

Supplemental Experimental Procedures

Natural History Information

Snapper (*Chrysophrys auratus*, Sparidae) are an abundant coastal, demersal sparid fish that are widely distributed in the temperate to sub-tropical western Pacific waters of the southern and northern hemispheres [1]. They support valuable coastal commercial and recreational fisheries, within the Hauraki Gulf in northeast New Zealand supporting the country's largest and most valuable stock. Snapper are long-lived (60 years) and maturity is reached at 3-5 years at which time around half the population changes sex from female to male [2]. Multiple tagging studies have shown that adult snapper display a range of movement patterns from highly territorial [3-5], to long-distance movements (418 km) [6]. They are serial broadcast spawners producing batches of eggs daily during an extended spring-summer season (November to February) [7]. Snapper larvae include rapid egg development (~1 day) and relatively short larval duration (18-32 days) [8]. The latter two factors equate to a relatively short planktonic larval duration, the most cited biological variable potentially affecting self-recruitment [9]. Low levels of genetic differentiation have been found among populations of adult snapper from both New Zealand and Australia using several highly variable genetic markers [10, 11]. This suggests relatively high gene flow between snapper populations. The current lack of knowledge of the scale and role of snapper larval dispersal from spawning grounds prevents us from ascertaining their contribution to the surrounding exploited population without the approaches used in this study.

Field sampling

The sampling targeted adult fish (>230 mm fork length FL) (spawners) inside the CROP MPA and their potential offspring sampled at multiple coastal locations either side of the reserve boundary (Figures 1A, 2A, Table S1). A total of 1053 adults were

non-lethally sampled from twenty-three (23) sites inside the marine reserve between September 2011 and February 2012 (Figure 1A). Sampling sites encompassed a wide range of habitats [12] and depths (2 -20 m) throughout the MPA. Each fish was caught on rod and reel with barbless circle hooks to minimise hook ingestion, and then measured to the nearest millimetre. Prior to their release, a uniquely coded 12mm-long plastic coated 'food-safe' PIT (Passive Integrated Transponder) tag (Hallprint, Australia) was injected in the gut cavity using a purpose-built gauge hypodermic syringe. This allowed us to re-identify fish (using a scanning wand), and thus prevent repeat sampling of individuals, while also obtaining recapture estimates. A small section of the pectoral fin was removed from each fish for genotyping.

For genetic tracing of offspring, we collected 933 juveniles non-lethally using squid-baited Sabiki rigs from March to May 2012. Four (4) sites sampled inside the MPA (Table S1) yielded a total of 78 samples, while 855 individuals were sampled from 33 sites within 7 areas up to 4 km north ($n=108$) and 40 km south ($n=747$) of the MPA boundary (Figure 2A, Table S1). Snapper juveniles were caught predominantly along the rocky reef/sand flat interface in 3-5m depth. Ross et al [13] found that juvenile snapper were most abundant in this type of habitat complexity, attributing this preferred distribution to a balance between the requirements of food acquisition and predator avoidance. However, compared to coral reef fish for which high levels of habitat specificity have been shown, snapper are more generalist, and typically recruit to various types of structured habitat, such as seagrass beds, horse mussels, sponge gardens and urchin barrens [14]. Each juvenile was measured to the nearest millimetre, and a small section of the pectoral fin was removed prior to release.

Adult snapper outside the CROP MPA (Non-MPA adults, $n = 1051$) were sampled concurrently as potential alternative parents from March 2011 to February 2012 (Figure 1C, Table S1). Fish samples were collected from local fishing competitions,

provided by supportive recreational fishers, and collected directly from boat ramps. Our sampling comprised 33 locations within 9 general areas spanning 55 km north (Bream Bay) to 45 km south (Whangaparaoa Peninsula) of the CROP marine reserve. All samples collected for this project were preserved in 95% ethanol and returned to the laboratory for subsequent genotyping.

DNA Extraction and microsatellite typing

Genomic DNA was extracted from ~2mm² of fin tissue from each sample (n = 3037) using the Gentra Puregene (Qiagen, Valencia, CA) blood and tissue kit and eluted with 200µl TE buffer (pH 8.0). A panel of 17 microsatellite loci previously described for congeneric species [15-18], or species-specific [19] were amplified using the Type-it Microsatellite PCR kit (Qiagen) following the manufacturer's protocol with annealing temperatures ranging from 57 °C to 62°C (Table S2). Primers were fluorescently labelled and pooled in five multiplex reactions (M1-M5) with up to four loci per reaction. Primer concentrations were adjusted for even amplification and ranged from 0.15 to 0.40 µM (Table S2). All reactions were in a total volume of 10 µL containing 5 µL of QIAGEN Multiplex Master Mix (2x), 3 µL of distilled water, 1 µL of primer premix, and 1 µL template DNA (10ng). The PCR thermal cycling program was: initial denaturation at 94°C for 15min, 35 cycles at 94°C for 30 s, primer specific Ta for 90 s (Table S2), 72°C for 90 s, and final extension at 72°C for 10 min. Two loci (CM278 & PaurGT2) negatively interacted when combined with other loci in a multiplex reaction. These were run in two singleplex reactions (S1 and S2). Each singleplex 10 µL PCR contained 10ng DNA template, primer concentrations as per Table S2, 50mM of Magnesium chloride (MgCl₂), 20mM dNTP mix, 0.125 U Platinum(R) Taq DNA Polymerase (Invitrogen, Life Technologies), with associated PCR buffer (Table S2). The PCR amplification program was: initial denaturation at 95°C for 1 min, 40 cycles at 95°C for 15 s, primer specific Ta for 30 s (Table S2), 72°C for 1 min, and final extension at 72°C for 5 min. Each 96-well tray run included

four positive controls known to consistently amplify well across all 17 loci, and a negative control (blank) to test for cross-contamination. Diluted amplicons were combined with formaldehyde and GeneScan LIZ-500 size standard (Life Technologies) and screened on an ABI 3730XL automated sequencer (Applied Biosystems) at Macrogen Inc. (Korea). Allele sizes were determined with the fragment analysis software GENEMAPPER v. 4.0 (Life Technologies).

Genotyping error quantification

Genotyping error was estimated by counting the error rate (number of mismatches between the reference genotype and the replicates) per allele [20] using the four positive control samples loaded into each 96-well tray, plus a proportion of re-run samples, giving a total of 204 repeated samples. These were amplified at 17 loci and screened on an ABI 3730XL automated sequencer following the protocol described above. The estimated mean allelic error rate was 1% across all loci.

Summary statistics

Of the 3037 snapper samples collected, a total of 2972 snapper samples were genotyped successfully (1048 MPA Adults, 892 juveniles, 1032 Non-MPA Adults) with a panel of 17 microsatellite markers grouped into five multiplexes and two singleplexes (Table S2). Summary statistics (number of alleles, number of genotyped individuals, observed and expected heterozygosities, and *F_{is}*) were generated for the three datasets in CERVUS v. 3.0 [21] and GenAlEx v. 6.5 [22] (Table S3).

The microsatellite markers were highly polymorphic with an average of 17.7 alleles per locus, ranging from 5 to 27 (Table S3). Mean expected heterozygosity was 0.73 +/-0.04 SE (0.356 – 0.910) per locus, and mean observed heterozygosity was 0.66 +/-0.04 SE (0.270 – 0.875) per locus. None of the locus pairwise comparisons showed evidence of significant linkage disequilibrium. Therefore all 17 loci were used in parentage analyses. This set of highly polymorphic markers provided a combined

probability of identity for the 17 loci dataset of $P = 7.4 \times 10^{-19}$, with a combined non-exclusion probability of $P = 1.4 \times 10^{-4}$ for the first parent, and $P = 1.26 \times 10^{-11}$ for a parent pair.

Parentage and relationship assignments

Categorical allocation of parent-offspring relationships was assessed using a maximum likelihood approach implemented in the software FaMoZ [23], and confirmed using CERVUS [21] and Colony [24]. Results from CERVUS and Colony are not presented here, but were comparable to those run in FaMoZ. FaMoZ computes log of the odds ratio (LOD) scores for assigning individuals to candidate parents based on the observed allelic frequencies at each locus. In the present study 10,000 simulated offspring were generated from genotyped parents and allele frequencies. These simulations allow the inclusion of an error term to take into account genotyping error. We used an error rate of 0.01% that minimised type I and II errors related to the parentage tests [25, 26]. The final LOD score threshold values were 6.8 for single-parent assignments and 20.0 for two-parent assignments. This parameter set was evaluated using the 'parentage test simulation' option to estimate the probability of excluding a true parent knowing that it was in the sample (Type I error) and the probability of assigning a false parent knowing that the true parent was not sampled (Type II error). The resulting Type II error was 1.17%, while the Type I error was 0.08%. All potential offspring were screened against the two pools of potential parents, MPA and non-MPA adults, to identify parent-offspring relationships. No missing data occurred across 17 loci. No more than 2 mismatches were ever allowed to occur to accommodate for genotyping error.

Assignment of other relationships "half-sib" between juvenile recruits and all adults were determined using COLONY, and checked using ML-Relate [24, 27]. These approaches assign relationships to this category using maximum likelihood

approaches, although the level of relatedness could also be assigned to alternative categories (e.g., 2nd generation offspring). Thus, the proportions of relationships allocated to this category (with $P > 0.99$) are used here as a proxy for the overall proportion of relationships that are significantly closer than random, rather than an indicator of that specific relationship (i.e., individuals that share 1 parent). An advantage of using these more distant relationships is that they are more common, less subject to sampling error than parent-offspring pairs, and may show relationships over more than one generation. The proportions of parent-offspring and “half-sib” pairs were calculated for each of the collection areas listed in Table S1.

Bayesian model for estimating population size and total contribution of the MPA to the fishery

A simple Bayesian model was used to obtain estimates and uncertainty intervals for the total proportion of juvenile snapper outside of the MPA that had parents from inside the MPA, i.e. the proportionate matched to MPA-based parents (P_C). As only a small sample of the spawning stock of the CROP MPA were sampled, many potential parentage matches with sampled larvae were missed. Our model corrected for this incomplete sampling using methods described by Harrison et al. [28], simultaneously accounting for uncertainty in the proportion of fish that were matched to MPA parents and a simple mark-recapture estimate of the proportion of the spawners inside the MPA that were sampled and genotyped. The population of adults in the MPA (N_R) was estimated using the number of MPA adults caught in Phase 1 during the initial fishing event inside the CROP MPA (September 2011 to February 2012, 13 fishing days, $M = 827$), the number of adults caught in Phase 2 during the second fishing event in the MPA (January to February 2014, 7 fishing days, $C = 483$), and the number of fish caught in Phase 2 that were recaptures from Phase 1 ($R = 34$). It was assumed that R was drawn from a binomial distribution,

$$R \sim \text{Bin}(\pi_R, C)$$

where π_R is an unknown parameter representing the expected proportion of Phase 2 fish that were previously caught during Phase 1. This provided the basis for estimating of the number of adult fish in the MPA, using the simple Lincoln-Petersen equation, while accounting for uncertainty in P_R .

$$N_R = \frac{M}{\pi_R}$$

The total number of adult MPA fish that were genotyped was $S = 1048$. The estimate of the proportion of MPA-based adults that were genotyped was thus

$$P_S = \frac{S}{N_R} = \frac{S \pi_R}{M}$$

In Phase 3, the number of juvenile fish that were caught and tested for parentage by MPA-based fish was given by n . The number of juvenile fish that were positively matched to parents in the MPA was assumed to have a binomial distribution

$$m \sim \text{Bin}(\pi_M, n)$$

π_M was thus a stochastic parameter representing the expected proportion of juvenile fish with parents in the MPA that were genotyped. This probability required a correction for the incomplete sampling of the potential parents in the MPA. Following Harrison et al. [28], the estimated corrected proportion of juvenile fish contributed by parents in the MPA was then calculated as

$$P_C = \frac{\pi_M}{1-(1-P_S)^2}$$

The probability parameters π_R and π_M were each given the flat prior distributions, $\pi \sim \text{beta}(1, 1)$. Treating these parameters as stochastic and unknown random variables allowed us to simultaneously account for the uncertainty in both, thereby providing rigorous estimates of uncertainty in the quantities of interest. The models were fit by MCMC using the software OpenBUGS [29]. Three independent chains were run for 10,000 iterations each after a burn-in of 1000.

Biophysical model development and data analysis

We constructed and parameterised an *a priori* dispersal model to simulate dispersal of the larval stage of *C. auratus* from the CROP Marine Reserve. A hybrid 2D/3D RMA-10 hydrodynamic model was used, with a 3D section near CROP, based on the model described in Le Port et al [30]. For the present work, the model simulated the time period 1st November 2011 to 28th December 2011, to fit in not only with the genetic field sampling (Oct 2011-Feb 2012), but also with the time of known peak snapper spawning activity. To simulate larval dispersal, the hydrodynamic model was again coupled to a particle-tracking model, but this was modified to allow for direct comparisons with the genetic field data. The particle tracking software was modified to handle multiple spawning events (egg releases) at arbitrary times during the simulation. This included daily releases of 500 particles (i.e., virtual larvae) at approximately 17:00 each day for the first 28 days of simulation, totalling 14000 particles released over the course of the simulation. No particles were released between days 29 to 56 of the simulation. Using a continuous release scenario and longer simulation time means our results are less dependent on specific weather events (e.g. storms). The present study simulated the vertical behaviour of snapper larvae using the 3 vertical behaviour scenarios described in Le Port et al [30] (none, "*C. auratus*" and "*C. major*"). As results did not vary greatly among the three behaviour models tested, the "*C. auratus*" behaviour [31] was selected. To get a more accurate measure of where larvae were positioned when ready to start settling, we looked at particle positions at age 28 days, regardless of release time. Sampling zones for counting particles were modified to correspond with genetic field sampling zones. These results from these simulations were compared with our empirical measurements of larval dispersal. All distances presented in this article were the shortest distances through water, to allow for larvae not being able to travel across land and thus having to travel around land masses such as peninsulas and islands. Through-water distances were calculated using the *gdistance* package for R [32].

Supplemental References

1. Paulin C.D. 1990 *Pagrus auratus*, a new combination for the species known as "snapper" in Australasian waters (Pisces: Sparidae). *New Zealand Journal of Marine and Freshwater Research* **24**, 259-265.
2. Francis M.P., Pankhurst N.W. 1998 Juvenile sex inversion in the New Zealand snapper *Chrysophrys auratus* (Bloch and Schneider, 1801) (Sparidae). *Australian Journal of Marine and Freshwater Research* **39**, 625-631.
3. Egli D.P., Babcock R.C. 2004 Ultrasonic tracking reveals multiple behavioural modes of snapper (*Pagrus auratus*) in a temperate no-take marine reserve. *ICES Journal of Marine Science* **61**, 1137-1143. (doi:10.1016/j.icesjms.2004.07.004).
4. Parsons D.M., Babcock R.C., Hankin R.K.S., Willis T.J., Aitken J.P., O'Dor R.K., Jackson G.D. 2003 Snapper *Pagrus auratus* (Sparidae) home range dynamics: Acoustic tagging studies in a marine reserve. *Marine Ecology Progress Series* **262**, 253-265.
5. Willis T.J., Parsons D.M., Babcock R.C. 2001 Evidence for long-term site fidelity of snapper (*Pagrus auratus*) within a marine reserve. *New Zealand Journal of Marine and Freshwater Research* **35**(3), 581-590.
6. Paul L.J. 1967 An evaluation of tagging experiments on the New Zealand snapper, *Chrysophrys auratus* (Forster), during the period 1952 to 1963. *New Zealand Journal of Marine and Freshwater Research* **1**, 455-463.
7. Crossland J. 1977 Seasonal reproductive cycle of snapper *Chrysophrys auratus* (Forster) in the Hauraki Gulf. *New Zealand Journal of Marine and Freshwater Research* **11**(1), 37-60.
8. Francis M.P. 1994 Duration of larval and spawning periods in *Pagrus auratus* (Sparidae) determined from otolith daily increments. *Environmental Biology of Fishes* **39**(2), 137-152.
9. Sponaugle S., Cowen R.K., Shanks A., Morgan S.G., Leis J.M., Pineda J., Boehlert G.W., Kingsford M.J., Lindeman K.C., Grimes C., et al. 2002 Predicting self-recruitment in marine populations: Biophysical correlates and mechanisms. *Bulletin of Marine Science* **70**, 341-375.
10. Bernal-Ramirez J.H., Adcock G.J., Hauser L., Carvalho G.R., Smith P.J. 2003 Temporal stability of genetic population structure in the New Zealand snapper, *Pagrus auratus*, and relationship to coastal currents. *Marine Biology* **142**(3), 567-574.
11. Sumpton W.D., Ovenden J.R., Keenan C.P., Street R. 2008 Evidence for a stock discontinuity of snapper (*Pagrus auratus*) on the east coast of Australia. *Fisheries Research* **94**(1), 92-98.

12. Leleu K., Remy-Zephir B., Grace R., Costello M.J. 2012 Mapping habitats in a marine reserve showed how a 30-year trophic cascade altered ecosystem structure. *Biological Conservation* **155**, 193-201.
13. Ross P.M., Thrush S.F., Montgomery J.C., Walker J.W., Parsons D.M. 2007 Habitat complexity and predation risk determine juvenile snapper (*Pagrus auratus*) and goatfish (*Upeneichthys lineatus*) behaviour and distribution. *Marine and Freshwater Research* **58**(12), 1144-1151.
14. Parsons D.M., Sim-Smith C.J., Cryer M., Francis M.P., Hartill B., Jones E.G., Le Port A., Lowe M., McKenzie J., Morrison M., et al. 2014 Snapper (*Chrysophrys auratus*): a review of life history and key vulnerabilities in New Zealand. *New Zealand Journal of Marine and Freshwater Research* **48**(2), 256-283. (doi:10.1080/00288330.2014.892013).
15. Brown R.C., Tsalavouta M., Terzoglou V., Magoulas A., McAndrew B.J. 2005 Additional microsatellites for *Sparus aurata* and cross-species amplification within the Sparidae family. *Molecular Ecology Notes* **5**(3), 605-607.
16. Chen S.L., Liu Y.G., Xu M.Y., Li J. 2005 Isolation and characterization of polymorphic microsatellite loci from an EST-library of red sea bream (*Chrysophrys major*) and cross-species amplification. *Molecular Ecology Notes* **5**(2), 215-217.
17. Hatanaka A., Yamada S.I., Sakamoto T., Mitsuboshi T. 2006 Isolation and application of microsatellite DNA markers for pedigree tracing of seedlings of red sea bream (*Pagrus major*). *Journal of the World Aquaculture Society* **37**(1), 139-143.
18. Takagi M., Taniguchi N., Cook D., Doyle R.W. 1997 Isolation and Characterization of Microsatellite Loci from Red Sea Bream *Pagrus major* and Detection in Closely Related Species. *Fisheries Science* **63**(2), 199-204.
19. Adcock G.J., Bernal Ramírez J.H., Hauser L., Smith P., Carvalho G.R. 2000 Screening of DNA polymorphisms in samples of archived scales from New Zealand snapper. *Journal of Fish Biology* **56**(5), 1283-1287.
20. Pompanon F., Bonin A., Bellemain E., Taberlet P. 2005 Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* **6**(11), 847-846.
21. Kalinowski S.T., Taper M.L., Marshall T. 2007 Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099-1106.
22. Peakall R., P.E. S. 2012 GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**, 2537-2539.

23. Gerber S., Chabrier P., Kremer A. 2003 FaMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Molecular Ecology Notes* **3**(3), 479-481.
24. Jones O.R., Wang J. 2010 COLONY: A program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**(3), 551-555.
25. Saenz-Agudelo P., Jones G.P., Thorrold S.R., Planes S. 2011 Connectivity dominates larval replenishment in a coastal reef fish metapopulation. *Proceedings of the Royal Society B: Biological Sciences* **278**(1720), 2954-2961.
26. Gerber S., Mariette S., Streiff R., Bodenes C., Kremer A. 2000 Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology* **9** (8), 1037-1048.
27. Kalinowski S.T., Wagner A.P., Taper M.L. 2006 ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes* **6**, 576-579.
28. Harrison H.B., Williamson D.H., Evans R.D., Almany G.R., Thorrold S.R., Russ G.R., Feldheim K.A., Van Herwerden L., Planes S., Srinivasan M., et al. 2012 Larval export from marine reserves and the recruitment benefit for fish and fisheries. *Current Biology* **22**(11), 1023-1028.
29. Lunn D., Spiegelhalter D., Thomas A., Best N. 2009 The BUGS project: Evolution, critique and future directions. *Statistics in Medicine* **28**(25), 3049-3067. (doi:10.1002/sim.3680).
30. Le Port A., Croucher A.E., Montgomery J.C. 2014 Biophysical modelling of snapper *Pagrus auratus* larval dispersal from a temperate MPA. *Marine Ecology Progress Series* **515**, 203-215. (doi:10.3354/meps10973 doi doi: 10.3354/meps10973).
31. Murphy H.M., Jenkins G.P., Hamer P.A., Swearer S.E. 2011 Diel vertical migration related to foraging success in snapper *Chrysophrys auratus* larvae. *Marine Ecology Progress Series* **433**, 185-194.
32. van Etten J. 2015 gdistance: Distances and Routes on Geographical Grids. R package version 1.1-7. <http://cran.r-project.org/package=gdistance>.