, G., & Wolf, J. B. (2016). Forecasting ecological genomics: high-tech animal instrumentation meets high-throughput sequencing. *PLoS Biology, 14*(1), e1002350.

Shaw, A. K., D'Aloia, C. C., & Buston, P. M. (2019). The evolution of marine larval dispersal kernels in spatially structured habitats: analytical models, individual-based simulations, and comparisons with empirical estimates. *The American Naturalist, 193*(3), 424-435.

Shlesinger, T., & Loya, Y. (2021). Depth-dependent parental effects create invisible barriers to coral dispersal. *Communications Biology, 4*(1), 1-10.

Swearer, S. E., Treml, E. A., & Shima, J. S. (2019). A review of biophysical models of marine larval dispersal. *Oceanography and Marine Biology: An Annual Review, 57,* 325-356.

Szulkin, M., & Sheldon, B. C. (2008). Dispersal as a means of inbreeding avoidance in a wild bird population. *Proceedings of the Royal Society B: Biological Sciences, 275*(1635), 703-711.

Taylor, M. S., & Hellberg, M. E. (2003). Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science, 299*(5603), 107-109.

Taylor, M. S., & Hellberg, M. E. (2005). Marine radiations at small geographic scales: speciation in neotropical reef Gobies (*Elacatinus*). *Evolution, 59*(2), 374-385.

Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews, 25*(1), 1-45.

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology, 13*(4), 921-935.

Wright, D., Bishop, J. M., Matthee, C. A., & von der Heyden, S. (2015). Genetic isolation by distance reveals restricted dispersal across a range of life histories: implications for biodiversity conservation planning across highly variable marine environments. *Diversity and Distributions*, *21*(6), 698-710.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114- 138.

2. Isolation-by-distance and genetic parentage analysis provide strikingly similar larval dispersal estimates

2.1 Introduction

The dispersal of offspring away from their parents is a vital driver of both ecological and evolutionary dynamics (Andrade-Restrepo et al., 2019; Roughgarden et al., 1988). In most marine species, dispersal is undertaken during a planktonic larval phase, which is followed by a relatively sedentary adult phase (Cowen & Sponaugle, 2009). A species' full larval dispersal pattern can be quantified by a dispersal kernel, which is a statistical distribution describing the probability that an individual born at a given location will successfully disperse to other locations at varying distances (Nathan et al., 2012). However, larval movements are difficult to observe directly, which makes measuring dispersal distances challenging. Consequently, full empirical dispersal kernels have been measured for only a few marine species. Various genetic methods can help overcome this challenge by measuring dispersal either directly, by assigning individuals to their parents or birth population, or indirectly, by inferring dispersal rates from spatial genetic patterns (Hedgecock et al., 2007).

Genetic parentage analysis has emerged as a popular and reliable method for directly measuring larval dispersal patterns (Almany et al., 2017; Catalano et al., 2021; Harrison et al., 2020). At its core, this approach applies the laws of Mendelian inheritance at polymorphic loci to assign parents of progeny from a pool of candidate parents. This method uncovers actual dispersal events. If individuals are geolocated, the distance between the parent(s) and offspring is taken as the net dispersal distance. These data, in

turn, can be used to construct a dispersal kernel (Bode et al., 2018). While this method has made important contributions to our understanding of the scale of marine dispersal, it is labor intensive and costly, rendering it difficult to apply broadly. Thus, parentage analysis has primarily been applied to coral reef fishes with relatively limited dispersal (e.g., D'Aloia et al., 2015; Jones et al., 2005). Another drawback of parentage analysis is that it captures a snapshot of dispersal probabilities at the time and place of sampling, and therefore may not capture temporal and spatial variability in dispersal patterns (but see Catalano et al., 2021). Given these constraints, there is clearly a need to evaluate the usefulness of other complementary methods for estimating marine larval dispersal (Manel et al., 2019; Pinsky et al., 2017).

Indirect genetic methods that use population genetic principles represent an alternative to parentage analysis with fewer logistical limitations, but there is a need to assess how reliably they quantify contemporary larval dispersal patterns. The isolation-by-distance (IBD) method is one promising indirect approach for drawing inferences about the scale of dispersal (Rousset, 1997; Vekemans & Hardy, 2004). An IBD pattern occurs when genetic similarity decreases between individuals or populations as the distance between them increases (Malécot, 1948; Wright, 1943). Patterns of IBD have been documented in diverse marine populations (Selkoe et al., 2016) including oysters (Lazoski et al., 2011), sea urchins (Wright et al., 2015), and several fishes (Ackiss et al., 2018; Benestan et al., 2021; Drinan et al., 2018).

IBD models predict that the decline in genetic similarity with distance is a function of both genetic drift and limited dispersal (Malecot, 1948; Rousset, 1997; Wright, 1943). These processes can be summarized by two key parameters: effective

population density (D_e) and the mean squared parent-offspring axial dispersal distance (σ^2) (Rousset, 1997). Thus, if measurements of IBD and effective population density are obtained, inferences about dispersal can be drawn. Notably, this indirect IBD approach can be applied at either the individual or population level (Vekemans & Hardy, 2004), and across both 1- and 2- dimensional habitats (Rousset, 1997). An advantage to this indirect approach is that it averages dispersal across generations, rather than taking a single spatio-temporal snapshot, and therefore represents mean dispersal patterns over a longer timescale. Moreover, because this method requires fewer samples, it is more feasible to apply this approach to a variety of species. The indirect IBD approach has been used to quantify dispersal in various terrestrial species (Broquet et al., 2006; Oddou-Moratio et al., 2010; Sumner et al., 2001), and the prevalence of IBD patterns in marine populations suggests that this indirect approach may offer a useful method for estimating dispersal in the ocean. Yet, few studies have applied this method to marine species. Limited examples have estimated the dispersal of anemonefishes (Pinsky et al., 2010; Pinsky et al., 2017) and several other coral reef fishes (Puebla et al., 2009; Puebla et al., 2012).

One major challenge in applying the indirect IBD method is obtaining accurate estimates of effective population density (D_e). Effective density measures the influence of genetic drift that is acting on the population. Most marine populations have large effective population sizes (N_e) (Marandel et al., 2019) and simulations show that the accuracy of N_e estimates declines as true N_e increases (Waples, 2016). High-throughput genetic data sets may help overcome this challenge, as large marker panels and sample sizes can improve N_e estimates (Marandel et al., 2019; Waples & Do, 2010). If accurate

N_e estimates can be obtained, the indirect IBD approach to estimating dispersal has the potential to be widely applied once its reliability is empirically evaluated. To date, only one study has explicitly compared a direct parentage analysis and an indirect IBD analysis in a marine species. Pinsky et al. (2017) observed similar estimates between these two methods in the clown anemonefish *Amphiprion percula*, a species with a relatively short larval duration (10-12 days; Almany et al., 2007) and a small effective population size. An average IBD dispersal distance of 12.1 km was estimated for *A. percula*, compared to 18.9 km and 13.3 km for average direct dispersal distances measured in 2009 and 2011, respectively. More comparative work is needed to assess the accuracy of the IBD method for a broader range of marine species that differ in key life history traits and population parameters.

Elacatinus lori, a sponge-dwelling goby endemic to the Mesoamerican Reef (Colin, 2002), is a good study species for the comparison of indirect and direct dispersal estimates. The larval stage of *E. lori* is approximately four weeks, after which time surviving larvae locate and settle on the outside of a sponge (Majoris et al., 2018). A complete empirical dispersal kernel has been constructed for this species using direct parentage analysis (D'Aloia et al., 2015). The parentage kernel found strongly limited dispersal over a 41 km stretch of reef, with an average unsigned dispersal distance of just 2.8 km. Patterns of population structure have also been described throughout the species' range, including a pattern of IBD along the Belizean barrier reef (D'Aloia et al., 2020). However, the indirect dispersal method has not yet been applied. In this study, I generate the first indirect IBD-based dispersal estimates for *E. lori* and evaluate them against direct parentage-based dispersal measurements.

2.2 Materials and Methods

2.2.1 Direct parentage estimate of σ

Direct and indirect dispersal estimates can be compared through σ^2, defined as the mean squared axial dispersal distance, or, the second noncentral moment of the dispersal distribution. To evaluate dispersal estimates obtained from the indirect IBD method, I first calculated σ^2 for the direct parentage data. Because these data are Euclidean distances *d* along a linear habitat (D'Aloia et al. 2015), they represent unsigned distances along a single axis. Following Rousset (1997, 2004) and Broquet & Petit (2009), σ^2 for Euclidean distance = $\mathbb{V}(d) = \mathbb{E}(d)^2 + \mathbb{E}(d^2)$. Given a symmetric dispersal distribution, $\mathbb{E}(d) = 0$ and σ^2 reduces to $\mathbb{E}(d^2)$. For the observed *E. lori* parentage data, $\sigma^2 = 15.42$ km. Thus, σ, which can be interpreted as the standard deviation, or spread, of the dispersal distribution, is 3.93 km. Having established this basis of comparison, I calculated σ indirectly via the IBD method, largely following the approach described by Pinsky et al. (2017).

2.2.2 IBD

To calculate the slope of the IBD regression (*m*), I used genotype data collected from adult *E. lori* in 2014. While previous research detected an IBD pattern using correlative Mantel tests (D'Aloia et al., 2020), here I re-analyzed these data because the indirect method requires the regression slope. Although previous descriptions of population structure for this species included sampling sites spanning the entire Belize reef system, in the present study I focused on the subset of sampling locations along the Belize Barrier Reef (Figure 3; 590 individuals total; $\bar{x} = 31 \pm 6$ S.D. individuals per site)

for two reasons. First, the previous parentage-based dispersal study was conducted along the barrier reef, making our IBD dispersal estimates directly comparable to the parentage estimates. Second, the linear structure of the barrier reef conforms to the classical 1-dimensional formulation for IBD dispersal models (Rousset, 1997).

My primary IBD analysis was based on adult genotype data for a genetic marker set of 2,418 single nucleotide polymorphisms (SNPs). To assess whether the strength of IBD—and its subsequent impact on dispersal estimates—was influenced by marker type, I also calculated the IBD relationship for two additional genetic marker panels: 82 microsatellites and 57 nonrepetitive nuclear loci. All marker panels were previously filtered for HWE and linkage disequilibrium (D'Aloia et al., 2020). For each marker panel, I calculated pairwise F_{ST} (Weir & Cockerham, 1984) and performed a linear regression to obtain the slope (m) of the relationship between linearized F_{ST} ($F_{ST}/(1-F_{ST})$) and geographic distance.

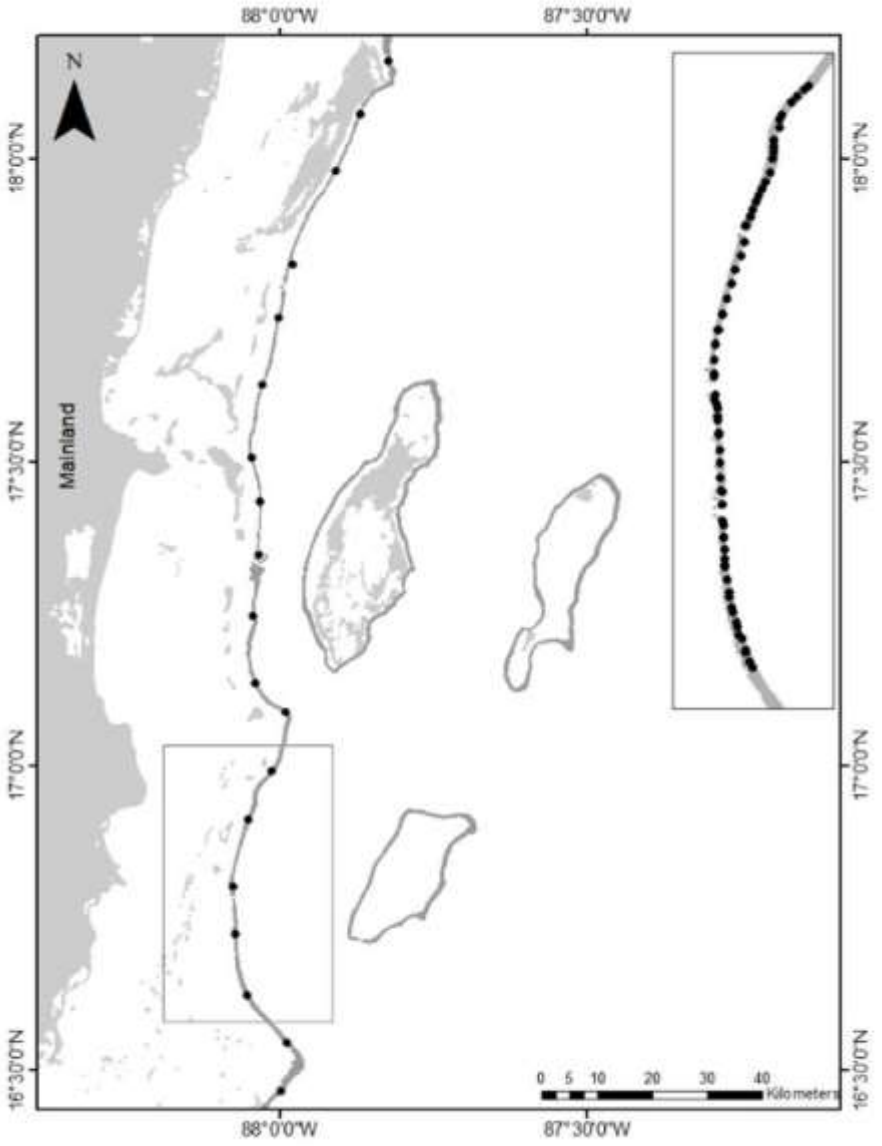

Figure 3. Sampling locations on the Belize Barrier Reef. Sites along the full barrier reef mark the locations of each adult *E. lori* collection site (D'Aloia et al., 2020). Sites indicated in the inset map mark the sampling sites for settler *E. lori* along a 41 km stretch of the barrier reef (D'Aloia et al., 2015).

2.2.3 Effective size and density estimates

To estimate effective size, I used genotype data collected from recently settled *E. lori* in 2013 (D'Aloia et al., 2015). These settlers were collected along a 41 km stretch of the Belize Barrier Reef (Figure 3 inset; ~ 100 individuals collected every km). All individuals were previously genotyped at a panel of 63 microsatellites (D'Aloia et al., 2018). A total of 4,074 individuals were used in this data set. I used these young individuals for effective size estimates because they represent a single cohort, which is an assumption of single-sample, linkage disequilibrium-based estimates of N_b (Waples & Do, 2010), where N_b reflects the number of breeders in the parent generation that produced the cohort (Waples, 2005).

Estimating N_b in continuously distributed populations that exhibit IBD is challenging. In particular, for the linkage disequilibrium method, when a species exhibits localized breeding, sampling within a continuously distributed population can mix genetically divergent individuals within the sample. This creates mixture linkage disequilibrium and can reduce empirical estimates of N_b depending on the spatial scale of sampling (Neel et al., 2013). Simulations suggest that while mixture linkage disequilibrium prevents estimates of a true global population-wide N_b, if the sampling scale is restricted to an area where potential parents are not genetically differentiated, the effective size estimate approximates the genetic neighborhood size (Neel et al. 2013). Thus, I subsampled settlers at approximately the spatial scale of the genetic neighborhood in 1- dimension, where neighborhood length is equal to $2\sqrt{\pi}\sigma$ (Wright, 1946). Using the parentage-based sigma estimate (σ = 3.93 km), I obtained a neighborhood length of 13.93 km. To estimate N_b, I applied a sliding window approach, moving down the 41-km settler

sampling area in windows of the neighborhood length, leading to a total of 28 overlapping windows.

I calculated N_b in each neighborhood window using the linkage disequilibrium method implemented in NeEstimator v 2.1 (Do et al., 2014). The linkage disequilibrium model requires an assumption about the species' mating system, with the two options being random mating or lifetime monogamy (following random mate selection). The mating system strongly influences effective size estimates, but the full genetic mating system of *E. lori* is not yet characterized for both sexes. To account for this uncertainty in the mating system and assess its influence on downstream dispersal estimates, I estimated N_b under both random and monogamous mating assumptions. I used a critical value of interest (P_{crit}: the value below which rare alleles are excluded from the calculation) of 0.02, as suggested by Waples & Do (2010). To obtain a point estimate of N_b, I took the weighted harmonic mean of the N_b estimates from the 28 neighborhood windows, where the weights were the effective degrees of freedom, and recorded the 95% jackknife confidence intervals.

Because *E. lori* has overlapping generations, I applied a bias adjustment to convert N_b to N_e, drawing on work which demonstrates that two life history traits—age at maturity (α) and adult lifespan (AL) —can be used to translate N_b to N_e (Waples et al., 2013), with an adjustment for age-structure bias in iteroparous species (Waples et al., 2014). Gobiids generally exhibit early maturation and short lifespans (Brandl et al., 2018; Herler et al., 2011) yet there are no published estimates of age at maturation or adult lifespan for *Elacatinus lori* or other *Elacatinus* species. Because congeners such as *E.*

oceanops are widely reared in the aquarium trade, I used aquarist accounts of α = 0.5 years and AL = 2 years as best estimates for these parameters.

I began by calculating a point estimate of adjusted N_b for *E. lori* using age at maturity (α) and adult lifespan (AL) following equations in Table 3 of Waples et al. (2014):

$$\widehat{N}_{b(Adj)} = \frac{\widehat{N}_b}{1.103-0.245 \times log(AL/\alpha)}, \quad \text{(eq. 1)}$$

Next, I calculated adjusted effective population size:

$$\widehat{N}_{e(Adj)} = \frac{\widehat{N}_{b(Adj)}}{0.485+0.758 \times log(AL/\alpha)}, \quad \text{(eq. 2)}$$

I divided $\widehat{N}_{e(Adj)}$ by the spatial scale of sampling (i.e., the neighborhood length) to obtain an effective density point estimate D_e. I assumed that this D_e was representative of the broader region over which IBD was estimated. To assess the impact of this correction on the *E. lori* D_e estimate, I also calculated D_e without implementing the life-history correction by using the uncorrected N_b as a proxy for N_e and dividing by the neighborhood length. The corrected values, however, were ultimately carried through to subsequent dispersal estimates.

While I had parentage data to inform my estimate of neighborhood size for *E. lori*, I acknowledge that such data would not be available for most species. To mimic a scenario in which I did not have these data, I also generated a single point estimate of N_b

using all 4,074 sampled individuals. I applied the same life history correction to translate N_b to N_e, and then divided by the total sampling area (41 km) to obtain D_e. I compared this D_e estimate to the D_e estimate measured at the neighborhood scale to assess the magnitude of their difference and its influence on dispersal estimates.

2.2.4 Indirect IBD estimate of σ

Using my estimates of IBD slope (*m*) and effective density (D_e), I calculated dispersal indirectly using the following equation for populations occupying 1-dimensional habitat (Rousset, 1997):

$$\sigma = \sqrt{\frac{1}{4D_e m}} \qquad \text{(eq. 3)}$$

where σ can be interpreted as the spread, or standard deviation, of the dispersal kernel. I compared this indirect σ estimate with the direct parentage-based estimate by calculating the percent change in σ.

2.2.5 Dealing with parameter uncertainty

Following the protocol described by Pinsky et al. (2017), I propagated uncertainty in my estimates for each parameter by resampling from probability distributions. For *m*, I took 100,000 draws from a normal distribution, with the mean and error taken from the output of the IBD regression. For N_b, I drew from a chi-squared distribution (following Waples & Do, 2010), parameterized with my point estimate of N_b and using the degrees of freedom from a chi-squared distribution with confidence intervals matching those

obtained from NeEstimator. Finally, to account for uncertainty in the life history parameter estimates used for the N_b/N_e correction, I generated distributions for AL and α. I used truncated normal distributions to bound the distributions with plausible values (i.e., to preclude negative values and values that exceeded aquarium observations): AL mean = 2, S.D. =0.5, range of values = 1-3; α mean = 0.5, S.D. = 0.1, range of values = 0.3-1. I took two draws from these distributions (one draw each for eqs. 1 and 2). I carried this uncertainty through eq. 3 to produce a distribution of σ values and recorded the 95% confidence intervals.

2.2.6 Visualizing dispersal kernels

To visualize dispersal kernels with my indirect σ estimates, I first assumed a Laplace distribution. Because this distribution was the best-fit functional form to the parentage data (D'Aloia et al., 2015), it is an appropriate assumption for the *E. lori* IBD-based kernels. Several other marine species with a pelagic larval phase exhibit dispersal patterns that conform to this distribution (Almany et al., 2017; Pinsky et al., 2017). I plotted indirect Laplace kernels for each effective size estimate: one for random mating and one for monogamous mating. I also estimated mean unsigned dispersal distance, as it is a useful metric for summarizing dispersal patterns. Under a Laplace distribution, the mean of unsigned dispersal distances is equal to $\frac{\sigma}{\sqrt{2}}$.

The functional form of the dispersal kernel influences the calculation of mean dispersal distance and its associated confidence intervals. While a Laplace distribution may be more appropriate for *E. lori* given what is already known about its dispersal pattern, for species without a priori dispersal data, it may be difficult to ascertain the most

appropriate distribution. A normal distribution may often be assumed by default. For this reason, I examined the effect that the assumption of a normal distribution would have on estimated mean dispersal distances and dispersal kernel shape. For a normal distribution the mean unsigned distance is equal to $\sigma\sqrt{2/\pi}$.

2.3 Results

2.3.1 IBD

The linear regressions for the adult goby data detected subtle but significant isolation by distance patterns for each genetic marker panel, and similar slopes were produced (Figure 4). The SNP dataset exhibited the strongest relationship between geographic and genetic distance while the regression of the nonrepetitive dataset had the poorest fit. Overall, these IBD regression results, which are needed to obtain the slope for the indirect dispersal estimates (eq. 3), are consistent with previously published correlative analyses of IBD, conducted in this same region, that control for spatial autocorrelation (D'Aloia et al., 2020). Unless otherwise stated, subsequent dispersal calculations were made based on the SNP slope values (Figure 4; m = 1.981e-05, SE = 1.601e-06, $p < 0.001$, $R^2 = 0.475$).

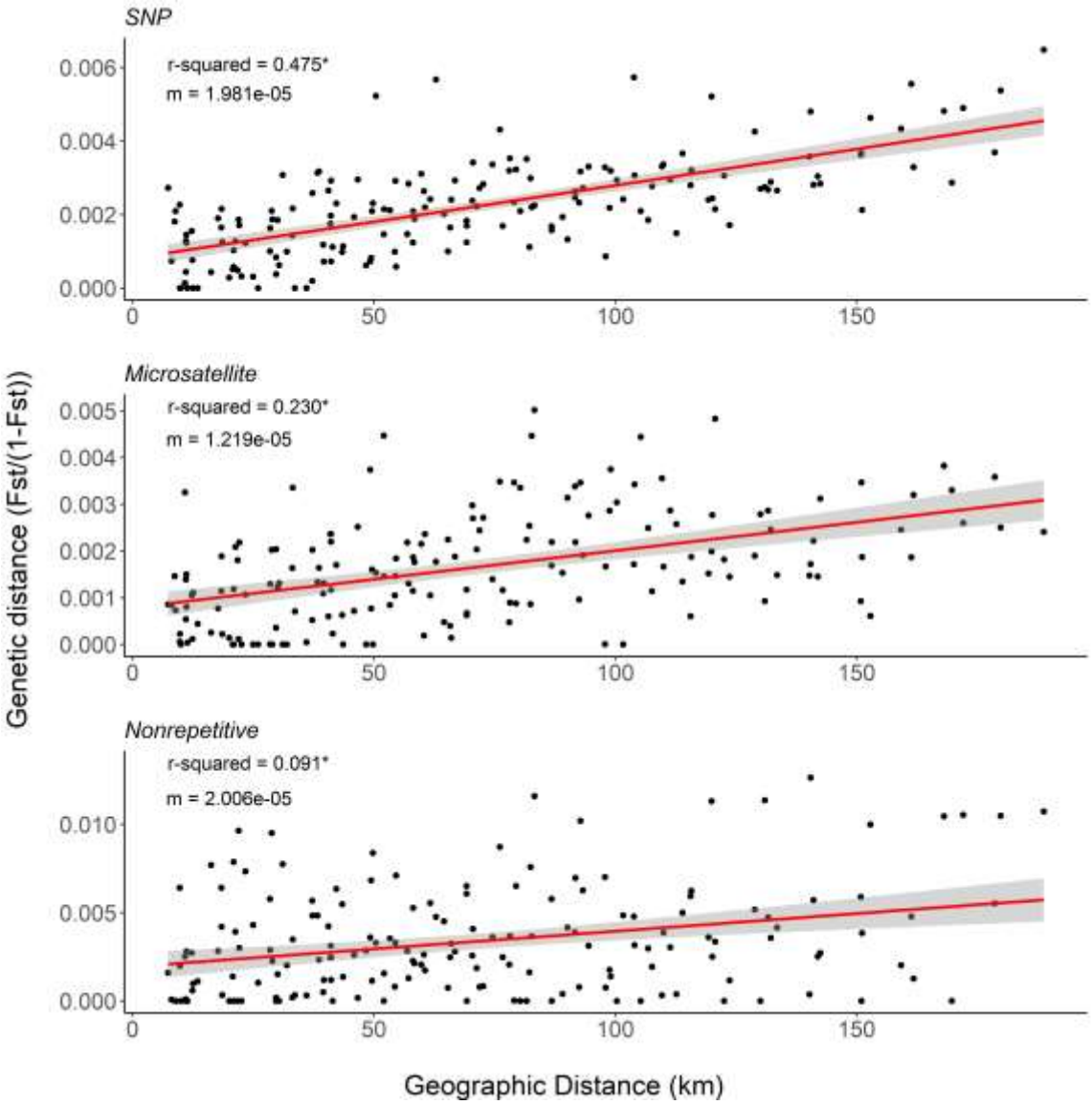

Figure 4. Linear regressions demonstrating isolation by distance trends based on three genetic data sets to calculate F_{ST}. The grey shading around each line represents a 95% confidence interval. Regression slope indicated by m; asterisk indicates p-value < 0.001. Note: the y-axis varies across panels.

2.3.2 Effective size and density estimates

At the spatial scale of the genetic neighborhood, the *E. lori* settler data produced a point estimate of N_b (random) = 9,825 individuals (95% CI: 7,091-14,321) and N_b (monogamous) = 19,652 individuals (95% CI: 13,929-29,504). As expected, the assumption of monogamy roughly doubles the N_b estimate relative to random mating. When converted to N_e using the life history correction, the local effective population size was 10,437 individuals (95% CI: 7,225 - 16,398) and 20,876 (95% CI: 14,171 - 33,586) under random and monogamous mating, respectively. Dividing adjusted N_e by the neighborhood length of 13.93 km yielded D_e = 749 (95% CI: 519 – 1,177) under random mating and D_e = 1,499 (95% CI: 1,017- 2,411) under monogamous mating. Finally, I found that the bias correction for age structure did not have a substantive effect on D_e estimates for this species, as there was only a ~10% reduction compared to the uncorrected values (Table 1).

Table 1. A comparison of D_e estimates with and without N_e life history correction for the SNP genetic marker panel.

	Random Mating	Monogamous Mating
D_e with correction	749	1,499
D_e no correction	674	1,348

2.3.3 Indirect dispersal estimates

The indirect dispersal spread estimates (σ) under assumptions of random and monogamous mating were both similar to the direct parentage-based dispersal estimate of σ = 3.93 km (95% CI: 3.29 - 4.71), with sigma values of σ (random) = 4.10 km (95% CI: 3.23 – 5.03) and σ (monogamous) = 2.90 km (95% CI: 2.26 – 3.59) (Figure 5). The percent change in standard deviation (σ) compared to the parentage estimate was 4.5% and 26.1% for the random and monogamous estimates, respectively. While the percent difference in standard deviation is greater under the assumption of monogamous mating compared to random mating, this large relative decrease is influenced by the small value of σ in the direct parentage data; the absolute difference in σ was just 1.03 km. Finally, I note that although these sigma values are based on the effective density estimates that are corrected by life history parameters (Table 1; drawing on eq. 1 and 2), using the uncorrected values translated to a small difference in sigma (~5% difference).

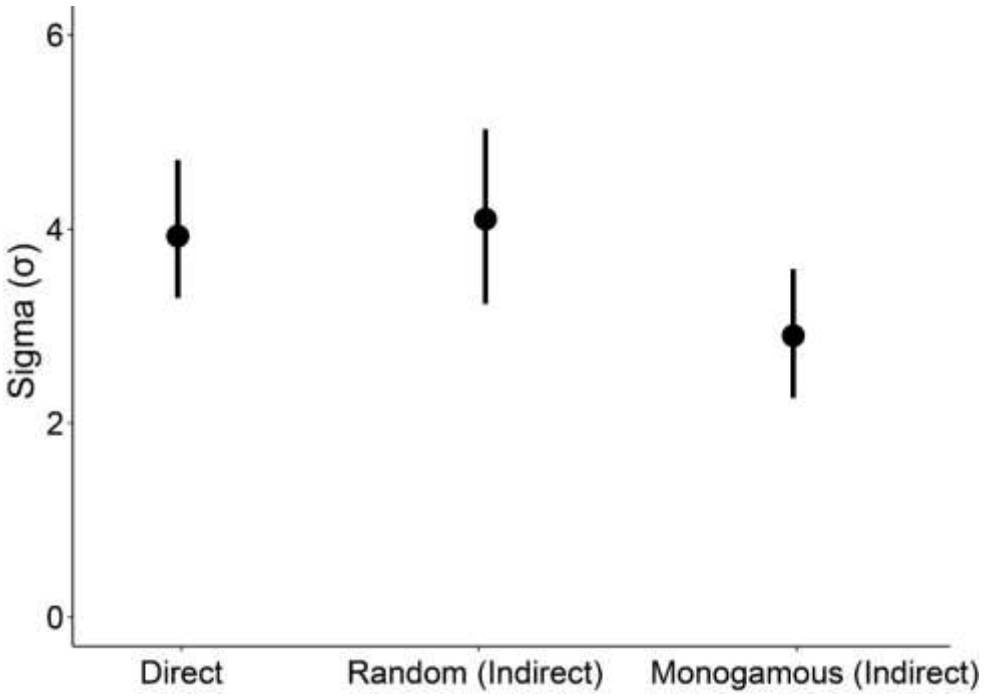

Figure 5. Comparison of direct and indirect sigma estimates assuming both random and monogamous mating in *E. lori*. Lines represent 95% confidence intervals.

Assuming a Laplace distribution, I found that the indirect IBD kernels are similar to each other, as well as to the parentage kernel, with considerable overlap of confidence intervals (Figure 6A). The direct parentage dispersal data have a mean unsigned dispersal distance of 2.8 km (95% CI: 2.3-3.3 km). I estimated mean unsigned dispersal distances of 2.9 km (95% CI: 2.3 - 3.6 km) and 2.1 km (95% CI: 1.6 - 2.5 km) for Laplace kernels generated from D_e estimates with the assumption of random and monogamous mating, respectively (Figure 6B). In summary, key assumptions about the mating system of *E. lori* did not ultimately have a strong effect on the indirect dispersal estimates.

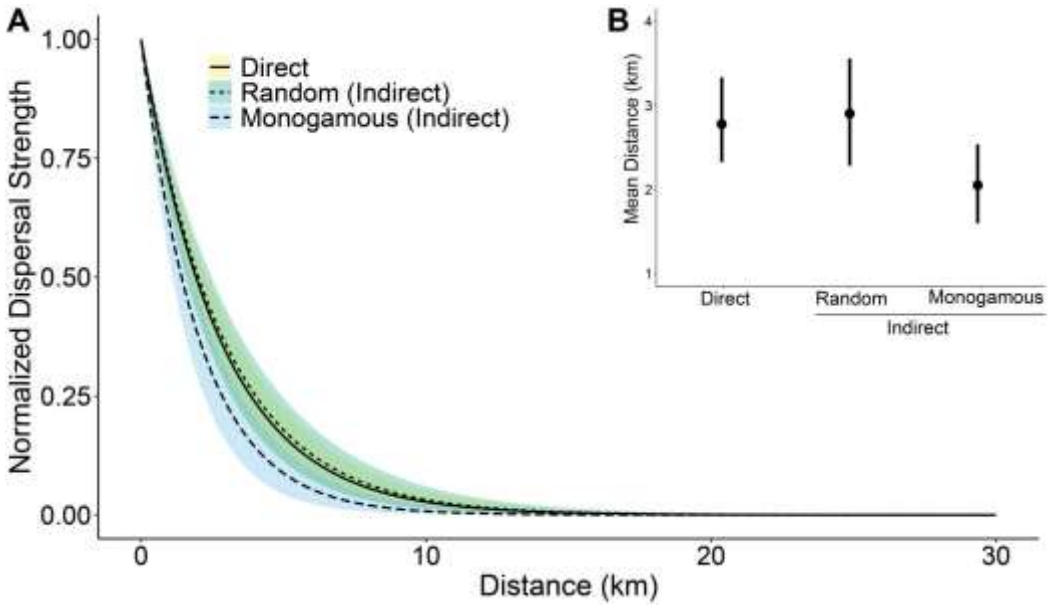

Figure 6. Comparison of direct and indirect dispersal estimates, assuming a Laplace distribution and considering both monogamous and random mating when calculating De. (A) Laplace dispersal kernels; Inset (B) mean unsigned dispersal distances with 95% confidence intervals.

2.3.4 Sensitivity analysis: Effect of genetic marker panel on dispersal estimates

The indirect kernels based on the three IBD genetic data sets are remarkably similar to each other, as well as to the parentage kernel, with considerable overlap of confidence intervals (Figure 7). Assuming a Laplace distribution and random mating, I estimated mean unsigned dispersal distances of 2.9 km (95% CI: 2.3 - 3.6 km), 3.7 km (95% CI: 2.9-4.7 km), and 2.9 km (95% CI: 2.1-4.2 km) for SNPs, microsatellites, and nonrepetitive loci, respectively. If monogamous mating is assumed, mean unsigned dispersal distances are 2.1 km (95% CI: 1.6 – 2.5 km), 2.6 km (95% CI: 2.0-3.4 km), and 2.0 km (95% CI: 1.5-2.9 km) for SNPs, microsatellites, and nonrepetitive loci,

respectively. This sensitivity analysis reveals that the indirect IBD dispersal estimates are robust to the choice of the genetic marker panel used to detect IBD.

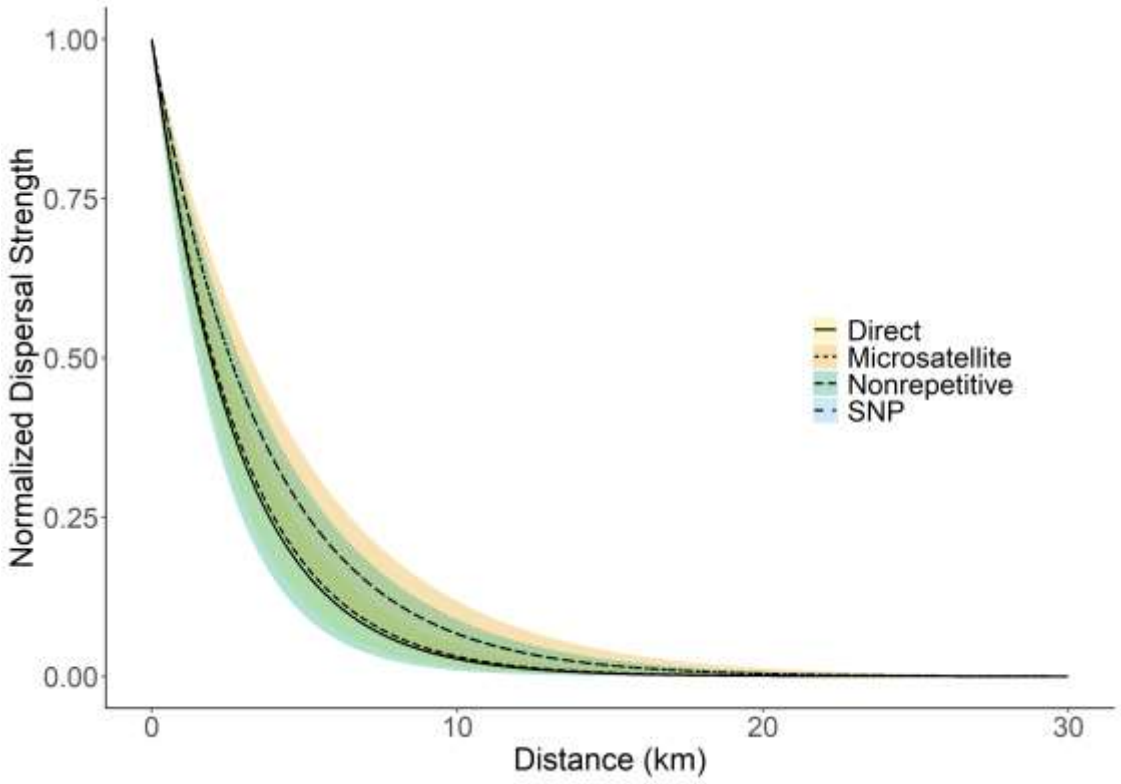

Figure 7. Laplace dispersal kernels visualized for all three genetic data sets, compared to the direct parentage kernel. For all indirect kernels, random mating is assumed.

2.3.5 Sensitivity analysis: Effect of kernel distribution assumption

When a normal distribution is assumed, in lieu of a Laplace distribution, indirect random and monogamous dispersal kernels are still similar to each other, as well as to the direct parentage kernel (Figure 8). Mean dispersal distances assuming a normal distribution are 3.3 km (95% CI: 2.6 – 4.0) and 2.3 km (95% CI: 1.8 – 2.9) for random and monogamous mating systems, respectively. This sensitivity analysis reveals that the

indirect IBD dispersal estimates are qualitatively consistent when using either a Laplace or a normal distribution.

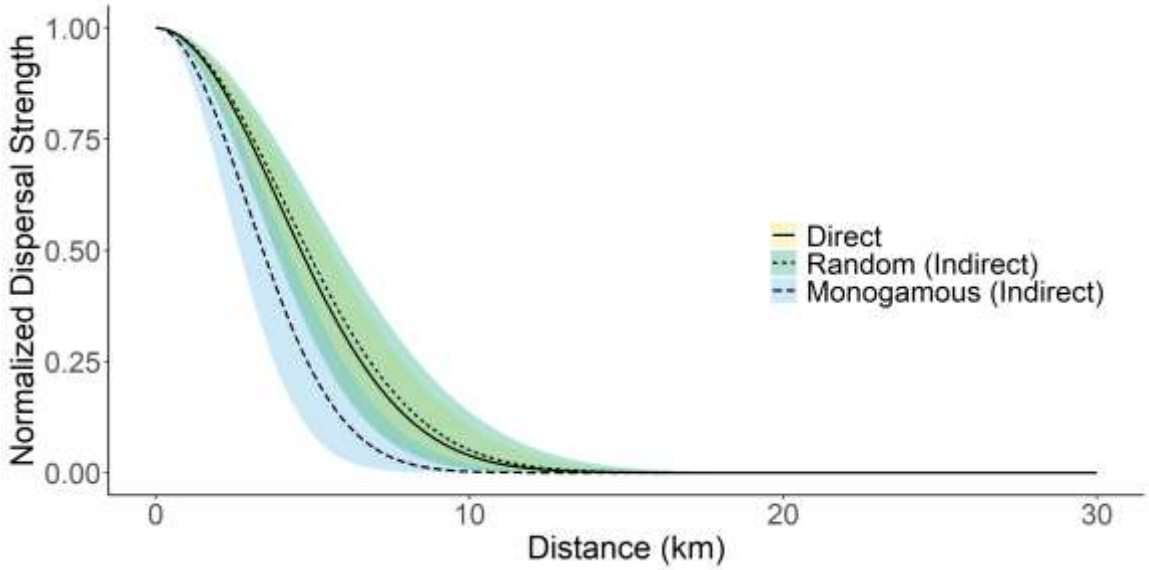

Figure 8. Comparison of direct and indirect dispersal estimates, assuming a normal distribution, using the SNP IBD slope, and considering both monogamous and random mating when calculating D_e.

2.3.6 Sensitivity analysis: Estimating D_e without prior data

To simulate circumstances where no prior information about neighborhood length is available, I ran NeEstimator using all settlers within the 41-km sampling area for a single point estimate. This produced a N_b (random) of 18,419 individuals (95% CI: 15,372 - 22,262) and N_b (monogamous) of 36,840 individuals (95% CI: 30,745 - 44,525). After applying the life history correction and dividing by the length of the entire sampling area, I obtained effective densities of 477 (95% CI: 371 - 652) and 955 (95% CI: 742 - 1,305)

under random and monogamous mating, respectively. The resulting sigma values were σ (random) = 5.14 km (95% CI: 4.33 - 5.98) and σ (monogamous) = 3.64 km (95% CI: 3.06 - 4.23). This sensitivity analysis demonstrates that, if the appropriate scale of observation is not considered, mixture linkage disequilibrium in continuously distributed populations can upwardly bias sigma estimates. However, these sigma values were still within 1-2 km of the sigma values that accounted for the neighborhood size, suggesting that estimates made without prior data may still provide reasonable approximations of sigma.

2.4 Discussion

Being able to accurately quantify dispersal is important for enabling predictions of how populations will change over time due to connectivity, environmental change, and other determining factors (Cowen & Sponaugle, 2009). Further, understanding larval dispersal can benefit conservation planning by revealing whether populations inside and outside protected areas are connected (Christie et al., 2017; Harrison et al., 2012). In the present study, I used an IBD model to indirectly parameterize dispersal kernels for the sponge-dwelling goby *Elacatinus lori* and found strong agreement with the species' directly measured parentage kernel. These results confirm that *E. lori* exhibits strongly limited dispersal and indicate that—when applied carefully—the indirect IBD method can accurately estimate larval dispersal distances.

2.4.1 Comparison of direct and indirect dispersal estimates in marine species

Our finding of strongly limited dispersal in *E. lori* through the indirect method is consistent with direct parentage analyses (D'Aloia et al., 2013; D'Aloia et al., 2015). The

prior parentage studies uncovered real dispersal events and, therefore, are clearly useful in verifying the accuracy of the IBD estimates. Alternatively, the strong agreement between the two methods could be interpreted as a test of the accuracy of the parentage kernel when considering dispersal over a broader time period and spatial extent. Parentage analysis provides a detailed snapshot of the dispersal patterns of a population at the time of sampling. However, dispersal can fluctuate across years and seasons in some marine species (Catalano et al., 2021; Harrison et al., 2020), and such temporal variability could influence the accuracy of single-year parentage studies. Given the similarity between dispersal estimates, my results suggest that contemporary dispersal patterns are relatively stable for *E. lori,* and therefore dispersal kernels representative of contemporary and long-term dispersal patterns in this species are equivalent.

The present study represents the second known validation of the IBD method with parentage data for a marine species, after Pinsky et al. (2017). The consistency I documented between methods for *E. lori* was similar to that observed in the clown anemonefish *Amphiprion percula* (Pinsky et al., 2017). While *E. lori* and *A. percula* are both small, coral reef microhabitat specialists, they differ in other traits. For instance, the pelagic larval duration for *E. lori* (PLD ≈ 26 days) is longer than that of *A. percula* (PLD ≈ 11 days). *E. lori* is also more continuously distributed along the seascape, exhibiting a subtler IBD pattern than *A. percula*. Additionally, *E. lori* has a larger population density than *A. percula,* both in terms of effective density (*E. lori*: $D_e \approx 749$ (random) - 1,499 (monogamous*)* [this study]*; A. percula:* $D_e = 6.1$ [Pinsky et al., 2017]) and census density estimates (*E. lori*: $D_{census} \approx 850$ adults/km [D'Aloia et al., 2013]; *A. percula*: $D_{census} = 188$

adults/km [Pinsky et al., 2017]). That such congruency was observed in both species underscores the promise of the IBD approach for marine species.

2.4.2 Implications for detecting IBD

A foremost consideration for the application of this indirect method is the true presence of —and ability to detect—an IBD relationship in the target population. While genetic structure is very subtle within this population, an IBD relationship was detected using all three nuclear genetic marker sets, as shown here and in D'Aloia et al. (2020). Nevertheless, the fit of each dataset to the IBD pattern was variable, with SNPs providing the strongest fit, and nonrepetitive nuclear loci providing the weakest (Figure 4). However, I note that the presence of IBD pattern may not hold for all potential marker types. Specifically, mtDNA has been shown to be ill suited for testing for IBD relationships, as it can fail to detect IBD patterns and can provide misleading estimates of species' dispersal potential (Teske et al., 2018).

With respect to the spatial sampling design, I intentionally measured IBD within the barrier reef, where the previous parentage study occurred, to avoid confounding issues such spatially variable dispersal across distinct reef habitats. While IBD is present in this species on the linear barrier reef (Figure 4), previous research has shown that this pattern is absent on the isolated, circular offshore atoll reefs of Belize (D'Aloia et al., 2020). As dispersal events have not been directly detected through parentage analysis for the atoll populations, it is unclear if and how dispersal patterns differ on non-linear *E. lori* habitats. In light of this observation, consideration of the potential influence of habitat

configuration on IBD presence should be taken into account when applying this indirect method.

IBD theory assumes that the populations being studied have reached an equilibrium between the rate of genetic drift and migration (Slatkin, 1993). The validity of this assumption is influenced by the scale of observation for the study species. When spatial genetic structure is measured over spatial scales ~10-50× the spread of the dispersal kernel, equilibrium is generally reached within a few generations (Hardy & Vekemans, 1999; Vekemans & Hardy, 2004). Based on my directly estimated value of σ (3.93 km), the IBD sampling scale in *E. lori* should fall within approximately 40 - 200 km. My sampling region spanned 180 km, thus satisfying this requirement. In other species with parentage or mark-recapture dispersal data, independent σ estimates could similarly be used to inform the design of IBD studies, although indirect estimates may be unnecessary in such species beyond serving to validate direct estimates. Discerning the appropriate scale over which to measure IBD is still possible for species lacking prior dispersal estimates, albeit more challenging. One option is to subsample the data to determine whether the presence and strength of IBD varies across spatial scales (Benestan et al., 2021; Bradbury & Bentzen 2007). If IBD changes substantially with spatial scale, an iterative approach can be used to estimate σ at decreasing scales until a stable σ is produced (following Puebla et al., 2009; Vekemans & Hardy, 2004).

2.4.3 Considerations for estimating N_e in marine populations

The accurate estimation of N_e is the second key component when applying the indirect method and is arguably a greater challenge than measuring IBD. True N_e is expected to be large in many marine populations and, consequently, accurate N_e estimates are difficult to obtain (Marandel et al., 2019; Waples, 2016). Indeed, more uncertainty surrounds my N_e estimates than my IBD slope estimate. However, simulations show that accuracy improves when large marker panels and large numbers of sampled individuals are used (Waples & Do, 2010), and my estimate was based on an informative panel of 63 microsatellites and 4,074 individuals. The large sampling effort produced non-infinite confidence intervals and limited uncertainty in N_e to within a single order of magnitude. Additionally, when this uncertainty was propagated through to dispersal estimates, the resulting confidence intervals around σ spanned only a few kilometers (Figure 5), even when comparing N_e estimates made at the genetic neighborhood scale versus the entire settler collection area. However, I acknowledge that my N_e estimates cannot account for variability in effective size beyond the bounds of my study area (Figure 3 inset). With that caveat, I conclude that uncertainty in N_e had little influence on the outcome of this study, but suggest it is still important to explore as erroneous estimates of N_e have been shown to skew the resulting σ estimates in other species (Pinsky et al., 2017). Differences across studies may be attributed to the non-linear relationship between σ and D_e (eq 3; Figure 9). Holding the IBD slope constant, changes in D_e at small values yield drastically different values of σ. As D_e increases, its relative effect on σ decreases, even with very large increases in D_e. The estimates of D_e generated in this study are sufficiently large that the resulting difference in σ is small.

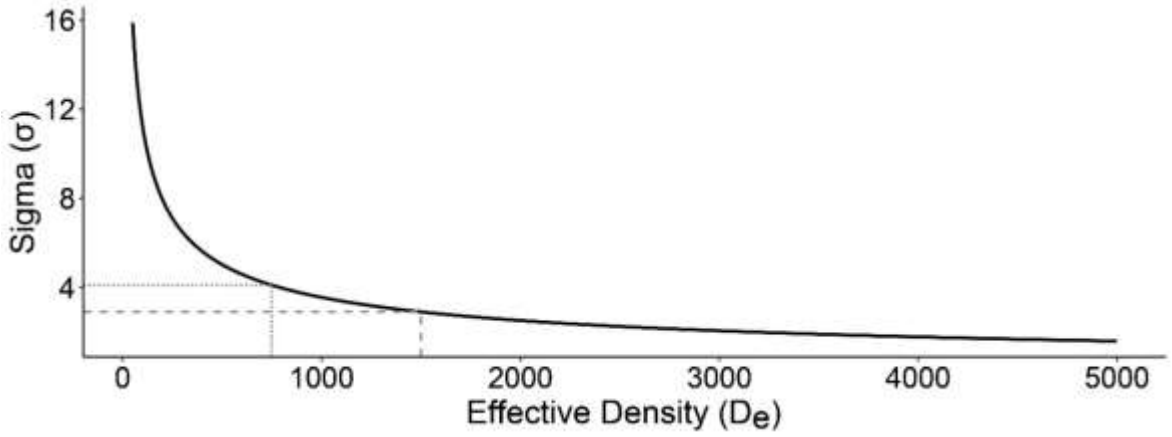

Figure 9. Non-linear relationship between sigma and effective population density for a given value of the IBD slope (*m* = 1.981e-05; SNP data set); see eq 3. Dotted and dashed grey intercepts represent sigma and D_e values under random and monogamous mating.

Because the assumption of mating system has a significant impact on N_e (Nunney, 1993; Waples, 2006), I explored the effect of both random and monogamous mating throughout the main analyses in this thesis. Tropical marine fishes exhibit remarkable diversity in mating system (Warner, 1984), and there are many species for which genetic mating systems have not been studied. In *E. lori*, the genetic mating system is not fully characterized for both sexes, but recent work indicates that the species does not conform to either mating system category in NeEstimator. Instead, some males may mate monogamously (within a single season), while others mate with multiple females, either sequentially or simultaneously (Francis et al. *in review*). Although mating system assumptions did not strongly influence my dispersal estimates, random mating did produce the closest match to the parentage data, with a difference in σ of only 0.17 km. For other species whose mating behaviors are not yet characterized, the influence of

opposing mating system assumptions on N_e should be evaluated and the resulting dispersal estimates compared.

2.4.4 On the limited uptake—and future potential— of the indirect IBD method

Foundational population genetic theory established that IBD patterns reflect limited dispersal (Malécot, 1948; Wright, 1943). However, relatively few empirical studies have taken the next step in using the IBD relationship to explicitly estimate dispersal patterns in the marine environment. Even in terrestrial plants, which have a propagule-driven dispersal life history similar to larval dispersal, the method is established, but not common. Comparative work has revealed consistent direct and indirect genetic dispersal estimates in some plant species (Oddou-Moratio et al., 2010), though not all (Oddou-Moratio & Klein, 2008). What is limiting the uptake of this method?

For species with spatial genetic structure that can be represented as an IBD pattern, the primary limitation is the difficulty in estimating effective (or census) size. While this challenge is not exclusive to marine species, the large N and N_e of many marine populations has likely stymied the uptake of the IBD approach. Therefore, I suggest that, moving forward, the IBD method may be best-suited to species with relatively small population sizes. However, I note that methodological improvements to N_e estimates, such as accounting for overlapping generations (Waples et al., 2014) and the effects of mixture linkage disequilibrium in continuous populations (Neel et al., 2013), coupled with large sampling efforts, could help overcome this hurdle. For example, Waples et al. (2018) demonstrated that a massive sampling effort yielded robust

N and N_e estimates in a large population of bluefin tuna. Additionally, much of the marine dispersal literature has focused on fishes, and this method is promising, but underexplored, in marine invertebrates that exhibit IBD patterns (e.g., Gazulla et al., 2021; Griffiths et al., 2021; Polato et al., 2010).

2.4.5 Conclusions

Quantifying dispersal for marine species is vital for predicting how populations may change over time, as well as for informing conservation decisions, such as planning networks of marine protected areas (Harrison et al., 2020; Jones et al., 2009). To date, it has been difficult to empirically verify the accuracy of indirect measures of larval dispersal in marine species. The present study shows that, with careful consideration of spatial sampling scale and effective population size, the indirect IBD method provides comparable dispersal estimates to those generated from direct parentage analysis. Additionally, the indirect dispersal estimates were relatively insensitive to assumptions about the mating system and the functional form of the kernel, as well as the choice of the genetic marker. This agreement between dispersal estimates in *E. lori* suggests that average dispersal patterns in this species have been consistent at both contemporary timescales and across generations. My results contribute to a limited body of evidence in support of the IBD method as a promising approach to measuring larval dispersal. As there is a wide diversity of marine species that exhibit IBD patterns of genetic differentiation, the implementation of this indirect method could provide a feasible means of investigating larval dispersal distances for other species.

Bibliography

Ackiss, A. S., Bird, C. E., Akita, Y., Santos, M. D., Tachihara, K., Carpenter, K. E. (2018). Genetic patterns in peripheral marine populations of the fusilier fish *Caesio cuning* within the Kuroshio Current. *Ecology and Evolution*, *8*(23), 11875-11886.

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. *Science*, *316*(5825), 742-744.

Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., Bonin, M.C., Frisch, A.J., Harrison, H.B., Messmer, V., Nanninga, G.B., Priest, M.A., Srinivasan, M., Sinclair-Taylor, T., Williamson, D.H., & Jones, G. P. (2017). Larval fish dispersal in a coral-reef seascape. *Nature Ecology & Evolution, 1(6),* 1-7.

Andrade-Restrepo, M., Champagnat, N., & Ferrière, R. (2019). Local adaptation, dispersal evolution, and the spatial eco-evolutionary dynamics of invasion. *Ecology Letters, 22(5),* 767-777.

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., Mouillot, D., Puebla, O., & Manel, S. (2021). Restricted dispersal in a sea of gene flow. *Proceedings of the Royal Society B*, *288*(1951), 20210458.

Bode, M., Williamson, D. H., Harrison, H. B., Outram, N., & Jones, G. P. (2018). Estimating dispersal kernels using genetic parentage data. *Methods in Ecology and Evolution, 9*(3), 490-501.

Bradbury, I.R., & Bentzen, P. (2007). Non-linear genetic isolation by distance: Implications for dispersal estimation in anadromous and marine fish populations. *Marine Ecology Press Series, 340,* 245-257.

Brandl, S. J., Goatley, C. H., Bellwood, D. R., & Tornabene, L. (2018). The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs. *Biological Reviews, 93*(4), 1846-1873.

Broquet, T., & Petit, E. J. (2009). Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution, and Systematics, 40*, 193-216.

Broquet, T., Ray, N., Petit, E., Fryxell, J. M., & Burel, F. (2006). Genetic isolation by distance and landscape connectivity in the American marten (*Martes americana*). *Landscape Ecology*, *21*(6), 877-889.

Catalano, K. A., Dedrick, A. G., Stuart, M. R., Puritz, J. B., Montes Jr, H. R., & Pinsky, M. L. (2021). Quantifying dispersal variability among nearshore marine populations. *Molecular Ecology*, *30*(10), 2366-2377.

Christie, M. R., Meirmans, P. G., Gaggiotti, O. E., Toonen, R. J., & White, C. (2017). Disentangling the relative merits and disadvantages of parentage analysis and assignment tests for inferring population connectivity. *ICES Journal of Marine*

Science, *74*(6), 1749-1762.

Colin, P. L. (2002). A new species of sponge-dwelling *Elacatinus* (Pisces: Gobiidae) from the western Caribbean. *Zootaxa*, *106*(1), 1-7.

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science, 1,* 443-466.

D'Aloia, C. C., Andrés, J. A., Bogdanowicz, S. M., McCune, A. R., Harrison, R. G., & Buston, P. M. (2020). Unraveling hierarchical genetic structure in a marine metapopulation: A comparison of three high-throughput genotyping approaches. *Molecular Ecology*, *29*(12), 2189-2203.

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. *Proceedings of the National Academy of Sciences*, *112*(45), 13940-13945.

D'Aloia, C. C., Xuereb, A., Fortin, M. J., Bogdanowicz, S. M., & Buston, P. M. (2018). Limited dispersal explains the spatial distribution of siblings in a reef fish population. *Marine Ecology Progress Series*, *607*, 143-154.

D'Aloia, C. C., Bogdanowicz, S. M., Majoris, J. E., Harrison, R. G., & Buston, P. M. (2013). Self-recruitment in a Caribbean reef fish: a method for approximating dispersal kernels accounting for seascape. *Molecular Ecology*, *22*(9), 2563-2572.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of softw ware for the estimation of

contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209-214.

Drinan, D. P., Gruenthal, K. M., Canino, M. F., Lowry, D., Fisher, M. C., & Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (*Gadus macrocephalus*). *Evolutionary Applications*, *11*(8), 1448-1464.

Francis, R. K., Catalano, K., Majoris, J. E., Rueger, T., D'Aloia, C. C., Bogdanowicz, S. M, & Buston, P.M. (in review) Characteristics of breeding habitat, genetic mating system, and determinants of mating success in the sponge-dwelling goby *Elacatinus lori*. *Behavioral Ecology and Sociobiology.*

Gazulla, C. R., López-Sendino, P., Antunes, A., Aurelle, D., Montero-Serra, I., Dominici, J. M., Linares, C., Garrabou, J., & Ledoux, J. B. (2021). Genetic Approach for the Conservation and Restoration of a Habitat-Forming Octocoral: The Case of Red Coral, *Corallium rubrum*, in the Réserve Naturelle de Scandola. *Frontiers in Marine Science, 8,* 633057.

Griffiths, S. M., Butler, M. J., Behringer, D. C., Pérez, T., & Preziosi, R. F. (2021). Oceanographic features and limited dispersal shape the population genetic structure of the vase sponge *Ircinia campana* in the Greater Caribbean. *Heredity, 126*(1), 63-76.

Hardy, O. J., & Vekemans, X. (1999). Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics

models. *Heredity, 83*(2), 145-154.

Harrison, H. B., Bode, M., Williamson, D. H., Berumen, M. L., & Jones, G. P. (2020). A connectivity portfolio effect stabilizes marine reserve performance. *Proceedings of the National Academy of Sciences, 117*(41), 25595-25600.

Harrison, H. B., Williamson, D. H., Evans, R. D., Almany, G. R., Thorrold, S. R., Russ, G. R., Feldheim, K.A., van Herwerden, L., Planes, S., Srinivasan, M., Berumen, M.L., & Jones, G. P. (2012). Larval export from marine reserves and the recruitment benefit for fish and fisheries. *Current Biology, 22*(11), 1023-1028.

Hedgecock, D., Barber, P. H., & Edmands, S. (2007). Genetic approaches to measuring connectivity. *Oceanography, 20(3),* 70-79.

Herler, J., Munday, P. L., & Hernaman, V. (2011). Gobies on coral reefs. In R. A. Patzner, & B. G. Kapoor (Eds.), *The Biology of Gobies*, (pp. 493-529). Jersey, British Isles, New Hampshire: Science Publishers.

Jones, G. P., Almany, G. R., Russ, G. R., Sale, P. F., Steneck, R. S., Van Oppen, M. J. H., & Willis, B. L. (2009). Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. *Coral Reefs, 28*(2), 307-325.

Jones, G. P., Planes, S., & Thorrold, S. R. (2005). Coral reef fish larvae settle close to home. *Current Biology, 15*(14), 1314-1318.

Lazoski, C., Gusmão, J., Boudry, P., & Solé-Cava, A. M. (2011). Phylogeny and phylogeography of commercially important Atlantic oyster species: evolutionary

history, limited genetic connectivity and isolation by distance. *Marine Ecology Progress Series*, *426*(3), 197-212.

Majoris, J. E., Francisco, F. A., Atema, J., & Buston, P. M. (2018). Reproduction, early development, and larval rearing strategies for two sponge-dwelling neon gobies, *Elacatinus lori* and *E. colini*. *Aquaculture*, *483*, 286-295.

Malécot, G. (1948). *Mathématiques de l'hérédité*. Paris: Masson et Cie.

Manel, S., Loiseau, N., Andrello, M., Fietz, K., Goñi, R., Forcada, A., Lenfant, P., Kininmonth, S., Marcos, C., Marques, V., Mallol, S., Pérez-Ruzafa, A., Breusing, C., Puebla, O., & Mouillot, D. (2019). Long-distance benefits of marine reserves: myth or reality?. *Trends in ecology & evolution*, *34*(4), 342-354.

Marandel, F., Lorance, P., Berthelé, O., Trenkel, V. M., Waples, R. S., & Lamy, J. B. (2019). Estimating effective population size of large marine populations, is it feasible?. *Fish and Fisheries*, *20*(1), 189-198.

Naaykens, T. (2021). Isolation-by-distance and genetic parentage analysis provide similar larval dispersal estimates, *Zonodo; 2.* doi: 10.5281/zenodo.6025350

Nathan, R., Klein, E. K., Robledo-Arnuncio, J. J., & Revilla, E. (2012). Dispersal kernels: review. J. Clobert, M. Baguette, T.G. Benton, & J.M. Bullock (Eds.), *Dispersal ecology and evolution* (pp. 187-210). Oxford, UK: Oxford University Press.

Neel, M. C., McKelvey, K., Ryman, N., Lloyd, M. W., Short Bull, R., Allendorf, F. W., Schwartz, F.W., & Waples, R. S. (2013). Estimation of effective population size

in continuously distributed populations: there goes the neighborhood. *Heredity, 111*(3), 189-199.

Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution, 47*(5), 1329-1341.

Oddou-Muratorio, S., & Klein, E.K. (2008). Comparing direct vs. indirect estimates of gene flow within a population of a scattered tree species. *Molecular Ecology*, *17*(11), 2743-2754.

Oddou-Muratorio, S., Bontemps, A., Klein, E. K., Chybicki, I., Vendramin, G. G., & Suyama, Y. (2010). Comparison of direct and indirect genetic methods for estimating seed and pollen dispersal in *Fagus sylvatica* and *Fagus crenata*. *Forest Ecology and Management, 259*(11), 2151-2159.

Pinsky, M. L., Montes, Jr, H. R., & Palumbi, S. R. (2010). Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution, 64*(9), 2688-2700.

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.

Polato, N. R., Concepcion, G. T., Toonen, R. J., & Baums, I. B. (2010). Isolation by distance across the Hawaiian Archipelago in the reef-building coral *Porites lobata*. *Molecular Ecology, 19*(21), 4661-4677.

Puebla, O., Bermingham, E., & Guichard, F. (2009). Estimating dispersal from genetic isolation by distance in a coral reef fish (*Hypoplectrus puella*). *Ecology*, *90*(11), 3087-3098.

Puebla, O., Bermingham, E., & McMillan, W. O. (2012). On the spatial scale of dispersal in coral reef fishes. *Molecular Ecology*, *21*(23), 5675-5688.

Roughgarden, J., Gaines, S., & Possingham, H. (1988). Recruitment dynamics in complex life cycles. *Science, 241(4872)*, 1460-1466.

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*(4), 1219-1228.

Rousset, F. (2004). *Genetic structure and selection in subdivided populations (Vol. 40).* Princeton, New Jersey: Princeton University Press.

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.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*(1), 264-279.

Sumner, J., Rousset, F., Estoup, A., & Moritz, C. (2001). 'Neighbourhood' size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandiae*) using individual genetic and demographic methods. *Molecular Ecology*, *10*(8), 1917-1927.

Teske, P.R., Golla, T.R., Sandoval-Castillo, J., Emami-Khoyi, A., van der Lingen, C.D., Von der Heyden, S., Chiazzari, B., van Vuuren, B.J., & Beheregaray, L.B. (2018). Mitochondrial DNA is unsuitable to test for isolation by distance. *Scientific Reports, 8*, 1-9.

Vekemans, X., & Hardy, O. J. (2004). New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, *13*(4), 921-935.

Waples, R. S. (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, *14*(11), 3335-3352.

Waples, R. S. (2006). A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conservation Genetics, 7*(2), 167-184.

Waples, R. S. (2016). Tiny estimates of the N_e/N ratio in marine fishes: Are they real? *Journal of Fish Biology*, *89*(6), 2479-2504.

Waples, R. S., & Do, C. H. I. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications, 3*(3), 244-262.

Waples, R. S., Grewe, P. M., Bravington, M. W., Hillary, R., & Feutry, P. (2018). Robust estimates of a high N e/N ratio in a top marine predator, southern bluefin tuna. *Science Advances, 4(7)*, 7759.

Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, *197*(2),

769-780.

Waples, R. S., Luikart, G., Faulkner, J. R., & Tallmon, D. A. (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1768), 20131339.

Warner, R. R. (1984). Mating behavior and hermaphroditism in coral reef fishes. *American Scientist*, *72*(2), 128-136.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358-1370.

Wright, D., Bishop, J. M., Matthee, C. A., & von der Heyden, S. (2015). Genetic isolation by distance reveals restricted dispersal across a range of life histories: implications for biodiversity conservation planning across highly variable marine environments. *Diversity and Distributions*, *21*(6), 698-710.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114- 138.

Wright, S. (1946). Isolation by distance under diverse systems of mating. *Genetics*, *31*(1), 39.

3. Curriculum Vitae

Candidate's full name: Taylor MacKenzie Naaykens

Universities attended (with dates and degrees obtained): University of Manitoba (BSc Hons Biological Sciences, May 2016)

Publications:

Naaykens T, D'Aloia CC. Isolation-by-distance and genetic parentage analysis provide similar dispersal estimates in a coral reef fish. *Molecular Ecology*. In review.

Naaykens T, Roth JD, Dupont D. (2016). Grey wolf selection for moose calves and factors influencing prey species consumption in southeastern Manitoba. *Proceedings of Manitoba's Undergraduate Science and Engineering Research 2*:1-17.

Conference Presentations:

Naaykens T, D'Aloia CC. 2021. Genetic parentage analysis and isolation-by-distance provide strikingly similar dispersal estimates in a coral reef fish. 14th International Coral Reef Symposium. Bremen, Germany (virtual).

Naaykens T, D'Aloia CC. 2021. Parentage analysis and isolation-by-distance provide comparable estimates for goby dispersal. Canadian Society for Ecology and Evolution. Victoria, BC (virtual).

Naaykens T, Roth JD, and Dupont D. 2016. Grey wolf selection for moose calves and factors influencing prey species consumption in southeastern Manitoba. 50th North American Moose Conference/Workshop. Brandon MB.

Naaykens T, Roth JD, and Dupont D. 2016. Wolves eating moose: selection for moose calves and influences on prey species consumption. Parks and Protected Areas Research Forum of Manitoba. Winnipeg MB.