



Using molecular markers for pedigree reconstruction of the greater amberjack (*Seriola dumerili*) in the absence of parental information

D. Rodríguez-Barreto*, S. Consuegra†, S. Jerez‡, J. R. Cejas‡, V. Martín‡ and A. Lorenzo*

*Departamento de Biología Animal, Facultad de Biología, Universidad de La Laguna, 38206, Tenerife, Spain. †Institute of Biological, Environmental & Rural Sciences, Aberystwyth University, Aberystwyth, SY23 3DA, UK. ‡Centro Oceanográfico de Canarias, Instituto Español de Oceanografía, Dársena Pesquera, 38180, Sta. Cruz de Tenerife, Spain.

Summary

Ensuring appropriate levels of genetic diversity in captive populations is essential to avoid inbreeding and loss of rare alleles by genetic drift. Pedigree reconstruction and parentage analysis in the absence of parental genotypes can be a challenging task that relies in the assignment of sibship relationships among the offspring. Here, we used eight highly variable microsatellite markers and three different assignment methods to reconstruct the most likely genotypes of a parental group of wild *Seriola dumerili* fish based on the genotypes of six cohorts of their offspring, to assess their relative contributions to the offspring. We found that a combination of the four most variable microsatellites was enough to identify the number of parents and their contribution to the offspring, suggesting that the variability of the markers can be more critical than the number of markers. Estimated effective population sizes were lower than the number of breeders and variable among years. The results suggest unequal parental contribution that should be accounted for breeding programs in the future.

Keywords aquaculture, captive breeding, inbreeding, microsatellites, parentage

Preserving genetic diversity within populations is needed for maintaining their evolutionary potential and fitness (Reed & Frankham 2003). The need for maintaining genetic diversity in cultured populations is also widely recognised (Gjerde *et al.* 2004; McLean *et al.* 2008). However, in fish species with mass communal spawning, unequal parental contribution (due to sperm competition or mate choice for example) and differential offspring survival at early stages may remain largely unnoticed unless an accurate pedigree reconstruction can be obtained (e.g., Frost *et al.* 2006). Important reductions in genetic diversity can happen even after a single generation in captivity (Porta *et al.* 2007) due to small effective population size resulting from a small number of parents and variance in reproductive success (Blonk *et al.* 2009). Molecular-based parentage analysis can be used to prevent inbreeding and maintain genetic diversity in cultured stocks (Sekino *et al.* 2003), although pedigree reconstruction in the absence of parental information remains challenging. New Bayesian or maximum-likelihood methods for parental and kinship assignment can

be useful approaches for parentage reconstruction when parental data are unavailable (Sefc & Koblmüller 2009; Massah *et al.* 2010), opening a range of possibilities for broodstock husbandry in species new to aquaculture (Garber *et al.* 2010).

We used a set of microsatellite markers and three different methods for parentage assignment and pedigree reconstruction to determine the number of parents and their relative contribution to six cohorts of a captive population of the greater amberjack (*Seriola dumerili*), a carangid fish of increasing interest for aquaculture (Nakada 2000), for which reproduction in captivity has proved particularly difficult (Jerez *et al.* 2006). *S. dumerili* reaches maturity at 3–5 years of age and reproduces several times during a spawning season with one seasonal peak per year (spring–summer). Females live longer than do males with a maximum lifespan around 17 years. Our aim was to reconstruct the contribution to the offspring of different *S. dumerili* parents reproducing by mass spawning in captivity over a number of years in order to estimate effective population size and the potential for inbreeding.

Fin clips were collected and preserved in ethanol at 4 °C from 94 *Seriola dumerili* 3–8 years of age resulting from the mass spawning of 11 wild breeders of unknown sex captured in 1996. The samples represented six different cohorts divided into four groups: 23 individuals from 2002, 23 individuals from 2003 to 2005 (pooled), 23 individuals

Address for correspondence

S. Consuegra, Institute of Biological, Environmental & Rural Sciences, Aberystwyth University, Edward Llwyd Building, Penglais Campus, Ceredigion, SY23 3DA, UK.
E-mail: skc@aber.ac.uk

Accepted for publication 07 February 2013

from 2006 and 25 individuals from 2007. Eggs from each spawning year were incubated separately, and juveniles from each cohort were raised in separate tanks, apart from those from 2003 to 2005, which were kept in the same tank before being individually tagged. Genomic DNA was extracted using the Wizard[®]SV96 DNA Purification Kit (Promega) and quantified using the NanoDrop1000 v.3.7 Spectrophotometer (Thermo Fisher Scientific). Nine microsatellite loci were amplified in one single multiplex (Sdu 1, Sdu 3, Sdu4, Sdu5, Sdu 6, Sdu7, Sdu22, Sdu31, Sdu 44; Renshaw *et al.* 2006, 2007) using the Qiagen Multiplex PCR Kit following the manufacturer's protocol. Amplification conditions consisted of 15 min at 95 °C followed by eight cycles of touchdown PCR starting at 64 °C and descending in 2 °C intervals to 56 °C and 24 cycles with an annealing temperature of 56 °C. Microsatellites were resolved on an ABI3130xl Genetic Analyser (Applied Biosystems), and fragment length was determined using GENEMAPPER v4.0. Sdu 31 and Sdu 7 displayed potential competition between the primers, and Sdu 31 was excluded from the analysis. To identify the most variable markers, genetic variability was estimated as the mean number of alleles (A) per locus, observed heterozygosity (Ho) and expected heterozygosity (He) using GENETIX v 4.03. Exact tests for Hardy–Weinberg equilibrium were carried out using GENEPOP v 4.0 and corrected by the Bonferroni procedure (Rice 1989). This allowed us to identify potential markers with null alleles. Ho per microsatellite ranged between 0.5 and 0.96, and A per locus ranged from 4 to 7 (Table S1). Four of the microsatellites (sdu6, sdu22, sdu44 and sdu1) deviated significantly from Hardy–Weinberg expectations after Bonferroni correction, albeit only in one group (Table S1).

For pedigree reconstruction, analyses were run initially using eight microsatellites. Combinations of the six and four most variable microsatellites were then assessed to identify the ideal set of markers (Sefc & Koblmuller 2009). We used three different approaches for parentage assignment and pedigree reconstruction; only results supported by the three methods were taken into account to ensure the highest accuracy of the results. A likelihood method, COLONY v 2.0 (Jones & Wang 2010), was initially used for pedigree reconstruction and for inferring parental genotypes. COLONY simultaneously estimates sibling relationships and parentage using multilocus genotype data. We first ran COLONY for independent cohorts, assuming random mating, no inbreeding, no typing error and mutation rate = 0. Ne was calculated by one run of medium length (heterozygote excess method). Further confirmation of parental groups was obtained by using PEDIGREE v 2.0 (<http://herbinger.biology.dal.ca:5080/Pedigree>), a likelihood pairwise score method, which allows pedigree reconstruction when parental information is not available, inferring sibship groups and the unknown parental genotypes (Smith *et al.* 2001; Butler *et al.* 2004). To assess the reliability of COLONY in retrieving parental genotypes, we used the simulation option in PAPA v 2.0 (Duchesne *et al.*

2002) to simulate 10 populations of 10 parents and 94 offspring based on the allelic distribution of our data for eight microsatellites with the following parameters: sexed parents, pre-parental assignment and uniform error (0.2). The pseudo-offspring simulated by PAPA assuming unknown parents was then run in COLONY, which was found to overestimate the number of breeders (an average of 28 parents per simulation were recovered), as had been seen before (Sefc & Koblmuller 2009). However, in all simulations, at least eight of the 10 parents simulated by PAPA were identified with a similarity of at least 87% and the two remaining with an allele similarity of at least 70%. Based on these simulation results, we considered for further analyses only those parental genotypes which appeared repeatedly in more than one cohort and shared at least 70% of the alleles. Seven candidate parents were inferred by COLONY, confirmed by PEDIGREE and then used in a likelihood-based assignment software, CERVUS v 3.0 (Kalinowski *et al.* 2007), to identify potential parental pairs (parent pairs with unknown sex, 10 000 replication cycles, proportion of parents and corresponding loci genotyped set to 1; typing error = 1%). The seven candidate parents allowed us to assign 66% of the offspring to a pair with at least 80% confidence level. Two of the tested genotypes did not contribute to the offspring. COLONY was run again for the unassigned offspring (full-sibling groups detailed in Fig. S1), this time incorporating the parental genotypes supported by PEDIGREE and CERVUS. With the new information provided, more parental genotypes could be inferred and were then added for parentage analysis in CERVUS. In total, 10 genotypes allowed assigning 99% of offspring with at least 80% confidence level. Pairing of the individuals in CERVUS identified two sex groups (A and B) of three and seven individuals (genotypes) respectively (Table 1) with unequal parental contribution among cohorts (Table 2) that could be due to mate competition or due to differential family survival during the early stages of development. The combination of the three approaches allowed for the identification of differences in parental contribution; these are common in species with mass spawning (Frost *et al.* 2006) and could result in reduced effective population size, inbreeding and loss of genetic diversity (Jackson *et al.* 2003).

The results using four microsatellites with the largest number of alleles ($A = 7$) and high heterozygosity ($Ho > 0.8$) (sdu4, sdu5, sdu6, sdu22; Table S1), chosen as the best combination, were consistent with those using all of the markers. The most likely number of parents was also 10; 87 individuals from the offspring were in the same full-sib groups, and the remaining eight were assigned to a different full-sib group. Eight of the parental genotypes identified with four and eight microsatellites were identical for the common markers, whereas the remaining ones were similar in 62.5% and 25% of the alleles (Table 1). Thus, despite minor differences, the four more variable microsatellites would be enough to identify the number of parents contributing to the offspring and their relative contribution

Table 1 Parental genotypes selected from the four and eight microsatellite pedigree reconstruction.

Parental ID	Sdu6		Sdu7		Sdu3		Sdu22		Sdu44		Sdu1		Sdu5		Sdu4	
	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b
Sex group 1																
A1	239	255	–	–	–	–	320	326	–	–	–	–	206	224	336	340
A1	239	255	355	355	215	221	320	326	120	124	321	321	206	224	336	340
A2	243	251	–	–	–	–	323	344	–	–	–	–	227	239	332	332
A2	243	251	358	358	215	221	323	344	120	122	325	341	227	239	332	332
A3	235	243	–	–	–	–	311	323	–	–	–	–	206	224	340	348
A3	235	243	355	358	221	221	311	323	120	122	321	321	206	224	340	348
Sex group 2																
B1	247	247	–	–	–	–	311	314	–	–	–	–	236	239	332	332
B1	247	247	355	361	215	221	311	314	120	124	325	333	236	239	332	332
B2	259	259	–	–	–	–	320	320	–	–	–	–	227	239	336	348
B2	259	259	340	355	221	224	320	320	120	124	321	325	227	239	336	348
B3	251	259	–	–	–	–	320	323	–	–	–	–	230	236	300	336
B3	251	259	355	355	215	221	320	323	120	120	321	333	230	236	300	336
B4	243	251	–	–	–	–	326	332	–	–	–	–	206	206	344	348
B4	243	251	340	340	215	227	326	332	114	120	333	333	206	206	344	348
B5	239	251	–	–	–	–	311	320	–	–	–	–	206	236	336	352
B5	239	251	340	355	215	221	311	320	114	120	349	349	206	236	336	352
B6	243	243	–	–	–	–	311	314	–	–	–	–	206	245	332	332
B6	243	255	340	355	215	221	320	323	120	120	305	333	206	245	332	348
B7	255	259	–	–	–	–	320	326	–	–	–	–	206	236	340	348
B7	243	251	340	355	215	224	314	326	114	120	305	333	206	245	332	344

Grey lines = most likely parents from four microsatellite reconstruction; white lines = most likely parents from eight microsatellite reconstruction.

Table 2 Relative contribution of parents from two sex groups (A and B) to the offspring of four cohorts of *Seriola dumerili*: (a) results from four microsatellite analysis, (b) results from eight microsatellite analysis.

	(a)															Individual contribution (%)
	2002			2003–2004–2005			2006			2007			Total			
	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	
B1	4.35	0.00	0.00	8.70	0.00	4.35	26.09	0.00	0.00	48.00	0.00	4.00	22.34	0.00	2.13	24.47
B2	8.70	0.00	0.00	34.78	0.00	4.35	60.87	0.00	0.00	4.00	0.00	16.00	26.60	0.00	5.32	31.91
B3	52.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	12.77	0.00	0.00	12.77
B4	0.00	0.00	0.00	4.35	0.00	0.00	13.04	0.00	0.00	16.00	0.00	0.00	8.51	0.00	0.00	8.51
B5	34.78	0.00	0.00	0.00	39.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.51	9.57	0.00	18.09
B6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	8.00	0.00	0.00	2.13	2.13
B7	0.00	0.00	0.00	4.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.00	1.06	0.00	1.06	2.13
Individual contribution (%)	100	0.00	0.00	52.17	39.13	8.70	100	0.00	0.00	68.00	0.00	32.00	79.79	9.57	10.64	–
	(b)															Individual contribution (%)
	2002			2003–2004–2005			2006			2007			Total			
	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	A1	A2	A3	
B1	4.35	0.00	0.00	8.70	0.00	4.35	26.09	0.00	0.00	48.00	0.00	4.00	22.34	0.00	2.13	24.47
B2	8.70	0.00	0.00	39.13	4.35	0.00	60.87	0.00	0.00	4.00	0.00	16.00	27.66	1.06	4.26	32.98
B3	56.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	13.83	0.00	0.00	13.83
B4	0.00	0.00	0.00	0.00	0.00	0.00	4.35	0.00	0.00	12.00	0.00	0.00	4.26	0.00	0.00	4.26
B5	30.43	0.00	0.00	0.00	39.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.45	9.57	0.00	17.02
B6	0.00	0.00	0.00	0.00	0.00	0.00	4.35	0.00	0.00	0.00	0.00	8.00	1.06	0.00	2.13	3.19
B7	0.00	0.00	0.00	4.35	0.00	0.00	4.35	0.00	0.00	4.00	0.00	4.00	3.19	0.00	1.06	4.26
Individual contribution (%)	100	0.00	0.00	52.17	43.48	4.35	100	0.00	0.00	68.00	0.00	32.00	79.79	10.64	9.57	–

to the offspring, suggesting that the variability of the markers can be more critical than the number of markers used (Sefc & Koblmüller 2009). Estimated effective population sizes per cohort [Ne 2002: 4; 95% CI (3–8); Ne 2003–05: 5; 95% CI (2–2*10⁹); Ne 2006: 4; 95% CI (3–8); Ne 2007: 5; 95% CI (3–12)] and for the whole population (Ne total: 6; 95% CI (4–13)) corresponded well to the estimated contribution of the parents to each cohort and to the total offspring and were lower than the number of breeders present, probably influenced by family variance. The large confidence interval observed in the 2003–2005 cohort could be the result of pooling three cohorts with different parental contribution. Our results contrast with previous observations in *Seriola quinqueradiata*, for which most of the broodstock was reported to spawn simultaneously (Nagakura *et al.* 2003), suggesting unequal parental contribution during mass communal spawning of *S. dumerili*. However, reasons for the observed unequal parental contribution could be attributed to a number of reasons, for example, sperm competition, mate choice, different fertility rates or different survival rates of the offspring in early stages of development, and deserve further investigation. Although all but one of the parents contributed to the offspring over all the years, annual unequal contribution of breeders could result in significant inbreeding over generations when numbers of breeders are small; thus, mechanisms such as active or cryptic mate choice (e.g., gametic incompatibility), common in natural populations (Pitcher & Neff 2007; Consuegra & García de Leániz 2008), should be taken into account in commercial breeding programs (Nordeide 2007), for example by dividing the broodstock in groups that could be redistributed during the spawning to guarantee the contribution of as much broodstock as possible to the offspring.

Acknowledgements

Carlos García de Leaniz provided helpful comments. This study was funded by the Ministerio de Ciencia e Innovación (Spain) (Grant Ref. AGL2008-05014-C02).

References

- Blonk R.J.W., Komen J., Kamstra A., Crooijmans R.P.M.A. & van Arendonk J.A.M. (2009) Levels of inbreeding in group mating captive broodstock populations of Common sole, (*Solea solea*), inferred from parental relatedness and contribution. *Aquaculture* **289**, 26–31.
- Butler K., Field C., Herbinger C.M. & Smith B.R. (2004) Accuracy, efficiency and robustness of four algorithms allowing full sibship reconstruction from DNA marker data. *Molecular Ecology* **13**, 1589–600.
- Consuegra S. & García de Leániz C. (2008) MHC-mediated mate choice increases parasite resistance in salmon. *Proceedings of the Royal Society B Biological Science* **275**, 1397–403.
- Duchesne P., Godbout M.H. & Bernatchez L. (2002) PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. *Molecular Ecology Notes* **2**, 191–3.
- Frost L.A., Evans B.S. & Jerry D.R. (2006) Loss of genetic diversity due to hatchery culture practices in barramundi (*Lates calcarifer*). *Aquaculture* **261**, 1056–64.
- Garber A.F., Tosh J.J., Fordham S.E., Hubert S., Simpson G., Symonds J.E., Robinson J.A.B., Bowman S. & Trippel E.A. (2010) Survival and growth traits at harvest of communally reared families of Atlantic cod (*Gadus morhua*). *Aquaculture* **307**, 12–9.
- Gjerde B., Terjesen B.F., Barr Y., Lein I. & Thorland I. (2004) Genetic variation for juvenile growth and survival in Atlantic cod (*Gadus morhua*). *Aquaculture* **236**, 167–77.
- Jackson T.R., Martin-Robichaud D.J. & Reith M.E. (2003) Application of DNA markers to the management of Atlantic halibut (*Hippoglossus hippoglossus*) broodstock. *Aquaculture* **220**, 245–59.
- Jerez S., Samper M., Santamaría F.J., Villamandos J.E., Cejas J.R. & Felipe B.C. (2006) Natural spawning of greater amberjack (*Seriola dumerili*) kept in captivity in the Canary Islands. *Aquaculture* **252**, 199–207.
- Jones O.R. & Wang J. (2010) COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**, 551–5.
- Kalinowski S.T., Taper M.L. & Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099–106.
- Massah N., Wang J., Russell J.H., Van Niejenhuis A. & El-Kassaby Y.A. (2010) Genealogical relationship among members of selection and production populations of yellow cedar (*Callitropsis nootkatensis* [D Don] Oerst.) in the absence of parental information. *Journal of Heredity* **101**, 154–63.
- McLean J., Seamons T., Dauer M., Bentzen P. & Quinn T. (2008) Variation in reproductive success and effective number of breeders in a hatchery population of steelhead trout (*Oncorhynchus mykiss*): examination by microsatellite-based parentage analysis. *Conservation Genetics* **9**, 295–304.
- Nagakura Y., Ohara E., Nakano S., Takahashi M., Ishibashi N., Sakamoto T. & Okamoto N.A. (2003) Parentage test confirmed that almost all yellowtail (*Seriola quinqueradiata*) parents mated randomly in induced spawning. *Saibai Gyogyo Gijutsu Kaihatsu Kenkyu* **30**, 43–7.
- Nakada M. (2000) Yellowtail and related species culture. In: *Encyclopedia of Aquaculture* (Ed. by R.R. Stickney), pp. 1007–36. Wiley, London.
- Nordeide J.T. (2007) Is there more in 'gamete quality' than quality of the gametes? A review of effects of female mate choice and genetic compatibility on offspring quality *Aquaculture Research* **38**, 1–16.
- Pitcher T. & Neff B. (2007) Genetic quality and offspring performance in Chinook salmon: implications for supportive breeding. *Conservation Genetics* **8**, 607–16.
- Porta J., Maria Porta J., Cañavate P., Martínez-Rodríguez G. & Carmen Alvarez M. (2007) Substantial loss of genetic variation in a single generation of Senegalese sole (*Solea senegalensis*) culture: implications in the domestication process. *Journal of Fish Biology* **71**, 223–34.

- Reed D.H. & Frankham R. (2003) Correlation between fitness and genetic diversity. *Conservation Biology* **17**, 230–7.
- Renshaw M.A., Patton J.C., Rexroad C.E. & Gold J.R. (2006) PCR primers for trinucleotide and tetranucleotide microsatellites in greater amberjack, *Seriola dumerili*. *Molecular Ecology, Notes* **6**, 1162–4.
- Renshaw M., Patton J., Rexroad C. & Gold J.R. (2007) Isolation and characterization of dinucleotide microsatellites in greater amberjack, *Seriola dumerili*. *Conservation Genetics* **8**, 1009–11.
- Rice W.R. (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223–5.
- Sefc K.M. & Koblmüller S. (2009) Assessing parent numbers from offspring genotypes: the importance of marker polymorphism. *Journal of Heredity* **100**, 197–205.
- Sekino M., Saitoh K., Yamada T., Kumagai A., Hara M. & Yamashita Y. (2003) Microsatellite-based pedigree tracing in a Japanese flounder *Paralichthys olivaceus* hatchery strain: implications for hatchery management related to stock enhancement program. *Aquaculture* **221**, 255–63.
- Smith B.R., Herbinger C.M. & Merry H.R. (2001) Accurate partition of individuals into full-sib families from genetic data without parental information. *Genetics* **158**, 1329–38.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Graphic representation of full-sib and half-sib relationships per cohort in six cohorts (four groups) of *Seriola dumerili* derived from 11 original breeders.

Table S1 Genetic variability of the offspring of 11 wild *Seriola dumerili* per cohort (or group) and locus.