

# Clownfish metapopulation persistence draft

Allison G. Dedrick<sup>a,\*</sup>

Katrina A. Catalano<sup>a</sup>

Michelle R. Stuart<sup>a</sup>

J. Wilson White<sup>b</sup>

Humberto Montes, Jr.<sup>c</sup>

Malin Pinsky<sup>a</sup>

a. Department of Ecology Evolution and Natural Resources, Rutgers University, 14 College Farm Road, New Brunswick, NJ 08901;

b. Oregon State University

c. Visayas State University

\* Corresponding author; e-mail: agdedrick@gmail.com

*(Author order not yet determined)*

## Introduction

Metapopulations exist along a continuum, with dynamics driven by the balance of

<sup>3</sup> extinction and colonization of local patches at one extreme and focused on the bal-

ance of immigration and emmigration at constantly-occupied local patches at the other (Kritzer and Sale, 2006). Terrestrial metapopulations often show extinction-  
6 colonization dynamics (e.g. Hanski, 1998), while marine metapopulations tend to exhibit immigration-emmigration dynamics where local extinction of patches is un-  
common (Kritzer and Sale, 2006). For these marine metapopulations, dynamics and  
9 persistence depend on connectivity among patches and the demographic rates at each patch (e.g. Hastings and Botsford, 2006; Hanski, 1998). Assessing levels of connectiv-  
ity and demographic parameters has been particularly challenging for marine  
12 species, where much of the mortality and movement happens at larval and juvenile stages when individuals are hard to track and have the potential to travel long distances with ocean currents (e.g. Kritzer and Sale, 2006; Cowen and Sponaugle, 2009;  
15 Roughgarden et al., 1988). A need to understand metapopulations for conservation and management, such as siting marine protected areas (e.g. Botsford et al., 2001;  
White et al., 2010), however, has led to a large body of theory describing how marine  
18 metapopulations might persist.

For any population to persist, individuals must on average replace themselves during their lifetime. Assessing replacement must take into account the demographic processes across the whole life cycle, including how likely individuals are to survive to the next age or stage, their expected fecundity at each stage, and the survival of any offspring produced to recruitment. In a spatially-structured population, in addition  
21 to assessing whether the reproductive output and survival of a population is sufficient,  
24 we must also consider how the offspring are distributed across space. Marine larvae were once thought to be well-mixed and dispersed far on ocean currents (e.g.

<sup>27</sup> Roughgarden et al., 1988), suggesting widespread connectivity among patches. Recent advances in estimating connectivity through natural tags and genetics, however, suggest that dispersal may be more limited (e.g. D'Aloia et al., 2013; Hameed et al.,  
<sup>30</sup> 2016; Almany et al., 2017), and local persistence of marine populations is seeming more possible.

Considering both the demographic processes within patches and the connectivity  
<sup>33</sup> among them, a metapopulation can persist in two ways: 1) at least one patch can achieve replacement in isolation, or 2) patches receive enough recruitment to achieve replacement through multi-generational loops of connectivity with other patches in  
<sup>36</sup> the metapopulation (Hastings and Botsford, 2006; Burgess et al., 2014). In the first case (termed self-persistence), enough of the reproductive output produced at one patch is retained at the patch for it to persist. If one patch persists, the metapopulation  
<sup>39</sup> persists. In the second (network persistence), closed loops of connectivity among at least some of the patches - where individuals from one patch settle at another and eventually send offspring back to the first in a future generation - provide the patch  
<sup>42</sup> with enough recruitment to persist within the network. Though it has been challenging to estimate the parameters necessary to understand how actual metapopulations persist, a large work of theory developed in part to guide marine protected area design  
<sup>45</sup> helps predict when each type of persistence is likely to occur (i.e., large patches relative to the mean dispersal distance are likely to be self-persistent, Botsford et al., 2001).

<sup>48</sup> New ways of identifying individuals and determining their origins, such as otolith microchemistry and genetic parentage analysis (e.g. Wang, 2004, 2014) are making

it increasingly possible to estimate both the demographic [ADD EXAMPLE CITATIONS] and the dispersal (e.g. Hameed et al., 2016; Almany et al., 2017) parameters necessary to assess persistence in real metapopulations. We might expect that populations on isolated islands are the most likely to be self-persistent, as they lack nearby populations with which to exchange larvae and would go locally extinct if they did not achieve replacement. At isolated Kimbe Island in Papua New Guinea, Salles et al. (2015) find that the population of orange clownfish (*Amphiprion percula*) can likely persist without outside immigration. In contrast, populations of bicolor damselfish (*Stegastes partitus*) at a set of reef patches across four isolated islands in the Bahamas do not appear able to persist without outside input (Johnson et al., 2018). For populations that exist in patches along a continuous coastline, rather than on separate islands, however, how patches interact and the scale of metapopulation persistence are still open questions. Flow differences along extended coastlines compared to isolated islands likely change the patterns of connectivity and ability of larvae to stay close to their natal site (NEED CITATIONS, EXPLANATION OF FLOW), possibly changing the scale of a metapopulation.

The number of studies estimating demographic rates and connectivity in marine metapopulations is growing (e.g. Salles et al., 2015; Johnson et al., 2018) (OTHER EXAMPLES), but most use data from one or a few years. Longer data sets enable better estimates of long-term average rates, rather than assuming the demographic and dispersal rates from a particular year are representative through time. More data is also useful for explicitly considering uncertainty, both to assess how well we understand persistence for a population and to see which parameters contribute

most to our uncertainty. Finally, sampling over many years provides the possibility of comparing abundance trends to persistence metrics to see if they tell a consistent  
75 story.

Here, we further our understanding of metapopulation dynamics in a network of patches along a coastline through a study of yellowtail clownfish (*Amphiprion*  
78 *clarkii*) in the Philippines. We assess persistence for all patches of habitat within a 30 km stretch of coastline, which exceeds estimates of the dispersal spread for this species (Pinsky et al., 2010), suggesting the network is likely to operate as a contained  
81 metapopulation. With seven years of annual sampling data, we are able to estimate persistence metrics and replacement over the longer term and investigate abundance through time to compare with the replacement-based persistence metrics. We use  
84 our long-term data set from habitat patches on a continuous section of coastline to understand persistence within a local network.

## Methods

### 87 Persistence theory and metrics

For a population to persist, individuals must be able to replace themselves on average at low abundance (CITATIONS). In non-spatially structured populations, we  
90 use criteria such as the average number of recruiting offspring each individual produces during its life (called  $R_0$  when the population is age-structured and density-independent) or the growth rate of the population (such as the dominant eigenvalue  $\lambda$  of an age-structured Leslie matrix) (Caswell, 2001; Burgess et al., 2014).  
93

For spatially-structured populations, we must also consider the spatial spread of offspring, often represented through a dispersal kernel or connectivity matrix (CITATIONS). At least one patch needs to achieve replacement for a metapopulation to persist (Hastings and Botsford, 2006).

We consider three primary metrics to assess whether and how the population is persistent: 1) lifetime recruit production (LRP), to assess whether the population has enough surviving offspring to achieve replacement 2) self-persistence (SP), to assess whether any individual patch can persist in isolation without input from other patches, and 3) network persistence (NP), to assess whether the metapopulation is persistent as a connected unit. We explain each metric below in detail. To represent the uncertainty in our estimates, we calculate each metric 1000 times, pulling each input parameter from a distribution or range. In our results, we show the range of values of each persistence metric as well as our best estimate.

### Lifetime production of recruits

We find the estimated number of recruits an individual recruit will produce (lifetime recruit production: LRP) by multiplying the total number of eggs a recruit-sized individual will produce in its lifetime (lifetime egg production: LEP) by the fraction of those eggs that will survive to become recruits ( $S_e$ ) (Fig. 1):

$$\text{LRP} = \text{LEP} * S_e. \quad (1)$$

If  $\text{LRP} \geq 1$ , the population has the possibility for replacement; individuals produce enough surviving offspring, before taking into account the probability of dispersal.

- <sup>114</sup> If  $LRP < 1$ , the individuals are not replacing themselves and the population cannot persist without input from outside patches.

### Self-persistence

- <sup>117</sup> A patch is able to persist in isolation (self-persistent) if individuals produce enough offspring that survive to recruitment (LRP) and disperse back to the natal patch (with probability of dispersal  $p_{i,i}$ ) to replace themselves. Burgess et al. (2014) use  
<sup>120</sup> LEP to represent offspring produced and local retention (LR) - the number of surviving recruits that disperse back to the natal patch over the number of eggs produced by the natal patch - to capture egg-recruit survival and dispersal in a criteria for  
<sup>123</sup> self-persistence:  $LEP \times LR \geq 1$ . We modify this to use include egg-recruit in the offspring term, using LRP in place of LEP, to assess whether a particular patch  $i$  is self-persistent:

$$SP_i = LEP \times \frac{\text{recruits}}{\text{egg}} \times \frac{p_{i,i} \times \# \text{ recruits from patch } i}{\frac{\text{recruits}}{\text{egg}} \times \# \text{ eggs produced by patch } i} \quad (2)$$

$$SP_i = LRP \times p_{i,i}.$$

- <sup>126</sup> A patch is self-persistent if  $SP \geq 1$ . If at least one patch is self-persistent, the metapopulation as a whole is persistent as well (Hastings and Botsford, 2006; Burgess et al., 2014).

<sup>129</sup> **Realized connectivity matrix and network persistence**

We find the probabilities of a recruit dispersing between each set of sites ( $p_{i,j}$ ) by integrating the dispersal kernel (eqn. 3) over the distances between sites. We then <sup>132</sup> create a realized connectivity matrix  $C$  by multiplying the dispersal probabilities by the expected number of recruits an individual produces:  $C_{i,j} = \text{LRP} \times p_{i,j}$  (Burgess et al., 2014, though we include egg-recruit survival in LRP, rather than in  $p_{i,j}$  as they <sup>135</sup> do). The diagonal entries of  $C$ , where the origin and destination are the same sites, are the values of self-persistence we calculate above.

Network persistence requires that the largest real eigenvalue of the realized connectivity matrix  $\lambda_C$  be greater than 1:  $\text{NP} = \lambda_C > 1$  (e.g. Hastings and Botsford, <sup>138</sup> 2006; White et al., 2010; Burgess et al., 2014).

### Defining recruit and census stage

<sup>141</sup> When assessing persistence, it is important to consider mortality and reproduction that occurs across the entire life cycle to determine whether an individual is replacing itself with an individual that reaches the same life stage (Burgess et al., 2014). We <sup>144</sup> define a recruit to be a juvenile individual that has settled on the reef within the previous year, which also encompasses the size we are first able to sample (3.5-6.0 cm for parentage studies). In theory, it does not matter how we define recruit as long <sup>147</sup> as we use that definition in our calculations of both egg-recruit survival and LEP. In our system, however, while it is straightforward to calculate LEP from any size, we do not have enough tagged recruits to reliably estimate survival to different recruit

<sub>150</sub> sizes. Instead, we choose the mean size of offspring matched in the parentage study  
as our best estimate of the size of a recruit ( $\text{size}_{\text{recruit}}$ ) and test sensitivity to different  
recruit sizes by pulling from a uniform distribution over the sizes the recruit stage  
<sub>153</sub> covers (3.5-6 cm, Table A1).

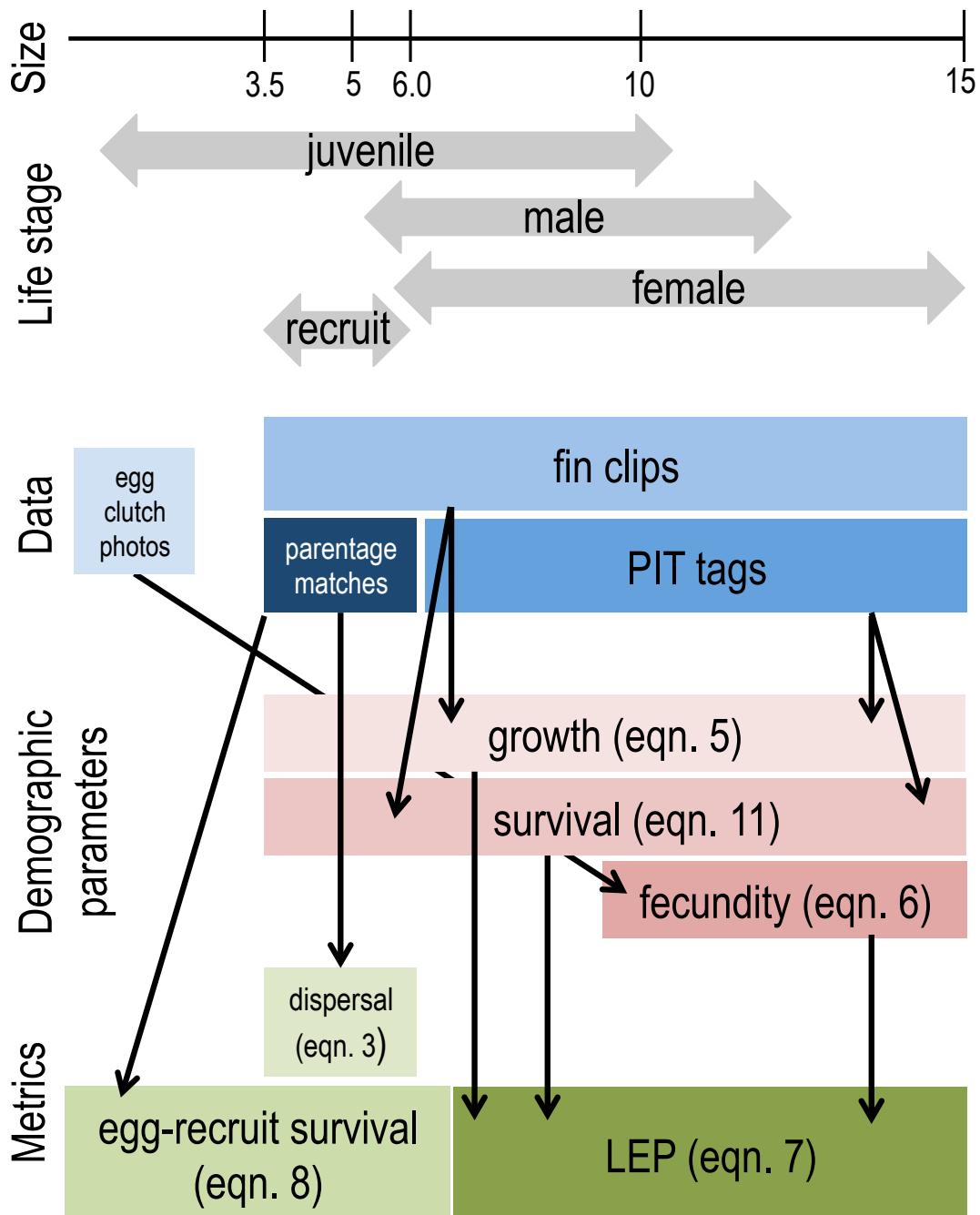


Figure 1: Here, we show the data collected for fish at each life stage (life stage boxes are not scaled by length of stage) and how the empirical data fit into our parameter  
10 and metric calculations.

## Study system

We focus on a tropical metapopulation of yellowtail clownfish (*Ampiprion clarkii*, Fig. 2c) on the coast of Leyte island facing the Camotes Sea in the Philippines (Fig. 2a). Like many clownfish species, yellowtail clownfish have a mutualistic relationship with anemones, where small colonies of fish live (Buston, 2003b; Fautin et al., 1992).  
Yellowtail clownfish are protandrous hermaphrodites and maintain a size-structured hierarchy; within an anemone, the largest fish is the breeding female, the next largest is the breeding male, and any smaller fish are non-breeding juveniles. The fish on an anemone maintain a strict social and size hierarchy (Buston, 2003b), with fish moving up in rank to become breeders only after the larger fish have died or left. In the tropical patch reef habitat of the Philippines, yellowtail clownfish spawn once per lunar month from November to May, laying clutches of benthic eggs that the parents protect and tend (Ochi, 1989; Holtswarth et al., 2017). Larvae hatch after about six days and spend 7-10 days in the water column before returning to reef habitat to settle in an anemone (Fautin et al., 1992).

Clownfish are particularly well-suited to metapopulation studies due to their limited movement as adults and clearly patchy habitat. Once fish have settled, they tend to stay within close proximity of their anemones [XX meters, CITATION]. This makes fish easier to relocate for mark-recapture studies and simiplifies the exchange between patches to only the dispersal during the larval phase. Patches, whether considered to be the reef patch or the anemone territory of the fish, are clearly discrete and easily delineated (Fig. 2a, b), which makes determining the spatial structure

of the metapopulation clear. Additionally, clear patches make it easier to assess  
177 how much of the site has been surveyed. These simplifying characteristics in habitat  
and fish behavior make clownfish and other similarly territory-based reef fish useful  
model systems for studies of metapopulation dynamics and persistence (e.g. Buston  
180 and DAloia, 2013; Salles et al., 2015; Johnson et al., 2018). Our focal species of  
yellowtail clownfish tends to behave more like larger reef fishes, with territories that  
can extend beyond single anemones (CITATION) and strong enough swimming skills  
183 that movement between patch reefs is possible though unusual (CITATION), than  
the smaller clownfish *A. percula* commonly used in previous metapopulation studies  
(e.g. Buston et al., 2011; Salles et al., 2015).

186 **Field data collection**

We focus on a set of seventeen patch reef sites spanning approximately 30 km along  
the western coast of Leyte island in the Philippines (Fig. 2a). The sites consist of  
189 rocky patches of coral reef and are separated by sand flats (Fig. 2b). Previous work  
using genetic isolation by distances estimated that yellowtail clownfish larvae have a  
dispersal spread of about 10 km (range 4-27 km, Pinsky et al., 2010), so our sites were  
192 selected to cover and exceed that range. On the north edge, the sites are isolated  
from nearby habitat with no substantial reef habitat for at least 20 km.

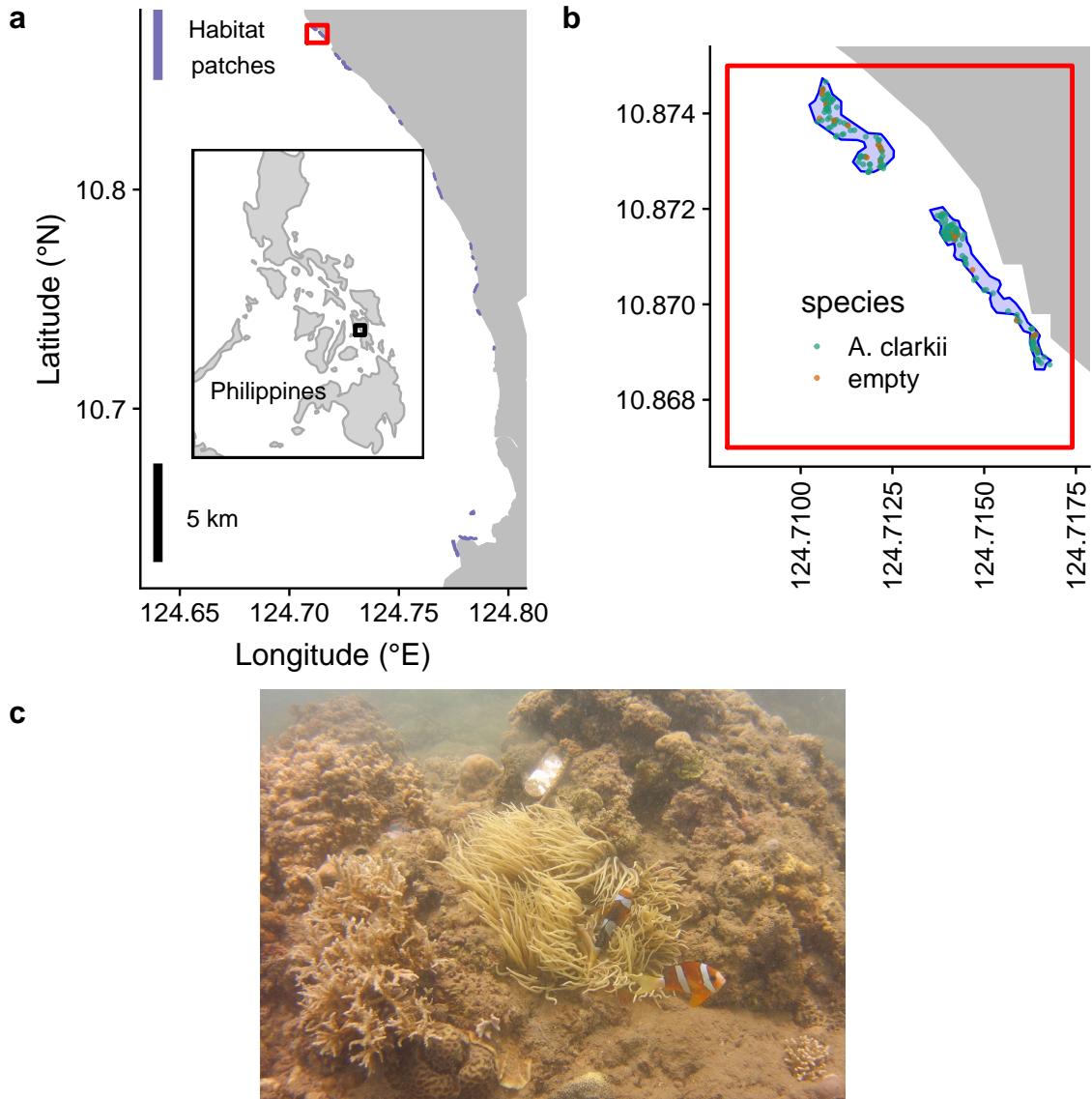


Figure 2: a) A map of the sites along the coast of Leyte in the Philippines. b) A zoomed-in map of the two northern-most sites, Palanas and Wangag, to show anemone arrangement, with anemones occupied by *A. clarkii* (green) or unoccupied by clownfish (orange). c) An example anemone occupied by *A. clarkii* in a typical habitat at the sites. The metal anemone tag is visible just above the anemone on the rock.

Since 2012, we have sampled fish and habitat at most of the sites annually (Table 195 A2). During sampling, divers using SCUBA and tethered to GPS readers swam the extent of each site. Divers visited each anemone inhabited by yellowtail clownfish, tagging the anemone to be able to track anemones through time. At each anemone, 198 the divers attempted to catch all of the yellowtail clownfish 3.5 cm and larger, taking a small tail fin-clip from each for use in genetic analysis, measuring the fork length, and noting the tail color (as an indicator of life stage). Starting in the 2015 field 201 season, fish 6.0 cm and larger were also tagged with a passive integrated transponder (PIT) tag, unless already tagged. Divers also looked for eggs around each anemone and measure and photographed any clutches found. In total, we took fin clips from 204 XX fish and PIT-tagged XX fish across all years and sites combined, with an average of XX fish clipped and XX fish tagged per year.

### Parentage analysis and dispersal kernel

207 We digested extracted genomic DNA from our tissue samples using the ddRAD-seq protocol (Peterson et al., 2012), filtered the sequences with the bioinformatics pipeline dDocent (Puritz et al., 2014), and called singled nucleotide polymorphisms 210 (SNPs) using the program FreeBayes (*is there a citation for this?*). In total, we genotyped XX fish. We used the SNPs to identify parent-offspring matches with the software program COLONY2 (Wang, 2012) (details on genotyping and parentage 213 analysis in Catalano et al., in prep).

Using the method described in (Bode et al., 2018), we fit a distance-based dispersal kernel (Catalano et al., in prep), where the relative dispersal is a function of

216 distance  $d$  as measured in kilometers and parameters  $\theta$  and  $k_d$ , which control the  
shape and scale of the kernel:

$$p(d) = e^k e^{-(e^k d)^\theta}. \quad (3)$$

219 We use a Laplacian dispersal kernel with shape parameters  $\theta = 1$  and scale parameter  
 $k_d = -1.84$  (Fig. 4a, estimated in (Catalano et al., in prep)). To account for uncertainty in the dispersal kernel, we keep the shape parameter  $\theta$  constant and pull the scale parameter  $k_d$  from a set capturing the 95% interval produced during kernal estimation in Catalano et al. (in prep).

225 The dispersal kernel is estimated using fish that have already recruited to a population and survived to be sampled so it gives the relative amount of dispersal given that a fish recruits somewhere, not the probability that a released larva will travel a particular distance. To find the probability of fish dispersing among our sites, we numerically integrate the dispersal kernel (eqn. 3) using the distance from the middle of the origin site to the closest and farthest bounds of the destination site as the upper and lower bounds. For example, the probability of dispersal from site A to B, where  $d_1$  is the distance from the middle of A to the closest edge of B and 228  $d_2$  is the distance from the middle of A to the far edge of B, is:

$$p_{A,B}(d) = \int_{d_1}^{d_2} e^k e^{-(e^k d)^\theta} dd. \quad (4)$$

## Estimating inputs from empirical data

### Growth and survival: mark-recapture analyses

- <sup>234</sup> We mark fish through both genetic samples and PIT tags, allowing us to estimate growth and survival through mark-recapture. After matching up recaptures of the same fish identified by genotype or tag, we have a set of encounters of XX marked fish that includes size and stage at each capture time.
- <sup>237</sup>

For growth, we estimate the parameters of a von Bertalanffy growth curve (Fabens, 1965) in the growth increment form relating the length at first capture  $L_t$  to the length at a later capture  $L_{t+1}$  (Hart and Chute, 2009), where  $L_\infty$  is the average asymptotic size across the population and  $K$  controls the rate of growth:

<sup>240</sup>

$$\begin{aligned} L_{t+1} &= L_t + (L_\infty - L_t)[1 - e^{(-K)}] \\ &= e^{(-K)}L_t + L_\infty[1 - e^{(-K)}]. \end{aligned} \tag{5}$$

We see from eqn. 5 that we would expect the first length  $L_t$  and the second length  $L_{t+1}$  to be related linearly (Hart and Chute, 2009). From the slope  $m = e^{(-K)}$  and y-intercept  $b = L_\infty[1 - e^{(-K)}]$ , we can estimate the von Bertalanffy parameters, such that  $K = -\ln m$  and  $L_\infty = \frac{b}{(1-m)}$ . We use the first and second capture lengths for fish that were recaptured after a year (within 345 to 385 days) to estimate  $L_\infty$  and  $K$ .

<sup>243</sup>

We have some fish that were recaptured multiple times so we randomly select only one pair of recaptures from each to use in estimating the parameters, then repeat this process 1000 times to generate a distribution.

<sup>246</sup>

<sup>249</sup>

We use the full set of marked fish to estimate annual survival  $\phi$  and probability of recapture  $p_r$  using the mark-recapture program MARK implemented in R (Laake, 2013). We consider several models with year, size, and site effects on the probability of survival and year and size effects on the probability of recapturing a fish on a log-odds scale (see full list in Table A3). For fish that are not recaptured in particular year, we estimate their size using our growth model (eqn. 5) and the size recorded or estimated in the previous year. Because fish are not well-mixed at our sites and instead stay quite close to their home anemones, we need to swim near an anemone to have a reasonable chance of capturing the fish on it. Therefore, we also consider a distance effect on recapture probability; we use the GPS tracks of divers to estimate the minimum distance between a diver and the anemone for each tagged fish in each sample year and include it as a factor in some of the models (Table A3).

## Fecundity

We use a size-dependent fecundity relationship, determined using photos of egg clutches and females (Yawdoszyn, in prep), where the number of eggs per clutch ( $E_c$ ) is exponentially related to the length in cm of the female ( $L$ ) with size effect  $\beta_l = 2.388$ , intercept  $b = 1.174$ , and egg age effect  $\beta_e = -0.6083$  dependent on if the eggs are old enough to have visible eyes:

$$\ln(E_c) = \beta_l \ln(L) + \beta_e[\text{eyed}] + b. \quad (6)$$

To get total annual fecundity  $f$ , we multiply the number of eyed eggs per clutch by the number of clutches per year  $c_e = 11.9$ , using the estimate from Holtswarth et al.

270 (2017).

We only consider reproductive effort once the fish has reached the female stage and use the average size of first observation as female for recapture fish as the transition size  $L_f = 9.32\text{cm}$ . To incorporate uncertainty, we draw from the full set of sizes at which recaptured fish were first captured as female (5.2 cm - 12.7 cm).  
273

### Lifetime egg production

We use an integral projection model (IPM) (e.g. Rees et al., 2014) with size as the continuous structuring trait  $z$  to estimate lifetime egg production (LEP), the total number of eggs produced by one individual, starting at the recruit stage. We initialize the IPM with one recruit-sized individual ( $\text{size}_{\text{recruit}}$ ) at the initial time step ( $t = 0$ ), then project forward for 100 time steps using the size-dependent survival (eqn. 12) and growth (eqn. 5) functions described in the probability density functions that make up the kernel to describe the survival and growth of the individual into the next time step. This gives us the size distribution at each time step, which represents the probability that the individual has survived and grown into each of the possible size categories, ranging from a minimum of  $L = 0\text{ cm}$  to a maximum of  $U = 15\text{ cm}$  (CITATION FOR MAX SIZE?). The probability that the individual is still alive and of any size decreases as the time steps progress; by using a large number of steps, we are able to avoid arbitrarily setting a maximum age and instead let the probabilities become essentially zero.  
276  
279  
282  
285  
288

We then multiply each size-distribution vector  $v_z$  in the matrix by the size-dependent fecundity function described above (eqn. 6) to get the total number of  
291

eggs produced at each time step. To get the total number of eggs one individual is likely to produce in its lifetime, we then sum across all time steps in the individual's  
294 potential life.

$$\text{LEP} = \sum_{t=0}^{t=100} \sum_{z=L}^{z=U} v_z, t f_z. \quad (7)$$

### Survival from egg to recruit

We estimate survival from egg to recruit ( $S_e$ ) with a method similar to that in  
297 Johnson et al. (2018) using parentage matches to estimate the number of surviving recruits produced by genotyped parents. We estimate the number of eggs produced by genotyped parent fish by multiplying the number of genotyped parents ( $N_g =$   
300  $XXX$ ) by the expected lifetime egg production for a fish of parent size ( $\text{LEP}_p$ ), using LEP calculated for a fish of 6 cm, the minimum size to be considered a parent in the parentage analysis (Catalano et al., in prep). We then scale the number of offspring we match to parents ( $R_m = XX$ , "tagged" recruits because they are able to be genetically matched back to their parents) to account potential recruits our sampling missed by including the proportion of habitat at our sites we cumulatively sampled  
303 across all sampling years ( $P_h = 0.34$ , see A.1 for details), the probability of capturing a fish if we sample its anemone ( $P_c$ , see A.2 for details), and the proportion of the total dispersal kernel area from each our of sites covered within our sampling region  
309 ( $P_d$ , calculation in A.2). Finally, because our dispersal kernel gives the probability of dispersal given that a recruit settled somewhere but our sampling region is not all habitat, we scale by the proportion habitat in our sampling region ( $P_s = XX$ , details

<sup>312</sup> in A.3). We then take this scaled number of estimated "tagged" recruits and divide it by the number of genetically tagged eggs produced by the genotyped parents to get an estimate of egg-recruit survival:

$$S_e = \frac{R_m}{N_g \text{LEP}_p} \cdot \frac{P_h P_c P_d P_s}{\text{XX}}. \quad (8)$$

<sup>315</sup> To incorporate uncertainty in our estimate of egg-recruit survival, we consider uncertainty in the number of offspring assigned to parents during the parentage analysis ( $R_m$ ) and in the probability of capturing a fish ( $P_c$ ). We generate a set <sup>318</sup> of values for the number of assigned offspring using a random binomial, where the number of trials is the number of genotyped offspring (XX) and the probability of success on each trial is the assignment rate of offspring from the parentage analysis <sup>321</sup> (XX) (Catalano et al., in prep). For the probability of capturing a fish, we pull values from a beta distribution that captures the mean and variance of capture probabilities across recapture dives (details in A.2).

### <sup>324</sup> Accounting for density-dependence

Ideally we would assess persistence metrics when the population is at low abundance and not limited by density-dependence, so we attempt to account for the effects of <sup>327</sup> density-dependence in the early life stages when it is likely strongest. Clownfish have strong social hierarchies and in other clownfish species, juveniles that have already settled on an anemone will prevent others from settling there as well (Buston, 2003a). <sup>330</sup> Each anenome, therefore, can only house one settling clownfish, so having anemones

already occupied by *A. clarkii* reduces the survival from egg to recruit if potential settlers get evicted by resident juveniles when they try to settle. We attempt to account for this density-dependent mortality by multiplying our estimate of settling recruits (the numerator of eqn. 8) by the proportional increase (DD) in unoccupied anemones at our sites if all of the *A. clarkii* anemones were unoccupied, where  $p_A$  is the proportion of anemones occupied by *A. clarkii* and  $p_U$  is the proportion of unoccupied anemones:  $DD = \frac{(p_U + p_A)}{p_U}$ . We present results both with and without this density-dependence modification.

### 339 Estimated abundance over time

We also consider trends in abundance of breeding females at each site over time to compare to our replacement-based estimates of persistence. Similarly to as we do for offspring, we scale up the number of females caught at each site  $i$  in each sampling year  $t$  by the proportion of habitat sampled in that site and year  $P_{h_{i,t}}$  and by the probability of capturing a fish  $P_c$ :

$$\# \text{ females}_{i,t} = \frac{\# \text{ females captured}_{i,t}}{P_{h_{i,t}} P_c}. \quad (9)$$

345 We then fit a linear model through the time series for each site  $i$  sampled in at least three years to assess whether the slope over time is positive or negative:

$$\# \text{ females}_i \sim \text{year} \quad (10)$$

## Results

<sup>348</sup> Our estimated abundance of females at each site over time does not suggest a clear trend in abundance (Fig. 3). In our very simple look at whether abundance seems to be increasing or decreasing at the sites we sampled in at least three years, eleven  
<sup>351</sup> sites had a positive slope over time and five had a negative slope (Fig. 3q). For the two largest sites, with a mean estimated number of females of between 150-200, one has a positive slope (Wangag, Fig. 3b) and one has a negative slope (Sitio Baybayon, Fig. 3p) and the next two largest sites are also split. Overall, there is not a clear directional change in abundance across the sites we sample over our sampling period.  
<sup>354</sup>

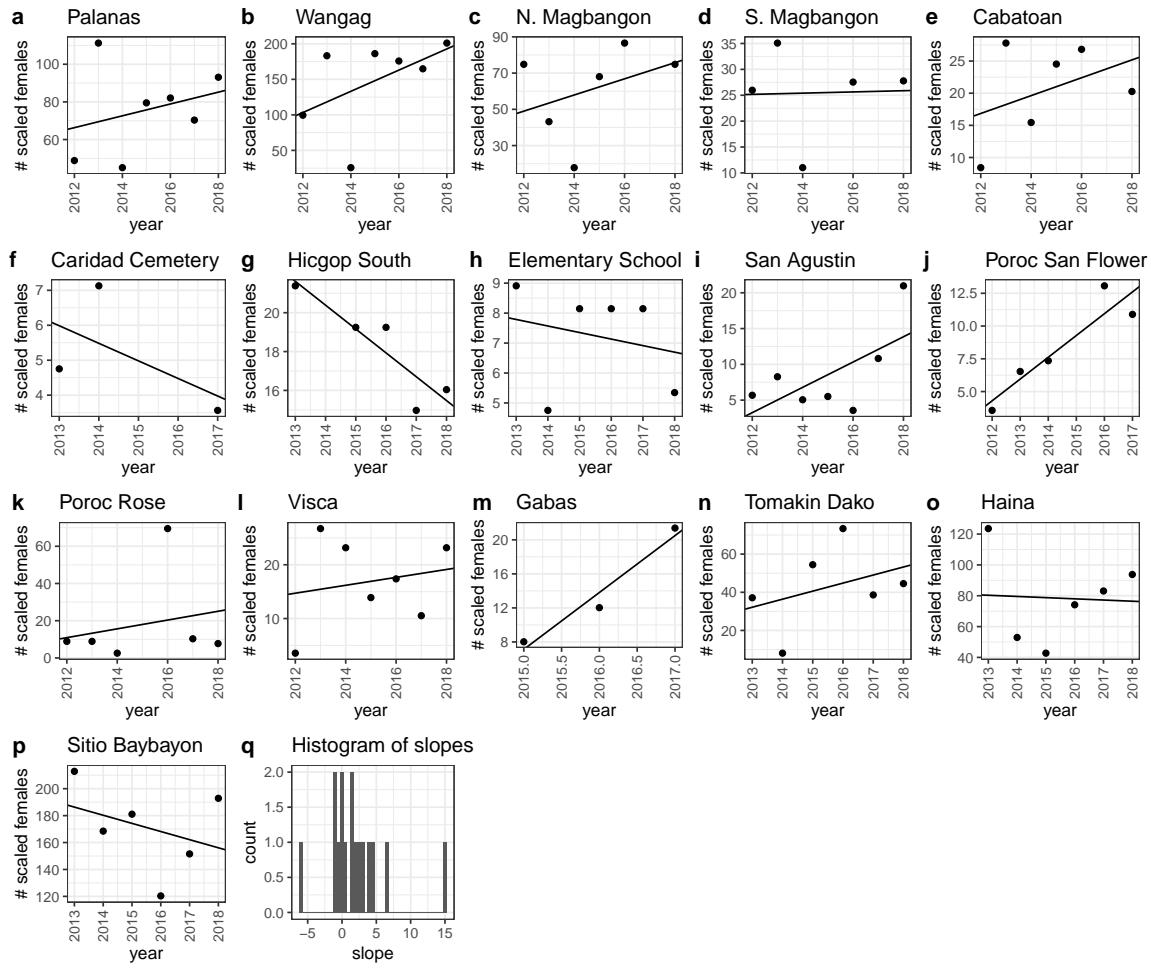


Figure 3: The estimated number of females at each site over the sampling years for sites sampled in at least three years. The total number of females at each site was estimated by scaling up the number of females captured at each site in each year by the proportion of habitat sampled at that site that season (see A.1 for details) and by the average probability of capturing a fish (see A.2). We show the estimated abundances and trend for each site individually (a-p) and a histogram of the slopes of abundance through time (q).

From the mark-recapture analysis of tagged and genotyped fish, we estimate mean  
 357 values of  $L_\infty = 10.58\text{cm}$  (range of estimates 10.39 - 10.75 cm) and  $K = 0.928$  (range  
 of estimates 0.854 - 1.025) for the von Bertalanffy growth curve parameters (Fig. 4b,  
 Table A1). For juvenile and adult (post-recruitment) survival on a log-odds scale, the  
 360 best-fit model has an effect of size, with coefficient  $b_a = 0.74 \pm 0.060$  SE and intercept  
 $b_\phi = -4.83 \pm 0.340$  SE. These results suggest that larger fish have higher annual sur-  
 vival, which is similar to survival estimates in other clownfish species (check Buston  
 363 paper). The accompanying best-fit model for log-odds recapture probability has a  
 size effect ( $b_1 = -1.816 \pm 0.080$  SE, Fig. A.2a) and a effect of diver distance from  
 the anemone ( $b_2 = -0.171 \pm 0.021$  SE, Fig. A.2b), with intercept  $b_{pr} = 17.93 \pm 0.858$   
 366 SE. The negative effect of both size and distance suggest that divers are less likely to  
 recapture larger fish and those at anemones far from areas sampled, with the chance  
 of recapturing an average-sized fish falling below 5% if a diver stays farther than XX  
 369 from its home anemone.

$$\log\left(\frac{\phi}{1-\phi}\right) = b_\phi + b_a \text{size} \quad (11)$$

$$\log\left(\frac{p_r}{1-p_r}\right) = b_{pr} + b_1 \text{size} + b_2 d. \quad (12)$$

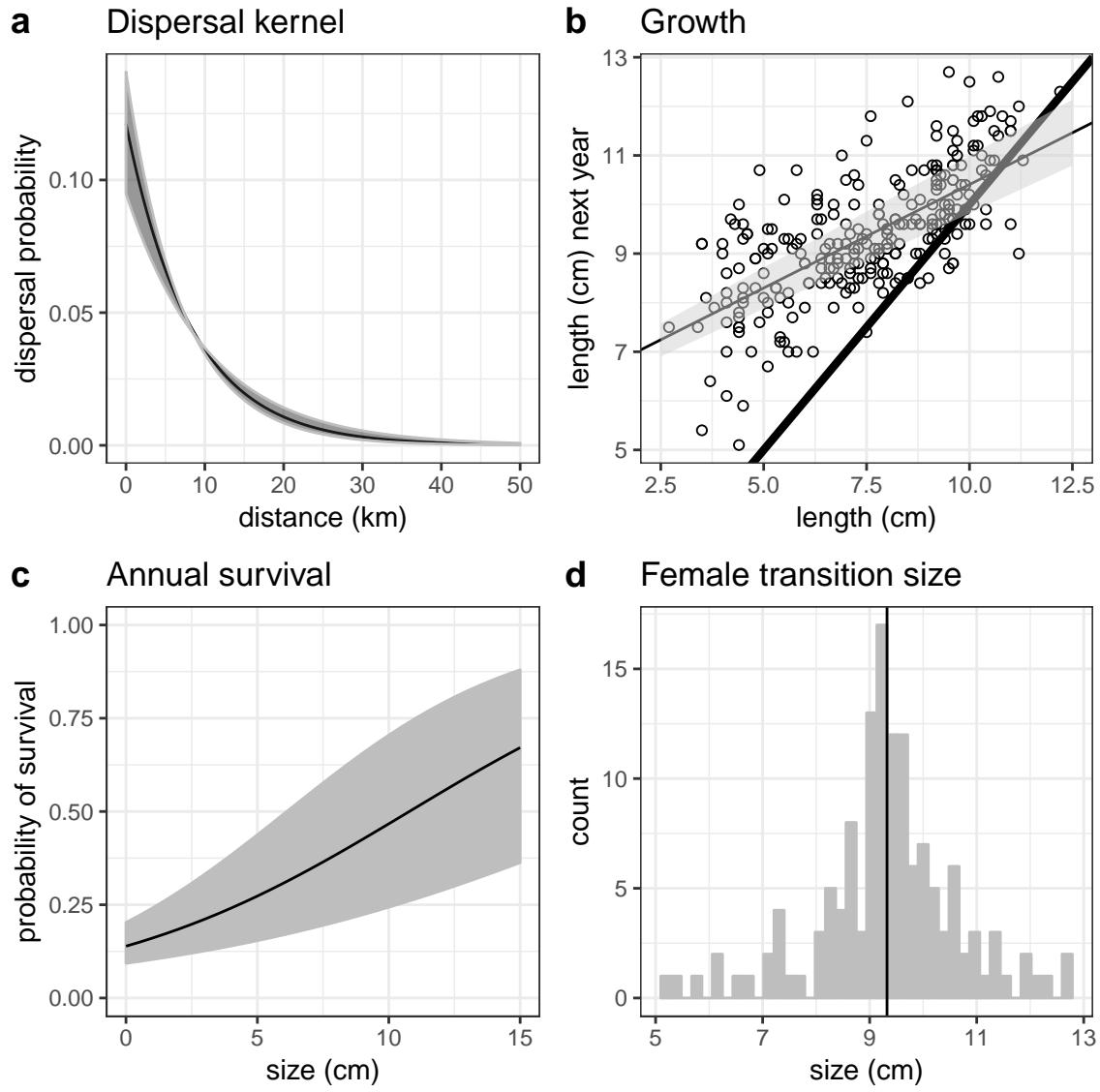


Figure 4: Best estimates (solid black line) and range included for uncertainty (gray) for dispersal (a), growth (b), post-recruit survival (c), and size at female transition (d) parameters.

Using our best estimates for growth, survival, and fecundity, we calculate a value

of LEP for 10876, ranging from XX to XX when we consider uncertainty in the inputs  
372 (Fig. 5a). The size at recruitment - the census point between egg-recruit survival  
and LEP - has the most effect on the value of LEP (Fig. B.4), with higher values of  
LEP the higher the size of recruitment as less mortality is included before reaching  
375 reproductive sizes.

We estimate egg-recruit survival  $S_e$  to be 1.82e-05, ranging from XX to XX  
when we include uncertainty in the number of offspring assigned to parents and  
378 the probability of catching a fish (Fig. 5b). Uncertainty in the size of transition  
to breeding female  $L_f$  has the largest effect on egg-recruit survival (Fig. B.7); we  
only consider reproduction from females, to avoid double-counting, so the larger  
381 the transition size to female, the fewer tagged eggs we estimate were produced by  
genotyped parents and the higher egg-recruit survival.

We estimate lifetime recruit production, the product of LEP and  $S_e$ , to be 0.20,  
384 below the value of 1 necessary for replacement. This suggests that even without  
considering connectivity, the individuals at our sample populations do not produce  
enough offspring that survive to recruitment to replace themselves. When we consider  
387 uncertainty in our parameter estimates, we do see a few cases where  $LRP > 1$ , but  
the majority are well below the threshold for replacement.

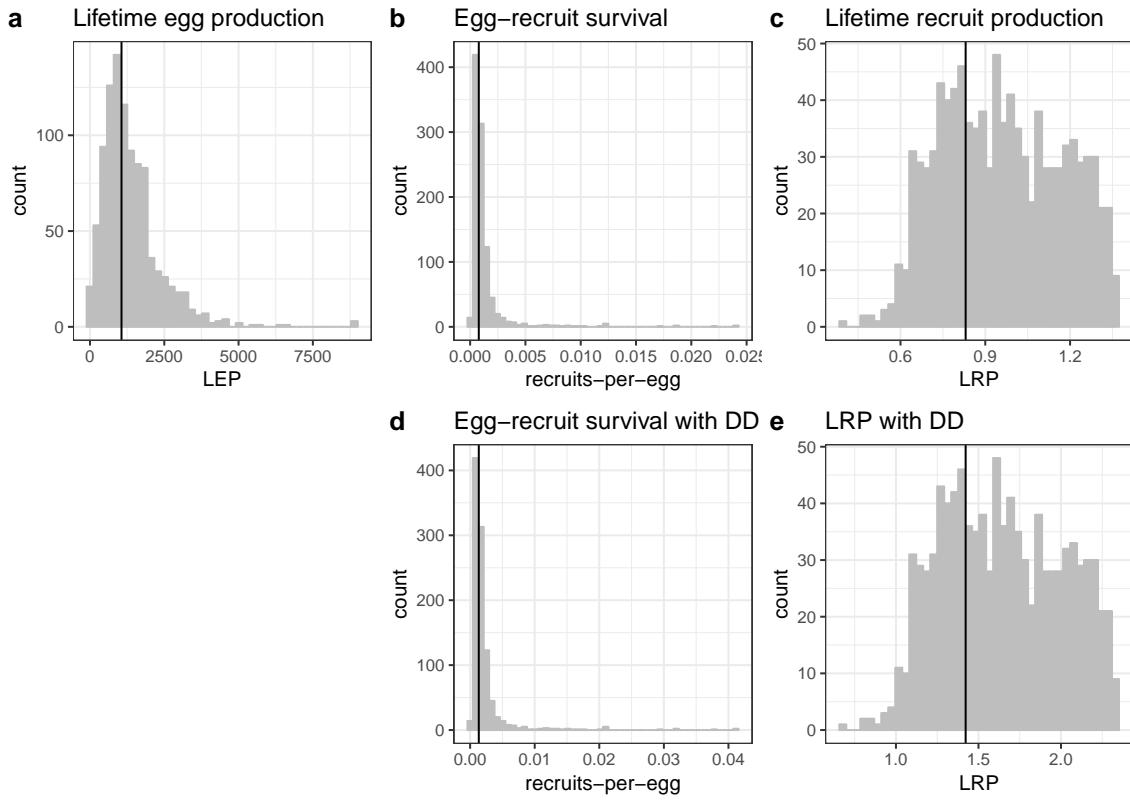


Figure 5: Estimates of a) LEP, b) egg-recruit survival, and c) LRP, showing the best estimate (black solid line) and range of estimates considering uncertainty in the inputs. The estimates in d-e include our attempt to remove density-dependence in the early life stages, while the estimates in b-c do not. We only consider density-dependence in egg-recruit survival so LEP (a) estimates are the same in both cases.

We do not find any sites with self-persistence values  $> 1$ , indicating that the  
 390 site could persist in isolation. Given that our estimate of LRP does not suggest replacement and only a fraction of that recruitment stays at the natal site, this

makes sense. We see the highest values of self-persistence at Haina ( $SP = 0.024$ ) and  
393 Wangag ( $SP = 0.010$ ), our two widest sites.

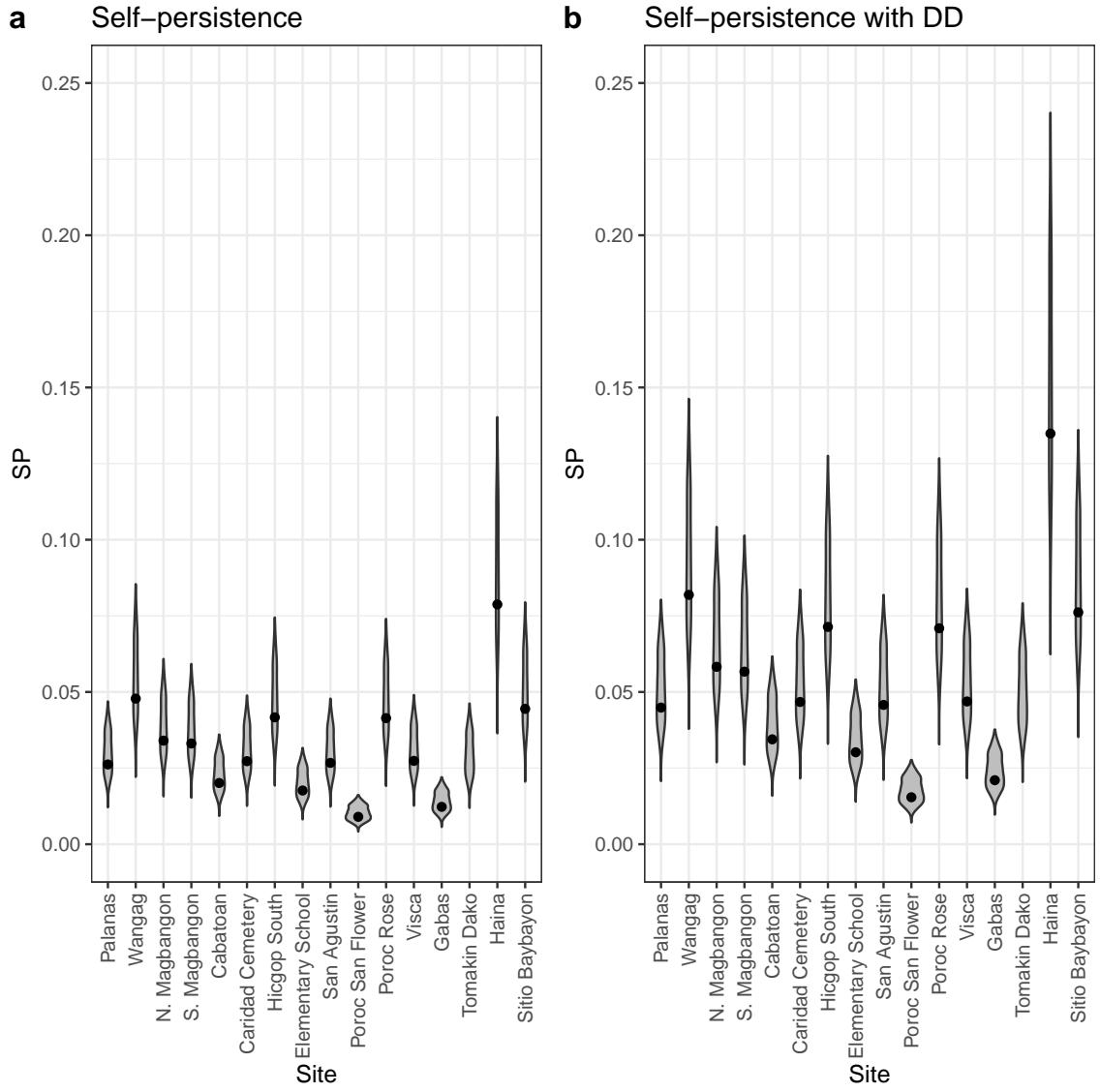


Figure 6: Values of self-persistence at each site, showing the best estimate (black point) and range of estimates considering uncertainty in the input parameters. No site reaches a value of  $SP > 1$ , necessary to be self-persistent. The estimates in (b) attempt to compensate for density-dependence in early life stages in our data, while the estimates in (a) do not.

We also do not find evidence of network persistence; the dominant eigenvalue of the realized connectivity matrix  $\lambda_c$  is 0.034, well below the value of 1 that indicates  
396 network persistence (Fig. 7a). The value of  $\lambda_c$  rises to XX when we compensate for density-dependence (Fig. 7c) but still below network persistent. We see that most of the connectivity occurs among the sites in the northern part of our sample area,  
399 from Palanas to Caridad Cemetery, and at the southern part of our sample area from Tomakin Dako to Sitio Baybayon (Fig. 7b).

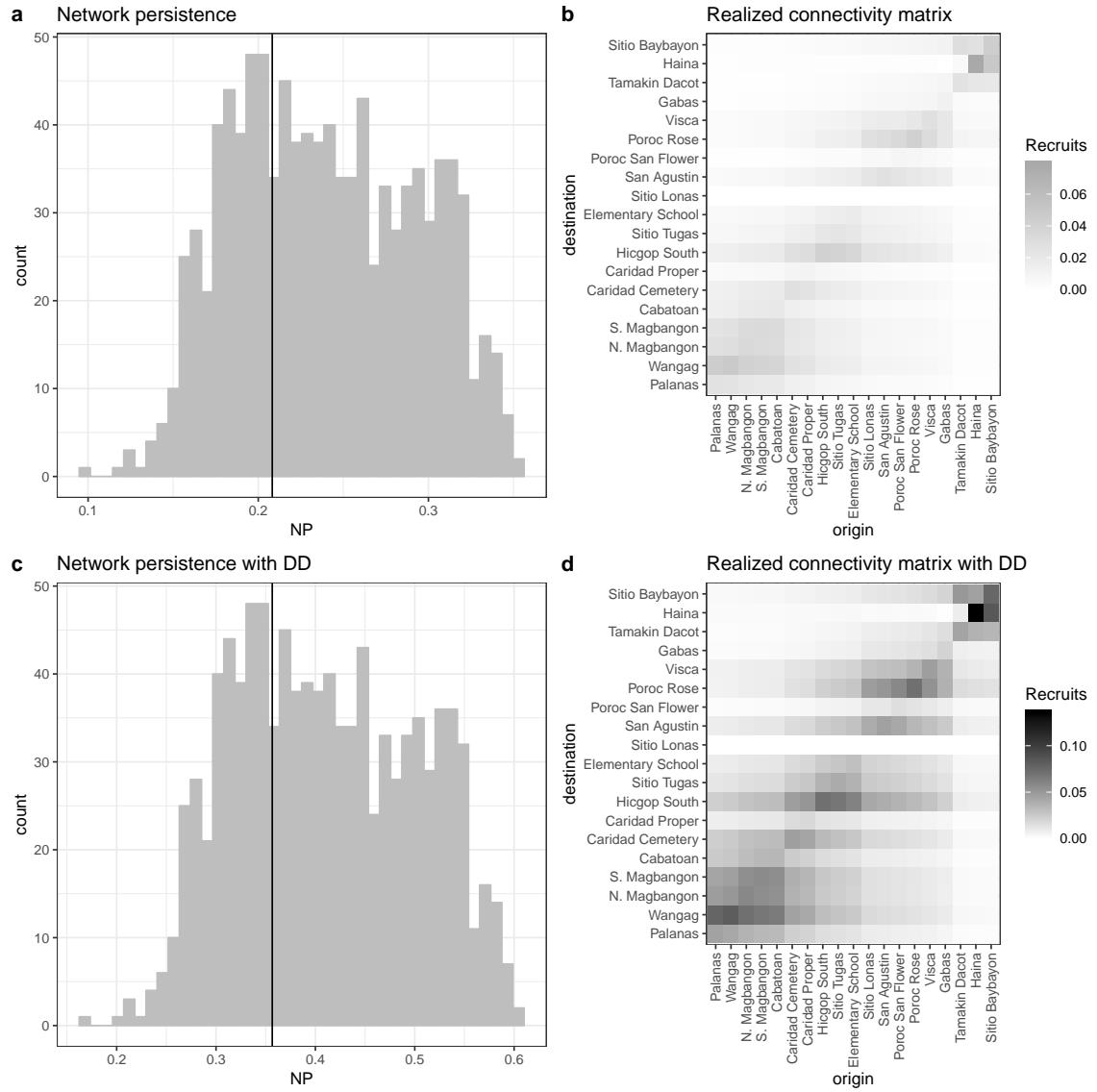


Figure 7: a) Network persistence values, showing the best estimate (black solid line) and range of estimates considering uncertainty, without accounting for density-dependence in our data. b) The realized connectivity matrix  $C$ , with sites arranged from north (Palanas) to south (Sitio Baybayon), without accounting for density-dependence in our data. Plots c) and d) show the network persistence values and realized connectivity matrix with attempts to compensate for density-dependence in early life stages in our data.

Based on our estimates of LRP, SP, and NP, we do not expect that our set of sites  
402 is able to persist in isolation as a closed system. To explore what would be required  
for persistence, we consider a hypothetical scenario in which we consider the system  
closed and assume that all of the recruits arriving at our sites came from adults at  
405 our sites. In this case, we find a value of  $LRP = 1.21$ , above the value of 1 necessary  
for replacement (Fig. B.1a). When we add in the connectivity, we see a higher  
value of  $\lambda_c$  in our best estimate ( $NP = 0.20$ ) but still not high enough to indicate  
408 network persistence (Fig. B.1b). We see more of the distribution of estimates above  
1, however, suggesting that network persistence is within our range of uncertainty  
in this case, though not likely. With our site configuration and dispersal kernel  
411 estimate, we would need a value of LRP of XX (an egg-recruit survival of XX with  
our estimated value of LEP or a value of LEP of XX with our estimated value of  
egg-recruit survival), to  $\lambda_c = 1$  and network persistence.

## 414 Discussion

We do not see evidence for persistence in our metric estimates, neither self-persistence  
where an individual site could persist alone nor network persistence accounting for  
417 exchange among sites. The abundances through time at our sites are relatively  
constant, however, suggesting that the population at our sites is stable but relies on  
input of recruits from outside sites to persist. The portion of coastline we sampled  
420 is likely a sink portion of a larger metapopulation.

For our sites to be able to persist as a network on their own, the number of  
surviving recruits produced by an average recruit - LRP - would need to be higher.

423 With our estimated connectivity among the sites, LRP would need to be at least XXX  
to see network persistence among our sites, about XXX times higher than our best  
estimate and XXX times higher than the top of our range of uncertainty. To achieve  
426 that value of LRP, either production of eggs would need to increase to an LEP of XX  
or egg-recruit survival would need to be at least XX. Our estimate of LEP is XXX  
compared to estimates for similar fish (e.g. CITATIONS) and our estimate of egg-  
429 recruit survival is XXX compared to other estimates (e.g. CITATIONS), making a  
rise of XX for either XX. Since LRP is less than one for our sites - the average recruit  
only replaces itself with XX of a surviving recruit of the same stage - no amount  
432 of increased retention or connectivity, even retaining all of the recruits produced  
from our sites, would lead to network persistence. Similarly, if other surrounding  
patch populations had a similar LRP, increasing the area of the network to include  
435 them would also not achieve network persistence. If nearby sites have higher egg  
production or survival to recruit, however, it might not take much of an increase in  
area considered to create a persistence network. Nearby sites such as Quatros Isles  
438 and XX have higher quality habitat and could be contributing recruits to our sites.

We see considerable uncertainty in our estimate of persistence metrics, depending  
on the particular input values we use, with LRP values ranging from XX to XX and  
441 NP values from XX to XX. We see a factor of XX between our highest and lowest  
LRP values and our best estimate of LRP is only XX times away from suggesting our  
population is network persistent. Measuring demographic and dispersal parameters  
444 in the field is challenging, particularly egg-recruit survival, which also has a large  
effect on our persistence metric estimates. and accounting for our uncertainty clearly

helps us understand XX. While none of our estimates even with uncertainty indicate  
447 persistence and our population is likely subsized by outside input and unable to persist as a network on its own,

### 3) Uncertainty in our estimates and why we need to consider it

- 450 • Try to consider uncertainty in a variety of our inputs, hard to estimate in the field
- Important to do this b/c not going to get a point estimate
- 453 • Think more about what the point of this paragraph is

We were surprised that our estimates of persistence are so far from the threshold for network persistence for our sites because previous estimates of the mean dispersal distance of *A. clarkii* (10 km Pinsky et al., 2010) and those from our 2012-2018 samples (Catalano et al., in prep, range from XX - XX km for individual-year kernels, XX for all-years kernel, which is used in this analysis) are well within the 30 km span of our sites. Our below-replacement estimate for LRP, however, suggests that lack of persistence in these sites is not due to excessive dispersal out of the area but due to low production and survival of offspring. SOME SORT OF COMPARISON  
459 WITH OTHER VALUES OF SURV, GROWTH, FECUNDITY. The reef health and habitat quality in our sites is generally low, due anthropogenic effects such as pollution and silt from a nearby gravel mining operations and habitat disturbance  
462 due to storms. Our sites are in an area that was hit by Typhoon XX, the strongest typhoon ever to make landfall, early in our sampling period, which destroyed much

of the reef habitat in some of our northern sampling areas. This recent disturbance  
468 and generally low habitat quality could contribute to low production of surviving  
recruits in our sites, necessitating subsidation by outside populations. RAP-UP,  
CONCLUDING SENTENCE OR TWO!

Persistence criteria, such as those detailed in Hastings and Botsford (2006) and  
Burgess et al. (2014), ask whether a population at low abundance can grow and  
recover rather than going extinct. Density-dependence is assumed to not exist at  
474 low abundances (CITATIONS, with the exception of xx density-dependence, like the  
allee effect) so is not explicitly considered in persistence metrics. In real populations,  
however, it can be challenging to estimate density-independent demographic rates,  
477 as density-dependence is occurring in the population as it is sampled. In *A. clarkii*,  
density-dependence is likely most important in early life stages, as for many fish  
species, but could play an important role throughout the life history due to the  
480 social hierarchies in colonies of clownfish (CITATIONS). In other species of clownfish,  
individuals on the same anemone maintain strict size spacing, restricting their food  
intake and growth to avoid encroaching on the position of another fish and being  
483 attacked or evicted (CITATIONS). This suggests that while fish are in the pre-  
reproductive queue, density-dependence may lower growth rates compared to the  
growth of fish alone on an anemone, as would be the case in a population at low  
486 abundance. We attempt to account for the primary effect of density-dependence  
on our estimate of egg-recruit survival but other estimates, particularly growth and  
survival, would also likely be higher in the absence of density-dependence and increase  
489 LRP.

Our estimates of persistence metrics do not suggest that the region of site we sample are able to persist as a network without outside input, despite covering an area  
492 larger than the estimated mean dispersal distance for our focal species. Our below-one estimate of LRP, even when we consider uncertainty, suggests that dispersal is not the primary reason our sites do not persist as a network: our sites do not  
495 produce enough offspring for replacement regardless of dispersal patterns, possibly due to worsening habitat quality. This is a reminder that dispersal is only part of the persistence story for metapopulations; even areas of habitat that seem large enough  
498 to contain a persistent network based on dispersal distance will not be able to persist in isolation if they have low production and survival of offspring. We do find recruits coming back to our region, and even to their natal site, but broader connectivity to  
501 more productive sites enables our sites to persist.

3) Uncertainty in our estimates and why we need to consider it

- Try to consider uncertainty in a variety of our inputs, hard to estimate in the  
504 field

- Important to do this b/c not going to get a point estimate
- Think more about what the point of this paragraph is

507 4) How do we reconcile this with the finding that mean dispersal distance is within the reach of our sites but they don't persist as a contained metapopulation?

- They can both be true at the same time - maybe fish from our sites don't  
510 disperse very far but others in sites nearby have access to different currents,

etc. send recruits farther and into our sites.

- It's a linear coastline but within a big bay - could have some interesting connectivity patterns via currents/eddies/etc. (reference how far the drifters went?)  
513 Are there some citations on this?

#### 5) Caveats paragraph

- Density-dependence - likely plays a larger role throughout the life cycle, not just at early life - particularly important in this species (and others with strong social hierarchies)  
516
- Site-specific demographic rates - coming in the future as get more of a handle on predator densities and such. Not have enough data from some of the sites and some of the params (fecundity) to get site-specific rates, though those have been interesting/important in other systems. If our larger sites had higher production/survival, might see persistence.  
519  
522

#### 6) Wrap-up paragraph

- How does this contribute to our understanding of metapopulation persistence?  
525 Where do we go from here?

We do not see evidence for persistence in our metric estimates, either self-persistence  
528 where an individual site could persist alone or network persistence with exchange among sites. The abundances through time at our sites are relatively constant, however, suggesting that the population at our sites is stable but relies on input of

<sup>531</sup> recruits from outside sites to persist. The portion of coastline we sampled is likely a portion of a larger metapopulation.

<sup>534</sup> Big picture: What do our results mean for persistence in this system and our understanding of metapopulations generally?

- So we don't see persistence in our metrics, either self-persistence or network persistence but our abundances don't seem to be changing. Suggests that this is just a portion of a larger metapopulation, rather than a self-contained metapopulation. Maybe it is a sink? Persistent in terms of constant abundance but relies on outside immigration to persist.

<sup>540</sup> • How does dispersal spread (estimated to be within our sites) interact with scale of a self-contained metapopulation? How do we reconcile this in our system, where we don't estimate dispersal that far but don't see network persistence in an area range that spans the estimated spread? (This point might change, depending on mean dispersal distance from the new kernels)

- Sensitivity - how would our parameters need to change to see persistence?

<sup>546</sup> Egg-recruit survival is a big one. Discuss limitations of how we calculated it (offspring going outside our pops not included - though we might change this), what we see for persistence when estimate recruits/recruits instead. Contextualize this with what other studies have found for these parameters, how reasonable it would be to get better estimates in the field.

Persistence criteria, such as those detailed in Hastings and Botsford (2006) and <sup>552</sup> Burgess et al. (2014), ask whether a population at low abundance can grow and

recover rather than going extinct. Density-dependence is assumed to not exist at low abundances (CITATIONS, with the exception of xx density-dependence, like the allee effect) so is not explicitly considered in persistence metrics. In real populations, however, it can be challenging to estimate density-independent demographic rates, as density-dependence is occurring in the population as it is sampled. In *A. clarkii*, density-dependence is likely most important in early life stages, as for many fish species, but could play an important role throughout the life history due to the social hierarchies in colonies of clownfish (CITATIONS). In other species of clownfish, individuals on the same anemone maintain strict size spacing, restricting their food intake and growth to avoid encroaching on the position of another fish and being attacked or evicted (CITATIONS). This suggests that while fish are in the pre-reproductive queue, density-dependence may lower growth rates compared to the growth of fish alone on an anemone, as would be the case in a population at low abundance. We attempt to account for the primary effect of density-dependence on our estimate of egg-recruit survival but other estimates, particularly growth and survival, would also likely be higher in the absence of density-dependence and increase LRP.

Our estimates of survival probabilities are similar to those estimated for other species of clownfish, particularly our relationship with size where small fish have a low annual survival and the largest fish have a high annual survival (CITATIONS, Buston paper, also compare to Salles et al. 2015). Our fecundity estimates are lower than those for *A. clarkii* in temperate areas, almost XX times lower (CITATIONS, Ochi papers - 17,500 eggs/yr/female, from Bell 1976).

576 More detailed discussion of our estimates, limitations, ways to move forward:

- Discuss density-dependence: not explicitly accounting for it, included in our egg-recruit survival estimate. But it's these metrics at low abundance, when DD isn't happening, that really matter for persistence. Egg-recruit-survival is probably higher in that case than our estimate of it here (b/c larvae able to settle without being chased off by already-settled recruits). But is it high enough?
- Discuss site-specific demographic rates, why we don't estimate them in our system, the importance they play in other studies, what we might need to go about resolving them, whether we would expect to see them.
- Contextualize our parameter estimates with those from other studies (esp. survival, growth, fecundity).

588 Broadening back out:

- What does this mean for moving forward in understanding metapopulation persistence more broadly? Stability in abundance doesn't mean the population would be able to persist in isolation. Area required seems to be much wider than dispersal kernel spread (particularly if LRP production is right around replacement). Even areas of habitat along a linear coastline seem to be drawing much of their recruitment from a larger surrounding area - even though we see some local retention, maybe broader connectivity is still the story in terms of receiving enough recruitment to persist.

## Appendix

### Summary of parameters

[Need to clarify somewhere what kind of distributions are going into the uncertainty  
 600 runs (drawn from data, uniform across a range, 95% confidence bounds, etc.)]

Parameter	Description	Best estimate	Range in uncertainty runs	Notes
$k_d$	scale parameter in dispersal kernel	-1.36	-2.03 to -0.96	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$\theta$	shape parameter in dispersal kernel	0.5	NA	estimated using methods in Bode et al. (2018) in Catalano et al. (in prep)
$L_\infty$	average asymptotic size in von Bertalanffy growth curve	10.58 cm	10.39 to 10.75 cm	

$K$	growth coefficient in von Bertalanffy growth curve	0.928	0.854 to 1.025	
$b_\phi$	intercept for adult survival	-4.83	$\pm 0.340$ standard error	
$b_a$	size effect for adult survival	0.74	$\pm 0.060$ standard error	
$b_{pr}$	intercept for recapture probability from mark-recapture analysis	17.93	$\pm 0.858$ standard error	not used in persistence estimates
$b_1$	size effect for recapture	-1.816	$\pm 0.080$ standard error	not used in persistence estimates
$b_2$	distance effect for recapture	-0.171	$\pm 0.021$ standard error	not used in persistence estimates
$\text{size}_{\text{recruit}}$	size (cm) of recruited offspring	mean of size of offspring in parentage analysis = 4.4 cm	3.5 - 6.0 cm	

$b_e$	coefficient for eyed eggs	-0.608		Yawdoszyn et al. (in prep)
$b_l$	size effect in eggs-per-clutch relationship	2.39		Yawdoszyn et al. (in prep)
$b$	intercept in eggs- per-clutch rela- tionship	1.17		Yawdoszyn et al. (in prep)
$L_f$	size at transition to female	9.32cm	5.2 - 12.7cm	
$P_c$	probability of capturing a fish	0.56	drawn from beta distri- bution with parameters $\alpha_{P_c} = 1.44$ and $\beta_{P_c} = 1.13$	details in A.2

Table A1:

## A Method details

### A.1 Proportion of habitat sampled

We used tagged anemones to estimate the proportion of habitat sampled at each site in each year ( $P_{h_{i,t}}$ ). We tagged each anemone that is home to *A. clarkii*, with a metal tag, which is relatively permanent and easy to re-sight (the anemone tag is visible above the anemone in Fig. 2c), so we consider the total number of metal tags at each site to be the total number of anemones that are habitat. We divide the number of tagged anemones visited each sampling year by the total number of metal tags at that site to get the proportion of habitat sampled. We use proportion of anemones rather than proportion of total site area because anemones, and therefore habitat quality, are unevenly distributed across the site; areas we did not visit are likely to have a lower density of anemones than the areas we did sample.

For scaling the number of tagged recruited offspring to account for areas of our sites we did not sample, we use the overall proportion habitat sampled across all sites and sampling years ( $P_h$ ). We sum the metal-tagged anemones we visited across all sites and years to get the total number of metal-tagged anemones we visited while sampling. We then divide that by the number of anemones we could have sampled, the sum of total metal-tagged anemones across all sites multiplied by the number of sampling years, to get the overall proportion habitat sampled across our sites and sampling years.

*Add details about how sometimes it is  $\neq 1$  if the site doesn't have metal tags?*

*Mention plastic tags?*

		% Habitat surveyed							
Site	# Total anems	2012	2013	2014	2015	2016	2017	2018	
Cabatoan	26	42	58	58	65	73	0	62	
Caridad Cemetery	4	0	75	50	0	50	50	50	
Elementary School	8	0	100	38	88	88	88	100	
Gabas	9	0	0	0	44	44	67	0	
Haina	104	0	6	13	13	10	56	80	
Hicgop South	18	0	67	22	28	56	83	78	
N. Magbangon	105	5	12	40	63	63	0	5	
S. Magbangon	34	41	56	32	0	65	0	71	
Palanas	137	29	58	47	63	85	86	86	
Poroc Rose	13	100	100	69	31	23	69	69	
Poroc San Flower	11	100	82	73	73	55	82	64	
San Agustin	17	94	65	71	65	100	82	76	
Sitio Baybaon	260	0	14	30	33	30	41	80	
Tomakin Dako	50	0	24	22	36	34	60	68	
Visca	13	100	100	23	38	62	85	62	
Wangag	296	18	32	42	34	26	49	68	

Table A2: Table showing the percent of metal-tagged anemones surveyed at each site in each sampling year.

## A.2 Probability of capturing a fish, from recapture dives

624 We use mark-recapture data from recapture dives done within a sampling season to  
estimate the probability of capturing a fish. During some of the sampling years (XX),  
portions of the sites were sampled again XX-XX weeks after the original sampling  
dives. We assume there is no mortality of tagged fish between the original sampling  
dives and the recapture dives because they are so close in time and that fish do not  
change their behavior or response to divers, so therefore assume that the probability  
627 of recapturing a fish is the same as the probability of capturing a fish on a sample dive.  
For each recapture dive, we use GPS tracks of the divers to identify the anemones  
covered in the recapture dive and the set of PIT-tagged fish encountered on those  
630 anemones during the original sampling dives. We estimate the probability of capture  
 $P_c$  as the number of tagged fish caught during the capture dive  $m_2$  divided by the  
total number of fish caught on the recapture dive  $n_2$ :  $P_c = \frac{m_2}{n_2}$ .

636 We use the mean  $P_c$  across all 14 recapture dives, covering XX sites in 3 sampling  
seasons (2016, 2017, 2018), as our best estimate. Because there are so few recapture  
dives compared to the number of times we calculate the metrics to show the range  
639 of uncertainty, we represent the probability of capture as a distribution, rather than  
pulling directly from the values calculated for each recapture dive. The distribution  
of capture probabilities across the 14 dives is quite skewed so we represent it as a  
642 beta distribution, using the mean  $\mu_{P_c}$  and variance  $V_{P_c}$  of the set of 14 values to find  
the appropriate  $\alpha_{P_c}$  and  $\beta_{P_c}$  parameters, where

$$\alpha_{P_c} = \left( \frac{1 - \mu_{P_c}}{V_{P_c}} - \frac{1}{\mu_{P_c}} \right) \mu_{P_c}^2 \quad (\text{A.1})$$

$$\beta_{P_c} = \alpha_{\mu_{P_c}} \times \frac{1}{\mu_{P_c} - 1}. \quad (\text{A.2})$$

The mean of the individual capture probability values is  $\mu_{P_c} = 0.56$ , with variance  
 645  $V_{P_c} = 0.069$ , which gives beta distribution parameters  $\alpha_{P_c} = 1.44$  and  $\beta_{P_c} = 1.13$ .  
 We sample 1000 values from the beta distribution, then truncate the sample to only  
 values larger than the lowest value of  $P_c$  estimated in an individual dive (0.20), to  
 648 avoid extremely low values that are sometimes sampled but are unrealistically low.  
 We then sample with replacement from the truncated set to get a vector of values  
 the length of the number of runs.

## 651 Proportion of dispersal kernel area sampled

[Add in description of calculation and equation]

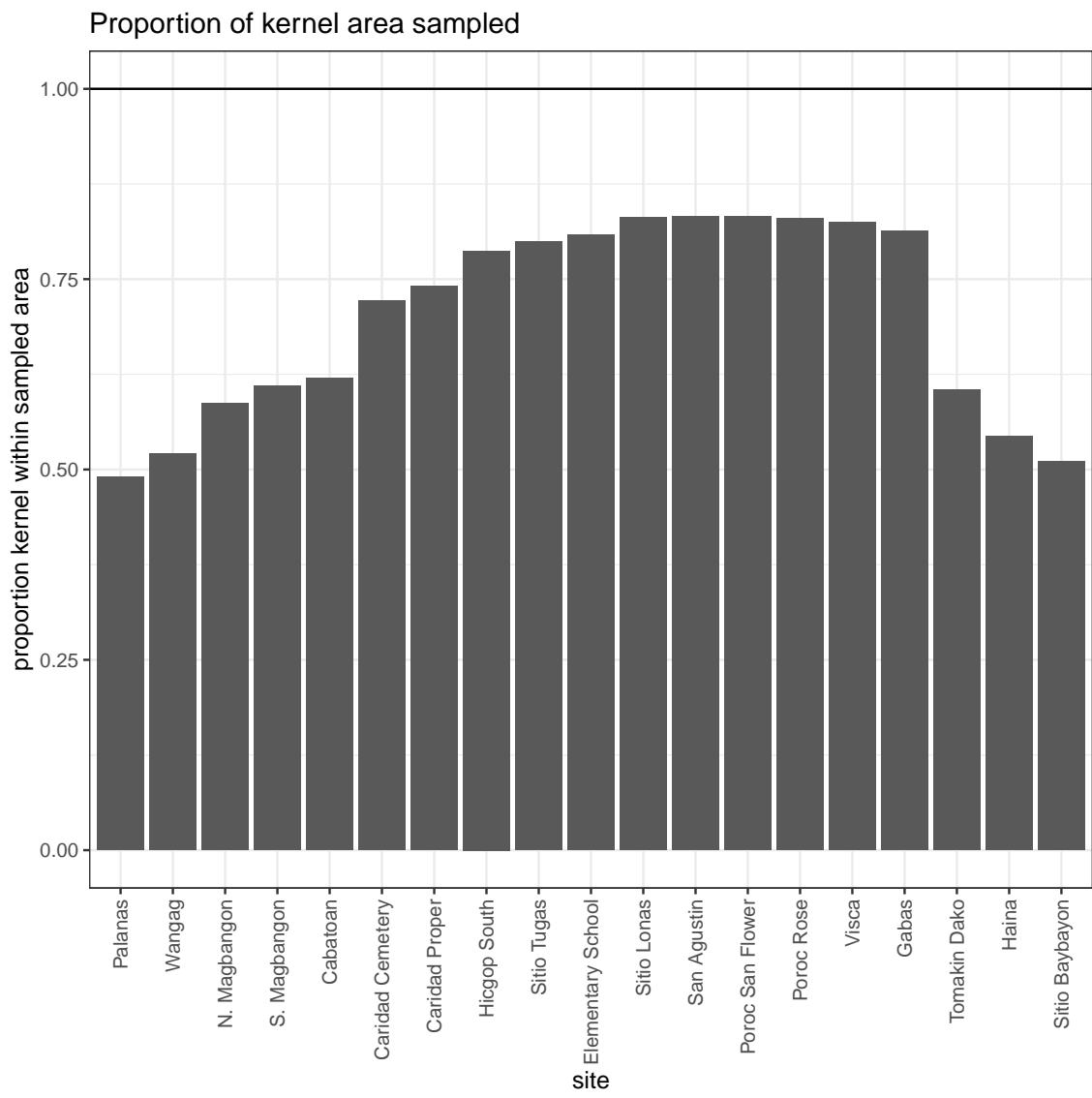


Figure A.1: Proportion of the dispersal kernel area from the center of each site covered by our sampling.

### A.3 Proportion habitat in sampling area

654 We assume that larvae are unable to navigate to habitat if they attempt to settle  
on an unsuitable patch, though clownfish larvae do likely have some ability both to  
sense habitat (CITATIONS) and move toward it (CITATIONS)). To avoid counting  
657 mortality due to settling on non-habitat twice - once in scaling up our matched  
recruits, which only includes those who settled on habitat, and once in integrating  
the dispersal kernel, we scale our estimate of total surviving recruits from our patches  
660 by the proportion of our sampling region that is habitat ( $P_s$ ). We find  $P_s$  by summing  
the lengths of all of our sites, which run approximately north-south, and dividing  
that by the total distance north-south of our sampling region, giving  $P_s = 0.20$ .

<sub>663</sub> **A.4 Full set of MARK models**

We consider the following set of models in MARK [*Need to add in models*]:

Model	Model description	AICc	dAICc
	survival size, recapture size+distance	3348.861	0
	survival size, recapture distance	3359.998	-11.1371
	survival constant, recapture dis- tance	3383.175	34.3141
	survival constant, recapture size+distance	3384.959	36.0981
	survival time, recapture constant	3408.342	59.4816
	survival site, recapture constant	3440.842	91.98112
	survival site, recapture size+distance	3440.842	91.98112
	survival constant, recapture time	3453.609	104.74839
	survival size, recapture size	3527.710	178.84940
	survival constant, recapture con- stant	3570.908	222.04690

Table A3:

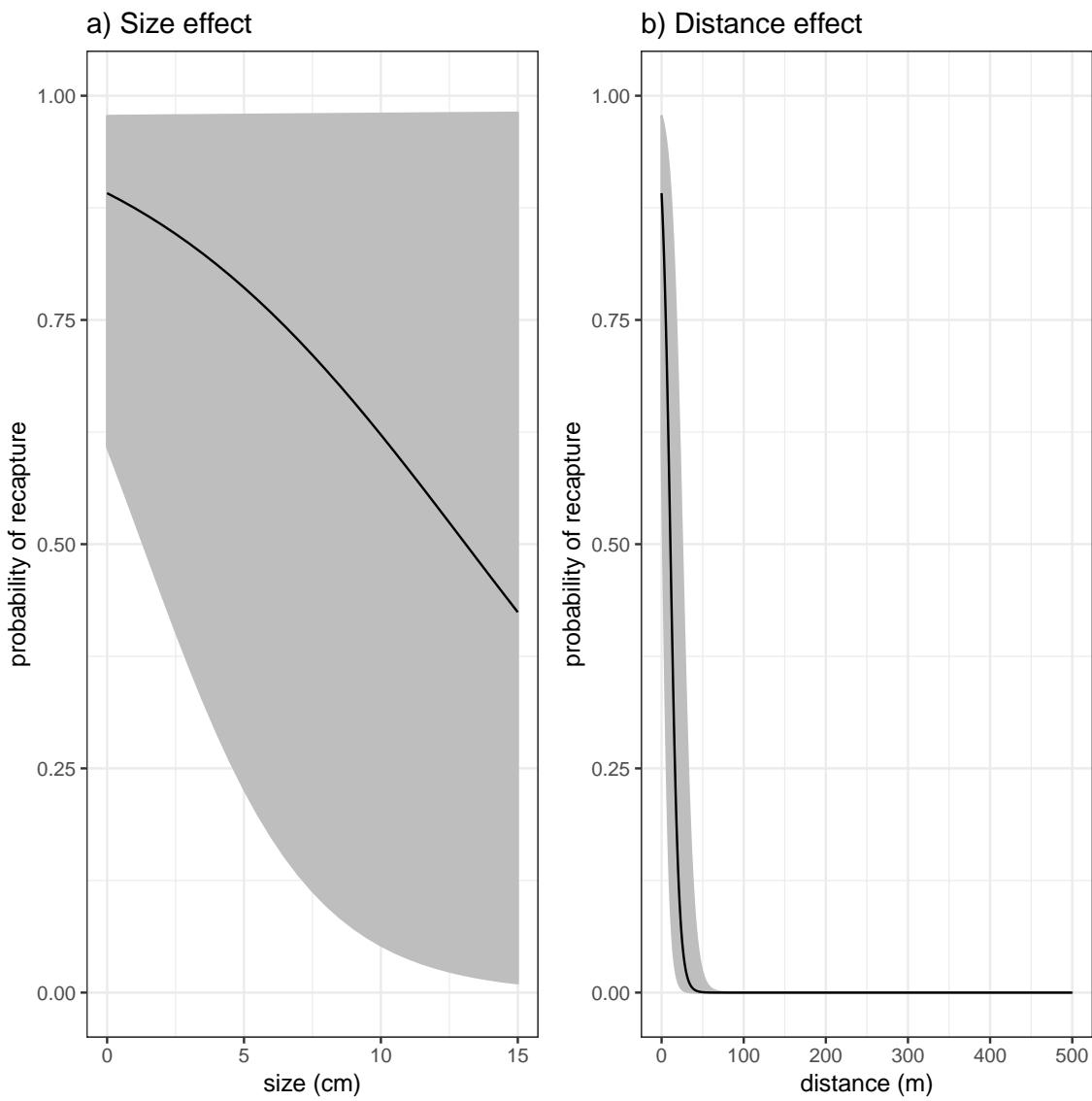


Figure A.2: Effects of a) fish size and b) minimum distance between divers and the anemone where the fish was first caught on the probability the fish will be recaptured, estimated along with survival in a mark-recapture analysis. Size has a slightly negative effect on the probability of recapture - larger fish are stronger swimmers and more likely to flee the anemone when divers approach. Distance also has a negative effect on recapture, as fish tend to stay close to their anemones and are not likely to be recaptured if divers do not get close to their anemone.

## B Uncertainty and sensitivity

### <sup>666</sup> B.1 What-if analyses

To compare our replacement-based persistence results, which do not suggest that our sites make up a persistent metapopulation, with our abundance trends (Fig. 3,  
<sup>669</sup> which suggest that population abundances at our site have been relatively stable over our sampling period, we estimate recruits arriving at our sites per recruit there, regardless of the origin of the arriving recruits. We repeat our metric estimates but  
<sup>672</sup> use all offspring genotyped at our sites, scaled by proportion habitat sampled within our sites  $P_h$  and the probability of capturing a fish  $P_c$ , as our estimate of recruited tagged offspring. We see XXX, which means YYY.

675 **B.1.0.1 All genotyped offspring at our sites originated from our sites**

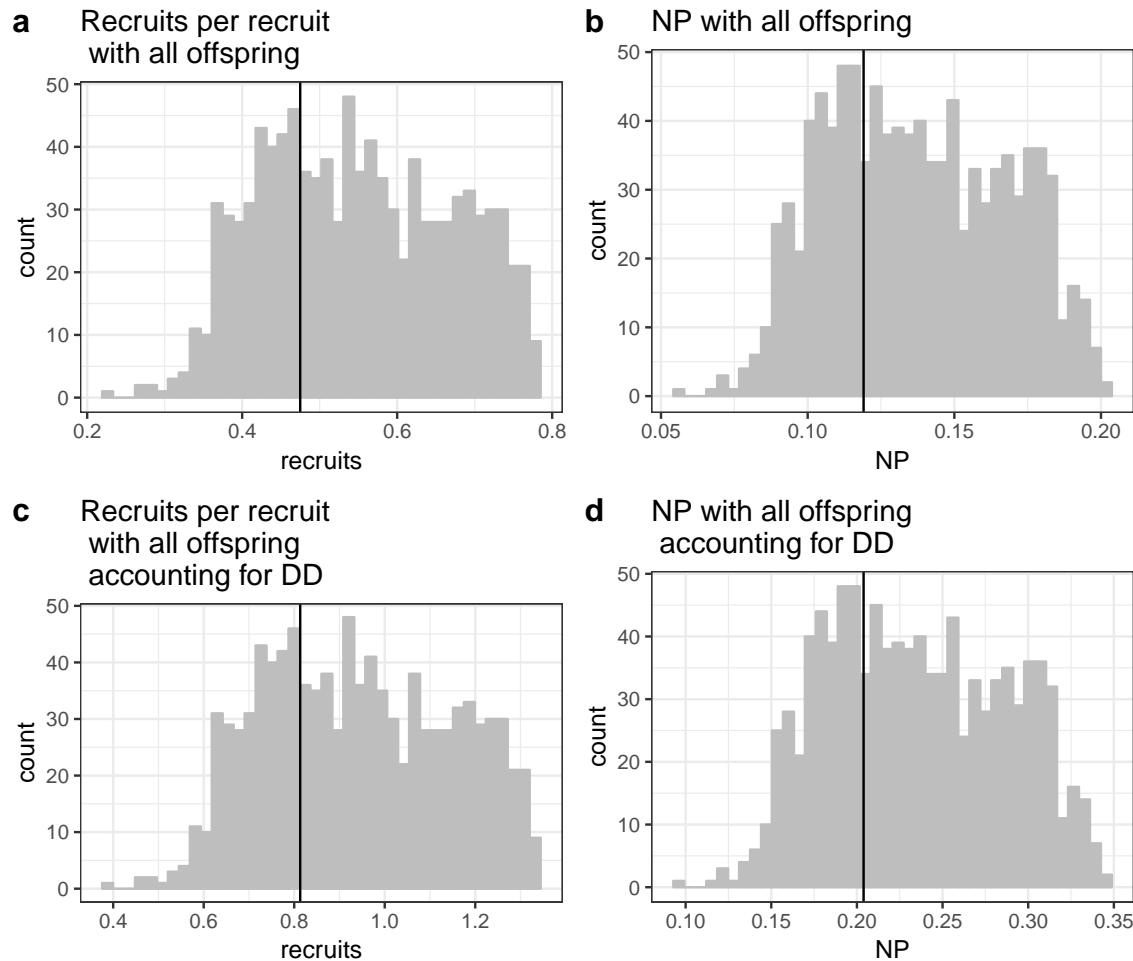


Figure B.1: a) Recruits per recruit when we consider all arriving recruits to have originated from our sites. b) Range of values of NP considering all arriving recruits to be offspring from our sites, with the best estimate in a black solid line.

## B.2 Sensitivity to parameters

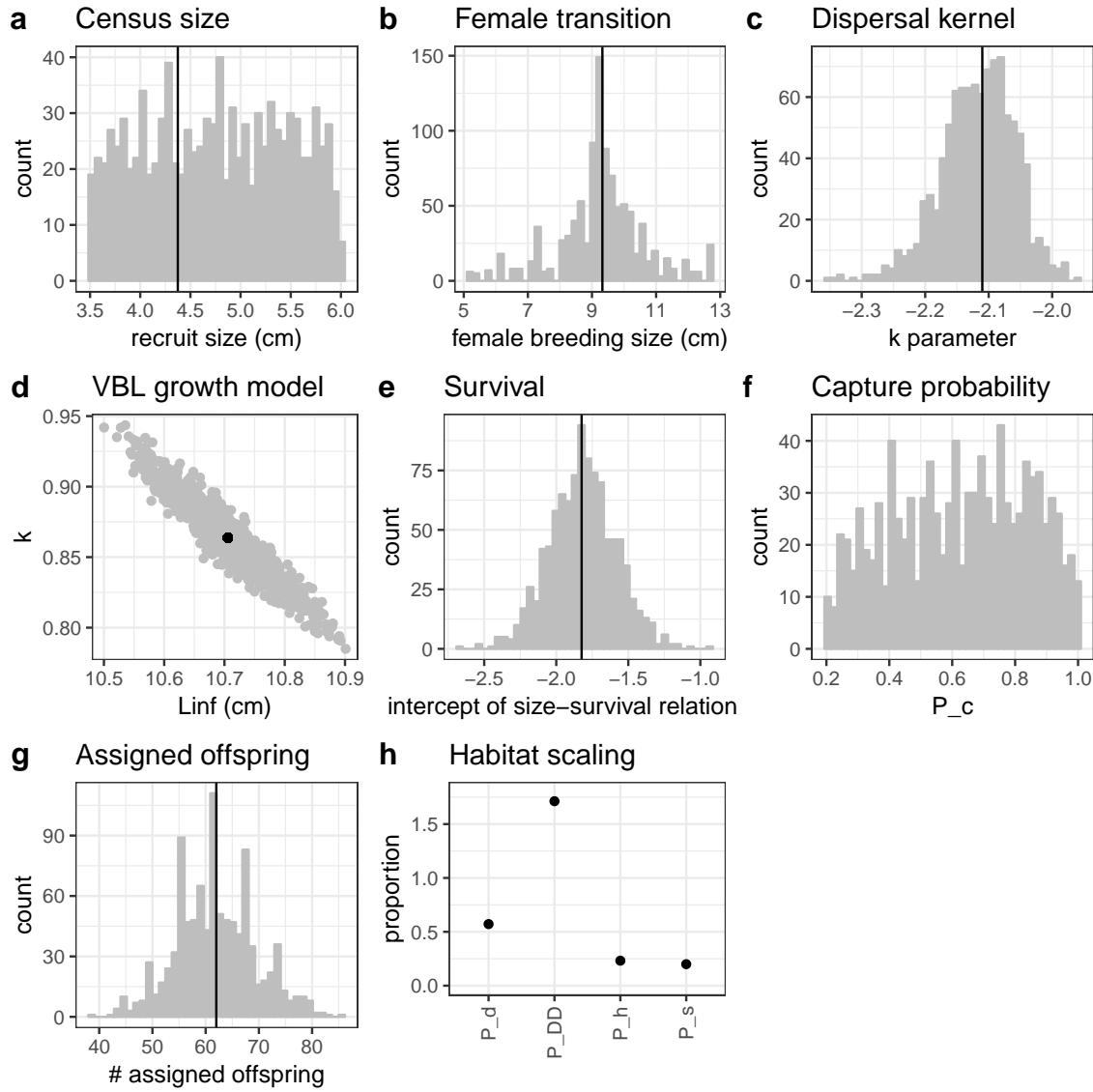


Figure B.2: Range of parameter inputs for uncertainty runs with all uncertainty included: a)  $\text{size}_{\text{recruit}}$ , the census size at which fish are considered to have recruited after egg-recruit survival occurs; b)  $L_f$ , the size at which fish transition from male to female and their reproductive output<sup>56</sup> included in the estimate of lifetime egg production (LEP); c)  $k_d$ , the scale parameter in the dispersal kernel; d) the parameters  $L_\infty$  and  $K$  of the von Bertalanffy growth model; e) the intercept  $b_\phi$  of the adult size-dependent survival relationship; f)  $P_c$ , the probability of capturing a fish; g) number of offspring assigned back to parents in the parentage analysis.

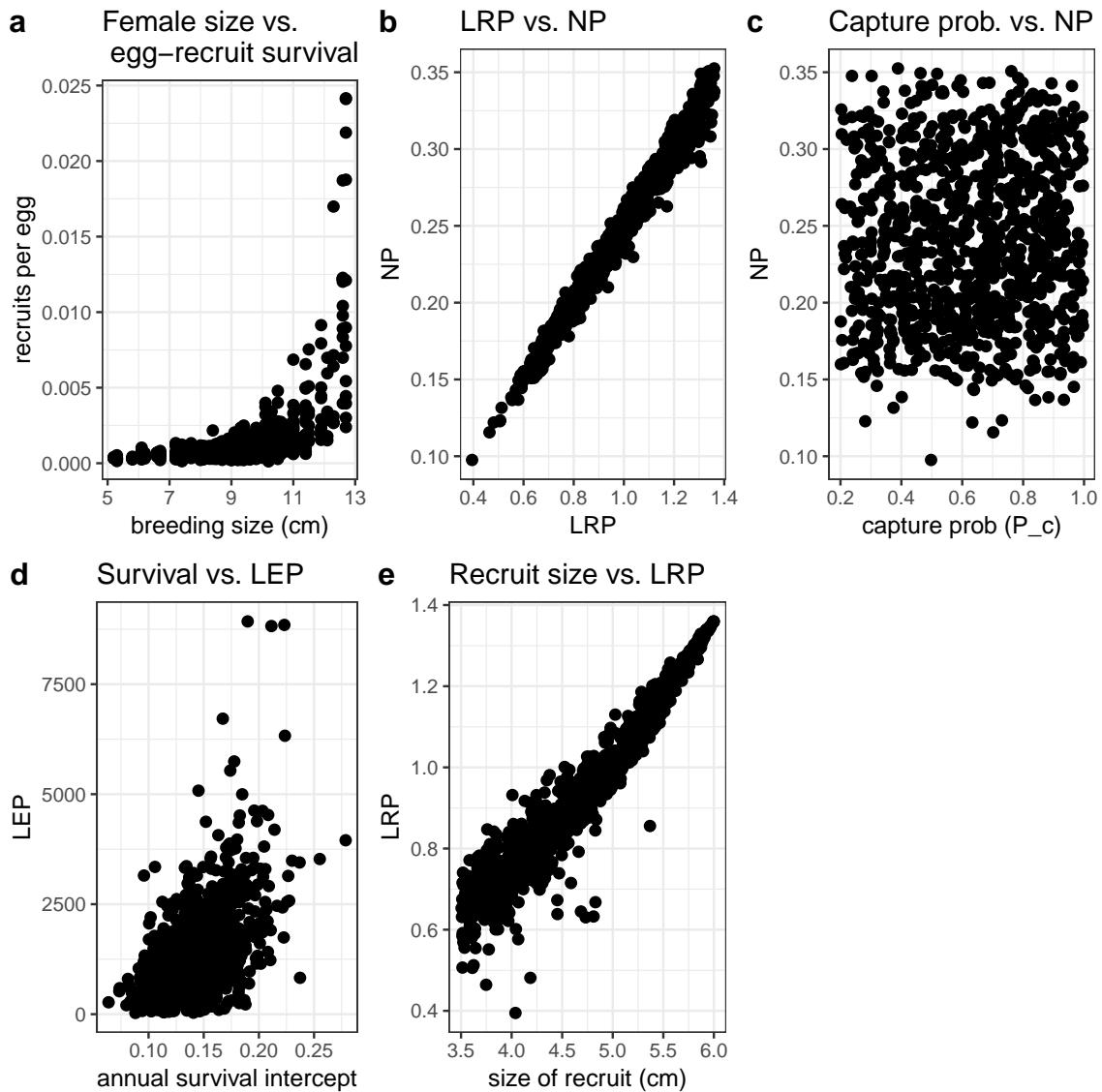


Figure B.3: Relationships among parameters and metrics. a) We only count reproductive effort by fish in the female stage so the higher the transition size to breeding female, the fewer eggs parents are considered to produce, which increases the estimated egg-recruit survival. b) LRP strongly affects NP by changing the number of potential recruits dispersed through the connectivity matrix. c) The probability of capturing a fish does not have a clear relationship to NP. d) LEP is higher with higher survival estimates because fish are more likely to survive longer as reproducing adults. e) The size we consider to be a recruit marks the transition of mortality included in egg-recruit survival to mortality being captured by annual adult survival. Because we do not have the data to change egg-recruit survival to account for

## B.3 Effects of different types of uncertainty on metrics

### <sup>678</sup> B.3.0.1 Lifetime egg production (LEP)

Annual survival post-recruitment provides drives most of the uncertainty in LEP, as lower survivals keep fish from reaching and staying at large breeding sizes, with <sup>681</sup> higher fecundity. The transition size to breeding female also drives uncertainty in LEP - the higher the transition size to female, the less time the fish has at a size where its reproduction is counted in LEP.

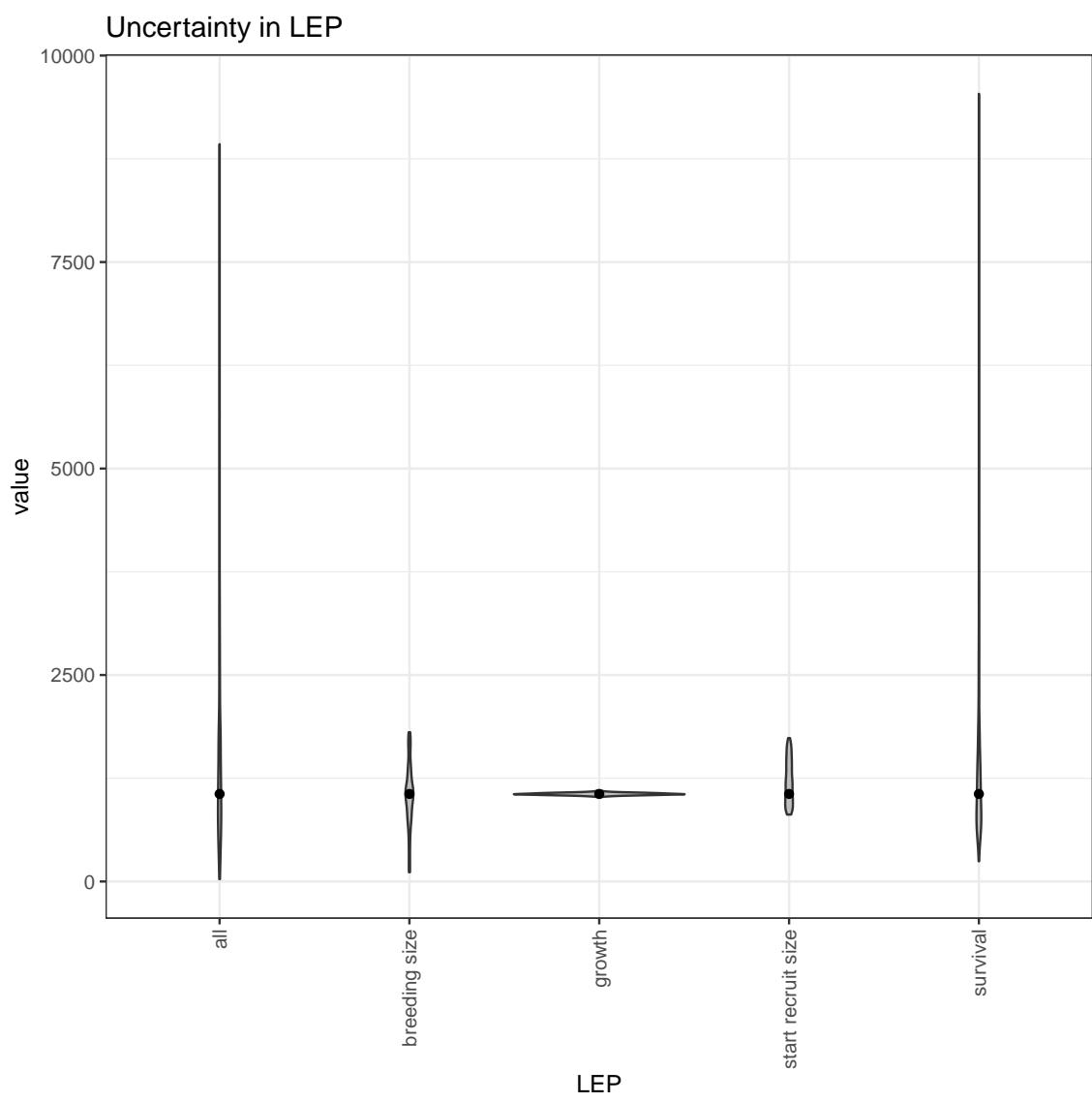


Figure B.4: The contribution of different sources of uncertainty in LEP.

<sup>684</sup> **B.3.0.2 Lifetime recruit production (LRP)**

Most of the uncertainty in LRP comes from uncertainty in the size of a recruit. This is an artifact of our sampling, where we are unable to estimate egg-recruit survival  
<sup>687</sup> differently to account for changes in the size of a recruit, so raising the size of a recruit reduces the mortality included in LRP without increasing the mortality included in egg-recruit survival, as it should in an ideal situation.

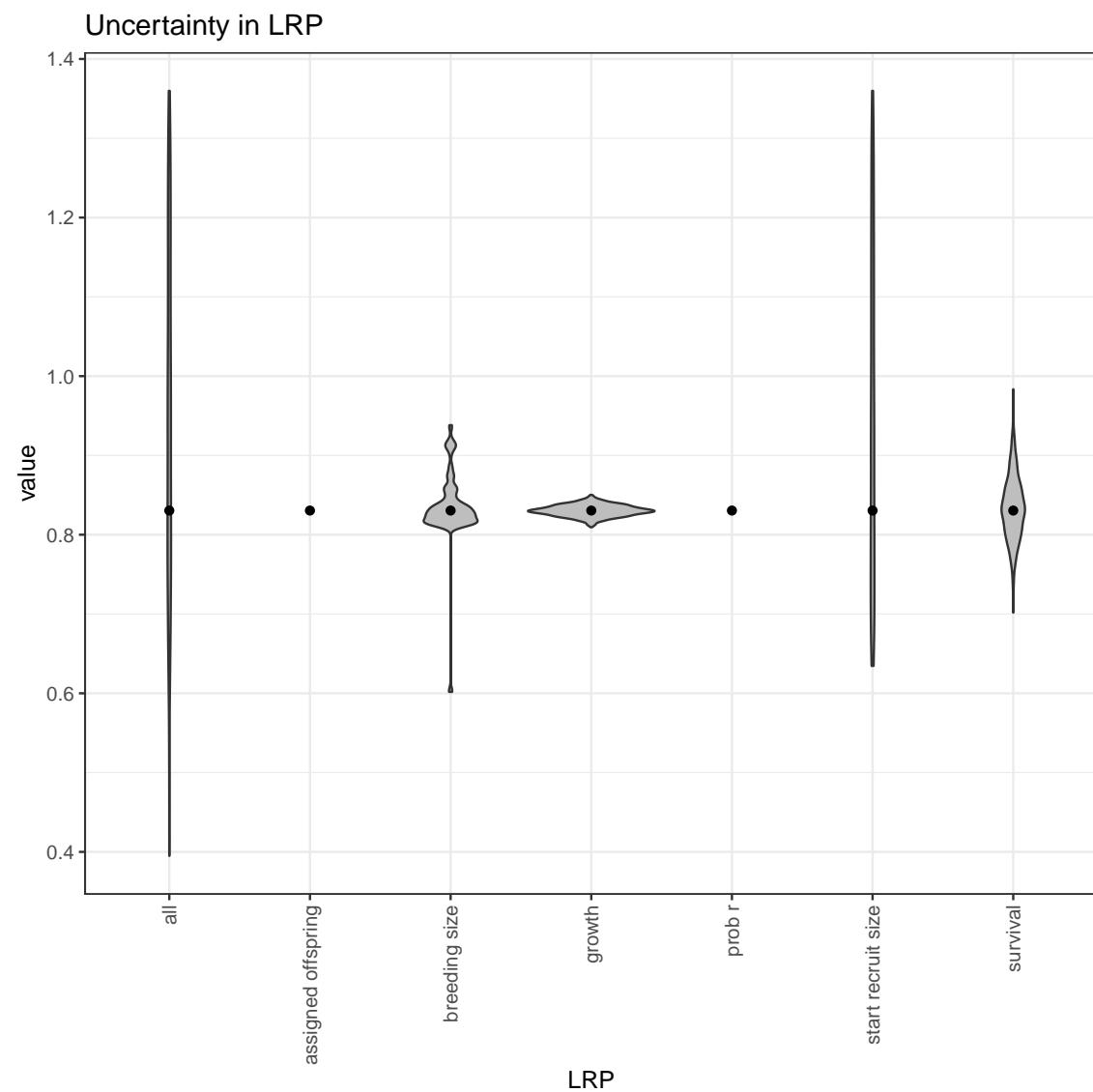


Figure B.5: The contribution of different sources of uncertainty in LRP.

### Uncertainty in LRP accounting for DD

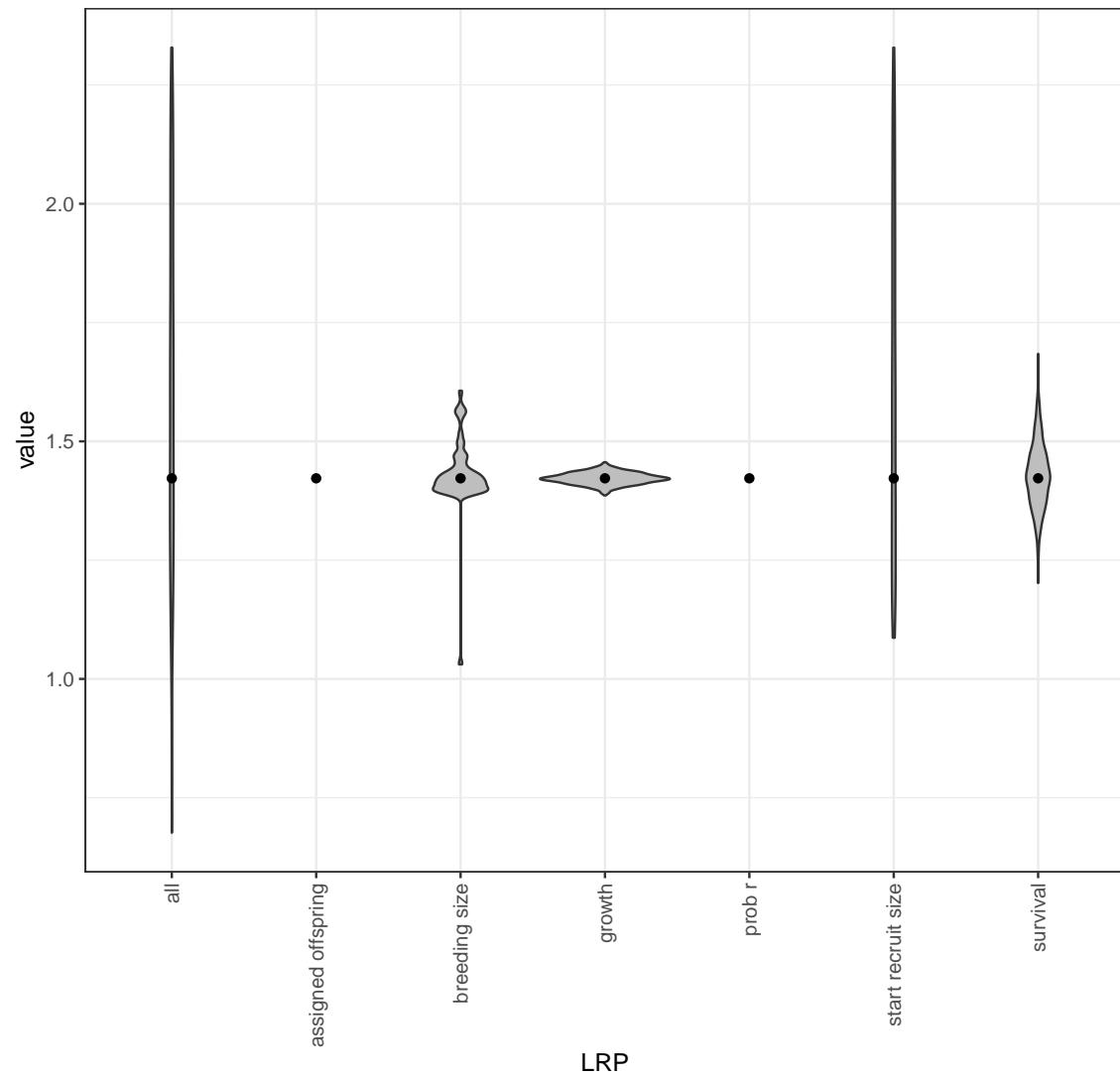


Figure B.6: The contribution of different sources of uncertainty in LRP, when we account for density-dependence in egg-recruit survival.

690 B.3.0.3 Egg-recruit survival ( $S_e$ )

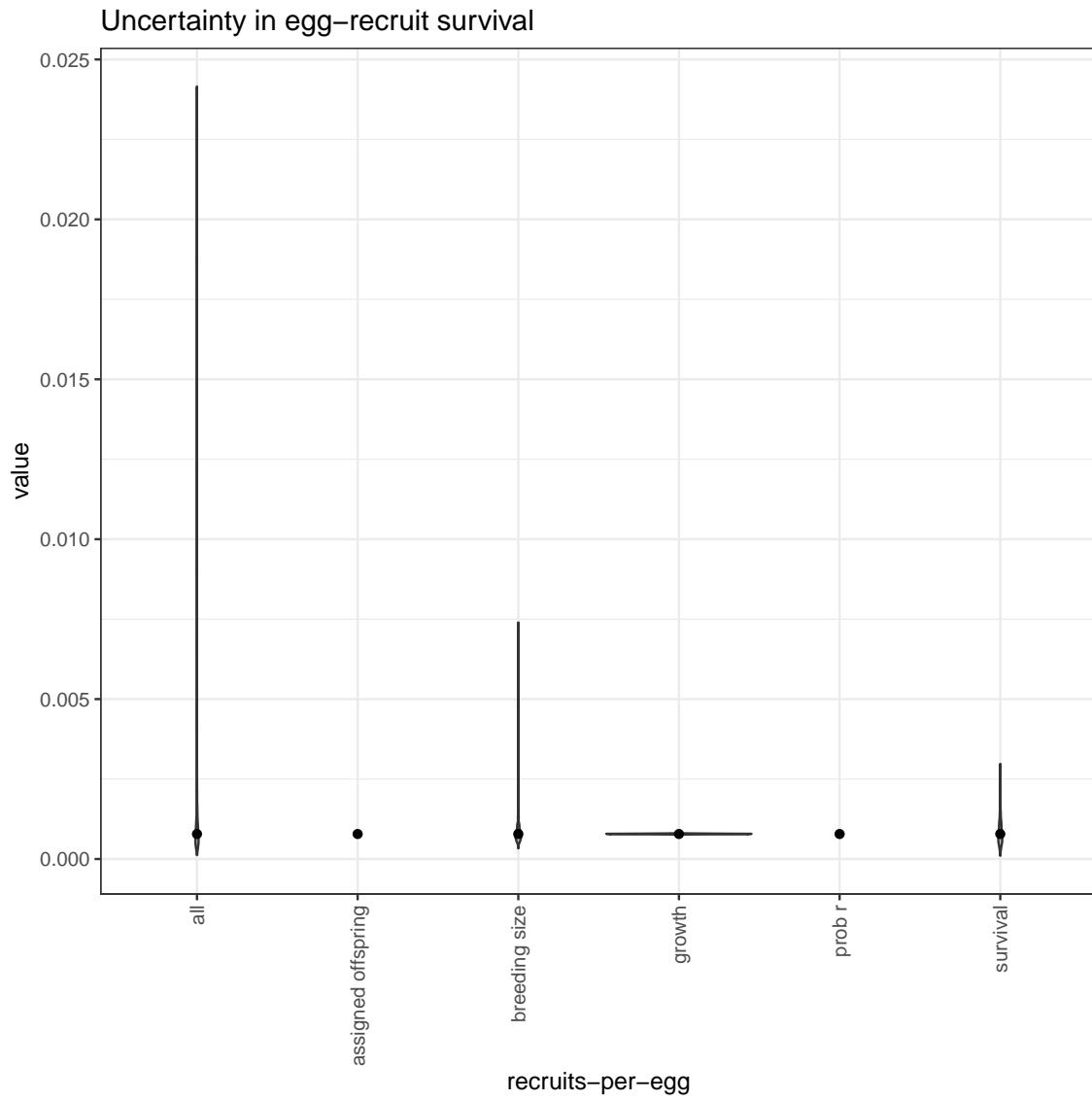


Figure B.7: The contribution of different sources of uncertainty in egg-recruit survival.

### Uncertainty in egg–recruit survival /n accounting for DD

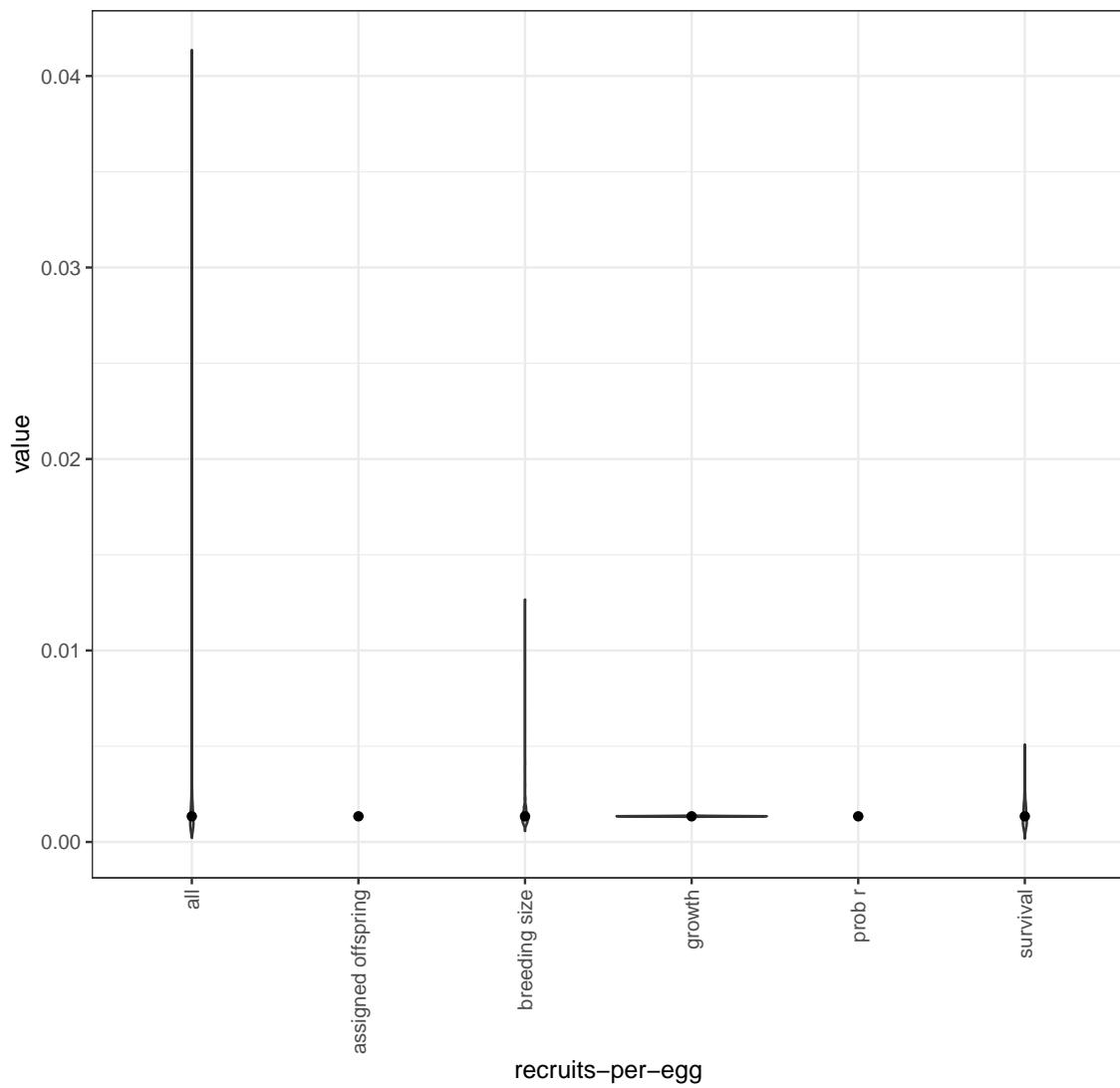


Figure B.8: The contribution of different sources of uncertainty in egg–recruit survival when we account for density-dependence in egg–recruit survival.

#### B.3.0.4 Network persistence (NP)

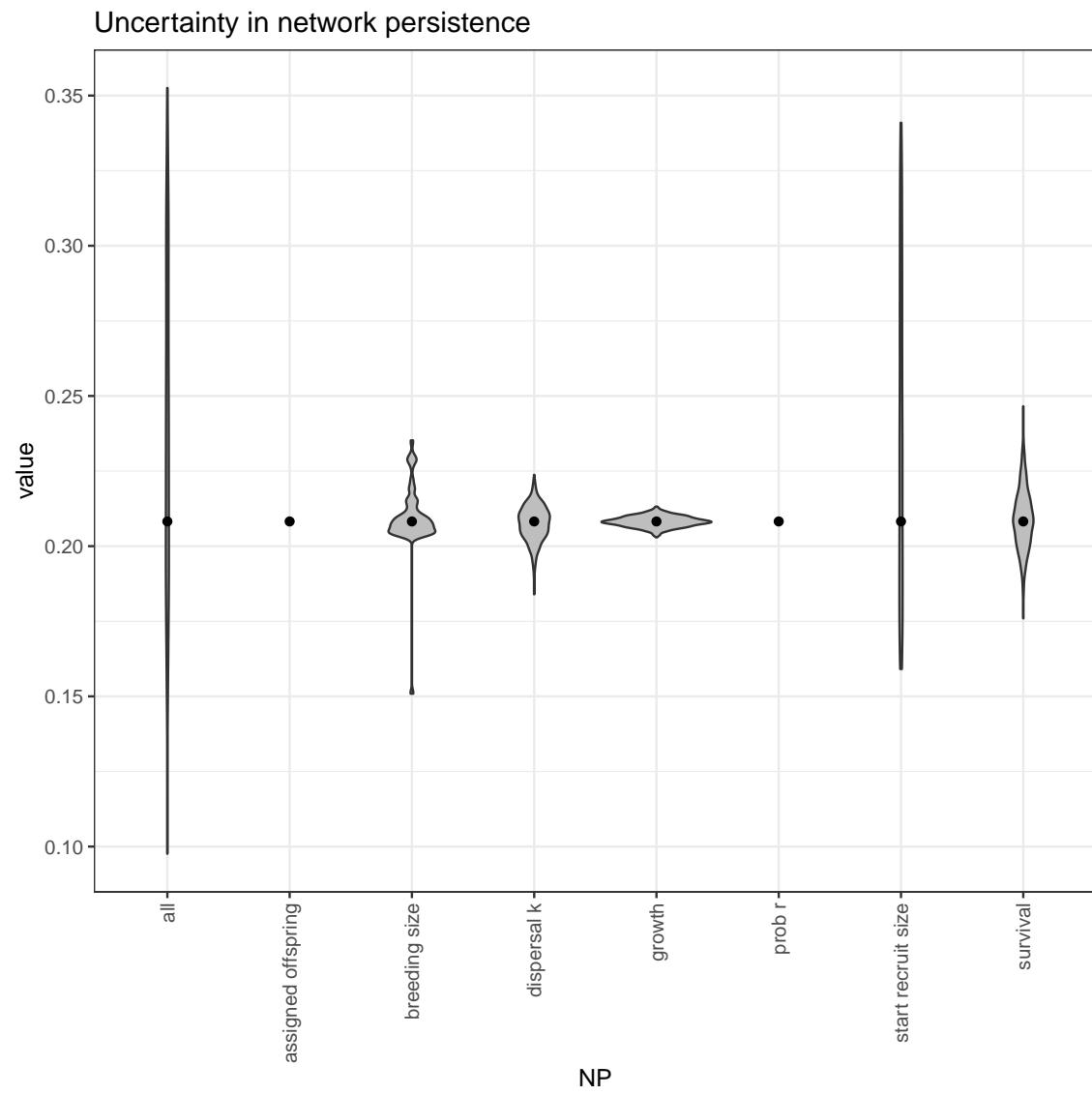


Figure B.9: The contribution of different sources of uncertainty in NP.

### Uncertainty in network persistence /n accounting for DD

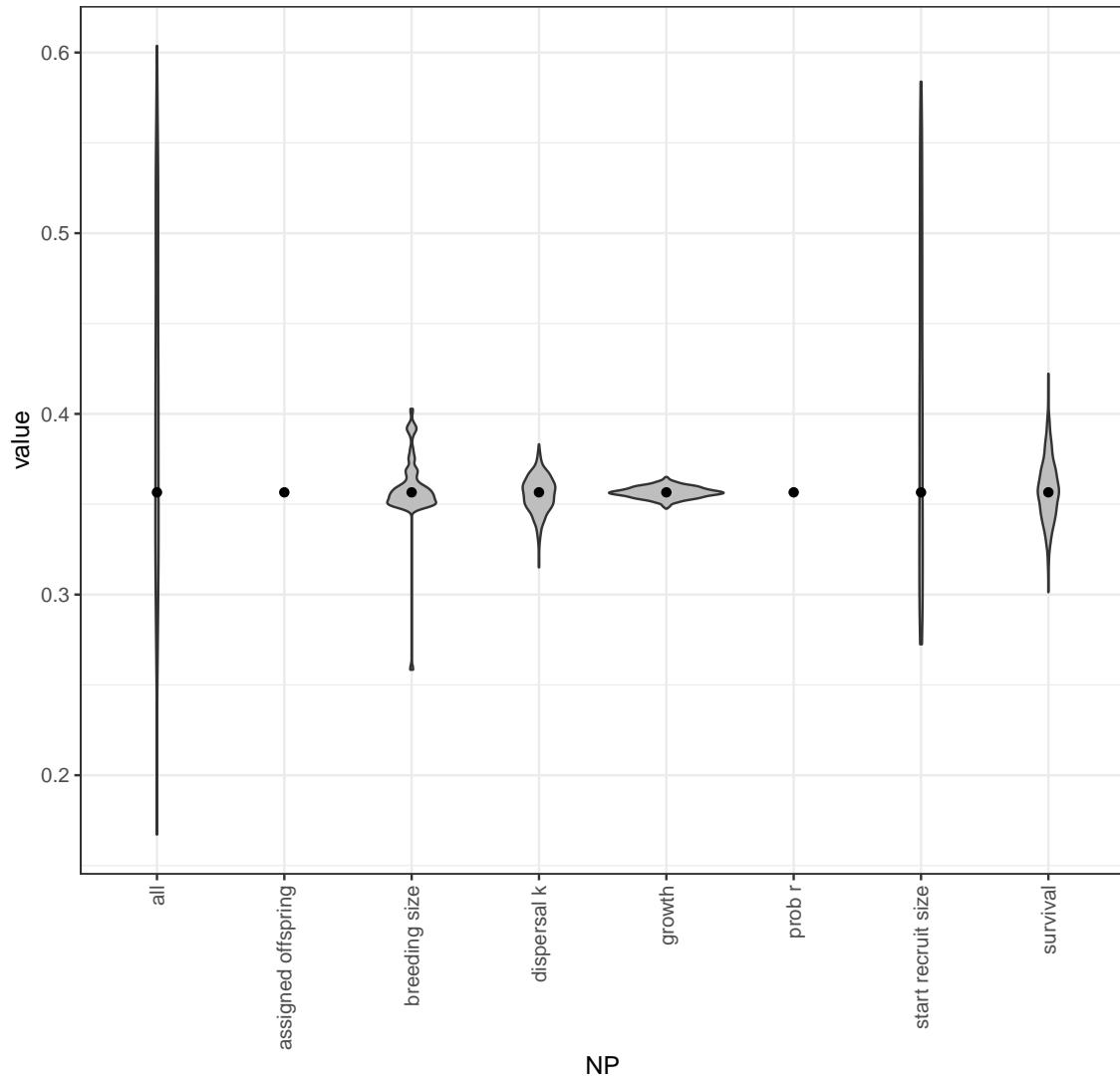


Figure B.10: The contribution of different sources of uncertainty in NP when we account for density-dependence in egg-recruit survival.

## References

- 693 Glenn R Almany, Serge Planes, Simon R Thorrold, Michael L Berumen, Michael  
Bode, Pablo Saenz-Agudelo, Mary C Bonin, Ashley J Frisch, Hugo B Harrison,  
Vanessa Messmer, et al. Larval fish dispersal in a coral-reef seascape. *Nature  
Ecology & Evolution*, 1:0148, 2017.
- 696 Michael Bode, David H Williamson, Hugo B Harrison, Nick Outram, and Geoffrey P  
Jones. Estimating dispersal kernels using genetic parentage data. *Methods in  
Ecology and Evolution*, 9(3):490–501, 2018.
- 699 Louis W. Botsford, Alan Hastings, and Steven D. Gaines. Dependence of sustainability  
on the configuration of marine reserves and larval dispersal distance. *Ecology  
Letters*, 4:144–150, 2001.
- 702 Scott C Burgess, Kerry J Nickols, Chris D Griesemer, Lewis AK Barnett, Alli-  
son G Dedrick, Erin V Satterthwaite, Lauren Yamane, Steven G Morgan, J Wilson  
White, and Louis W Botsford. Beyond connectivity: how empirical methods can  
quantify population persistence to improve marine protected-area design. *Ecolog-  
ical Applications*, 24(2):257–270, 2014.
- 705 Peter Buston. Forcible eviction and prevention of recruitment in the clown anemone-  
fish. *Behavioral Ecology*, 14(4):576–582, 2003a.
- 708 Peter Buston. Social hierarchies: size and growth modification in clownfish. *Nature*,  
424(6945):145–146, 2003b.

Peter M Buston and Cassidy C DAloia. Marine ecology: reaping the benefits of local dispersal. *Current Biology*, 23(9):R351–R353, 2013.

<sup>714</sup> Peter M Buston, Geoffrey P Jones, Serge Planes, and Simon R Thorrold. Probability of successful larval dispersal declines fivefold over 1 km in a coral reef fish. *Proceedings of the Royal Society of London B: Biological Sciences*, page rspb20112041,  
<sup>717</sup> 2011.

Hal Caswell. *Matrix population models: construction, analysis, and interpretation*. Sinauer Associates Inc., Sunderland, Massachusetts, 2nd edition, 2001.

<sup>720</sup> Katrina A Catalano, Allison G Dedrick, Michelle Stuart, Jonathan Purtiz, Humberto Montes, Jr., and Malin Pinsky. Interannual variability of genetic connectivity in a coral reef fish *Amphiprion clarkii*. in prep.

<sup>723</sup> Robert K. Cowen and Su Sponaugle. Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, 1(1):443–466, January 2009. ISSN 1941-1405, 1941-0611. doi: 10.1146/annurev.marine.010908.163757. URL <http://www.annualreviews.org/doi/abs/10.1146/annurev.marine.010908.163757>.

<sup>729</sup> C. C. D’Aloia, S. M. Bogdanowicz, J. E. Majoris, R. G. Harrison, and P. M. Buston. Self-recruitment in a Caribbean reef fish: a method for approximating dispersal kernels accounting for seascape. *Molecular Ecology*, 22(9):2563–2572, May 2013. ISSN 09621083. doi: 10.1111/mec.12274. URL <http://doi.wiley.com/10.1111/mec.12274>.

<sup>732</sup> Augustus J. Fabens. Properties and fitting of the von bertalanffy growth curve. *Growth*, 29:265–289, 1965.

Daphne Gail Fautin, Gerald R Allen, Gerald Robert Allen, Australia Naturalist, <sup>735</sup> Gerald Robert Allen, and Australie Naturaliste. Field guide to anemonefishes and their host sea anemones. 1992.

Sarah O Hameed, J Wilson White, Seth H Miller, Kerry J Nickols, and Steven G <sup>738</sup> Morgan. 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, 2016.

<sup>741</sup> Ilkka Hanski. Metapopulation dynamics. *Nature*, 396(6706):41–49, 1998.

Deborah R Hart and Antonie S Chute. Estimating von bertalanffy growth parameters from growth increment data using a linear mixed-effects model, with an application <sup>744</sup> to the sea scallop *Placopecten magellanicus*. *ICES Journal of Marine Science*, 66 (10):2165–2175, 2009.

Alan Hastings and Louis W. Botsford. Persistence of spatial populations depends on <sup>747</sup> returning home. *Proceedings of the National Academy of Sciences*, 103:6067–6072, 2006.

Jordan N. Holtswarth, Shem B. San Jose, Humberto R. Montes Jr., James W. Morley, <sup>750</sup> and Malin. L Pinsky. The reproductive seasonality and fecundity of yellowtail clownfish (*Amphiprion clarkii*) off the Philippines. *Bulletin of Marine Science*, 93, 2017.

- 753 Darren W Johnson, Mark R Christie, Timothy J Pusack, Christopher D Stallings,  
and Mark A Hixon. Integrating larval connectivity with local demography reveals  
regional dynamics of a marine metapopulation. *Ecology*, 99(6):1419–1429, 2018.
- 756 Jacob P Kritzer and Peter F Sale. *Marine metapopulations*. Elsevier Academic Press,  
2006.

J.L. Laake. RMark: An r interface for analysis of capture-recapture data with  
759 MARK. AFSC Processed Rep. 2013-01, Alaska Fish. Sci. Cent., NOAA,  
Natl. Mar. Fish. Serv., Seattle, WA, 2013. URL [http://www.afsc.noaa.gov/  
Publications/ProcRpt/PR2013-01.pdf](http://www.afsc.noaa.gov/Publications/ProcRpt/PR2013-01.pdf).

- 762 Haruki Ochi. Mating behavior and sex change of the anemonefish, amphiprion clarkii,  
in the temperate waters of southern japan. *Environmental Biology of Fishes*, 26  
(4):257–275, 1989.

765 Brant K Peterson, Jesse N Weber, Emily H Kay, Heidi S Fisher, and Hopi E Hoek-  
stra. Double digest radseq: an inexpensive method for de novo snp discovery and  
genotyping in model and non-model species. *PloS one*, 7(5):e37135, 2012.

- 768 Malin L Pinsky, Humberto R Montes Jr, and Stephen R Palumbi. Using isolation  
by distance and effective density to estimate dispersal scales in anemonefish. *Evo-  
lution*, 64(9):2688–2700, 2010.

771 Jonathan B Puritz, Christopher M Hollenbeck, and John R Gold. ddocent: a radseq,  
variant-calling pipeline designed for population genomics of non-model organisms.  
*PeerJ*, 2:e431, 2014.

- 774 Mark Rees, Dylan Z Childs, and Stephen P Ellner. Building integral projection  
models: a user's guide. *Journal of Animal Ecology*, 83(3):528–545, 2014.
- 777 J Roughgarden, S Gaines, and H Possingham. Recruitment dynamics in complex  
life cycles. *Science*, 241(4872):1460–1466, September 1988. ISSN 0036-8075, 1095-  
9203. doi: 10.1126/science.11538249. URL <http://www.sciencemag.org/cgi/doi/10.1126/science.11538249>.
- 780 Ocane C. Salles, Jeffrey A. Maynard, Marc Joannides, Corentin M. Barbu, Pablo  
Saenz-Agudelo, Glenn R. Almany, Michael L. Berumen, Simon R. Thorrold, Ge-  
offrey P. Jones, and Serge Planes. Coral reef fish populations can persist without  
783 immigration. *Proceedings of the Royal Society B: Biological Sciences*, 282(1819):  
20151311, November 2015. ISSN 0962-8452, 1471-2954. doi: 10.1098/rspb.2015.  
1311. URL <http://rspb.royalsocietypublishing.org/lookup/doi/10.1098/rspb.2015.1311>.
- 786 Jinliang Wang. Sibship reconstruction from genetic data with typing errors. *Genetics*,  
166(4):1963–1979, 2004.
- 789 Jinliang Wang. Computationally efficient sibship and parentage assignment from  
multilocus marker data. *Genetics*, 191(1):183–194, 2012.
- 792 Jinliang Wang. Estimation of migration rates from marker-based parentage analysis.  
*Molecular ecology*, 23(13):3191–3213, 2014.
- Jw White, Lw Botsford, A Hastings, and Jl Largier. Population persistence in ma-  
rine reserve networks: incorporating spatial heterogeneities in larval dispersal.

795      *Marine Ecology Progress Series*, 398:49–67, January 2010. ISSN 0171-8630, 1616-  
1599. doi: 10.3354/meps08327. URL <http://www.int-res.com/abstracts/meps/v398/p49-67/>.

798    Adam Yawdoszyn. Fecundity in clownfish. in prep.