Vol. 721: 161–167, 2023
https://doi.org/10.3354/meps14419

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published October 19

NOTE

Neither larval duration nor dispersal distance predict spatial genetic diversity in planktonic dispersing species

Elizabeth A. Esser[1,3], James M. Pringle[2], James E. Byers[1,*]

[1]Odum School of Ecology, University of Georgia, Athens, GA 30602, USA
[2]Department of Earth Sciences, and the Institute of Earth, Oceans, and Space, 142 Morse Hall, 8 College Road, University of New Hampshire, Durham, NH 03824, USA
[3]*Present address:* College of Forest Resources, Mississippi State University, Starkville, MS 39762, USA

ABSTRACT: The increase in genetic distance between marine individuals or populations with increasing distance has often been assumed to be influenced by dispersal distance. In turn, dispersal distance has often been assumed to correlate strongly with pelagic larval duration (PLD). We examined the consistency of these assumptions in species with long planktonic durations. Reviewing multiple marine species, Selkoe & Toonen (2011; Mar Ecol Prog Ser 436:291–305) demonstrated significant fit of a species' PLD with metrics of genetic distance between sampling sites. However, for long dispersers (PLD >10 d) whose dispersal is more influenced by ocean currents, the fit of PLD and genetic connectivity metrics was not significant. We tested if using realistic ocean currents to determine simulated dispersal distances would produce an improved proxy for larval dispersal that correlates more strongly with genetic connectivity metrics. We estimated the dispersal distance of propagules for locations in the genetic studies compiled by Selkoe and Toonen with a global ocean model (Mercator, 1/12° resolution). The model-derived estimates of dispersal distance did not correlate better than PLD against the genetic diversity metrics global F_{ST} km^{-1} and isolation-by-distance (IBD) slope. We explored 2 explanations: (1) our ocean circulation-based dispersal distance estimates are too simple to produce biologically meaningful improvement over PLD, and (2) IBD slope is not a powerful predictor of variation in dispersal distance between species with long PLD. Exploring these explanations suggests directions for future research which will enable better quantitative understanding of genetic diversity and its spatial distribution in coastal marine organisms.

KEY WORDS: Advection · Oceanographic modeling · Pelagic dispersal · Planktotrophic larvae · Broadcast spawning · Downstream drift

1. INTRODUCTION

Connectivity among populations drives many ecological and evolutionary processes including local adaptation, population dynamics, and interspecific interaction strengths (Hanski 1998). Connectivity is often inferred from spatial gradients in gene frequency such as isolation by distance (IBD), whose strength is assumed to be controlled in part by the dispersal of individuals and their offspring (Rousset 1997). There are 2 potential problems with these inferences. First, spatial gradients in genetic pat-

---

[*] Corresponding author: jebyers@uga.edu

terns, like IBD, are also influenced by the effective population density (Rousset 1997), and the relationship between IBD and dispersal distance will only be accurate if the species' system and sampling meets certain restrictive assumptions (including spatially and temporally stationary dispersal statistics, limited distance between sampled individuals, and genetic equilibrium between drift and dispersal). This casts doubt on the reliability of the relationship between dispersal distance and IBD.

Second, even if IBD is governed predominantly by dispersal distance, dispersal distance itself is hard to estimate for species with long planktonic durations. For many benthic species, dispersal occurs primarily during a planktonic stage that is controlled by ocean currents and by biological traits such as larval development time and vertical behavior (Byers & Pringle 2006, Oyarzun & Strathmann 2011). Because dispersal by a steady current is proportional to the current speed multiplied by the time a larvae drifts in that current, investigators often assume ocean currents are spatially uniform and use an organism's pelagic larval duration (PLD) as a proxy for dispersal distance (e.g. Shanks et al. 2003, Cowen & Sponaugle 2009). Therefore, many investigators have explored quantitative relationships between PLD and IBD to infer how dispersal distance affects population structure and connectivity (e.g. Siegel et al. 2003), and have used IBD to estimate dispersal (e.g. Palumbi 2003).

To scrutinize the validity of drawing such inferences, we used currents from an ocean model to convert PLD into dispersal distances so that a proxy variable is no longer needed. We then tested the quantitative strength of the direct link between dispersal distance and IBD. To do so, we explored analyses and data compiled by Selkoe & Toonen (2011) (hereafter S&T). S&T used PLD as a proxy for dispersal distance and compared the PLD of organisms to the genetic spatial diversity metrics IBD and the closely related F_{ST} per kilometer of sampling region (F_{ST} km^{-1}) for species with planktonic dispersal from many studies. The correlations S&T found between PLD and F_{ST} km^{-1} and between PLD and IBD slope were both significant. However, most of the correlation strength was driven by the bimodality of PLD values in their data set, with PLDs either less than a few days or >10 d (Fig. 1). Our analysis of S&T's data set reveals that when only species with a PLD greater than 10 d (long dispersers) are included, the correlation between PLD and the genetic metrics is either not, or only marginally, significant.

One of the key steps in our approach is to use numerical ocean circulation models to make independent, place-based estimates of dispersal distance. To that end, we used high-resolution (1/12 of a degree) circulation models to compute the path of propagules, quantifying dispersal distance. Although realistic larval dispersal including depth changes and settlement behavior is still difficult to accurately simulate, this method gives more information about dispersal than PLD by including the spatial and temporal variation of ocean currents. When estimated dispersal distance is correlated against PLD for long dispersers, PLD explains only 34% of the variability (Fig. 2), suggesting that a perfect measure of dispersal could explain up to 66% more of the variability in dispersal than a model with PLD alone.

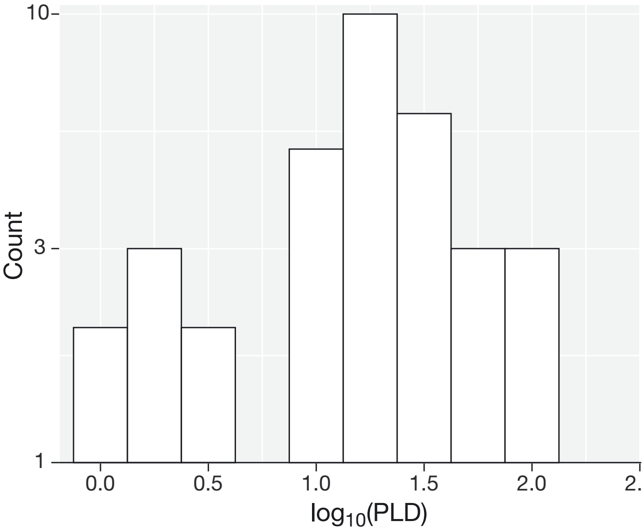

Fig. 1. Distribution of pelagic larval duration (PLD) lengths in the data set of Selkoe & Toonen (2011)

We compared IBD and F_{ST} km^{-1} from S&T for species with PLD >10 d against both PLD and the ocean model estimate of dispersal. If the oceanographically realistic dispersal distance estimates are more correlated to these measures of spatial genetic diversity than PLD, it will strongly suggest that these genetic metrics are significantly driven by dispersal distance (and could even be used to estimate dispersal distance). If, conversely, the correlation is not improved, it suggests that either the genetic metrics are not strongly influenced by changes in dispersal distance for species with PLD >10 d, or that our circulation-based estimates of dispersal have no skill beyond PLD in estimating dispersal.

2. METHODS

Dispersal was estimated for surface-trapped particles advected using the high-resolution 1/12 degree

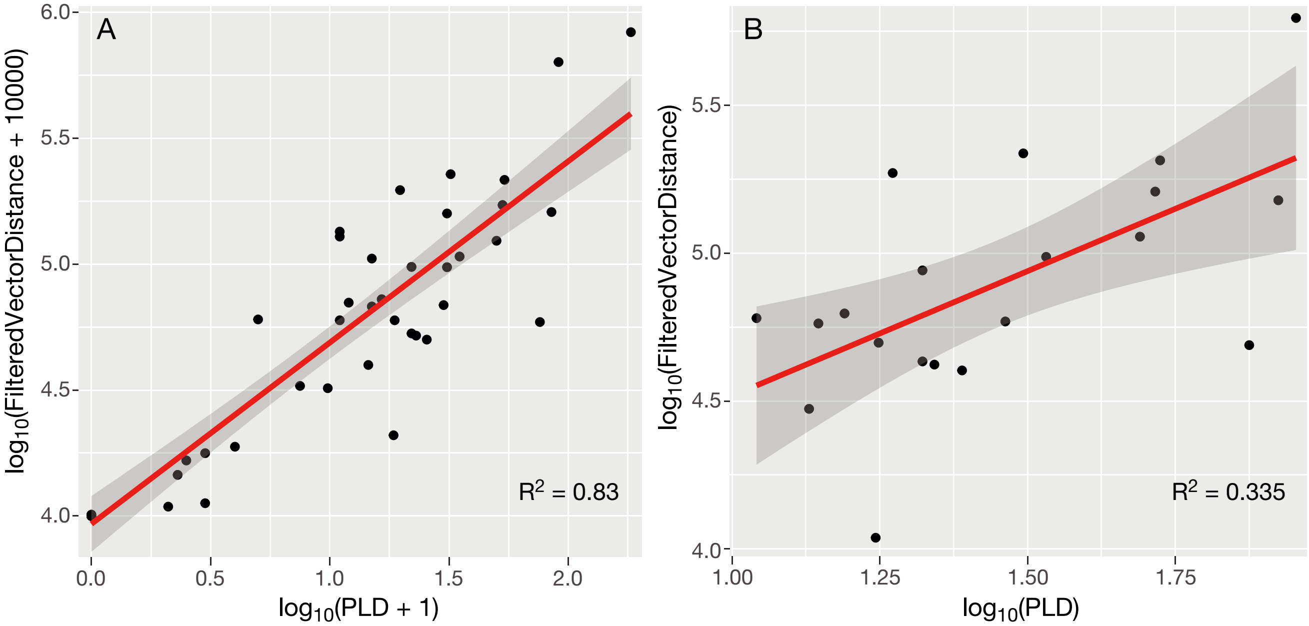

Fig. 2. Log pelagic larval duration (PLD) versus log filtered vector-mean distance for (A) the full range of PLD and (B) only PLD >10 d. Dark gray regions: 95 % confidence interval

Mercator model of ocean currents. The model is forced with surface fluxes from numerical atmospheric models and assimilates multiple sources of ocean data to get a statistically optimal estimate of ocean currents (Cummings & Smedstad 2013, Lellouche et al. 2013, 2018). The Mercator model was found to be superior in the coastal ocean to other global models at all time scales (Wilkin & Hunter 2013).

Sampling sites from the studies compiled by S&T were used as starting points for the dispersal estimates. The location of each sampling site not listed in the studies was estimated in Google Earth based on the site name or figures depicting site location. For 19 studies where the genetic metric was calculated by grouping data from proximate sampling sites together, the first sampling site of the group listed in the study represented that sampling location. One species was excluded because explicit sampling locations were not reported.

Particles were released at 5 evenly spaced points within 1/28 of a degree (~4 km) around each sampling site. For each species and each sampling location, particles were released every half hour from 2007 to 2017 and allowed to disperse for the PLD given by S&T. The Lagrangian particle tracks were calculated with Ocean Parcels 2.2 (Delandmeter & Van Sebille 2019). Since nearly all species in these studies are coastal, ending locations >50 km from shore were removed from our analysis (Fig. 3).

Three dispersal distances were calculated for each species. The distance between the mean start point and the mean end location for all larvae from each release site was calculated (Fig. 4). The 'filtered vector-mean distance' is the average of this distance for all sampling sites for a species. 'Mean absolute distance' is the average great circle distance between the end point of every particle and its corresponding start point for each species (Fig. S1 in Supplement 1 at www.int-res.com/articles/suppl/m721p161_supp1.pdf). 'Vector standard deviation' is the root mean square of the distance between the mean end location and the actual end points of each particle for each release site and is averaged over all release sites (Fig. S2 in Supplement 1).

Vector-mean distance was also calculated for particles released only during spring, the most common spawning season for marine organisms (April, May, and June in the Northern hemisphere; September, October, and November in the Southern Hemisphere). This did not improve the fit of dispersal distance on genetic connectivity (Fig. S3 in Supplement 1).

To directly compare our dispersal measures to those of S&T, we re-created S&T's plots of PLD against the F_{ST} km^{-1} and IBD slope reported by S&T. All species are shown in Fig. S4 in Supplement 1; only those with a PLD >10 d are shown in Fig. 5. Following S&T, we performed linear regression to determine the correlation between the log-transformed dispersal distance estimate and the log-transformed genetic metrics using R version 3.6.2 (all code and data are included in Supplements 2 & 3 at www.int-res.com/articles/suppl/m721p161_supp2.xlsx and www.int-res.com/articles/suppl/m721p161_supp3.pdf). We report adjusted R^2 values, using Dunn and Clarke's *z*-test to determine if R^2 values differed between analyses. Multiple linear regression was used to determine if the combination of both PLD and distance data was more strongly correlated to the genetic connectivity metrics. For each genetic response variable, the full and reduced versions of each model were compared using Akaike's information criterion (AIC). Regressions and correlations using all PLDs are reported in Supplement 1.

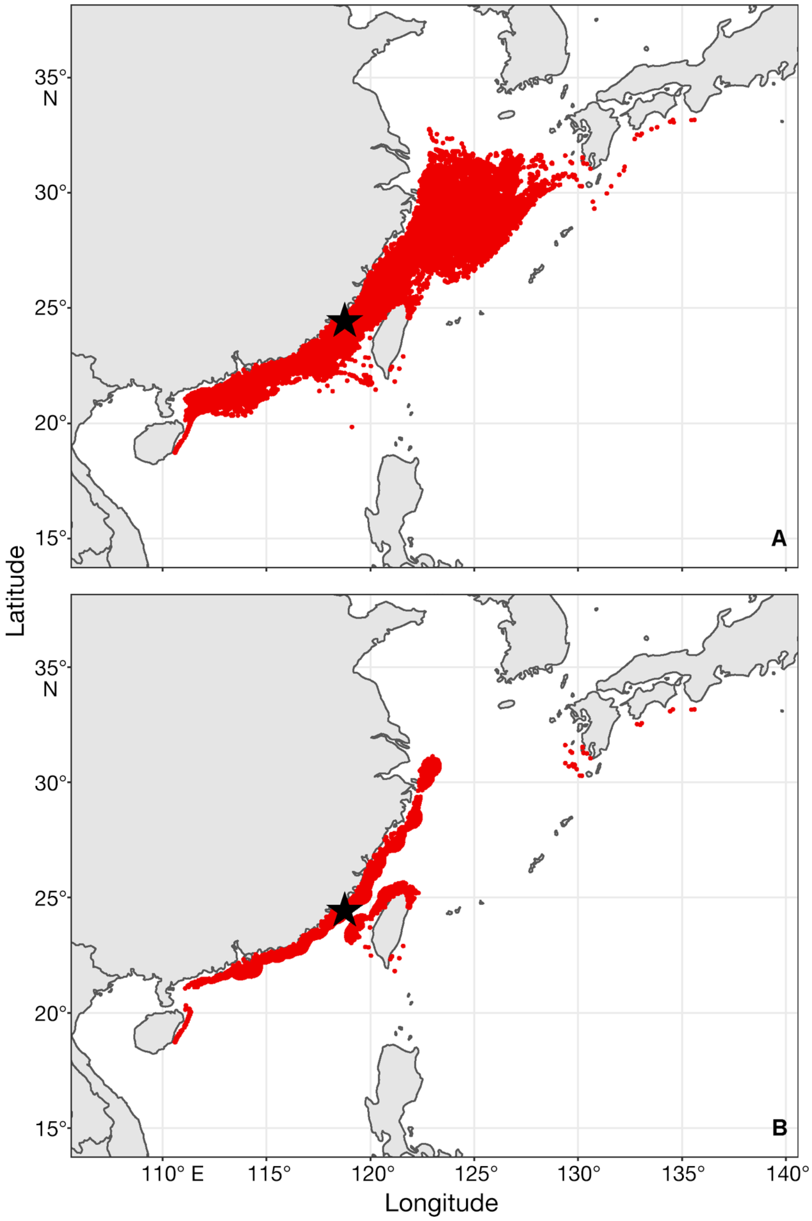

Fig. 3. Dispersal end points for example larval dispersal simulation released from 1 starting point (black star) (A) before and (B) after the 50 km offshore-filter was applied. Red represents the final location of a particle

3. RESULTS

3.1. F_{ST} km^{-1}

S&T's correlation between log-transformed PLD and log-transformed F_{ST} km^{-1} for all species was $R^2 = 0.609$ (all correlations here and below are between log-transformed variables). Our re-created correlation between PLD and F_{ST} km^{-1} using S&T's data was $R^2 = 0.567$ ($p < 0.0001$) (Fig. S4A in Supplement 1), which differed slightly from the results of S&T because we excluded species without larvae that ended within 50 km offshore and excluded 1 species that lacked spatially explicit sampling sites. Using only species with a PLD >10 d, the correlation between PLD and F_{ST} km^{-1} was $R^2 = 0.179$ ($p = 0.0225$) (Fig. 5A). Filtered vector-mean distance had a weaker correlation with F_{ST} km^{-1} ($R^2 = 0.0158$, $p = 0.254$), but the difference between the 2 models was not significant (Dunn and Clark's *z*-test, $p = 0.156$) (Fig. 5B). In multiple linear regression correlating F_{ST} km^{-1} to PLD and filtered vector-mean distance, the interaction term between PLD and distance did not significantly improve the model (with interaction term: $R^2 = 0.124$, $p = 0.134$; without interaction term: $R^2 = 0.153$, $p = 0.068$). The most parsimonious model explaining F_{ST} km^{-1} included PLD only, with an AIC value 1.64 lower than the next best model, which included PLD and filtered vector-mean distance (Table S1 in Supplement 1). The correlation between PLD and F_{ST} km^{-1} was stronger than the correlation between F_{ST} km^{-1} and any of the other distance metrics (Figs. S3, S5 & S6 in Supplement 1).

3.2. IBD slope

For species with a PLD >10 d, the correlation between PLD and IBD slope was $R^2 = -0.0326$ (p =

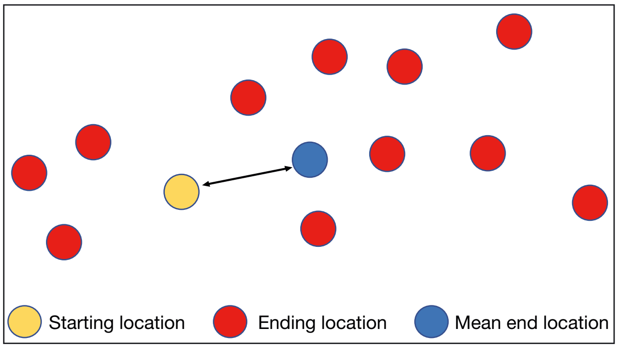

Fig. 4. Filtered vector-mean distance calculation. The blue dot is the mean ending location of all the red dots. The arrow between the starting location and the mean ending location represents the vector-mean distance. These vector-mean distances were averaged across all discrete sampling sites in a study to yield an overall mean for each species in a study

0.5352) (Fig. 5C). Filtered vector-mean distance had a similar correlation to IBD slope, with $R^2 = -0.005$ ($p = 0.356$) (Fig. 5D). The difference in R^2 was not statistically significant (Dunn and Clark's *z*-test, $p = 0.732$). AIC values of both models differed by only 0.55, indicating that the models have equal parsimony (Table S2 in Supplement 1). Including the interaction term in a multiple linear regression correlating both PLD and filtered vector-mean distance to IBD slope did not improve the correlation (with interaction term: $R^2 = -0.041$, $p = 0.536$; without: $R^2 = -0.064$, $p = 0.659$), and AIC values confirmed that models without the interaction term were more parsimonious (Table S2). None of the other distance metrics had significant correlations (Figs. S3, S5 & S6 in Supplement 1).

3.3. PLD and distance correlation

Using the full data set, PLD and filtered vector-mean distance were strongly correlated ($R^2 = 0.83$, $p < 0.0001$) (Fig. 2A). When using only species with a PLD >10 d, the correlation was much weaker ($R^2 = 0.34$, $p = 0.004$) (Fig. 2B).

4. DISCUSSION

Estimated dispersal distance did not correlate with F_{ST} km^{-1} or IBD slope more strongly than PLD. Neither PLD nor dispersal distance explains much of the variability in metrics of spatial genetic diversity for species with long-lived larvae (PLD >10 d). This suggests 2 non-exclusive explanations:

(1) Circulation-based estimates of dispersal have no more skill at predicting dispersal than PLD.

(2) Dispersal distance does not relate simply to spatial genetic diversity in the ocean, so even improved dispersal estimates will not improve correlation to the genetic metrics.

Explanation (1) suggests that our model of dispersal is not sufficient to improve upon other proxies. Larvae drift in the real ocean, not a numerical ocean, and the two can differ. However, results from comparison of observations to Mercator surface current in the Mid-Atlantic Bight (Wilkin & Hunter 2013) and the European Shelf Seas (Lellouche et al. 2013, Brennan et al. 2016) suggest that the Mercator circulation model does a reasonable job of estimating coastal currents.

Our model also did not include any larval behavior that could affect dispersal distance. The larval analogues were kept 1 m from the surface, while larvae for many species preferentially seek other depths (Queiroga 1996, Miller & Morgan 2013). The propagules were released into our numerical model in all seasons, whereas most marine species spawn only during certain months. However, calculations of dispersal in only spring months, when many species spawn (Reitzel et al. 2004), did not improve correlations (Fig. S3 in Supplement 1). It seems unlikely that larval vertical behavior and phenology would eliminate all effects of regionally varying ocean currents on the vector-mean dispersal; plankton would still be expected to be moved farther by currents in regions where currents are stronger.

Explanation (2) suggests that dispersal distance might not control genetic connectivity metrics as straightforwardly as often assumed; equivalently, these genetic metrics may be a poor proxy for dispersal. This work, S&T, and many others assume that dispersal is the primary control on F_{ST} km^{-1} and IBD, but theoretical work suggests that these are controlled by many factors. In the work of Rousset (1997), IBD was governed by the product of dispersal distance and population density, and population density would be expected to vary considerably between different species in the data set of S&T. Rousset (1997) also assumed temporal equilibrium of genetic diversity (which can take a long time to achieve relative to, for example, the last ice-age; Hellberg 2009). To the extent that IBD is influenced by non-equilibrium dynamics, selection, or temporal variability in dispersal or reproductive success, its use as a proxy for dispersal (or its estimate from dispersal) is suspect.

Even in temporal equilibrium, it is not clear that Rousset's (1997) theory of IBD applies to advective environments like the ocean. This theory assumes no pre-

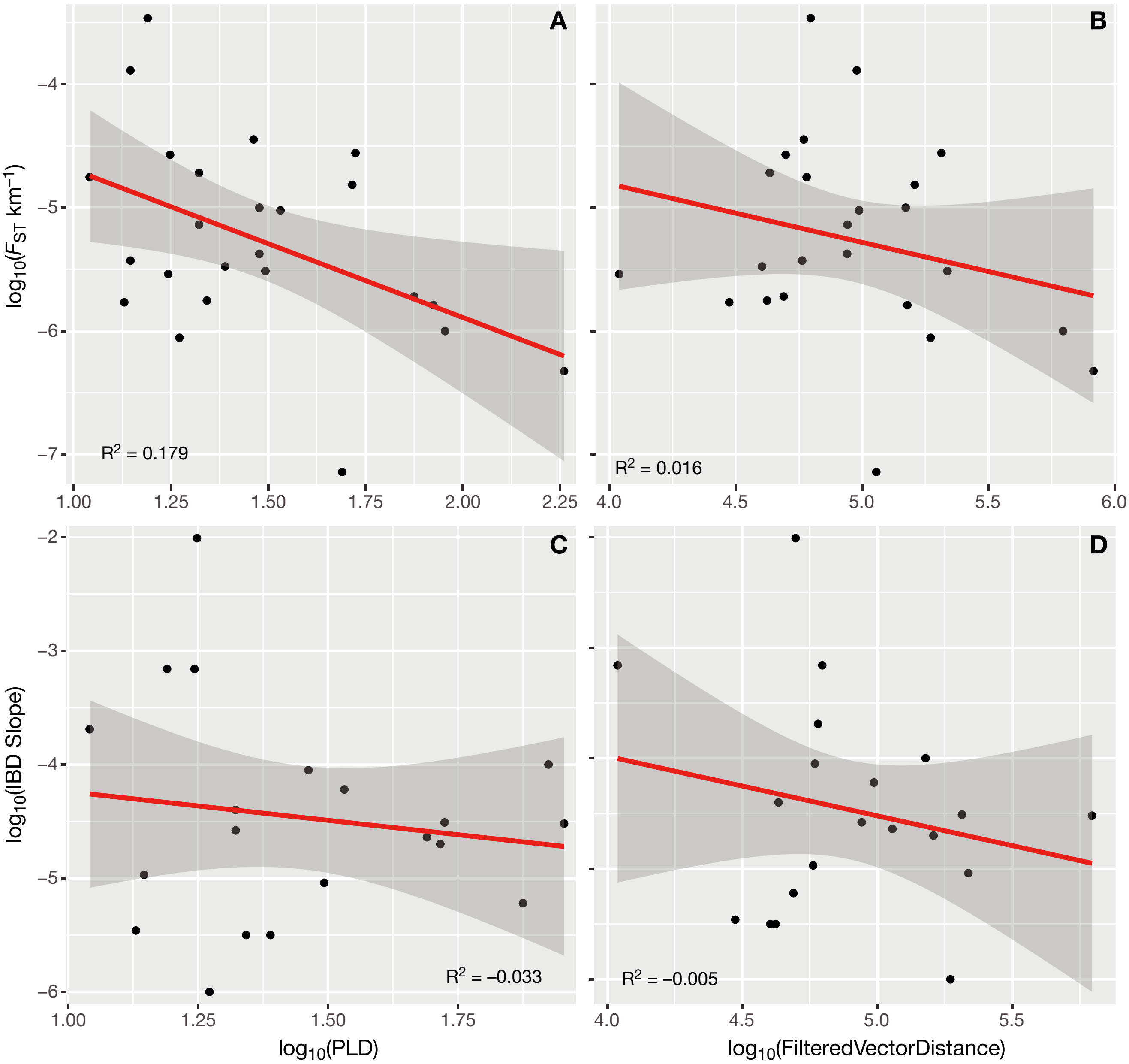

Fig. 5. Top row: log F_{ST} km^{-1}, which is the F_{ST} standardized by the length of the sampling region, versus (A) log pelagic larval duration (PLD) and (B) log filtered vector-mean distance. Bottom row: log isolation by distance (IBD) slope versus (C) log PLD and (D) log filtered vector-mean distance. Dark gray regions: 95 % confidence interval

ferred directionality to larval dispersal, spatially homogeneous dispersal statistics, and no nested structure in the population. These are not good assumptions in the ocean, where currents have strong directional biases and larvae are preferentially moved in one direction (Largier 2003, Siegel et al. 2003, Byers & Pringle 2006). When larval dispersal is biased in a single direction, a source/sink metapopulation structure can develop in which the magnitude of the genetic diversity is not governed by the vector-mean dispersal but by a combination of the vector-mean dispersal, the vector standard deviation of dispersal, and alongshore variation in habitat or interruptions to alongshore transport (Wares & Pringle 2008, Pringle et al. 2011). This leads to genetic structure that consists of regions of relatively homogeneous genetics separated by spatially localized clines in gene frequency (e.g. Altman et al. 2013 and citations therein). In these systems, there is no simple relation between IBD and any single measure of dispersal. This makes IBD a poor proxy for dispersal.

The need for estimates of dispersal distance of marine species with planktonic larval dispersal, combined with the difficulty in measuring this distance, has led to a desire for proxies for dispersal. However, these proxies are not carefully validated, and they lead us astray. The assumption that PLD is a proxy for dispersal distance, that dispersal is simply related to IBD, and that dispersal can be estimated from IBD, are rampant in the literature despite little empirical or theoretical support. Our results suggest that these assumptions are problematic. We suggest that more biologically and oceanographically realistic estimates of planktonic dispersal, and theories of how this dispersal influences spatial genetic diversity, are necessary before any of these can be used as proxies for each other in species with long planktonic durations.

Acknowledgements. We thank Rob Toonen and several reviewers for valuable feedback, and NSF funding from OCE-1947954 and OCE-1947884.

LITERATURE CITED

Altman S, Robinson JD, Pringle JM, Byers JE, Wares JP (2013) Edges and overlaps in Northwest Atlantic phylogeography. Diversity 5:263–275

Brennan CE, Bianucci L, Fennel K (2016) Sensitivity of northwest North Atlantic shelf circulation to surface and boundary forcing: a regional model assessment. Atmos-Ocean 54:230–247

Byers JE, Pringle JM (2006) Going against the flow: retention, range limits and invasions in advective environments. Mar Ecol Prog Ser 313:27–41

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Annu Rev Mar Sci 1:443–466

Cummings JA, Smedstad OM (2013) Variational data assimilation for the global ocean. In: Park S, Xu L (eds) Data assimilation for atmospheric, oceanic and hydrologic applications (Vol II). Springer, Berlin, p 303–343

Delandmeter P, Van Sebille E (2019) The Parcels v2. 0 Lagrangian framework: new field interpolation schemes. Geosci Model Dev 12:3571–3584

Hanski I (1998) Metapopulation dynamics. Nature 396:41–49

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Annu Rev Ecol Evol Syst 40: 291–310

Largier JL (2003) Considerations in estimating larval dispersal distances from oceanographic data. Ecol Appl 13:71–89

Lellouche JM, Le Galloudec O, Drévillon M, Régnier C and others (2013) Evaluation of global monitoring and forecasting systems at Mercator Océan. Ocean Sci 9:57–81

Lellouche JM, Greiner E, Le Galloudec O, Garric G and others (2018) Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system. Ocean Sci 14:1093–1126

Miller SH, Morgan SG (2013) Interspecific differences in depth preference: regulation of larval transport in an upwelling system. Mar Ecol Prog Ser 476:301–306

Oyarzun FX, Strathmann RR (2011) Plasticity of hatching and the duration of planktonic development in marine invertebrates. Integr Comp Biol 51:81–90

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. Ecol Appl 13:146–158

Pringle JM, Blakeslee AMH, Byers JE, Roman J (2011) Asymmetric dispersal allows an upstream region to control population structure throughout a species' range. Proc Natl Acad Sci USA 108:15288–15293

Queiroga H (1996) Distribution and drift of the crab *Carcinus maenas* (L.) (Decapoda, Portunidae) larvae over the continental shelf off northern Portugal in April 1991. J Plankton Res 18:1981–2000

Reitzel AM, Miner BG, McEdward LR (2004) Relationships between spawning date and larval development time for benthic marine invertebrates: a modeling approach. Mar Ecol Prog Ser 280:13–23

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

Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Mar Ecol Prog Ser 436:291–305

Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. Ecol Appl 13:159–169

Siegel DA, Kinlan BP, Gaylord B, Gaines SD (2003) Lagrangian descriptions of marine larval dispersion. Mar Ecol Prog Ser 260:83–96

Wares JP, Pringle JM (2008) Drift by drift: effective population size is limited by advection. BMC Evol Biol 8:235

Wilkin JL, Hunter EJ (2013) An assessment of the skill of real-time models of Mid-Atlantic Bight continental shelf circulation. J Geophys Res Oceans 118:2919–2933

Editorial responsibility: Rob Toonen,
Kāne'ohe, Hawai'i, USA
This and a previous version reviewed in MEPS by:
5 anonymous referees in total

Submitted: May 6, 2023
Accepted: August 21, 2023
Proofs received from author(s): October 13, 2023