Genet. Res., Camb. (2001), **78**, *pp.* 259–270. *With 3 figures.*
DOI: 10.1017/S0016672301005353 Printed in the United Kingdom

Population bottleneck and effective size in *Bonamia ostreae*-resistant populations of *Ostrea edulis* as inferred by microsatellite markers

SOPHIE LAUNEY*, MARC BARRE, ANDRE GERARD AND YAMAMA NACIRI-GRAVEN†
IFREMER, Laboratoire Génétique et Pathologie (LGP), BP 133 Ronce les Bains, 17390 La Tremblade, France

(*Received 12 May 1999 and in revised form 17 May 2000 and 22 June 2001*)

SUMMARY

Genetic variability at five microsatellite loci was analysed in three hatchery-propagated populations of the flat oyster, *Ostrea edulis*. These populations were part of a selection programme for resistance to the protozoan parasite *Bonamia ostreae* and were produced by mass spawns, without control of the genealogy. Evidence for population bottlenecks and inbreeding was sought. A reduction in the number of alleles, mainly due to the loss of rare alleles, was observed in all selected populations, relative to the natural population from which they were derived. Heterozygote excesses were observed in two populations, and were attributed to substructuring of the population into a small number of families. Pedigree reconstruction showed that these two populations were produced by at most two spawning events involving a limited number of parents. Most individuals within these populations are half or full-sib, as shown by relatedness coefficients. The occurrence of population bottlenecks was supported by estimates of effective number of breeders derived by three methods: temporal variance in allelic frequencies, heterozygote excess, and a new method based on reduction in the number of alleles. The estimates from the different methods were consistent. The evidence for bottleneck and small effective number of breeders are expected to lead to increasing inbreeding, and have important consequences for the future management of the three *O. edulis* selected populations.

1. Introduction

The effective population size (N_e) is a critical parameter to monitor for the management of genetic resources, because it determines the rate of increase in inbreeding (ΔF), hence the rate of loss of genetic variability in a population ($\Delta F = 1/2N_e$) (Crow & Kimura, 1970). Conservation programmes are often based on captive breeding, which, if based on a small number of founders and breeders, can cause an increase in the inbreeding coefficient and a subsequent decrease in fitness known as inbreeding depression (see for instance Backus *et al.*, 1995).

In the case of aquatic species and especially marine bivalves, improvements in hatchery technology in the past decade have allowed aquaculture broodstocks to be kept as closed populations, without input of individuals from other wild or captive populations. Moreover, aquatic animals can have very high fecundity but variable fertilities, and there is evidence that hatchery seed may come from only a few successfully spawning individuals (Gaffney *et al.*, 1992). Estimates of N_e for cultivated species suggest that closed aquatic broodstocks are small populations in which random drift can be important (Sbordoni *et al.*, 1986; Hedgecock & Sly, 1990; Hedgecock *et al.*, 1992).

The direct assessment of effective size requires measurement of many demographic parameters that are difficult to carry out in natural populations. Many different indirect methods have been used to evaluate N_e, based on genetic variability at marker loci. Several studies have focused on estimating N_e indirectly from temporal changes in allelic frequencies in finite

* Corresponding author. Laboratoire de Génétique des Poissons, INRA-CRJ, 78352 Jouy-en-Josas, France. Tel: +33 1 34 65 27 57. Fax: +33 1 34 65 23 90. e-mail: launey@jouy.inra.fr
† Current address: Conservatoire et Jardin Botaniques, 1, chemin de l'Impératrice, CP 60, 1292 Chambésy, Geneva, Switzerland.

populations (Pollack, 1983). This temporal method has been shown to be robust and well suited to the study of small captive populations (Waples, 1989), and particularly of aquaculture broodstock (Hedgecock & Sly, 1990; Hedgecock *et al*., 1992). Hill (1981) has developed a statistical method for estimating effective population size from data on linkage disequilibrium among gene loci. However, this method requires prior knowledge of linkage relationships among loci, and may produce a large standard error if only a small number of loci are considered (Hill, 1981). Finally, Pudovkin *et al*. (1996) have developed a method to estimate the effective numbers of breeders, based on the heterozygote excess in a progeny born from a small number of parents; this heterozygote excess is caused by differences, due to binomial sampling errors, in allelic frequencies between male and female parents.

A reduction in the effective population size following a population bottleneck is correlated with a decrease in heterozygosity and a loss of rare alleles (Wright, 1931). Nei *et al*. (1975) even predicted that during a bottleneck allele diversity would decrease faster than heterozygosity. However, the evolution of the actual number of alleles at a locus has not been used to estimate effective size. We developed a method, modelling the reduction in the number of alleles in a population as a function of the sample size, which can be used to estimate the effective number of founders in a population that has undergone a bottleneck.

In this paper we investigate effective size and inbreeding in three hatchery-maintained populations of *Ostrea edulis* selected for resistance to *Bonamia ostreae*, a widespread protozoan parasite that has endangered the flat oyster *O. edulis* in France, the Netherlands, Spain, Denmark, England and Ireland, and in North America (see review in Naciri-Graven *et al*., 1998). Our purpose is to assess *a posteriori* the effect of bottlenecks and a small effective number of breeders on the genetic variability in the three selected *O. edulis* populations. We also compare estimates of N_e obtained from temporal variance in allelic frequencies, heterozygote excess and reduction in number of alleles. Microsatellites markers have been chosen as the molecular tool to investigate these issues; not only are they suitable for genetic diversity studies (Jarne & Lagoda, 1996), but their high polymorphism makes them particularly adapted to studies dealing with relatedness and kinship assessment (Queller *et al*., 1993; Blouin *et al*., 1996).

2. Materials and methods

(i) *Biological material*

Three different hatchery populations of *Ostrea edulis*, namely S85, S89W and S89I, have been developed since 1985 and 1989 respectively. These populations have shown a significant increase in 'resistance' to *Bonamia ostreae* (Naciri-Graven *et al*., 1998). Fig. 1 shows the relationship between the populations. The first generation (G1) of S85 and S89W/S89I and the second generation (G2) of S85 were produced in 1985, 1989 and 1990 respectively. All populations were derived from wild animals from Quiberon Bay (Brittany, France), an area where *B. ostreae* is known to be endemic. The first generation of S85 was divided in three different groups that were reared in different locations in Brittany (namely Paimpol, Quiberon and La Trinité); the three groups were not subsequently mixed for the production of the G2. S85-G1, S85-G2, S89I-G1 and S89W-G1 have been produced by mass spawning: around 100 animals were placed in a raceway, and larvae were collected over a short period of time (1–3 days). Such a technique does not allow for pedigree management because the number and the identity of the animals that actually spawn is unknown.

All animals still alive in 1995 from S85-G2, S89I-G1 and S89W-G1 (67, 38 and 52 animals respectively) were sampled. These were the last generations produced by mass spawning and were used as parents in a biparental crossing experiment in 1995 (Naciri-Graven *et al*., 2000). A sample of 49 individuals recruited in Quiberon Bay in 1993 was used as a control wild population, representative of the wild population from which the founders of the selected lines were drawn. For all animals sampled, a piece of gill or mantle tissue was collected by biopsy and stored either in 100% ethanol at room temperature, or dried at −80 °C.

(ii) *DNA extraction and microsatellite amplification*

After removal of the ethanol, 400 μl of 5% Chelex, 30 μl of TE buffer (Tris 0·01 mM, EDTA 1 mM) and 10 μl of Proteinase K (10 mg/ml, Boehringer-Mannheim) were added to each sample, which was then incubated at 55 °C for 4 h. The samples were then boiled for 15 min, vortexed for 30–60 s and centrifuged at 8000 rpm for 5 min. The supernatant was collected for storage at −20 °C and the Chelex beads were discarded.

Five microsatellite loci specific to *O. edulis* (*Oedu*J12, *Oedu*U2, *Oedu*H15, *Oedu*O9 and *Oedu*T5; Launey, 1998) were used in this study.

Amplifications were performed individually in a 10 μl total volume containing 2 μl DNA, 0·4 μM of each primer (one labelled with ^{33}P), 1·5 mM $MgCl_2$, 75 μM of each dNTP and 0·35 unit of red Goldstar DNA polymerase (Eurogentec, Liège, Belgium). Thirty amplification cycles (1 min at 94 °C, 1 min at the optimum hybridization temperature, 1 min 15 s at

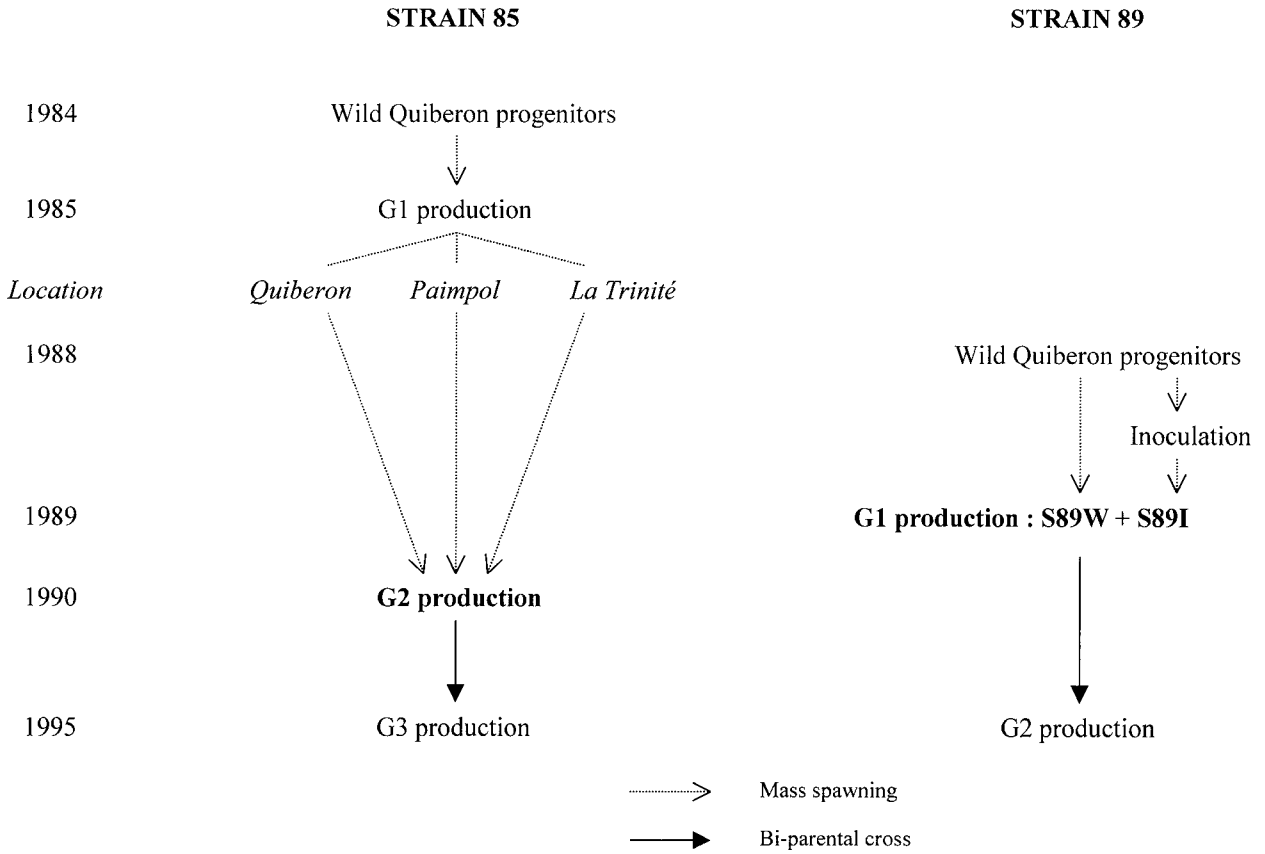

Fig. 1. Chronological record of selected *O. edulis* strains development. The generations studied in this paper are indicated in bold.

72 °C) were run in a PTC-100 Thermocycler (MJ Research, Watertown, MA). PCR products were electrophoresed on 10% PAGE gels (acrylamide: bisacrylamide, 29:1, 7 M urea) using 0·5× TBE (Tris borate EDTA) buffer. Results were visualized by autoradiography after exposing the dry gel overnight.

(iii) *Analysis of genetic parameters*

Standard descriptors of the genetic variability were calculated for each locus in each selected population and the control: the number of alleles (N_a), allelic frequencies, observed heterozygosity (H_0) and unbiased expected genic diversity ($H_{exp} = (1-\sum p_i^2)\ 2N/(2N-1)$, (Nei, 1978) where N is the sample size and p_i the frequency of the i^{th} allele). Departures from Hardy–Weinberg equilibrium were estimated by calculations of Wright's *Fis* according to Weir & Cockerham (1984). The null hypothesis, $Fis = 0$, was tested by bootstrapping over alleles. Calculations were performed using the program Genetix (Belkhir *et al.*, 1996).

Genetic relatedness, r, between each pair of individuals within each selected population and the control, was calculated according to Queller & Goodnight (1989) using the software Relatedness developed by these authors. r was computed as follow:

$$r = \frac{\sum_x \sum_k \sum_l (P_y - P^*)}{\sum_x \sum_k \sum_l (P_x - P^*)}$$

where x indexes individuals in the data set, k indexes loci and l indexes allelic position (1 or 2 for a diploid individual). The variables used here can be defined as follow: P_x is the frequency within the current x individual of the allele found at x's locus k and allelic position l; P_y is the frequency of that same allele in the selected line from which x comes; and P^* is the frequency of the allele in the Quilberon population, used here as a baseline population.

The mean relatedness value r was computed for each selected population (r_{85}, r_{89I}, r_{89W}) and the Quiberon population (r_Q); standard errors were estimated by jackknifing over loci. Differences between mean relatedness were tested by a t-test. Groups of related individuals were identified by UPGMA clustering, using $(1-r)$ as a measure of genetic distance. This technique has proved useful for identifying family relationships. For instance, Blouin *et al.* (1996) used relatedness estimates and UPGMA clustering to recover without misclassifications the true family relationships among the offspring of four sets of

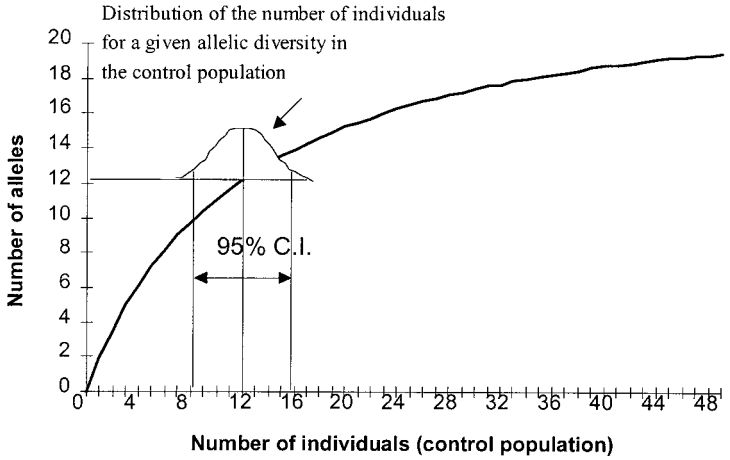

Fig. 2. Estimation of the number of founders in a selected strain, from the reduction in the number of alleles in comparison with a control population. The number of founders is estimated by (i) calculating using successive bootstrapping the number of alleles in 1000 subsamples of size p from a population of size N ($1 < p < N$), (ii) calculating the number of alleles (N_a) in the selected strain, (iii) estimating the mean size of a subsample of the control population that would have the same allelic richness N_a. In this particular example, the control population ($N = 49$) has a total number of alleles of 19. The selected population has 12 alleles, which is the allelic diversity of a subsample of 12 individuals from the control population. An estimate of the number of founders for this selected strain is therefore 12 (95% CI: 9·1–15·7).

independent maternal half-sibs in the mouse, *Mus musculus*, using 10 microsatellite loci.

On the basis of the presence of alleles shared between individuals, tentative pedigrees were reconstructed for the different selected populations. These putative pedigrees were then compared with the UPGMA clustering. This moreover provides a first estimate of the number of founder parents.

(iv) *Estimation of the effective numbers of founders*

In the absence of genealogical information, we used different methods to estimate the number of founders at the origin of each selected population, including one that we developed for this study.

(a) *Temporal variance in allelic frequencies* (Waples, 1989)

Allelic frequencies and an estimate of temporal variance in allelic frequency, F, which is standardized to compensate for differences among loci in initial allelic frequencies, were calculated. According to Hedgecock *et al.* (1992), we used Pollack's (1983) estimator for F, F_k. An estimator of the harmonic mean effective population number over the time interval is:

$$N_k = t(2[F - 1(2S_0) - 1/(2S_t)]) \quad (1)$$

where F is the population mean, estimated by the mean of F_k across the different loci, weighted by the number of alleles; and t is the number of generations in the time interval between the initial sample of S_0 individuals and the second sample of S_t individuals. The Quiberon population was used as the initial sample; the number of generation was 2 for S85-G2 and 1 for both S89I-G1 and S89W-G. In addition, in order to estimate the effective size and the effect of drift in the control population, N_k was estimated for the Quiberon population using two samples from different years (animals recruited in 1989 and in 1993, considered to be separated by 2 generations). Standard errors for N_k were calculated by using a χ^2 approximation (Waples, 1989). Calculations were performed with the program EPS, provided by D. Hedgecock and V. Chow (Bodega Marine Laboratory, UC Davis).

(b) *Heterozygote excess* (Pudovkin *et al.*, 1996)

H_0 being the observed heterozygosity in the progeny, and H_{exp} being the expected heterozygosity under Hardy–Weinberg equilibrium, Pudovkin *et al.* (1996) showed that an estimate of the number of breeders, N_{eb}, which in our case is equivalent to the effective size N_e, is

$$N_{eb} = 1/(2D) + 1/(2(D+1)) \quad (2)$$

where $D = (H_0 - H_{exp})/H_{exp}$.

Confidence intervals were obtained by jackknifing over loci.

(c) *Reduction in the number of alleles*

We developed a simple method to estimate the effective number of breeders in a population that had undergone a bottleneck, if samples from this population before and after the bottleneck are available, based on the reduction in the number of alleles. The basic idea is to model bottlenecks of various intensities in a population of known allelic composition, and estimate the number of alleles in the population immediately

following this bottleneck. From a population of N individuals of known genotypes, a sample of size p is drawn, and the number of alleles in that sample is computed. The mean and variance of the number of alleles for a given sample size p are calculated by doing 1000 successive samplings. It is then possible to graphically represent the expected number of alleles (N_a) as a function of the size of a subsample of the population (N_i). Using this algorithm, the curves $N_a = f(N_i)$ were constructed for each locus for the Quiberon population. They were modelled by:

$$N_a = C \cdot (1 - e_i^{-kN}) \quad (3)$$

Parameters C and k were estimated using the *nlin* procedure from the SAS software. If N_a is the number of alleles in a selected population, then we can determine N_i from (3). This can be used as an estimate of the effective number of founders (and 95% confidence interval) for each selected population (see Fig. 2).

3. Results

(i) *Genetic variability in the selected populations*

The genetic variability for the selected populations is shown in Table 1, in comparison with the Quiberon population. Allelic frequencies for the different populations are given in the Appendix. The mean number of alleles by locus ($\pm$standard deviation) is $19{\cdot}8 \pm 2{\cdot}2$ significantly in the control, and is reduced in all the selected populations: $12{\cdot}8 \pm 2{\cdot}2$ alleles/locus in S85-G2 ($P < 0{\cdot}05$), $9{\cdot}4 \pm 0{\cdot}9$ for S89I-G1 ($P < 0{\cdot}01$) and $5{\cdot}4 \pm 0{\cdot}5$ for S89W-G1 ($P < 0{\cdot}01$). This reduction in allele number is due in large part to the loss of alleles that were at low frequency in the control population. In the control population the mean frequency of the alleles that are not shared with S85-G2 is $0{\cdot}025 \pm 0{\cdot}019$ vs $0{\cdot}068 \pm 0{\cdot}054$ for the alleles present both in the control and in S85-G2 ($P < 0{\cdot}001$); similarly, these frequencies are $0{\cdot}033 \pm 0{\cdot}028$ vs $0{\cdot}076 \pm 0{\cdot}058$ for S89I-G1 ($P < 0{\cdot}001$) and $0{\cdot}037 \pm 0{\cdot}039$ vs $0{\cdot}076 \pm 0{\cdot}062$ for S89W-G1 ($P < 0{\cdot}01$).

The mean observed heterozygosity level is lower in S85-G2 ($H_0 = 0{\cdot}716 \pm 0{\cdot}087$) than in the control ($H_0 = 0{\cdot}840 \pm 0{\cdot}089$; $P < 0{\cdot}001$). On the contrary, the mean heterozygosity level in S89I-G1 ($H_0 = 0{\cdot}859 \pm 0{\cdot}143$) and S89W-G1 ($H_0 = 0{\cdot}969 \pm 0{\cdot}051$) are respectively not significantly different from ($P = 0{\cdot}40$), and significantly higher than ($P < 0{\cdot}001$), that of the control.

The *Fis* values are given in Table 2. The control population shows a slight overall heterozygote de-

Table 1. *Genetic variability in the selected populations and the control population*

Locus		Control ($n = 49$)	S85 ($n = 67$)	S891 ($n = 38$)	S89W ($n = 52$)
*Oedu*J12	N_a	22	14	9	5
	H_0	0·918	0·582	0·649	0·961
	H_{exp}	0·926	0·877	0·790	0·766
*Oedu*U2	N_a	25	13	11	6
	H_0	0·896	0·812	1	1
	H_{exp}	0·939	0·849	0·848	0·787
*Oedu*H15	N_a	14	10	9	5
	H_0	0·698	0·688	0·778	1
	H_{exp}	0·884	0·817	0·776	0·777
*Oedu*O9	N_a	18	11	9	5
	H_0	0·878	0·746	0·919	0·882
	H_{exp}	0·921	0·772	0·779	0·746
*Oedu*T5	N_a	20	15	9	6
	H_0	0·809	0·750	0·947	1
	H_{exp}	0·900	0·804	0·809	0·803

N_a is the number of alleles, H_0 the observed heterozygosity, H_{exp} the unbiased expected heterozygosity assuming Hardy–Weinberg equilibrium (Nei, 1978).

Table 2. *Conformity to Hardy–Weinberg equilibrium as estimated by* Fis *values*

Fis	Multilocus	*Oedu*J12	*Oedu*U2	*Oedu*H15	*Oedu*O9	*Oedu*T5
Control	0·082***	0·008 (NS)	0·047 (NS)	0·213***	0·048 (NS)	0·102*
S85	0·104***	0·319***	−0·045 (NS)	0·144**	0·036 (NS)	0·047 (NS)
S89I	−0·074*	0·181*	−0·182***	−0·003 ns	−0·183**	−0·174**
S89W	−0·239***	−0·244***	−0·260***	−0·278***	−0·173**	−0·237***

Fis values estimated according to Weir & Cockerham (1984). Ns: $P > 0{\cdot}05$; *$P < 0{\cdot}05$; **$P < 0{\cdot}01$; ***$P < 0{\cdot}001$.

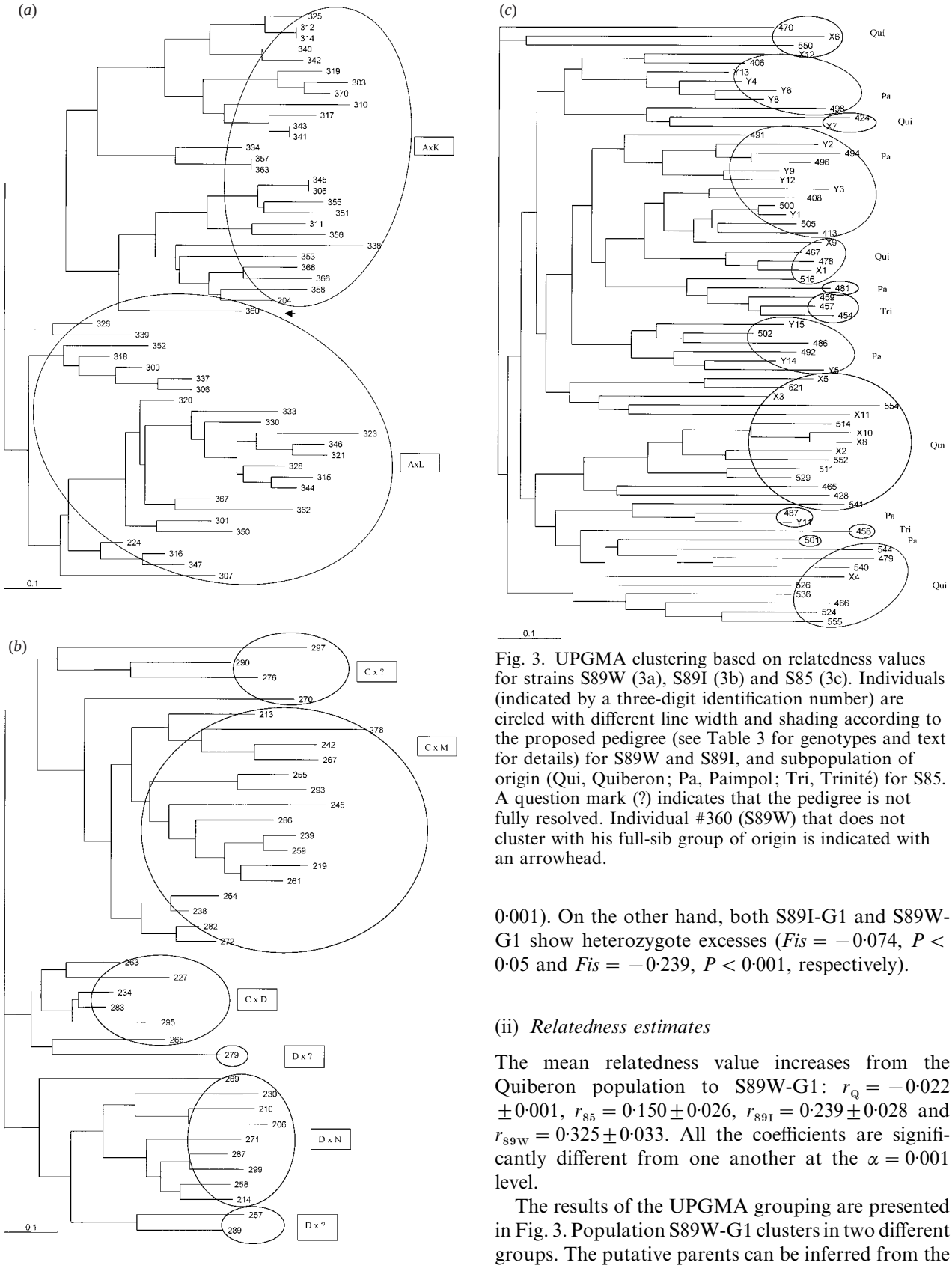

Fig. 3. UPGMA clustering based on relatedness values for strains S89W (3a), S89I (3b) and S85 (3c). Individuals (indicated by a three-digit identification number) are circled with different line width and shading according to the proposed pedigree (see Table 3 for genotypes and text for details) for S89W and S89I, and subpopulation of origin (Qui, Quiberon; Pa, Paimpol; Tri, Trinité) for S85. A question mark (?) indicates that the pedigree is not fully resolved. Individual #360 (S89W) that does not cluster with his full-sib group of origin is indicated with an arrowhead.

ficiency ($Fis = 0{\cdot}082$, $P < 0{\cdot}001$) which seems to be due mainly to *Oedu*H15. S85-G2 shows a slightly higher heterozygote deficiency ($Fis = 0{\cdot}104$, $P < 0{\cdot}001$). On the other hand, both S89I-G1 and S89W-G1 show heterozygote excesses ($Fis = -0{\cdot}074$, $P < 0{\cdot}05$ and $Fis = -0{\cdot}239$, $P < 0{\cdot}001$, respectively).

(ii) *Relatedness estimates*

The mean relatedness value increases from the Quiberon population to S89W-G1: $r_{Q} = -0{\cdot}022 \pm 0{\cdot}001$, $r_{85} = 0{\cdot}150 \pm 0{\cdot}026$, $r_{89I} = 0{\cdot}239 \pm 0{\cdot}028$ and $r_{89W} = 0{\cdot}325 \pm 0{\cdot}033$. All the coefficients are significantly different from one another at the $\alpha = 0{\cdot}001$ level.

The results of the UPGMA grouping are presented in Fig. 3. Population S89W-G1 clusters in two different groups. The putative parents can be inferred from the individual genotypes for each of these groups, and are given in Table 3. A very likely hypothesis is that S89W-G1 is constituted by two biparental half-sib families that share one parent. *O. edulis* is a brooding

Table 3. *Probable genotypes for the identified parents at the origin of S891 and S89W*

Pop.	Family	Parent1	Parent2	*Oedu*J12	*Oedu*U2	*Oedu*H15	*Oedu*O9	*Oedu*T5
S89W		A		230/250	182/206	193/213	160/168	106/134
			K	230/258	164/176	173/189	166/174	142/146
			L	232/242	162/176	173/201	160/162	124/128
S89I	Fam1	C		224/248	168/212	189/213	154/162	106/124
			M	216/224	162/204	173/221	156/170	132/–
	Fam2	D		238/244	146/148	173/201	154/156	124/140
			N	242/244	146/162	173/181	152/160	106/122

oyster: the male gametes are released in the seawater and filtered by the female, and fertilisation takes place inside the female's shell. Given this reproductive biology, the results suggest that the population issued from a single spawning event involving one female and two different males. The proposed pedigree is listed for each individual on the UPGMA tree. The clustering reflects perfectly the family structure, with only one misclassified individual (#360).

S89I-G1 also clusters in three main groups. By analyzing individual genotypes, and based on shared alleles (data not shown), two of these clusters can be attributed to two independent half-sib families; in each of these families, the majority of the individuals can be attributed to a single full-sib group. For each of these families, the genotypes of the common parents and of one other parent are given in Table 3 (respectively C and M, and D and N, for each family group). At least two other parents must have contributed to each of the families to account for the existing allelic richness, but their genotypes remain uncertain due to the small sample size. On the basis of these reconstructed genotypes, the individuals that form the intermediate cluster can be attributed to a single full-sib family CxD. Therefore two individuals drawn at random from S89I-G1 can be either full-sib, half-sib or unrelated. The proposed pedigrees are listed for individuals on the UPGMA tree. Again, the clustering is a good image of the family structure. This family structure gives us a first estimate of the number of parents of at least 8.

The situation is different for S85-G2: due to subsequent mixing of genotypes between the first and second generation, and possibly to a greater number of founders (see below), it is difficult to identify groups of related individuals or family structure. On the UPGMA tree we listed the subpopulation (Quiberon, Paimpol, Trinité) to which each individual belongs. These subpopulations share the same grandparents (G0), but G1 individuals were not mixed for the production of the G2 (see Fig. 1). The UPGMA clustering is not very informative, although the ends of the nodes tend to cluster individuals from the same subpopulation. Because of a small number of individuals in the Trinité subpopulations, subsequent relatedness analyses were conducted only for the Quiberon and Paimpol subpopulations (respectively 35 and 28 individuals), still using allelic frequencies in the wild Quiberon sample as a baseline. The mean relatedness values within the Paimpol subpopulation ($r = 0{\cdot}374 \pm 0{\cdot}010$) is significantly higher than within the Quiberon subpopulation ($r = 0{\cdot}134 \pm 0{\cdot}008$, $P < 0{\cdot}001$).

(iii) *Estimation of the number of founders at the origin of the selected populations*

Estimates of the number of founders are given in Table 4. Although we already determined that S89W-G1 derived from exactly 3 parents, we made the calculation for this population as a way of testing for the accuracy of our methods.

In the case of S89I-G1 and S89W-G1, the estimates of the effective number of parents are consistent for the different loci and the different methods, even though the heterozygote excess methods gives the lower estimates. The mean ($\pm$standard deviation) of all estimates is $8{\cdot}3 \pm 1{\cdot}4$ for S89I-G1 and $4{\cdot}2 \pm 1{\cdot}6$ for S89W-G1. These values are concordant with the estimate that we previously derived from the pedigree reconstruction attempt. A narrower confidence interval was found for the heterozygote excess method, but for S89W-G1 the confidence interval does not include 3. With this exception, each confidence interval includes the estimates derived from the pedigree reconstruction.

No pedigree information could be used for S85-G2 and the heterozygote excess method is not applicable. The other two methods give concordant multilocus estimates; the mean of all estimates is $18{\cdot}7 \pm 2{\cdot}2$. Results are consistent across loci with the exception of *Oedu*T5, which gives a higher estimate (27·7) and a broader confidence interval (17·4–64·8) through the reduction in the number of alleles method.

Effective size in the wild Quiberon population was estimated through the temporal variance in allelic frequencies method, and we found $N_e = 464$, with a negative value as the upper bound of the confidence

Table 4. *Estimated number of founders for each selected population*

Pop.	Locus	Pedigree	Temporal variance in allelic frequencies		Heterozygote excess		Reduction in the no. of alleles
			F	N_k [95% CI]	D	N_{eb} [95% CI]	N_i [95% CI]
S85	*Oedu*J12		0·058	24·6	na	na	16·3 [10–30·7]
	*Oedu*U2		0·066	20·7	–	–	14·5 [8·8–24·3]
	*Oedu*H15		0·073	18·3	–	–	13·5 [7·6–31·5]
	*Oedu*O9		0·065	20·9	–	–	11·9 [6·8–23·2]
	*Oedu*T5		0·072	18·4	–	–	27·7 [17·4–64·8]
	Multilocus	**na**	0·067	**20·4 [12·8–31·9]**	na	**na**	**16·8 [10·1–34·9]**
S89I	*Oedu*J12		0·096	6·9	na	na	8·2 [4·5–12·6]
	*Oedu*U2		0·090	7·5	–	–	10·0 [5·8–15]
	*Oedu*H15		0·071	10·9	–	–	11·0 [6–22·8]
	*Oedu*O9		0·115	5·5	–	–	8·6 [4·5–14·5]
	*Oedu*T5		0·099	6·6	–	–	10·0 [5·1–19·6]
	Multilocus	**⩾ 8**	0·094	**7·1 [4·4–10·9]**	0·085	**6·3 [3·8–9·2]**	**9·6 [5·2–16·9]**
S89W	*Oedu*J12		0·291	1·8	na	na	3·9 [2·0–5·7]
	*Oedu*U2		0·314	1·7	–	–	4·6 [2·1–6·3]
	*Oedu*H15		0·157	3·7	–	–	4·6 [2·1–7·8]
	*Oedu*O9		0·178	3·2	–	–	3·9 [2·1–5·8]
	*Oedu*T5		0·065	11·3	–	–	6·2 [2·2–9·5]
	Multilocus	**3**	0·201	**2·8 [1·3–5·1]**	0·248	**2·4 [2·3–2·5]**	**4·6 [2·1–7·0]**

N_k, N_{eb} and N_i are estimates of effective size N_e according to the following methods respectively: temporal variance in allele frequencies (Waples, 1989), heterozygote excess (Pudovkin *et al*., 1996) and reduction in the number of alleles (this study). *F* is the estimate of the variance in allelic frequencies between the selected population and the control population, D is the mean heterozygote excess $(H_0 - H_{exp})/H_{exp}$.
CI, confidence interval; na, not applicable. See text for details of calculation.

interval, which was interpreted as infinity according to Waples (1989).

4. Discussion

(i) *Usefulness of microsatellites for genetic variability management in the selected populations*

This study demonstrates the usefulness of highly polymorphic markers such as microsatellites for answering the numerous questions related to a selection programme. With only five markers we were able to assess the genetic variability within the selected populations, determine the number of families and reconstruct pedigrees *a posteriori*. However, a larger number of markers would be necessary to identify with more accuracy the parents of S89I-G1 and the number of families in S85-G2.

There is evidence, especially from the number of alleles, that the selected populations have been subjected to bottlenecks. As predicted (Nei *et al*., 1975), bottlenecks have a stronger and more immediate effect on allelic diversity than on heterozygosity. Whereas heterozygosity is still high, the number of alleles in the selected populations is significantly reduced compare with the control population, and this is due in large part to the loss of rare alleles. All selected populations showed departure from expected Hardy–Weinberg equilibrium, but in two opposite directions. Heterozygote excess was seen in both S89I-G1 and S89W-G1, and can be correlated with sampling bias of alleles because of a very small number of founders and non-random mating. A temporary increase in heterozygosity level has been shown to occur in the generation immediately following a bottleneck, due to allelic evening out (Maruyama & Fuerst, 1985; Hedgecock & Sly, 1990; Leberg, 1992), as seen in S89W-G1. On the contrary, S85-G2, which may have undergone two successive bottlenecks, showed a deficit in heterozygotes over all loci compared with Hardy–Weinberg expectations. Pervasive heterozygote deficiencies at marker loci (mostly allozymes) have often been reported for marine bivalve populations (see for instance Gaffney, 1994), but the origin of these deficiencies is still under debate (technical artefacts, functionally null alleles, selection, aneuploidy, Wahlund effect, etc.). In S85-G2, only two microsatellite loci taken individually showed heterozygote deficiencies (*Oedu*H15 and *Oedu*J12). Different hypotheses, which are not mutually exclusive, can be put forward to account for the observed heterozygote deficiencies: (a) the population is in equilibrium, and deficiencies at *Oedu*J12 and *Oedu*H15 are due to null alleles, which could also explain the heterozygote deficiency in the control

population at *Oedu*H15; (b) the population is substructured, since the different G1 subpopulations were kept separate for the production of the G2 (see Fig. 1), creating the possibility of differentiation (through founder effect and drift) between the G2 subpopulations, and hence a possible ground for Wahlund effect; and (c) inbreeding level is increasing.

(ii) *Estimation of effective size*

The other aim of this study was to obtain and compare different estimates of the effective size of bivalve hatchery populations, as a predictor of the rate of loss of neutral variation. Our estimates indicated that the effective size of the selected *O. edulis* populations is very low, between 3 and 20 at the most.

(a) *Assumption of the models*

The different methods for estimating the effective size – temporal variance in allelic frequencies (1), heterozygote excess (2) and reduction in the number of alleles (3) – are based on a number of assumptions: (a) no (or negligible) mutation (1, 3), (b) no immigration from neighbouring subpopulations (1 to 3), (c) neutrality of alleles studied (1 to 3), (d) random mating (2), and (e) random sampling in the population (1 to 3). In addition, methods 1 and 3 require data for at least two samples drawn at different times whereas method 2 requires data from a single cohort only. Assumptions (b) and (c) are obviously fulfilled in our study.

Mutation. Mutation rate for a microsatellite sequence is expected to be higher than for other random or coding sequences (Jarne & Lagoda, 1996). However, no data are available on the mutation rate of the loci used in this study. Some alleles are found at a relatively high frequency in a selected population and are not found in the control population. This is the case for instance for allele *Oedu*U2-212 (frequency of 0·211 in population S89I-G1), or for allele *Oedu*T5-146 (frequency of 0·147 in S89W-G1). This could be due to a mutation in one of the founding parents; alternatively, this could be explained by drift in the selected populations and sampling bias in the control population. In S89W-G1, where all the putative parents' genotypes could be inferred, no alleles were seen in the progeny that could not be traced back to one of the parents. We will assume that in our study (limited sample size, limited number of generations), microsatellite alleles are not affected by mutation.

Random mating. Pudovkin *et al.* (1996) noted that the heterozygote excess that would be seen in the progeny of a small number of parents could be reduced or even cancelled if the union of gametes is not at random, because of the Wahlund effect. There is evidence from the pedigree reconstruction that S89I-G1 and S89W-G1 were not produced by random mating of the putative parents. We expect, then, that the heterozygote excess method will underestimate the effective number of parents compared with the direct estimation obtained through pedigree reconstruction (especially for S89W-G1). This could account for the lower estimates obtained with this method.

Random sampling. Method 1 depends on random sampling, and therefore on reliable estimates of allelic frequencies in a closed population at different points in time. Here, the wild Quiberon Bay population departs from an 'ideal' control in that it was not sampled at $t = 0$ but was sampled at time t, (i.e. at the same time as the selected populations). With a small t, the values for N_k given in Table 4 could be interpreted as the mean number of parents for both the selected population and the control, across t generations (Hedgecock *et al.*, 1992). In our case, since the estimated effective size for the control population is 1 to 2 order of magnitude higher than that of the selected populations, drift can be considered to be limited in the control population, and N_k can be taken as a reasonable estimate of the effective size of the selected populations.

(b) *Comparison of the methods*

N_e values are traditionally estimated from the increase in the inbreeding rate of population (N_{eI}) and from the variance of allele frequencies (N_{eV}). In this paper, we used both methods. N_{eV} was estimated using the temporal variance method (Pudovkin *et al.*, 1996), the method based on the reduction of the number of alleles and the method based on departure from Hardy–Weinberg equilibrium. The estimate of N_e based on the calculation of genetic relatedness can be assimilated to N_{eI}. Although in a population of constant size the two parameters N_{eV} and N_{eI} are equivalent, in a population that changes size very frequently N_{eI} and N_{eV} can differ greatly. In this paper, the different estimates of N_{eV} are very congruent, and no discrepancy was found between N_{eV} and N_{eI} when estimates have been obtained (S89I and S89W). This might be attributed to the small number of bottlenecks (1) experienced in each of these two populations.

All three numerical methods, as well as pedigree reconstruction when applicable, give estimates of effective population size of the same order of magnitude. The confidence interval is larger for the

reduction in the number of alleles method and narrower for the heterozygote excess method. As it stands, the method based on the reduction in the number of alleles can be used instantly to estimate the effective size of cohort produced from a subsample of a larger population. It is also straightforward and easy to compute. Non-random mating in the subsample should have little effect, but loss of genetic diversity through drift will. It might be useful, then, to estimate N_e by comparing two samples that are close in time, whereas the temporal method relies on the assumption that all the variance in allelic frequencies is due to drift, and an important number of generations between samples is generally needed for a precise estimate of N_e (Waples, 1989). Possible extensions of our programme could include the effect of sampling variation among the individuals and loci, of allele-frequency profiles in both samples, of heterozygosity level, and the reliability and power of this estimate compared with the other methods. For instance, for the same total number of alleles, the profile of the curve $N_a = f(N_i)$, as defined in Section 2, will be very different when all the alleles have equal frequency or when there is a large number of rare alleles. The control population has more rare alleles at *Oedu*T5 than at the other loci and this could explain the discrepancy found at this locus in the estimate of N_e in S85-G2 and S89W-G1.

(c) *Low N_e:N ratios*

The effective sizes of the selected populations were found to be very small: between 15 and 20 for S85, around 8–10 for S89I and 3 for S89W. This is much smaller than the number of animals that were actually reared together in the tank during each mass spawning experiment (around 100). Such strong founder effects (and correlated important genetic changes) have often been reported in hatchery populations of *O. edulis* and other bivalves. For instance, Saavedra & Guerra (1996) collected over 5 days the spawn from 120 flat oysters, and estimated the effective number of parents to be around 3·5 (see also Alvarez *et al.*, 1989; Hedgecock *et al.*, 1992; Gaffney *et al.*, 1992). Among marine bivalves, however, the flat oyster is particularly sensitive to founder effects, and an important reduction in effective size can be observed as soon as the first generation. Saavedra & Guerra (1996) suggest that this might be due to the flat oyster's life cycle, and to the fact that, in contrast to the cupped oyster *Crassostrea gigas* for instance, there is no synchronicity in sexual maturation and gamete release.

Similarly, the estimate of N_e in the control population is far smaller than might be assumed from the abundance of flat oysters, even in populations that might be depressed by disease. In France, the remaining stock is estimated to be in the thousands, especially in Brittany (H. Grizel, personal communication). The order of magnitude of the discrepancy between census size and calculated N_e (ratio $N_e:N$ lower than 10^{-3}) is similar to that previously reported by Saavedra (1997) for a flat oyster population from Spain. Such a discrepancy has already been noted for other marine bivalves (Hedgecock *et al.*, 1992; Hedgecock, 1994). A proposed hypothesis was that bivalves have a large variance in reproductive success, V_k, reducing effective population numbers to a small fraction of breeding numbers (see also Li & Hedgecock, 1998).

(iii) *Consequences for the selection programme*

Small effective size and high relatedness have numerous consequences for the selection programme that deal mainly with inbreeding management. Biparental crosses had been made using some of these individuals in 1995, before these results were obtained. The pedigree information is now available for both S891 and S89W. All the individuals within S89W-G1 are at least half-sibs, and any biparental cross within this population will lead to a family with an inbreeding coefficient f of at least 0·125 (Falconer & McKay, 1996). When crossing the S89I-G1 to itself, this will happen in approximately half the cases. We can expect a decrease in performance in the offspring as soon as the G2, as regards both survival (Ibarra *et al.*, 1995) and growth (Beattie *et al.*, 1987). The offspring performances are expected to vary depending on the degree of kinship of the parents, and Naciri-Graven *et al.* (2000) have observed negative correlations between growth performances of the offspring and relatedness of their parents.

Biparental crosses will allow precise control of genealogy and inbreeding level. In order to avoid the increase in inbreeding level, it is probably wise to stop crosses within S89W and S85 Paimpol subpopulation. On the other hand, it is possible to drive the selected populations towards inbred lines and select for a better hybrid value, as there seems to be a high potential for heterosis in marine bivalves (Hedgecock *et al.*, 1995, 1996), and particularly in these flat oyster selected lines (Baud *et al.*, 1997; Naciri-Graven *et al.*, 1999, 2000).

We wish to thank all the staff members of the various laboratories for their help in maintenance and sampling, especially C. Ledu, P. Phélipot, S. Heurtebise and A.-G. Martin. We also thank D. Hedgecock for helpful comments and proof-reading of the manuscript and, together with V. Chow, for providing the program EPS for N_e calculation. The manuscript has been improved by helpful remarks from three anonymous reviewers. This work was supported by grants from the French State, and the Brittany, the Pays de la Loire and the Poitou-Charentes Regions.

Appendix. *Allelic frequencies in* O. edulis *selected populations* (*S85−G2*, *S89I−G1 and S89W−G1*) *and the Quiberon control population. Allele name indicates allele size in bp.*

Locus	Allele	S85	S89I	S89W	Quiberon
*Oedu*J12	216	0.143	0.068		0.082
	218	0.007			
	224	0.093	0.300		0.194
	226				0.041
	228	0.007			0.020
	230			0.353	0.010
	232	0.043		0.147	0.061
	234	0.050	0.014		0.092
	236		0.014		0.020
	238	0.100	0.054		0.071
	240	0.143			0.031
	242	0.036	0.054	0.118	0.092
	244	0.050	0.284		0.031
	246	0.236			0.041
	248	0.036	0.203		0.020
	250	0.043		0.177	0.051
	252				0.041
	254	0.014	0.014		0.010
	256				0.041
	258			0.206	0.020
	262				0.010
	264				0.010
	272				0.010
*Oedu*U2	146		0.224	0.020	
	148		0.013		
	152				0.010
	154	0.056			
	156	0.229			0.167
	158	0.188			0.104
	160	0.007	0.026		0.042
	162	0.007	0.105	0.100	0.031
	164	0.007		0.220	0.052
	166				0.010
	168		0.118		0.042
	170				0.052
	172	0.007			0.042
	174	0.049	0.040		0.063
	176			0.160	0.042
	178	0.278			0.063
	180	0.069			0.073
	182		0.040	0.210	0.021
	184				0.010
	186		0.013		0.031
	188				0.031
	190	0.007	0.013		0.021
	192	0.014			
	196	0.076			0.042
	198	0.007			0.010
	200				0.010
	204		0.197		
	206			0.290	0.010
	210				0.010
	212		0.211		
	222			0.010	
*Oedu*O9	140	0.021			0.010
	142				0.010
	144				0.020
*Oedu*O9 (*cont.*)	146	0.036			
	148				0.031
	150	0.121			0.041
	152	0.007	0.068		0.071
	154	0.064	0.297		0.041
	156	0.043	0.338		0.082
	158				0.061
	160	0.114	0.068	0.353	0.102
	162	0.100	0.122	0.137	0.133
	164	0.414	0.014		0.163
	166		0.014	0.049	0.071
	168	0.050	0.014	0.265	0.071
	170	0.027	0.068		0.020
	172				0.020
	174			0.196	0.041
	176				0.010
*Oedu*H15	165				0.012
	173	0.293	0.389	0.212	0.267
	177				0.023
	181	0.129	0.042		
	185	0.007			0.058
	189		0.208	0.096	0.093
	193			0.202	0.023
	197				0.047
	201	0.043	0.014	0.183	0.116
	205	0.036			0.070
	209	0.014			0.047
	213	0.229	0.153	0.308	0.070
	217	0.157	0.014		0.035
	221	0.079	0.111		0.105
	225	0.014	0.042		0.035
	229		0.028		
*Oedu*T5	106	0.357	0.237	0.294	0.213
	112	0.007			
	118	0.036	0.013		0.149
	120		0.013		
	122		0.079		
	124	0.043	0.263	0.098	0.106
	126				0.011
	128	0.014		0.177	0.085
	130	0.014			0.021
	132	0.007	0.211		0.075
	134	0.064	0.013	0.206	0.107
	136	0.057			0.021
	138	0.114			0.053
	140	0.207	0.158		0.011
	142	0.029		0.078	0.011
	144				0.011
	146			0.147	
	148				0.021
	150	0.014	0.013		0.021
	152	0.021			0.032
	154	0.014			0.021
	158				0.011
	162				0.011
	166				0.011

References

Alvarez, G., Zapata, C., Amaro, R. & Guerra, A. (1989). Multilocus heterozygosity at protein loci and fitness in the European oyster, *Ostrea edulis* L. *Heredity* **63**, 359–372.

Backus, V. L., Bryant, E. H., Hughes, C. R. & Meffert, L. M. (1995). Effect of migration or inbreeding followed by selection on low-founder-number populations: implications for captive breeding programs. *Conservation Biology* **9**, 1216–1224.

Baud, J. P., Gérard, A. & Naciri-Graven, Y. (1997). Comparative growth of *Bonamia ostreae* resistant and wild flat oysters, *Ostrea edulis*, in an intensive upwelling system. I. First year of experiment. *Marine Biology* **130**, 71–79.

Beattie, J. H., Purdue, J., Hershberger, W. & Chew, K. (1987). Effect of inbreeding on growth in the Pacific oyster *Crassostrea gigas*. *Journal of Shellfish Research* **6**, 25–28.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. & Bonhomme, F. (1996). Genetix, logiciel sous Windows pour la génétique des populations. http://www.univ-montp2.fr/~genetix/genetix.htm

Blouin, M. S., Parsons, M., Lacaille, V. & Lotz, S. (1996). Use of microsatellite loci to classify individuals by relatedness. *Molecular Ecology* **5**, 393–401.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. Minneapolis, USA: Alpha Editions.

Falconer, D. S. & McKay, T. F. C. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow, UK: Longman.

Gaffney, P. M. (1994). Heterosis and heterozygote deficiencies in marine bivalves: more light? In *Genetics and Evolution of Aquatic Organisms* (ed. A. R. Beaumont), pp. 146–152. London: Chapman & Hall.

Gaffney, P. M., Davis, C. V. & Hawes, R. O. (1992). Assessment of drift and selection in hatchery populations of oysters (*Crassostrea virginica*). *Aquaculture* **105**, 1–20.

Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms? In *Genetics and Evolution of Aquatic Organisms* (ed. A. R. Beaumont), pp. 122–134. London: Chapman & Hall.

Hedgecock, D. & Sly, F. (1990). Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster, *Crassostrea gigas*. *Aquaculture* **88**, 21–38.

Hedgecock, D., Chow, V. & Waples, R. S. (1992). Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies. *Aquaculture* **108**, 215–232.

Hedgecock, D., McGoldrick, D. J. & Bayne, B. L. (1995). Hybrid vigor in Pacific oysters: an experimental approach using crosses among inbred lines. *Aquaculture* **137**, 285–298.

Hedgecock, D., McGoldrick, D. J., Manahan, D. T., Vavra, J., Appelmans, N., *et al.* (1996). Quantitative and molecular genetic analysis of heterosis in bivalve molluscs. *Journal of Experimental Marine Biology and Ecology* **203**, 49–59.

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**, 229–239.

Ibarra, A. M., Cruz, P. & Romero, B. A. (1995). Effect of inbreeding on growth and survival of self fertilized catarina scallop larvae, $\pm$. *Aquaculture* **134**, 37–47.

Jarne, P. & Lagoda, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends in Evolution and Ecology* **11**, 424–429.

Launey, S. (1998). Marqueurs microsatellites chez l'huître plate *Ostrea edulis* (L.): caractérisation et application à un programme de sélection pour une résistance au parasite *Bonamia ostreae* et à l'étude de populations naturelles. Doctoral thesis, l'INA-PG, Paris.

Leberg, P. (1992). Effects of populations bottlenecks on genetic diversity as measured by allozyme electrophoresis. *Evolution* **46**, 477–494.

Li, G. & Hedgecock, D. (1998). Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Canadian Jounal of Fisheries and Aquatic Sciences* **55**, 1025–1033.

Maruyama, T. & Fuerst, P. A. (1985). Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* **111**, 675–690.

Naciri-Graven, Y., Martin, A. G., Baud, J. P., Renault, T. & Gérard, A. (1998). Selecting flat oyster *Ostrea edulis* for survival when infected with the parasite *Bonamia ostreae*. *Journal of Experimental Marine Biology and Ecology* **224**, 91–107.

Naciri-Graven, Y., Haure, J., Gérard, A. & Baud, J. P. (1999). Comparative growth of *Bonamia ostreae* resistant and wild flat oyster *Ostrea edulis* in an intensive system. II. Second year of experiment. *Aquaculture* **171**, 195–208.

Naciri-Graven, Y., Launey, S., Lebayon, N. & Baud, J. P. (2000). Influence of parentage upon growth in *Ostrea edulis*: evidence for inbreeding depression. *Genetical Research* **76**, 159–168.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590.

Nei, M., Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**, 1–10.

Pollack, E. (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics* **104**, 531–548.

Pudovkin, A. I., Zaykin, D. V. & Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote-excess in progeny. *Genetics* **144**, 383–387.

Queller, D. C. & Goodnight, K. F. (1989). Estimating relatedness using genetic markers. *Evolution* **3**, 258–275.

Queller, D. C., Strassman, J. E. & Hughes, C. R. (1993). Microsatellites and kinship. *Trends in Ecology and Evolution* **8**, 285–288.

Saavedra, C. (1997). Low effective size in hatchery populations of the european oyster (*Ostrea edulis*): implications for the management of genetic resources. *Journal of Shellfish Research* **16**, 441–446.

Saavedra, C. & Guerra, A. (1996). Allozyme heterozygosity, founder effect and fitness traits in a cultivated population of the European oyster, *Ostrea edulis*. *Aquaculture* **139**, 203–224.

Sbordoni, V., De Matthaeis, E., Cobolli Sbordoni, M., La Rosa, G. & Mattoccia, M. (1986). Bottleneck effects and the depression of genetic variability in hatchery stocks of *Penaeus japonicus* (Crustacea, Decapoda). *Aquaculture* **57**, 239–251.

Waples, R. S. (1989). A generalized approach of estimating effective population size from temporal changes in allelic frequency. *Genetics* **121**, 379–391.

Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.