

Scale-dependent variability in *Forsterygion lapillum* hatchling otolith chemistry: implications and solutions for studies of population connectivity

Philipp Neubauer^{1,*}, Jeffrey S. Shima¹, Stephen E. Swearer²

¹School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand

²Department of Zoology, University of Melbourne, Parkville, Victoria 3010, Australia

ABSTRACT: Elemental signatures of the otoliths of fishes are increasingly used in connectivity studies to infer the natal origins of settlers or recruits. We evaluated the utility of this approach by assessing variability in trace element signatures within otoliths of hatchlings of the common triplefin *Forsterygion lapillum*. We sampled hatchling otoliths from eggs collected in a spatially hierarchical design spanning a range of environmental conditions (open coasts, bays and sounds, offshore islands) around Cook Strait, New Zealand. Our results indicate that trace element signatures vary among clutches within sites, among sites within regions and between the North Island and South Island of New Zealand. Because sites within some regions are similar to sites in other regions, we tested a statistical grouping framework based on simulated annealing, which aimed to maximize the power to make robust inferences at a given spatial scale with respect to classification error rate. We further adapted and evaluated a statistical exclusion test framework as an alternative to assignment tests when not all putative source populations could be sampled. For our study system, we found that this exclusion method performed well for some individual sites with sufficiently unique signatures, but did not perform well for groups of sites at larger scales. Overall, our work has highlighted some of the challenges that may limit the utility of hatchling otoliths when used alone for inference of natal origins of fish, and we have presented a set of statistical procedures that may improve the strength of inferences for some ecological questions.

KEY WORDS: Larval dispersal · Otolith core · Otolith chemistry · Connectivity · *Forsterygion lapillum*

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Patterns of connectivity among spatially discrete local populations have important consequences for the dynamics and behaviour of local populations and the larger metapopulation (Caley et al. 1996, Cowen et al. 2000). Nonetheless, our understanding of patterns of connectivity in coastal marine systems remains limited. Connectivity in many marine metapopulations is generally facilitated by dispersal of larvae. While larval durations can last weeks to months for some species, recent evidence suggests that long distance dispersal may be a relatively rare occurrence (Cowen et al. 2006), and self-recruitment could be important in some

populations (see Swearer et al. 2002 for a review). Obtaining more complete knowledge of dispersal patterns in marine metapopulations is of critical importance if we hope to understand some of the key drivers of marine population dynamics.

In teleost fishes, otoliths (fish 'ear stones') are increasingly used to address dispersal questions (e.g. Campana & Thorrold 2001, Thorrold et al. 2002, Elsdon & Gillanders 2003, Shima & Swearer 2009). Otoliths are known to act as environmental recorders during a fish's life (Campana & Thorrold 2001); trace elements from the developmental environment are incorporated into growing otoliths and may reflect local environmental conditions (Campana 1999). Otoliths are meta-

*Email: philipp.neubauer@yahoo.com

bologically inert, such that once deposited, trace element signatures form a permanent record. Trace element signatures can facilitate reconstruction of dispersal histories (e.g. Elsdon & Gillanders 2003, Hamilton et al. 2008, Shima & Swearer 2009) and classification of individuals to putative natal source populations (e.g. Thorrold et al. 2002, Gillanders 2005, Ruttenberg et al. 2005).

Brophy et al. (2004) documented evidence for a peak of Mn in cores (the natal region) of otoliths in 2 species of fish, and Ruttenberg et al. (2005) extended these findings to 6 other species. In otoliths sampled from older individuals, the distinct Mn peak may thus serve as a marker of the core which may in turn act as a marker of natal conditions (Ruttenberg et al. 2005, 2008, Barbee & Swearer 2007). Comparing trace element signatures from the otolith cores of settlers or recruits to an 'atlas' of trace element signatures (e.g. constructed from the otoliths of hatchlings collected from known locations) could facilitate the assignment of settlers and/or recruits to natal populations of origin. The utility of this approach relies upon sufficient spatial variation in natal (core) otolith chemistry over the scale of this 'atlas' to enable settlers of unknown origin to be assigned to putative natal source populations with a high level of confidence.

Little is known about scales of variability of hatchling otolith trace element signatures in coastal environments. To date, most studies that have employed this approach in marine environments have focused on island chains (Warner et al. 2005, Ruttenberg & Warner 2006, Ruttenberg et al. 2008) and have provided evidence of significant between-site variation within islands for trace element signatures. At larger spatial scales, among-site differences in signatures may overlap with those of other regions, and the discrimination of sites on these scales may prove problematic (Ruttenberg & Warner 2006, Ruttenberg et al. 2008). In addition, natal signatures may be influenced by maternal or physiological effects (Campana 1999, Thorrold et al. 2006, S. Swearer unpubl. data), such that clutches from different mothers within sites may differ markedly in signatures. This source of variation could mask contributions from environmental signals to limit inferences relating to natal origins. Thus, before this overall approach can be successfully applied to address many questions of population connectivity, the geographical resolution and population specificity of hatchling otolith signatures need to be established (Ruttenberg et al. 2008).

Additionally, statistical approaches to explore and utilize patterns in hatchling otoliths have been limited to statistical testing and linear discriminant analysis (LDA). Though LDA is generally practical for the assignment of recruits to possible sources (White &

Ruttenberg 2007), it may be limiting when the atlas is incomplete (Munch & Clarke 2008, White et al. 2008). Alternatives have been suggested (Munch & Clarke 2008), but their utility has not been empirically evaluated (but see White et al. 2008 for a related treatment of this problem). Furthermore, the proportion of correctly assigned individuals in a discrimination problem will diminish with an increasing number of classes (Smouse et al. 1982, Gillanders 2005, Barbee & Swearer 2007). This has prompted some investigators to consider spatially aggregating sites to counteract this effect (e.g. Gillanders 2002). While spatial aggregation of sites can be straightforward in simple geographical contexts (e.g. linear coastlines), the same is not necessarily true for more complex coastal geometries. In such contexts, no methods currently exist to find optimal groupings and therefore optimize the usefulness of hatchling otoliths as proxies for natal origins of fish.

The purpose of this study was 2-fold. (1) We explored patterns of variation in hatchling otolith signatures using a hierarchical (i.e. spatially nested) sampling design around Cook Strait, New Zealand, spanning a wide range of environmental conditions in the region. This approach enabled us to evaluate sources of variation and infer plausible spatial scales of resolution that may be achievable using hatchling otolith microchemistry for our study system and species. (2) We described and evaluated 2 approaches that may be used to improve the resolution and strength of inferences derived from trace element signatures. First, we employed a simulated annealing algorithm (e.g. Dupanloup et al. 2002) that statistically optimised the spatial groupings of sites in order to increase the probability of correctly assigning individuals of unknown origin to their natal location. Second, we tested an 'exclusion test' approach (Cornuet et al. 1999, Munch & Clarke 2008, Standish et al. 2008), which has been suggested for situations where not all possible source locations can be referenced within the natal atlas. We tested the utility of this approach for our data by performing exclusion tests on otolith signatures of known origin to identify the spatial scales at which this approach can provide the most robust assignment of larval origins.

MATERIALS AND METHODS

Study species and sampling. In our study, we focussed on the common triplefin *Forsterygion lapillum*, a small reef fish that is ubiquitous on relatively sheltered rocky reefs and tide pools around New Zealand to a depth of ~10 m (Clements 2003). Adults are strongly site-attached, with a home range limited

to a few m² (Clements 2003, Wellenreuther et al. 2007). Females lay benthic eggs almost year-round within our study area (Wellenreuther & Clements 2007, P. Neubauer pers. obs.), typically on the underside of smooth cobbles, and nests are guarded by territorial males until hatching. Larvae hatch after ~10 d (P. Neubauer unpubl. data) and settle back to subtidal reefs in the region after a ~52 d pelagic larval duration (PLD) (Shima and Swearer 2009).

We employed a hierarchical sampling design that included locations on both sides of the Cook Strait (the divide between the North and South islands of New Zealand, Fig. 1). The spatial extent of our study area was determined by knowledge of mean currents in the region (Bowman et al. 1983, Heath 1986) and the probable local distribution of *Forsterygion lapillum*, derived from a boosted regression tree model of *F. lapillum* abundance in relation to important habitat variables (Smith 2008).

We collected eggs during the Austral summer from 11 to 28 December 2007 on the North Island and 1 to 10 January 2008 in the Marlborough Sounds region of the South Island. We visited a total of 30 sites and found eggs at 17 of them (Fig. 1). From each site we attempted to collect 5 hatchlings from each of 5 discrete egg clutches. Since this was not always possible (due to the scarcity of clutches or the failure of some clutches to hatch), some of the sites in this study are represented by 3 to 4 clutches but with 6 to 8 hatchlings per clutch to have comparable number of hatchlings from each site.

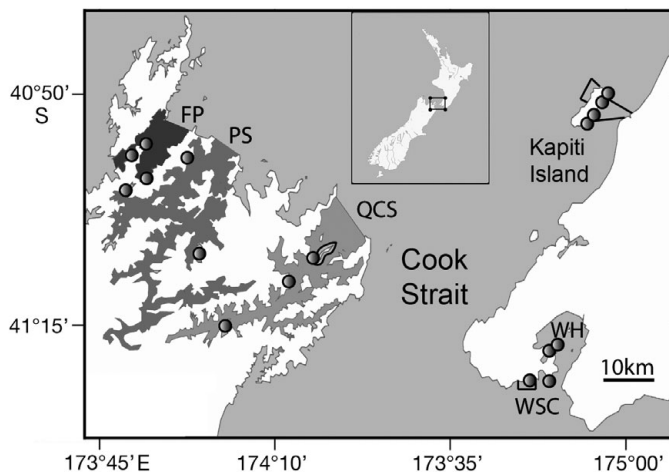


Fig. 1. New Zealand and the sampling domain. Sampling sites (gray circles) are scattered either side of Cook Strait. Black outlines mark borders of regional marine reserves. South Island regions (QCS: Queen Charlotte Sound; PS: Pelorus Sound; FP: French Pass) encompass 3 sites each. Kapiti Island was sampled at 4 sites along its eastern side, the south coast (WSC) of Wellington and Wellington's harbour (WH) regions consist of 2 sites each

Clutches were collected by divers using SCUBA or snorkelling in <3 m depth. Cobbles or rocks containing late-stage eggs (indicated by embryos with visible eye spots) were collected by hand and placed in a sterile plastic bag along with ~2 l of water from the location of collection. On shore, the contents of bags were emptied into clean PVC buckets with lids, such that individual clutches were submerged and maintained in ambient water from the natal site of collection. All but 2 clutches (both from a single site at Kapiti Island) hatched within 30 min after collection. The remaining 2 clutches were kept in the lab (in sea water from the site of collection, supplemented with bubbled air) for 36 and 48 h, respectively. These clutches were excluded from later statistical analysis because signatures were substantially different from clutches that hatched immediately at this site, indicating possible handling effects. After hatching, all larvae were transferred into 0.2 ml Eppendorf microcentrifuge tubes containing 95% analytical grade ethanol and transported to the lab for otolith extraction and analysis.

Otolith extraction, preparation and analysis. Hatchling otoliths were extracted individually from fish, in a randomised order, under a laminar flow fume hood. To further minimise the risk of contamination (which could confound trace element signatures measured within otoliths), all material used during extraction and preparation of the otoliths was acid-rinsed for 24 h in 6 N HCL, then rinsed 3 times in Millipore 18.2 MΩ water before use. Individual hatchlings (i.e. larvae) were placed in a droplet of Millipore water on a clean slide and a sagittal otolith was extracted from each larva. Otoliths were subjected to cleaning and rinsing baths as described in Barbee & Swearer (2007) before being embedded in resin on a gridded slide. A Varian Laser Ablation Inductively Coupled Plasma Mass Spectrometer (LA-ICPMS) at the University of Melbourne was used to analyse 11 elements contained within the hatchling otoliths (⁷Li, ¹¹B, ²⁴Mg, ³¹P, ³⁴S, ⁵⁵Mn, ⁶³Cu, ⁶⁶Zn, ⁸⁸Sr, ¹³⁸Ba, ²⁰⁸Pb). Otoliths were ablated vertically in a deepening pit that resulted in a series of readings obtained from the surface of the otolith through the core. For further details of the extraction, preparation and analytical methods see Barbee & Swearer (2007).

All data were postprocessed with calculations being identical to those used in Barbee & Swearer (2007). All time series resulting from this processing were inspected visually to check for signs of surface contamination. Specifically, we inspected the first and last readings corresponding to the otolith surface and looked for elevated peaks in elements such as Zn and Pb, which are prone to contamination. Ten scans to either side of the Mn peak (interpreted as the otolith primordium) were averaged to obtain the natal trace element signatures for subsequent statistical analyses.

Descriptive statistics, significance testing and discrimination. We employed a nested model to explore patterns and sources of variation in trace element signatures across all levels of our hierarchical sampling design. 'Clutches' (Level 1: random effect) were nested within 'sites' (Level 2: random effect), which were in turn nested within the different (*a priori* defined) 'regions' (Level 3: fixed effect) distributed across both main 'islands' (Level 4: fixed effect) of New Zealand, separated by the Cook Strait. Note that by making such arbitrary groupings, we made implicit assumptions about processes that govern geographical patterns in otolith chemistry; i.e. we assumed that all sites within a region are influenced by a region-specific generative process. We address this issue in more detail below. We performed nonmetric multidimensional scaling (MDS) to visually inspect similarities at different levels of our hierarchy (e.g. sites in our sampling domain and clutches that hatched at different times; see previous section 'Study species and sampling').

To test significance at different levels of the sampling design we opted for permutation-based multivariate ANOVA (Anderson 2001). All data were standardized to unit variance to equalize the relative influence of elements on the statistic. All levels were tested with 9999 permutations. Significance testing was conducted using PRIMER software with PERMANOVA add-on (Anderson 2001, Anderson et al. 2008). As this was a mixed model, components of variation were calculated as sums of squares divided by degrees of freedom for fixed effects and as variance components derived from expected mean squares for random effects (Anderson et al. 2008).

To facilitate comparisons with earlier studies, we used LDA of log-transformed data to discriminate amongst classes in levels of the sampling hierarchy. LDA has been widely used to classify and assign hatchling otoliths to potential source populations, though other classifiers are preferred in certain situations. The usefulness of our hatchling otolith 'atlas' is ultimately measured by its ability to successfully discriminate amongst putative source populations. We chose to evaluate the performance of this method based on the overall classification success of hatchling otoliths (with known geographic origins), using 'leave-one-out' cross validation as an indicator of performance of the technique at different levels of the sampling hierarchy. Since this measure represents an estimate from the sample at hand, we used 1000 parametric bootstrap samples (drawing from a multivariate normal distribution with mean and covariance matrix of log-transformed data) for each source to calculate confidence limits of cross-validation results as 2.5 and 97.5% quantiles of the bootstrap distribution. We also

calculated the relative contribution of each element to class separation by calculating their contribution to the *F*-ratio (the ratio of between class variability to within class variability).

To visually illustrate the signal-to-noise ratio of levels of the hierarchy, we plotted the mean Mahalanobis distance at each level, which is the Euclidian distance between groups (the 'signal') weighted by intragroup variance (the 'noise'), against number of groups and expected assignment success (Smouse et al. 1982, White & Ruttenberg 2007). All analyses in this and the following sections were conducted using the statistical computing language R (R Development Core Team 2007).

Optimal grouping of sites by simulated annealing. We initially imposed a hierarchical spatial structure on our sampling programme (e.g. defining discrete regions within islands), but this structure might not represent natural underlying scales of variability in trace element signatures. To identify optimal, non-*a priori* groupings of sites that might improve the allocation success of unknown signatures (and, thus, improve the power of the approach), we used a simulated annealing (SA) algorithm. SA is a general optimization heuristic that can locate global optima in combinatorial problems. Our procedure essentially follows a simple SA algorithm (e.g. Burkard & Rendl 1984) except that the proposal at each iteration conforms to constraints in a contiguity matrix (only adjacent sites can be grouped) as well as problem-related constraints if needed (see also Dupanloup et al. 2002 for a related simulated annealing algorithm in genetics). The procedure is summarized as follows:

1. Propose any starting configuration, such as our *a priori* groupings.
2. Randomly choose to either merge 2 adjacent sites or clusters of sites or remove a site from a cluster. This involves a Bernoulli trial (merge/split) with probability 0.5 followed by 2 multinomial trials $T = (i, j)$ with probability $1/N_T$ (Sites or Clusters I and J to merge or site J to remove from Cluster I, where I and J are the outcomes of trials i and j respectively), where N_T is the number of possible choices in each trial. Note that N_T for trial j is dependent on the outcome of trial i .
3. Calculate the value E of the 'cost function'. In our case this is the expected assignment accuracy for the new configuration (Smouse et al. 1982).
4. Accept the new grouping if $E_j > E_{best}$, otherwise accept with probability $\exp(\Delta E/\tau)$. If accepted, set $E_{best} = E_j$.
5. Return to step 2 for k iterations.

The variable τ is called the 'temperature' with analogy to related physical problems. It is decreased (logarithmically in our case) with the number of iterations such that, at the initiation of the procedure, suboptimal

configurations are accepted with high probability. In later iterations, suboptimal configurations become increasingly unlikely. This progressive decrease in the acceptance of suboptimal configurations enables the algorithm to overcome local optima and identify a global optimum. A more conventional clustering approach would yield a single (but potentially suboptimal) cluster configuration; our approach facilitates a more thorough exploration of all possible configurations to increase the likelihood of arriving at a globally optimal solution.

We ran the algorithm with 3 different additional constraints: (1) any spatial configuration of sites was allowed except regrouping across islands; this restriction was always applied since we felt that the reasonable discriminatory power at this scale (see 'Results') made regroupings across islands unnecessary and counterproductive. (2) the optimal solution must differentiate at least 4 SA groups (SA4) and (3) the optimal solution was biased in favour of many small groupings (SA MS). For this last constraint the cost function was modified by multiplying the expected allocation success by the number of groupings (favours many groups) and then dividing by the size of groups (favours small groups). Note that by introducing a constant in this fraction, the cost function can be arbitrarily modified to give selection criteria that favour resolution or overall allocation success. Post hoc permutational multivariate ANOVA was performed on groupings obtained from the different constraint scenarios.

Exclusion test. For some applications and research questions, it may not be possible or necessary to construct a definitive atlas of natal signatures (e.g. Elsdon et al. 2008). In such instances, it may still be useful to differentiate between individuals that probably originated from a particular location (e.g. a natal source population within a marine reserve) versus all other locations. Hence, we applied an 'exclusion test' framework to a subset of our data. Here, our focus was on subgroups within the sampling area, and we applied a statistical approach to evaluate the probability that individuals of unknown origin do not belong to a grouping of interest (Cornuet et al. 1999, Barbee & Swearer 2007, Munch & Clarke 2008, Standish et al. 2008). Using log-transformed trace element signatures from hatchling otoliths, we calculated Mahalanobis distances from the centroid of the focal group to samples known to originate from all other sites (i.e. individuals with natal origins other than the focal population; a benchmark test). The distribution of squared Mahalanobis distances from a multivariate normal distribution follows a χ^2 distribution with p degrees of freedom, where p is the dimension of the data. We can thus use this distribution as the null distribution and look up the critical value (at $\alpha = 0.05$) for belonging to a focal

group/site (represented by this distribution) in a standard χ^2 table. We then reject those samples with a squared Mahalanobis distance to the centroid of the focal group or site that is greater than the value of the chi-squared distribution corresponding to the critical value (Munch & Clarke 2008).

To illustrate the potential utility of this approach, we applied this exclusion test to a subset of our data that included individual sites and groups of sites within Pelorus Sound and Kapiti Island. Additionally, we investigated exclusion potential from 2 marine reserves in our data set: Long Island Marine Reserve in the Marlborough Sounds and Island Bay (Taputearanga Marine Reserve) on the Wellington south coast. Since these results may be an artefact of sampling (i.e. unsampled locations may have similar signatures to sampled ones), we repeated these tests for 1 case (Pelorus Sound), each time dropping 2, 5 or 10 randomly selected sites from our benchmark test. This approach enabled us to evaluate the power of the exclusion potential for this site or subset for a given sampling effort.

RESULTS

Descriptive statistics, significance testing and discrimination

Statistics relevant to chemical analyses, detection limits, as well as precision estimates, are given in Table 1. Only Li and Pb were routinely below detection limits, and neither element contributed to differences between sites (Table 2). We identified significant variation among *Forsterygion lapillum* clutches within sites ($p = 0.0001$), among sites within *a priori* defined regions ($p < 0.0001$) and between islands ($p = 0.0002$) (Table 2a). Our *a priori* regions within islands were not significantly different ($p = 0.43$), and their difference explained 0.2% of the overall variance in the data set. Small-scale differences in trace element signatures explained the largest portion of the variance, with sites accounting for 7.9%, whereas clutch differences within these locations explained 11.8%. A total of 72.8% of the total variance between signatures remained unexplained by the terms in our model (i.e. variation among eggs within clutches).

Jackknife cross-validation on LDA with equal prior probabilities for all classes indicated that 76.2% (CI: 74.4 to 78.9%) of hatchlings could be correctly assigned to their natal island despite overlap of distributions (Table 2a, Fig.2a). Also, 34.9% (CI: 33.4 to 43.8%) of hatchlings were correctly assigned to our predefined regions (Table 2a), and the separation between these regional groupings was relatively poor

Table 1. Distribution of trace element concentrations (in mol per mol Ca; note multiplier for each element column) relative to detection limits. Given are elemental concentrations for the quantiles that bracket 95% of all observations made from sampled otoliths as well as the median value. Detection limits (DL) in mol per mol Ca and external precision estimates (consistency standard, CS) for analyses of hatchling otoliths by LA-ICP-MS. DL estimates are based on 70 blank analyses. Estimates of external precision are given in relative SEs (%RSD) and are based on 31 blocks of samples (a block is a series of samples bracketed by standards), treating the National Institute of Standards and Technology (NIST) standard that is just below the standard used for calibration for each element as unknown and calculating the element:Ca ratio for this standard for each block of samples

Element	⁷ Li (× 10 ⁻⁶)	¹¹ B (× 10 ⁻³)	²⁴ Mg (× 10 ⁻³)	³¹ P (× 10 ⁻³)	³⁴ S (× 10 ⁻³)	⁵⁵ Mn (× 10 ⁻⁶)	⁶³ Cu (× 10 ⁻⁶)	⁶⁶ Zn (× 10 ⁻⁶)	⁸⁸ Sr (× 10 ⁻³)	¹³⁸ Ba (× 10 ⁻⁶)	²⁰⁸ Pb (× 10 ⁻⁶)
Sample	4.50	0.14	0.40	0.69	0.54	175.27	0.34	6.63	1.69	3.73	0.0005
quantiles	8.20	0.23	0.54	0.87	0.67	427.35	1.77	33.24	2.04	12.09	0.046
	34.62	0.45	0.83	1.29	0.78	1108.58	25.08	295.45	2.67	48.86	0.41
DL	11.44	0.030	0.029	0.095	0.14	3.82	1.57	2.36	0.00041	0.031	0.096
CS	8.54	9.14	7.83	13.27	6.16	7.32	8.37	8.26	7.48	7.77	7.21
NIST(CS)	614	614	612	612	612	612	614	614	612	614	614

Table 2. Variance explained (as sums of squares divided by appropriate degrees of freedom for fixed effects and variance components for random effects) and p-values (proportion of permuted pseudo *F*-values (Ps*F*) > model pseudo *F*) from nested permutational MANOVA for (a) *a priori* groupings, (b) sites within regions and (c) optimal groupings obtained with simulated annealing (SA4 for conditions specifying a maximum number of 4 groups, SA MS for conditions favouring small groups). Mean CV shows overall mean percentage of correctly assigned individuals using leave-one-out cross-validation. Elements are the 3 most important elements in the LDA in terms of their contributions to separation amongst groups as indicated by their relative contribution to the *F*-statistic (in brackets)

Term	Levels	Variance (%)	p (>Ps <i>F</i>)	Elements			Mean % CV
(a) A priori groupings							
Island	2	7.3	0.0002	Sr (0.35)	B (0.23)	Mg (0.13)	76.2
Region[Island]	6	0.2	0.43	Sr (0.30)	B (0.19)	S (0.13)	34.9
Site(Region[Island])	17	7.9	0.0001	Sr (0.23)	S (0.21)	B (0.20)	32.5
Clutch{Site(Region[Island])}	74	11.8	0.0001				
Residual		72.8					
(b) Sites within regions							
Kapiti	4	10.7	0.13	S (0.28)	Mg (0.16)	Sr (0.15)	56.4
Residual		89.3					
Pelorus	3	9.6	0.0001	S (0.23)	Sr (0.22)	P (0.17)	79.8
Residual		90.4					
(c) After simulated annealing (SA)							
SA4 (Island)	4		0.55	S (0.30)	Sr (0.22)	B (0.20)	51.7
SA MS (Island)	9		0.69	Sr (0.28)	B (0.20)	S (0.20)	35

(Fig. 2b,c). Sites of natal origin were correctly assigned for only 32.5% (CI: 29.7 to 33.2%) of hatchlings (Table 2a) in the total data set, but showed the highest signal-to-noise ratio (Fig. 3) of all levels in the hierarchy. Additionally, sites from different regions were often more similar than were sites within regions (Fig. 4); i.e. sites often overlapped in their distributions. When restricting the domain to just a single region, we could correctly assign 79.8% (CI: 73.1 to 86.0%) of hatchlings to their natal site of origin within Pelorus Sound (Table 2b, Fig. 2d), and 56.4% (CI: 50.1 to 63.5%) within Kapiti (Table 2b), even though trace element signatures for the sites within Kapiti were not significantly different.

Discrimination at all scales was driven by a suite of elements, the most important of which were Sr, B and S for global discrimination. Narrowing the scope to sites within regions revealed that elements such as S, P and Mg determined the discrimination at this scale. No elements generally dominated the LDA over all analyses.

Optimal grouping of sites by simulated annealing

Simulated annealing for the 2 different cost functions produced distinct grouping patterns which were found repeatedly on different runs of the algorithm. With no

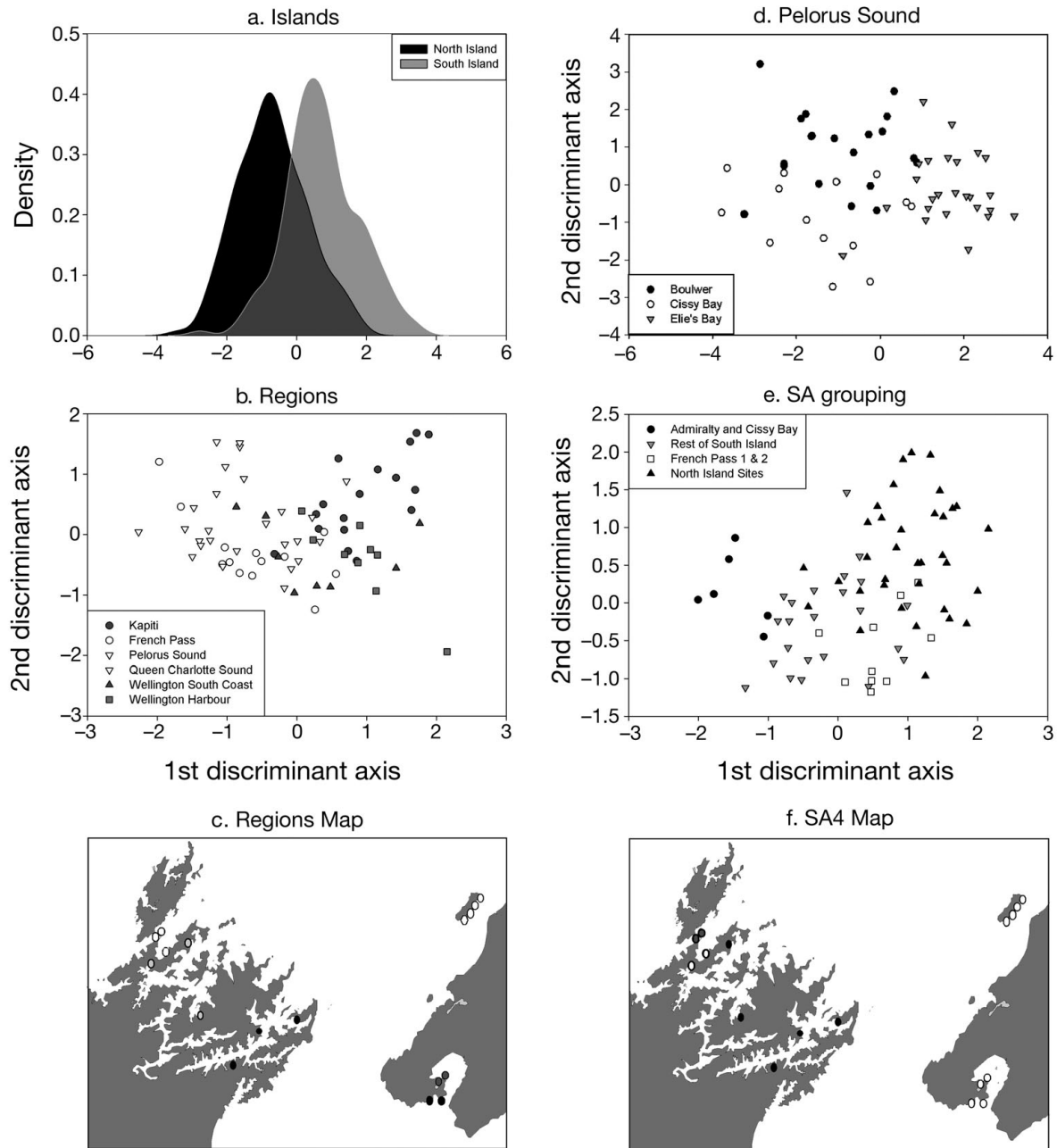


Fig. 2. Top left to bottom right panel: (a) Kernel-smoothed histograms of trace element signatures from the 2 main islands along the linear discriminant function. (b) Regional differences along the first 2 linear discriminant axes. Only clutch means are plotted for clarity. (c) *A priori* defined regions; symbols of similar shading define discrete regions. (d) Site differences in Pelorus Sound. (e) Four groups obtained with simulated annealing (SA4) to optimize allocation success in a linear discriminant analysis; clutch means are shown along the first 2 axes. (f) Groups obtained by SA; symbol shadings define groupings

restrictions, the algorithm optimising for overall allocation success consistently converged to 2 groups representing the 2 islands. When restricted to at least 4 groups, all sites within the North Island were grouped together as a single entity, while the South Island sites formed 3 distinct groupings. This spatial configuration (informed by the data rather than ad hoc by us as

observers) resulted in 51.7% (CI: 49.8 to 58.1%) of hatchlings correctly assigned to their natal regions (Table 2c & Fig. 2e,f). Varying the cost function to favour more groups produced 6 groupings and left 3 sites ungrouped. Thus, a total of 9 groups were entered in the LDA, which correctly allocated 35.0% (CI: 33.3 to 40.3%) of hatchlings to natal regions.

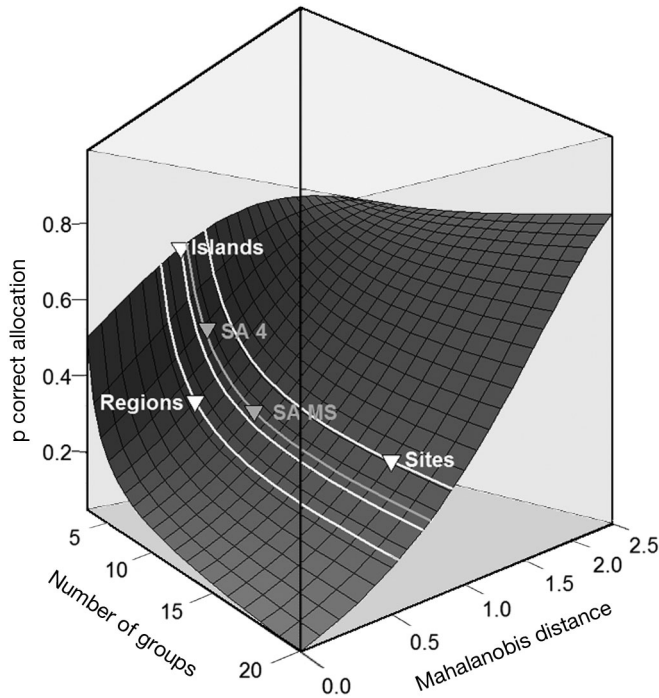


Fig. 3. Perspective plot showing predicted probability (p) of correct assignment in a linear discriminant analysis in relation to number of groups and intergroup Mahalanobis distance. With no separation between groups and equal group sample sizes, with all else being equal, assignment success in a linear discriminant analysis will be $1/I$, with I being the number of groups in the analysis. With increasing distance between group centroids, assignment success increases nonlinearly. At the 'site' level of our sampling hierarchy we obtain the highest signal-to-noise ratio (as indicated by higher average distance); yet, overall success is low due to the elevated number of groups (17 sites). Groupings from simulated annealing (SA4 and SA MS) achieve better signal-to-noise ratios than *a priori* groupings of sites into 'regions'

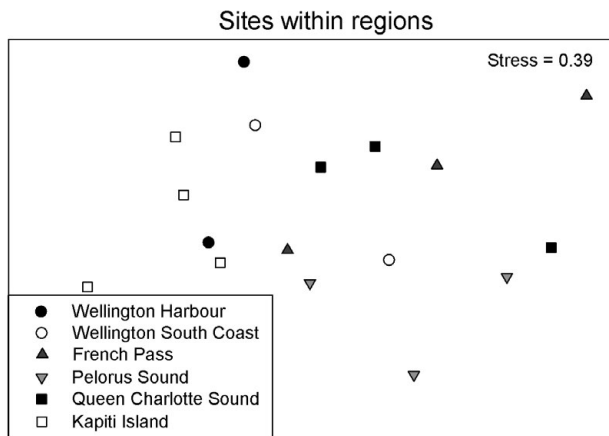


Fig. 4. Nonmetric multidimensional scaling of sites within the total sampling domain. *A priori* defined regions are represented by symbols of same colour and shape

Exclusion test for groups of sites

Exclusion tests for Pelorus Sound rejected 98, 88 and 61 % of hatchlings from all other sites for the 3 sampled sites in this embayment. Of all signatures from other sites, 76 %, on average, could be correctly excluded from Kapiti Island sites. For the sites of Long Island Marine Reserve and Island Bay (Taputearanga Marine Reserve), we could exclude 95 and 85 % of remaining signatures in the data set respectively (Fig. 5a,b). When pooling sites within regions for the test, the percentage rejected decreased dramatically, only 39 and 23 % of the remaining hatchlings could be rightfully excluded for Pelorus Sound and Kapiti Island (Fig. 5c), respectively.

Repeating these tests for Pelorus Sound sites showed that variability in estimates of our ability to correctly exclude signatures from other sites was directly related to sampling effort (Table 3): the variability in our results increased with the number of sites discarded. For instance, for Site 2 (Table 3), we estimated that we could exclude about 88 % of foreign signatures from this site based on our remaining data set of 16 remaining sites. When we estimated the same quantity from only 6 remaining sites, our estimate was as high as 99 % of foreign signatures correctly excluded (or as low as 79 % depending on which sites made up the final sample) (Table 3).

DISCUSSION

Scaling of variability in hatchling otolith signatures

A complete description of connectivity in marine metapopulations requires an understanding of all pairwise larval exchanges between local populations. Inferring such patterns from hatchling otolith micro-chemistry depends upon unique trace element signatures across all sites within a domain of interest (and across all sites in a natal atlas when recruits of

Table 3. Exclusion test results in percent of hatchlings correctly excluded from the focal location using full and truncated data sets with 2, 5 and 10 randomly chosen sites dropped from the analysis. For the truncated scenarios only minimum and maximum percentages out of 20 trials are shown

Sites omitted	0	2	5	10
Pelorus all	39	35–41	29–44	22–52
Site 1	98	98–99	98–100	96–100
Site 2	88	87–92	83–95	79–99
Site 3	61	59–65	56–67	50–79

unknown origin are to be assigned to putative natal populations).

Despite the dissimilarity in geography, we observed scales of variability in hatchling otolith signatures in

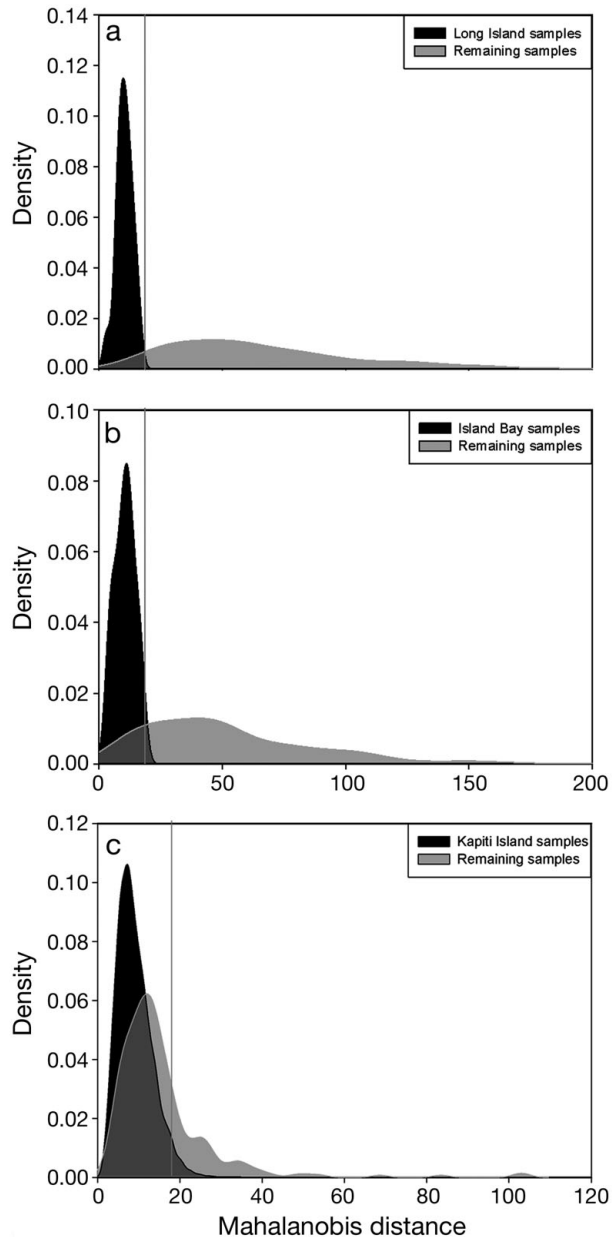


Fig. 5. Kernel-smoothed histograms (black) of Mahalanobis distances to the centroid of the focal site from chemical signatures of individual *Forsterygion lapillum* spawned at: (a) Island Bay (Taputearanga marine reserve) on the Wellington South Coast; (b) Long Island Marine Reserve in Queen Charlotte Sound; (c) Kapiti region (pooled across sites within this region). The light polygons are Mahalanobis distances of remaining signatures in the data set to the centroid of the focal site, the grey line is the 95% cut-off value for belonging to the site (i.e. the portion of the light histogram to the left of this line represents signatures incorrectly classified to the focal population, at $\alpha = 0.05$)

Forsterygion lapillum comparable with those found for other species in the Galapagos Islands and Hawaiian Islands (Ruttenberg & Warner 2006, Ruttenberg et al. 2008). Though there were clear differences in trace element signatures of hatchlings across small scales (e.g. between clutches within sites, and between sites within regions) and large scales (e.g. between islands), we failed to observe significant variability in signatures among regions within islands. We used simulated annealing to redefine regions post hoc, and while this improved the overall power of our approach (i.e. capturing additional variability at regional scales), differences among regions within islands were still not particularly strong in our system. This suggests to us that, for our system, much of the variability in trace element signatures recorded within the otoliths of hatchlings is explained by processes occurring at large spatial scales (e.g. temperature and salinity gradients between north and south islands) and at smaller spatial scales (e.g. variation in trace elements among sites and/or females). When we restrict our focus to individual regions for instance, sites are relatively distinct: within the Pelorus Sound embayment and for the offshore island of Kapiti, sites can be discriminated with some confidence, despite some sites (e.g. Kapiti) being relatively close geographically and not significantly different from each other.

Our *a priori* regional groupings reflected our beliefs about processes that may contribute to spatial variation in trace element signatures, such as 'polluted' and 'pristine' river inputs for Wellington Harbour and Pelorus Sound, respectively. However, our results suggest that this variation (and presumably also the processes that shape this variation) act on scales that differ from our *a priori* defined regions. That our data do not allow for accurate estimation of natal origins at the regional level of our natal reference atlas may seem an important limitation of the hatchling otolith approach for our system. We suggest that this limitation could be addressed by adopting a Bayesian approach (Munch & Clarke 2008) that incorporates prior knowledge of local hydrodynamics, genetic information and distributional maps of species, for example. Explicitly and probabilistically integrating such supplementary information could reduce uncertainty around natal sources by adding confidence in assignments.

Our analyses suggest that clutches collected from a common site vary substantially in trace element signatures, and this is consistent with reports from other systems (Warner et al. 2005, Ruttenberg & Warner 2006, Ruttenberg et al. 2008). These differences possibly represent differential maternal provisioning, and/or they may reflect fine scale variability in the environment (Ruttenberg et al. 2005, Chittaro et al. 2006). Some trace metals (e.g. Zn and Mn) act as cofactors in

enzymatic reactions of phytoplankton and may thus exhibit small scale variability with respect to depth or sources of input and associated planktonic communities (Morel & Price 2003).

A suite of elements was useful for discrimination at all scales. Since no one element dominated the F -statistics at any scale, or even over multiple scales, our data suggest that multi-elemental signatures may enhance discrimination. Variation in many of these elements may reflect differences in water chemistry (e.g. Sr, Ba; (Elsdon & Gillanders 2005, Walther & Thorrold 2006), whereas variation in other elements may indicate differences in physiology (perhaps mediated by spatial variation in water temperature or genetic variation) or differential maternal investment (e.g. Chittaro et al. 2006, Thorrold et al. 2006). We note that a large part of the variance in our data set remains unexplained by our model. This may be due to an inadequate model (e.g. our model structure does not reflect the structure in the data), measurement noise, nonobvious contamination and/or residual variation between hatchlings from a common clutch. Additional research is needed to determine the underlying sources of observed variability in trace element signatures across different spatial scales (see also Elsdon et al. 2008).

Furthermore, our data do not allow us to examine temporal stability in these signatures. Seasonal changes in environmental variables, such as temperature and freshwater runoff, may alter site-specific patterns through time by influencing fish physiology (e.g. higher metabolic rates at higher temperatures) and bioavailability of some elements (e.g. Elsdon & Gillanders 2006). Although not a strict limitation to the use of core signatures as markers of natal origin, temporal stability of some useful elements would greatly facilitate the practical application of this approach. Otherwise, atlases of natal signatures will have to be regenerated for each subsequent recruitment event.

Optimal regional groupings by SA

Nearby sites in our atlas have considerably different signatures (as evidenced from Fig. 4), and pooling these sites (without attention to this underlying heterogeneity) would result in a group with large variance and, hence, low overall classification success. In our case, the best groupings are not obvious from a MDS plot of site centroids (Fig. 4). We thus aimed to optimize such clusters with respect to specific criteria, which can be modified to suit a given ecological question at hand. One could, for example, leave out a location of interest, such as a reserve or site of impact, and then optimally group remaining sites to optimize resolution with respect to desired predictive success.

While SA groupings did provide improvement for regional scale allocation success and optimised signal-to-noise ratio (Fig. 3), even these optimal groupings did not achieve satisfactory allocation success at this scale. We emphasize nevertheless that this approach allows one to explore the limits of resolution at a given scale and within a given data set or a given ecological question. For the modified cost function, for instance, the algorithm favoured small groups and, thus, a high resolution, which is generally synonymous with a loss in predictive power in the LDA. Yet, the overall predictive success for the 9 selected groupings was similar to that of the 6 *a priori* defined regions. Thus, the groupings seem to be more homogenous: we do not lose predictive power by increasing the resolution with respect to *a priori* defined regions.

Exclusion tests

For some applications, such as dispersal from and/or to protected populations whose viability might depend on external larval supply (Stobutzki 2000, Halpern et al. 2006), it may be of interest to estimate levels of self-recruitment versus external larval supply instead of explicitly identifying precise source locations. This may be achieved by using an exclusion test such as the one we employed here. Taking a subset of locations from the total data set, we are able to show that exclusion of nonlocal signatures is possible in certain settings where locations are sufficiently unique in their signatures. For Long Island, Island Bay and Kapiti Island sites, 3 marine reserves within the sampling domain, we can exclude 95, 85 and 76% of remaining signatures, respectively, and thus provide a way to estimate larval supply to these marine reserves. If all other possible natal sites are sampled within the range of dispersal of the species, we may, by reciprocity, also estimate larval supply from a marine reserve to surrounding localities by applying this test.

In our case, when we employed such a test on a large scale, we quickly became confronted with the problem of overlapping site distributions between regions: the proportion of individuals correctly rejected becomes very small as soon as we enlarge our focal area (location) to more than 1 site (e.g. multiple sites within a region). In such circumstances the application of an exclusion test would be misleading: failure to exclude a large proportion of 'foreign' fish from a larger subset of focal sites would result in an overestimate of self-recruitment. Standish et al. (2008) used a similar approach to ours for excluding recruits from a set of 3 potential source locations. However, no experimental tests were conducted on their data set and it is difficult to know how well their exclusion test performed.

Our repeated tests with truncated data sets further revealed how sampling effort is directly related to our ability to estimate the exclusion potential of the test. In our case, we could drop 2 sites from our data set without dramatic changes in the percentage of correctly excluded individuals. If we dropped 10 sites, however, these estimates diverged by as much as 30%, depending on which sites remained in the sample. For cases in which many sites with similar trace element signatures remained in the sample, estimates of correctly excluded individuals were poor relative to cases that included more heterogeneous sites. Therefore, the likelihood of such extreme samples depends on the number of sites sampled and the extent of spatial covariance relative to the spacing of the sampled sites. To make any claims about the ability of an exclusion procedure to correctly exclude foreign signatures, the sample thus needs to encompass as much of the variability in signatures found across the dispersal range of the focal species as possible. Carefully considered sampling designs, e.g. informed by preliminary pilot studies, can be extremely useful in this respect and may lend increased confidence in the inferences derived from studies using otolith chemistry.

Conclusion

Our aim with this paper was to investigate spatial scales of variability in hatchling otolith signatures in *Forsterygion lapillum* for our system in the light of ecological questions involving larval dispersal. We demonstrated that site differences in an atlas of hatchling otolith signatures are exploitable at certain scales, the usefulness of which depends on the ecological question at hand. On some scales (regional scales in our system) supplementary information may help to improve the strength of inferences of dispersal pathways. The exclusion test provides a useful tool to assess larval subsidies from specific locations (e.g. marine reserves) on small scales. On a large scale, hatchling otolith microchemistry may yield insights into the question of present day dispersal between the main islands of New Zealand. In the light of such potential applications, further research into determinants of variability in hatchling otolith microchemical signatures seems warranted, and could ultimately facilitate valuable insights into dispersal patterns in reef fish metapopulations.

Acknowledgements. Our research was funded through a Marsden grant (Royal Society of New Zealand) to J. Shima, S. Swearer and S. Chiswell. An Education New Zealand Post-graduate Study Abroad Award to P. Neubauer facilitated additional travel. The Victoria University Coastal Ecology

Laboratory provided logistic support for our work. We thank J. Vola, J. Ford, N. Barbee and D. McNaughten for their help in conducting this research. Helpful comments from B. Ruttenberg and 3 anonymous reviewers helped to increase the quality of the final paper.

LITERATURE CITED

- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32–46
- Anderson MJ, Gorley RN, Clarke KR (2008) *Permanova+ for primer: guide to software and statistical methods*. PRIMER-E, Plymouth
- Barbee NC, Swearer SE (2007) Characterizing natal source population signatures in the diadromous fish *Galaxias maculatus*, using embryonic otolith chemistry. *Mar Ecol Prog Ser* 343:273–282
- Bowman MJ, Kibblewhite AC, Murtagh RA, Chiswell SM, Sanderson BG (1983) Circulation and mixing in greater Cook Strait, New Zealand. *Oceanol Acta* 6:383–391
- Brophy D, Jeffries TE, Danilowicz BS (2004) Elevated manganese concentrations at the cores of clupeid otoliths: Possible environmental, physiological, or structural origins. *Mar Biol* 144:779–786
- Burkard RE, Rendl F (1984) A thermodynamically motivated simulation procedure for combinatorial optimization problems. *Eur J Oper Res* 17:169–174
- Caley MJ, Carr MH, Hixon MA, Hughes TP, Jones GP, Menge BA (1996) Recruitment and the local dynamics of open marine populations. *Annu Rev Ecol Syst* 27:477–500
- Campana SE (1999) Chemistry and composition of fish otoliths: pathways, mechanisms and applications. *Mar Ecol Prog Ser* 188:263–297
- Campana SE, Thorrold SR (2001) Otoliths, increments and elements: keys to a comprehensive understanding of fish populations? *Can J Fish Aquat Sci* 58:30–38
- Chittaro PM, Hogan JD, Gagnon J, Fryer BJ, Sale PF (2006) In situ experiment of ontogenetic variability in the otolith chemistry of *Stegastes partitus*. *Mar Biol* 149:1227–1235
- Clements KD (2003) Triplefins. In: Andrew N, Francis M (eds) *The living reef: The ecology of New Zealand's rocky reefs*. Craig Potton Publishing, Nelson, p 160–167
- Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 153:1989–2000
- Cowen RK, Lwiza KMM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or closed? *Science* 287:857–859
- Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations. *Science* 311:522–527
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Mol Ecol* 11:2571–2581
- Elsdon TS, Gillanders BM (2003) Reconstructing migratory patterns of fish based on environmental influences on otolith chemistry. *Rev Fish Biol Fish* 13:219–235
- Elsdon TS, Gillanders BM (2005) Consistency of patterns between laboratory experiments and field collected fish in otolith chemistry: an example and applications for salinity reconstructions. *Mar Freshw Res* 56:609–617
- Elsdon TS, Gillanders BM (2006) Temporal variability in strontium, calcium, barium, and manganese in estuaries: implications for reconstructing environmental histories of fish from chemicals in calcified structures. *Estuar Coast Shelf Sci* 66:147–156

- Elsdon T, Wells B, Campana S, Gillanders B and others (2008) Otolith chemistry to describe movements and life-history parameters of fishes. In: Gibson RN, Atkinson RJA, Gordon JDM (eds) *Oceanography and marine biology, an annual review*, Vol 46. CRC Press, Boca Raton, p 297–330
- Gillanders BM (2002) Connectivity between juvenile and adult fish populations: Do adults remain near their recruitment estuaries? *Mar Ecol Prog Ser* 240:215–223
- Gillanders BM (2005) Using elemental chemistry of fish otoliths to determine connectivity between estuarine and coastal habitats. *Estuar Coast Shelf Sci* 64:47–57
- Halpern BS, Regan HM, Possingham HP, McCarthy MA (2006) Accounting for uncertainty in marine reserve design. *Ecol Lett* 9:2–11
- Hamilton SL, Regetz J, Warner RR (2008) Postsettlement survival linked to larval life in a marine fish. *Proc Natl Acad Sci USA* 105:1561–1566
- Heath RA (1986) In which direction is the mean flow through Cook Strait, New Zealand – evidence of 1 to 4 week variability. *N Z J Mar Freshw Res* 20:119–137
- Morel FMM, Price NM (2003) The biogeochemical cycles of trace metals in the oceans. *Science* 300:944–947
- Munch SB, Clarke LM (2008) A Bayesian approach to identifying mixtures from otolith chemistry data. *Can J Fish Aquat Sci* 65:2742–2751
- Ruttenberg BI, Warner RR (2006) Spatial variation in the chemical composition of natal otoliths from a reef fish in the Galapagos Islands. *Mar Ecol Prog Ser* 328:225–236
- Ruttenberg BI, Hamilton SL, Hickford MJH, Paradis GL and others (2005) Elevated levels of trace elements in cores of otoliths and their potential for use as natural tags. *Mar Ecol Prog Ser* 297:273–281
- Ruttenberg BI, Hamilton SL, Warner RR (2008) Spatial and temporal variation in the natal otolith chemistry of a Hawaiian reef fish: prospects for measuring population connectivity. *Can J Fish Aquat Sci* 65:1181–1192
- Shima JS, Swearer SE (2009) Larval quality is shaped by matrix effects: implications for connectivity in a marine metapopulation. *Ecology* 90:1255–1267
- Smith ANH (2008) Predicting the distribution and relative abundance of fishes on shallow subtidal reefs around New Zealand. National Institute of Water and Atmospheric Research (NIWA), Wellington
- Smouse PE, Spielman RS, Park MH (1982) Multiple-locus allocation of individuals to groups as a function of the genetic variation within and differences among human populations. *Am Nat* 119:445–463
- Standish JD, Sheehy M, Warner RR (2008) Use of otolith natal elemental signatures as natural tags to evaluate connectivity among open-coast fish populations. *Mar Ecol Prog Ser* 356:259–268
- Stobutzki IC (2000) Marine reserves and the complexity of larval dispersal. *Rev Fish Biol Fish* 10:515–518
- Swearer SE, Shima JS, Helberg ME, Thorrold SR and others (2002) Evidence of self-recruitment in demersal marine populations. *Bull Mar Sci* 70:251–271
- Team RDC (2007) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Thorrold SR, Jones GP, Helberg ME, Burton RS and others (2002) Quantifying larval retention and connectivity in marine populations with artificial and natural markers. *Bull Mar Sci* 70:291–308
- Thorrold SR, Jones GP, Planes S, Hare JA (2006) Transgenerational marking of embryonic otoliths in marine fishes using barium stable isotopes. *Can J Fish Aquat Sci* 63:1193–1197
- Walther BD, Thorrold SR (2006) Water, not food, contributes the majority of strontium and barium deposited in the otoliths of a marine fish. *Mar Ecol Prog Ser* 311:125–130
- Warner RR, Swearer SE, Caselle JE, Sheehy M, Paradis G (2005) Natal trace-elemental signatures in the otoliths of an open-coast fish. *Limnol Oceanogr* 50:1529–1542
- Wellenreuther M, Clements KD (2007) Reproductive isolation in temperate reef fishes. *Mar Biol* 152:619–630
- Wellenreuther M, Barrett PT, Clements KD (2007) Ecological diversification in habitat use by subtidal triplefin fishes (Tripterygiidae). *Mar Ecol Prog Ser* 330:235–246
- White JW, Ruttenberg BI (2007) Discriminant function analysis in marine ecology: some oversights and their solutions. *Mar Ecol Prog Ser* 329:301–305
- White JW, Standish JD, Thorrold SR, Warner RR (2008) Markov chain Monte Carlo methods for assigning larvae to natal sites using natural geochemical tags. *Ecol Appl* 18:1901–1913

Editorial responsibility: Matthias Seaman, Oldendorf/Luhe, Germany

*Submitted: January 8, 2010; Accepted: July 20, 2010
Proofs received from author(s): September 9, 2010*