Running head: The formation of marine kin structure

Title: **The formation of marine kin structure: effects of dispersal, larval cohesion, and variable reproductive success**

Authors: Cassidy C. D'Aloia[*1] and Michael G. Neubert[1]

[1] Biology Department, Woods Hole Oceanographic Institution, 266 Woods Hole Road, MS # 34, Woods Hole, MA 02543-1049

[*] Corresponding author current contact information:

Address: Department of Ecology & Evolutionary Biology, University of Toronto, 25 Willcocks St., Toronto, ON M5S 3B2; e-mail: cassidy.daloia@gmail.com

Abstract

The spatial distribution of relatives has profound effects on kin interactions, inbreeding, and inclusive fitness. Yet, in the marine environment, the processes that generate patterns of kin structure remain understudied because larval dispersal on ocean currents was historically assumed to disrupt kin associations. Recent genetic evidence of co-occurring siblings challenges this assumption and raises the intriguing question of how siblings are found together after a (potentially) disruptive larval phase. Here, we develop individual based models to explore how stochastic processes operating at the individual level affect expected kinship at equilibrium. Specifically, we predict how limited dispersal, sibling cohesion, and variability in reproductive success differentially affect patterns of kin structure. All three mechanisms increase mean kinship within populations, but their spatial effects are markedly different. We find that: (1) when dispersal is limited, kinship declines monotonically as a function of the distance between individuals; (2) when siblings disperse cohesively, kinship increases within a site relative to between sites; and (3) when reproductive success varies, kinship increases equally at all distances. The differential effects of these processes therefore only become apparent when individuals are sampled at multiple spatial scales. Notably, our models suggest that aggregative larval behaviors, such as sibling cohesion, are not necessary to explain documented levels of relatedness within marine populations. Together, these findings establish a theoretical framework for disentangling the drivers of marine kin structure.

Keywords: individual based model, relatedness, collective dispersal, aggregated dispersal, sweepstakes reproductive success, kinship, larval dispersal, marine ecology

Introduction

When kin are close in space, they can cooperate, compete, and mate with each other. In turn, these interactions influence individual fitness (Crnokrak and Roff 1999), population dynamics (Keller and Waller 2002), and the evolution of social behavior via inclusive fitness (Hamilton 1964, Queller 1994). Understanding when and how kin are structured in space is therefore of broad interest in ecology and evolutionary biology.

Historically, kinship has been understudied in marine environments relative to terrestrial environments. This lag has been attributed to the assumption that the biphasic life cycle of many marine species, wherein larvae disperse and adults are relatively sedentary, should disrupt kin associations (Kamel and Grosberg 2013). Several aspects of larval ecology contribute to the intuition that first-degree relatives (e.g., parent-offspring pairs or siblings) should not be in close proximity. First, larvae disperse in a fluid environment and the dispersive phase can last upwards of several months. Larvae therefore have the potential to travel far from their parents, and may also become separated from their siblings, depending on the movement of the water they are traveling in. Second, multiple cohorts can mix in the plankton, homogenizing the larval pool and diluting the fraction of the pool that is closely related. Third, larval mortality rates are high. If mortality rates are independent of family line, any given individual will have few surviving kin. Collectively, these ideas were taken as indirect support for the hypothesis that few relatives would be found in close proximity within marine populations after settlement (Victor 1984, Leis 1991).

Despite the potential for kin disruption, genetic studies have shown that across major taxonomic groups, pairs of putative siblings, and even family groups, are sometimes found within the same microhabitat or sampling location after settlement (e.g., Iacchei et al. 2013, Salles et al. 2016, Selwyn et al. 2016, Riquet et al. 2017). These findings challenge conventional views of kinship in the sea and raise questions about how kin associations persist despite the potential for dispersal to disrupt them. Theoretical work can help answer these questions by elucidating how various mechanisms could maintain sibling associations and, consequently, shape spatial patterns of kinship. In particular, empirical find-

ings have raised the intriguing question of whether observing first-degree relatives nearby in space requires aggregative larval dynamics, or whether such observations are consistent with a more parsimonious explanation such as limited dispersal.

Population genetic models provide some insight into how limited dispersal can create spatial kin structure (Wright 1943, Malécot 1969). If dispersal success declines with distance, an isolation by distance pattern emerges, where genetic distance increases with geographic distance. Malécot (1969) quantified genetic distance as the kinship coefficient ϕ_{ij} between individuals i and j, defined as the probability that two homologous alleles, one sampled from each individual, are identical by descent. At migration-drift equilibrium, with a small mutation rate, kinship coefficients decline with distance. Subsequent papers have extended this theory to consider stepping-stone models and nonequilibrium populations (e.g., Kimura and Weiss 1964, Slatkin 1993), but the basic idea has remained a cornerstone of both theoretical and empirical population genetics.

Despite the longstanding impact of isolation by distance theory, limited dispersal has not gained much traction in marine ecology as a probable explanation for the observation of close kin within populations. Marine populations often exhibit weak signals of isolation by distance, and sometimes exhibit alternative patterns of genetic structuring, such as chaotic genetic patchiness, regional structuring, or panmixia (Johnson and Black 1982, Selkoe et al. 2014). These patterns all support the notion that long distance dispersal and gene flow can be common in the marine environment. Nevertheless, genetic parentage studies have revealed limited dispersal in some species (D'Aloia et al. 2015, Salles et al. 2016); therefore, it remains a plausible explanation for marine kin structure.

An alternative explanation is that variability in reproductive success among individuals contributes to the co-occurrence of close kin. The hypothesis of sweepstakes reproductive success posits that highly fecund marine species exhibit variability in reproductive success due to the fortuitous matching of reproductive output with "oceanographic conditions conducive to gamete maturation, fertilization, larval development, settlement, and recruitment to the adult spawning population" (Hedgecock and Pudovkin 2011). When re-

4

productive success is extremely skewed, the sweepstakes hypothesis predicts that kinship within larval cohorts could be higher than kinship between larval cohorts.

A third hypothesis is that some larvae remain cohesive during the dispersive phase via passive retention in the same water mass or active aggregative behavior. Recent papers have invoked this cohesion hypothesis as an explanation for empirical findings of siblings pre- or post-settlement (Selwyn et al. 2016, Riquet et al. 2017). While there is evidence from otolith microchemistry that some individuals within larval cohorts have shared dispersal pathways for part of the larval phase (Ben-Tzvi et al. 2012, Shima and Swearer 2016), the kinship of cohesive larvae remains unclear. To affect kin structure, some cohesive larvae must be kin, and their aggregation must persist throughout the larval phase.

Because it is not yet possible to track larvae from spawning to settlement site, the dynamics of dispersing larvae are not directly observable, and disentangling the potential drivers of kin structure remains a major challenge. To contextualize the growing number of empirical kinship studies, we need clearer predictions about how stochastic demographic processes and aggregative behaviors can give rise to spatial patterns of kinship. Here, we use individual-based models to explore alternative mechanisms that have been proposed to impact marine kinship. As a starting point, we develop a general model with a simplified life cycle and non-overlapping generations. Our primary goal is to elucidate how limited dispersal, sibling cohesion, and variability in reproductive success differentially affect post-settlement patterns of kin structure at increasing spatial scales. In developing these theoretical predictions, our secondary goal is to make general recommendations about spatiotemporal sampling designs that can disentangle the drivers of marine kin structure.

Methods

Model overview

We developed a set of discrete-time, individual-based models to explore the effects of dispersal, cohesion, and variable reproductive success on the formation of marine kin structure. Each model includes the following sequential steps: breeding pairs form; breeding

pairs reproduce; adults die; larvae disperse; a fraction of larvae successfully settle, including a small number of immigrants from outside of the system; and successful settlers mature into the next generation of breeding adults (Fig. 1). For a full description of the numerical implementation of the model, see Appendix S1. We simulated the models on a finite one-dimensional lattice of uniformly-sized sites with equal carrying capacity (a coarse approximation of marine habitats such as linear coastlines or reefs). Each individual was characterized by an individual identifier; site identifier; sex; and identifiers for its mother, father, mate, and whether it was a resident or immigrant.

Pair Formation: We consider a monogamous system where reproductive pairs randomly form within each site and produce broods of full siblings. The number of pairs at a site is the lesser of the number of males and the number of females at the site.

Reproduction: After breeding pairs form, every pair produces f offspring, so fecundity represents the total number of fertilized eggs (larvae) produced by each female. At birth, each larva is assigned a sex at random, i.e., it has a probability of 0.5 of being either male or female. We assume that f is an independent and identically distributed lognormal random variable with mean μ and standard deviation σ for each breeding pair (Ripley and Caswell 2006). In most simulations, fecundity is fixed (with $\mu = \ln 100$ and $\sigma = 0$). To explore the effect of variable reproductive success, we held μ at $\ln 100$ and varied the standard deviation. In all simulations, adults die after reproduction.

Dispersal: We model two forms of stochastic dispersal. In the first form, larvae disperse independently. A larva born in site j moves to site i with probability c_{ij} given by

$$c_{ij} = \begin{cases} \frac{1}{2}p(1-p)^{|j-i|}, & \text{if } i \neq j, \\ p, & \text{otherwise,} \end{cases} \tag{1}$$

with $0 < p \leq 1$. Here, p represents the probability of *staying* in a site, so as p increases, dispersal becomes more limited. For $p < 1$, there is a positive probability that a larva born in site j travels beyond the edge of the habitat and dies. The dispersal distributions of the remaining larvae are centered at site j. The shapes of these distributions are determined by p (Appendix S1: Fig. S1). When p is large, the probability of dispersal drops quickly

6

and monotonically with distance. For smaller p, the distributions are flatter and a significant fraction of larvae disperse out of the system. For completeness, we also consider the extreme case of a well-mixed population: here, the dispersal kernel is uniform within the study domain and there is an equal probability of dispersing to each of the m sites (i.e., $c_{ij} = 1/m$). Dispersal does not remove any larvae from the system in this case.

In the second form of dispersal, sibling larvae form groups that then disperse independently from other groups. Groups stay together for the duration of the larval phase and travel to a site based on the dispersal kernel (either uniform or Eq. 1). Offspring are split into groups through a sequence of binomial trials that determines the number of groups and their sizes. Let G_i be a random variable for the number of larvae in group i, and g_i be a realization of G_i. G_1 has a binomial distribution with f trials (unassigned larvae) and probability ψ (identical for all breeding pairs) that an individual larva will join the group. That is, $G_1 \sim B(f, \psi)$. This leaves $f - g_1$ unassigned larvae to potentially join the second group. Thus, $G_2 \sim B(f - g_1, \psi)$. This process continues until all larvae have been assigned. The dependence of the expected group size distribution on ψ is illustrated in Appendix S1: Fig. S2. For small ψ siblings disperse approximately independently from each other; when $\psi = 1$ all siblings disperse as one cohesive group.

Settlement: After dispersal, k individuals are chosen at random at each site; these individuals successfully settle and recruit to the next generation. Any additional larvae die, representing settlement mortality. If fewer than k larvae disperse to a given site, all of those larvae settle. As a final step, we randomly select 5% of successful settlers from across the entire population and replace them with unrelated immigrants. These immigrants can be thought of as rare long-distance dispersers into the relatively 'closed' focal population. They are necessary because in the absence of new unrelated individuals, pedigree-based kinship coefficients tend towards fixation ($\phi \to 1$) as the number of generations in the pedigree increases (Appendix S1: Fig. S4).

Modifying the breeding system

The core models used to explore kin structure formation assume a monogamous mating system (Fig. 1), yet real populations exhibit a range of breeding systems. To briefly explore the effect of an alternative mating system, we consider another extreme of the reproductive spectrum in marine populations: a form of broadcast spawning, wherein all males and females within a site release gametes that randomly fuse in the water column. The number of unique partners that an individual reproduces with is influenced by the site's sex ratio and carrying capacity. We assume that all eggs are fertilized as long as there is at least one male in the site.

Simulations

Limited dispersal, sibling cohesion, and variable reproductive success have each been proposed to explain empirical observations that kin are sometimes found close to one another after settlement. To quantify their effects, it is necessary to first establish a basis for comparison by running the model with none of these mechanisms turned on. In these simulations, breeding pairs produce the same number of larvae and those larvae disperse independently, with an equal probability of reaching any site (the 'well-mixed' scenario). Next, we compared these baseline simulations to simulations in which one mechanism operates but the other two do not. For example, to isolate the effects of limited dispersal, we turned off both sibling group formation and variability in reproduction (the 'limited dispersal' scenario). The same logic underlies the 'sibling cohesion' and 'variable reproductive success' scenarios. These simulations allowed us to explore each mechanism's independent effect on kin structure. In reality, however, these mechanisms may operate simultaneously; thus, after isolating each mechanism, we ran simulations where two or three mechanisms operated.

We initialized the environment with 21 sites, each containing 10 founding individuals. Founders were randomly assigned a sex and assumed to be unrelated. We simulated 30 generations, and sensitivity analyses revealed this was enough time for the population

to reach an equilibrium kinship value in environments of various sizes (Appendix S1: Fig. S5). Because population size can modulate the absolute value of kinship coefficients, we also explored larger site-level carrying capacities ($k = 20$ and 50) (Appendix S1: Fig. S6 and S8). Other parameter values were drawn from Table 1 in combinations that represent the mechanisms described above (note: to ensure our findings were generalizable, we tested a broad range of values for each parameter). For each set of parameter values, we ran 100 simulations. We tracked the entire pedigree over all generations, used the pedigree to calculate pairwise kinship coefficients between all individuals, and recorded pairwise kinship coefficients in the final generation based on a standard algorithm that accounts for the distribution of kinship coefficients in the preceding generations (*sensu* Lange 1997, Sinnwell et al. 2014). In brief, the algorithm uses an identity by descent approach for diploid individuals at autosomal loci. It is recursive, in that each individual is successively added to the kinship matrix, which enables inbreeding to be incorporated into kinship coefficients.

Results

Pedigree Structure

Limited dispersal, sibling cohesion, and variability in reproductive success have distinct effects on pedigree structure (Fig. 2). Here, we consider each pedigree to be a spatiotemporal network, where nodes represent individuals and links represent connections between parents and their offspring. An individual's vertical position represents its generation and its horizontal position represents its location along the one-dimensional environment. For visualization purposes, only females and maternal links are plotted.

In a well-mixed population, many mothers contribute offspring to the next generation and those offspring are distributed uniformly across space (Fig. 2a). In contrast, when dispersal is limited, many mothers contribute offspring to the next generation, but offspring remain close to their natal site. This leads to a pedigree with relatively vertical lineages (Fig. 2b). When larvae disperse as cohesive siblings, many mothers contribute offspring to the next generation, but they settle in groups. This creates a pattern where

sisters aggregate within sites (Fig. 2c). Finally, when there is large variability in reproductive success, few mothers contribute offspring to the next generation, and their offspring disperse uniformly across space. This generates a distinct shape, with very few mothers linked to most of the daughters in the subsequent generation (Fig. 2d). These examples demonstrate that limited dispersal, sibling cohesion, and variability in reproductive success have unique effects on the shape of the pedigree itself.

Kinship within a site

Next, we investigated kinship between individuals living within a site. This represents the scale at which most empirical studies are conducted. In our model, a site is defined as the range within which individuals find mates, so it is also the scale relevant to kin interactions. We define $\bar{\phi}_{site}$ as the expected within-site kinship coefficient at equilibrium.

Expected within-site kinship ($\bar{\phi}_{site}$) increases as dispersal becomes more limited (Fig. 3a). The relationship between $\bar{\phi}_{site}$ and dispersal is nonlinear; as the probability of staying in a site tends to one, the rate of change in $\bar{\phi}_{site}$ increases. As the strength of sibling cohesion increases, $\bar{\phi}_{site}$ also increases monotonically (Fig. 3b). However, the maximum within-site kinship value predicted by cohesion is nearly three times smaller than the maximum value predicted by limited dispersal. This occurs because when siblings disperse as cohesive groups, they mix with other unrelated groups at their destination and compete for settlement space. In contrast, when individuals stay in their natal site, sites become isolated, leading to an increase in inbreeding. These patterns were consistent across populations with different carrying capacities, but population size attenuated the effects of limited dispersal and cohesion on the magnitude of $\bar{\phi}_{site}$ (Appendix S1: Fig. S6).

Expected within-site kinship also increases as the standard deviation in fecundity increases (Fig. 3c). Variable reproductive success predicts the largest values of $\bar{\phi}_{site}$. Because fecundity is modeled as a lognormal random variable, one or a few females dominate the larval output in most simulations, though occasionally fecundity is more evenly distributed. The effect of variable reproductive success on $\bar{\phi}_{site}$ remained strong even as car-

10

rying capacity increased (Appendix S1: Fig. S6).

These results highlight the challenge of disentangling the drivers of within-site kinship, even when they are modeled in isolation: different mechanisms can predict the same within-site kinship value (Fig. 3a-c). Our focus here is the mean of the kinship distribution, but the variance is also of interest. The coefficient of variation ($\frac{\sigma^2}{\mu} \times 100$) represents the relative variability in within-site kinship across models and demonstrates that the rate of increase in the mean exceeds the rate of increase in the variance (although the greatest relative variability occurs at intermediate values for the dispersal limitation and cohesion models) (Fig. 3d-f). Notably, while variance in pairwise kinship increases monotonically along with mean pairwise kinship in all models, the magnitude of the variance is always small (Appendix S1: Fig. S7), suggesting that observed kinship values will tend to be close to the mean. Taken together, these results demonstrate that the mechanistic inferences that can be drawn from kinship data collected at a single spatial location, at a single point in time, are limited.

Kinship at increasing distances

Next, we explored kinship between individuals at increasing distances apart to quantify the complete pattern of kin structure, and found that each mechanism generates a distinct spatial pattern (Fig. 4a-c). When dispersal is limited, expected kinship declines as the distance between individuals increases, and this decline directly reflects the shape of the dispersal kernel (Fig. 4a). When the probability of dispersal drops quickly as a function of distance from source (e.g., $p = 0.9$), expected kinship also drops quickly as a function of distance between individuals. At the other extreme, when the kernel is nearly flat (e.g., $p = 0.1$), expected kinship is relatively constant across distance classes. Thus, as long as dispersal is limited to some degree, kinship will decline as the distance between individuals increases.

In contrast, sibling cohesion has a relatively strong effect on kinship within sites, compared to kinship between sites. In other words, kinship increases markedly when the

distance between individuals is zero, and is relatively constant for individuals living in different sites, regardless of how far apart the sites are (Fig. 4b). The magnitude of the relative increase in kinship within sites depends on the strength of the cohesion parameter. However, the baseline level of kinship at distances greater than 0 also increases when cohesion occurs due to increased inbreeding.

Finally, fecundity variability generates consistent expected kinship values across space for a given set of parameters (Fig. 4c). It has a strong effect on the magnitude of expected kinship (both within and between sites), but no spatial effect on kinship. Collectively, these simulations reveal that dispersal, sibling cohesion, and variable reproductive success are predicted to have markedly different effects on spatial patterns of kin structure when they operate in isolation. Although expected kinship values decrease as carrying capacity increases, the overall predicted kin structure patterns remain consistent (Appendix S1: Fig. S8).

When multiple mechanisms operate simultaneously, their individual predicted spatial patterns are still apparent (Fig. 4d-f). For example, the interaction between sibling cohesion and limited dispersal results in increased within-site kinship, i.e., the cohesion effect, along with a decline in kinship as the distance between individuals increases, i.e., the limited dispersal effect (Fig. 4d). Turning on fecundity variability and limited dispersal increases the magnitude of expected kinship, but the spatial pattern still reflects the dispersal kernel (Fig. 4e). When all three mechanisms operate, the spatial effects of both limited dispersal and cohesion can still be observed (Fig. 4f).

Effect of mating system

All else equal, the form of broadcast spawning that we implemented reduces mean within-site kinship compared to monogamous mating (Fig. 5a). The reduction in kinship is greatest when dispersal is strongly limited (i.e., when p is high). Broadcast spawning also reduces kinship between individuals separated in space (Fig. 5b), and this reduction is smaller at large distances compared to small distances. These results are consistent across differ-

ent carrying capacities (Appendix S1: Fig. S9 and S10), revealing that this form of broadcast spawning has a predictable effect on expected kinship relative to monogamous mating: kinship declines at all spatial scales, with the magnitude of the change mediated by carrying capacity and the dispersal kernel parameter p.

Discussion

Widespread larval dispersal was once assumed to disrupt kin associations and diminish the number of siblings found close to each other after settlement, but recent advances in molecular ecology have fueled an interest in the formation of marine kin structure. We developed a general model for a population with a simple biphasic life cycle, and demonstrated how three commonly invoked mechanisms – limited dispersal, sibling cohesion, and variability in reproductive success – are predicted to affect the relative magnitude and spatial pattern of kinship. Our central finding is that these mechanisms are only distinguishable when individuals are sampled at multiple spatial scales. This general finding was insensitive to the specific parameter values tested. These results have important implications for empirical work; mainly, any metric of kinship measured within a single site (e.g., the absolute number of sibling pairs or mean pairwise relatedness) cannot be used to accurately infer the larval process that generated it.

Linking pattern and process, at the appropriate spatial scale, is a fundamental topic in ecology (Levin 1992). In the marine environment, patterns of first-order relatives nearby in space have prompted inferences about generating processes during the larval phase – a phase that is consequential to the ecology of marine populations, but that is not directly observable. Marine ecologists have long sought answers to questions about the dispersive phase (e.g., *How far do larvae go? How do they behave as they travel? Who are they dispersing with?*) and kinship data may provide new insights into these questions. While the consequences of kin structure for inbreeding and inclusive fitness are also of great interest (Hamilton 1964, Crnokrak and Roff 1999), our focus here is building a theoretical understanding of how alternative mechanisms give rise to kinship patterns, which

may, in turn, guide the collection of empirical data and clarify its interpretation.

Limited dispersal as a foundational explanation of marine kin structure

A baseline mechanism for marine kin structure should arguably be dispersal limitation. Our simulations revealed that even when dispersal is relatively extensive, kinship increases within sites (Fig. 3a), and some siblings are predicted to be found nearby. By exploring kinship over increasing distances, we also found that the pattern of kin structure directly reflects the shape of the dispersal kernel (Fig. 4a), consistent with allelic-based population genetic models (Malécot 1969). These results generate the testable prediction that, unless the dispersal kernel is approximately uniform, the probability that any two individuals are first-degree kin should decrease as the distance between them increases. Although these findings rest on the assumption that dispersal is equally likely in both directions, enforcing directional dispersal to reflect a prevailing current does not change the predicted pattern of kin structure as long as the functional form of the kernel is consistent (Appendix S1: Fig. S11). Given the ubiquity of dispersal across marine taxa, we suggest that the level of kinship predicted by dispersal, over a specified spatial scale, should be used as a null expectation when interpreting empirical data. Such an approach requires an estimate of the dispersal kernel, which is challenging for marine species, but may become more accessible if reliable dispersal estimates can be extracted from population genetic data (Pinsky et al. 2017).

The scaling of larval dispersal has been of interest for decades, and empirical estimates suggest that dispersal capacity is heterogeneous across marine taxa (Kinlan and Gaines 2003). Nevertheless, dispersal is almost certainly limited to some degree in natural populations, and can be quite limited. For instance, ascidians with short lecithotrophic larval phases, such as *Didemnum molle*, may travel only tens of meters (Olson 1985). Site-attached reef fishes with longer planktotrophic larval phases also exhibit short-distance dispersal. The neon goby *Elacatinus lori*, for example, disperses for approximately 26 days, but most individuals travel less than 2 km (D'Aloia et al. 2015). In such viscous popula-

tions, we expect to find many relatives in close spatial proximity.

The potential consequences of collective sibling dispersal

The hypothesis of collective dispersal in the marine environment has gained traction in both theoretical (Broquet et al. 2013, Burgess et al. 2018), and empirical (Selwyn et al. 2016, Riquet et al. 2017) research. Generally, collective dispersal refers to correlated dispersal paths without specific reference to kinship. In this manuscript, because mating was monogamous in our core models, we focused on sibling cohesion as a specific type of collective dispersal, and our models revealed two interesting results. First, the effect of cohesion on within-site kinship was weak: even when siblings dispersed as completely cohesive groups, the expected kinship value was equivalent to that predicted by intermediate parameter values for dispersal or fecundity (Fig. 3a-c). Thus, in the absence of dispersal limitation or strong demographic variability, extreme collective dispersal had a relatively small impact on site-level kinship because multiple groups ultimately mixed at settlement. Group mixing at settlement was a construct of our model, but it aligns with predictions from some particle-tracking models that show how oceanographic features can accumulate larvae from many sources before settlement (Harrison et al. 2013). The second finding is that sibling cohesion creates a distinct kin structure pattern, with a strong relative increase in kinship within sites, compared to between sites. This prediction could be used to empirically test for cohesion by asking whether the number of siblings within a site exceeds expectations based on dispersal.

These ecological findings complement recent population genetic models that illustrate how collective dispersal can increase spatial genetic structure between sites (Broquet et al. 2013, Yearsley et al. 2013, Eldon et al. 2016). Genetic drift modulates the effect of collective dispersal on structure, such that it is strongest in small populations and/or when reproductive skew is high (Yearsley et al. 2013). Similarly, we found that cohesion effects on kinship are strongest in small populations, and that interactions between cohesion and other processes such as fecundity variability lead to the highest within-site kinship coef-

ficients (Fig. 4f). The flexibility of our demographic approach enabled us to add further ecological complexity, e.g., by allowing dispersal probabilities to vary with the distance between sites, we showed that cohesion should be detectable even when dispersal is limited (Fig. 4d). Together, these studies provide insight into both the ecological and evolutionary effects of non-independent larval dispersal.

Future empirical work may elucidate whether some closely related larvae actually stick together throughout the entire dispersive phase. Although point estimates have identified putative siblings within a larval pool (Riquet et al. 2017), combined genetic-otolith microchemistry analyses revealed that individuals identified as cohesive larvae were not closely related (Ben-Tzvi et al. 2012). Our point is not that sibling cohesion is impossible; rather, we note that the proportion of siblings that spend the entire larval phase together remains unclear, and will likely be influenced by spatiotemporal patterns of larval release, hydrodynamics, and larval behavior. Further, our models suggest that detecting some sibling pairs within a recruitment pulse, or an estimated age class post-settlement, should not be interpreted as *direct* evidence of collective dispersal because such pairs could also be generated through limited dispersal or strong fecundity variability.

Variable reproductive success increases post-settlement kinship

Variation in demographic traits among adults can also contribute to the formation of kin structure by altering the makeup of the larval pool. We found that, of the three mechanisms tested, extreme variability in fecundity had the strongest effect on within-site kinship (Fig. 3c); this effect was sustained across space (Fig. 4c); and it remained strong even when population size increased (Appendix S1: Fig. S6). While our model implements variability in fecundity, it was inspired by the more general sweepstakes reproduction model, which posits that relatively few individuals can produce a majority of successful recruits due to chance (Hedgecock and Pudovkin 2011). Some simplifications in our model could affect our findings. For example, we modeled non-overlapping generations with a single reproductive event, but in real marine populations there are often overlapping genera-

tions and multiple reproductive events per mating season. Therefore, if the 'lucky' adults vary over time, reproductive success may be more evenly distributed across the cumulative reproductive season, as evidenced recently in a brooding mollusc (Riquet et al. 2017). These factors would all weaken the strong effect of variable reproductive success on kinship, which was predicted by our model.

Effect of mating system

While the monogamous mating system we adopt in this model is an approximation of pair mating in some species, particularly fishes (Whiteman and Côté 2004), diverse mating systems are observed across marine taxa (Thresher 1984, Carlon 1999, Wong and Buston 2013). Alternative mating systems could impact kinship through two mechanisms. First, they can change the proportion of different relationships within larval cohorts. While monogamous pairs will produce clutches of full siblings, other systems such as broadcast spawning, brooding, or polygynandry can produce a mix of full siblings, half siblings, and non-relatives, depending on the relative frequency of unique mate pairings. All else equal, as the proportion of full siblings in the larval pool declines, kinship will also decline. This effect was evident when we implemented a form of broadcast spawning, in which the gametes produced by all individuals in a site fused randomly (Fig. 5).

A second mechanism by which mating system can affect kinship is through increased variability in reproductive success across individuals. For example, in haremic systems one male (or female) has multiple mates, creating a skew in reproductive success based on factors such as harem size and reproduction frequency. This effect is implicitly captured in our simulations of skewed fecundity, and it affects the magnitude of kinship, but not the pattern of kin structure (Fig. 4c). Thus, unless mating system or reproductive mode is linked to another mechanism that affects spatial kin structure (e.g. clonal reproduction linked to limited dispersal [Jackson 1986]), its independent effect will manifest in the overall magnitude of kinship, but not the pattern of kin structure.

17

Implications for empirical sampling

Before translating these findings into suggestions for empirical sampling, it is important to acknowledge the model's simplifying assumptions and limitations. For example, we tracked all surviving individuals, enabling us to reconstruct complete pedigrees and determine how each mechanism affects pedigree shape (Fig. 2); such a comprehensive description of family lineages will be unattainable in large marine populations. Our assumption of non-overlapping generations also limits the observable relationships to intragenerational siblings and cousins. Real populations with overlapping generations can also contain intergenerational relationships such as parent-offspring, grandparent-grandchild, and avuncular pairs. We also focused exclusively on three mechanisms, while other factors could certainly play a role (e.g., dispersal costs; post-settlement selection; larval behavior). Still, the core message of this paper has important implications for empirical studies as interest in understanding marine kin aggregations continues to grow: an absolute metric of kinship within a single site can be explained by several hypotheses. Therefore, if the goal is to infer process(es) from kinship patterns, a continuously-sampled transect, scaled to the species' movement capabilities, will be more informative than data from a single site.

Our focus here is spatial, but if species have overlapping generations or a protracted reproductive season, the temporal dimension of sampling may also be important. Indeed, many empirical studies have addressed time by comparing relatedness within and between recruitment cohorts to test for kinship within groups of recently-settled individuals (e.g., Selkoe et al. 2006). Sampling at or near settlement could provide a better estimate of kinship between similarly-aged individuals before any effects of post-settlement processes accumulate. However, the same limitations apply: one must show that the number of putative siblings exceeds expectations from a null model. If feasible, spatiotemporal sampling strategies that explore kinship along continuous spatial transects, over multiple time points, would shed light on the stability of kin structure and the mechanisms that drive it.

Conclusions

This study lays a conceptual foundation for understanding of how commonly-invoked demographic and behavioral processes impact the formation of kin structure in the sea. Using an individual based modeling approach, we delineate how these mechanisms are predicted to affect spatial patterns of kinship. By testing a broad range of parameter values, we produce the general finding that these mechanisms have markedly different spatial effects on kin structure, which are only observable when individuals are sampled at multiple spatial scales. Based on our results, we suggest that continuous, multi-scale sampling strategies are better equipped than single-site strategies for disentangling these hypotheses, and that the scale of dispersal should be used as a null model for interpreting empirical kinship data. While the mechanisms tested were motivated by marine ecological theory, these same considerations hold for terrestrial species with similar traits. Moving forward, we hope these model predictions inform empirical data collection in this burgeoning field, and facilitate the integration of theory and data to provide new insights into how kin structure forms in diverse marine populations.

Literature Cited

Ben-Tzvi, O., A. Abelson, S. D. Gaines, G. Bernardi, R. Beldade, M. S. Sheehy, G. L. Paradis, and M. Kiflawi. 2012. Evidence for cohesive dispersal in the sea. PLoS One 7:e42672.

Broquet, T., F. Viard, and J. M. Yearsley. 2013. Genetic drift and collective dispersal can result in chaotic genetic patchiness. Evolution 67:1660–1675.

Burgess, S. C., R. E. Snyder, and B. Rountree. 2018. Collective dispersal leads to variance in fitness and maintains offspring size variation within marine populations. The American Naturalist 191:000–000.

Carlon, D. B. 1999. The evolution of mating systems in tropical reef corals. Trends in Ecology & Evolution 14:491–495.

Crnokrak, P., and D. A. Roff. 1999. Inbreeding depression in the wild. Heredity 83:260–270.

D'Aloia, C. C., S. M. Bogdanowicz, R. K. Francis, J. E. Majoris, R. G. Harrison, and P. M. Buston. 2015. Patterns, causes, and consequences of marine larval dispersal. Proceedings of the National Academy of Sciences 112:13940–13945.

Eldon, B., F. Riquet, J. Yearsley, D. Jollivet, and T. Broquet. 2016. Current hypotheses to explain genetic chaos under the sea. Current Zoology 62:551–566.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. Journal of Theoretical Biology 7:1–16.

Harrison, C. S., D. A. Siegel, and S. Mitarai. 2013. Filamentation and eddy-eddy interactions in marine larval accumulation and transport. Marine Ecology Progress Series 472:27–44.

Hedgecock, D., and A. I. Pudovkin. 2011. Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. Bulletin of Marine Science 87:971–1002.

Iacchei, M., T. Ben-Horin, K. A. Selkoe, C. E. Bird, F. J. García-Rodríguez, and R. J. Toonen. 2013. Combined analyses of kinship and FST suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. Molecular Ecology 22:3476–3494.

Jackson, J. B. C. 1986. Modes of dispersal of clonal benthic invertebrates: consequences

for species' distributions and genetic structure of local populations. Bulletin of Marine Science 39:588–606.

Johnson, M., and R. Black. 1982. Chaotic genetic patchiness in an intertidal limpet, Siphonaria sp. Marine Biology 70:157–164.

Kamel, S. J., and R. K. Grosberg. 2013. Kinship and the evolution of social behaviours in the sea. Biology Letters 9:20130454.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. Trends in Ecology & Evolution 17:230–241.

Kimura, M., and G. H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576.

Kinlan, B. P., and S. D. Gaines. 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology 84:2007–2020.

Lange, K. 1997. Mathematical and statistical methods for genetic analysis. Springer-Verlag, New York.

Leis, J. M. 1991. The pelagic stage of reef fishes: the larval biology of coral reef fishes. Pages 183–230 *in* P. F. Sale, ed. The ecology of fishes on coral reefs, 1st ed. Acadmic Press, San Diego.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943–1967.

Malécot, G. 1969. The mathematics of heredity. W.H. Freeman and Company, San Francisco.

Olson, R. R. 1985. The consequences of short-distance larval dispersal in a sessile marine invertebrate. Ecology 66:30–39.

Pinsky, M. L., P. Saenz-Agudelo, O. C. Salles, G. R. Almany, M. Bode, M. L. Berumen, S. Andréfouët, S. R. Thorrold, G. P. Jones, and S. Planes. 2017. Marine dispersal scales are congruent over evolutionary and ecological time. Current Biology 27:149–154.

Queller, D. C. 1994. Genetic relatedness in viscous populations. Evolutionary Ecology 8:70–73.

Ripley, B. J., and H. Caswell. 2006. Recruitment variability and stochastic population growth of the soft-shell clam, Mya arenaria. Ecological Modelling 193:517–530.

Riquet, F., T. Comtet, T. Broquet, and F. Viard. 2017. Unexpected collective larval dispersal but little support for sweepstakes reproductive success in the highly dispersive brooding mollusc Crepidula fornicata. Molecular Ecology 26:5467–5483.

Salles, O. C., B. Pujol, J. A. Maynard, G. R. Almany, M. L. Berumen, G. P. Jones, P. Saenz-Agudelo, M. Srinivasan, S. R. Thorrold, and S. Planes. 2016. First genealogy for a wild marine fish population reveals multigenerational philopatry. Proceedings of the National Academy of Sciences 113:13245–13250.

Selkoe, K. A., O. E. Gaggiotti, ToBo Laboratory, B. W. Bowen, and R. J. Toonen. 2014. Emergent patterns of population genetic structure for a coral reef community. Molecular Ecology 23:3064–3079.

Selkoe, K. A., S. D. Gaines, J. E. Caselle, and R. R. Warner. 2006. Current shifts and kin aggregation explain genetic patchiness in fish recruits. Ecology 87:3082–3094.

Selwyn, J. D., J. D. Hogan, A. M. Downey-Wall, L. M. Gurski, D. S. Portnoy, and D. D. Heath. 2016. Kin-aggregations explain chaotic genetic patchiness, a commonly observed genetic pattern, in a marine fish. PLOS ONE 11:e0153381.

Shima, J. S., and S. E. Swearer. 2016. Evidence and population consequences of shared larval dispersal histories in a marine fish. Ecology 97:25–31.

Sinnwell, J. P., T. M. Therneau, and D. J. Schaid. 2014. The kinship2 R package for pedigree data. Human heredity 78:91–93.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264–279.

Thresher, R. 1984. Reproduction in reef fishes. T.F.H. Publications, Neptune City.

Victor, B. C. 1984. Coral reef fish larvae: patch size estimation and mixing in the plankton. Limnology and Oceanography 29:1116–1119.

Whiteman, E. A., and I. M. Côté. 2004. Monogamy in marine fishes. Biological Reviews

79:351–375.

Wong, M. Y. L., and P. M. Buston. 2013. Social systems in habitat-specialist reef fishes. BioScience 63:453–463.

Wright, S. 1943. Isolation by distance. Genetics 28:114–138.

Yearsley, J. M., F. Viard, and T. Broquet. 2013. The effect of collective dispersal on the genetic structure of a subdivided population. Evolution 67:1649–1659.

Tables

Table 1: Parameters and values used in the model

Notation	Description	Values
g	number of generations	30
m	number of sites	21
k	carrying capacity per site[*]	10, 20, 50
μ	mean number of larvae produced by each breeding pair	$\ln 100$
σ	standard deviation of number of larvae produced by each breeding pair[†]	1, 1.125, 1.25 ... 2.5
p	probability of leaving a site[‡]	0.05, 0.1, 0.15 ...1
ψ	strength of larval cohesion	0.05, 0.1, 0.15 ...1

Notes: [*] We present $k = 10$ in the main text, but see Appendix S1 for simulations with larger values. [†]In most simulations, fecundity is fixed (i.e., $\sigma = 0$); however, when testing variable reproductive success, individual fecundity is assigned based on a lognormal distribution. [‡] This parameter controls the shape of the dispersal kernel.

Figure Captions

Figure 1. Sequential steps of the model in a focal site. (1) Monogamous pairs form; here, the grey male did not find a mate due to an unequal sex ratio. (2) Depending on the model, pairs produce a fixed or variable number of offspring. (3) Adults die immediately after reproducing. (4) Larvae produced by a breeding pair in the focal site disperse to other sites, either independently or cohesively, depending on the model. (5) A larval pool forms above the focal site based on dispersal from all sites. Only a fraction of those larvae settle, and a small percentage of settlers may be unrelated immigrants from other populations. (6) Successful settlers (no more than k) mature into the next generation of adults. These processes are simulated in all sites simultaneously.

Figure 2. Effects of dispersal, sibling cohesion, and variable reproductive success on pedigree shape. For visualization, maternal links over the first five generations, after the founding generation, are plotted for a single simulation. (a) A well-mixed population; (b) Limited dispersal ($p = 0.9$); (c) Larval cohesion ($\psi = 1$); and (d) Variable reproductive success ($\sigma = 2.5$). These pedigrees are spatial, such that in each generation, individuals nearby in the pedigree are also nearby in space.

Figure 3. Kinship within a site. *Top row:* Expected kinship within a site ($\bar{\phi}_{site}$) as a function of the degree of (a) limited dispersal, (b) sibling cohesion, and (c) variable reproductive success. *Bottom row:* The relative variability in within-site kinship changes as a function of (d) limited dispersal, (e) sibling cohesion, and (f) variable reproductive success. Across all panels, per-site carrying capacity is 10. Violin plots (light gray) show kernel density estimates of either $\bar{\phi}_{site}$ or CV_{site} for 100 simulations; black dots represent medians. Note the different y-axis scaling.

Figure 4. Kinship between individuals separated by increasing distances. *Top row:* The effect of three independent mechanisms: (a) limited dispersal, (b) sibling cohesion, and (c) variable reproductive success when per-site carrying capacity is 10. *Bottom*

row: The effect of multiple mechanisms: (d) limited dispersal + cohesion, (e) limited dispersal + variability in reproductive success, and (f) all three mechanisms. In the bottom row, limited dispersal was set to $p = 0.5$, and the gray dots represent the dispersal-only scenario for comparison. Note the different y-axis scaling.

Figure 5. Effect of monogamous mating versus broadcast spawning. (a) Expected within-site kinship and (b) Expected kinship between individuals separated by increasing distances under two mating systems and with different degrees of dispersal limitation.
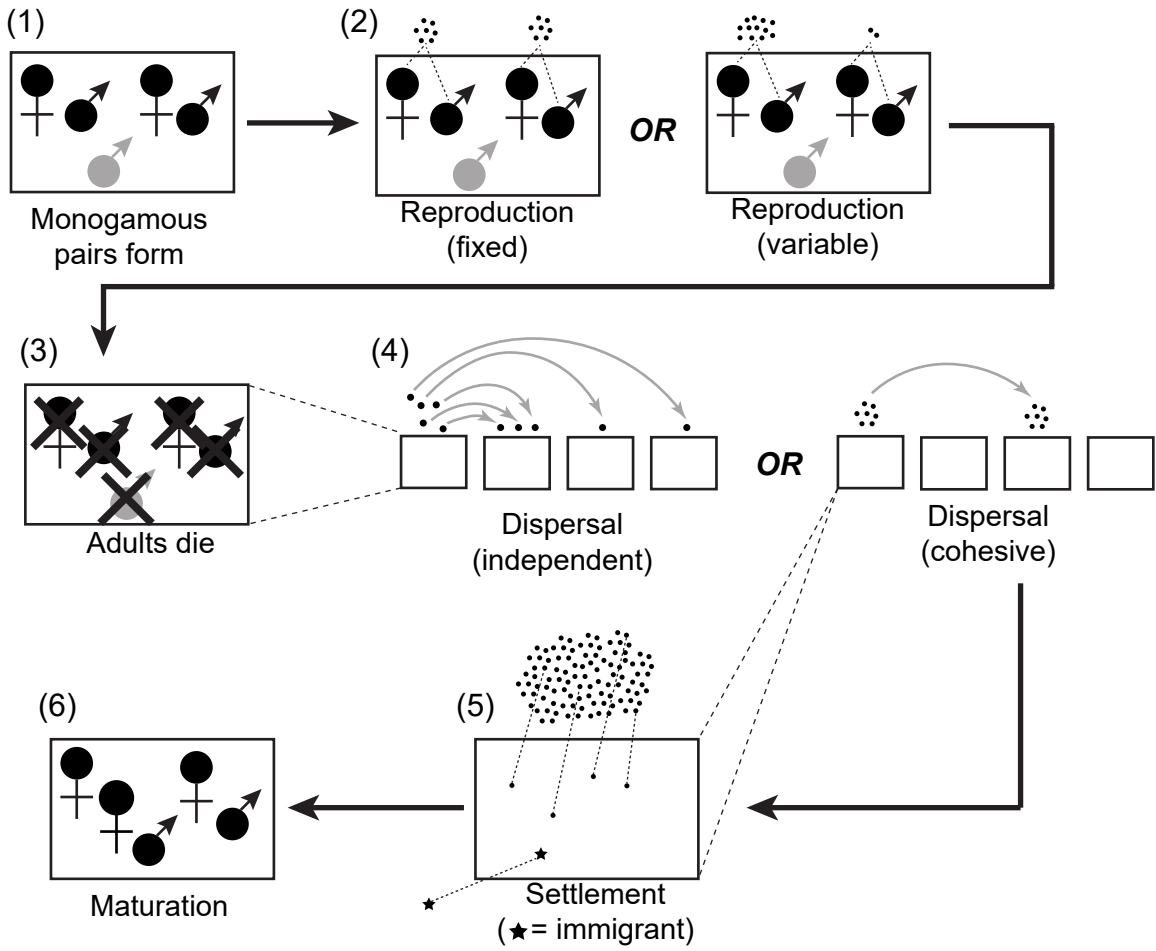
Figure 1:
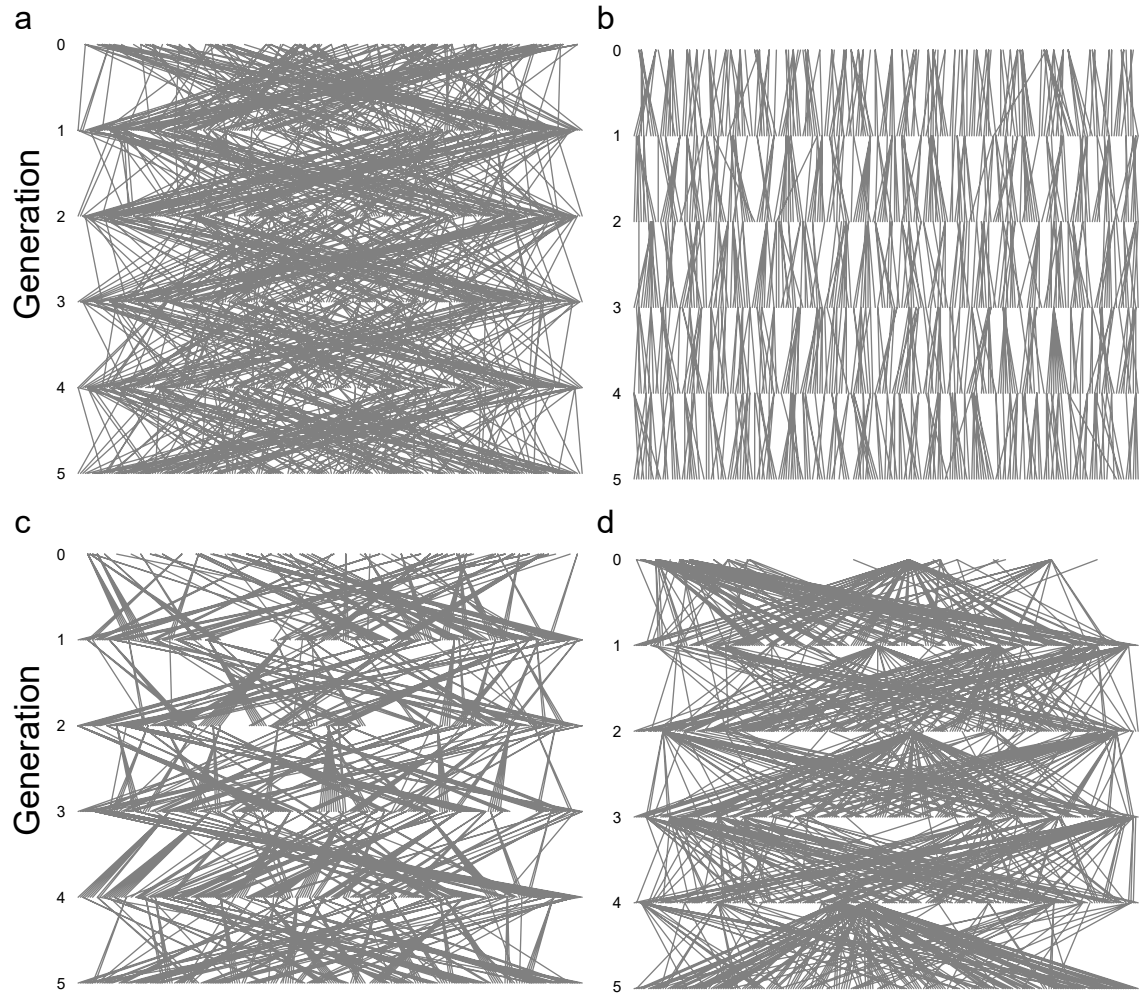
Figure 2:
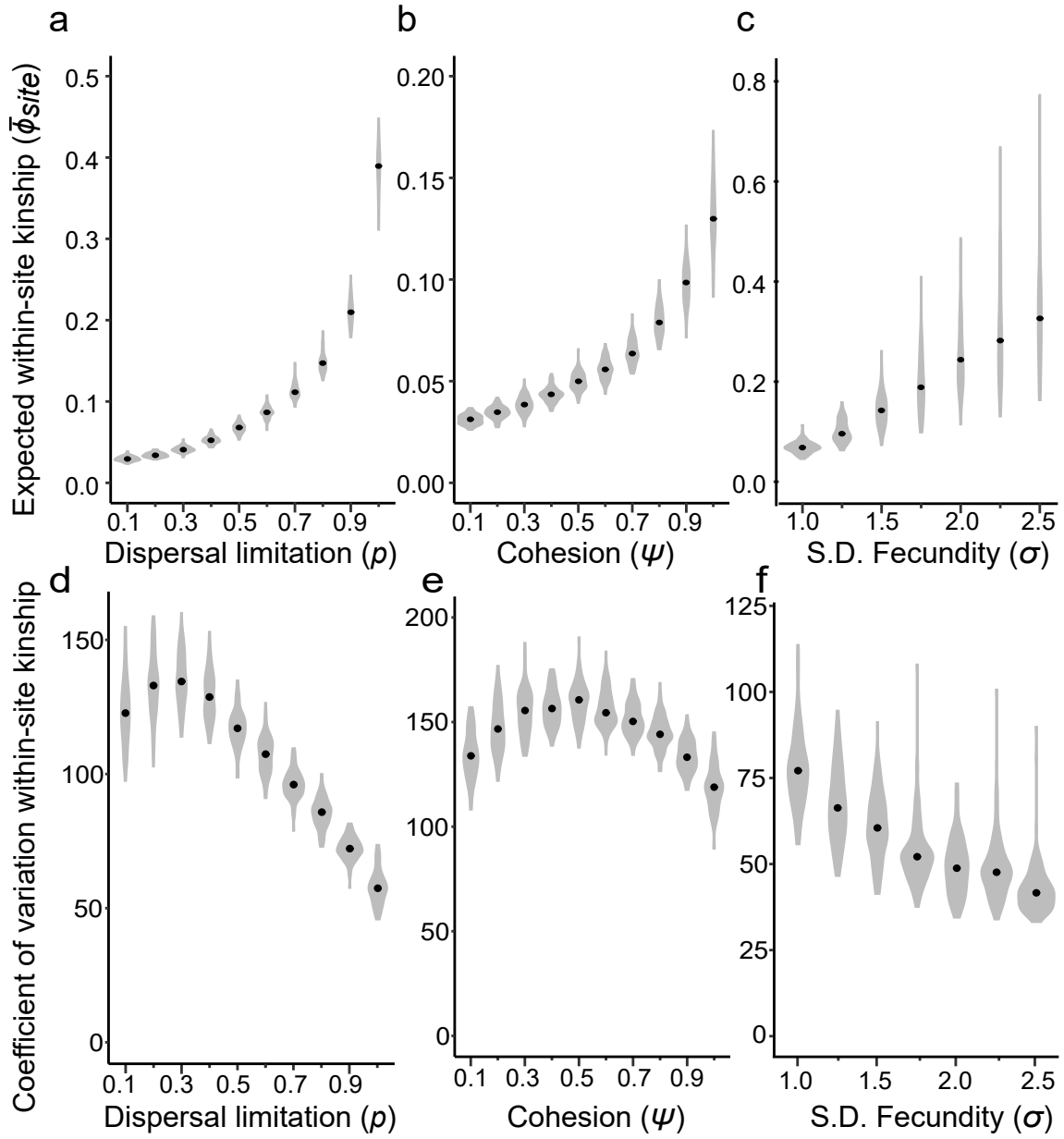
Figure 3:
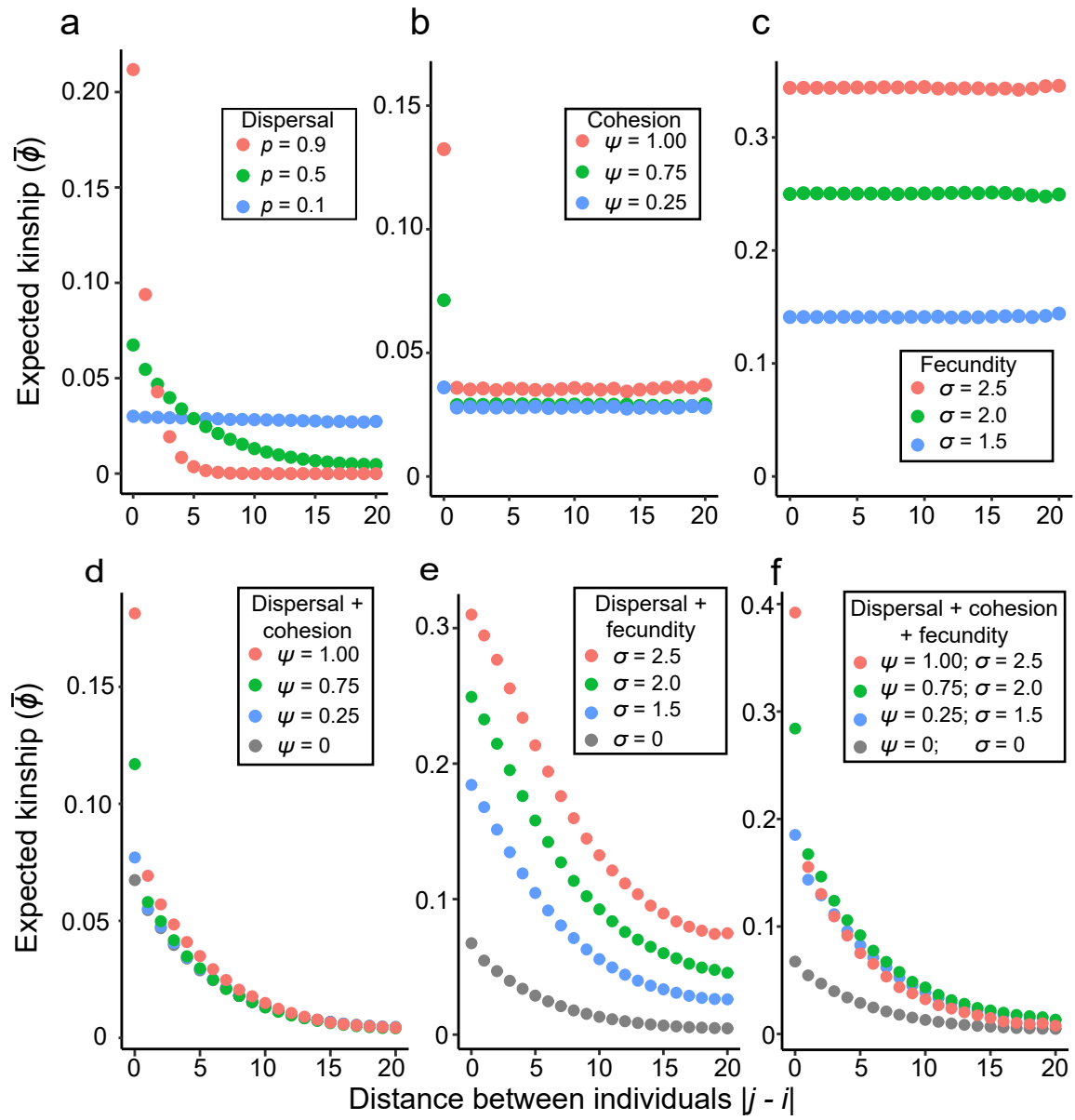
Figure 4:
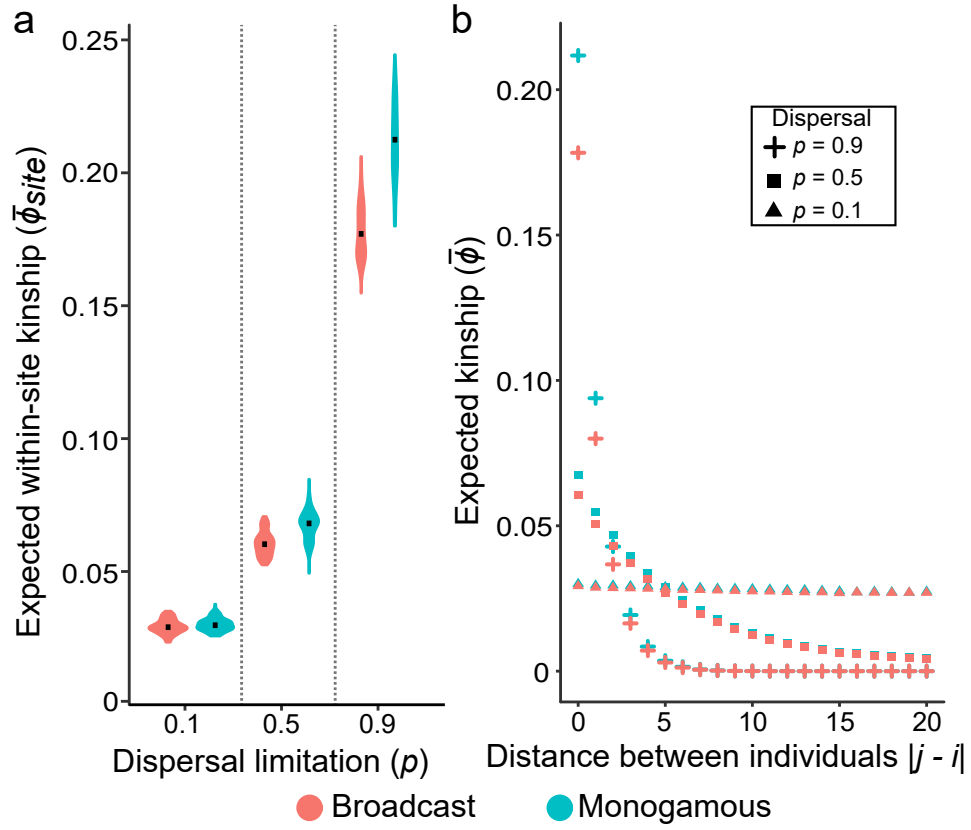
Figure 5:

Appendix S1

The formation of marine kin structure: effects of dispersal, larval cohesion, and variable reproductive success.

D'Aloia CC and Neubert MG

S1.1 Numerical Implementation of the Models

In this appendix we describe the numerical implementation of the three models: the limited dispersal model, the sibling cohesion model and the variable reproductive success model. In each model, we begin by generating the connectivity matrix $\mathbf{C}$. The columns of $\mathbf{C}$, denoted as $\mathbf{c}_{.j}$, are the dispersal kernels centered on site j. The elements of $\mathbf{C}$, denoted as c_{ij}, are the probability that a larva born in site j moves to site i. When dispersal is well-mixed, $c_{ij} = 1/m$. When dispersal is limited, c_{ij} is calculated from Eq. 1 (cf., main text). As an example, we illustrate the difference between well-mixed and limited dispersal when p, the probability of *staying* in a site, is 0.5 (Fig. S1). As p increases, the kernel becomes more peaked, and as p decreases, the kernel becomes flatter.

For each model we also pre-allocate a data frame with a row for each of $k \times m \times (g + 1)$ individuals. This data frame includes individuals from a founding generation and g subsequent generations. Each individual begins with a unique identifier, a site identifier, the generation in which it was born, and a randomly assigned sex.

Limited Dispersal Model

In the limited dispersal model, the next step is to randomly join males and females within a site into monogamous breeding pairs in generation 0. The number of breeding pairs in site j is

$$q_j = \min \left(\#\text{females in site } j, \ \#\text{males in site } j \right). \tag{S1}$$

Each paired female takes her mate's identifier as an additional attribute. Unpaired females do not have a mate and therefore are not considered as potential mothers.
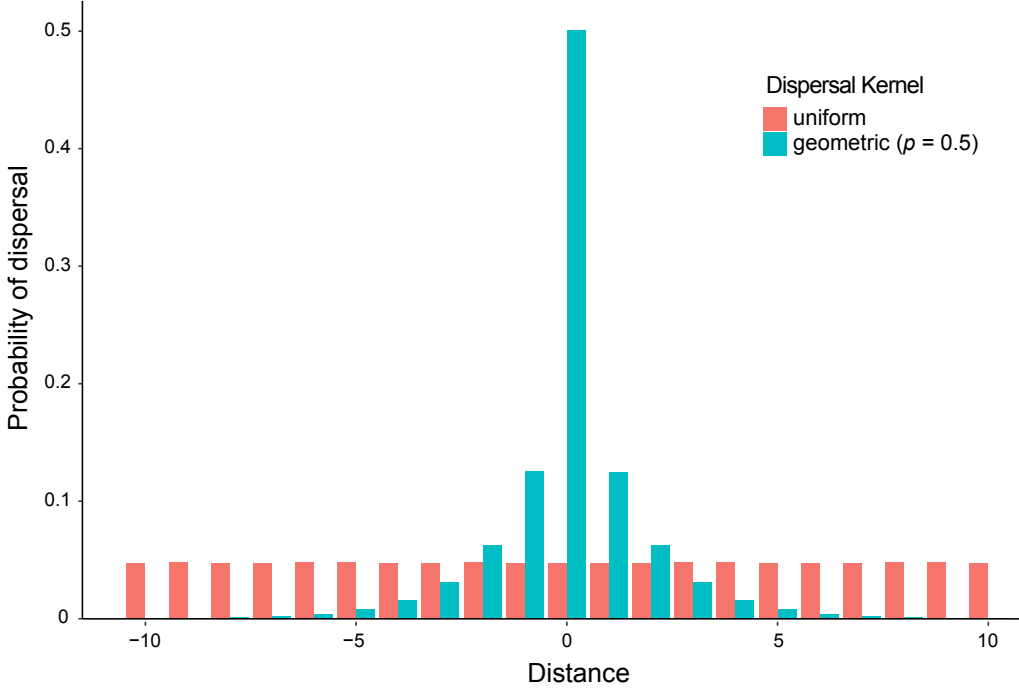
Figure S1: Probability of dispersing up to ten sites away, in two directions, when released from a focal site (distance = 0). A uniform and an incomplete back-to-back geometric dispersal kernel (Eq. 1, $p = 0.5$) are compared.

Next, we generate larval pools that 'hover' above each of the m sites after dispersal. Here, we create an $(m + 1) \times m$ matrix of larval immigrants $\mathbf{L}$. The elements of $\mathbf{L}$, ℓ_{ij}, give the number of larvae that arrived in site i that were born in site j. We treat dispersal as a stochastic process, so each column of $\mathbf{L}$, denoted as $\ell_{\cdot j}$, is an $(m + 1) \times 1$ multinomial random variable: $\ell_{\cdot j} \sim M(q_j f, \mathbf{c}_{\cdot j})$, where $q_j f$ is the pool of all offspring dispersing *from* site j and $\mathbf{c}_{\cdot j}$ are the probabilities that larvae travel from j to any site. The last row of $\mathbf{L}$ represents the number of larvae that disperse beyond the boundaries of the environment.

Next we choose k successful settlers randomly from the larval pool at each site. To do so, we create an $m \times m$ matrix of successful settlers $\mathbf{S}$. The elements of $\mathbf{S}$, s_{ij}, give the number of larvae that settle in site i that were born in site j. Because settlement is another stochastic process, the i-th row of $\mathbf{S}$, $\mathbf{s}_{i\cdot}$, is a multivariate hypergeometric random variable: $\mathbf{s}_{i\cdot} \sim MH(\ell_{i\cdot}, k)$.

Each settler must next be assigned to two parents. To assign a settler's parents, a breeding female from the settler's natal site is randomly chosen and assigned as the mother; her monogamous mate is assigned as the father. To prevent kinship fixation, we randomly select 5% of the settlers from the whole environment and replace them with unrelated immigrants.

Finally, all settlers mature into adults. Simulation continues with the formation of new breeding pairs from these adults.

Sibling Cohesion Model

In the sibling cohesion model, we again start by joining males and females within a site into monogamous breeding pairs as described in the limited dispersal model. In this model, however, we now track the dispersal of *groups* of siblings from distinct breeding pairs.

After mating pairs are formed, the offspring of each mated female are gathered into cohesive dispersal groups. The size of the first group is a binomial random variable, $G_1 \sim B(f, \psi)$, with f 'trials' (the number of offspring) and probability of success—i.e., successfully joining the group—equal to ψ. We take g_1 to be a realization of G_1. The second group is formed in a similar way, but with the number of trials reduced by g_1 (as long as $g_1 < f$); i.e., $G_2 \sim B(f - g_1, \psi)$. Continuing in this manner, the i-th group, G_i, is a random variable with distribution $B(u_i, \psi)$, where

$$u_i = f - \sum_{n=1}^{i-1} g_n \tag{S2}$$

is the number of unassigned larvae available to join group i. The process continues until all offspring are assigned to a group (possibly of size 1). The expected group size distribution as a function of ψ is illustrated in Fig. S2. At low values of ψ (e.g., $\psi = 0.25$), there are many singletons and the mean number of groups declines as group size increases. At larger values of ψ (e.g., $\psi = 0.5$ and 0.75), there are fewer singletons and multiple peaks in the mean number of groups. When $\psi = 1$, all offspring are placed into one cohesive group.
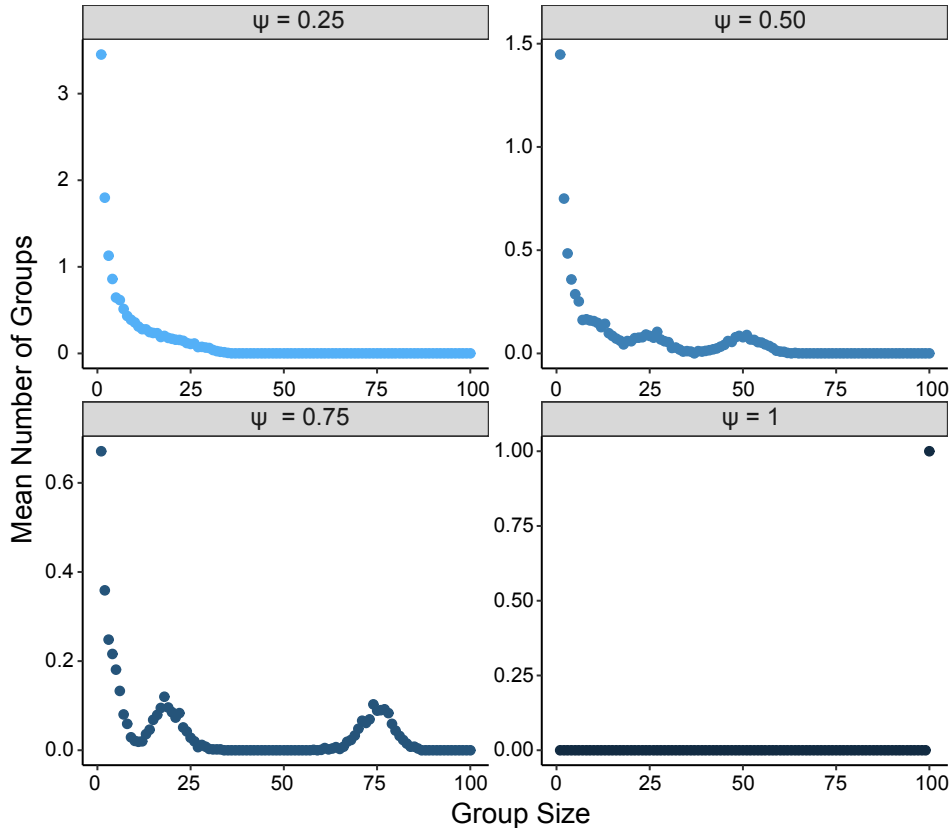
3

Figure S2: Mean number of groups of each size (from groups of 1 to 100 individuals) at different values of ψ. Means are calculated from 1,000 simulations.

Once all of a mother's f offspring are placed into cohesive groups, each group disperses independently with an equal probability of landing in any site. This process of sibling group formation and dispersal is repeated for every breeding female, and results in the the larval immigrant matrix $\mathbf{L}$. In this model, each column of $\mathbf{L}$, denoted as $\ell_{.j}$, gives the number of larvae that arrived in site i that were born to the j-th breeding female.

Next, at each site, k successful settlers are chosen from the larval pool as described in the limited dispersal model. The resulting matrix is $\mathbf{S}$; its elements s_{ij}, give the number of larvae that settle in site i who were born to the j-th breeding female.

Each settler is already assigned a mother. The mother's mate is assigned as the father. Finally, the introduction of unrelated immigrants and the formation of new breeding pairs occurs as described in the limited dispersal model.

Variable Reproductive Success Model

Variability in reproductive success is typically a multiplicative process that can be influenced by variation in parental fecundity as well as mortality at the larval, settlement and post-settlement stages. In our code, we implemented this variability as stochasticity in female fecundity. However, because our model does not directly track individuals from birth through settlement, we would have obtained identical results by alternatively implementing this variability in the processes of survival or settlement. In particular, we assumed that the fecundity of each female was an independent lognormal random variable with mean μ and standard deviation σ (Fig. S3).

Next, we generate the larval pool that forms at each site after dispersal by creating the matrix of larval immigrants $\mathbf{L}$. Each column of $\mathbf{L}$, denoted as $\ell_{\cdot j}$, is an $m \times 1$ multinomial random variable: $\ell_{\cdot j} \sim M(f_i, \mathbf{c}_{\cdot j})$, where f_i is the number of larvae produced by the i-th female, and $\mathbf{c}_{\cdot j}$ are the probabilities that her larvae travel from their origin site j to any site. In this model there is an equal probability of traveling to any site; i.e., $c_{ij} = 1/m$.

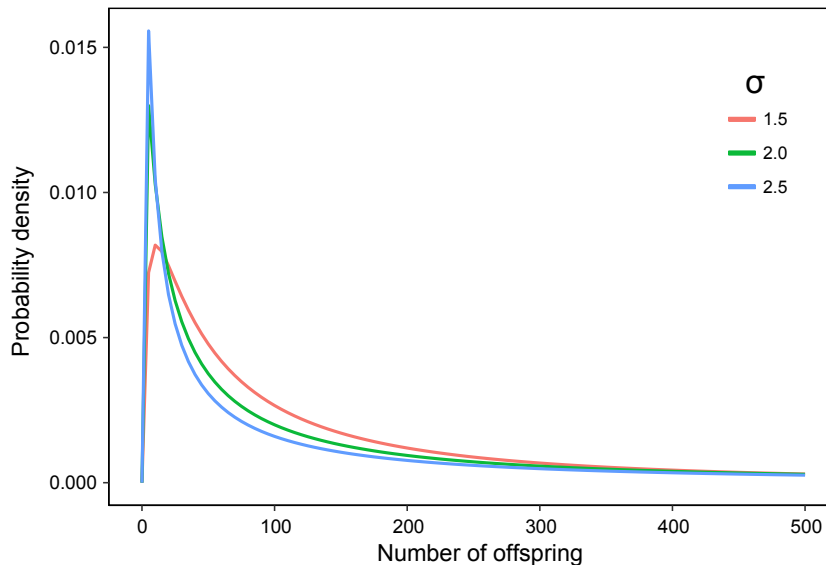The remaining steps proceed exactly as described in the sibling cohesion model.

Figure S3: Lognormal distributions for fecundity. We held μ at $ln(100)$ and varied σ. This held the median at 100 offspring as the standard deviation changed.

5

S1.2 Preventing kinship fixation

Pedigree-based metrics of coancestry tend towards fixation ($\phi \rightarrow 1$) as the depth of the pedigree increases. We introduced a low level of immigration in our model to prevent fixation. This approach is analogous to the introduction of mutation to prevent fixation in allelic-based models of coancestry. We explored whether the mean kinship coefficient stabilized over time when immigration was included.

Without any immigration, mean kinship increased linearly as a function of the number of generations in the model (Fig. S4a). When we included 5% immigration, mean kinship began to stabilize by 30 generations (Fig. S4b). Note that this time to equilibrium was consistent across other parameter values (see 'Time to Equilibrium').
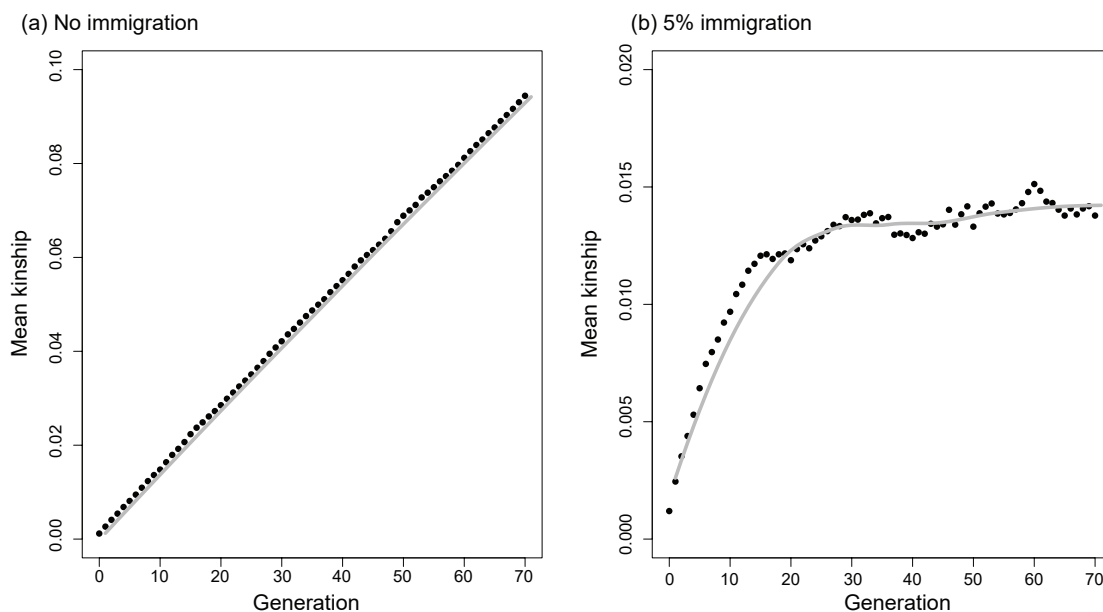
Figure S4: Mean population-level kinship per generation, when (a) there is no immigration; and (b) 5% of settlers per generation are unrelated immigrants. Black dots represent mean kinship, grey lines represent loess curves. Note the different y-axis scales. In these simulations, $m = 21$; $k = 20$; $p = 0.5$; $\psi = 0$; and $\sigma = 0$.

S1.3 Time to Equilibrium

Because the model is initialized with unrelated founders, we tested the number of generations required to reach an equilibrium kinship value in the population. We also explored whether population size affected the time to equilibrium, as the magnitude of kinship and inbreeding are mediated by population size. We tested the effect of increasing the number of sites m, and increasing the carrying capacity k per site while holding m constant. We found that kinship came to equilibrium by approximately 30 generations under all scenarios tested (Fig. S5). Adding more generations would not change the overall level of inbreeding, but would increase the amount of memory needed to run the model.
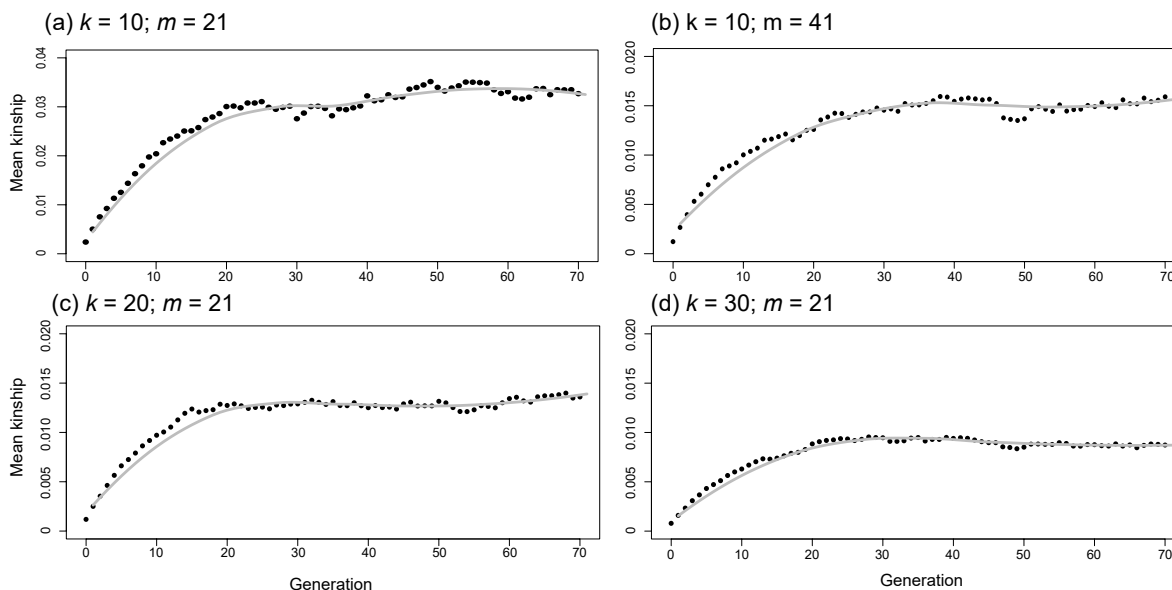
Figure S5: Mean kinship over time in (a) a small reference population; (b) a population with more sites; and (c-d) in populations with larger carrying capacities per site. Black dots represent mean kinship across the whole population, grey lines represent loess curves. For all simulations, we used a dispersal kernel (Eq. 1) with $p = 0.5$, and assumed fixed fecundity and no sibling cohesion.

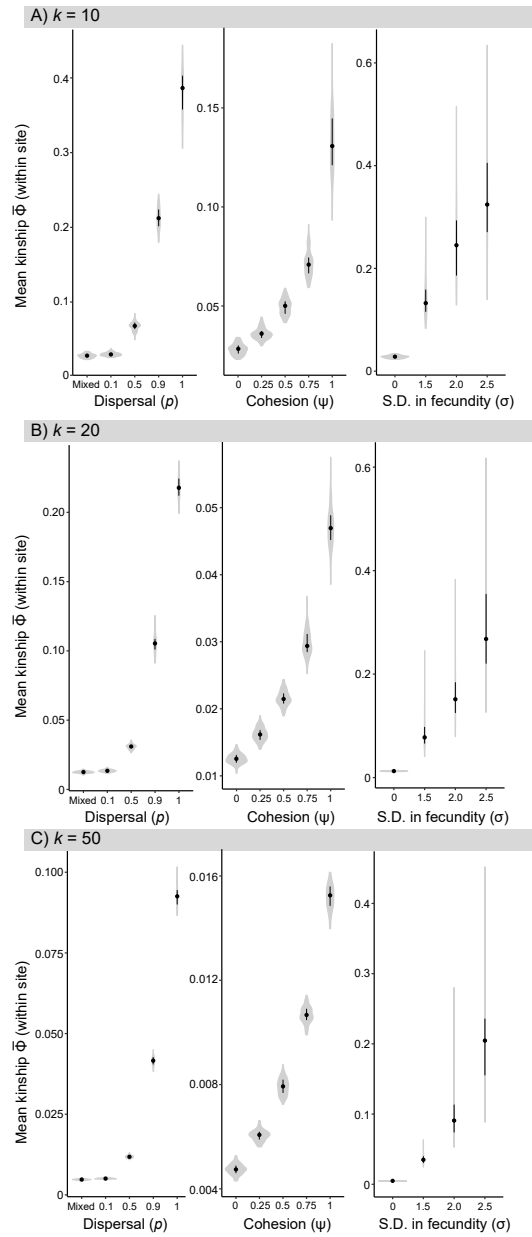## S1.4 Effect of carrying capacity on within-site kinship

Figure S6: Effects of limited dispersal, sibling cohesion, and variable reproductive success on mean kinship within a site at different carrying capacities. Violin plots (light gray) show kernel density estimates of $\bar{\phi}_{site}$ for 100 simulations. Black dots represent medians and black vertical lines represent interquartile ranges.
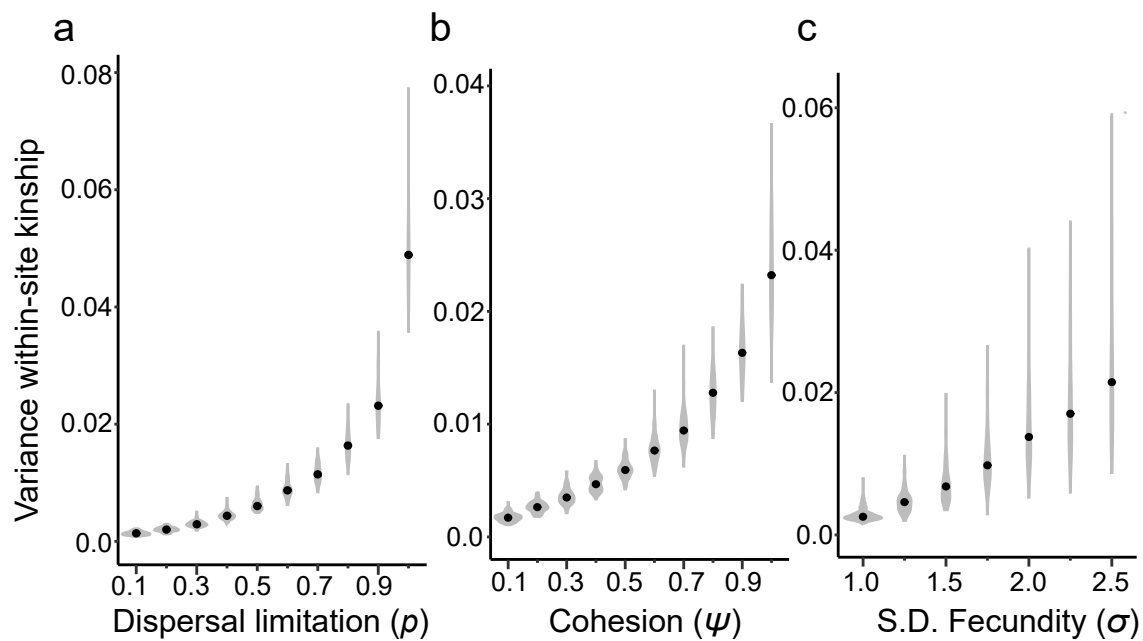
S1.5   Variance in within-site kinship

Figure S7: Effects of (a) dispersal, (b) cohesion, and (c) fecundity variability on the variance of within-site kinship. Violin plots (light gray) show kernel density estimates of the variance in ϕ_{site} for 100 simulations. Black dots represent medians.

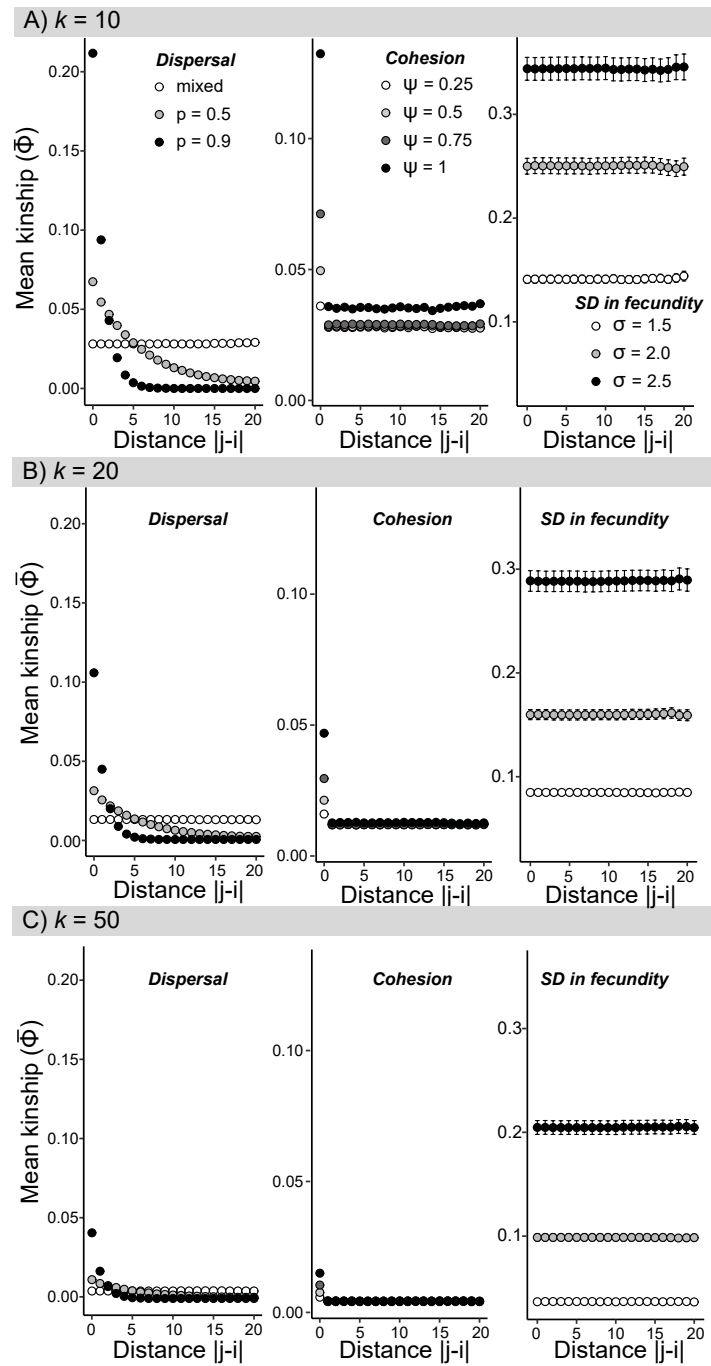## S1.6 Effect of carrying capacity on kinship at increasing distances

Figure S8: Effects of dispersal, cohesion, and variable reproductive success on mean kinship across spatial scales at different carrying capacities. Error bars are based on 100 simulations (when error is small, bars are not visible).

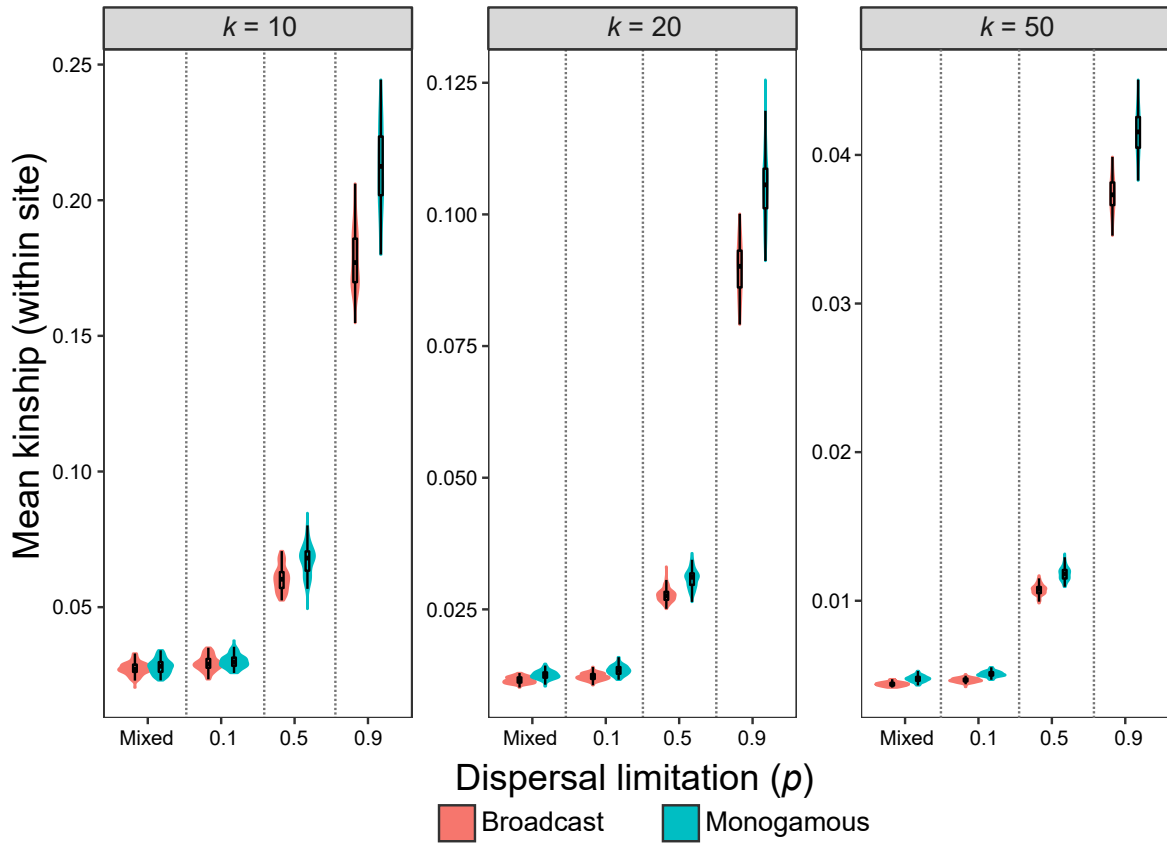## S1.7 Broadcast spawning at different carrying capacities

Figure S9: Violin plots of the distribution of mean kinship coefficients within sites ($n = 100$ simulations). Groupings on the x-axis refer to the dispersal kernel (either mixed, or limited with a given value of p). Note the different scales on y-axes. Boxplots are overlaid onto the violin plots.
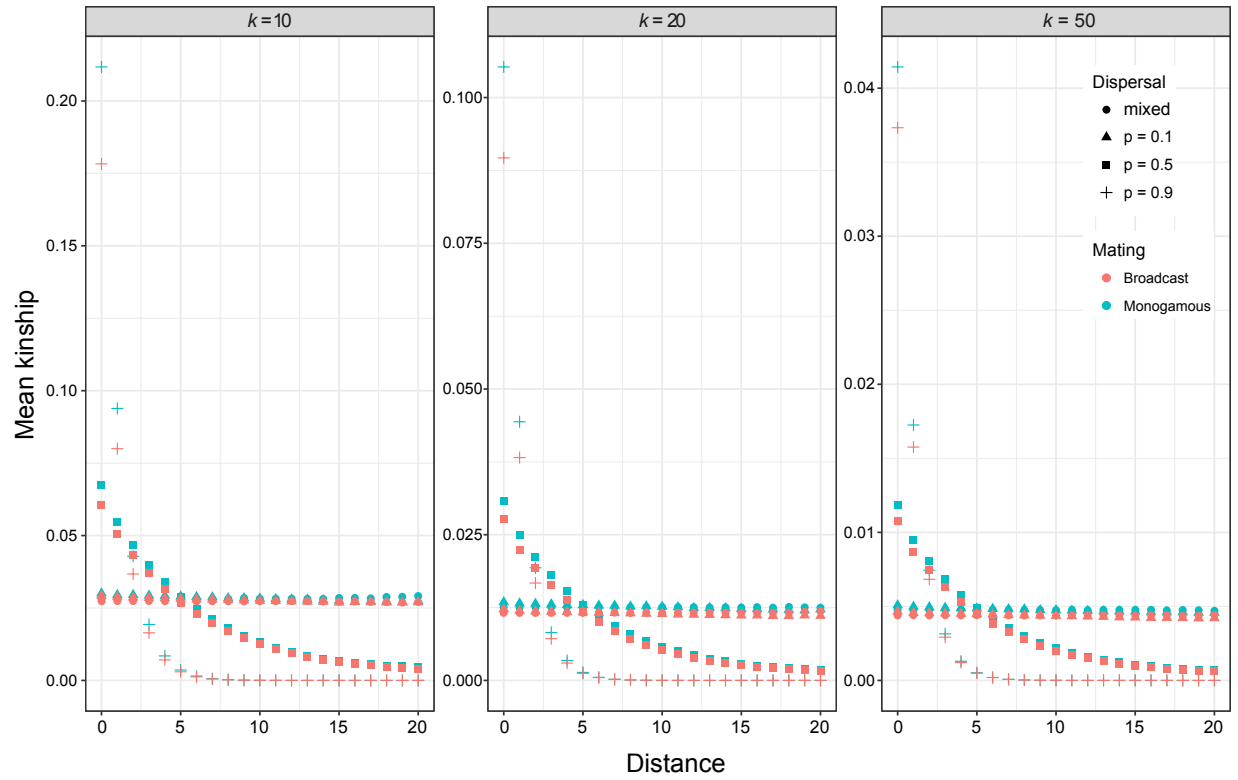
Figure S10: Mean kinship coefficients at increasing distance classes ($n = 100$ simulations). Distance on the x-axis represents the distance $|j - i|$ between every pair of individuals. Note the different scales on y-axes.

S1.8 Directional dispersal

In the main text we assume that dispersal is symmetric, i.e., when dispersal is limited we use a back-to-back geometric kernel (Eq. 1, cf., main text). Here, we briefly explore the effect of asymmetric dispersal, such that larvae can only travel in one direction along the lattice. The elements of one triangle of the connectivity matrix $\mathbf{C}$ are set to 0, while the elements in the other triangle are calculated from a geometric distribution: $c_{ij} = p(1 - p)^{j-i}$, for $i \leq j$, but $c_{ij} = 0$ otherwise. Asymmetric dispersal increases kinship overall, but does not change the general finding that expected kinship declines as the distance between sites increases (Fig. S11). Note, however, that in this relatively closed system, upstream sites are only sustained by local retention, leading to elevated inbreeding. Site-level extinction is also inversely related to p and the carrying capacity. In the extreme, populations in upstream sites could successively go extinct.
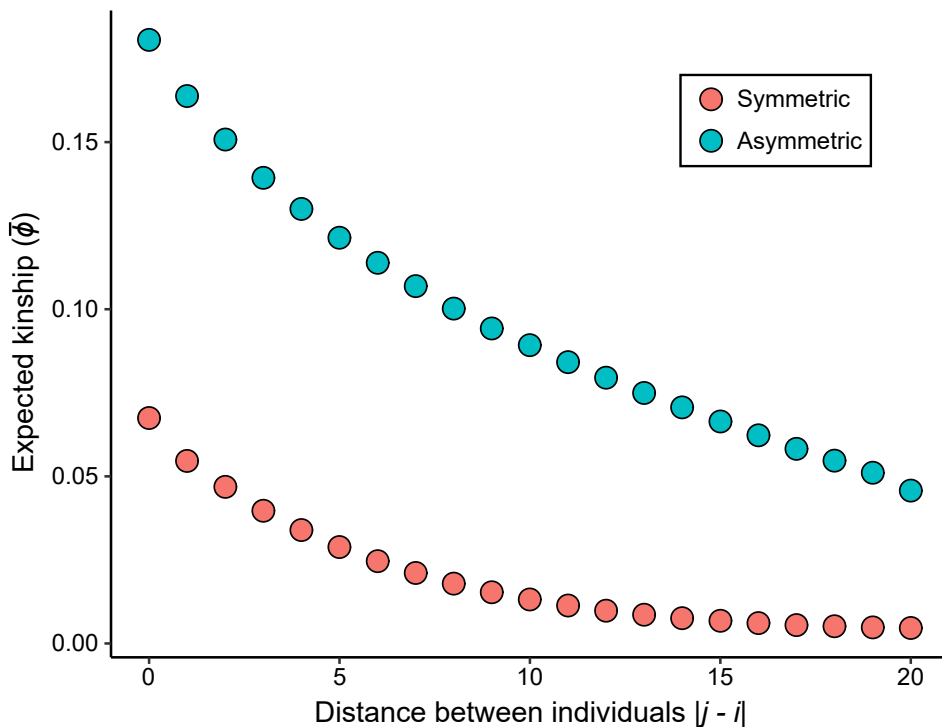

Figure S11: Effect of dispersal directionality on expected kinship at increasing spatial scales ($n = 100$ simulations).