

## Aberystwyth University

### *Integrating genetic and otolith microchemistry data to understand population structure in the Patagonian Hoki (*Macrurus magellanicus*).*

McKeown, Niall; Arkhipkin, Alexander I.; Shaw, Paul

*Published in:*  
Fisheries Research

*DOI:*  
[10.1016/j.fishres.2014.10.004](https://doi.org/10.1016/j.fishres.2014.10.004)

*Publication date:*  
2015

*Citation for published version (APA):*

McKeown, N., Arkhipkin, A. I., & Shaw, P. (2015). Integrating genetic and otolith microchemistry data to understand population structure in the Patagonian Hoki (*Macrurus magellanicus*). *Fisheries Research*, 164, 1-7. <https://doi.org/10.1016/j.fishres.2014.10.004>

#### **General rights**

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400  
email: [is@aber.ac.uk](mailto:is@aber.ac.uk)

1 **Title:** Integrating genetic and otolith microchemistry data to understand population structure  
2 in the Patagonian Hoki (*Macruronus magellanicus*).

3 **Author names and affiliations:** Niall J. McKeown<sup>a</sup>, Alexander I. Arkhipkin<sup>b</sup>, Paul W. Shaw<sup>a</sup>

4 <sup>a</sup>Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University,  
5 Penglais, Aberystwyth, SY23 3FG, UK

6 <sup>b</sup>Falkland Islands Fisheries Department, FIPASS, Stanley FIQQ 1ZZ, Falkland Islands

7 **Corresponding author:** Niall J. McKeown (Email: [njm2@aber.ac.uk](mailto:njm2@aber.ac.uk))

8

### 9 **Abstract**

10 Information from genetic (microsatellites and mtDNA Control Region) and previously  
11 collected otolith (trace element fingerprinting of otolith core and edge) markers was jointly  
12 interpreted to describe dispersal and gene flow in the Patagonian hoki (*Macruronus*  
13 *magellanicus*), an intensively harvested marine fish with seasonal migrations between  
14 spawning and feeding grounds. Spawning adults from a Chilean (Pacific) spawning site and  
15 three feeding ground samples (one from Chile and two temporal samples from the Falkland  
16 Islands (Atlantic)) were analysed. The data indicated a high level of Atlantic/Pacific  
17 connectivity by means of non-natal homing of individuals to spawning aggregations. Against  
18 this background of regional connectivity however, genetic data support the existence of a  
19 reproductively isolated population within the overwintering stock. Otolith core results are  
20 compatible with reproductive isolation being effected by natal homing to an Atlantic  
21 spawning site and/or local adaptation. The discordance between geopolitically defined  
22 Atlantic and Pacific management stocks and underlying biocomplexity, and implications for  
23 sustainability, are discussed. The study highlights the importance of intraspecific variation in  
24 homing behaviours in shaping population structure and the merit of employing  
25 complementary analytical approaches.

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

**Keywords:** population genetics; homing; fisheries; food security; biodiversity; conservation; management

## 1. Introduction

The Patagonian hoki, *Macruronus magellanicus* (hereafter hoki), is a migratory pelagic species inhabiting water depths of 60-600 m throughout its range from 33°S in the Southwest Atlantic, and 29°S in the Southeast Pacific, to 57°S around Cape Horn (Wöhler and Giussi, 2001). From austral spring to autumn, adult hoki are dispersed throughout their feeding grounds south of 48°S on the Patagonian shelf (Atlantic) and southern Chile (Pacific) (Wöhler and Giussi, 2001). In austral winter part of the stock migrates to more northern spawning areas, but a substantial proportion of adults remain on feeding grounds and skip spawning (Rideout et al., 2005). Large spawning aggregations have been reported around Guamin Island, Chile, between 43°S and 48°S (Galleguillos et al., 1996; Paya et al., 2002), while in the southwest Atlantic smaller aggregations of spawners and juveniles have been reported in the Gulf of San Matias (42°S) and Gulf of San Jorge (46°S) in Argentina (Wöhler and Giussi, 2001).

51

52 Exploitation of hoki intensified in the late 1980s as an alternative to the overfished common  
53 hake *Merluccius hubbsi* (Wöhler et al., 1999), with peak annual catches of 473,900 t reported  
54 in 1999 (FAO, 2008). The species is currently managed as two separate geographical stocks  
55 in Pacific and Atlantic waters. Pacific stocks have declined in recent years (Chong et al.,  
56 2007) while abundance in the Atlantic has increased (Wöhler et al., 2007), interpreted by  
57 some as supporting the Pacific / Atlantic stock distinction. However, it seems that known San  
58 Matias and San Jorge (Atlantic) spawning aggregations cannot sustain the observed biomass  
59 in the Atlantic region, implicating connectivity between the regions and/or additional high  
60 seas spawning in the southwest Atlantic (Wöhler and Giussi, 2001). There are also  
61 uncertainties regarding stock structuring on finer geographical scales within regions with  
62 spatial and temporal patterns in the distribution of juvenile and mature or post-spawned fish  
63 suggesting complex demographic stock heterogeneity (Giussi, 1996; Perier and Di Giacomo,  
64 1999). Spawning site fidelity in hoki, as suggested for the closely related New Zealand hoki  
65 (*M. novaezelandiae*, Hicks et al., 2003), could restrict gene flow. Initial population genetic  
66 studies have suggested subtle genetic differentiation in hoki between Pacific and Atlantic  
67 stocks (Machado-Schiaffino and Garcia-Vazquez, 2011) and within Atlantic waters  
68 (D'Amato, 2006), however robust inferences on stock structure are prevented for a number of  
69 reasons. Firstly, reported levels of genetic differentiation are so low that their biological  
70 significance could be questioned (Hedrick, 1999). Secondly, for species with  
71 spatially/temporally partitioned spawning and feeding periods the nature of sampling (i.e.  
72 spawning vs. non-spawning individuals) may be vital to the resolution of population structure  
73 (Hauser and Ward, 1998). None of the genetic studies of hoki to date have included samples  
74 from spawning populations and in such cases mechanical admixture, as opposed to  
75 hybridisation (Nielsen et al., 2003), may compromise estimates of population structure. The

76 relevance of this for hoki is emphasised by the findings of Schuchert et al. (2010) who  
77 reported extensive admixture of Atlantic and Pacific spawned individuals, determined by  
78 trace element analyses of otoliths, in both areas.

79

80 Dispersal and gene flow, due to their respective influences on population structuring, are key  
81 processes affecting both short-term population dynamics and long-term evolutionary change.  
82 Dispersal mediates the abundance and exchange of individuals among subpopulations and the  
83 extent to which local populations may fluctuate independently. Gene flow, through dispersal  
84 and consequent interbreeding, determines how populations are bound together as  
85 evolutionarily cohesive units. Spectrometric trace element analysis of otoliths permits  
86 elucidation of ontogenetic movements of individuals between habitats with different water  
87 chemistry (Campana, 1999) and has been used to study stock structure in Southwest Atlantic  
88 fish such as southern blue whiting *Micromesistius australis* (Arkhipkin et al., 2009). Genetic  
89 markers, which may also be applied to study ‘real-time’ dispersal (Castric & Bernatchez  
90 2004), are the only tools that can describe effective dispersal across generations (i.e.  
91 interbreeding). Combining genetic and otolith trace element approaches may confer  
92 synergistic insights into population structure (Svendäng et al., 2010). In this study a primary  
93 objective was to compare patterns of genetic variation among hoki samples collected from  
94 Pacific and Atlantic waters. While the spatial arrangement of samples was similar to that of  
95 Machado-Schiaffino and Garcia-Vazquez, (2011) an important distinction is that in this study  
96 both spawning and overwintering aggregations were analysed. As the genotyped individuals  
97 were collected along with those used in the otolith trace element study by Schuchert et al.  
98 (2010) (i.e. identical sampling time and location) an implicit additional objective was to  
99 combine both types of information towards a more informed description of dispersal and gene  
100 flow in the species.

101

## 102 **2. Materials and methods**

### 103 2.1 Sample collection and molecular analyses

104 Samples of adult hoki were obtained from commercial and research trawl catches. Hoki were  
105 collected from a known spawning site in Chile during the spawning period in austral winter  
106 (July 2007 - CSG1), and from two geographically distant feeding grounds in the southeast  
107 Pacific (southern Chile - CFG1) and in the southwest Atlantic (Patagonian Shelf southwest of  
108 the Falkland Islands - FFG1) during austral spring (October 2007) (Fig.1 and Table 1). As a  
109 temporal comparison, a second sample, of overwintering fish, was collected on the northern  
110 Falkland Islands feeding grounds (FFG2) in late austral summer 2008. For each fish pre-anal  
111 length, weight, sex and maturity stage was recorded, and samples taken for otolith chemistry  
112 (detailed in Schuchert et al. 2010) and genetic analysis (muscle fixed in 95% ethanol).

113

114 Total DNA was extracted using a standard CTAB-chloroform/isoamylalcohol method  
115 (Winnepenninckx et al., 1993). Nuclear genetic variation was assessed at two tetranucleotide  
116 (*Mm 5-4* and *Mm 14-IT4*) and four dinucleotide (*Mm 9-2*, *Mm 18-1*, *Mm 110-8*, *Mm 110-13*)  
117 microsatellite loci described by D'Amato et al. (1999). Hansen et al. (2001) demonstrated that  
118 misclassification of 4% of genotypes could produce an apparent  $F_{ST}$  of 0.001 to 0.003 when  
119 true  $F_{ST} = 0$ . Given that low  $F_{ST}$  might be expected between the hoki samples a number of  
120 steps were taken to maximise accuracy of genotyping: (i) PCR products of four individuals  
121 with known genotypes were run for every locus in every gel; (ii) all genotyping was  
122 performed independently by two experienced individuals with any mismatching genotypes  
123 being included in the repeat analysis (step iii); (iii) ~20% of all individuals were re-assayed  
124 (i.e. PCR, electrophoresis and genotyping) to assess rates of genotyping error.

125

126 Previous population genetic studies of hoki mtDNA variation have assayed variation in  
127 coding genes by either RFLP (ND5/6 - D'Amato and Carvalho, 2005) or direct sequencing  
128 (COI - Machado-Schiaffino and Vazquez, 2011) and have reported low variation and star-  
129 shaped genealogies with a single ancestral haplotype being found in the majority of  
130 individuals. The mtDNA Control Region does not code for a functional gene and therefore is  
131 under fewer functional and structural constraints, leading to a high average substitution rate  
132 (Saccone et al., 1987). As it is usually the fastest evolving region in the mtDNA of  
133 vertebrates, and therefore potentially more sensitive to fine scale population structuring, this  
134 region was targeted in this study. Predicting that hoki adhered to the ancestral mtDNA gene  
135 order of gadoids wherein the control region is flanked by the cytochrome b and 12S genes  
136 (Roques et al., 2006), GenBank sequences for Patagonian hoki and *M. novaezelandiae* were  
137 used to design primers rooted in the cytochrome B (HokiCR-F 5'-  
138 CAGCCTTTTCATCTGTTGTCC-3') and 12S (Hoki CR-R5'-  
139 GGCGACGGTGGTATATAAGC-3') genes to PCR amplify a fragment containing the entire  
140 Control Region. PCR reactions were performed in a total volume of 30ul, containing ~100ng  
141 of template DNA, 1 µM of each primer, 1X PCR Buffer, 2.0MM MgCl<sub>2</sub> and 0.5U Taq DNA  
142 polymerase (Bioline UK). The PCR thermoprofile was 3min at 95 °C, followed by 35 cycles  
143 of 30s at 95 °C, 30s at 55 °C and 45 s at 72 °C; followed by a final 5 min extension at 72 °C.  
144 PCR products were purified using ExoSAP IT and sequenced from both ends with the PCR  
145 primers on an ABI 3130 DNA sequencer. Sequence chromatograms were examined and  
146 edited in CHROMAS. Initially a small number of samples were sequenced, then following  
147 confirmation using BLAST that the Control Region was being sequenced internal primers  
148 (HokiCR-F2 5'-AGAGCACCAGCCTTGTAAG-3' and HokiCR-R2 5'-  
149 GGGGTTTTCTAGGTCCCATC-3') were designed to amplify and sequence (from both  
150 ends), using the same conditions as the initial primers, the Control Region in a larger number

151 of individuals. Sequence alignment was performed using the CLUSTAL W (Thompson et al.,  
152 1994) program executed in BIOEDIT (Hall, 1999) with adjustments made by eye where  
153 necessary.

154

## 155 2.2 Statistical analysis of microsatellite data

156 Numbers of alleles ( $N_A$ ), allelic richness ( $A_R$ ; El Mousadik and Petit, 1996), observed  
157 heterozygosity ( $H_O$ ), and expected heterozygosity ( $H_E$ ), were calculated using FSTAT 2.9.3.2  
158 (Goudet, 1995). Genotype frequency conformance at individual loci to Hardy-Weinberg  
159 equilibrium (HWE) expectations and genotypic linkage equilibrium between pairs of loci  
160 were tested using exact with default parameters in GENEPOP 3.3 (Raymond and Rousset,  
161 1995). Multilocus values of significance for HWE tests were obtained using Fisher's method  
162 (Sokal and Rohlf, 1995) to combine probabilities of exact tests. Locus-by-sample  
163 combinations were tested for the presence of null alleles using MICROCHECKER (van  
164 Oosterhout et al. 2004) with significant effects adjusted for using the van Oosterhout  
165 algorithm. Genetic structuring was assessed using a number of approaches. Single- and multi-  
166 locus values of the unbiased  $F_{ST}$  estimator,  $\theta$  (Weir and Cockerham, 1984), were calculated  
167 using FSTAT, with the significance of estimates tested by 10 000 permutations of genotypes  
168 among samples (Goudet et al., 1996). Genotypic differentiation was tested using the log  
169 likelihood (G) based exact test, and genic differentiation by Fishers exact test, both  
170 implemented in GENEPOP (with default settings). Genetic structure was also investigated  
171 without *a priori* sample information included using two clustering methods. Firstly, the  
172 Bayesian clustering analysis implemented in the program STRUCTURE (Pritchard et al.,  
173 2000) was used to identify the number of clusters,  $K$  (from a range of 1-4), with the highest  
174 posterior probability. Both the 'no admixture model' (as recommended for low  $F_{ST}$ ; Pritchard  
175 et al., 2000) and 'admixture model with correlated allele frequencies' were employed. Each



176 MCMC run consisted of a burn in of  $10^6$  steps followed by  $5 \times 10^6$  steps. Three replicates  
177 were conducted for each  $K$  to assess consistency. The  $K$  value best fitting the data set was  
178 estimated by the log probability of data [ $\Pr(X/K)$ ]. The second clustering method used was the  
179 discriminant analysis of principal components (DAPC) implemented in ADEGENET  
180 (Jombart et al. 2010). Whereas STRUCTURE assigns cluster memberships by minimising  
181 Hardy-Weinberg and linkage disequilibria within clusters DAPC has less assumptions and  
182 simply maximises differences between groups while minimising differences within groups.  
183 The optimal model (i.e. number of genetic clusters) was identified by the lowest associated  
184 Bayesian information criterion (BIC) after  $10^6$  iterations for models of  $K = 1$  to 5.

185

### 186 2.3 Statistical analysis of mtDNA data

187 All analysis was performed using ARLEQUIN 3.1 (Excoffier et al., 2005) unless stated  
188 otherwise. Genetic variation was described using indices of haplotype and nucleotide  
189 diversity ( $h$  and  $\pi$  respectively; Nei and Tajima, 1981; Nei, 1987) and their variances. A  
190 minimum spanning network was constructed in NETWORK ([www.fluxus-](http://www.fluxus-engineering.com/sharenet.htm)  
191 [engineering.com/sharenet.htm](http://www.fluxus-engineering.com/sharenet.htm)). Fu's  $F_s$  (Fu, 1997) and Tajima's  $D$  (Tajima, 1989) tests were  
192 used to test for deviations from mutation-drift equilibrium that could be attributed to selection  
193 and/or population size changes. Mismatch distributions (Harpending, 1994), the frequency  
194 distribution of numbers of pairwise differences between haplotypes within a sample, and  
195 simulated distributions under a model of demographic expansion were compared with the sum  
196 of squared deviations (SSD) between observed and expected distributions (significance  
197 assessed after 10 000 bootstrap replicates) used as a test statistic, and the expansion parameter  
198  $\tau$  estimated. Rough dates of population expansion were estimated with the formula  $T = \tau/2u$   
199 (Rogers and Harpending, 1992) assuming a mutation rate of 11% per million years  
200 (Bargelloni et al. 2003) and an average generation time of 3.8 years (Argentinean hoki

201 (*Macruronus magellanicus* ) Fishery assessment report 2011). The partitioning of variation  
202 was analysed using AMOVA (Excoffier et al., 1992) derived estimates of various  $\Phi$ -statistics  
203 (and their variance components), the significance of which were assessed by 10 000  
204 permutations. Differentiation between pairs of samples was further tested by exact tests of  
205 haplotype frequency homogeneity and pairwise  $\Phi_{ST}$  (significance assessed by permutation).

206

#### 207 2.4 Estimation of Type I and Type II error rates

208 For both the microsatellite and mtDNA markers the sample size-dependent probability of  
209 Type I and Type II errors was estimated using the simulation method in POWSIM (Ryman  
210 and Palm, 2006). For microsatellite markers the observed global allele frequencies were used  
211 as representative of the ancestral population in the analysis. As the detected number of  
212 mtDNA haplotypes ( $n = 75$ ) exceeded the maximum number of alleles ( $n = 50$ ) permitted in  
213 POWSIM the analysis for mtDNA was performed assuming 50 alleles at equal frequencies  
214 (0.02 - the observed average global haplotype frequency was 0.013).

215

### 216 3. Results

217 Levels of single- and multi-locus microsatellite variability were similar across samples  
218 (Supplementary Table 1). Number of alleles per locus ranged from 3 to 21 (average = 13), and  
219 of 78 alleles resolved 13 were private alleles (CSG1 = 3; CFG1 = 3; FFG1 = 4; FFG2 = 1)  
220 with an average intra-sample frequency of 0.013 (range 0.006 – 0.031). No significant  
221 deviations from random associations of genotypes between loci were detected, either across  
222 all samples (data pooled) or in any single sample, indicating that the loci are independent.  
223 Tests for conformity to Hardy Weinberg equilibrium expectations revealed a number of  
224 deviations, in all cases due to deficits of heterozygotes. With the exception of *Mm* 18 and *Mm*  
225 110-8, all loci exhibited significant global deviations from HWE. Significant multi-locus

226 heterozygote deficits were also detected for each of the samples, however the number of  
227 individual loci exhibiting such deficits was lower for the putative spawning site sample  
228 (CSG1 – only *Mm9-2*) than for the 3 feeding aggregation samples (4 loci in CFG1 – *Mm 5-4*,  
229 *Mm 9-2*, *Mm 14-1T4*, *Mm 110-13*; 3 loci in FFG1 – *Mm 5-4*, *Mm 9-2*, *Mm 14-1T4*, *Mm 110-8*;  
230 3 loci in FFG2 – *Mm 5-4*, *Mm 9-2*, *Mm 110-13*). Locus *Mm9-2* exhibited a significant  
231 heterozygote deficit in all samples. Microchecker identified underlying null alleles for all  
232 cases of locus/sample heterozygote deficits.

233

234 Both the STRUCTURE and DAPC clustering analyses reported no evidence for more than  
235 one genetic cluster within the data. However, all global tests of population structure among  
236 samples yielded significant outcomes :  $F_{ST} = 0.005$  ( $P = 0.001$ ), and exact tests for genic ( $P <$   
237  $0.0001$ ) and genotypic ( $P = 0.0004$ ) differentiation. Analysis of pairwise tests between  
238 samples (Table 2) identified the main contribution to the global structuring to be the  
239 differentiation of the FFG2 sample, which was significantly differentiated from all other  
240 samples according to  $F_{ST}$  and tests for genic and genotypic differences. Similar results were  
241 obtained after correcting for null alleles: significant global differentiation ( $F_{ST} = 0.005$ ,  $P =$   
242  $0.001$ ) and pairwise differentiation of FFG2 from all samples with nonsignificant results for  
243 all other pairwise comparisons (Table 2).

244 Pruning of mtDNA sequences permitted comparison of 1125 sites across 101 individuals. The  
245 sequenced region was AT rich (A= 33.82%, T = 33.11%) and contained 81 polymorphic sites  
246 (51 transitions, 21 transversions, 11 indels) defining 75 haplotypes (GenBank accession  
247 numbers x to x). Sixty-five haplotypes were found within only single samples (private  
248 haplotypes), with 62 (unique haplotypes) being represented by single individuals (Table 3,  
249 Fig. 3). Overall haplotype diversity was 0.9846 (SD = 0.0065) and nucleotide diversity was  
250 0.0040 (SD = 0.0022), with levels of variability similar among the four samples (Table 3).

251 Adjacent haplotypes in the network were separated by an average of 1.86 mutations  
252 (maximum 6 mutations) and there was no obvious phylogeographic structure in their  
253 distribution among samples (Fig. 2). Tajima's  $D$  and Fu's  $F_s$  statistics were significantly  
254 negative for each sample (Table 3) and for global analyses (global  $D = -2.327$ ,  $P = 0.002$ ;  
255 global  $F_s = -25.683$ ,  $P < 0.001$ ). Mismatch distributions were compatible with a model of  
256 rapid population expansion with similar values of  $\tau$  for each sample (Table 3). The global  $\tau$   
257 was 4.409 resulting in an estimated expansion occurring 17,823.86 years ago. AMOVA  
258 reported nearly all the variation (99.3%) to be contained within samples, with a non-  
259 significant amount partitioned among samples ( $\Phi_{ST} = 0.006$ ,  $P = 0.06$ ). Pairwise  $\Phi_{ST}$  and exact  
260 tests of haplotype frequency homogeneity were non-significant in all cases (Table 2).

261

262 POWSIM analysis indicated that the microsatellite data (average sample size = 74) had a low  
263 Type I error (Fisher  $P = 0.03$ ) and a high probability (Fisher  $P = 0.998$ ) for detecting  
264 differentiation at  $F_{ST} = 0.010$ . The employed sample sizes for mtDNA (average  $N_{mtDNA} =$   
265 26) conferred a low Type I error probability (Fisher  $P = 0.03$ ) but also a low power (Fisher  $P$   
266 = 0.62) to detect differentiation at  $F_{ST} = 0.010$ .

267

#### 268 **4. Discussion**

269 The genetic data reported here adds to the number of studies indicating that Patagonian Hoki  
270 around southern South America do not belong to a single panmictic unit (D'Amato, 2006,  
271 Machado-Schiaffino and Garcia-Vazquez, 2011). However, combining genetic data with  
272 associated information on individual fish natal area and adult movements derived from otolith  
273 trace element analyses (Schuchert et al., 2010) provides new insights into the biological  
274 significance of, and underlying mechanisms driving, this population structure.

275

276 MtDNA control region polymorphism was among the highest reported for a marine species  
277 (McMillen-Jackson and Bert, 2004) highlighting the potential utility of the control region as  
278 an informative marker in future hoki studies. However, simulation analysis indicated that the  
279 large number of low frequency haplotypes conferred a high Type II error probability for  
280 pairwise tests with the sample sizes employed here. A salient feature of the nuclear data was  
281 the significant differentiation of the austral summer Falkland Islands feeding sample (FFG2) ,  
282 hereafter referred to as the overwintering sample, from all other samples. Statistical  
283 differentiation of FFG2 was supported by all pairwise tests employed with power analysis  
284 indicating a low probability of Type I error. This differentiation was not revealed by the  
285 clustering analysis, although such analyses have been shown to lack resolution at low levels  
286 of interpopulation divergence (Latch et al., 2006).

287

288 While the relationship between statistical and biological significance is complicated (e.g.  
289 Jorde and Ryman, 1996) a number of features support the biological significance of the  
290 differentiation for FFG2. Firstly, sampling of adults rather than younger individuals reduces  
291 the probability that the differentiation is linked to non-random sampling within sites due to  
292 family aggregations (Hansen et al., 1997). Secondly, although FFG2 was sampled at a later  
293 time than the other samples the intervening period would be insufficient to introduce  
294 intergenerational noise. Furthermore, identical results were obtained when pairwise tests  
295 among feeding ground samples were performed including only 2 year old fish (the most  
296 abundant cohort, Supplementary figure 1). Therefore, the differentiation of FFG2 from the  
297 other feeding ground samples can not be attributed to temporal genetic changes within a  
298 single population. Thirdly, mtDNA variation revealed evidence of demographic fluctuations  
299 concordant with those suggested by D'Amato and Carvalho (2005). The mismatch  
300 distributions, high haplotype diversity and shallow phylogenetic structure support a post-last

301 glacial maximum (LGM = 20KYA) population expansion, with neutrality tests indicating non  
302 equilibrium signatures in the genetic diversity. This demography has implications for the  
303 detection of population differentiation using genetic markers. In the case of hoki, where there  
304 is no evidence of historical divergence, loci that are not at migration-drift equilibrium may  
305 retain signatures of historical gene flow and underestimate contemporary population isolation.  
306 Fourthly,  $F_{ST}$  reflects the proportion and not absolute number of migrants. Therefore, when  
307 populations are large, even very low  $F_{ST}$  values may reflect contemporary migration rates that  
308 are so low that populations may be reciprocally autorecruiting on time scales of relevance to  
309 fishery management (Palumbi, 2003; Hauser & Carvalho 2008). Finally, the sample was  
310 composed of overwintering fish and revealed a distinctive pattern of otolith core and edge  
311 trace element concentrations (Schuchert et al. 2010) indicating that the genetic differences for  
312 this sample are associated with life history differences.

313

314 The data therefore reveal a population within the overwintering stock that exhibits a degree of  
315 reproductive isolation. The companion otolith core results indicate that this population is  
316 largely composed of Atlantic spawned individuals. These features are compatible with the  
317 findings of D'Amato (2006) who reported evidence of four genetic groups occurring in  
318 Atlantic waters with the most divergent samples postulated to belong to an overwintering  
319 stock. Other features of the data indicate that the differentiation of the overwintering  
320 population is seemingly maintained against a background of high Atlantic/Pacific  
321 connectivity. Otolith core fingerprints revealed most CSG individuals to be Atlantic spawned  
322 (63.3%; Schubert et al., 2010) demonstrating a high level of dispersal from Atlantic to Pacific  
323 spawning sites. Concordant with this was the lack of genetic differentiation between the FFG1  
324 sample and both Pacific samples. Although both the genetic and ontogenetic patterns could be  
325 generated by mechanical mixing without interbreeding this must be considered unlikely: the

326 CSG sample was collected at spawning time, suggesting that the presence of Atlantic  
327 spawned individuals does reflect reproductive dispersal. Furthermore, there was no evidence  
328 of cryptic admixture of genetically distinct units revealed by the clustering analysis or tests of  
329 Hardy-Weinberg equilibrium. The high degree of adult mediated dispersal and presumed gene  
330 flow from Atlantic to Pacific spawning grounds reported here is concordant with patterns of  
331 parasite diversity between both regions (Mackenzie et al. 2013) but contrasts with the  
332 Atlantic/Pacific differentiation suggested by Machado-Schiaffino & Garcia-Vasquez (2011).  
333 However, based on the results of this study it seems likely that within Atlantic structuring  
334 may have confounded estimates of interregional divergence by Machado-Schiaffino &  
335 Garcia-Vasquez (2011).

336

337 A central discussion in marine population structure is the relative roles of physical structuring  
338 and behaviour (Heath et al., 2008). The high levels of mixing of Atlantic and Pacific spawned  
339 adults at feeding grounds, however, emphasises the potential importance of homing  
340 behaviours in shaping population connectivity. Here, the distinction between homing and  
341 natal homing is important. Homing, where adults return to spawning grounds irrespective of  
342 whether they were hatched there has been widely reported in a number of species (Lundy et  
343 al., 2000). Natal homing, where fish return to spawn at their natal site, though more difficult  
344 to demonstrate, has also been reported (Svendäng et al., 2010). The distinction is vital as  
345 homing may not result in genetic differentiation, and may actually effect gene flow where  
346 there is non-natal recruitment of individuals to spawning aggregations. The identification of  
347 large numbers of Atlantic-spawned individuals spawning at the Pacific spawning site reveals  
348 a high level of non-natal homing recruitment of individuals, presumably through social  
349 learning of spawning behaviour within feeding assemblages (McQuinn 1997). In contrast, the  
350 genetic differentiation reported for the overwintering sample indicates restricted allo-

351 recruitment. The high proportion of Atlantic spawned individuals within the overwintering  
352 sample would be compatible with natal homing as a mechanism maintaining reproductive  
353 isolation through spatial/temporal isolation of spawning. Selection against member-vagrant  
354 hybrids (Sinclair 1988) could also act as a postzygotic reproductive isolating mechanism.

355

356 In conclusion, otolith chemistry and genetic marker analyses provided complementary  
357 insights into population structure in Patagonian hoki, and in agreement with both Machado-  
358 Schiaffino & Garcia-Vazquez (2011) and Mackenzie et al. (2013) confirmed that current  
359 management policy based on separate national regulations (Chile/Argentina/Falkland Islands)  
360 is discordant with underlying species biocomplexity. The data indicate a high level of  
361 connective Atlantic / Pacific gene flow within a system of non-natal spawning site homing.  
362 Although such a system is expected to buffer populations against stochastic demographic  
363 change (McQuinn 1997) an important consideration in light of reported declines in the Pacific  
364 hoki population is the possibility that the predominance of Atlantic individuals at the CSG  
365 sample may reflect a reduction in Pacific self recruitment. The reproductively isolated and  
366 potentially locally adapted population within the Atlantic overwintering stock may be  
367 particularly susceptible to population declines, due to cryptic overfishing within the mixed  
368 stock fishery. Improved understanding of species ecology, and additional genetic and  
369 ontogenetic marker analysis of short interval time-series samples of spawning and feeding  
370 grounds will be needed to confidently match hoki recruitment dynamics to an appropriate  
371 management strategy. Advances in molecular techniques allowing genome wide analysis  
372 (Moen et al., 2008) and genotyping of markers under directional selection may prove to be  
373 particularly insightful.

374

375 **Acknowledgements**



376 This research was funded by the Falkland Islands Fisheries Department.

377

378 **References**

379 Arkhipkin AI, Schuchert PC, Danyushevsky L (2009). Otolith chemistry reveals fine  
380 population structure and close affinity to the Pacific and Atlantic oceanic spawning  
381 grounds in the migratory southern blue whiting (*Micromesistius australis australis*). *Fish Res*  
382 96: 188–194.

383

384 Bargelloni L, Alarcon JA, Alvarez MC, *et al.* (2003) Discord in the family Sparidae  
385 (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *J*  
386 *Evol Bio* 16: 1149-1158

387

388 Campana SE (1999) Chemistry and composition of fish otoliths: pathways, mechanisms, and  
389 applications. *Mar Ecol-Prog Ser* 188: 263–297.

390

391 Castric V, Bernatchez L (2004) Individual assignment tests reveals differential restriction to  
392 dispersal between two salmonids despite no increase of genetic differences with distance. *Mol*  
393 *Ecol* 13: 1299-1312.

394

395 Chong JV, Aguyo M, Paya I (2007) Estimación de edad, crecimiento, y mortalidad natural de  
396 Merluza de Cola, *Macruronus magellanicus* Lönnberg, 1907 (Macruronidae, Gadiformes) en  
397 el Océano Pacífico Suroriental. *Revista de Biología Marina y Oceanografía* 42: 311–333.

398

399 D'Amato ME, Lunt DH, Carvalho GR (1999) Microsatellite markers for the hake *Macruronus*  
400 *magellanicus* amplify other gadoids. *Mol Ecol* 8: 1086–1088.

401

402 D'Amato ME, Carvalho GR (2005) Population genetic structure and history of the  
403 long-tailed hake, *Macrurus magellanicus*, in the SW Atlantic as revealed by  
404 mtDNA RFLP analysis. *ICES J Mar Sci* 62: 247–255.

405

406 D'Amato ME (2006) Demographic expansion and subtle differentiation in the longtailed hake  
407 *Macrurus magellanicus*: evidence from microsatellite data. *Mar Biotechnol* 8: 189–201.

408

409 Donaldson KA, Wilson RR Jr. (1999) Amphi-panamic germinates of snook (Percoidae:  
410 Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA  
411 control region of fishes. *Mol Phylogenet Evol* 13: 208-213.

412

413 El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness  
414 among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco.  
415 *Theor Appl Genet* 92: 832–839.

416

417 Excoffier, L., Smouse, P.E., and Quattro, J.M. 1992. Analysis of molecular variance inferred  
418 from metric distances among DNA haplotypes: application to human mitochondrial DNA  
419 restriction data. *Genetics* 131: 479-491.

420

421 Excoffier, L, Laval G, Schneider S (2005) Arlequin ver,3.0: An integrated software package  
422 for population genetic data analysis. *Evol Bioinform Online* 1: 47-50.

423

424 Fisheries Department F.A.O. (2008) .FAO Yearbook. Fishery and Aquaculture Statistics  
425 2006.

426

427 Fu YX (1997) Statistical tests of neutrality of mutations against population growth,  
428 hitchhiking and background selection. *Genetics* 147: 915-925.

429

430 Galleguillos G, Montoya R, Troncoso L, Oliva M, Oyarzún C (1996) Identificación de  
431 Unidades de Stock en el Recurso Merluza de Cola en el Área de Distribución  
432 de la Pesquería. Informe Final (FIP 96-30), IFOP, 81 pp.

433

434 Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered*  
435 86: 485-486.

436

437 Goudet J, Raymond M, de Meeus T, Rousset F (1996) Testing differentiation in diploid  
438 populations. *Genetics* 144: 1933-1940.

439

440 Giussi AR (1996) Estudio de algunos aspectos del ciclo vital de la merluza de cola  
441 *Macrurus magellanicus*, Lönnberg, 1907. Tesis Doctoral, Universidad Nacional de Mar del  
442 Plata.

443

444 Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis  
445 program for windows95/98/NT. *Nucl Acid Series* 41: 95-98.

446

447 Hansen MM, Nielsen EE, Møller KLD (1997) The problem of sampling families rather  
448 than populations: relatedness among individuals in samples of juvenile brown trout *Salmo*  
449 *trutta* L. *Molecular Ecology* 6, 469-474.

450

451 Hansen MM, McPherson A, Smedbol K, Kenchington E (2001) review and report on the  
452 importance of different kinds of genetic population structure in relation to human impact.  
453 Report of the working group on the Application of Genetics in Fisheries and Management.  
454 ICES, CM 2001/F:03  
455

456 Harpending HC (1994) Signature of ancient population growth in a low-resolution  
457 mitochondrial DNA mismatch distribution. *Hum Biol*, 66, 591-600.  
458

459 Hauser L, Ward RD (1998) Population identification in pelagic fish: the limits of molecular  
460 markers. In: *Advances in Molecular Ecology* 98-144. IOS press, London.  
461

462 Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries: ugly hypotheses slain by  
463 beautiful facts. *Fish Fish*, 9 (4), 333-362.  
464

465 Heath MR, Kunzlik Pa, Gallego A, Holmes SJ, Wright PJ (2008) A model of meta-population  
466 dynamics for North Sea and West of Scotland cod- the dynamic consequences of natal  
467 fidelity. *Fish Res* 93: 92-116.  
468

469 Hedrick PW (1999) Perspective: Highly variable microsatellite loci and their interpretation  
470 in evolution and conservation. *Evolution*, 53, 313-318.  
471

472 Hicks AC, Smith PJ, Horn PL, Gilbert DJ (2003). Differences in otolith measurements and  
473 gill raker counts between the two major spawning stocks of hoki (*Macruronus*  
474 *novaezelandiae*) in New Zealand. *New Zealand Fisheries Assessment Report 2003/7*.  
475

476 Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal component: a  
477 new method for the analysis of genetically structured populations. *BMC Genet* 11, 94.

478

479 Jorde PE, Ryman N (1996) Demographic genetics of brown trout (*Salmo trutta*) and  
480 estimation of effective population size from temporal change of allele frequencies. *Genetics*,  
481 143, 1369-1381.

482

483 Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE Jr (2006) Relative performance of  
484 Bayesian clustering software for inferring population substructure and individual assignment  
485 at low levels of population differentiation. *Conserv Genet*, 7, 295–302.

486

487 Lundy C J, Rico C, Hewitt GM (2000) Temporal and spatial genetic variation in spawning  
488 grounds of European hake (*Merluccius merluccius*) in the Bay of Biscay. *Mol Ecol*, 9, 2067-  
489 1079.

490

491 Machado-Schiaffino G, Garcia-Vazquez E (2011) Population structure of long tailed hake  
492 *Macruronus magellanicus* in the Pacific and Atlantic Oceans: Implications for fisheries  
493 management. *Fish Res*, 111, 164-169.

494

495 MacKenzie K, Brickle P, Hemmingsen W, George-Nascimento M (2013) Parasites of hoki,  
496 *Macruronus magellanicus*, in the Southwest Atlantic and Southeast Pacific Oceans, with an  
497 assessment of their potential value as biological tags. *Fish Res*, 145, 1-5.

498

499 McMillen-Jackson A, Bert Theresa M (2004) Genetic diversity in the mtDNA control region  
500 and population structure in the pink shrimp *Farfantepenaeus duorarum*. J Crust Bio 24: 101-  
501 109.

502

503 McQuinn IH (1997) Metapopulations and the Atlantic herring. Rev Fish Biol Fisher, 7, 297-  
504 329.

505

506 Moen T, Hayes B, Nielsen F, Delghandi M, Fjalestad KT, Fevolden SE, Berg PR, Lien S  
507 (2008) identification and characterization of novel SNP markers in Atlantic cod: Evidence for  
508 directional selection. BMC Genetics, 9:18.

509 Nei, M. and Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases.  
510 Genetics 97: 145-163.

511 Nei, M. (1987) Molecular evolutionary genetics. New York: Columbia University Press.

512 Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Gronkjaer P (2003) Evidence of a  
513 hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by  
514 individual admixture analysis. Mol Ecol 12: 1497-1508.

515

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

518

519 Payá I, Rubilar P, Pool H, Céspedes R, Reyes H, Ehrhardt N, Adarme L, Hidalgo H, (2002)  
520 Evaluación de la merluza de cola y merluza tres aletas. IFOP, Informe  
521 Final Proyecto FIP 2000-15, 163 pp.

522

523 Perier MR, Di Giacomo EE (1999). La merluza de cola en la pesquería del Golfo San  
524 Matías: un recurso alternativo. In Avances en Métodos y Tecnología Aplicados a la  
525 Investigación Pesquera. Seminario Final del Proyecto INIDEP-JICA Sobre Evaluación y  
526 Monitoreo de Recursos Pesqueros (1994e1999). 209e211.  
527  
528 Pritchard JK, Stephen M, Donnelly PJ (2000) Inference of population structure using  
529 multilocus genotype data. *Genetics* 155: 945-959.  
530  
531 Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for  
532 exact tests and ecumenicism. *J Hered* 86: 248-249.  
533  
534 Rideout RM, Rose GA, Burton MPM (2005) Skipped spawning in female iteroparous fish.  
535 *Fish Fisheries* 6: 50-72.  
536  
537 Rogers AR, Harpending H. (1992). Population growth makes waves in distribution of  
538 pairwise genetic differences. *Mol Biol Evol* 9: 552–569.  
539  
540 Roques S, Fox CJ, Villasana MI, Rico C (2006) The complete mitochondrial genome of the  
541 whiting, *Merlangius merlangus* and the haddock, *Melanogrammus aeglefinus*: a detailed  
542 genomic comparison among closely related species of the gadidae family. *Gene* 383: 12-23.  
543  
544 Ruzzante DE, Walde SJ, Gosse JC, Cussac VE, Habit E, Zemplak TS, Adams EDM (2008)  
545 Climate control on ancestral population dynamics: insight from Patagonian fish  
546 phylogeography. *Mol Ecol* 17: 2234–2244.  
547

548 Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when  
549 testing for genetic differentiation. *Mol Ecol* 6: 600-602.  
550

551 Saccone C, Attimonelli M, Sbisa E (1987) Structural elements highly preserved during the  
552 evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *J Mol Evol* 26:  
553 205–211.  
554

555 Schuchert PC, Arkhipkin AI, Koenig AE (2010) Traveling around Cape Horn: otolith  
556 chemistry reveals a mixed stock of Patagonian hoki with separate Atlantic and Pacific  
557 spawning grounds. *Fish Res* 102: 80-86.  
558

559 Sinclair M (1988) *marine populations: an essay on population regulation and speciation.*  
560 University of Washington Press, Seattle.  
561

562 Sokal RR, Rohlf, FJ (1995) *Biometry: the principles and practise of statistics in biological*  
563 *research.* 3<sup>rd</sup> Edition.  
564

565 Svendäng H, André C, Jonsson P, Elfman M, Limburg KE (2010) Migratory behaviour and  
566 otolith chemistry suggest fine-scale sub-population structure within a genetically homogenous  
567 Atlantic Cod population. *Environ Biol Fish*, 89: 383-397.  
568

569 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA  
570 polymorphism. *Genetics* 123: 585-595.  
571

572 Thacker CE (2004) Population structure in two species of the reef goby *Gnatholepis*



573 (Teleostei: Perciformes) among four South Pacific island groups. Coral Reefs 23:  
574 357–366.  
575  
576 Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of  
577 progressive multiple sequence alignment through sequence weighting, position-specific gap  
578 penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.  
579  
580 Van Oosterhout C, Hutchinson WF, Derek PMW, Shipley P (2004) MICRO\_CHECKER:  
581 software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol  
582 Notes 4, 535-538.  
583  
584 Winpenninckx B, Backeljau T, DeWachter R (1993) Extraction of high-molecular-weight  
585 DNA from molluscs. Trends Genet 9: 407-407.  
586  
587 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population  
588 structure. Evolution 38: 1358-1370.  
589  
590 Wöhler O C, Hansen JE, Cordo HD, Giussi AR (1999) Evaluacion del estado de explotacion  
591 y recomendaciones para el manejo de la merluza de cola, *Macrurus magellanicus*, del  
592 Atlantico Sudoccidental. Periodo 1985e1997. Informe Interno INIDEP, 6/99. 16 pp.  
593  
594 Wöhler OC, Giussi AR (2001) L Merluza de Cola (*Macrurus magellanicus*) en el Mar  
595 Argentino. Taller Internacional de Investigacion y Desarrollo Pesquero (INIDEP), Argentina  
596

597 Wöhler OC, Hansen EJ, Giussi AR, Cordo HD (2007) Evaluación de meruza de cola  
598 (*Macrurus magellanicus*) en el Atlántico Sudoccidental. Período 1985-  
599 2001. Instituto Nacional de Investigación y Desarrollo Pesquero (INIDEP),  
600 Argentina. Technical Report No. 62/07

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

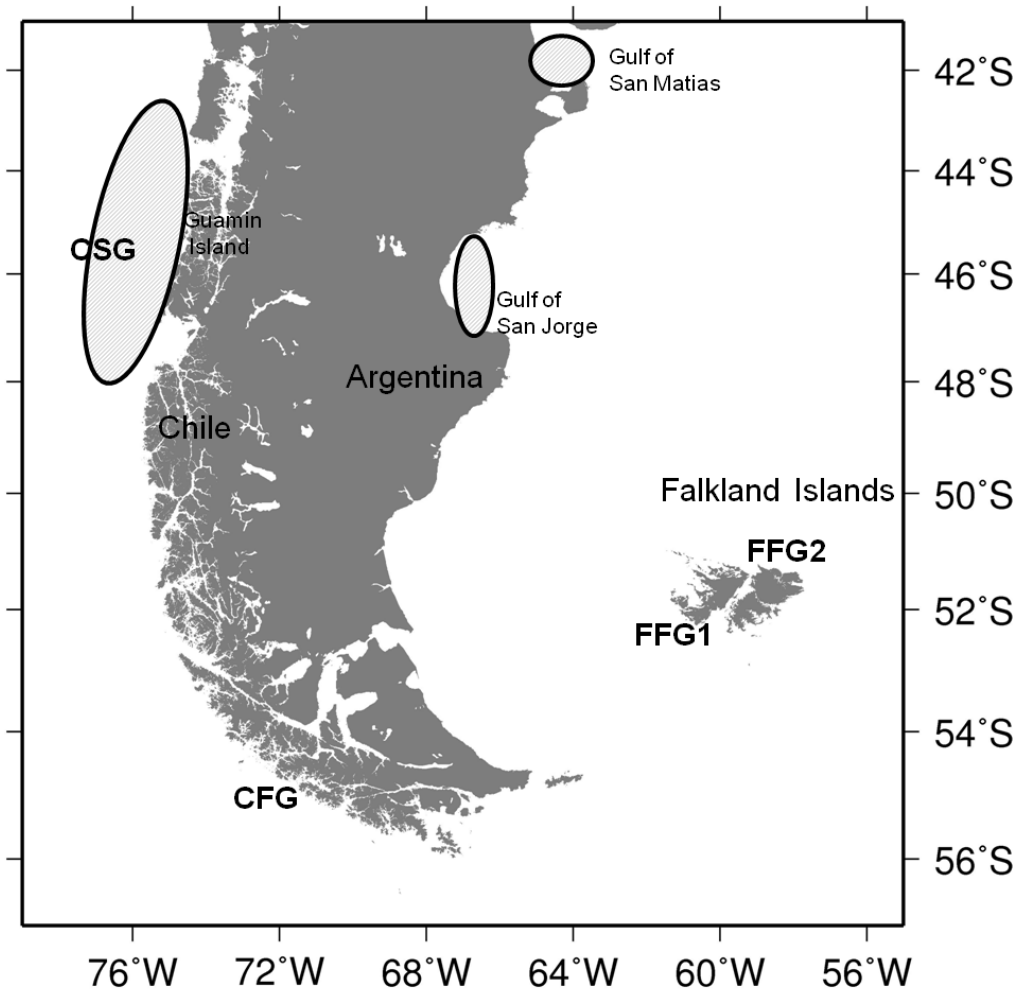
617

618

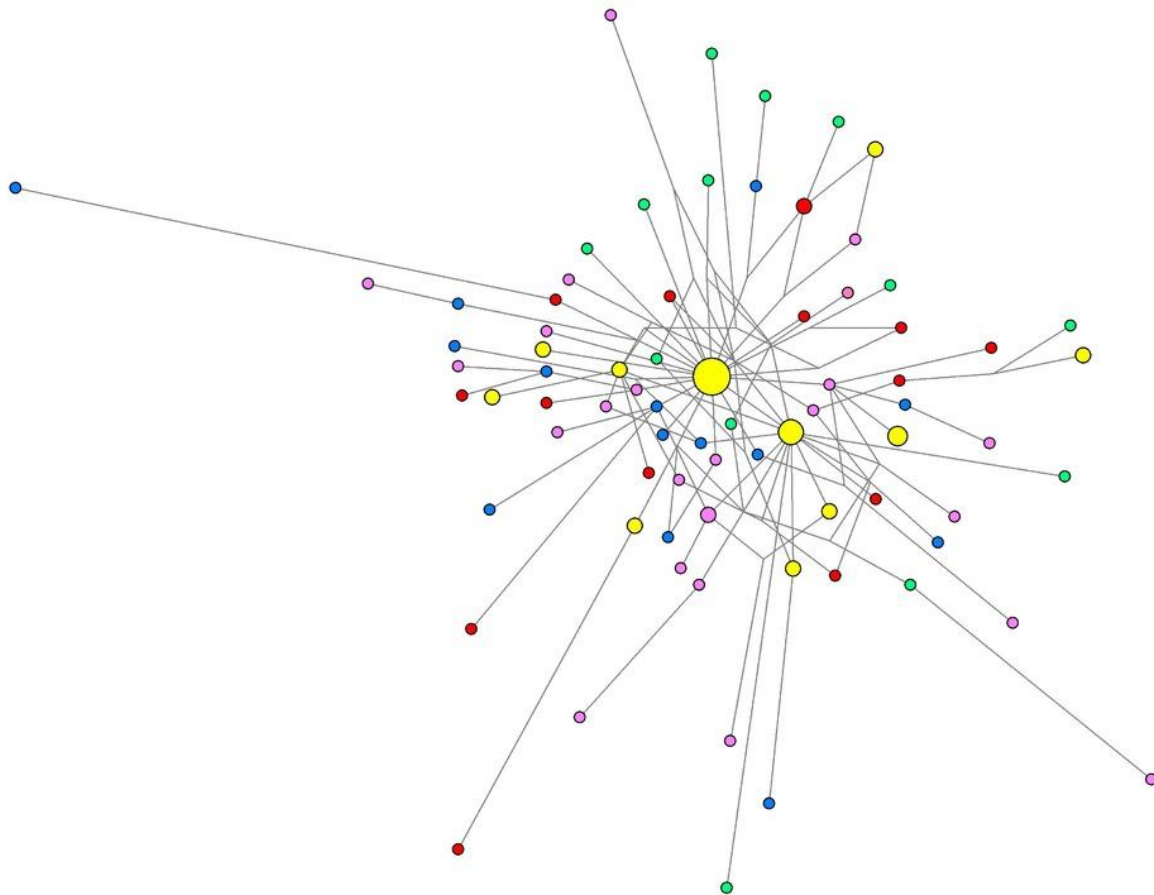
619

620

621



622  
 623 Figure 1. Sampling locations (CSG, CFG, FFG1 and FFG2) around the South Atlantic and  
 624 Pacific, confirmed spawning grounds in Chile and the Gulfs of San Matias and San Jorge in  
 625 Argentina indicated by light grey shaded areas.  
 626  
 627  
 628  
 629  
 630



631 Figure 2. Median joining haplotype network, with private haplotypes marked as pink (CSG),  
 632 blue (CFG), red (FFG1) and green (FFG2) and haplotypes detected in more than one sample  
 633 as yellow. Disc sizes are proportional to overall frequency. Median vectors are not shown.  
 634  
 635  
 636

637 Table 1. Details of *M. magellanicus* samples included in this study, including sample names  
 638 corresponding to locations indicated in Figure 1, time of sampling, sample sizes for  
 639 microsatellite ( $N$ ) and sub-sample sizes for mtDNA ( $N_{mtDNA}$ ) analyses. Also presented are  
 640 admixture proportions of Atlantic and Pacific spawned individuals derived from otolith core  
 641 analysis reported in Schuchert et al. (2010).  
 642

Sample	Date	$N$ ( $N_{mtDNA}$ )	Atlantic:Pacific admixture
CSG	Jun-07	49(31)	63.3:36.7
CFG	Oct-07	60(21)	52.8:47.2
FFG1	Oct-07	91(22)	77.6:22.4
FFG2	Feb-08	95(27)	80.9:19.1

643  
 644  
 645  
 646  
 647  
 648

649 Table 2. Pairwise tests of differentiation in allele/haplotype (Exact) and genotype (G test)  
650 frequencies, and estimates of  $F_{ST}$  for which significance ( $F_{ST} P$ ) was tested by 10 000  
651 permutations. NAC denotes test results with correction for null alleles. Significant values in  
652 bold.

	Microsatellite				MtDNA		
	Exact	G test	$F_{ST}$ (NAC)	$F_{ST} P$ (NAC)	Exact	$\Phi_{ST}$	$\Phi_{ST} P$
CSG v CFG	0.136	0.325	0.0003(0.0036)	0.243(0.150)	0.755	0.015	0.068
CSG v FFG1	<b>0.031</b>	0.132	0.0012(0.0002)	0.113(0.183)	0.429	0.017	0.056
CSG v FFG2	<b>0.002</b>	<b>0.009</b>	<b>0.0071(0.0082)</b>	<b>0.001(0.025)</b>	0.755	-0.005	0.679
CFG v FFG1	0.059	0.232	0.0029(0.0042)	0.110(0.075)	0.435	-0.003	0.584
CFG v FFG2	<b>0.009</b>	<b>0.023</b>	<b>0.0049(0.0062)</b>	<b>0.031(0.008)</b>	0.479	0.0067	0.396
FFG1 v FFG2	<b>0.0004</b>	<b>0.005</b>	<b>0.0094(0.0075)</b>	<b>0.007(0.031)</b>	0.065	-0.002	0.527

653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663

Table 3. Descriptive statistics for the samples analysed for mtDNA variation including the numbers of private singleton haplotypes and occurrence of non-private haplotypes identified. Haplotype ( $h$ ) and nucleotide ( $\pi$ ) diversities and associated standard deviations, and results of demographic tests (mismatch distribution, Fu's  $F_s$ , Tajima's  $D$ ), all obtained using ARLEQUIN 3.1 (Excoffier et al. 2005).  $P(SSD)$  denotes the probability that the empirical distribution of mismatches was significantly different than the distribution simulated under a demographic expansion model. Probabilities for Fu's  $F_s$  and Tajima's  $D$  estimated following 10 000 bootstrap replicates.

	CSG	CFG	FFG1	FFG2
Singleton private haplotypes	22	14	14	14
Hap_3	1		2	
Hap_6	2			
Hap_10	1	1		
Hap_16	1	2		2
Hap_17		3	2	6
Hap_25		1	1	
Hap_31	1			1
Hap_35	1			1
Hap_36	1			1
Hap_40	1		1	
Hap_41			1	1
Hap_43			2	
$h$ (SD)	0.9978 (0.0089)	0.9810 (0.0225)	0.9870 (0.0175)	0.9516 (0.0320)
$\pi$ (SD)	0.004475 (0.002486)	0.003545 (0.002061)	0.003924 (0.002246)	0.003748 (0.002139)
Fu $F_s$ (P)	-25.53307 ( $>0.001$ )	-13.97864 ( $>0.001$ )	-14.42352 ( $>0.001$ )	-13.11913 ( $>0.001$ )
Tajima's $D$ (P)	-2.07731 (0.007)	-2.18863 (0.002)	-1.77490 (0.023)	-2.21836 (0.001)
$P(SSD)$	0.4501	0.3435	0.3912	0.3302
$\tau$	4.984	3.582	4.445	3.76

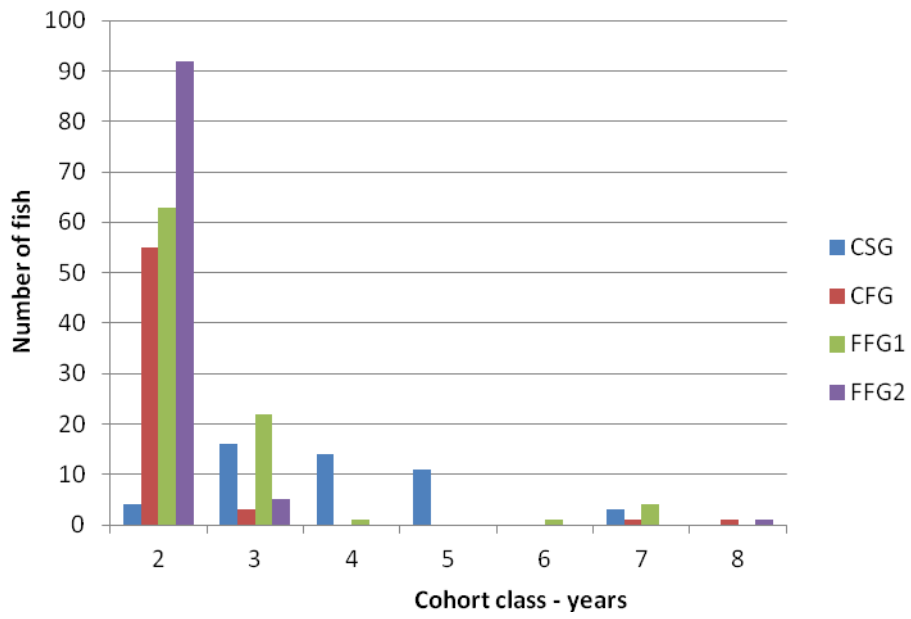
664  
665  
666  
667

668 Supplementary table 1. Summary statistics of microsatellite intrasample diversity, including  
669 allele number (Na), allelic richness (Ar), observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity and  
670 P- values for tests of Hardy-Weinberg equilibrium (PHW) for which significant deviations are  
671 indicated by values in bold.  
672

Locus	Index	CSG	CFG	FFG1	FFG2	Overall
<i>Mm</i> 5-4	Na	11	11	12	11	14
	Ar	10.96	10.66	11.01	9.68	10.784
	$H_E$	0.854	0.832	0.822	0.808	0.857
	$H_O$	0.783	0.623	0.667	0.697	0.692
	PHW	0.296	<b>0.0099</b>	<b>0.0049</b>	<b>0.0443</b>	<b>0.0004</b>
<i>Mm</i> 9-2	Na	14	11	12	9	15
	Ar	13.78	10.57	10.89	8.87	10.89
	$H_E$	0.818	0.831	0.82	0.783	0.728
	$H_O$	0.617	0.714	0.662	0.623	0.654
	PHW	<b>0.0061</b>	<b>0.0007</b>	<b>0.0084</b>	<b>0.0047</b>	<b>&lt;0.0001</b>
<i>Mm</i> 14-1T4	Na	7	9	7	6	10
	Ar	6.93	8.27	6.76	5.725	6.81
	$H_E$	0.545	0.474	0.521	0.597	0.821
	$H_O$	0.563	0.456	0.42	0.628	0.517
	PHW	0.8255	<b>0.0366</b>	<b>0.0008</b>	0.054	<b>0.0007</b>
<i>Mm</i> 18-1	Na	1	1	2	2	3
	Ar	1	1	1.75	1.87	1.69
	$H_E$	0	0	0.022	0.033	0.665
	$H_O$	0	0	0.022	0.033	0.0138
	PHW	-	-	1	1	1
<i>Mm</i> 110-8	Na	12	10	13	11	15
	Ar	11.68	9.3	10.43	9.52	10.123
	$H_E$	0.652	0.573	0.655	0.547	0.504
	$H_O$	0.646	0.667	0.637	0.603	0.638
	PHW	0.2522	0.2146	<b>0.0221</b>	0.8531	0.0879
<i>Mm</i> 110-13	Na	13	15	16	15	21
	Ar	12.42	15	11.64	11.75	12.75
	$H_E$	0.543	0.639	0.571	0.657	0.456
	$H_O$	0.49	0.444	0.518	0.561	0.503
	PHW	0.1459	<b>0.0025</b>	0.0641	<b>0.0037</b>	<b>0.0001</b>

673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687

688  
689



690  
691  
692  
693

Supplementary Figure 1. Cohort composition of the samples included in genetic analysis.

694