

Genetic structure at different spatial scales in the pearl oyster (*Pinctada margaritifera cumingii*) in French Polynesian lagoons: beware of sampling strategy and genetic patchiness

Sophie Arnaud-Haond^{1, 4, *}, V. Vonau¹, C. Rouxel¹, F. Bonhomme³, Jean Prou¹, E. Goyard¹ and P. Boudry²

¹ Laboratoire d'Aquaculture tropicale, IFREMER/COP, IFREMER, BP 7004, 98719 Taravao, Tahiti, French Polynesia, France

² Laboratoire de Génétique et Pathologie, IFREMER, BP 133, 17390 La Tremblade, France

³ Département de Biologie Intégrative, Institut des Sciences de L'Evolution de Montpellier (ISEM) UMR 5554, Station Méditerranéenne de l'Environnement Littoral, 34200 Sète, France

⁴ Present address: Laboratoire Environnement Profond, Centre de Brest, IFREMER, BP 70, 29280 Plouzané, France

*: Corresponding author : Sophie Arnaud-Haond, email address : sarnaud@ifremer.fr

Abstract:

In order to study further the genetic structure of the pearl oyster *Pinctada margaritifera* in French Polynesia with a special consideration for the sampling scale, we analyzed or re-analyzed sets of data based on nuclear DNA markers obtained at different spatial scales. At a large scale (several 1,000 km), the remote Marquesas Islands were confirmed to be significantly differentiated from Tuamotu–Gambier and Society archipelagos, with a marked difference however for the two main islands that are different from each other. At a medium scale (several 10 to several 100 km), overall homogeneity was observed within and between these two archipelagos, with some exceptions. This could be attributed both to large-scale larval dispersal and to human-driven spat translocations due to pearl oyster cultivation. These results contrast with those observed (1) at a small scale (less than 10 km) in a lagoon heavily impacted by translocation and cultural practices, where significant genetic differentiation was detected among three laying beds, and (2) at a micro scale where we detected an important variability of the genetic composition of young spat recruited on artificial collectors. Such patterns could result from a high variance in the number of genitors at the origin of each cohort, or from pre- or post-settlement selection on linked loci. Altogether, our data support the hypothesis that under certain conditions populations of bivalves may exhibit patterns of chaotic genetic patchiness at local scale, in line with the increasing report of such patchiness in marine benthic organisms. This underlines the importance of sampling scale that should be rigorously defined depending on the questions to be answered. Nevertheless, a survey of about 80 articles dealing with population genetics of marine invertebrates showed that only 35% of those studies disclosed details about the sampling strategy (particularly the area explored). These results emphasize the need for cautious interpretation of patterns of genetic structure at medium scale when rigorous sampling strategies are not deployed.

Keywords: genetic patchiness; sampling strategy; benthic organisms; pearl oyster; bivalve; Polynesia; population genetics

53 **Introduction**

54 Genetic homogeneity over large geographic scales has been long expected, and
55 sometimes observed, due to adult or larval mobility and to the absence of physical barriers
56 to dispersal (Vermeij 1987). A large number of marine invertebrates exhibit large
57 population sizes, external fertilization, high fecundity, an extensive pelagic larval phase,
58 and a benthic adult stage. These characteristics lead us to an expectation of Hardy-
59 Weinberg equilibrium as well as low genetic divergence, due to extensive gene flow
60 during the larval stage, whereas the low mobility at the adult stage may favor local
61 adaptation at the latest life stages.

62

63 Studies of the distribution of genetic variability of the black-lipped pearl oyster,
64 *Pinctada margaritifera* from the Central Pacific, performed on samples collected in the
65 1980's, suggested a natural pattern of restriction to gene flow at both large (more than
66 1000km) and medium scales (200 to 400 km). For example, populations from the Society
67 archipelagos are significantly differentiated, although separated by less than 200 km
68 (Arnaud-Haond et al. 2003a). Previous studies, based on allozymes, even suggested the
69 possible existence of genetically differentiated laying beds coexisting within the atoll of
70 Takapoto (Blanc et al. 1985; Durand and Blanc 1986). However, studies performed on
71 more recent samples showed homogenization of genetic pools of the archipelagos of
72 Society and Tuamotu-Gambier (Arnaud-Haond et al. 2004). At the archipelago scale, this
73 homogenization was attributable to the massive spat translocation in that area during
74 1990s. At the intra archipelago scale, within the Tuamotu-Gambier archipelago, where few
75 samples were analyzed before the translocation events, it is more difficult to distinguish
76 between the hypothesis of extensive natural gene flow at small spatial scale (several tens
77 of kilometers) and the hypothesis of artificial gene flow linked to farming practices

78 (Arnaud-Haond et al. 2003b). In any event, further screening of isolated islands should be
79 performed to identify possibly still divergent natural populations in atolls where no pearl
80 culture is developed or no translocation has been performed, and which may represent
81 interesting genetic resources in the perspective of future hatchery produced stock and
82 selection programs. For *P. margaritifera*, at the intra lagoon spatial scale, as for other
83 sessile species presenting pelagic larval stage (Roughgarden et al. 1988; Gaines and
84 Bertness 1992), a very stochastic dynamic of spat recruitment has been reported in both
85 space and time (Friedman et al. 1998; Friedman and Bell 1999). Whether this chaotic
86 pattern is accompanied by spatial or temporal variance in the genetic composition of spat
87 is still not known. Although no such data exist on *P. maxima* recruitment, some genetic
88 data suggest the occurrence of localized heterogeneity in the genetic constitution of
89 recruits, interpreted by the authors as resulting from large variance in the local
90 reproductive success (Benzie and Smith-Keune 2006). Recent studies on *P. margaritifera*
91 showed no significant difference in the genetic variability of wild samples and spat-
92 collected samples from farms (Arnaud-Haond et al. 2003b). To what extent this
93 observation reflects high genetic variability of locally collected spat, or high genetic
94 variability of admixtures of distinct groups of collected spats in farms remains to be
95 determined.

96
97 In benthic species for which migration is restricted to the larval stage, most species
98 with lecithotrophic larvae exhibit more restricted gene flow compared with those with
99 planctotrophic larvae, who remain longer in the plankton (Hunt 1993; Hellberg 1996;
100 Poulin and Féral 1996; Hoskin 1997; Arndt and Smith 1998; Bonhomme and Planes
101 2000). However, in several invertebrate species that show substantial larval dispersal
102 capability and large-scale genetic homogeneity, small-scale spatial and temporal genetic

103 patchiness has been reported (Johnson and Black 1984; Watts et al. 1990; David et al.
104 1997b; Johnson and Wernham 1999). The hypotheses invoked to explain such sub-
105 structure of populations at small scale include both pre- or post-settlement selection of
106 genotypes, as well as the different genetic origins of settling larvae (Johnson and Black
107 1982; Johnson and Black 1984; David et al. 1997a; David et al. 1997b) due to the
108 stochastic recruitment in the sea (Roughgarden et al. 1988; Gaines and Bertness 1992).
109 High variance in reproductive success, implying low effective number of genitors at the
110 origin of a cohort, have already been reported in bivalves (Hedgecock 1994; Li and
111 Hedgecock 1998; Boudry et al. 2002) and may favor the occurrence of genetically distinct
112 pools of recruits at small geographic or temporal scales. A complementary explanation is
113 the occurrence of differential selection in space and time favoring genetic differentiation of
114 recruited cohorts. This hypothesis requires a strong genetic linkage between the markers
115 used and some fitness component. The occurrence of post-settlement selection hypothesis
116 was supported in *Littorina saxatilis* (Johannesson et al., 1995), and pre-settlement
117 selection hypothesis has never been tested due to the technical difficulties in isolating and
118 scoring very young larvae. It is therefore hard to distinguish between the hypotheses of
119 distinct genetic origin of recruits: the synthesis of both hypotheses can be summarized as
120 the “recruitment history” (Johnson and Black 1984; Watts et al. 1990; David et al. 1997b;
121 Johnson and Wernham 1999). The existence of mosaic patterns at small scale, independent
122 of the possible large scale homogeneity of population and gene flow among distant sites,
123 underlines the importance of sampling scale for population genetics and biogeographic
124 studies (Benzie 2000). The collection of samples representative of the populations studied
125 is a pre-requisite to the interpretation of data in terms of gene flow, and may be influenced
126 by the scale of sample collection compared to the area where the species collected is
127 locally distributed. Yet, for most species studied so far, the difficulty of access to marine

128 environments and the lack of knowledge of the existence and scale of possible genetic
129 patchiness often hamper the application of a strict sampling strategy. In this context, the
130 interpretation of population genetics data in terms of gene flow may therefore often be a
131 difficult and risky exercise (Johansson et al., 1995).

132

133 We used four anonymous nuclear markers to assess the pattern of genetic structure of the
134 Polynesian black-lipped pearl oyster populations at four spatial scales: among
135 archipelagos, among islands within archipelagos, among distinct sampling sites within an
136 island, and among artificial spat collectors within sampling sites. We addressed the
137 following questions: 1) At large and medium scale, among archipelagos and among islands
138 within archipelagos, can some genetically divergent populations still be detected in
139 isolated or non harvested atolls? 2) When sampled at micro-scale, do *P. margaritifera*
140 laying beds exhibit genetic heterogeneity? , 3) what are the possible consequences of
141 genetic patchiness for sampling strategy in population genetics and biogeographic studies?

142

143 **Material and Methods**

144

145

146

147 *2.1. Sample collection and DNA extraction*

148 For large and medium-scale analysis, wild samples from fifteen geographic lagoons
149 in French Polynesia, ranging from the western Society (Manuae) to the South-Eastern
150 Tuamotu-Gambier (Mangareva) and the Northern Marquesas (Nuku-Hiva) were sampled by
151 SCUBA-diving during 1999-2002 (Figure 1; Figure 2; Table 1). When possible, samples
152 were collected in areas as large as possible, encompassing different laying beds, in order to
153 collect a sample as representative as possible of the population studied. Details of the area
154 encompassed by sampling are given in Table1 when available. Considering the difficulties in
155 gathering samples from so many dispersed islands, we were able to get some samples thanks
156 to the kindness of inhabitants, and information as to the detailed sampling area was not
157 always possible to obtain. Eight of the fifteen samples (Manuae, Maupihaa, Arutua, Apataki,
158 Manihi, Takaroa, Mangareva and Hiva Oa) have already been analyzed for all nuclear
159 markers in recent studies (Arnaud-Haond et al. 2003a; Arnaud-Haond et al. 2004).

160 Small-scale study was performed within Takapoto lagoon (Tuamotu-Gambier): wild
161 samples from three distinct zones were sampled by diving in May 2002 (Figure 2; Table 1),
162 each in restricted areas of several m² in order to avoid mixing distinct beds. Finally, six spat
163 samples of 42 to 50 individuals each were collected in February 2002 in three distinct
164 collection stations (Figure 2; Table 1), representing different levels of spat density (high
165 density and homogeneous distribution of spat on the station and collectors, mid-density and
166 relative homogeneity, very low density and heterogeneity).

167 A piece of adductor muscle or gill (Raroïa) was removed from each specimen and
168 preserved in 80% ethanol. The procedure of DNA extraction, precipitation and storage
169 were similar to those described in Sambrook et al. (1989); we used approximately 0.5g of
170 chopped and subsequently air-dried tissue. The nucleic acid pellet obtained after
171 precipitation in 100% ethanol was washed with 70% ethanol, air-dried, resuspended in 100
172 to 200µl of deionised water and preserved at -20°C. DNA concentrations, obtained by
173 fluorimetry, were found to be about 300ng/µl.

174

175 *2.2. Polymerase Chain Reaction (PCR) and electrophoresis of anonymous nuclear loci*

176 All samples were analyzed with four markers (Arnaud-Haond et al. 2002) developed
177 using the DALP (Desmarais et al. 1998) and the EPIC (Palumbi 1995) methods. Although
178 they have never been tested formally on controlled crosses and progeny, these markers are
179 co-dominant and their previous use in natural laying beds supported the hypothesis they are
180 Mendelian.

181 PCR was performed in a 20µl reaction volume with final concentrations of 300µM
182 each dNTP, 1.8 mM MgCl₂, 0.4µM of each primer, about 30ng of template DNA, 1X Taq
183 buffer and 0.75 units of Taq polymerase. In order to resolve length polymorphism, PCR
184 products were separated through 6% denaturing polyacrylamide gels
185 (acrylamide:bisacrylamide, 29:1, 7M Urea) using 1X Tris-Borate-EDTA buffer. The gels
186 were then silver stained (Bassam et al. 1991). Ambiguities in genotype reading were checked
187 by rerunning either the same or a new PCR product.

188

189 *2.3. Genetic diversity analyses, Hardy Weinberg equilibrium, linkage disequilibrium and* 190 *genetic structure*

191

192 Genetic diversity within populations was estimated by unbiased (H_{nb}) and observed
 193 (H_{obs}) gene diversity (Nei 1987). We estimated the overall values for the inbreeding
 194 coefficient (F_{is}) as described by Weir and Cockerham (1984) and we used a permutation
 195 procedure (1000 permutations) to test whether a particular F_{is} value was significantly
 196 different from 0. The two-locus correlation coefficient R^2 (Weir 1979) was estimated with
 197 the procedure of Black and Krafur (1985), and its departure from zero was tested by a
 198 permutation approach.

199 Genetic differentiation (F_{st}) was estimated between pairs of populations with the
 200 estimator θ of Weir & Cockerham (Weir and Cockerham 1984). The significance of the θ
 201 values was tested by randomly permuting 3000 X the individuals between samples. Those
 202 calculations were performed using the GENETIX 4 package (Belkhir et al. 1996-2001).

203

204

205 2.4. Variance of allelic frequencies

206 In order to estimate the effective number of genitors at the origin of each spat
 207 sampled, the standardized variance in allele frequency change was estimated according to
 208 the theory of selectively neutral alleles in finite populations (Waples 1989):

$$\hat{F}_k = \frac{1}{k-1} \sum_{i=1}^k \frac{(x_i - y_i)^2}{(x_i + y_i)/2}$$

209

210 Where k is the number of alleles and x_i and y_i the allelic frequencies of the i^{th} of k alleles. A
 211 mean F_k across loci was weighted by the number of alleles at each locus, and N_e was
 212 estimated using the formula:

$$\hat{N}_e = \frac{t}{2[F_k - (1/2n_o) - (1/2n_t)]}$$

213

214

215 where t is the number of generations (we assumed one generation separated the spat from the
 216 putative pool of genitors), n_0 the sample size at the generation 0 and n_t the sample size at the
 217 generation t . The confidence interval at 95% is estimated as:

$$218 \left[\frac{n\hat{F}}{\chi^2_{\alpha/2}[n]}; \frac{n\hat{F}}{\chi^2_{1-\alpha/2}[n]} \right]$$

219 The estimation was made using the software *NeEstimator* (Peel et al. 2004) with all putative
 220 genetic pools of genitors: each of the three wild samples from Takapoto, the synthetic
 221 population obtained while pooling these three samples, and the synthetic population obtained
 222 while pooling the four samples from Tuamotu (Arutua, Apataki, Manihi and Takaroa).

223

224

225 *2.5 Literature screening*

226

227 A general search was performed in ISI (Web of Science) for the keywords [(“population
 228 genetics” OR “genetic structure”) and (benthic OR “marine invertebrate” OR “mollusks”
 229 OR “echinoderm” OR “bivalve” OR “gastropods”)], which returned a list of 170 published
 230 studies. Although we were aware that this does not represent an exhaustive sampling of the
 231 existing work, we considered that this ‘sampling strategy’ was able to deliver a
 232 representative sample of the existing literature. Among those, 88 were found to actually
 233 deal with population genetics or biogeography of benthic organisms and 66 could be
 234 gathered and screened for details of sampling strategy. Information retained was classified
 235 as follows: A) as for the area explored 1/no information disclosed, 2/ details given
 236 about the area explored for each site and B) as for the strategy for choosing samples 3/
 237 mention of the strategy as “random or haphazard” 4/ exhaustive sampling 5/ else when

238 samples collected were chosen according to a particular strategy such as the use of a
239 transect or a grid.

240

241

242 **Results**

243

244

245

246 *3.1 Level of genetic variability, Hardy Weinberg equilibrium, linkage disequilibrium*

247

248 The level of unbiased gene diversity H_{nb} ranged between 0.30 (Manuae) and 0.45
249 (Kauehi) in natural laying beds, where the number of alleles was of 2.5 (Hao) to 5
250 (Hereheretue), and on collectors, the heterozygosity ranged from 0.29 to 0.40 and the
251 number of alleles from 3.5 to 4.00 (Table 1).

252 Significant heterozygote deficiency values were observed in four of the six
253 collector samples, and in five of the eighteen samples from natural laying beds (Table 1).

254 Linkage disequilibrium was observed to be significant between pinucl3 and pinaldo
255 in one sample from natural population: Takapoto 3. All the remaining significant values
256 concerned samples from spat collectors, and locus pinucl2: with pinucl3 for collector 3
257 and 4, with pinaldo for collectors 2 and 3, and with pinU4 for collectors 5 and 6.

258

259 *3.2 Pairwise genetic differentiation*

260 For samples from wild laying beds, pairwise genetic differentiation was analyzed at
261 three geographic scales, among lagoons among archipelagos, among lagoons within
262 archipelagos, and within lagoon. At the among archipelagos scale, despite the inclusion of
263 samples from remote atolls from Tuamotu-Gambier that supposedly received no income of
264 spat, no significant genetic differentiation was detected between the Society and Tuamotu
265 archipelagos. As for the samples from the Marquesas archipelagos, the sample from Hiva
266 Oa was significantly different from all other samples, except the ones from Makemo and

267 Hao. The sample from Nuku-Hiva, despite being more northern (but less eastern), was less
268 differentiated from samples from the Northern Tuamotu-Gambier, but remained distinct
269 from the southern part of the archipelago (Hereheretue, Marutea and Mangareva), and
270 Society samples. At the within-archipelago scale, despite the inclusion of samples from
271 remote atolls the values of F_{st} between pairs of samples for the nine wild populations from
272 Tuamotu-Gambier were very low, and most were null (Table 4). The significant values
273 mostly reflected the differentiation of the sample from the very isolated atoll of
274 Hereheretue (Tuamotu-Gambier), one sample from the center of Tuamotu-Gambier
275 (Makemo) as well as two samples from southern Tuamotu-Gambier (Marutea and
276 Mangareva). The differentiation between the northern sample from Arutua (Northern
277 Tuamotu-Gambier) and the samples from Takapoto3 and Marutea (Southern Tuamotu-
278 Gambier) is also noteworthy. Finally, at the within-lagoon scale, Takapoto 1 and 3 are
279 significantly differentiated from each other, despite being the geographically nearest
280 samples and not showing any significant differentiation with the remote sample of
281 Takapoto 2.

282 As for the collectors, more differentiation is observed, particularly for the collector
283 1, which is almost completely differentiated from all the other samples from laying beds or
284 collectors from Takapoto (except collector 6), and from most samples from other Tuamotu
285 islands. Collector 4, 3 and 6 are respectively differentiated of one (Takapoto 3), two
286 (Takapoto 2 and 3) and all three samples from Takapoto natural population, whereas all
287 collectors from 2 to 6 are not different one from the others, and not significantly different
288 from any other Tuamotu samples.

289

290 *3.3 Variance in allelic frequencies and estimation of the effective number of genitors.*

291 The effective number of genitors estimated for each collector was low in any case,
292 being estimated by considering as a putative mother population either : i) each of the three
293 samples from Takapoto (from 3 to 93, except for one case with collector 2, for which no
294 upper limit could be estimated if Takapoto 2 would be the origin of recruits), ii) the pool of
295 those three samples (5 to 136), or iii) the pool of all other natural laying beds sampled in
296 Tuamotu archipelago. However, significantly more important values can be observed
297 while considering this last case (from 22 to ∞).

298

299 *3.4 Results of the literature survey*

300 Results of the literature survey are summarized in Figure 3. Among 66 articles
301 screened about 65% did not report any indication as to the sampling area explored or
302 distance among samples; among these none gave any indication of the strategy for
303 choosing individuals (according some particular criteria like size for example), and only 8
304 of those 65% acknowledged a random or haphazard strategy for picking samples. Now,
305 35% of studies screened gave indication as to the area encompassed by sampling, or the
306 distance among collected samples, and 90% of those documented studies also reported the
307 strategy for picking up samples (17% reported an exhaustive collection, 65% a random
308 choice and 9% other kind of strategy such as sampling along a transect or in a grid).

309

310

311 **Discussion**

312

313 *Spatial distribution of genetic variability at large scale*

314 At the large scale (more than 1500km), in agreement with previous studies
315 (Arnaud-Haond et al., 2003a; Arnaud-Haond et al., 2004), the Marquesas samples, and
316 particularly the one from Hiva Oa, display a significant level of differentiation with those
317 from the two other archipelagos, reaching 8 to 10 % when comparing Hiva Oa with
318 Society samples. Genetic differentiation of Marquesas populations is also evident in other
319 species, and particularly well described for the surgeonfish *Acanthurus triostegus* (Planes
320 et al. 1996; Planes and Fauvelot 2002). Two main phenomena are commonly invoked to
321 explain this isolation. The first is the Marquesas countercurrent, which is regular
322 throughout the year and which is opposed by the South Equatorial current; this might
323 constitute a barrier preventing larval dispersal (Vermeij 1987; Planes and Fauvelot 2002).
324 A second possibility is the influence of glaciations and co-occurring sea level drops and
325 lagoon drainage (Paulay 1990), have led to extinction-recolonization of most inner-reef
326 species in the lagoons; whereas the Marquesas islands might have represented refugia for
327 some of those species. In the case of the pearl oyster, it is likely that populations persisted
328 in Marquesas during the glacial episodes, whereas extinction and post-glacial (re)
329 colonisation is suspected for other Polynesian archipelagos (Arnaud-Haond et al. 2003a).
330 Although no pattern of monophyly of Marquesas haplotypes was observed with
331 mitochondrial DNA analysis (Arnaud-Haond et al. 2003a), the present data support
332 previous findings that suggest a present day restriction to gene flow that may be partly
333 attributed to the existence of current limiting exchange with other archipelagos.

334 The distribution of genetic variation between the populations of the Tuamotu-
335 Gambier archipelago supports the proposition that very little genetic differentiation exists

336 within this area. In part, this phenomenon probably reflects the occurrence of larval
337 dispersal during the three-week pelagic larval stage of *P. margaritifera*. Yet, it is also
338 partly attributable to the transfers practiced over the ten years preceding our last sampling
339 campaign (Arnaud-Haond et al. 2004). The interpretation of the results obtained in terms
340 of natural patterns of gene flow must then be performed with caution. Nevertheless, our
341 data suggest the existence of only a slight restriction to gene flow at a medium scale (more
342 than five hundred kilometers), and the occurrence of genetically differentiated stocks.
343 Indeed, some isolated atolls do exhibit indices of genetic differentiation, for example: the
344 atolls of Southern Tuamotu-Gambier (Marutea and Mangareva), and the isolated atoll of
345 Hereheretue (Southern Tuamotu-Gambier). Given the scale of sampling for those
346 populations (see Table 1), we are confident the samples are likely to be representative of
347 the populations inhabiting those islands, and structure observed reflects restriction to gene
348 flow rather than genetic patchiness.

349 It is striking that at the smaller scale, of less than 300 km (Table 4), much more
350 differentiation (20% of the sample pairs) is observed than at medium scale (Table 2; 8% of
351 the sample pairs). This discrepancy is however mostly driven by the genetic heterogeneity
352 of samples gathered in very restricted areas: the ones from collectors, and one of the three
353 Takapoto samples (Takapoto 3) that shows indices of slight differentiation with both
354 Arutua and Takapoto 1 samples. These findings support the occurrence of recruitment
355 patchiness at that scale, rather than long term and stable restriction to gene flow between
356 those sampling locations.

357

358 *Spatial genetic heterogeneity of recruitment*

359 Both biotic and abiotic factors have been proposed as possible sources of the
360 observed stochasticity as well as of the spatial and temporal patchiness of cohort recruits,

361 but the factors that may explain the genetic patchiness are still unknown (Johnson and
362 Black 1984; Watts et al. 1990; McShane and Smith 1991; David et al. 1997a). The main
363 observations that can be retained from the collector data include (1) the positive
364 relationship between the spat density and its departure from HWE and (2) the negative
365 relationship between spat density and genetic differentiation. All spat samples from high
366 density collectors (C4, C5, C6) exhibit significant heterozygote deficiencies and linkage
367 disequilibrium, but are not significantly differentiated from most other samples. On the
368 other hand, two of the three spat samples from low-medium density collectors (C1, C3)
369 show HWE and are responsible for most genetic differentiation observed, particularly the
370 very low density one (C1 where all individuals have been sampled).

371 Recruitment is a spatially and temporally stochastic process: what is observed at
372 one site might represent only a single recruited cohort; whereas, at the neighbour site,
373 several distinct cohorts might have settled. When combined with the heterogeneity in the
374 genetic composition of recruits (due to difference in the source of recruits or to pre-
375 settlement selection), this suggests HWE and genetic originality may occur mostly in
376 infrequent recruiting sites. In contrast recruited pools made-up of an admixture of distinct
377 cohorts will exhibit linkage disequilibrium and departure from Hardy-Weinberg
378 equilibrium due to Wahlund effect, but allelic frequencies more homogeneous with that of
379 the population they come from.

380

381 The estimation of the putative number of genitors N_g at the origin of a given spat
382 sample gives low values for all collectors, would we consider Takapoto or the whole
383 Tuamotu Gambier as a putative mother population. Trying to interpret those values in
384 terms of an absolute numerical estimation of N_g would be risky because these are
385 calculated on the basis of F_{st} estimates that may have a large error variance. However,

386 interesting information can be extracted from those data if one compares the different N_g
387 estimates involving different putative parents' genetic pools: a larger number is
388 systematically estimated while considering the whole Tuamotu- Gambier genetic pool as
389 putative parents' pool. This suggests that spat recruited in Takapoto represent an admixture
390 more representative of islands other than Northern Tuamotu-Gambier, and indicates that
391 the low number of parents at the origin of a given cohort is probably one of the factors
392 responsible for the heterogeneity of the genetic composition of recruits, as observed for the
393 closely related species *P. maxima* in Australia (Benzie and Smith-Keune 2006).

394 In conclusion, our data indicate that the spatio-temporal variability of spat
395 recruitment is accompanied by high spatio-temporal variability of the genetic composition
396 of recruits at local scale. Would this be an isolated observation, one may question the
397 generality of such finding, wondering about the possible influence of the large densities of
398 farmed stocks in Takapoto. However this fine scale "genetic patchiness", as described by
399 Johnsson et al. (1982, 1984), has been reported on a large range of non human-influenced
400 populations of marine invertebrates including gastropods (Campton et al. 1992; Johnson et
401 al. 1993; Holborn et al. 1994; Tatarenkov and Johannesson 1994), sea urchin (Watts et al.
402 1990; Edmands et al. 1996) and bivalves (Hedgecock 1994; David et al. 1997b; Benzie
403 and Smith-Keune 2006). Hedgecock proposed a "sweepstakes-chance matching
404 hypothesis", suggesting that in a heterogeneous and changing environment, young
405 succeeding in recruiting and surviving are the products of spawning by only a small
406 fraction of the adult population that spawned in the right windows of time and
407 environmental conditions. This hypothesis favours the importance of pre-settlement factors
408 at the origin of genetic mosaics in natural populations, rather than micro-environmental
409 post-settlement selection that has been demonstrated on one species of gastropods
410 (Johannesson et al. 1995). In other cases, the variation of the pattern over time as well as

411 the chaotic distribution of genetic structure in space, suggested that pre-settlement factors
412 were more likely than post-settlement micro-environmental selection (Johnson and Black
413 1982; Johnson and Black 1984; Watts et al. 1990). In our case, the lack of consistence in
414 space and the genetic differentiation of collectors located close to each other suggest that
415 the “sweepstakes-chance matching hypothesis” is more likely to explain the genetic
416 patchiness observed than a post settlement selective process.

417

418

419 *Implication for sampling and interpreting data in population genetics and biogeographic*
420 *studies*

421

422 This study illustrates a finding reported in several other benthic marine species: local
423 genetic structuring (several metres to several kilometres) sometimes reaching or exceeding
424 the differentiation reported at larger scale (hundreds to thousands kilometres). We have
425 highlighted that this apparent inconsistency is strongly related to the sampling design of
426 the study. The detection of genetic patchiness was only achieved by sampling in very
427 restricted areas compared to the ones that were explored (when such information was
428 available) when assessing large scale patterns and processes.

429 Chaotic genetic patchiness is now increasingly reported in marine invertebrates
430 (Jolly et al. 2003; Juinio-Menez et al. 2003; Casu et al. 2005; Virgilio and Abbiati 2006;
431 Virgilio et al. 2006; Andrade and Solferini 2007) and on fishes (Doherty et al. 1995;
432 Exadactylos et al. 1998; Planes et al. 2002; McPherson et al. 2003; Selkoe et al. 2006;
433 Burford and Larson 2007; Gonzalez-Wanguemert et al. 2007). Besides the evolutionary
434 and biological causes and implications of chaotic genetic patchiness, this phenomenon
435 becomes an issue when it comes to collecting a sample ‘representative of the population

436 studied'. All marine biologists face the problem of access to samples that is increasingly
437 difficult when studying a large-scale metapopulation system. It is seldom realistic to
438 expect an extremely sophisticated, hierarchically designed, and standardized sampling
439 scheme, particularly when working on sub-tidal and deep-sea organisms. Most of us have
440 to be content with a "statistically satisfying" number of samples collected, sometimes
441 including samples kindly provided by other researchers or local people, without the
442 minimum information as to the habitat, area covered, density of populations, etc. As a
443 result, only 33% of 68 studies dealing with population genetics of marine benthic
444 organisms gave indication as to the area and strategy of sampling in each population, and
445 57% did not give any indication of either the area explored, nor the choice of samples
446 collected. Technical and logistical difficulties may not allow much improvement in the
447 sampling we, as researchers, will have access to. However, the occurrence of genetic
448 patchiness should warn us against conclusions of genetic composition or limitations to
449 gene flow based on blind samples that depending on some characteristics of the sampling
450 sites (such as area explored or density) may reflect transient and localized genetic
451 patchiness rather than significant and long-term restriction to gene flow among localities
452 analyzed.

453

454

455

456

457 **Acknowledgements**

458

459 We would like to thank all the people who helped us to collect samples: Terii Seaman and
460 Patrick Jilcot from the Pearl-Culture Service, Peva Levy, Claude Soyez, Marcel Matarere,
461 Destremeau Poroi, Herle Gorager and Johan Hamblin from Ifremer and Philippe Shoun.

462 We wish to address a very special thank to Gaby Hauamani, without whom none of the
463 work on spat would have been possible, and Vincent Prasil for useful discussions and
464 information about spat translocations. We thank Helen Boudry for help with the English in
465 this manuscript, and three anonymous referees for their useful comments and suggestions.

466 This work was supported in part by the "Contrat de développement n°2 2000-2003 entre
467 l'Etat et le Territoire de Polynésie Française" and the BRG (Bureau des Ressources
468 Génétiques)

469

470

471 **References**

472

473

474 Andrade SCS, Solferini VN (2007) Fine-scale genetic structure overrides macro-scale
475 structure in a marine snail: nonrandom recruitment, demographic events or
476 selection? *Bio J of the Linn Soc* 91: 23-36

477 Arnaud-Haond S, Bonhomme F, Blanc F (2003a) Large discrepancies in differentiation of
478 allozymes, nuclear and mitochondrial DNA loci in recently founded Pacific
479 populations of the pearl oyster *Pinctada margaritifera*. *J Evol Biol* 16: 388-398

480 Arnaud-Haond S, Boudry P, Saulnier D, Seaman T, Vonau V, Bonhomme F, Goyard E
481 (2002) New anonymous nuclear DNA markers for the pearl oyster *Pinctada*
482 *margaritifera* and other *Pinctada* species. *Mol Ecol Notes* 2: 220-222

483 Arnaud-Haond S, Vonau V, Bonhomme F, Boudry P, Prou J, Seaman T, Goyard E (2004)
484 On the impact of cultural practices on genetic resources: evolution of the genetic
485 composition of wild stocks of pearl oyster (*Pinctada margaritifera cumingii*) in
486 French Polynesia after ten years of spat translocation. *Mol Ecol* 13: 2001-2007

487 Arnaud-Haond S, Vonau V, Bonhomme F, Boudry P, Prou J, Seaman T, Veyret M,
488 Goyard E (2003b) Sustainable management of local genetic resources of the pearl
489 oyster (*Pinctada margaritifera cumingii*) in French Polynesia: an evaluation of the
490 potential impact of the cultural practice of spat collection. *Aquaculture* 219: 181-
491 192

492 Arndt A, Smith MJ (1998) Genetic diversity and population structure in two species of sea
493 cucumber: differing pattern according to mode of development. *Mol Ecol* 7: 1053-
494 1064

- 495 Bassam BJ, Caetano-Anolles G, Greshof PM (1991) Fast and sensitive silver-staining of
496 DNA in polyacrylamide gels. *Annals of Biochemistry* 196: 80-83
- 497 Belkhir K, Borsa P, Chikhi L, N.Raufaste, Bonhomme F (1996-2001) GENETIX 4.02,
498 logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome
499 et Populations, Interactions, Adaptations, CNRS UMR5000, Université Montpellier
500 II, Montpellier (France)
- 501 Benzie JAH (2000) The detection of spatial variation in widespread marine species:
502 methods and bias in the analysis of population structure in the crown of thorns
503 starfish (Echinodermata : *Astroidea*). *Hydrobiologia* 420: 1-14
- 504 Benzie JAH, Smith-Keune C (2006) Microsatellite variation in Australian and Indonesian
505 pearl oyster *Pinctada maxima* populations. *Mar Ecol Prog Ser* 314: 197-211
- 506 Black WC, Krafur ES (1985) A FORTRAN program for the calculation and analysis of
507 two-locus linkage disequilibrium coefficients. *Theor App Gen* 70: 491-496
- 508 Blanc F, Durand P, Shine-Milhaud M (1985) Genetic variability in populations of black
509 pearl oyster *Pinctada margaritifera* (Mollusque, Bivalve) de Polynésie. Proc. Fifth
510 International Coral Reefs Congress, Tahiti, 1985 4: 113-118
- 511 Bonhomme F, Planes S (2000) Some genetic arguments about what maintains the pelagic
512 phase in reef fishes. *Env Biol of Fishes* 59: 365-383
- 513 Boudry P, Collet B, Cornette F, Hervouet V, Bonhomme F (2002) High variance in
514 reproductive success of the Pacific oyster (*Crassostrea gigas*, Thunberg) revealed
515 by microsatellite-based parentage analysis of multifactorial crosses. *Aquaculture*
516 204: 283-296
- 517 Burford MO, Larson RJ (2007) Genetic heterogeneity in a single year-class from a
518 panmictic population of adult blue rockfish (*Sebastes mystinus*). *Mar Biol* 151:
519 451-465

- 520 Campton DE, Berg CJ, Robison LM, Glazer RA (1992) Genetic patchiness among
521 populations of queen conch *Strombus gigas* in the Florida keys and Bimini. Fish
522 Bull 90: 250-259
- 523 Casu M, Maltagliati F, Cossu P, Lai T, Galletti MC, Castelli A, Commito JA (2005) Fine-
524 grained spatial genetic structure in the bivalve *Gemma gemma* from Maine and
525 Virginia (USA), as revealed by Inter-Simple Sequence Repeat markers. J Exp Mar
526 Biol Ecol 325: 46-54
- 527 David P, Berthou P, Noel P, Jarne P (1997a) Patchy recruitment patterns in marine
528 invertebrates: a spatial test of the density-dependent hypothesis in the bivalve
529 *Spisula ovalis*. Oecologia 111: 331-340
- 530 David P, Perdieu M-A, Pernod A-F, Jarne P (1997b) Fine-Grained Spatial and Temporal
531 population genetic structure in the marine bivalve *Spisula ovalis*. Evolution 51 (4):
532 1318-1322
- 533 Desmarais E, Lanneluc I, Lagnel J (1998) Direct Amplification of Length Polymorphism
534 (DALP), or how to get and characterise new genetic markers in many species. Nucl
535 Ac Res 26(6): 1458-1465
- 536 Doherty PJ, Planes S, Mather P (1995) Gene Flow and Larval Duration in 7 Species of
537 Fish from the Great-Barrier-Reef. Ecology 76: 2373-2391
- 538 Durand P, Blanc F (1986) Divergence génétique chez un bivalve marin tropical: *Pinctada*
539 *margaritifera*. Coll Nat CNRS "Biologie des populations": 323-330
- 540 Edmands S, Moberg PE, Burton RS (1996) Allozyme and mitochondrial DNA evidence of
541 population subdivision in the purple sea urchin *Strongylocentrotus purpuratus*. Mar
542 Biol 126: 443-450
- 543 Exadactylos A, Geffen AJ, Thorpe JP (1998) Population structure of the Dover sole, *Solea*
544 *solea* L., in a background of high gene flow. J Sea Res 40: 117-129

- 545 Friedman KJ, Bell JD (1999) Variation in abundance of blacklip pearl oyster (*Pinctada*
546 *margaritifera* Linne.) spat from inshore and offshore reefs in Solomon islands.
547 *Aquaculture* 178: 273-291
- 548 Friedman KJ, Bell JD, Tiroba G (1998) Availability of wild spat of the blacklip pearl
549 oyster, *Pinctada margaritifera*, from "open" reef systems in Solomon islands.
550 *Aquaculture* 167: 283-299
- 551 Gaines SD, Bertness MD (1992) Dispersal of juveniles and variable recruitment in sessile
552 marine species. *Nature* 360: 579-580
- 553 Gonzalez-Wanguemert N, Perez-Ruzafa A, Canovas F, Garcia-Charton JA, Marcos C
554 (2007) Temporal genetic variation in populations of *Diplodus sargus* from the SW
555 Mediterranean Sea. *Mar Ecol Prog Ser* 334: 237-244
- 556 Hedgecock D (1994) Does variance in reproductive success limit effective population size
557 of marine organisms? *Genetics and Evolution of Aquatic Organisms*: 122-134
- 558 Hellberg ME (1996) Dependence of gene flow on geographic distance in two solitary
559 corals with different larval dispersal abilities. *Evolution* 50(3): 1167-1175
- 560 Holborn K, Johnson MS, Black R (1994) Population genetics of the corralivorous
561 gastropod *Drupella cornus* at Ningaloo reef, Western Australia. *Coral Reefs* 13:
562 33-39
- 563 Hoskin MG (1997) Effects of contrasting modes of larval development on the genetic
564 structures of populations of three species of prosobranch gastropods. *Mar Biol* 127:
565 647-656
- 566 Hunt A (1993) Effects of contrasting patterns of larval dispersal on the genetic
567 connectedness of local populations of two intertidal starfish, *Patiriella calcar* and
568 *P. exigua*. *Mar Ecol Prog Ser* 92: 179-186

- 569 Johannesson K, Johannesson B, Lundgren U (1995) Strong natural-selection causes
570 microscale allozyme variation in a marine snail. Proc Nat Ac Sci USA 92: 2602-
571 2606
- 572 Johnson MS, Black R (1982) Chaotic Genetic Patchiness in an Intertidal Limpet,
573 *Siphonaria* sp. Mar Biol 70: 157-164
- 574 Johnson MS, Black R (1984) Pattern beneath the chaos: the effect of recruitment on
575 genetic patchiness in an intertidal limpet. Evolution 38: 1371-1383
- 576 Johnson MS, Holborn K, Black R (1993) Fine-scale patchiness and genetic heterogeneity
577 of recruits of the corralivorous gastropod *Drupella cornus*. Mar Biol 117: 91-96
- 578 Johnson MS, Wernham J (1999) Temporal variation of recruits as a basis of ephemeral
579 genetic heterogeneity in the western rock lobster *Panulirus cygnus*. Mar Biol 135:
580 133-139
- 581 Jolly MT, Viard F, Weinmayr G, Gentil F, Thiebaut E, Jollivet D (2003) Does the genetic
582 structure of *Pectinaria koreni* (Polychaeta : *Pectinariidae*) conform to a source-
583 sink metapopulation model at the scale of the Baie de Seine? Helgoland Mar Res
584 56: 238-246
- 585 Juinio-Menez MA, Magsino RM, Ravago-Gotanco R, Yu ET (2003) Genetic structure of
586 *Linckia laevigata* and *Tridacna crocea* populations in the Palawan shelf and shoal
587 reefs. Mar Biol 142: 717-726
- 588 Li G, Hedgecock D (1998) Genetic heterogeneity, detected by PCR-SSCP, among samples
589 of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large
590 variance in reproductive success. Can J Fish Aquat Sci 55: 1025-1033
- 591 McPherson AA, Stephenson RL, Taggart CT (2003) Genetically different Atlantic herring
592 *Clupea harengus* spawning waves. Mar Ecol Prog Ser 247: 303-309

- 593 McShane PE, Smith MG (1991) Recruitment variation in sympatric populations of
594 *Haliotis rubra* (Mollusca: Gastropoda) in southeast Australian waters. Mar Ecol
595 Prog Ser 73: 203-210
- 596 Nei M (1987) Molecular evolutionary genetics Columbia Univ. Press, New York
- 597 Palumbi SR (1995) Nucleic acids II: The polymerase chain reaction. In: D. Hillis and C.
598 Moritz e (ed) Molecular systematics, Sinauer, Sunderland, Massachusetts, pp 205-
599 247
- 600 Paulay G (1990) Effects of late Cenozoic sea-level fluctuations on the bivalve faunas of
601 tropical oceanic islands. Paleobiology 16: 401-434
- 602 Peel D, Ovenden JR, Peel SL (2004) NeEstimator: software for estimating effective
603 population size,. Department of Primary Industries and Fisheries., Queensland
604 Government
- 605 Planes S, Fauvelot C (2002) Isolation by distance and vicariance drive genetic structure of
606 a coral reef fish in the Pacific Ocean. Evolution 56: 378-399
- 607 Planes S, Galzin R, Bonhomme F (1996) A genetic metapopulation model for reef fishes
608 in oceanic islands: the case of the surgeonfish, *Acanthurus triostegus*. J Evol Biol
609 9: 103-117
- 610 Planes S, Lecaillon G, Lenfant P, Meekan M (2002) Genetic and demographic variation in
611 new recruits of *Naso unicornis*. J Fish Biol 61: 1033-1049
- 612 Poulin E, Féral J-P (1996) Consequences of brood protection on the diversity of antarctic
613 echinoids. Evolution 50: 820-830
- 614 Roughgarden J, Gaines S, Possingham H (1988) Recruitment dynamics in complex life
615 cycles. Science 241: 1460-1466
- 616 Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, 2nd edition. Cold Spring
617 Harbor laboratory. Cold Spring Harbor, N. Y.

- 618 Selkoe KA, Gaines SD, Caselle JE, Warner RR (2006) Current shifts and kin aggregation
619 explain genetic patchiness in fish recruits. *Ecology* 87: 3082-3094
- 620 Tatarenkov A, Johannesson K (1994) Habitat related allozyme variation on a
621 microgeographic-scale in the marine snail *Littorina mariae* (Prosobranchia,
622 *Littorinacea*). *Bio J Linn Soc* 53: 105-125
- 623 Vermeij G (1987) The dispersal barrier in the tropical Pacific: implications for molluscan
624 speciation and extinction. *Evolution* 41: 1046-1058
- 625 Virgilio M, Abbiati M (2006) Temporal changes in the genetic structure of intertidal
626 populations of *Hediste diversicolor* (Polychaeta : *Nereididae*). *J Sea Res* 56: 53-58
- 627 Virgilio M, Backeljau T, Abbiati M (2006) Mitochondrial DNA and allozyme patterns of
628 *Hediste diversicolor* (Polychaeta : *Nereididae*): the importance of small scale
629 genetic structuring. *Mar Ecol Prog Ser* 326: 157-165
- 630 Waples RS (1989) A generalized approach for estimating effective population size from
631 temporal changes in allele frequency. *Genetics* 121: 379-391
- 632 Watts RJ, Johnson MS, Black R (1990) Effects of recruitment on genetic patchiness in the
633 urchin *Echinometra mathaei* in Western Australia. *Mar Biol* 105: 145-151
- 634 Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics* 35: 235-254
- 635 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of populations
636 structure. *Evolution* 38(6): 1358-1370
- 637
- 638
- 639

640 **Table 1** : Sample size (N), geographic location in the lagoon (see Fig 2), number of
 641 alleles, expected (H_{nb}) and observed (H_{obs}) heterozygosity, and heterozygote deficiency
 642 (F_{is}) * significant after a permutation test (1000 permutations).

643

archipelagos	sample	Samples	area	N	Alleles Nb	H_{nb}	H_{obs}	F_{is}
<i>Society</i>	Manuae	MA	-	23	3.00	0.304	0.268	0.118
	Maupihaa	MP	-	40	4.25	0.364	0.266	0.273*
<i>Tuamotu-Gambier</i>	Arutua	AR	-	28	3.25	0.376	0.314	0.167*
	Manihi	MH	-	22	3.25	0.377	0.263	0.307*
	Apataki	AI	1 km ²	28	3.75	0.359	0.372	0.092
	Hereheretue	HERE	>10 km ²	27	5.00	0.397	0.369	0.071
	Takaroa	TA	-	30	3.50	0.339	0.303	0.108
	Raroia	RA	-	15	3.25	0.406	0.293	0.285*
	Kauehi	KA	-	19	3.75	0.451	0.3691	0.188
	Makemo	MK	-	28	3.25	0.410	0.366	0.109
	Hao	HAO	0.06 km ²	12	2.50	0.355	0.275	0.234
	Marutea	MRT	0.4 km ²	29	3.00	0.323	0.294	0.094
	Mangareva	MG	4 km ²	40	3.75	0.368	0.323	0.125*
<i>Marquesas</i>	Nuku Hiva	NH	>10 km ²	42	3.00	0.396	0.353	0.111
	Hiva Oa	HO	>10 km ²	31	2.75	0.389	0.378	0.029
<i>Takapoto beds</i>	Takapoto 1	Tkp1	1	37	3.75	0.342	0.212	0.383*
	Takapoto 2	Tkp2	2	54	3.75	0.346	0.300	0.138
	Takapoto 3	Tkp3	3	29	4.25	0.366	0.315	0.142
<i>Takapoto collectors</i>	Collecteur 1	C1	C	37	4.00	0.294	0.252	0.144
	Collecteur 2	C2	B	52	3.75	0.365	0.242	0.338*
	Collecteur 3	C3	B	41	3.50	0.382	0.374	0.021
	Collecteur 4	C4	A	50	3.50	0.360	0.240	0.283*
	Collecteur 5	C5	A	50	4.00	0.392	0.282	0.148*
	Collecteur 6	C6	A	50	4.00	0.404	0.344	0.336*

644

645

Table 5 : Estimation of the average (N_e), maximum (max, $P<0.975$) and minimum (min, $P<0.025$) number of genitors of the cohorts sampled in the different collectors $N_{e \min}^{\max}$, using as a putative parental group the samples from Takapoto 1,2,3, the average of these three samples, and the average of all samples from samples from Tuamotu (Apataki, Manihi, Arutua and Takaroa).

sample	N_e	Tkp 1	Tkp 2	Tkp 3	Takapoto	Tuamotu
Collecteur 1	37	16_5^{93}	13_4^{47}	11_4^{43}	18_6^{59}	57_{13}^{∞}
Collecteur 2	52	14_5^{59}	12_4^{36}	10_4^{33}	18_7^{48}	41_{12}^{382}
Collecteur 3	41	14_4^{79}	9_3^{23}	8_3^{21}	12_5^{31}	132_{17}^{∞}
Collecteur 4	50	15_{43}^{69}	12_4^{43}	11_3^{40}	20_7^{65}	79_{16}^{∞}
Collecteur 5	50	14_4^{49}	35_6^{∞}	15_5^{75}	30_{10}^{136}	37_{12}^{270}
Collecteur 6	50	11_4^{36}	9_3^{24}	9_3^{28}	13_5^{34}	46_{22}^{14381}

Figure 1: Localization of the Polynesian islands where samples were collected, on the three principals archipelagoes from French Polynesia: Society, Tuamotu-Gambier and Marquesas.

Figure 2: Map of the Takapoto lagoon, indicating the spat collection stations (with high and homogeneous density: A, and low and heterogeneous density : B and C), where samples were collected (respectively 3, 2, and 1 spat collector sampled). Sampling zones of wild beds are also indicated (1, 2 and 3).

Figure 3: Results of the literature survey for sampling details and strategy.

Appendix : list of articles screened for sampling strategy

Figure 1
[Click here to download high resolution image](#)

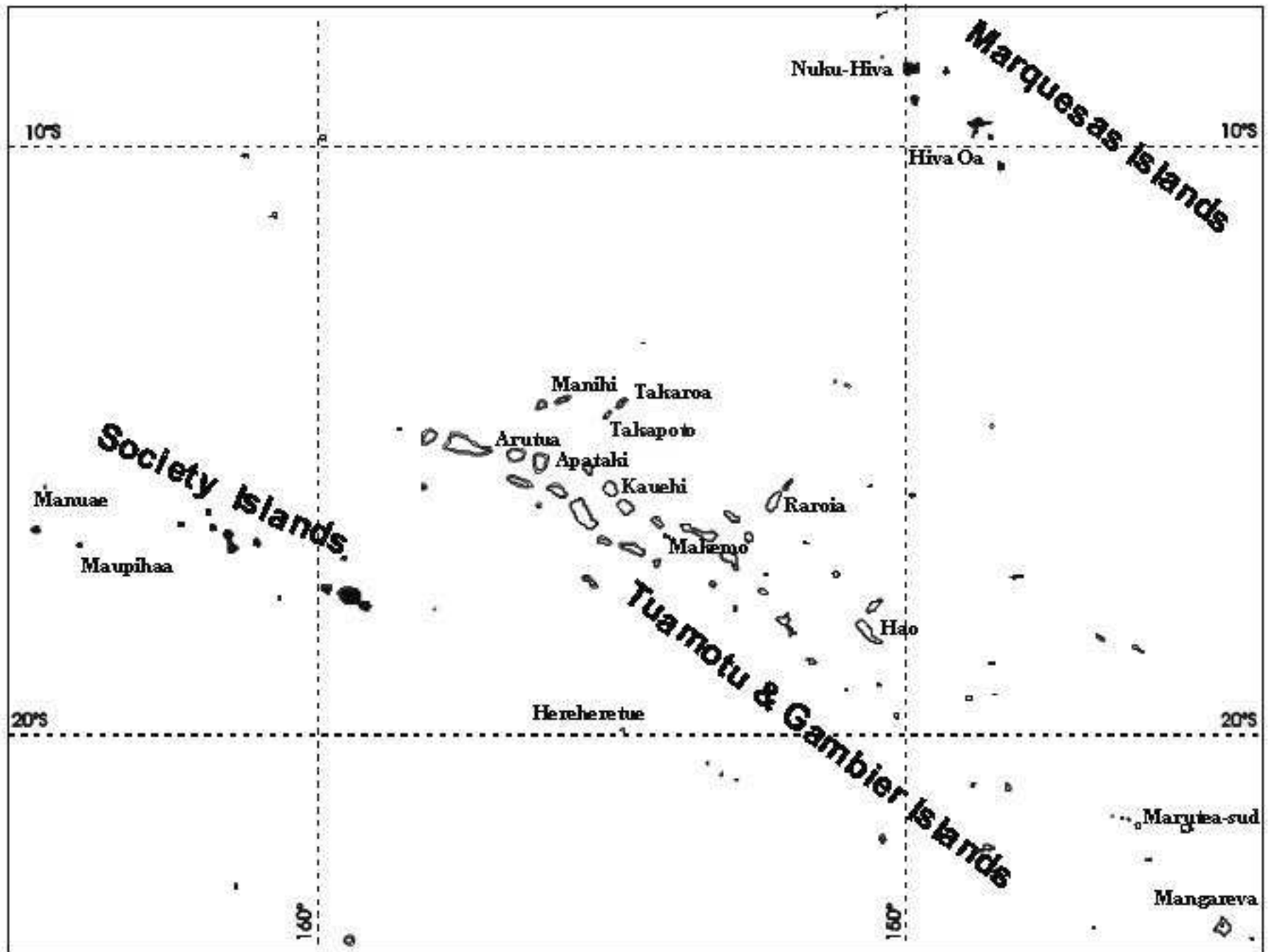


Figure 2
[Click here to download high resolution image](#)

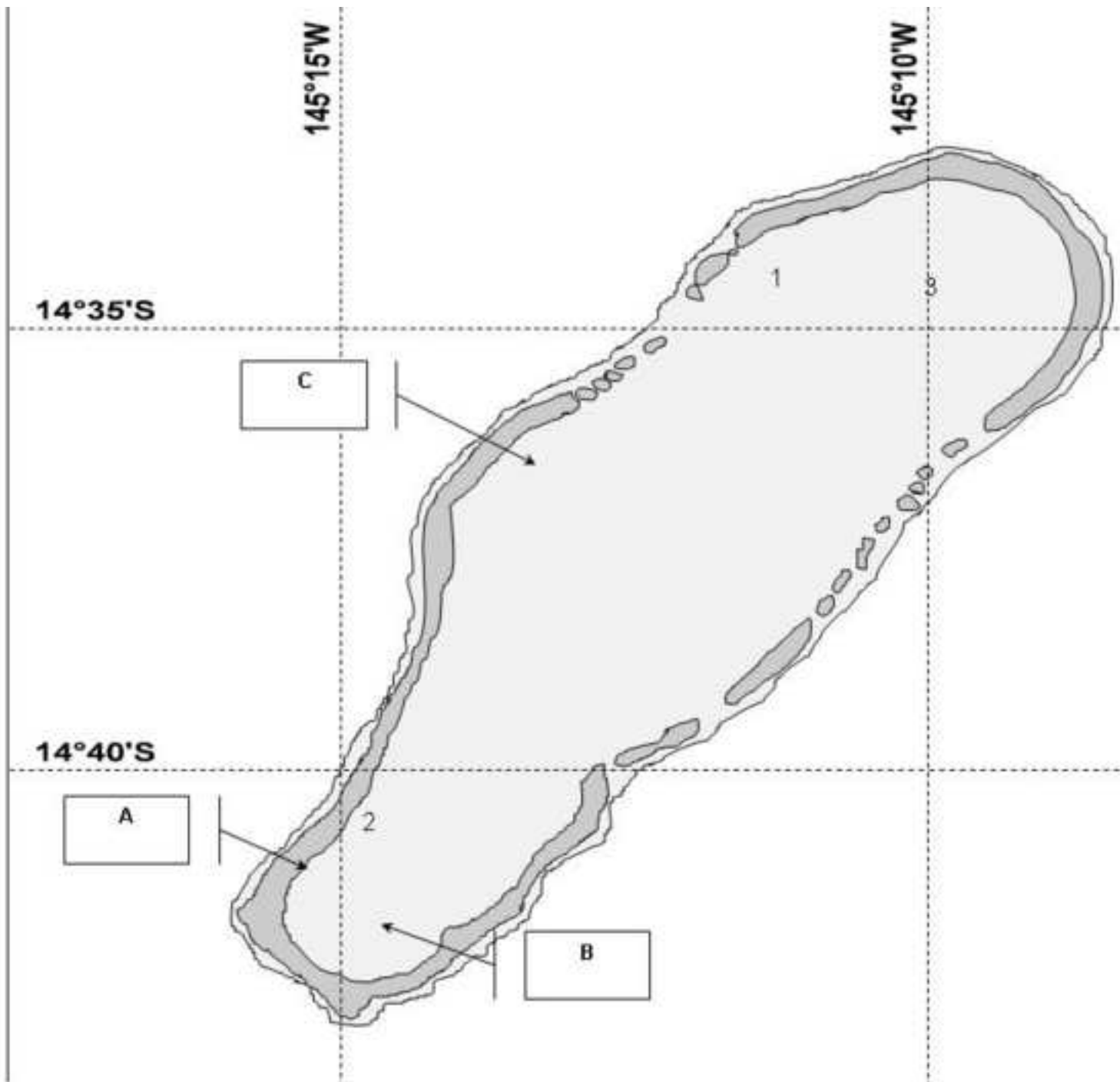


Figure 3
[Click here to download high resolution image](#)

