SELECTING SOLE

Breeding programs for natural - mating populations

Robbert Blonk

THESIS COMMITTEE

Thesis supervisor

Prof. dr. ir. J.A.M. van Arendonk
Professor of Animal Breeding and Genetics
Wageningen University

Thesis co-supervisor

Dr. ir. J. Komen
Associate-Professor of Animal Breeding and Genetics
Wageningen University

Other members

Dr. B. Gjerde
NOFIMA, Ås, Norway

Prof. dr. A.D. Rijnsdorp
IMARES, Wageningen University and Research centre

Prof. dr. J.A.J. Verreth
Wageningen University

Prof. dr. E. Verrier
AgroParisTech, Paris, France

This research was conducted under the auspices of the Wageningen Institute of Animal Sciences (WIAS) graduate school

SELECTING SOLE:

Breeding programs for natural - mating populations

Robbert Blonk

Thesis

submitted in fulfilment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus

Prof. dr. M.J. Kropff,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 5 November 2010

at 4 p.m. in the Aula.

Robbert J.W. Blonk
Selecting Sole: Breeding programs for natural - mating populations,
174 pages.

Thesis, Wageningen University, Wageningen, NL (2010)
With references, with summaries in Dutch and English.

ISBN 978-90-8585-712-9

Abstract

The aim of this thesis was to design a breeding program for increased productivity of farmed common sole, *Solea solea*, 1) using natural mating in groups to obtain offspring and 2) using present farm infrastructures as much as possible. Parental allocation with DNA marker data on offspring from natural mating parents showed that parental contributions were highly skewed. This indicates that few animals contribute to majority of the offspring. Levels of coancestry in offspring populations were high (2-4%) showing that selection methods need to restrict rates of inbreeding in future generations.

Genetic variances of body weight and body length in common sole were measured at harvest and it was shown that genetic improvement of these traits is possible. It was pointed out that selection on growth of common sole needs to be accompanied by selection for shape to compensate for undesired correlated responses in shape. Further, it was demonstrated that in populations with skewed contributions, use of marker data can be more efficient once breeding values are estimated with continuous molecular relatedness rather than with a reconstructed pedigree. To decrease costs of breeding programs, selection of parents may be from production populations directly. Results indicated that estimation of genetic parameters is affected by husbandry practices as grading, but that effects are predictable. To optimise breeding programs with natural mating of parents, 2-stage selection schemes with optimal contribution selection from mass selected and genotyped fractions were compared with mass selection schemes. Using 2-stage selection schemes, rates of inbreeding can be restricted with a smaller nucleus than with mass selection schemes. However, response is lower as well. The different findings in this thesis are put into context of a breeding program for common sole and discussed in relation to profitability, benefits of natural mating and correlated responses to selection for growth.

Contents

1

General Introduction

Aquaculture and genetic improvement

An increasing part of the worldwide aquaculture sector is currently based on genetically improved fish. In Europe, the main species with large scale breeding programs are Atlantic salmon, *Salmo salar*, rainbow trout, *Oncorhynchus mykiss*, gilthead seabream, *Sparus aurata* and European sea bass, *Dicentrarchus labrax* (Chavanne et al., 2008). All of these species have a large market and volumes produced by grow out farms are massive, from 60 to more than 900 thousand tonnes in 2008 (FAO, 2010).

In contrast, aquaculture enterprises which either are in start up phase, culture a "new" species or which serve an exclusive niche market, have relatively small production volumes. Examples of such initiatives in the Netherlands are culture of pike perch, *Sander lucioperca* and soles, *Solea spp*. Such small scale aquaculture ventures often rely entirely on wild broodstocks and do not have programs to select genetically superior breeders and to control rates of inbreeding. However, lack of genetic improvement of populations greatly obstructs improvement of economical viability of farms. For example, genetic improvement of Atlantic salmon resulted in increased growth rates up to 50% compared to wild founders (Gjedrem, 2005). Among other causes, breeding programs triggered and promoted development of salmon industry worldwide.

Costs of genetic improvement

In breeding programs, genetically superior breeders are selected from present populations in order to improve the next generation. Here, reproduction of parents is followed by genetic evaluation of offspring to obtain breeding values. After this, offspring with the best breeding values are selected to serve as parents for the next generation (figure 1.1).

In many breeding programs, family relationships between animals are required for genetic evaluation. In captive species such as e.g. cattle and pigs, these relationships are generally inferred from known pedigrees and to recognise individuals within groups over time, animals are tagged at birth. However, in many fish species, early tagging of families is impossible due to small size of animals at hatching. Consequently, families are often reared separately in family rearing tanks until tagging size. Alternatively, individuals of different families are subsequently reared communally until tagging size, tagged and genotyped for pedigree reconstruction.

Both separate rearing of families and genotyping greatly increases costs of consumables and infrastructure. To increase profitability, aquaculture production enterprises that run a breeding program should improve efficiency. This can be achieved through increased scale of production or by specialisation of activities. In the latter, reproduction and breeding programs are located in few specialised breeding companies whereas others have a focus on production of consumption fish. A well known example of this is culture of Atlantic salmon.

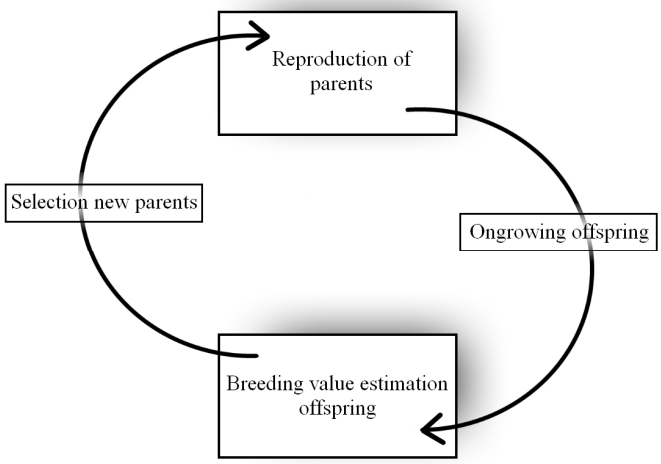


Figure 1.1 Overview of steps in a breeding program.

In contrast, small scale farms have a primary commercial focus on production activities. In addition, they need to select and reproduce their own broodstock. In order to facilitate implementation of breeding programs into such small scale farms, there is a severe need for breeding programs that are cheap and use available farm infrastructure.

Natural mating

For many cultured fish species, offspring can be obtained using artificial reproduction. Examples are salmonids, *Salmonidae spp*. (Billard, 1992), carps, *Cyprinus spp*. (Billard, 1995), turbot, *Scophthalmus maximus* (Chereguini et al., 1999) and African catfish, *Clarias gariepinus* (Goos and Richter, 1995). In contrast, controlled reproduction is

difficult or impossible for several other species (Mylonas et al., 2010) and natural mating in groups is needed to obtain offspring. The use of natural mating can also be a matter of efficiency as artificial reproduction can be very labour intensive. Examples may be found in species such as European sea bass, *Dicentrarchus labrax* (Massault et al., 2009), gilthead seabream, *Sparus aurata* (Brown et al., 2005) and Atlantic cod, *Gadus morhua* (Bekkevold et al., 2002; Herlin et al., 2007).

With natural mating in groups, males and females are housed in one large tank. During spawning seasons, animals spawn naturally and fertilised eggs are collected from the tank; there is no control on parental contributions to offspring populations. This typically leads to mixing of families, unknown pedigrees of individuals, unbalanced mating designs and skewed parental contributions with few animals producing most of the offspring (Bekkevold et al., 2002; Brown et al., 2005; Fessehaye et al., 2006b). Without doubt, the need to use natural mating in groups to obtain offspring impedes genetic improvement programs. For example, mass selection (selection on phenotype) on such populations is expected to yield high rates of inbreeding when not correcting for parental contributions (Bentsen and Olesen, 2002). Although skewness of parental contributions has been determined for several natural mating species, it remains unclear how to efficiently prevent excess of inbreeding and how to optimise response while minimising genotyping costs.

Aim of this thesis

Common sole, *Solea solea* (in Dutch: “Noordzeetong”), is a highly valued flatfish species in the North-Eastern Atlantic. Culture of common sole, but also of its relative Senegalese sole, *Solea senegalensis*, serves a substantial niche market and shows high potential as natural stocks are under pressure of fisheries. Reproduction of common sole still relies on natural mating of broodstocks (Dinis, 1999; Imsland et al., 2003b; Howell et al., 2006). At the start of this project, no breeding program for genetic improvement of *Solea spp.* was available. To achieve improvement of production while controlling rates of inbreeding in future generations of cultured common sole, this project was initiated by Solea B.V. (the Netherlands) and the Animal Breeding and Genomics centre (Wageningen University). Aim of the project was to design a breeding program for increased productivity of farmed common sole that 1) uses natural mating of parents to obtain offspring and 2) is suitable for small scale farms by using present farm

infrastructures as much as possible. This project was financed by the Netherlands Organisation for Scientific Research (NWO, Casimir personal grant) and received financial support from Solea B.V. and Wageningen University.

Outline

To design and run a breeding program, several standard steps should be undertaken (see figure 1.2). First, a founder population with sufficient genetic variation is required. Then, a breeding goal should be defined including economically and biologically relevant traits. For traits in the breeding goal, heritability and correlations need to be estimated.

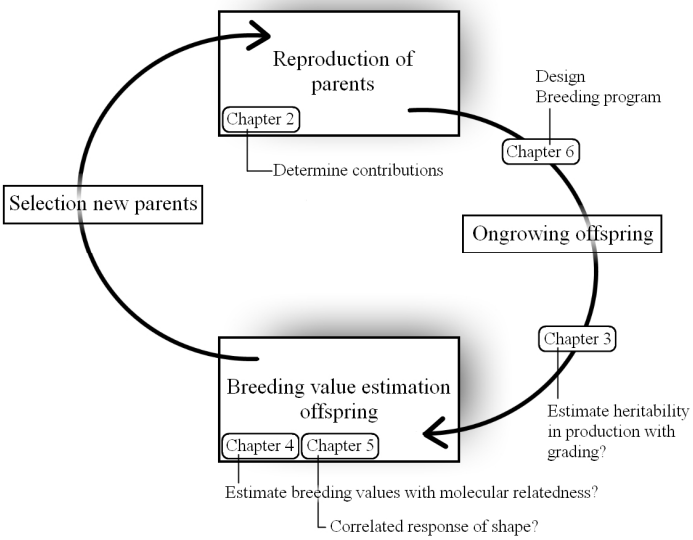


Figure 1.2 Overview of steps in a breeding program. The different chapters of this thesis are indicated at the applicable step.

To optimise response to selection, rates of inbreeding and profitability, decisions have to be made on the number of selection candidates, selected animals and methods for breeding value estimation and selection. Finally, breeding values of offspring are estimated and based on this, animals are selected for the new parental population (Falconer and Mackay, 1996; Gjerde et al., 1996b; Gjedrem, 2000; Gjedrem, 2005). At the start of this thesis, no such data was available for common sole.

To optimise a breeding program for common sole with natural mating, skewness of parental contributions, as well as its effects on rates of inbreeding and response to selection, should be determined. Additionally, with natural mating of parents, estimation

of genetic parameters and BLUP breeding values requires costly genotyping. Hence, the question is how to optimise and increase profitability of breeding programs with natural mating of parents while restricting genotyping. Further, when a breeding program aims at reduction of costs by using present farm infrastructures, selection may be from production stocks rather than from separate test populations. However, most production stocks undergo repeated grading and this may bias estimation of genetic parameters. Little is known about the extent to which grading affects genetic parameter estimation and this should be characterised when developing a breeding program. In this thesis, these questions are analysed and discussed in relation to common sole with natural mating of parents (figure 1.2).

In **chapter 2**, parental contributions and their effects on rates of inbreeding were characterised using parental allocation with genetic markers on offspring from a cultured common sole population. In addition, the difference between a generally used population genetic approach and a quantitative approach to estimate inbreeding is shown.

Chapter 3 demonstrates effects of grading of fish populations on estimated heritability of body weight and body length using phenotypic distributions over tanks at different time points after grading. Results from simulated and real populations are presented and compared.

When some parental genotypes are missing, methods to reconstruct explicit pedigrees lead to loss of data whereas methods that estimate molecular relationships between animals do not require parental genotypes and may be more efficient. In **chapter 4**, genetic markers are used to estimate molecular relationships between offspring from natural mating parents. Accuracies of estimated breeding values from this method are compared with results from a method with explicit pedigree reconstruction.

In **chapter 5**, it is determined if selection for improved production results into correlated responses of body shape of common sole. Genetic correlations between shape and production traits (e.g. body weight) are estimated. Moreover, heritabilities and standard errors from manual analysis and from digital image analysis are compared.

In **chapter 6**, breeding programs are optimised for the use of natural mating of parents. Stochastic simulation is used to estimate rates inbreeding and response to selection in breeding programs. Schemes with optimal contribution selection from mass selected and genotyped fractions (2-stage selection) are compared with schemes with mass selection.

In **chapter 7 (**general discussion), the use of natural mating and selection from production populations in breeding programs is integrated and further discussed with relation to implications for profitability of breeding programs, benefits of natural mating and other correlated responses.

References

Bekkevold, D. X., Hansen, M. M. and Loeschcke, V., 2002. Male reproductive competition in spawning aggregations of cod (*Gadus morhua* , L.). Molecular Ecology 11, 91-102.

Bentsen, H. B. and Olesen, I., 2002. Designing aquaculture mass selection programs to avoid high inbreeding rates. Aquaculture 204, 349-359.

Billard, R. R., 1992. Reproduction in rainbow trout: sex differentiation, dynamics of gametogenesis, biology and preservation of gametes. Aquaculture 100, 263-298.

Billard, R. R., 1995. Biology of sperm and artificial reproduction in carp. Aquaculture 129, 95-112.

Brown, C. R., Woolliams, J. A. and McAndrew, B. J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247, 219-225.

Chavanne, H., Norris, A., Haffray, P., Sonesson, A. K., Vandeputte, M., Chatain, B. and Boudry, P., 2008. Survey on the breeding practices in the European aquaculture industry, Reprofish Aquabreeding Workshop 2008.

Chereguini, O., de la Banda, I. G., Rasines, I. and Fernandez, A., 1999. Artificial fertilization in turbot, *Scophthalmus maximus* (L.): different methods and determination of the optimal sperm-egg ratio. Aquaculture Research 30, 319-324.

Dinis, M. T., 1999. A review on the cultivation potential of Solea senegalensis in Spain and in Portugal. Aquaculture 176, 27-38.

Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to quantitative genetics. Pearson Prentice Hall Harlow, Essex.

FAO, 2010. www.fao.org.

Fessehaye, Y., El-Bialy, Z., Rezk, M. A., Crooijmans, R., Bovenhuis, H. and Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. Aquaculture 256, 148-158.

Gjedrem, T., 2000. Genetic improvement of cold-water fish species. Aquaculture Research 31, 25-33.

Gjedrem, T., 2005. Selection and Breeding Programs in Aquaculture. Springer, Dordrecht.

Gjerde, B., Gjøen, H. M. and Villanueva, B., 1996. Optimum designs for fish breeding programmes with constrained inbreeding mass selection for a normally distributed trait. Livestock Production Science 47, 59-72.

Goos, H. and Richter, C., 1995. Internal and external factors controlling reproduction in the African catfish, *Clarias gariepinus*. Aquatic Living Resources 9, 45-58.

Herlin, M., Taggart, J. B., McAndrew, B. J. and Penman, D. J., 2007. Parentage allocation in a complex situation: A large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. Aquaculture 272, S195-S203.

Howell, B., Cañavate, P., Prickett, R. and Conceição, L., 2006. The Cultivation of Soles. Report of a 3rd workshop held at Cifpa El Toruño, Cadiz, Spain.

Imsland, A. K., Foss, A., Conceiçao, L. E. C., Dinis, M. T., Delbare, D., Schram, E., Kamstra, A., Rema, P. and White, P., 2003. A review of the culture potential of *Solea solea* and *S. senegalensis*. Reviews in Fish Biology and Fisheries 13, 379-408.

Massault, C., Hellemans, B., Louro, B., Batargias, C., Houdt, J. K. J. V., Canario, A., Volckaert, F. A. M., Bovenhuis, H., Haley, C. and Koning, D. J. d., 2009. QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. Animal Genetics 9999.

Mylonas, C. C., Fostier, A. and Zanuy, S., 2010. Broodstock management and hormonal manipulations of fish reproduction. General and Comparative Endocrinology 165, 516-534.

2

Levels of inbreeding in group mating captive broodstock populations inferred from parental relatedness and contribution.

Robbert J.W. Blonk
Hans Komen
Andries Kamstra
Richard P.M.A Crooijmans
Johan A.M. van Arendonk

Published in Aquaculture 289 (2009), 26-31

Abstract

In this paper, we estimate levels of inbreeding with parental relatedness and contribution inferred from microsatellites in groups of common sole that reproduce by natural mating. We present results on spawning patterns during one entire reproductive season of wild common sole, *Solea solea*, kept in two broodstock groups (28 animals in broodstock A; 20 animals in broodstock B) under semi-natural conditions. Batches of eggs were collected daily and incubated separately. First, we performed a parentage analysis on parents and samples of 24 newly hatched larvae from all batches, using 10 polymorphic microsatellite markers. As expected, contribution of parents to offspring was highly skewed. In both broodstocks, 5 or less parental pairs produced more than half of the total progeny. Natural spawning and unequal contributions of parents to offspring resulted in significant deviations from Hardy-Weinberg equilibria. Furthermore, few alleles were lost and levels of heterozygosity in offspring population increased. Next, we calculated relatedness between parents that mated successfully based on estimates of molecular similarity. Using parental relatedness and contributions, mean coefficients of coancestry in offspring were determined. Levels of coancestry in progeny were substantially high. These results show that due to different parental contributions, natural mating in groups can result in significant inbreeding in future generations despite of limited loss of alleles and high levels of heterozygosity in first generation progeny. This shows that using loss of alleles and levels of heterozygosity alone can be misleading for estimation of genetic diversity.

Introduction

Most livestock breeding programmes use controlled reproduction methods like artificial insemination, or natural mating of couples to control family structures and pedigree. This is of major importance when executing breeding programmes with restriction on level of inbreeding (Henderson, 1984; Bijma et al., 2000; Dupont-Nivet et al., 2006).

In some fish species controlled reproduction is a relatively easy matter. For example, in salmonids (Billard, 1992), carps, *Cyprinus spp.* (Billard, 1995), turbot, *Scophthalmus maximus* (Chereguini et al., 1999) and African catfish, *Clarias gariepinus* (Goos and Richter, 1995) artificial fertilization is used. A problem occurs when reproduction is dependent on natural mating of animals kept in groups. Natural mating in groups often results in production of massive and variable family sizes of unknown pedigree (Gjerde et al., 1996a; Komen et al., 2006). Furthermore, only a restricted number of animals (especially in males) contributes to majority of the descendants. Such skewed contributions have been shown in Nile tilapia, *Oreochromis niloticus* in hapas (Fessehaye et al., 2006b), Atlantic cod, *Gadus morhua* (Bekkevold et al., 2002; Bekkevold, 2006; Rowe, 2007) and in Gilthead seabream, *Sparus aurata* (Brown et al., 2005).

In sole it is still difficult to induce spawning by artificial means. Therefore, natural mating in groups is used (Dinis, 1999; Imsland et al., 2003b; Howell et al., 2006). Natural mating behaviour of common sole is described by Baynes et al. in 1994.

Consequences of natural mating in fish species are usually analyzed with classical population genetic approaches e.g. Exadactylos (1999), Perez-Enriquez et al. (1999) and Porta (2006a). In these cases, inbreeding and genetic variability are analyzed using loss of alleles and levels of heterozygosity such as F_{IS}. However, when performing directional selection, these methods can not be used to predict consequences of natural mating for populations. Additive genetic relationships on the other hand, are widely used to estimate genetic parameters and breeding values of individuals. Further, they can be used to optimize selection designs in terms of minimizing inbreeding, e.g. through optimal contribution (Sanchez et al., 2003). When pedigree information is absent, relatedness can be inferred from molecular markers (reviewed in Oliehoek et al. (2006)). In this paper we present a method to analyze levels of inbreeding in species that reproduce by natural

mating in groups, using estimates of parental relatedness together with contribution. Our results show that natural mating in groups leads towards an increased mean coefficient of (molecular) coancestry in offspring of sole. This is in contrast with traditional population genetic analysis where heterozygosity in progeny increases.

Methods

Broodstock

Two common sole broodstock populations A (n=28) and B (n=20) were collected from the Dutch North-Western coastal area during 2003 through 2005. From collection up to start of the experiment both broodstocks were conditioned indoor in separate tanks. Mean body lengths and weights are shown in table 2.1. Broodstocks were fed a diet of moist pellets and polychaetes to 0.5 % of their bodyweight every second day. During the spawning period the diet was given ad libitum.

Table 2.1 Number of parents, mean body weight (BW) (± sd), body length (BL) (± sd) , mean hatching rate (Hr) and estimated number of produced larvae (L) per broodstock (BS).

BS		n	BW (g)		BL (mm)		Hr (%)	L (106)
A	**Males**	14	504.9	(102.5)	36.2	(2.7)	36	1.2
	Females	14	836.5	(218.6)	42.6	(2.9)		
B	**Males**	10	476.8	(128.2)	36.5	(2.7)	30	2.5
	Females	10	977	(143.1)	45.7	(3.1)		

Broodstock management

Each broodstock tank had a diameter of approximately 3 m and a height of 1.40 m. Salinity was 34 ppt. Each tank had a sand bottom of approximately 5 cm. Both tanks were connected to one recirculation system (total volume 70 m^3). Artificial temperature and light regimes were controlled per tank separately and simulated those of natural circumstances (52 ° N 2.5 ° E) with a six months difference between broodstocks. In broodstock A spawning commenced on June the 28th of 2006 and lasted until October the 12th. Spawning of broodstock B lasted from January the 2nd of 2007 until May the 5th.

Every three weeks spawning was suspended artificially for one week by lowering water temperature.

Egg collection

Pelagic eggs were continuously collected at the out flow of the spawning tank. Every morning at approximately 09.00 hours all eggs were harvested from the egg collector and kept as one separate batch during subsequent incubation and larval rearing. Each batch was weighed and egg quality was evaluated as “low”, “moderate” or “good”. Batches were incubated subsequently in a conical incubation tank of approximately 80 l. Each incubation tank was connected to a recirculating system with UV-treatment. Incubation temperatures were 10 °C and hatching took place after 3 days incubation.

Every day from day 0 until hatching, sinking eggs were drawn off and weighed. From these data the approximate number of hatched eggs per batch was calculated:

$$He_i = (Wt_i - \sum_{j=1}^{n} (Ws_{ij})) \cdot \rho e \qquad 2.1$$

where He_i is number of hatched eggs of batch i, Wt_i is weight (g) of eggs of batch *i* at time of collection and Ws_{ij} is weight (g) of sinking eggs of batch *i* at day *j*. *Pe* is number of eggs per g which was determined at start of the spawning season by counting numbers of eggs in approximately 1 g of eggs.

Due to restricted capacity in the incubation system several small batches had to be incubated in small floating 1 l tubes within the same incubation system. From small batches sinking eggs could not be measured as these quantities were generally too small for accurate weighing. Instead hatching percentages were estimated from average hatching percentages in large batches with similar evaluated quality (i.e. low, moderate and good) according to:

$$He_{ik} = Wt_{ik} \cdot Hp_k \cdot \rho e \qquad 2.2$$

where He_{ij} is number of hatched eggs of batch *i* with quality *k*, *Wt* is total weight (g) of eggs in the small batch at day of collection *i* and *Hp* is predicted hatching rate at quality *k* in large batches. *Pe* is number of eggs per g.

DNA sampling and analysis

Before the experiment, a blood sample of each parent was taken for DNA analysis. Samples were stored in EDTA 0.27 M in physiological salt (0.8% NaCl) solution at -80 °C. For DNA isolation of parental blood, Puregene DNA purification kit for non-mammalian whole blood samples (Gentra Systems) was used. Sampling of larvae for DNA analysis was performed before first feeding at 3-4 days post hatch. From each batch, 24 larvae were sampled at random from the incubator and processed individually for DNA isolation using nucleospin tissue columns (96 procedure, Machery-Nagel). To test if DNA was extracted successfully, DNA concentrations were measured from several samples in all plates using a spectrophotometer (Nanodrop technologies ND-1000). For all samples DNA concentration was diluted to 5-10 ng/µl for further analysis.

The following 10 microsatellite markers were used for DNA analysis: AF173855, AF173854, AF173852, AF173849 (Iyengar et al., 2000), AY950593, AY950592, AY950591, AY950589, AY950588, AY950587 (Garoia et al., 2006).
PCR amplification involved 5 minutes of denaturation at 95 °C followed by 36 cycles of consecutively 30 s denaturation at 95 °C, 45 s annealing at 55 °C and 90 s elongation at 72 °C. After 36 cycles a final elongation step of 10 minutes at 72 °C was applied. After PCR amplification, marker samples were pooled per individual and analyzed on a ABI 3730 automatic sequencer. Fragment sizes were set relatively to Genescan LIZ 500 size standard (Applied Biosystems). Output data was analyzed using Genemapper software (Applied Biosystems) in order to determine allele profiles at each locus. Parental allocation was performed with PAPA 2.0 software (Duchesne et al., 2002). Papa 2.0 is a software package which performs parental allocation by calculating breeding likelihood of a parental pair of (multilocus) genotypes producing a given offspring genotype. It allows a certain degree of genotyping error or mutation. The breeding couple with highest likelihood is defined as the most likely parental pair. Results were manually checked for correct allocation according to Mendelian inheritance afterwards.

Contribution of parents

Using genotyped larvae, contribution of parents and parental pairs to total offspring was calculated relatively to the total number of larvae produced. This was done for both broodstocks separately.

Population genetics

Expected and observed heterozygosities, resulting fixation rates (F_{IS}) after Weir and Cockerham (1984), Hardy-Weinberg exact tests and probability tests for differentiation of allelic distribution were calculated with Genepop web application (Raymond and Rousset, 1995). Calculations were performed for parental and offspring populations separately. Further, broodstock A and B were analyzed separately as well as pooled. Expected heterozygosities were calculated using observed allele frequencies from the same population.

Parental relatedness

To determine relatedness between broodstock animals, a relationship matrix (A matrix) was constructed using genotypic information. This was done for broodstock A and B separately. Construction of the A matrix involved two steps. At off diagonal elements, relatedness coefficients between individuals x and y (r_{xy}) were estimated using molecular similarity index $S_{xy,l}$ (Lynch, 1988; Caballero and Toro, 2000; Eding and Meuwissen, 2001). Here $S_{xy,l} = ¼[I_{ac} + I_{ad} + I_{bc} + I_{bd}]$ (Li and Horvitz, 1953) where at locus l, a and b are alleles of individual x and c and d are alleles of individual y. When at I_{ac} allele a is identical to allele c, I_{ac} equals one, and zero otherwise, etc. According to Lynch (1988) expected $S_{xy,l}$ as $E[S_{xy,l}] = f_{xy} + (1 - f_{xy})s_l$ with f_{xy} expressing the probability of alleles being *identical by descent* (IBD), $(1 - f_{xy})s_l$ the probability of being *alike in state* (AIS) where s_l is mean similarity (sum over squared allele frequencies, Σp^2) at locus l in the base population. Since in diploid species $r = 2f$ (Malécot, 1948; Falconer and Mackay, 1996), relatedness coefficient (r_{xy}) is then calculated as (Oliehoek et al., 2006):

$$r_{xy} = \frac{2}{L}\sum_{l=1}^{L}\frac{S_{xy,l} - s_l}{1 - s_l}$$

where L is number of loci.

Diagonal elements (r_{xx}) of the A matrix are defined as $r_{xx} = 1 + F_x$ where F_x is the coefficient of inbreeding for animal x. The estimator of relatedness in its simplest form ignores AIS and thus $s_l = 0$. Subsequently $f_{xy} = \bar{S}_{xy}$ leads to $2S_{xx} = 1 + F_x$ which can be rearranged to $F_x = 2(S_{xx} - ½)$. From this, the coefficient of inbreeding was derived:

$$F_x = \frac{1}{L}\sum_{l=1}^{L}\frac{2(S_{xx,l} - 1/2) - s_l}{1 - s_l} \quad .$$

These estimators form an A matrix where mean relatedness and mean coefficient of inbreeding are zero. To account for overestimation of diversity, relatedness estimates at all elements of the A matrix were corrected using regression coefficient (β) (Oliehoek et al., 2006) where β = 0.079[ln(no.loci)-1][ln(no. alleles)+1.22]. All elements were then regressed to their mean using the following equation $\hat{r}_{xy}^{*} = \bar{r}_{xy} + \beta(\hat{r}_{xy} - \bar{r}_{xy})$. As $r_{xx} = 1 + F_x$, a value of 1 was added to diagonal elements.

Mean relatedness $\bar{r}$ and coefficients of coancestry were determined for offspring populations as proposed by Meuwissen (1997). This was done for both offspring populations separately. Here $\bar{r}$ = c'Ac where $\bar{r}$ is mean coefficient of relatedness, c is the vector with actual parental contributions and A is the relationship matrix of parents in the analyzed broodstock as explained above. Contributions were scaled to add up to ½ for each gender. Calculation of mean coancestry was done according to $\bar{r} = 2\bar{f}$ (Malécot, 1948; Falconer and Mackay, 1996).

Results

Density of eggs was estimated at approximately 550 eggs per g. Estimated cumulative larvae production was 1.2 million larvae in broodstock A and 2.5 million larvae in broodstock B. In broodstock A, 39 batches were collected of which 29 were large and 10 small. Fifty-one batches were collected in broodstock B of which 40 were large and 11 small. Small batches from broodstock A contained in total 210,000 larvae. For broodstock B this was a total of 140,000 larvae. Average hatching percentages were 36% in broodstock A and 30% in broodstock B. Average hatching rates of large batches increased proportionally to corresponding quality evaluation: *low:* 21% ± 12%; *moderate:* 45% ± 20%; *good:* 62% ± 17%.

Average number of parents contributing to a batch was 3.7 in broodstock A and 5.7 in broodstock B. Other broodstock parameters can be found in table 2.1. Some larval samples were lost due to poor DNA quality. All parents and 2103 offspring were genotyped successfully. In total 134 alleles were found in the parental population whereas 125 alleles in the offspring population. Unequivocal parental allocation success of larvae genotypes was high (>99%) in both broodstocks.

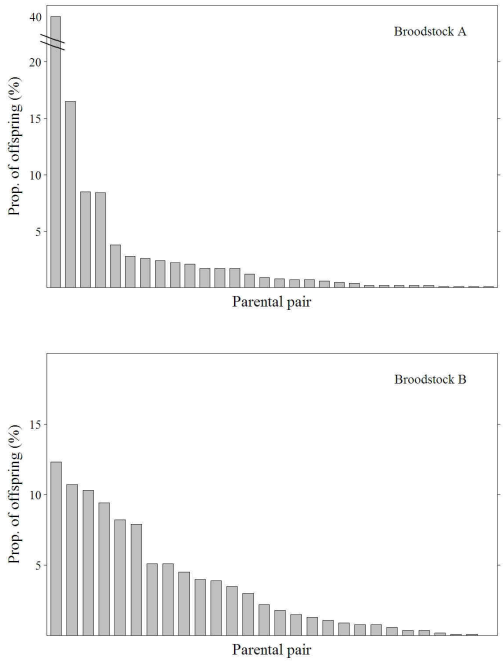


Figure 2.1 Contribution (%) of parental pairs to total number of offspring in one spawning season.

Contribution of parents

The contribution of parental pairs to total offspring population is given in figure 2.1. In both broodstocks, the distribution is skewed; only few parental pairs contributed to most of the offspring. In broodstock A the largest offspring group resulting from one parental pair comprised nearly 40% of the total offspring population. The distribution of contributions of individual parents to the total offspring population was found to be skewed (figure 2.2): in broodstock A male 8 sired 46% of the offspring. The second most productive male produced 16% of the offspring. In broodstock A, 5 males did not participate in spawning; in broodstock B this was 3 males. Of the females, 4 did not participate in spawning in broodstock A; this was 3 in broodstock B.

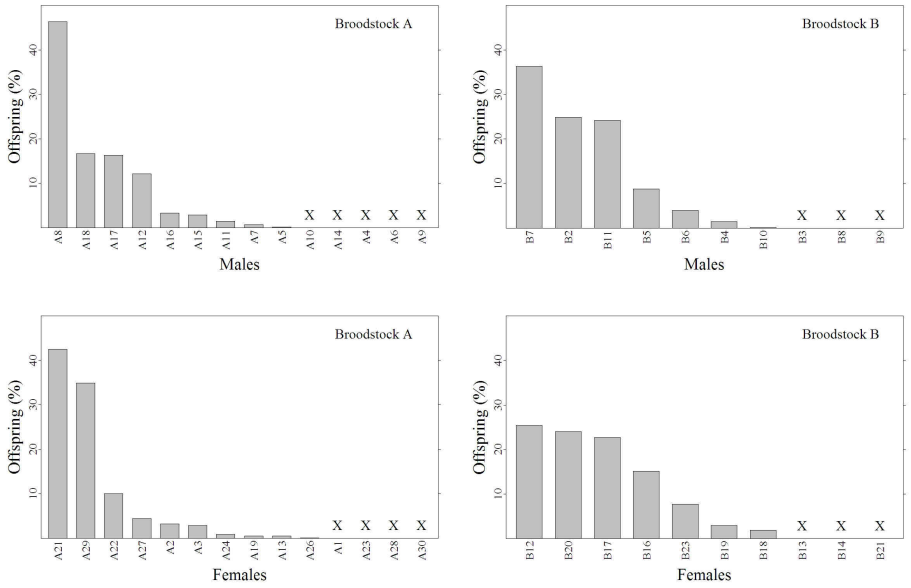


Figure 2.2 Contributions (%) of males and females to total number of offspring in one spawning season. Crosses (x) refer to non-contributing animals.

Population genetics

Expected and observed heterozygosities, fixation indices (F_{IS}), tests for deviation from Hardy-Weinberg equilibrium and chi-square tests for differentiation of allele frequencies between parent and offspring populations (pooled data) are shown for each locus in table 2.2. Analysis was performed for separate broodstocks as well, but results were not different. At some markers alleles were lost, particularly at locus 7. In the parental populations, loci 4, 7, 9 and 10 (40%) were not in Hardy-Weinberg equilibrium. In the offspring population, no loci were in Hardy-Weinberg equilibrium. In general, differentiation of allelic distribution across generations was significant, indicating a change in allele frequencies.

Observed heterozygosity was lower than expected in both the parental and offspring populations. Compared to the parental population, relatively more heterozygotes were observed in the offspring population. This resulted in lower F_{IS} values in the offspring population.

Table 2.2 Number of alleles, frequencies of expected (He) and observed (Ho) heterozygosity, F_{IS} fixation index, Hardy-Weinberg equilibrium P-value at 10 microsatellite marker loci[1] in the parental and the offspring population[2].

	Parents					Offspring					
Locus	**na**	**He**	**Ho**	**FIS**	**HW**	**na**	**He**	**Ho**	**FIS**	**HW**	**Diff**
1	10	0.77	0.75	0.02		10	0.79	0.81	-0.02	***	***
2	6	0.48	0.40	0.17		5	0.44	0.47	-0.07	***	***
3	8	0.72	0.69	0.05		7	0.70	0.79	-0.12	***	***
4	9	0.70	0.70	0.01	***	9	0.70	0.69	0.01	***	
5	16	0.91	0.83	0.08		14	0.89	0.95	-0.06	***	***
6	19	0.93	0.94	-0.01		19	0.90	0.95	-0.06	***	*
7	20	0.82	0.79	0.03	*	16	0.85	0.90	-0.06	***	***
8	13	0.85	0.75	0.12		13	0.81	0.69	0.15	***	***
9	12	0.87	0.77	0.12	*	12	0.82	0.74	0.10	***	***
10	21	0.94	0.80	0.15	**	20	0.91	0.74	0.19	***	***
Total	**134**	**0.8**	**0.74**	**0.07**		**125**	**0.78**	**0.77**	**0.01**		

[1]Loci accession names: 1: AF173855; 2: AF173854; 3: AF173852; 4: AF173849; 5: AY950593; 6: AY950592; 7: AY950591; 8: AY950589; 9: AY950588; 10: AY950587.

[2]Results from pooled broodstock and offspring populations (A and B)

*[3]Per locus, the P value for differentiation of allelic distribution across the parental and offspring populations is given. Significance levels: * for P < 0.050, ** for P < 0.010,*** for P << 0.000.*

Relatedness

Values of relatedness in the A matrix for broodstock A ranged from -0.26 to 0.26. These values reflect pairs being relatively unrelated to related. In broodstock B this was -0.24 to 0.25. In the offspring, mean relatedness was 0.1 for broodstock A and 0.04 for broodstocks B. Mean coancestry in the offspring population from broodstock A thus was higher (4.9%) compared to offspring from B (2.0%).

Discussion

Our data show that natural mating in groups of common sole produces a typical non-uniform mating pattern: some animals mated more frequently with certain partners than

with others. Contribution of parents and pairs to offspring was found to be skewed as well. Both broodstocks roughly showed the same pattern, although broodstock A was more extreme.

In the parental population, most loci were in Hardy-Weinberg equilibrium. Due to participation of most parents many parental alleles were retained in the next generation: 125 out of 134. Nevertheless, the offspring population was not in HW-equilibrium. A decrease in F_{IS} was found implying relatively more heterozygotes in offspring than in the parental population. From a population genetics point of view this can be interpreted as an indication that genetic diversity, in terms of number of alleles and a minimum level of heterozygosity, was relatively well preserved. However, different contributions of parents engendered a significant change in frequency for most alleles in offspring.

Two studies performing population genetic analysis on juvenile *Solea solea* (Exadactylos, 1999) and on grown out *Solea senegalensis* (Porta et al., 2006b; Porta et al., 2007) concluded a loss of genetic diversity from parents to offspring due to a loss of alleles and heterozygosity. In our study, it was found that on average 5 parents contributed to one batch (data not shown) and Porta et al. (2006b) found only one female and two males contributing during an entire spawning season of *Solea senegalensis*. Therefore, it is very likely that in the mentioned studies only few parents contributed to the tested populations. A limited number of batches obviously results in a low genetic diversity, as very likely only few parents contribute to one batch. In this study, an entire production season was sampled and therefore covered maximal potential genetic diversity.

Despite of a broad potential genetic diversity covered, our calculations showed that the mean coefficient of coancestry increased in the offspring population. This means that in future generations, there could be considerable loss of genetic variability due to inbreeding.

In this study, we used molecular coancestry for estimation of relatedness and coancestry as explained in Oliehoek et al. (2006). In their paper, the authors tested several relatedness estimators based on genetic similarity under different population structures, numbers of loci and of alleles. Among estimators compared, molecular coancestry turned out to be relatively robust. We therefore decided to use this estimator.

To our knowledge this is the first time where calculation of average coancestry in offspring using molecular similarity (Oliehoek et al., 2006) and parental contributions (Meuwissen, 1997) is published in aquaculture literature before. Parental contributions

were used by Brown et al. (2005) for estimating change in level of inbreeding, but they did not include relatedness estimators. More commonly, traditional population genetic analyses are used (Perez-Enriquez et al., 1999; Jackson et al., 2003; Porta et al., 2007). Here, the number of alleles e.g., is frequently used as measure of genetic diversity. However, these analyses do not take into account effect of parental contributions (if present) and therefore might not be accurate. Skewed contributions of parents, when combined with low polymorphism of loci or with high number of parents participating, might by chance lead to few alleles being lost. Loss of few alleles generally is interpreted as a limited decrease of genetic diversity, whereas in fact many alleles in the offspring population can be *identical by descent*. Thus the mean coancestry will be high. Apparently, traditional population genetic analyses do not fully describe development of genetic diversity, especially in case of selected populations. The method in our paper puts together relatedness between parents and parental contributions to future generations to predict levels of coancestry.

Consequences for broodstock management

Excessive inbreeding has been shown to lead to inbreeding depression in fertility traits (Frommen, 2008) and commercial traits (Gallardo, 2004). Clearly, attempts should be made to avoid inbreeding. By analyzing samples of every single batch produced, our study covered maximum available genetic diversity within broodstocks. Despite this extensive range, increase in mean coefficients of coancestry exceeded 1% per generation. With random mating and equal contributions, the mean coefficient of coancestry equals the mean level of inbreeding in the next generation. In this study mean levels of inbreeding would therefore exceed 1% under random mating. This is too high if one wishes to preserve fitness and prevent inbreeding depression (Bijma, 2000; Dupont-Nivet et al., 2006). Levels of inbreeding are determined by both broodstock population sizes and difference in parental contribution. When calculating level of inbreeding based on broodstock population size alone and assuming random mating and unselected populations (Falconer and Mackay, 1996) values of 1.8% are found for broodstock A and 2.5% are found for broodstock B, indicating that both broodstocks are too small. In broodstock A the estimated level of inbreeding according to the method of Meuwissen (1997) is 4.9%. This is higher compared to inbreeding based on broodstock population size alone. In broodstock B the level of inbreeding is more in line with expectations under

random mating, 2.0%, which is probably due to more equal contributions from parents to offspring.

It can be concluded that when implementing natural mating of Sole in groups for a breeding program, levels of inbreeding should be reduced by restricting parental contributions to the next generation. This can be done by increasing broodstock population sizes, genotyping selection candidates and composing broodstocks based on restriction of molecular coancestry.

Acknowledgements

This study was funded by Casimir NWO (the Netherlands Organization for Scientific research), Solea bv and Wageningen University. The authors are greatly indebted to the team of Solea and the ABGC laboratory technicians for their assistance. We thank Piter Bijma for his helpful suggestions.

References

Baynes, S. M., Howell, B. R., Beard, T. W. and Hallam, J. D., 1994. A description of spawning behaviour of captive dover sole, Solea solea (L.). Netherlands Journal of Sea Research 32, 271.

Bekkevold, D. X., 2006. Male size composition affects male reproductive variance in Atlantic cod Gadus morhua L. spawning aggregations. Journal of fish biology 69, 945.

Bekkevold, D. X., Hansen, M. M. and Loeschcke, V., 2002. Male reproductive competition in spawning aggregations of cod (Gadus morhua , L.). Molecular ecology 11, 91.

Bijma, P., 2000. Long-term genetic contributions: prediction of rates of inbreeding and genetic gain in selected populations. Wageningen.

Bijma, P., Van Arendonk, J. A. M. and Woolliams, J. A., 2000. A General Procedure for Predicting Rates of Inbreeding in Populations Undergoing Mass Selection. Genetics 154, 1865-1877.

Billard, R. R., 1992. Reproduction in rainbow trout: sex differentiation, dynamics of gametogenesis, biology and preservation of gametes. Aquaculture 100, 263.

Billard, R. R., 1995. Biology of sperm and artificial reproduction in carp. Aquaculture 129.

Brown, C. R., Woolliams, J. A. and McAndrew, B. J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, Sparus aurata. Aquaculture 247, 219.

Caballero, A. and Toro, 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. Genetical research 75, 331-343.

Chereguini, O., de la Banda, I. G., Rasines, I. and Fernandez, A., 1999. Artificial fertilization in turbot, Scophthalmus maximus (L.): different methods and determination of the optimal sperm-egg ratio. Aquaculture Research 30, 319-324.

Dinis, M. T., 1999. A review on the cultivation potential of Solea senegalensis in Spain and in Portugal. Aquaculture 176, 27.

Duchesne, P., Godbout, M. H. and Bernatchez, L., 2002. PAPA (Package for the Analysis of Parental Allocation) : A computer program for simulated and real parental allocation. Molecular Ecology Notes 2, 191-194.

Dupont-Nivet, M., Vandeputte, M., Haffray, P. and Chevassus, B., 2006. Effect of different mating designs on inbreeding, genetic variance and response to selection when applying individual selection in fish breeding programs. Aquaculture 252, 161-170.

Eding, H. and Meuwissen, 2001. Marker-based estimates of between and within population kinships for the conservation of genetic diversity. Journal of animal breeding and genetics 118, 141-159.

Exadactylos, A., 1999. Growth and genetic variation in hatchery-reared larval and juvenile Dover sole, Solea solea (L.). Aquaculture 176, 209-226.

Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to Quantitative Genetics. Fourth Edition. Pearson Prentice Hall Harlow.

Fessehaye, Y., El-Bialy, Z., Rezk, M. A., Crooijmans, R., Bovenhuis, H. and Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (Oreochromis niloticus) in breeding hapas: a microsatellite analysis. Aquaculture 256, 148-158.

Frommen, J. G., 2008. Inbreeding depression affects fertilization success and survival but not breeding coloration in threespine sticklebacks. Behaviour 145, 425-441.

Gallardo, J., 2004. Inbreeding and inbreeding depression of female reproductive traits in two populations of Coho salmon selected using BLUP predictors of breeding values. Aquaculture 234, 111-122.

Garoia, F., Marzola, S., Guarniero, I., Trentini, M. and Tinti, F., 2006. Isolation of polymorphic DNA microsatellites in the common sole Solea vulgaris. Molecular Ecology Notes 6, 144-146.

Gjerde, B., Gjoen, H. H. M. and Villanueva, B. B., 1996. Optimum designs for fish breeding programmes with constrained inbreeding: Mass selection for a normally distributed trait. Livestock production science 47, 59.

Goos, H. and Richter, C., 1995. Internal and external factors controlling reproduction in the African catfish, *Clarias gariepinus*. Aquatic living resources/Ressources vivantes aquatiques 9, 45-58.

Henderson, C. R., 1984. Applications of Linear Models in Animal Breeding. Guelph Univ. Press, Guelph, Canada.

Howell, B., Cañavate, P., Prickett, R. and Conceição, L., 2006. The Cultivation of Soles. Report of a 3rd workshop held at Cifpa El Toruño, Cadiz, Spain.

Imsland, A. K., Foss, A., Conceiçao, L. E. C., Dinis, M. T., Delbare, D., Schram, E., Kamstra, A., Rema, P. and White, P., 2003. A review of the culture potential of Solea solea and S. senegalensis. Reviews in fish biology and fisheries 13, 379.

Iyengar, A., Piyapattanakorn, S., Stone, D. M., Heipel, D. A., Howell, B. R., Baynes, S. M. and Maclean, N., 2000. Identification of Microsatellite Repeats in Turbot (Scophthalmus maximus) and Dover Sole (Solea solea) using a RAPD-Based Technique: Characterization of Microsatellite Markers in Dover Sole. Marine Biotechnology 2, 49-56.

Jackson, T. R., Martin-Robichaud, D. J. and Reith, M. E., 2003. Application of DNA markers to the management of Atlantic halibut (Hippoglossus hippoglossus) broodstock. Aquaculture 220, 245-259.

Komen, J., Bovenhuis, H. and Van Arendonk, J. A. M., 2006. Consequences of reproductive characteristics for fish breeding schemes. 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte-MG, Brazil, August 13-18, 2006.

Li, C. C. and Horvitz, D. G., 1953. Some methods of estimating the inbreeding coefficient. Amer J hum Genet 5, 107-117.

Lynch, M., 1988. Estimation of relatedness by DNA fingerprinting. Molecular Biology and Evolution 5, 584-599.

Malécot, G., 1948. Les Mathématiques de l'Hérédité. Masson, Paris.

Meuwissen, T. H. E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. Journal of animal science 75, 934-940.

Oliehoek, P. A., Windig, J. J., van Arendonk, J. A. M. and Bijma, P., 2006. Estimating Relatedness Between Individuals in General Populations With a Focus on Their Use in Conservation Programs. Genetics 173, 483-496.

Perez-Enriquez, R., Takagi, M. and Taniguchi, N., 1999. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (Pagrus major) used for

stock enhancement, based on microsatellite DNA markers. Aquaculture 173, 413-423.

Porta, J., Porta, J. M., Martínez-Rodríguez, G. and del Carmen Alvarez, M., 2006a. Development of a microsatellite multiplex PCR for Senegalese sole (Solea senegalensis) and its application to broodstock management. Aquaculture 256, 159-166.

Porta, J., Porta, J. M., Martinez-Rodriguez, G. and Alvarez, M. C., 2006b. Genetic structure and genetic relatedness of a hatchery stock of Senegal sole (Solea senegalensis) inferred by microsatellites. Aquaculture 251, 46.

Porta, J., Porta, J. M., Canavate, P., Martinez-Rodriguez, G. and Alvarez, M. C., 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, 223-234.

Raymond, M. and Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86, 248-249.

Rowe, S. S., 2007. Nonrandom mating in a broadcast spawner: mate size influences reproductive success in Atlantic cod (Gadus morhua). Canadian journal of fisheries and aquatic sciences 64, 219.

Sanchez, L., Bijma, P. and Woolliams, J. A., 2003. Minimizing inbreeding by managing genetic contributions across generations. Genetics 164, 1589-1595.

Weir, B. S. and Cockerham, C. C., 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution 38, 1358-1370.

3

Effects of grading on heritability estimates under commercial conditions

Robbert J.W. Blonk
Hans Komen
Andries Kamstra
Johan A.M. van Arendonk

Published in Aquaculture 300 (2010): 43-49

Abstract

Breeding programs in aquaculture typically use family rearing of fry until tagging size is reached. After tagging, fish are preferably reared communally. However, communal rearing promotes social interactions such as competition for access to food. In commercial conditions, negative social effects are reduced through grading of similar sized fish into one tank. In breeding programs with selection on size related traits, grading of communally reared untagged fish potentially introduces bias in estimates of genetic parameters as tank effects might be confounded with genetic potential.

We evaluated effects of grading on estimated heritability of the graded trait body length using data from simulated populations with different heritabilities and a "real" dataset of common sole, *Solea solea*. Datasets were sorted into eight "classes" of equal "size range". Six size ranges were chosen to represent different time points after grading. Here it was assumed that after grading, size range of classes increases with time. Consequently, overlap of classes increases, diluting confounding tank effects and genetic effects. Ranges of classes varied from 100% total size range (full overlap of classes) to 12.5% total size range (no overlap of classes). In addition, the real dataset was analyzed with the observed size distribution of animals over tanks. In the simulated datasets heritability of body length was underestimated in similar patterns when the size range of classes comprised less than 65% of the total size range. In the real dataset, estimated heritability for body length was 0.28 (±0.11). However, extrapolation with results from simulation showed that a heritability of 0.40 (±0.13) can be expected. Grading had a minor effect on estimated heritability of the correlated trait body weight (h^2 = 0.21 ±0.09). This study demonstrates that upper estimates of heritabilities can be obtained in conditions with grading when size distributions of fish over tanks are known.

Introduction

Breeding programs in aquaculture typically use family rearing of fry until tagging size is reached. After tagging, fish are mixed and reared in tanks or cages until selection. Breeding values and other genetic parameters are then estimated using BLUP procedures. Without individual (tag) information, fish are preferably raised and grown out communally to minimize tank effects and to provide more accurate estimates of genetic parameters (Wilson and McDonald, 2003; Vandeputte et al., 2004; Saillant et al., 2007; Vandeputte et al., 2008).

However, communal rearing until harvest of fish promotes social interactions such as agonistic behaviour and competition for access to food. These interactions are widely recognized as a source of phenotypic variation, a cause of mortality in fish stocks (Fessehaye et al., 2006a) and a source of unexpected responses in breeding programs (Bijma et al., 2007). To reduce social interactions and to improve growth in commercial grow out conditions, many fish species are graded on size into separate tanks (Barki et al., 2000; Dou et al., 2004; Ahvenharju and Ruohonen, 2007). For example, in the breeding program "PROSPER" multiple grading with rechallenging of selection candidates was implemented to minimise bias due to social interactions when selecting on own phenotype (Chevassus et al., 2004). However, when performing BLUP to estimate genetic parameters as heritability, grading of untagged selection candidates can lead to biased estimates as tank effects are likely to be confounded with genetic effects. Little is known about the extent to which grading affects estimated heritability when untagged selection candidates are graded.

In this paper we explore the consequences of grading on heritability estimates in a commercial population of common sole, *Solea solea*. To minimize detrimental effects of social interactions, commercial populations are graded. Besides, in common sole, parents typically spawn naturally in groups of 20 to 30 animals and produce batches of fry with mixed families. In these cases, family rearing followed by tagging is not possible. Farmers are forced to raise offspring with communal rearing and pedigrees are only reconstructed from molecular data at time of selection.

To evaluate effects of grading on estimated heritability of the graded trait body length, we analyzed simulated populations with different distributions of animals over tanks to

approximate distributions at different time points after grading. Here it was assumed that after grading, the phenotypic range of tanks increases with time as graded cohorts of fish are likely to rebuild variance. Consequently, overlap of classes increases. This effect dilutes confounding of tank effects and genetic effects.

To determine if effects of grading derived from simulations are consistent with "real life", we analyzed a commercial population of animals alike. The major difference between simulations and real data was their mating design: simulated populations had balanced mating designs. Populations produced by natural spawning of broodstocks typically yield unbalanced family structures (Brown et al., 2005; Blonk et al., 2009) and might therefore behave differently. Results show that upper estimates of heritabilities can be obtained when size distribution of animals over tanks are known.

Materials and Methods

Simulated dataset

Simulated populations with one generation offspring and two correlated traits "body length" and "body weight" were created using stochastic simulation in FORTRAN. Offspring phenotypes of both traits were constructed using parental breeding values which were simulated from a bivariate distribution. Parameters for simulations were as follows: heritabilities (h^2) for both traits: 0.15, 0.3 and 0.45; phenotypic and genetic correlations ($\rho_p = \rho_a$): 0.50 and 0.9. Each simulation included a full factorial mating design with 30 sires and 30 dams producing ten offspring per mating resulting in 9000 offspring. Phenotypic variances (σ^2_p) for both traits were set to 1. For each set of parameters ten simulations were run using a random seed.

Sire and dam breeding values for body length were simulated as $z\sqrt{\left(h^2 \cdot \sigma_p^2\right)}$ with z as a normal deviate generated from a simulated normal distribution *(N~(0,1))* and σ^2_p as phenotypic variance. Offspring phenotypic record P_k for body length of offspring k resulted from sire and dam breeding values plus a Mendelian sampling term and a random error term as follows:

$$P_k = \frac{1}{2}A_i + \frac{1}{2}A_j + z_1 \cdot \sqrt{\frac{1}{2} \cdot h^2 \cdot \sigma_p^2} + z_2 \cdot \sqrt{\left(1 - h^2\right) \cdot \sigma_p^2}$$

where A_i and A_j are breeding values for sire *i* and dam *j* and z_1 and z_2 normal deviates generated from a simulated normal distribution *(N~(0,1))*. Subsequently, the obtained values for the trait body length were used to simulate breeding values and phenotypes for the correlated trait body weight. Parents were assumed to be non inbred.

Real dataset

Two common sole broodstock populations A (*n*=28) and B (*n*=23) were kept indoor in separate tanks for natural spawning. Spawning in broodstock A commenced in June and lasted until October 2004. Spawning of broodstock B lasted from January until May 2005. Per broodstock, batches of larvae were reared in separate tanks. After weaning, offspring from both broodstocks were mixed as soon as possible and grown out in eight tanks (raceways). Animals were graded and redistributed on regular basis by manual grading into the eight tanks. The exact schedule for grading was not known. The criterion for grading was body length.

In January 2008, 1338 animals with pedigree obtained by pedigree reconstruction (described in Blonk et al. (2010a)) were randomly sampled from the eight tanks. The sampled group of offspring consisted of 59 full sib families, with a highly skewed contribution of parents; six parental pairs produced 70% of the offspring. In total, 21 sires and 17 dams contributed to the sampled offspring. All animals were anaesthetized with 2-phenoxyethanol and measurements were taken by three teams of two persons. From each animal body weight (g), body length (mm), tank and processing team was recorded. Gender of each animal was examined using ultrasound (System: Esaote Pie Medical MyLab30Vet; Transducer: Esaote LA435 6-18 MHz) by a seventh person.

Sorting

Both simulated data and real data were sorted on body length into eight “classes” of equal “size range”. Six “size ranges of classes” were chosen to approximate distributions of animals over tanks at different time points after grading (figure 3.1). Here it was assumed that after grading, size range of classes increases with time as cohorts of animals in graded classes have chance to rebuild variance; larger classes imply a longer period. Consequently, overlap of classes increases. This effect dilutes confounding of tank effects and genetic effects. To avoid bias of selection when estimating heritabilities, no animals were discarded.

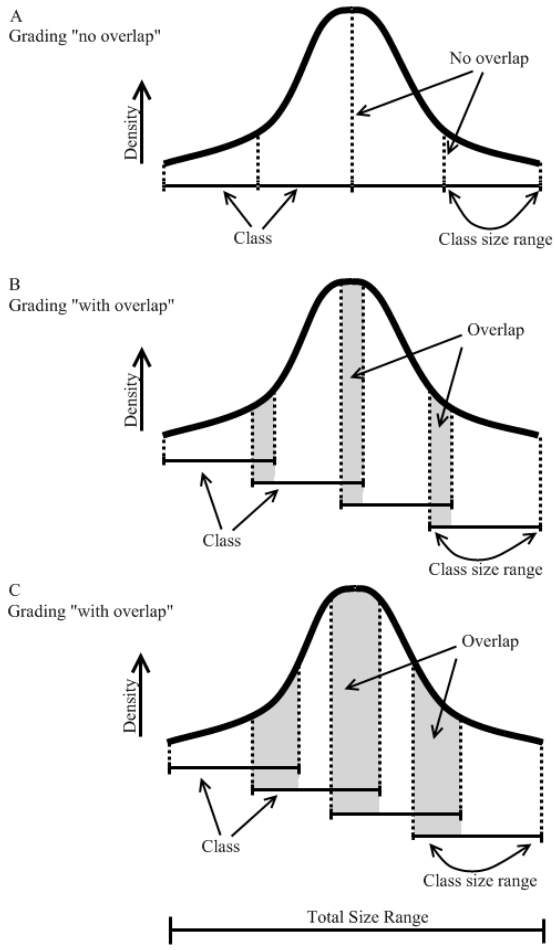


Figure 3.1 Example of simulated grading of a population on phenotype (body length) with four classes of equal class size range. A: class size range = 25%; B: class size range = 30%; C: class size range = 35%.

When sorting a dataset into eight classes, the smallest possible size range of classes is $1/8^{th}$ or 12.5% of the total size range. Size range of classes was expressed as a percentage of the total size range. In this study, six different size ranges of classes were simulated: 100, 82.5, 65.0, 47.5, 30.0 and 12.5% of the total size range. With a size range of classes comprising 12.5% of the total size range there is no overlap of classes whereas a size range of 100% implies that all classes fully overlap. Intermediate situations had partly

overlapping classes. Essentially, a size range of classes of 100% implies random allocation of fish to eight tanks. In practical situations, this would not occur. However, we included this size range to gain insight in extreme situations.

Animals that could be sorted into multiple classes due to overlap were assigned to an applicable class at random. Sorting was done with R software (R Development Core Team, 2008).

Statistical analysis

Average body length, body weight and numbers of animals in the real dataset are shown per tank in table 3.1. No significant effect of broodstock was found (P = 0.2490).

Table 3.1 Number of animals, mean body weight (BW) in g and mean body length (BL) in mm with standard deviations (sd) per tank in real data.

		BW		BL	
Tank	**n**	**mean**	**sd**	**mean**	**sd**
A	182	122.3	45.0	20.2	2.1
B	94	134.3	78.9	20.7	3.4
C	160	141.1	39.8	21.6	2.1
D	284	169.0	61.7	22.9	2.7
E	175	171.4	58.4	22.6	2.2
F	111	220.3	100.9	24.0	3.4
G	161	256.0	88.8	25.9	2.8
H	171	294.2	104.1	26.5	2.8

For genetic analysis, heritability, phenotypic and genotypic correlations were estimated with ASReml (Gilmour et al., 2006) using a bivariate model with some adjustments for simulated data:

$$y_{ijkl,BL}\, y_{ijkl,BW} = \mu + Team_i + Tank_j + Gender_k + u_l + \varepsilon_{ijkl}$$

where y is response variable body length and body weight, μ is mean, $Team_i$, is fixed effect of team i, $Tank_j$ is fixed effect of tank j (corresponds to class j with sorting), $Gender_k$ is fixed effect of gender k, u_l is random animal effect of animal l and ε is the random error term. Due to non normal distribution of residuals, bodyweight (BW) was log-transformed before sorting.

Heritabilities (h^2) were calculated from the additive genetic variance (V_A) and phenotypic variance (V_P) obtained from ASREML as $V_A/V_P = h^2$. For analysis of sorting effects in simulated data, team and gender effects were excluded from the model. The real dataset was also analyzed using the observed size distribution of animals over tanks. Here tank effects corresponded to the tank from which animals were sampled.

Results

Sorting

Size distributions of animals over classes after sorting are shown for simulated data in figure 3.2. Overlap of classes after sorting reduced with decreasing size range of classes (figure 3.2A-F). In figure 3.2A e.g., all eight classes fully overlap and animals are equally distributed over classes. With the smallest size range of classes i.e. 12.5% of the total size range (figure 3.2F), there is no overlap. Results after sorting of real data showed very similar patterns and are therefore not shown. With the observed size distribution of animals over tanks in real data (figure 3.3), overlap between tanks was large and looked similar to sorting with size ranges of classes approximately 65.0% to 47.5% of the total size range (figure 3.2C and D). Distribution of families over tanks in the real dataset is shown in figure 3.4. It can be seen that large families are found in multiple tanks, while small families are present in only few tanks.

Statistical analysis

Results of heritability estimation in simulated data with heritability 0.3 are shown in table 3.2. It can be seen that heritability is underestimated after sorting with decreasing size range of classes. However, estimated heritabilities remained relatively stable when sorting with size ranges of classes 100% to 65% of the total size range. Here, estimated

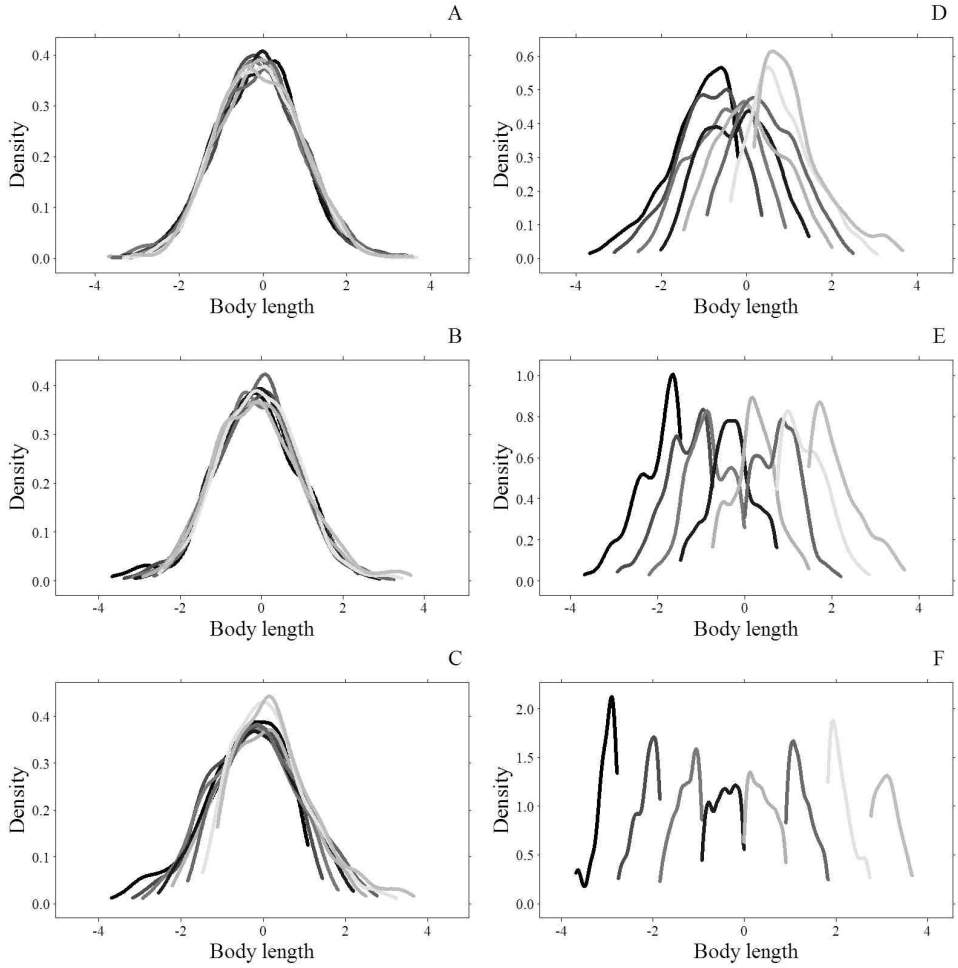


Figure 3.2 Simulated distribution of a population of animals in eight classes of equal size range, according to body length. In A) each class comprises the total size range of the population. In this case there is full overlap in classes. In F) each class covers a distinct range comprising 12.5% (1/8) of the total size range of the population (no overlap in classes). Figure B, C, D, and E represent intermediate situations with partly overlapping size classes comprising 82.5% (B), 65.0% (C), 47.5% (D) and 30.0% (E) of the total size range in the population. Populations were simulated assuming h^2 *for body length 0.3 and* r_g *and* r_p *between body length and weight 0.90.*

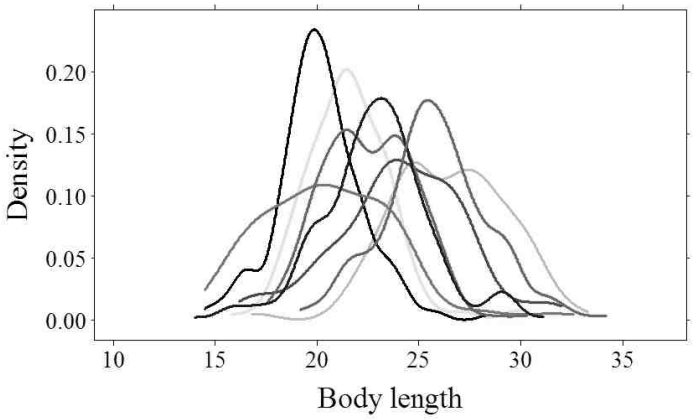
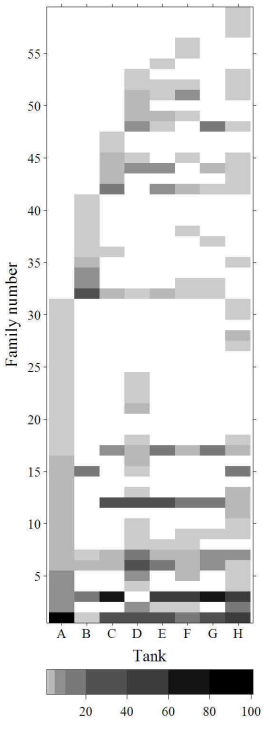


Figure 3.3 Observed distribution of body length (in cm) over eight tanks in the commercial population ("real dataset").

Figure 3.4 Observed distribution of families (n = 59) over tanks (A : H) in the real dataset. The colour key indicates different numbers of sibs in a tank. The data was ordered on decreasing number of sibs/family per tank (from bottom to top), starting with A, followed by tank B, C, etc. For example, family no. 1 is present with the largest number of sibs (>100) in tank A, and also present in all other tanks. Family no. 5 is present in tank A with 20-40 sibs and also present in tanks D, F and H, but not in tanks B, C, E and G. Family no. 57, 58 and 59 were only observed in tank H.



heritability after sorting was 0.30 (±0.01) to 0.29 (±0.01) for body length, which was not different from the simulated heritability. With size range of classes 30.0% and 12.5% of the total size range, heritability estimates ranged from 0.10 (±0.02) to 0.03 (±0.01). Here, overlap between classes was small or absent causing confounding of tank effects with genetic effects.

Figure 3.5 and 3.6 show similar patterns of estimated heritabilities in simulated datasets with heritability 0.15 and 0.45 (3.5A) and when correlations were set to 0.50 (3.6A). Mean standard errors in simulated datasets with heritabilities 0.15 and 0.45 ranged from 0.01 to 0.03; approximately similar to those of simulations with heritability 0.3. For this reason ten replications per simulation were considered enough.

Estimated genetic and phenotypic correlations with simulated heritability 0.3 remained as defined in simulations with a size range of classes of 100% to 65.0% (see table 3.2). However, levels dropped when size range of classes decreased. Genetic correlations reached 0.22 with a high standard error (±0.15) with simulated correlations of 0.90. Phenotypic correlations dropped to approximately 0.50 (±0.02). Estimated correlations in simulations with heritability 0.15 and 0.45 and with correlations set on 0.5 showed the same patterns and standard errors (results not shown).

Table 3.2 Mean estimated heritabilities (h^2), genetic (r_g) and phenotypic (r_p) correlations ± mean standard error (and standard deviation of the parameter over ten replications) for bivariate models with body length (BL) and body weight (BW) in simulated data, after sorting with size ranges of classes 100% to 12.5% of the total size range. Populations were simulated assuming h^2 for BL 0.3 and r_g and r_p between BL and BW 0.90.

	100%	**82.5%**	**65.0%**	**47.5%**	**30.0%**	**12.5%**
h^2 BL	0.30 ±0.03 (0.01)	0.30 ±0.03 (0.01)	0.29 ±0.03 (0.01)	0.20 ±0.02 (0.02)	0.10 ±0.02 (0.02)	0.03 ±0.01 (0.01)
h^2 BW	0.31 ±0.03 (0.01)	0.31 ±0.03 (0.01)	0.30 ±0.03 (0.01)	0.24 ±0.02 (0.02)	0.20 ±0.02 (0.02)	0.24 ±0.02 (0.02)
r_g	0.91 ±0.01 (0.01)	0.91 ±0.01 (0.01)	0.90 ±0.01 (0.01)	0.81 ±0.03 (0.03)	0.55 ±0.07 (0.05)	0.22 ±0.15 (0.11)
r_p	0.90 ±0.00 (0.00)	0.90 ±0.00 (0.00)	0.90 ±0.00 (0.00)	0.85 ±0.00 (0.01)	0.73 ±0.01 (0.02)	0.48 ±0.01 (0.02)

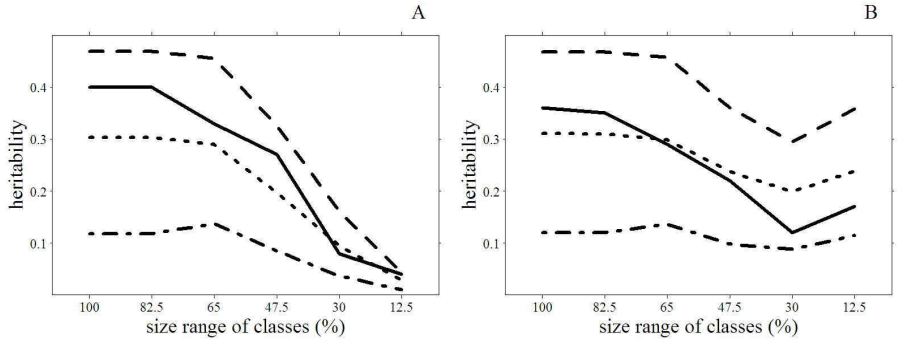


Figure 3.5 Heritability estimates for body length (A) and body weight (B) from a bivariate model with body length and body weight as correlated traits in simulated populations and a commercial population. All populations were sorted using different size ranges of classes. Heritabilities for both traits in simulated populations were 0.15 (— •), 0.3 (•••) and 0.45 (— —). Correlations r_g and r_p for body weight and length were 0.9. (———) = commercial population. Values on the x-axis represent size ranges of classes (100% to 12.5%) as percentages of the total size range (see figure 3.1 for explanation).

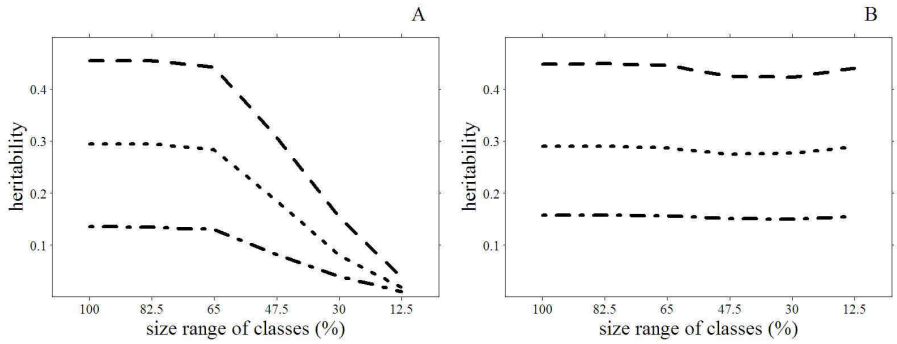


Figure 3.6 Heritability estimates for body length (A) and body weight (B) from a bivariate model with body length and body weight as correlated traits in simulated populations. All populations were sorted using different size ranges of classes. Heritabilities for both traits in simulated populations were 0.15 (— •), 0.3 (•••) and 0.45 (— —). Correlations r_g and r_p for body weight and length were 0.5. Values on the x-axis represent size ranges of classes (100% to 12.5%) as percentages of the total size range (see figure 3.1 for explanation).

Estimated heritabilities of the correlated trait (body weight) dropped less than estimated heritabilities for the graded trait (body length) in simulated datasets with correlations set on 0.90 (3.5B). When simulating correlations of 0.50, body weight heritability was not affected by sorting (3.6B). Mean standard deviations with simulated correlations of 0.9 ranged from 0.01 to 0.03, whereas with simulated correlations of 0.5 mean standard deviations ranged from 0.01 to 0.04. With sorting of real data, estimated heritabilities were between 0.40 (±0.13) and 0.33 (±0.12) in size range of classes 100% to 65.0% (table 3.3). Estimated heritabilities after sorting simulated data and real data are compared in figure 3.5A.

In accordance with results of simulated data, sorting of real data with size ranges of classes 30.0% and 12.5% of the total size range resulted in low heritabilities: 0.08 (±0.05) and 0.04 (±0.03). Genetic and phenotypic correlations (table 3.3) were estimated at 0.95 (±0.03) to 0.92 (±0.01) in a size range of classes of 100% to 65.0% of the total size range. However, genetic and phenotypic correlations dropped to lower levels as size ranges of classes decreased beyond 30.0%. Here genetic correlations approached zero with a high standard error (0.04 ±0.45) whereas phenotypic correlations remained 0.47 (±0.03). Estimated heritabilities for the correlated trait (body weight) in the real dataset decreased when the size range of classes was small (table 3.3), though to a larger extend than in simulated data with comparable correlations (figure 3.5B).

Table 3.3 Estimated heritabilities (h^2), genetic (r_g) and phenotypic (r_p) correlations ± standard error for bivariate models with body length (BL) and body weight (BW) in real data with observed size distribution and sorting with size ranges of classes 100% to 12.5% of the total size range.

	Actual dist	**Sorting**					
		100.0%	**82.5%**	**65.0%**	**47.5%**	**30.0%**	**12.5%**
h^2 BL	0.28 ±0.11	0.40 ±0.13	0.40 ±0.13	0.33 ±0.12	0.27 ±0.11	0.08 ±0.05	0.04 ±0.03
h^2 BW	0.21 ±0.09	0.36 ±0.12	0.35 ±0.12	0.29 ±0.11	0.22 ±0.09	0.12 ±0.06	0.17 ±0.07
r_g	0.89 ±0.07	0.95 ±0.03	0.95 ±0.03	0.94 ±0.04	0.87 ±0.08	0.47 ±0.30	0.04 ±0.45
r_p	0.87 ±0.01	0.93 ±0.01	0.93 ±0.01	0.92 ±0.01	0.87 ±0.01	0.76 ±0.02	0.51 ±0.03

In the real dataset with the observed size distribution of animals over tanks, estimated heritability for body length was 0.28 (±0.11) (table 3.3). Estimated heritability for body weight was 0.21 (±0.09). Genetic and phenotypic correlations were respectively 0.89 (±0.07) and 0.87 (±0.01).

Discussion and Conclusion

Effect of sorting

In this study we explored consequences of grading on heritability estimates of a trait on which grading was based (body length) and on a correlated trait (body weight) in common sole. As expected, heritability for body length was underestimated when the data was sorted into considerable small size ranges of classes. This was found after sorting both simulated data and real data. The reason for decrease of estimated heritability with smaller size ranges of classes i.e. with genetic parameter estimation after relatively recent grading, is that genetic effects for body length become increasingly confounded with tank effects. Heritabilities for body length drop particularly when relatively small size ranges of classes (< 65.0% of the total size range) were used (figure 3.5).

Estimated heritabilities of traits with high heritability drop relatively fast after grading than traits with low heritability. This is explained by the fact that between family variance at high heritabilities is larger (Lynch and Walsh, 1998); families with larger body length are more likely to end up in the same tank. Consequently, tank effects are confounded with family, i.e. genetic effects.

The magnitude to which correlated traits are affected by grading on another trait, depends on the degree of correlation between the two traits. With high correlations such as 0.90 (figure 3.5B), heritabilities of the correlated trait were underestimated. However, in simulations with initial correlations of 0.50 (figure 3.6B), no effect of sorting on estimated heritability of correlated traits was found. Nevertheless, in all cases underestimation of heritability due to high correlations was relatively small.

Yet, the shown effect is relevant especially for species which are graded on traits that have high correlations with economically important traits as growth or bodyweight at harvest. For example, in Nile tilapia, *Oreochromis niloticus*, Rutten *et al.* (2005) showed

high genetic and phenotypic correlations (approximately 0.9) between head width and body weight at harvest. Meanwhile, in several grading machines or devices animals are graded on head width. Both phenotypic and genotypic correlations dropped when relatively small size ranges of classes were used (< 65.0%). Decrease of correlations after selection becomes clear when realizing that correlation in a selected subset drops compared to the correlation in the full dataset. A similar effect is described by Cochran (1951).

It was assumed that sorting animals with class size ranges approximates distributions at different time points after grading. However, distributions of families over tanks from sorting with large class size ranges might differ from distributions observed after real grading. In our method, animals are distributed to tanks to simulate subsequent time points after grading. Hence, distribution of families over tanks is different for each size class range and thus for different for time points. However, in reality representation of families in tanks is fixed from the moment of grading. Theoretically, confounding of families with tanks at later time points after grading will therefore be stronger than in our simulations. This is especially the case with high heritabilities, were between family variance is large. However, in practice, families show relatively equal distribution over tanks at commonly observed heritabilities. A reason for this is that grading in real situations is never perfect. Farmers usually aim at equal stocking densities and “tails” of distributions, if not discarded, will be mixed with other size classes. Moreover, the “middle” of a normal distribution contains say 60 to 70% of the population and here, relatively many families are presented. Grading these animals into tanks leads to overlapping size classes and restricts confounding of families with tanks. Effect of number of size classes was not tested in this study. Very likely, their effects are in line with effects of class size ranges.

With more classes, class size ranges decrease as they need to cover the total size range. Consequently estimated heritabilities will decrease. When grading with fewer classes, the effect is reversed.
In the “PROSPER” breeding program, beneficial effects of multiple grading and rechallenging on selection response in trout have been shown (Chevassus et al., 2004). Here, after a growth challenge, “small” size classes are culled and “large” and “medium”

classes are recombined to similar size classes. After this the recombined groups are challenged again. This is repeated until the desired selection intensity is reached. With this method, animals are selected on own phenotype while negative effects of social interactions are minimized. Our study shows that estimation of genetic parameters also can be done within populations that are graded. However, with the aim of estimation of breeding values and heritabilities, culling of smaller size classes, as is common practice at many farms, should be minimized. Discarding specific size classes introduces selection of data and this may provoke biased estimation of genetic parameters (Falconer and Mackay, 1996).

Heritability real data

The observed size distribution of animals over tanks in real data (figure 3.3 and 3.4) appeared comparable to size distributions after sorting with size ranges of classes approximately 65.0% to 47.5% of the total size range (figure 3.2C and D). Therefore, it is expected that heritability for body length as estimated in this study with real data (h^2 = 0.28 (± 0.1)) was underestimated. Following the trend of results found with sorting of simulated data and real data (figure 3.5A), it is expected that heritability for body length in common sole is likely to be 0.40. This means that the heritability was underestimated by approximately 30%.

Analysis of real data resulted in an estimated heritability of 0.21 (± 0.09) for body weight (non-graded trait) with high genetic and phenotypic correlations (table 3.3). From simulations with correlations of 0.90 and similar size distributions of animals over tanks (i.e. similar size range of classes), it can be inferred that the observed heritability of the correlated trait is also likely to be underestimated. Extrapolation of the estimated heritability in figure 3.5B would lead to a “true” heritability of approximately 0.36. We are the first to report estimated heritabilities of common sole, so no comparison with other data can be made. However, the observed value is well within the range of values found for other species (Gjedrem, 2000).

Implications

Response to selection may be dependent on circumstances under which animals are selected, especially when selection environments and commercial environments are

different. The magnitude of GxE interactions is estimated from the correlation between genetic values for a trait in different environments.

In pigs, observed correlations between breeding values for growth in test farms and commercial farms were between 0.3 and 1.0 (Merks, 1988). In tilapia, *Oreochromis niloticus*, no significant GxE interactions were found between two farming conditions: test or "high input", and commercial or "low input" (Khaw et al., 2009) whereas in European sea bass, *Dicentrarchus labrax*, GxE interactions, although low, were observed (Dupont-Nivet et al., 2008). Although dependent on species and environment, GxE interactions may lead to different ranking of animals. Consequently, realized selection responses in commercial environments may be different from expected when selection is under a testing environment rather than under the same commercial environment (Lynch and Walsh, 1998). This means that it can be necessary to either implement grading into testing environments of breeding programs or to select parents from production stocks directly and account for effects of grading on genetic parameter estimation afterwards.

Estimation of genetic parameters in graded populations may also be relevant from a perspective of cost efficiency. Especially in slow growing species, it can be more efficient for breeding programs to select parents from the production stock itself. As these stocks generally are graded (Barki et al., 2000; Dou et al., 2004; Martins et al., 2005; Ahvenharju and Ruohonen, 2007), estimation of genetic parameters is dependent on this type of data. In breeding programs with natural spawning species, BLUP procedures are entirely dependent on molecular genotyping of fish to verify pedigrees. Few cases are known where pedigrees are completely based upon genetic markers (Chavanne et al., 2008). Genotyping is expensive and the number of animals to be genotyped should preferably not be larger than the number of selection candidates. Animals are therefore mostly genotyped at time of selection i.e. when phenotypic information is collected. Until that time untagged selection candidates should be reared with grading to avoid detrimental social effects.

Our findings imply that estimation of heritabilities under commercial conditions with grading is likely to result in underestimated values of both graded and correlated traits. However, by reconstructing size distributions of animals over tanks it is possible to estimate the degree of underestimation and thus to extrapolate results to a more likely heritability. To estimate heritability of correlated traits, there should also be an indication

of correlations between the graded and non graded trait. Our study shows that estimating heritability and selecting under conditions with grading is possible.

Acknowledgements

This study was funded by Casimir NWO (the Netherlands Organization for Scientific research), Solea bv and Wageningen University. The authors thank team Solea and Bart Ducro for their support.

References

Ahvenharju, T. and Ruohonen, K., 2007. Agonistic behaviour of signal crayfish (*Pacifastacus leniusculus* Dana) in different social environments: Effect of size heterogeneity on growth and food intake. Aquaculture 271, 307-318.

Barki, A., Harpaz, S., Hulata, G. and Karplus, I., 2000. Effects of larger fish and size grading on growth and size variation in fingerling silver perch. Aquaculture International 8, 391-401.

Bijma, P., Muir, W.M. and Van Arendonk, J.A.M., 2007. Multilevel Selection 1: Quantitative Genetics of Inheritance and Response to Selection. Genetics 175, 277-288.

Blonk, R.J.W., Komen, H., Kamstra, A. and Van Arendonk, J.A.M., 2010. Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations of Common sole, Solea solea. Genetics 184, 1-7.

Blonk, R.J.W., Komen, J., Kamstra, A., Crooijmans, R.P.M.A. and 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.

Brown, C.R., Woolliams, J.A. and McAndrew, B.J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247, 219-225.

Chavanne, H., Norris, A., Haffray, P., Sonesson, A.K., Vandeputte, M., Chatain, B. and Boudry, P., 2008. Survey on the breeding practices in the European aquaculture industry. Reprofish Aquabreeding Workshop 2008.

Chevassus, B., Quillet, E., Krieg, F., Hollebecq, M.G., Mambrini, M., Faure, A., Labbe, L., Hiseux, J.P. and Vandeputte, M., 2004. Enhanced individual selection for selecting fast growing fish: the "PROSPER" method, with application on brown trout (*Salmo trutta fario*). Genetics Selection Evolution 36, 643-661.

Cochran, W.G., 1951. Improvement by means of selection. Proceedings of the 2nd Berkely symposium on mathematics, statistics and probability, edited by J. Neyman., University of California Press, Berkely.

Dou, S.-Z., Masuda, R., Tanaka, M. and Tsukamoto, K., 2004. Size hierarchies affecting the social interactions and growth of juvenile Japanese flounder, *Paralichthys olivaceus*. Aquaculture 233, 237-249.

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H. and Chatain, B., 2008. Heritabilities and GxE interactions for growth in the European sea bass (*Dicentrarchus labrax* L.) using a marker-based pedigree. Aquaculture 275, 81-87.

Falconer, D.S. and Mackay, T.F.C., 1996. Introduction to quantitative genetics. Pearson Prentice Hall Harlow, Essex.

Fessehaye, Y., Kabir, A., Bovenhuis, H. and Komen, H., 2006. Prediction of cannibalism in juvenile *Oreochromis niloticus* based on predator to prey weight ratio, and effects of age and stocking density. Aquaculture 255, 314-322.

Gilmour, A.R., Gogel, B.J., Cullis, B.R. and Thompson, R., 2006. ASReml User Guide Release 2.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Gjedrem, T., 2000. Genetic improvement of cold-water fish species. Aquaculture Research 31, 25-33.

Khaw, H.L., Bovenhuis, H., Ponzoni, R.W., Rezk, M.A., Charo-Karisa, H. and Komen, H., 2009. Genetic analysis of Nile tilapia (Oreochromis niloticus) selection line reared in two input environments. Aquaculture 294, 37-42.

Lynch, M. and Walsh, B., 1998. Genetic and Analysis of Quantitative traits. Sinauer Associates Inc., Sunderland, USA.

Martins, C.I.M., Aanyu, M., Schrama, J.W. and Verreth, J.A.J., 2005. Size distribution in African catfish (*Clarias gariepinus*) affects feeding behaviour but not growth. Aquaculture 250, 300-307.

Merks, J.W.M., 1988. Genotype x environment interactions in pig breeding programmes. Ph.D. Thesis, Landbouwuniversiteit Wageningen, Wageningen, the Netherlands.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, URL http://www.R-project.org.

Rutten, M.J.M., Bovenhuis, H. and Komen, H., 2005. Genetic parameters for fillet traits and body measurements in Nile tilapia (Oreochromis niloticus L.). Aquaculture 246, 125-132.

Saillant, E., Ma, L., Wang, X.X., Gatlin, D.M. and Gold, J.R., 2007. Heritability of juvenile growth traits in red drum (Sciaenops ocellatus L.). Aquaculture Research 38, 781-788.

Vandeputte, M., Kocour, M., Mauger, S., Rodina, M., Launay, A., Gela, D., Dupont-Nivet, M., Hulak, M. and Linhart, O., 2008. Genetic variation for growth at one and two summers of age in the common carp (*Cyprinus carpio* L.): Heritability estimates and response to selection. Aquaculture 277, 7-13.

Vandeputte, M., Kocour, M., Mauger, S., Dupont-Nivet, M., De Guerry, D., Rodina, M., Gela, D., Vallod, D., Chevassus, B. and Linhart, O., 2004. Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). Aquaculture 235, 223-236.

Wilson, A.J. and McDonald, 2003. Marker-assisted estimation of quantitative genetic parameters in rainbow trout, *Oncorhynchus mykiss*. Genetical Research 81, 145-156.

4

Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations

Robbert J.W. Blonk
Hans Komen
Andries Kamstra
Johan A.M. van Arendonk

Published in Genetics 184 (2010): 1-7

Abstract

Captive populations where natural mating in groups is used to obtain offspring typically yield unbalanced population structures with highly skewed parental contributions and unknown pedigrees. Consequently, for genetic parameter estimation, relationships need to be reconstructed or estimated using DNA marker data. With missing parents and natural mating groups, commonly used pedigree reconstruction methods are not accurate and lead to loss of data. Relatedness estimators however, infer relationships between all animals sampled. In this study, we compared a pedigree relatedness method and a relatedness estimator ("molecular relatedness") method using accuracy of estimated breeding values. A commercial dataset of common sole, *Solea solea*, with 51 parents and 1953 offspring ("full dataset") was used. Due to missing parents, for 1338 offspring a pedigree could be reconstructed with ten microsatellite markers ("reduced dataset").

Cross validation of both methods using the reduced dataset showed an accuracy of estimated breeding values of 0.54 with pedigree reconstruction and 0.55 with molecular relatedness. Accuracy of estimated breeding values increased to 0.60 when applying molecular relatedness to the full dataset. Our results indicate that pedigree reconstruction and molecular relatedness predict breeding values equally well in a population with skewed contributions to families. This is probably due to the presence of few large full sib families. However, unlike methods with pedigree reconstruction, molecular relatedness methods ensure availability of all genotyped selection candidates which results in higher accuracy of breeding value estimation.

Introduction

To estimate genetic parameters, additive genetic relationships between individuals are inferred from known pedigrees (Falconer and Mackay, 1996; Lynch and Walsh, 1997). However, in natural populations (Ritland, 2000; Thomas et al., 2002) and in captive species where natural mating in groups is used to obtain offspring (Brown et al., 2005; Fessehaye et al., 2006b; Blonk et al., 2009) pedigrees are reconstructed. In these populations there is no control on mating structure and typically unbalanced population structures with highly skewed parental contributions are obtained (Bekkevold et al., 2002; Brown et al., 2005; Fessehaye et al., 2006b; Blonk et al., 2009). To reconstruct pedigrees, parental allocation methods are often used (Marshall et al., 1998; Avise et al., 2002; Duchesne et al., 2002). These methods require that all parents are known. For situations where parental information is not available, numerous DNA marker based methods to estimate molecular relatedness have been developed (Lynch, 1988; Queller and Goodnight, 1989; Ritland, 2000; Toro et al., 2002). These relatedness estimators determine relationship values between individuals on a continuous scale. Evaluation of relatedness estimators within real and simulated data in both plants and animals (e.g. see Van de Casteele et al., 2001; Milligan, 2003; Oliehoek et al., 2006; Rodríguez-Ramilo et al., 2007; Bink et al., 2008) has generally focused on bias and sampling error of estimated genetic variances or relatedness values. Relatively little attention has been paid to their efficiency for estimation of breeding values.

Two types of relatedness estimators are currently available: method-of-moments estimators and maximum likelihood estimators. Method-of-moments estimators (e.g. Queller and Goodnight, 1989; Li et al., 1993; Ritland, 1996; Lynch and Ritland, 1999; Toro et al., 2002) determine relationships while calculating sharing of alleles between pairs in different ways. A variant of method-of-moments estimators is the transformation of continuous relatedness values to categorical genealogical relationships using “explicit pedigree reconstruction“ (Fernández and Toro, 2006) or thresholds (Rodríguez-Ramilo et al., 2007).

However, correlations of transformed coancestries with known genealogical coancestries are low (Rodríguez-Ramilo et al., 2007). Several studies have compared different

method-of-moments estimators but none revealed one single best estimator (Van de Casteele et al., 2001; Oliehoek et al., 2006; Rodríguez-Ramilo et al., 2007; Bink et al., 2008).

Maximum likelihood approaches classify animals to a limited number of relationship classes (Mousseau et al., 1998; Thomas et al., 2002; Wang, 2004; Herbinger et al., 2006; Anderson and Weir, 2007). For each pair a likelihood to fall into a possible relatedness class (e.g. full sib vs. unrelated) is calculated given their genotype and phenotype. Maximum likelihood techniques combined with a Markov Chain Monte-Carlo approach reconstruct groups with specific relationships jointly and are therefore more efficient than other ML approaches. To minimize standard errors, all discussed ML methods require balanced population structures, large sibling groups and large variance of relatedness (Thomas et al., 2002; Wang, 2004; Anderson and Weir, 2007). Therefore, these methods may not be suitable for natural mating systems.

Unlike parental allocation methods, a benefit from relatedness estimators is that essentially all selection candidates are maintained for breeding value estimation, even with missing parents. The question is however, if such relatedness estimators also give accurate breeding values in order to perform selection.

In this study, we test suitability of a relatedness estimator to obtain breeding values in a population of common sole, *Solea solea* (n = 1953) obtained by natural mating. First, we estimate breeding values using pedigree relatedness of animals for which a pedigree could be reconstructed (using parental allocation). This dataset (n = 1338) is further referred to as “reduced dataset”. We compare results with estimated breeding values using a simple but robust method-of-moments relatedness estimator: “molecular relatedness” (Toro et al., 2002; 2003). Next, we estimate breeding values using molecular relatedness in the full dataset (n = 1953). Results show that accuracies of estimated breeding values obtained with molecular relatedness and pedigree relatedness are comparable. Accuracy increases when breeding values are estimated with molecular relatedness in the full dataset. This implies that a molecular relatedness estimator can be used to estimate breeding values in captive natural mating populations.

Materials and methods

Dataset

Two thousand animals (age three years old) from a commercial production population were sampled randomly from eight tanks. All animals have been produced by natural spawning and families were mixed over tanks. The parents (n = 51) had been collected from the Dutch North-Western coastal area and were sampled at the same time as the offspring. A blood sample for genotypic analysis was taken from all offspring and available parents. However, no samples were available on six potential parents due to mortality before sampling. Of all offspring, body weight, gender, tank location and measuring team were recorded. All animals were genotyped with ten microsatellite markers: *AF173855*, *AF173854*, *AF173852*, *AF173849*, *AY950593*, *AY950592* (Iyengar et al., 2000), *AY950591*, *AY950589*, *AY950588*, *AY950587* (Garoia et al., 2006). See Blonk *et al.* (2009) for details on DNA isolation and PCR amplification protocols.

Reconstructed pedigree relatedness

From the offspring dataset, only individuals with more than five successfully amplified markers were used for analysis. Pedigree reconstruction was performed with PAPA 2.0 software (Duchesne et al., 2002). A uniform error of 0.02 on all markers was used. To set up a reliable pedigree, allocation results were checked as follows:

Offspring that were not allocated to any parent were removed from the dataset. Also, animals that were allocated to more than two parents due to equal likelihood of breeding of these parents (termed as “ambiguous” in PAPA) were removed from the dataset. Further, allocated parent-offspring pairs having the single highest likelihood of breeding among all other possible pairs were checked for consistency of Mendelian allele inheritance. Allocations were only considered correct when more than five markers were allocated consistently with Mendelian inheritance. Checking of allocation results was performed using R software (R Development Core Team, 2008). The remaining dataset was further referred to as “reduced dataset”.

Molecular relatedness

Genetic relationships between animals were also estimated using a relatedness estimator. To estimate molecular relatedness, we used the method described by Toro *et al.* (2002;

2003). Here, coancestry (*f*) is calculated from similarity (*S*) of alleles between two individuals *x* and *y*. Here $S_{xy,l}$ = ¼[I_{ac} + I_{ad} + I_{bc} + I_{bd}] where at locus *l*, *a* and *b* are alleles of individual *x* and *c* and *d* are alleles of individual *y* (Li et al., 1993). When at I_{ac} allele *a* is identical to allele *c*, I_{ac} equals one, and zero otherwise, etc. According to Lynch (1988) molecular coancestry is then calculated as:

$$f_{xy} = \frac{1}{L}\sum_{l=1}^{L}\frac{S_{xy,l} - s_l}{1 - s_l} \qquad 4.1$$

where *L* is the number of markers (markers) and s_l is mean similarity (sum of squared allele frequencies, Σp^2) at locus *l* in the base population. Estimated relatedness *r* between two animals is calculated as $r = 2f$. When ignoring alleles which are alike in state (AIS), molecular relatedness between two individuals is calculated from coancestry as follows:

$$r_{xy} = \frac{2}{L}\sum_{l=1}^{L} S_{xy,l} \qquad 4.2$$

Consequently, values of molecular relatedness are continuous and may range between zero (no alleles are similar) and two (all alleles are similar).

Genetic analysis

To estimate heritabilities and breeding values, genetic analysis was performed with relatedness inferred from either pedigree reconstruction or molecular relatedness. The restricted maximum likelihood method was used with ASReml software (Gilmour et al., 2006) for analysis of the following genetic univariate linear model:

$$BW_{ijkl} = \mu + T_i + Ta_j + G_k + u_l + \varepsilon_{ijkl} \qquad 4.3$$

where BW is the response variable for bodyweight, μ is the mean, T_i, is the fixed effect of team *i*, Ta_j is the fixed effect of tank *j*, G_k is the fixed effect of gender *k*, u_l is the random animal effect of animal *l* and ε is the random error term.

For estimation of breeding values from the pedigree reconstruction method, the relationship matrix was calculated within ASReml software using the reconstructed pedigree. The relationship matrix from the molecular relatedness method, however, was prepared separately and offered to ASReml as a generalized inversed matrix ("GIV-file"). To obtain convergence in ASReml, before inverting the relationship matrix was made positive definite using the bending method. Bending and inverting was done using R

software (R Development Core Team, 2008). Heritabilities where then calculated using the estimated additive genetic variance and the residual variance as $\sigma^2_A / (\sigma^2_A + \sigma^2_E)$.

Comparison of molecular relatedness and reconstructed pedigree relatedness

Pearson correlations were calculated between relatedness values as well as between breeding values obtained from molecular relatedness and reconstructed pedigree. To evaluate how well both methods fitted the model to the data, the accuracy in predicting the phenotype was calculated using cross validation of each method. For cross validation, the dataset was randomly divided into 100 equal sized subsets. Genetic analysis with the genetic model was then performed 100 times on all subsets jointly while successively omitting phenotype records for one subset at a time. In each analysis, for the subsets where observed phenotypes were omitted, phenotypes were predicted based on available observed phenotypes in the other subsets.

Observed and predicted phenotypes from genetic analysis were corrected for fixed effects (further referred to as "corrected phenotypes"). Accuracy of estimated breeding values (r_{HI}) was calculated as follows:

$$\frac{r_{\hat{P}P}}{\sqrt{h^2}} \qquad 4.4$$

where $r_{\hat{P}P}$ is the Pearson correlation coefficient between predicted corrected phenotypes and observed corrected phenotypes observed from cross validation and h^2 the heritability estimated using pedigree reconstruction. All calculations were performed using R (R Development Core Team, 2008).

Results

Datasets and relatedness

DNA isolation and marker analysis was successful, with only 30 samples giving no signal. Out of 2000 samples, finally 1953 animals could be genotyped with more than five markers. This is further referred to as the "full dataset". In the full dataset, per marker on average 14.3 alleles were found in the offspring. The parental dataset, per marker 13.4 alleles were found.

After parental allocation, all allocated parent-offspring pairs were tested for consistency of Mendelian inheritance on each marker. The number of markers following correct Mendelian inheritance patterns was counted for each pair. The frequency distribution of allocated parent-offspring pairs over the number of consistent markers is shown in figure 4.1. Forty five percent of the pairs were allocated with all ten markers. To construct the final pedigree, only parent-offspring pairs with more than five markers consistent with Mendelian inheritance patterns (66%) were taken into account. This resulted in a reconstructed pedigree for 1338 offspring, further referred to as "reduced dataset". In the reduced dataset, 21 males and 17 females contributed to the offspring producing 59 full sib families. Parental contribution was highly skewed with six parental pairs producing 70% of the offspring. In the reduced dataset, on average 13.7 alleles per marker were found in the offspring.

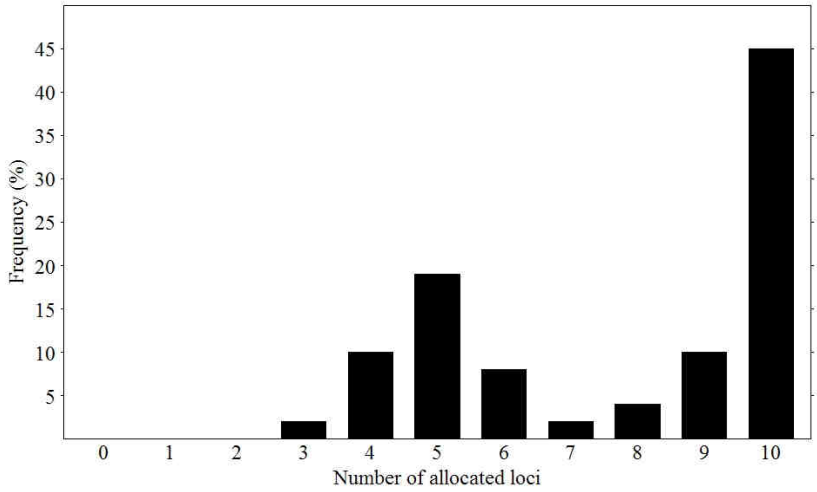


Figure 4.1 Frequency distribution of offspring allocated to two parents versus number of markers which were allocated following Mendelian inheritance after pedigree reconstruction in cultured common sole (full dataset, n = 1953).

Parents were assumed to be unrelated. Consequently relatedness values between offspring pairs after pedigree reconstruction were 0, 0.25 and 0.5. This corresponds to unrelated, half and full sib pairs. The average relatedness in the reconstructed pedigree was 0.1483 (σ_{PR} = 0.1865). As expected, molecular relatedness values in the reduced

dataset were continuous and ranged between 0 and approximately 1.5. Average relatedness was 0.5588 (σ_{MR} = 0.2100). The continuous molecular relatedness values in the full dataset ranged between 0 and 1.6. Average relatedness was 0.50 (σ_{MR} = 0.2000).

Genetic analysis

In order to obtain normally distributed residuals, the trait body weight was log-transformed. In the reduced dataset, heritability for body weight was estimated at 0.23 (± 0.09) in a linear univariate model using relatedness inferred from pedigree reconstruction. The genetic variance was $0.56 \cdot 10^{-2}$ g^2. With molecular relatedness in the reduced dataset a heritability of 0.13 (± 0.04) was observed with a genetic variance of $0.31 \cdot 10^{-2}$ g^2. In the full dataset using molecular relatedness, heritability was estimated at 0.11 (± 0.03). Here, the observed genetic variance was $0.27 \cdot 10^{-2}$ g^2.

Comparison of molecular relatedness and reconstructed pedigree relatedness

Per pedigree relationship class, the mean, standard deviation, minimum and maximum values of molecular relatedness were calculated. Results are shown in table 4.1. A large overlap of values obtained from molecular relatedness can be seen between relatedness classes of pedigree reconstruction.

PR		**MR**			
class		**mean**	**sd**	**min**	**max**
Unrelated	0	0.4337	0.1254	0.0000	1.2500
Half sib	0.25	0.6354	0.1410	0.1000	1.4286
Full sib	0.5	0.8737	0.1548	0.1875	1.5000
Self	1	1.2385	0.1216	1.0000	1.6000

Table 4.1 Mean, standard deviation (sd), minimum (min) and maximum (max) of molecular relatedness (MR) between pairs of offspring per class of reconstructed pedigree relatedness (PR) in one generation of cultured common sole (reduced dataset, n=1338).

The Pearson correlation coefficient between molecular and reconstructed pedigree relatedness was 0.8 (P<<0.000). Estimated breeding values from pedigree reconstruction

were positively correlated with breeding values obtained from molecular relatedness (figure 4.2). The estimated Pearson correlation coefficient was 0.77 (P<<0.000).

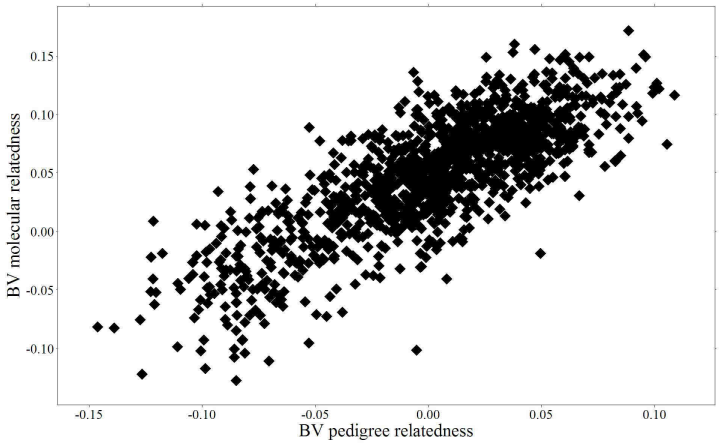


Figure 4.2 Relationship between breeding values estimated from molecular relatedness and reconstructed pedigree relatedness of offspring in cultured common sole (reduced dataset, n=1338).

Cross validation with 100 equal sized randomly chosen subsets resulted in similar correlations between predicted and observed corrected phenotypes in both methods. In figure 4.3 relationships between predicted and observed phenotypes in both methods are shown for the reduced dataset. Note that the stratified shapes of figure 4.3 can be explained by presence of large sib groups in the inferred pedigree. With pedigree reconstruction, predicted and observed phenotypes correlated with 0.2577 after correction for fixed effects. Accuracy of estimated breeding values with pedigree reconstruction was 0.54. Molecular relatedness resulted in a correlation of 0.2619 (P<<0.000) and an accuracy of 0.55. In the full dataset, the correlation between predicted and observed phenotypes was 0.2936 (P<<0.000) after correction for fixed effects. Consequently, the accuracy in the full dataset was 0.60.

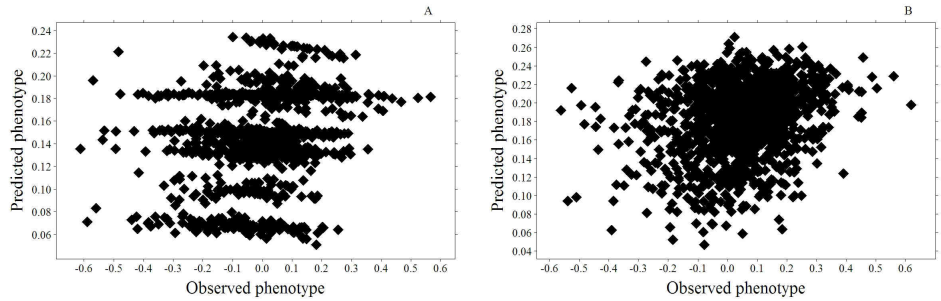


Figure 4.3 Relationship between predicted and observed corrected phenotypes in a linear model with reconstructed pedigree relatedness (A) and molecular relatedness (B) in one generation of cultured common sole (reduced dataset, n=1338).

Discussion

In this study we tested the suitability of molecular relatedness to estimate breeding values in a commercial population of common sole, *Solea solea*, obtained by natural spawning of parents. Progeny and parents were genotyped with ten microsatellite markers. Correlations of estimated relatedness and of breeding values between both methods were high. However, although relations were positive and significant, estimated heritabilities were different between methods.

In the reduced dataset, accuracy of the univariate linear model was 0.54 with pedigree reconstruction and 0.55 with molecular relatedness (figure 4.3). This result implies that both methods fitted the model to the data with nearly equal precision. In the full dataset, accuracy with molecular relatedness increased to 0.60. These results show that in cases with missing pedigree data, molecular relatedness can be used to obtain breeding values. There was a highly skewed contribution of full sib families to the offspring population. Therefore, the observed accuracies should be put in perspective of the maximum accuracy attainable with full sib information (0.71). This implies that the observed levels of accuracy are high.

Prior to drawing conclusions on the optimal indicator of relationships between animals, it is useful to determine if the used forms of relationships add information to breeding value estimation at all. To this purpose, we calculated breeding values using the model (equation 4.3) without any relatedness information. Estimated breeding values from this

model and from the model with relationships inferred from reconstructed pedigree correlated with 0.66 (P<<0.000, results not shown). This suggests that inclusion of relationships between animals alters estimation of breeding values.

Relatedness estimator

Many authors compared relatedness estimators under different population structures, numbers of markers and alleles (e.g. Van de Casteele et al., 2001; Oliehoek et al., 2006; Bink et al., 2008). Maximum likelihood estimators (Mousseau et al., 1998; Thomas et al., 2002; Wang, 2004; Herbinger et al., 2006) or pedigree reconstruction methods that do not need parental information (Berger-Wolf et al., 2007; Ashley et al., 2009) were less suitable for our situation. Assumptions made by these methods, as balanced mating structures and large sibling groups, would be violated in our dataset. In the used dataset skewed parental contributions were observed (data not shown). This is typical for natural spawning species in e.g. aquaculture (Brown et al., 2005; Blonk et al., 2009).
From literature it remains unclear which method-of-moments relatedness estimator performs best; differences are small. In this study we used molecular coancestry (Toro et al., 2002; Toro et al., 2003): among all estimators compared, molecular coancestry turned out to be relatively simple but robust.

The importance of the number of markers for the power of a relatedness estimator has been emphasized by several authors. Oliehoek *et al.* (2008) and Bink *et al.* (2002) showed that reasonable correlations (>0.7) between estimated and pedigree relatedness were only found in simulated data when using at least 50 markers and 4 to 5 alleles per marker. In contrast, in this study a correlation of 0.8 was found between estimated and pedigree relatedness using ten markers with on average 13.7 alleles per marker. The difference in results might also be caused by the characteristics of the population under study. Relatedness estimators generally are used for natural populations where variation of relationship structures is low. In these cases discriminative power of relatedness estimators is low with few markers. For example, Thomas *et al.* (2000) and Wilson and McDonald (2003) showed that 12 to 15 markers were not enough to produce reliable heritability estimates with a relatedness estimator. The tested populations (Soay sheep, *Ovis aries*, and rainbow trout, *Oncorhynchus mykiss*) showed low variation of

relationship structures. Larger variation of relationship structures, i.e. larger families but also larger sample sizes, decreases bias and sampling errors of estimated relatedness (Thomas et al., 2000; Thomas et al., 2002; Rodriguez-Ramilo et al., 2007). In aquaculture populations with natural spawning in groups, generally large variation of relationship structures is found (Gjedrem, 2000; Vandeputte, 2005; Saillant et al., 2006). As contribution of families to the offspring population in this study was highly skewed, this may explain the success with relatively few markers.

Heritability

In this study heritability for bodyweight in common sole was estimated at 0.23 (±0.09) when using pedigree relatedness. The estimated heritability is in line with values found for body weight in other aquatic species (Gjedrem, 2000). Heritability estimated from molecular relatedness however, was much lower: 0.11 (±0.02) to 0.13 (±0.04). This suggests underestimation of genetic variance which is likely to be caused by erroneous ascribing of relationships between offspring due to too few markers. Thomas *et al. (2002)* showed that, among other factors, use of fewer markers increases variance and bias of sampling error variance of estimated relatedness. With higher variance of relatedness, genetic variance and heritability decrease as $\hat{\sigma}_A^2$ $r = cov(P_i,P_j)$ (Falconer and Mackay, 1996). This is supported by a study on genomic selection where fewer markers led to underestimation of heritability estimates in an Angus cattle population (Hayes and Goddard, 2008). A similar effect was demonstrated by Shikano (2007) in Japanese flounder, *Paralichthys olivaceus*. It is expected that heritability will be estimated more accurately with a higher number of markers and consequently will approach the value from pedigree reconstruction.

In our particular case of skewed contribution of parents, the observed heritabilities from molecular relatedness will have little effect on estimated breeding values and accuracies. This is because breeding values were based on family information (i.e. family means) rather than on heritabilities.

Applications for breeding programs

Our study illustrates that pedigree reconstruction using genotyping data is a inefficient method when parental genotypes are missing. Difficulty with allocation is reflected by figure 4.1. Loss of selection candidates due to ambiguous parental allocation was 33%.

Comparable results (10 to 30%) were found by several authors in other species (Herlin et al., 2007; Herlin et al., 2008; Pierce et al., 2008). Loss of selection candidates has major effects on costs of selection procedures in breeding programs and may lead to lower realized selection responses and increased rates of inbreeding.

As parental allocation programs often allow uncertainties to the performed allocations (Duchesne et al., 2002), mistakes in reconstructed pedigrees may occur. This increases the risk that related animals are considered as unrelated. Selection of animals using optimal contribution theory (Meuwissen, 1997) may therefore unintentionally increase rates of inbreeding.

This study shows that the molecular relatedness estimator circumvents problems with pedigree reconstruction in an aquaculture population with natural spawning of groups of parents. Moreover, due to preservation of all selection candidates, a higher accuracy of breeding value estimation is achieved. The increasing amount of genetic information (e.g. number of SNP's and microsatellites) made available due to current developments in genomics (Hayes et al., 2009) will further enhance accurate estimation of genetic parameters.

Acknowledgements

This study was funded by Casimir NWO (the Netherlands Organization for Scientific research), Solea bv and Wageningen University. The authors thank team Solea for their help in collecting the data, and Piter Bijma for his helpful suggestions and advice on the genetic analysis with relatedness estimators. We thank the reviewers of Genetics for their useful comments.

References

Anderson, A. D. and Weir, B. S., 2007. A Maximum-Likelihood Method for the Estimation of Pairwise Relatedness in Structured Populations. Genetics 176, 421-440.

Ashley, M. V., Caballero, I. C., Chaovalitwongse, W., Dasgupta, B., Govindan, P., Sheikh, S. I. and Berger-Wolf, T. Y., 2009. KINALYZER, a computer program for reconstructing sibling groups. Molecular Ecology Resources 9, 1127-1131.

Avise, J. C., Jones, A. G., Walker, E., DeWoody, J. A. and collaborators, 2002. Genetic mating systems and reproductive natural histories of fishes: lessons for ecoloy and evolution. Annual Review of Genetics 36, 19.

Bekkevold, D. X., Hansen, M. M. and Loeschcke, V., 2002. Male reproductive competition in spawning aggregations of cod (*Gadus morhua* , L.). Molecular Ecology 11, 91-102.

Berger-Wolf, T. Y., Sheikh, S. I., DasGupta, B., Ashley, M. V., Caballero, I. C., Chaovalitwongse, W. and Putrevu, S. L., 2007. Reconstructing sibling relationships in wild populations. Bioinformatics 23, i49-56.

Bink, M., Anderson, A., van de Weg, W. and Thompson, E., 2008. Comparison of marker-based pairwise relatedness estimators on a pedigreed plant population. TAG Theoretical and Applied Genetics 117, 843-855.

Blonk, R. J. W., Komen, J., Kamstra, A., Crooijmans, R. P. M. A. and 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.

Brown, C. R., Woolliams, J. A. and McAndrew, B. J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247, 219-225.

Duchesne, P., Godbout, M. H. and Bernatchez, L., 2002. PAPA (Package for the Analysis of Parental Allocation) : A computer program for simulated and real parental allocation. Molecular Ecology Notes 2, 191-194.

Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to quantitative genetics. Pearson Prentice Hall Harlow, Essex.

Fernández, J. and Toro, M. A., 2006. A new method to estimate relatedness from molecular markers. Molecular Ecology 15, 1657-1667.

Fessehaye, Y., El-Bialy, Z., Rezk, M. A., Crooijmans, R., Bovenhuis, H. and Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. Aquaculture 256, 148-158.

Garoia, F., Marzola, S., Guarniero, I., Trentini, M. and Tinti, F., 2006. Isolation of polymorphic DNA microsatellites in the common sole Solea vulgaris. Molecular Ecology Notes 6, 144-146.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. and Thompson, R., 2006. ASReml User Guide Release 2.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Gjedrem, T., 2000. Genetic improvement of cold-water fish species. Aquaculture Research 31, 25-33.

Hayes, B. J. and Goddard, M. E., 2008. Technical note: Prediction of breeding values using marker-derived relationship matrices. Journal of Animal Science 86, 2089-2092.

Hayes, B. J., Visscher, P. M. and Goddard, M. E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91, 47-60.

Herbinger, C. M., O'Reilly, P. T. and Verspoor, E., 2006. Unravelling first-generation pedigrees in wild endangered salmon populations using molecular genetic markers. Molecular Ecology 15, 2261-2275.

Herlin, M., Taggart, J. B., McAndrew, B. J. and Penman, D. J., 2007. Parentage allocation in a complex situation: A large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. Aquaculture 272, S195-S203.

Herlin, M., Delghandi, M., Wesmajervi, M., Taggart, J. B., McAndrew, B. J. and Penman, D. J., 2008. Analysis of the parental contribution to a group of fry from a single day of spawning from a commercial Atlantic cod (*Gadus morhua*) breeding tank. Aquaculture 274, 218-224.

Iyengar, A., Piyapattanakorn, S., Stone, D. M., Heipel, D. A., Howell, B. R., Baynes, S. M. and Maclean, N., 2000. Identification of microsatellite repeats in turbot (*Scophthalmus maximus*) and Dover sole (*Solea solea*) using a RAPD-based technique: characterization of microsatellite markers in Dover sole. Marine Biotechnology 2, 49-56.

Li, C. C., Weeks, D. E. and Chakravarti, A., 1993. Similarity of DNA fingerprints due to chance and relatedness. Human Heredity 43, 45-52.

Lynch, M., 1988. Estimation of relatedness by DNA fingerprinting. Molecular Biology and Evolution 5, 584-599.

Lynch, M. and Walsh, B., 1997. Genetic and Analysis of Quantitative traits. Sinauer Associates Inc., Sunderland, USA.

Lynch, M. and Ritland, K., 1999. Estimation of pairwise relatedness with molecular markers. Genetics 152, 1753-1766.

Marshall, T. C., Slate, J., Kruuk, L. E. B. and Pemberton, J. M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7, 639-655.

Meuwissen, T. H. E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. Journal of Animal Science 75, 934-940.

Milligan, B. G., 2003. Maximum-Likelihood Estimation of Relatedness. Genetics 163, 1153-1167.

Mousseau, T. A., Ritland, K. and Heath, D. D., 1998. A novel method for estimating heritability using molecular markers. Heredity 80, 218-224.

Oliehoek, P. A., Windig, J. J., van Arendonk, J. A. M. and Bijma, P., 2006. Estimating relatedness between individuals in general populations with a focus on their use in conservation programs. Genetics 173, 483-496.

Pierce, L. R., Palti, Y., Silverstein, J. T., Barrows, F. T., Hallerman, E. M. and Parsons, J. E., 2008. Family growth response to fishmeal and plant-based diets shows genotype × diet interaction in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 278, 37-42.

Queller, D. C. and Goodnight, 1989. Estimating relatedness using genetic markers. Evolution 43, 258-275.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.

Ritland, K., 1996. A marker-based method for inferences about quantitative inheritance in natural populations. Evolution 50, 1062-1073.

Ritland, K., 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. Molecular Ecology 9, 1195-1204.

Rodriguez-Ramilo, S. T., Toro, M. A., Caballero, A. and Fernandez, J., 2007. The accuracy of a heritability estimator using molecular information. Conservation Genetics 8, 1189-1198.

Rodríguez-Ramilo, S. T., Toro, M. A., Martínez, P., Castro, J., Bouza, C. and Fernández, J., 2007. Accuracy of pairwise methods in the reconstruction of family relationships, using molecular information from turbot (Scophthalmus maximus). Aquaculture 273, 434-442.

Saillant, E., Dupont-Nivet, M., Haffrey, P. and Chatain, B., 2006. Estimates of heritability and genotype-environment interactions for body weight in sea bass (*Dicentrarchus labrax* L.) raised under communal rearing conditions. Aquaculture 254, 139-147.

Thomas, S. C., Pemberton, J. M. and Hill, W. G., 2000. Estimating variance components in natural populations using inferred relationships. Heredity 84, 427-436.

Thomas, S. C., Coltman, D. W. and Pemberton, J. M., 2002. The use of marker-based relationship information to estimate the heritability of body weight in a natural population: a cautionary tale. Journal of Evolutionary Biology 15, 92-99.

Toro, M., Barragán, C., Óvilo, C., Rodrigañez, J., Rodriguez, C. and Silió, L., 2002. Estimation of coancestry in Iberian pigs using molecular markers. Conservation Genetics 3, 309-320.

Toro, M. A., Barragan, C. and Ovilo, C., 2003. Estimation of genetic variability of the founder population in a conservation scheme using microsatellites. Animal genetics 34, 226-228.

Van de Casteele, T., Galbusera, P. and Matthysen, E., 2001. A comparison of microsatellite-based pairwise relatedness estimators. Molecular Ecology 10, 1539-1549.

Vandeputte, M., 2005. Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). Aquaculture 247, 31-32.

Wang, J., 2004. Sibship Reconstruction From Genetic Data With Typing Errors. Genetics 166, 1963-1979.

Wilson, A. J. and McDonald, 2003. Marker-assisted estimation of quantitative genetic parameters in rainbow trout, *Oncorhynchus mykiss*. Genetical Research 81, 145-156.

5

Heritability of shape in common sole, estimated with digital image analysis

Robbert J.W. Blonk
Hans Komen
Amabel Tenghe
Andries Kamstra
Johan A.M. van Arendonk

Published in Aquaculture 307 (2010), 6-11

Abstract

In fish, selection on production traits may alter traits such as shape through genetically correlated response to selection. In species which are sold un-filleted, it may be desired to maintain "natural" shape while selecting for increased growth. Several methods have been proposed to describe shape of fish, but most of these are either sensitive to observer perception or time consuming. Digital image analysis is objective, relatively free of bias, rapid and widely used to describe shape in e.g. crops and fruits. To our knowledge there are no reports on its use in selection programs for fish. In this study we compare estimated genetic parameters of shape and body measurements obtained with digital image analysis (DIA) and manual measurements in a captive commercial population of common sole, *Solea solea*, with 1222 pedigreed animals. Results show that estimated genetic parameters were similar across methods and that DIA resulted in significantly lower standard errors ($\bar{se}_{DIA}$ = 0.0772, $\bar{se}_{manual}$ = 0.0880, P = 0.01244). Heritability for shape using DIA was 0.34 ± 0.11. The genetic correlation between shape and body weight was –0.44. Consequently, selecting for increased body weight at harvest yields more circular shaped fish. However, by using a selection index for shape and body weight, it is possible to maintain shape while improving body weight at harvest. This study shows that digital image analysis is an accurate and time efficient method to estimate genetic parameters for shape traits in fish .

Introduction

Appearance traits are important for marketing of many fish species and may directly influence consumer willingness to pay the price of fish products. This effect is clearly documented for fillet colour in salmon (Steine et al., 2005; Alfnes et al., 2006) but also the importance of shape has been pointed out (Kause et al., 2003a). Despite this, only few aquaculture breeding programs incorporate shape traits in the breeding goal (Chavanne et al., 2008). In case farmers are paid for filleted products, selective breeders can deliberately alter shape by selecting on body weight or on body measurements that have a significant genetic correlation with fillet yield (Rutten et al., 2004). Yet, in species which are sold un-filleted, such as common sole, *Solea solea,* a correlated response in shape is undesired and selection methods should include some kind of body measurement to maintain "natural" shape.

In several cultured species including Nile tilapia, *Oreochromis niloticus* (Rutten et al., 2005), rainbow trout, *Oncorhynchus mykiss* (Gjerde and Schaeffer, 1989) and Atlantic salmon, *Salmo salar* (Powell et al., 2008), substantial genetic variances have been estimated for body length, body height and body width (estimated heritabilities between 0.25 and 0.46). Also for fish shape, genetic effects were demonstrated in e.g. cultured rainbow trout, *Oncorhynchus mykiss* (Gjerde, 1989; Gjerde and Schaeffer, 1989; Kause et al., 2003a), gilthead seabream, *Sparus auratus* (Navarro et al., 2009) and in a natural population of Northern red belly dace, *Phoxinus eos* (Toline and Baker, 1997). For example, Kause *et al*., (2003a) showed a heritability of 0.46 for shape in rainbow trout when scored as a categorical trait. These studies indicate that sufficient genetic variation is available to maintain or alter shape of several fish species by means of selective breeding.

To improve efficiency of selection, recording of traits should preferably be unbiased (i.e. not affected by person), accurate and with low input of financial resources as e.g. labour (Gjedrem, 1997). Most studies employ manually determined measurements of traits as e.g. body length and body height (Rutten et al., 2005), and characterize shape as a ratio of these traits (Gjerde and Schaeffer, 1989). These methods are labour intensive, especially with large numbers of animals. Further, differences between observers e.g. when determining the optimal point for measurement of body height, increase error in analyses.

Other studies quantified shape using categorical and subjective scoring (Rye and Gjerde, 1996; Kause et al., 2003a; Kause et al., 2004). Although simple, these manual methods are sensitive to observer perceptions and significant differences between results of scorers were found (Norris and Cunningham, 2004).

To objectively characterize shape, more elaborate methods have been developed using image analysis and e.g. geometric morphometrics (Currens et al., 1989; Loy et al., 1999; Monet et al., 2006; Park et al., 2007; Ambrosio et al., 2008; Russo et al., 2009). Digital image analysis methods (further referred to as "DIA") describe shape relatively free of bias, can easily be automated and are widely used in sorting procedures in (agricultural) industry (Gonzalez et al., 2008; Meyer and Neto, 2008; Burgos-Artizzu et al., 2009). However, to our knowledge, no such method has been used to describe shape in fish with the purpose of using automated data recording on shape in a breeding program (Chavanne et al., 2008).

In this study, the goal was to quantify genetic variation of shape of the saggital plane of common sole, using DIA. The saggital plane divides individuals into the left and right portions, which corresponds to the blind and the eyed side of soles. The saggital plane of common sole has an elliptic shape. However, abnormal "turbot-like" (circular) shapes have been observed in several commercial grow-out populations (Imsland et al., 2003b). In this study we show that with DIA of shape in common sole, similar genetic parameters and slightly lower standard errors are obtained when compared to manual analysis. We also show that it is possible to maintain shape while improving body weight at harvest by selecting on shape and body weight using DIA.

Materials and methods

Data collection

A group of 1338 animals with pedigree (Blonk et al., 2010a) was randomly sampled from eight tanks at a commercial farm (Solea b.v., IJmuiden, The Netherlands). The population had been produced by natural mating of 21 sires and 17 dams. This resulted in a complex full / half sib structure where nearly all sires and dams had multiple partners. The mean number of offspring was 67.8 (sd = 165.7) for sires and 78.8 (sd = 140.4) for dams indicating a very skewed contribution of parents. The sex ratio of the offspring was approximately 2.3 : 1 (male : female). As all families were mixed at spawning, the

moment of production and age of animals is unknown. However, each spawning season lasted for three months suggesting a maximum age difference of three months as well. Relatively to the age of the animals (2-3 years) this is a small difference and it was therefore assumed that this would not affect analysis.

Prior to measurements, each animal was anaesthetized with 2-phenoxyethanol. Sex was examined using ultrasound (System: Esaote Pie Medical MyLab30Vet; Transducer: Esaote LA435 6-18 MHz). Next, a photo of the blind side of each animal was taken with a digital camera (Creative webcam, picture size 800 x 600 pixels) at a standard height perpendicular to the animal. The blind side of the animal was chosen to ease analyzing shape of the saggital plane. All photos were provided with a reference scale length. After this, body length (to the nearest mm), body height (to the nearest mm) and body weight (to the nearest g) were recorded ("Manual method"). Determination of body length and body height is shown schematically in figure 5.1. Measurements were performed by three teams, each consisting of two persons. Per team, measurements were carried out by one person and written down by a second person. Each team measured approximately one third of the dataset.

Digital image analysis

Because of poor quality, 116 photos were discarded. The final dataset consisted of 1222 animals with pictures, manual measurements and pedigree. Body measurements of the saggital plane at the blind side of each photographed fish were quantified using ImageJ software (Rasband, 2008). Using a macro, photos were loaded sequentially into ImageJ. From each photo, the outline of the animal, without tail and side fins, was traced manually. Using a standard algorithm of the software, an ellipse was then fitted to the outline to determine the longest and the smallest diameter (major and minor axis, see figure 5.1). Here the minor axis is always exactly perpendicular (90° angle) to the major axis. The major axis was used as a measure for body length; the minor axis was used as a measure for body height. After this, the recorded data was stored and a new photo was loaded automatically. Using this method, each hour approximately 170 pictures were analysed.

To provide an indication of repeatability of DIA and of effects of persons selecting the outline, 25 randomly chosen animals were analyzed three times repeatedly in random order by two persons. Results were tested using a mixed linear model with a fixed effect

for person and a random effect of repeated measurement in R (R Development Core Team, 2008). Repeatability was calculated as the correlation between estimates after correction for the fixed effect of person.

Quantification of shape

We used ellipticity as a measure for shape of common sole as the shape of the saggital plane of this species is nearly elliptic. We calculated ellipticity (adapted from Merigot et al., 2007) using body length (*BL*) and body height (*BH*) for both the manual method and the DIA method:

$$Ellipticity = \frac{(BL - BH)}{(BL + BH)} \qquad 5.1$$

From equation 1 it can be derived that larger shape values (maximum = 1) reflect more elongated shapes whereas smaller values represent more circular shapes. In a perfect circle, ellipticity is zero.

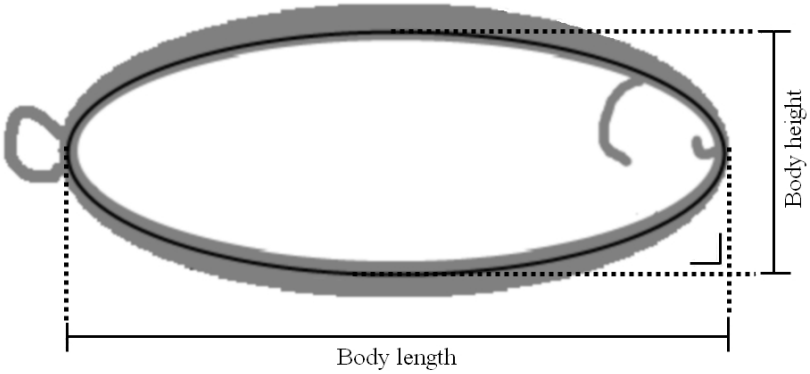


Figure 5.1 Body length and body height as measured on the saggital plane at the blind side of common sole, for manual analysis and digital image analysis. For image analysis, the fitted outline is indicated.

Genetic analysis

To determine effects to be included in genetic analyses of both methods, preliminary ANOVA was performed using R software (R Development Core Team, 2008) for body length, body height, body weight and shape. Due to non normal distribution of residuals, body weight (BW) was log-transformed for analysis. Effects of sex and tank were found significant for all traits (P<<0.0001) across both methods and were thus included in genetic analysis. Effects of team were not relevant for DIA and were only included in analysis of the manual method (P<<0.0001). No significant interactions of fixed effects were found. The skewed contribution of parents resulted into 4 parents with a contribution of 1. These records did not affect results of genetic analysis and were therefore kept in the dataset.

Genetic analysis of data (n = 1222) was performed using ASReml software (Gilmour et al., 2006). To estimate heritability of all traits, variance components were obtained using the following univariate linear animal model:

$$Y_{ijkl} = \mu + Sex_i + Tank_j + Team_k + a_l + e_{ijkl} \quad 5.2$$

where Y_{ijkl} is the response variable of animal l, μ is the population mean, Sex_i is the fixed effect of sex i, $Tank_j$ is the fixed effect of tank j, $Team_k$ is the fixed effect of team k in the manual method, a_l is the random animal effect of animal l, and e_{ijkl} is the random residual error of animal l. For genetic analysis of traits obtained by the DIA method, the effect of team was not present. Heritability was obtained using estimated variance components from ASReml as $h^2 = \sigma^2_a / \sigma^2_p$ where σ^2_a is the genetic variance and σ^2_p is the phenotypic variance.

To estimate genetic and phenotypic correlations between all traits measured, a bivariate linear model was used:

$$Y_{1,ijkl}\, Y_{2,ijkl} = \mu + Sex_i + Tank_j + Team_k + a_l + e_{ijkl} \quad 5.3$$

where $Y_{1,ijkl}$ and $Y_{2,ijkl}$ are the first and second response variables of animal l and $Team_k$ is the fixed effect of team k if the corresponding response variable was obtained using the manual method.

Accuracy and response of selection

To determine the optimal selection index to maximize gain of body weight and to minimize response in shape, we simulated different selection indices in SelAction

software (Rutten et al., 2002) using parameters obtained with DIA. SelAction software predicts selection responses, rates of inbreeding and accuracy for various breeding program designs with a crossed mating design while accounting for Bulmer effects.

Each simulated breeding program contained 200 sires and 200 dams with each dam producing 100 male and 100 female full sib offspring. Per sex, a selection proportion of 1% was used. In all simulations, the breeding goal was for body weight and shape. Selection indices included information from BLUP (Best Linear Unbiased Prediction) and sibs. Several selection indices with exclusively body weight, body length, body height ("single trait index") were tested. Also, various combinations of body weight, body length, body height and shape ("multi trait indices") were tested. We used a desired gains approach to determine optimal weights for the index traits whereby the aim was to maximise gain in body weight with zero response in shape. Response of body weight and shape as well as accuracy (r_{IH}) were recorded.

Table 5.1 Mean, minimum (min), maximum (max) and coefficient of variation (CV) per sex and in total for body length (BL), body height (BH), shape (SH) and body weight in one generation of common sole (n_{males} = 851; $n_{females}$ = 371; n_{total} = 1222) for both manual and digital image analysis methods.

	BL			BH			SH		
Manual	**m**	**f**	**tot**	**m**	**f**	**tot**	**m**	**f**	**tot**
mean	22.23	24.79	23.01	8.38	9.78	8.80	0.45	0.44	0.45
min	14.50	15.70	14.50	5.20	5.80	5.20	0.30	0.25	0.25
max	30.90	34.20	34.20	13.80	13.80	13.80	0.54	0.58	0.58
CV	13.42	13.58	14.43	15.72	16.40	17.60	6.31	7.00	6.76
Digital image analysis									
mean	19.48	21.71	20.16	7.55	8.66	7.88	0.44	0.43	0.44
min	13.01	13.64	13.01	4.64	4.95	4.64	0.30	0.34	0.30
max	28.17	29.84	29.84	12.08	12.10	12.10	0.53	0.52	0.53
CV	13.13	13.58	14.24	15.01	16.01	16.71	6.72	6.49	6.75
Body weight									
mean	164.69	241.72	188.08						
min	45.00	52.00	45.00						
max	650.00	691.00	691.00						
CV	44.08	45.71	49.38						

Results

Mean, minimum and maximum values and coefficients of variation for all traits measured with the manual method and with DIA are presented in table 5.1. Mean body weight of all fish in the population was 188.1 g (range 45.0 - 691.0 g). As is normal for common sole, mean values of body length, body height and body weight were higher for females than for males.

Generally, mean values measured with DIA were smaller than the corresponding manual measurements. For example, mean body length obtained from manual measurements was 23.01 cm whereas for DIA this was 20.16 cm. For all traits, there was little difference of coefficients of variation among methods; coefficients of variation for body length and body height were between 13 and 18%. Repeatability for DIA was estimated on 0.96 – 0.99, proving consistency of the method within each person. However, results showed a significant effect of person ($P < 0.05$).

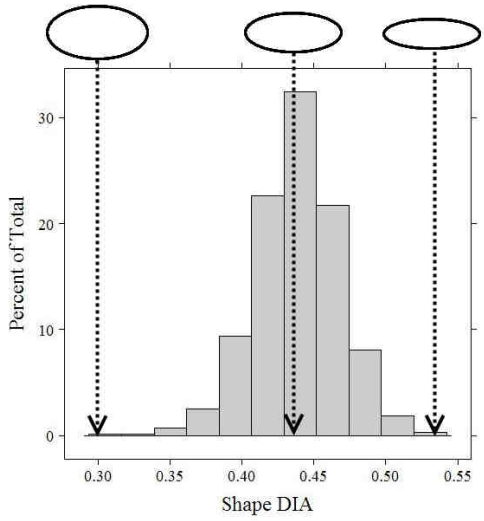


Figure 5.2 Distribution of shape measured with digital image analysis in one generation of common sole. The ellipses above visualize the shape corresponding the value on the x-axis.

The distribution of shape obtained by DIA is shown in figure 5.2. In DIA, shape ranged from 0.30 to 0.53 with a mean value of 0.44. The coefficient of variation for shape from DIA was 6.75%, indicating relatively low variability in the population for shape. On average, females had a very similar shape when compared to males.

Genetic analysis

Estimated variance components and heritabilities from univariate analysis are shown in table 5.2. In nearly all cases, estimated variance components obtained with data from DIA were lower than values obtained with data from the manual method. However, differences were not significant (results not shown).

Table 5.2 Genetic variance, phenotypic variance and estimated heritability ± standard error (h^2± S.E.) for body length (BL), body height (BH), shape (SH) and body weight (BW) for manual and digital image analysis methods in one generation of common sole.

Manual	σ^2_A	σ^2_P	h^2 ± S.E.
BL	2.2454	6.2373	0.36 ± 0.12
BH	0.4128	1.3316	0.31 ± 0.11
SH	0.0004	0.0009	0.45 ± 0.14
Digital image analysis			
BL	1.2425	4.6019	0.27 ± 0.10
BH	0.3501	1.0610	0.33 ± 0.11
SH	0.0003	0.0009	0.34 ± 0.11
BW (log)	0.0061	0.0245	0.25 ± 0.09

Heritabilities for body length and body height ranged from 0.27 to 0.36 with standard errors from 0.10 to 0.12, indicating that estimated heritabilities were not significantly different across methods (confidence interval = mean ± 2 · s.e.m.). Estimated heritability of shape using manual measurements was 0.45 (± 0.14). When using DIA, estimated heritability of shape was somewhat, although not significantly, lower: 0.34 (± 0.11).

Estimated heritability of body weight was 0.25 (± 0.09). This is similar to the value obtained in a previous study with the full dataset containing 1338 individuals (Blonk et al., 2010a).

Correlations between methods

Generally, correlations between the manual method and DIA were high with low standard errors; for shape, the genetic correlation was 0.96 (± 0.04). For body length and body height, high correlations lead to difficulty with convergence of REML. Estimated correlations were 0.99 ± 0.00. As expected, the phenotypic correlation between shape obtained with manual measurements and obtained with DIA was lower (0.66 ± 0.03).

Table 5.3 Genetic correlation (below diagonal in italics) and phenotypic correlation (above diagonal) ± standard error between body length (BL), body height (BH), shape (SH) and body weight (BW) for manual and digital image analysis methods in one generation of common sole.

	Manual			
	BL	**BH**	**SH**	**BW**
BL		0.83 ± 0.02	0.01 ± 0.06	0.90 ± 0.01
BH	*0.77 ± 0.12*		-0.53 ± 0.04	0.93 ± 0.01
SH	*0.23 ± 0.27*	*-0.43 ± 0.23*		-0.32 ± 0.05
BW	*0.87 ± 0.07*	*0.98 ± 0.02*	*-0.26 ± 0.28*	
	Digital image analysis			
	BL	**BH**	**SH**	**BW**
BL		0.82 ± 0.02	-0.02 ± 0.05	0.91 ± 0.01
BH	*0.81 ± 0.10*		-0.58 ± 0.03	0.91 ± 0.01
SH	*-0.12 ± 0.29*	*-0.68 ± 0.16*		-0.30 ± 0.04
BW	*0.91 ± 0.05*	*0.96 ± 0.03*	*-0.44 ± 0.25*	

Phenotypic and genetic correlations between traits are shown in table 5.3 for both methods separately. Mean standard errors from parameters obtained with DIA were significantly lower than those obtained with manual analysis ($\bar{se}_{DIA}$ = 0.0772; $\bar{se}_{Manual}$ = 0.0880; P = 0.01244). Although differences were small, this indicates that genetic parameters are more accurately estimated with DIA than with the manual method.

Correlations with shape

Correlations between body length and shape were not different from zero in both methods. With DIA, a phenotypic correlation of -0.02 ± 0.05 was found. The estimated genetic correlation was -0.12 ± 0.29. Given the large standard errors, correlations estimated from the manual measurements were not different from those obtained with DIA.

Phenotypic and genetic correlations between body height and shape from DIA were negative: respectively -0.58 ± 0.03 and -0.68 ± 0.16 (table 5.3). For the manual method, similar values were obtained. This indicates that shape and body height are related, with more circular shaped fish (lower values of shape) having higher body height.

Correlations with body weight

Estimated genetic correlations between body weight and body height were high: 0.98 ± 0.02 for manual measurements and 0.96 ± 0.03 for DIA. Phenotypic correlations were respectively 0.91 ± 0.01 and 0.93 ± 0.01. Genetic correlations between body weight and body length were somewhat lower: 0.87 ± 0.07 for manual measurements and 0.91 ± 0.05 for DIA. This suggests that body weight has a less strong relationship with body length than with body height although differences are small. Correlations between shape and body weight were lower and negative (-0.26 to -0.44) but standard errors for genetic correlations were high (0.25 to 0.28).

Accuracy and response of selection

Predicted response per generation and accuracy from different selection indices are shown in table 5.4 for genetic parameters obtained with DIA. In all scenarios, responses of body weight and shape were in opposite directions. For example, with selection on body weight alone, shape decreased with 2.27% (response expressed to initial mean) whereas response in body weight was +23.1% per generation.

Compared to other single trait indices, response of shape was most affected when selection was on body height alone. Selecting on body length affected response of shape the least. Here, response of body weight was lowest as well (+20.8% per generation).

Table 5.4 Response of traits in the breeding goal (H) [body weight (BW) in % and shape (SH) in % ellipticity relative to the initial mean] and accuracy of selection (r_{IH}) when selecting with single trait indices or multi trait indices containing body weight (BW), body length (BL), body height (BH) and shape (SH) from digital image analysis.

	Response (H)		r_{IH}
	SH	**BW**	
Single Trait Index (I)			
BW	-2.27	23.1	0.558
BL	-0.68	20.8	0.544
BH	-3.64	22.1	0.489
Multi Trait Index (I)			
SH + BW	-1.14	22.3	0.574
SH + BL	-1.36	21.6	0.550
SH + BH	-1.14	22.6	0.579
Multi Trait Index (I) - weighed			
SH + BW	0.00[a]	20.3	0.573
SH + BL	0.00[a]	19.6	0.548
SH + BH	0.00[a]	20.6	0.579

[a] Weights SH: 2; BW, BL and BH: 1

In multi trait indices, responses of shape were very similar (-1.14 to -1.36%). The highest response in body weight (+22.6 % /generation) was obtained when body height was included in the index as a second trait. To set response of shape to zero, a weighed index with weights of “two” on shape and “one” on body weight was required. Responses of body weight in weighed selection indices were lower than when no weighing of selection indices was used. For example, when body height was included in the index as a second trait, response of body weight was 20.6 % per generation.

Accuracies of single trait indices ranged from 0.489 to 0.558. In multi trait indices, the highest accuracies (0.573 and 0.579) were achieved when body weight or body height were in the index with shape.

Discussion

In this study, we used DIA to determine genetic variation of shape in a captive commercial population of common sole. To describe shape of the saggital plane,

ellipticity (equation 1) was calculated. Measuring shape with DIA provides a relatively objective way to quantify shape as a continuous trait. The estimated heritability of shape in the analyzed population (0.34 ± 0.11) is higher than heritability estimates in rainbow trout reported by Gjerde and Schaeffer (1989). These authors obtained values from 0.00 to 0.25 using ratios of body measurements (length, height, and width) to describe shape. The estimates presented in this study are in the range of those reported by Kause et al. (2004) who showed heritabilities from 0.31 to 0.40, based on subjective categorical scoring.

Digital image analysis

To quantify shape or other traits, manual or subjective scoring methods have been used in several studies (Gjerde and Schaeffer, 1989; Rye and Gjerde, 1996; Kause et al., 2003a; Kause et al., 2004; Norris and Cunningham, 2004; Rutten et al., 2005). However, these type of methods are sensitive to errors for several reasons. First, bias in estimates due to different observers has been shown by Norris and Cunningham (2004). The effect of observers is also demonstrated by significant effects of teams in our data. Second, the discriminating power of scoring is highly influenced by type and variation of the trait under study. For example, when determining colour intensity, distinguishing different colour tones often becomes a matter of personal interpretation, especially when relative differences between samples are small (Norris and Cunningham, 2004). The same applies to traits such as shape. Also, when estimating body height or width in fish, it is often difficult to determine the exact position and angle for the measurement. This leaves, again, room for between observer variation. Although such (subjective) scoring methods may seem easy under certain circumstances, accurately evaluating traits might require more objective methods.

In this study, shape was also measured with DIA to reduce between observer variance. This type of analysis has been widely used in agricultural research and has proven a high degree of accuracy in measurements (Meyer and Neto, 2008; Burgos-Artizzu et al., 2009; Gonzalo et al., 2009). Several studies have applied DIA to quantify colour of fillets or skin, (Bencze Rørå et al., 1998; Kause et al., 2008; Powell et al., 2008; Shikano, 2008) or fish shape using geometric morphometrics (Currens et al., 1989; Loy et al., 1999; Park et al., 2007; Ambrosio et al., 2008; Russo et al., 2009). In this study, we used a relatively simple method to characterize shape of flatfish. The use of this method was

straightforward as flatfishes, in particular *Solea spp*., have relatively simple elliptic shapes.

In this study, correlations of measurements across DIA and manual methods were generally high and estimated genetic parameters were in the same range. Moreover, genetic parameters estimated from DIA data had significantly lower mean standard errors ($\overline{se}_{DIA}$ = 0.0772) than genetic parameters estimated from manual analysis ($\overline{se}_{manual}$ = 0.0880), although the difference was small. These results suggest that both methods produce at least very similar results and that DIA can replace manual analysis.

One advantage of DIA is that this method is far more time efficient than the manual method. In this study, it took eight hours (one person working for one day) to analyse the dataset with DIA while 16 hours (two persons, one day) were effectively required to do the manual measurements, excluding supporting work. Another advantage of DIA is that images can be stored for later use, e.g. for data checking or to serve as reference for later generations.

Recent developments in DIA software programs enable fully automated detection of outlines (Rasband, 2008; Gonzalo et al., 2009). This allows DIA to be used in automated data recording, as is routinely done in industry. Unfortunately, in our dataset, the quality of the images was poor, which was mainly due to lack of contrast and inconsistent light intensities. This precluded the use of automated tracing and data recording. Instead, the outline of each fish had to be selected by hand and fitted to a ellipse, which probably introduced a bias in the measurements. Consequently, body length and body height are approximately 10% lower with DIA than with manual measurements (table 5.1). Nevertheless, repeatability of the DIA method was very high for body length and body height, indicating that robustness of the analysis was not affected. We are currently optimising the conditions of image capture so that it will be suited for automated detection of outline. This will likely reduce bias and increase accuracy even further.

Implications for selection

Correlations between shape and body weight were moderate and negative (table 5.3). This implies that direct selection on body weight will lead to decreased shape in the population (Lynch and Walsh, 1998). In other words, there will be more circular shaped animals.

Moreover, correlations between body weight and body height were high. Correlations between body weight and body length were slightly lower but still high (≈ 0.90). This implies that body height and body length can predict body weight accurately. For other species as e.g. gilthead seabream and Nile tilapia, comparable correlations were shown (Rutten et al., 2005; Navarro et al., 2009). These results suggest that in common sole, indirect selection on body weight through one of these traits will lead to decreased shape in the population (Falconer and Mackay, 1996; Lynch and Walsh, 1998). However, both for body length and body height, genetic and phenotypic correlations with shape were moderate at most whereas the phenotypic variance of shape was very small. This suggests that the correlated response of shape is likely to be limited.

Predicted responses (table 5.4) from different selection indices in SelAction software (Rutten et al., 2002) show relationships in accordance with conclusions from the correlation structures as found in this study. In all tested selection indices, responses in the breeding goal traits were negative for shape and positive for body weight. Including shape in the index as a second trait generally corrects response for shape towards zero. The highest response for body weight was found when selecting on body weight or on body height (Falconer and Mackay, 1996). Although differences were small, response of body weight was higher when selecting on body height rather than when selecting on body weight itself. The reason for this is that heritability of body height is higher than heritability of body weight (table 5.2). Consequently, this enables improvement of production through selecting for traits such as shape and body height, e.g. using DIA.
Accuracies of selection in multi trait indices were lowest when body length was included; 0.548 to 0.550. Higher accuracies (0.574 to 0.579) were obtained in an index which included shape with body weight or body height. This shows that when a multi trait selection index is used, response of body weight as well as accuracy of selection increase. Similar effects were described by Mrode (2005) and Falconer and Mackay (1996). To obtain zero response in shape, the weight given to shape had to be twice the weight for body weight or body height. This resulted in a reduced response of body weight.

In this study, we showed that digital image analysis is a promising tool to phenotype animals for genetic analysis and to quantify shape, yielding at least the same genetic parameters as manual methods, but being more accurate and far more time efficient. Our

results also suggest that, due to high correlations between body weight and body height, selection on body height can be used to improve production of common sole. In this way, image analysis can aid automation of all-in-one recording systems for traits as shape concomitantly with body weight. This method can also easily be integrated with current automated equipment for grading of animals, thus enabling low cost accurate recording which is important for effective implementation of breeding programs.

Acknowledgements

This study was funded by Casimir NWO (the Netherlands Organization for Scientific research), Solea bv and Wageningen University. The authors thank team Solea for their help in collecting the data.

References

Alfnes, F., Guttormsen, A. G., Steine, G. and Kolstad, K., 2006. Consumers' willingness to pay for the color of salmon: A choice experiment with real economic incentives. American Journal of Agricultural Economics 88, 1050-1061.

Ambrosio, P. P., Costa, C., Sanchez, P. and Flos, R., 2008. Stocking density and its influence on shape of Senegalese sole adults. Aquaculture International 16, 333-343.

Bencze Rørå, A. M., Kvåle, A., Mørkøre, T., Rørvik, K.-A., Hallbjoørn, S., Steien, M. and Thomassen, S., 1998. Process yield, colour and sensory quality of smoked Atlantic salmon (*Salmo salar*) in relation to raw material characteristics. Food Research International 31, 601-609.

Blonk, R. J. W., Komen, H., Kamstra, A. and Van Arendonk, J. A. M., 2010. Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations of Common sole, *Solea solea*. Genetics 184, 1-7.

Burgos-Artizzu, X. P., Ribeiro, A., Tellaeche, A., Pajares, G. and Fernández-Quintanilla, C., 2009. Improving weed pressure assessment using digital images from an experience-based reasoning approach. Computers and Electronics in Agriculture 65, 176-185.

Chavanne, H., Norris, A., Haffray, P., Sonesson, A. K., Vandeputte, M., Chatain, B. and Boudry, P., 2008. Survey on the breeding practices in the European aquaculture industry, Reprofish Aquabreeding Workshop 2008.

Currens, K. P., Cameron, S. S., Hjort, R., Schreck, C. B. and Li, H. W., 1989. Effects of Different Feeding Regimes on the Morphometrics of Chinook Salmon (*Oncorhynchus tshawytscha*) and Rainbow Trout (*O. mykiss*). Copeia 1989, 689-695.

Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to quantitative genetics. Pearson Prentice Hall Harlow, Essex.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. and Thompson, R., 2006. ASReml User Guide Release 2.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Gjedrem, T., 1997. Flesh quality improvement in fish through breeding. Aquaculture International 5, 197-206.

Gjerde, B., 1989. Body traits in rainbow trout. I. Phenotypic means and standard deviations and sex effects. Aquaculture 80, 7-24.

Gjerde, B. and Schaeffer, L. R., 1989. Body traits in rainbow trout: II. Estimates of heritabilities and of phenotypic and genetic correlations. Aquaculture 80, 25-44.

Gonzalez, E. B., Nagasawa, K. and Umino, T., 2008. Stock enhancement program for black sea bream (*Acanthopagrus schlegelii*) in Hiroshima Bay: Monitoring the genetic effects. Aquaculture 276, 36-43.

Gonzalo, M. J., Brewer, M. T., Anderson, C., Sullivan, D., Gray, S. and van der Knaap, E., 2009. Tomato Fruit Shape Analysis Using Morphometric and Morphology Attributes Implemented in Tomato Analyzer Software Program. J. Amer. Soc. Hort. Sci. 134, 77-87.

Imsland, A. K., Foss, A., Conceiçao, L. E. C., Dinis, M. T., Delbare, D., Schram, E., Kamstra, A., Rema, P. and White, P., 2003. A review of the culture potential of *Solea solea* and *S. senegalensis*. Reviews in Fish Biology and Fisheries 13, 379-408.

Kause, A., Ritola, O. and Paananen, T., 2004. Breeding for improved appearance of large rainbow trout in two production environments. Aquaculture Research 35, 924-930.

Kause, A., Ritola, O., Paananen, T., Eskelinen, U. and Mantysaari, E., 2003. Big and beautiful? Quantitative genetic parameters for appearance of large rainbow trout. Journal of Fish Biology 62, 610-622.

Kause, A., Stien, L. H., Rungruangsak-Torrissen, K., Ritola, O., Ruohonen, K. and Kiessling, A., 2008. Image analysis as a tool to facilitate selective breeding of quality traits in rainbow trout. Livestock Science 114, 315-324.

Loy, A., Bronzi, P. and Molteni, S., 1999. Geometric morphometries in the characterisation of the cranial growth pattern of Adriatic sturgeon *Acipenser naccarii*, pp. 50-53.

Lynch, M. and Walsh, B., 1998. Genetic and Analysis of Quantitative traits. Sinauer Associates Inc., Sunderland, USA.

Merigot, B., Letourneur, Y. and Lecomte-Finiger, R., 2007. Characterization of local populations of the common sole *Solea solea* (*Pisces, Soleidae*) in the NW Mediterranean through otolith morphometrics and shape analysis. Marine Biology 151, 997-1008.

Meyer, G. E. and Neto, J. C., 2008. Verification of color vegetation indices for automated crop imaging applications. Computers and Electronics in Agriculture 63, 282-293.

Monet, G., Uyanik, A. and Champigneulle, A., 2006. Geometric morphometrics reveals sexual and genotypic dimorphisms in the brown trout. Aquatic Living Resources 19, 47-57.

Mrode, R. A., 2005. Linear models for the prediction of animal breeding values. CABI, Wallingford, UK.

Navarro, A., Zamorano, M. J., Hildebrandt, S., Ginés, R., Aguilera, C. and Afonso, J. M., 2009. Estimates of heritabilities and genetic correlations for growth and carcass traits in gilthead seabream (*Sparus auratus L.*), under industrial conditions. Aquaculture 289, 225-230.

Norris, A. T. and Cunningham, E. P., 2004. Estimates of phenotypic and genetic parameters for flesh colour traits in farmed Atlantic salmon based on multiple trait animal model. Livestock Production Science 89, 209-222.

Park, I. S., Woo, S. R., Song, Y. C. and Cho, S. H., 2007. Effects of starvation on morphometric characteristics of olive flounder, *Paralichthys olivaceus*. Ichthyological Research 54, 297-302.

Powell, J., White, I., Guy, D. and Brotherstone, S., 2008. Genetic parameters of production traits in Atlantic salmon (*Salmo salar*). Aquaculture 274, 225-231.

R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org.

Rasband, W. S., 2008. ImageJ. U.S. National Institutes of Health, Bethesda, Maryland, USA.

Russo, T., Pulcini, D., Bruner, E. and Cataudella, S., 2009. Shape and Size Variation: Growth and Development of the Dusky Grouper (*Epinephelus marginatus Lowe, 1834*). Journal of Morphology 270, 83-96.

Rutten, M. J. M., Bovenhuis, H. and Komen, H., 2004. Modeling fillet traits based on body measurements in three Nile tilapia strains (*Oreochromis niloticus L.*). Aquaculture 231, 113-122.

Rutten, M. J. M., Bovenhuis, H. and Komen, H., 2005. Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus L.*). Aquaculture 246, 125-132.

Rutten, M. J. M., Bijma, P., Woolliams, J. A. and van Arendonk, J. A. M., 2002. SelAction: Software to Predict Selection Response and Rate of Inbreeding in Livestock Breeding Programs. J Hered 93, 456-458.

Rye, M. and Gjerde, B., 1996. Phenotypic and genetic parameters of body composition traits and flesh colour in Atlantic salmon, *Salmo salar L*. Aquaculture Research 27, 121-133.

Shikano, T., 2008. Estimation of Quantitative Genetic Parameters Using Marker-Inferred Relatedness in Japanese Flounder: A Case Study of Upward Bias. J Hered 99, 94-104.

Steine, G., Alfnes, F. and Bencze Rørå, M., 2005. The Effect of Color on Consumer WTP for Farmed Salmon. Marine Resource Economics 20, 211-219.

Toline, C. A. and Baker, A. J., 1997. Evidence of a heritable component to body shape in the northern redbelly dace (*Phoxinus eos*). Can. J. Zool.-Rev. Can. Zool. 75, 1247-1253.

6

Minimising genotyping in BLUP selection schemes with natural mating

Robbert J.W. Blonk
Cécile Massault
Hans Komen
Johan A.M. van Arendonk

Submitted to Aquaculture

Abstract

In aquaculture, natural mating of parents in groups leads to highly skewed parental contributions and offspring with unknown pedigree. Mass selection in breeding schemes with natural mating of parents is therefore prone to yield high rates of inbreeding. To suppress rates of inbreeding, breeding schemes can use optimal contribution selection (OCS). With OCS, relationships between selection candidates are required. These can be reconstructed using genetic markers in populations with natural mating of parents. However, genotyping is expensive. Costs for genotyping can be minimised by performing OCS on a pre-selected fraction of the population. Animals in this fraction are selected on own performance (mass selection) and genotyped to establish pedigrees ("2-stage selection").

In this paper, we use stochastic simulation to determine the optimal size of the genotyped fraction in terms of response to selection (ΔG) and rates of inbreeding (ΔF) when using populations with natural mating of parents. Results are compared with mass selection. Our simulations show that genotyping in 2-stage selection can be minimised and that the optimal size of the genotyped fraction is 2.5% of the total population. With this size, ΔF can be restricted to 1%. With equal rates of inbreeding (1%), expected ΔG of mass selection (0.336 σ_p) slightly exceeds 2-stage selection (0.330 σ_p). However, the required number of selected parents (300) is larger than with 2-stage selection schemes where the required number of selected parents is 150 animals. We conclude that both mass selection and 2-stage selection can be used to genetically improve populations where natural mating of parents is used to obtain offspring, but it depends on the economic importance of the trait whether the increasing costs for broodstock maintenance outweighs costs for genotyping.

Introduction

In many aquaculture populations (e.g. Atlantic cod, *Gadus morhua*, Nile tilapia, *Orecochromis niloticus* and common sole, *Solea solea*), highly skewed parental contributions are observed when natural mating in groups is used to obtain offspring (Bekkevold, 2006; Fessehaye et al., 2006; Blonk et al., 2009). In such situations, breeding schemes using mass selection are prone to yield high rates of inbreeding (Blonk et al., 2009) unless many parents are used (Gjerde et al., 1996; Bentsen and Olesen, 2002). To restrict rates of inbreeding, breeding schemes can also use optimal contribution selection (OCS) theory (Meuwissen, 1997) combined with Best Linear Unbiased prediction of breeding values (BLUP). With OCS, animals are selected to obtain maximum response under a pre-set restriction to the rate of inbreeding; the difference between the mean relationship of the parental and offspring populations should not exceed a given restriction. Optimisation is achieved by varying the number of selected parents and calculating parental contributions to the next generation.

For OCS, relationships between selection candidates are required. As relationships between selection candidates are unavailable in natural mating populations, these should be reconstructed using genotypic information from DNA markers as e.g. microsatellites or SNP's (Duchesne et al., 2002; Fessehaye et al., 2006; Blonk et al., 2010). With large numbers of selection candidates, costs can increase dramatically. Sonesson (2005) showed that the number of selection candidates to be genotyped can be restricted using 2-stage selection including OCS on a mass selected fraction of the total population of selection candidates. In this method, only individuals with the best phenotypes are selected and genotyped (first stage: mass selection). From the genotyped fraction, animals are selected using OCS (second stage). In this way cost-return ratios of genotyping and selection response can be optimized.

Previous studies that used OCS included populations with controlled parental contributions and balanced mating designs (Bijma et al., 2001; Sonesson, 2005; Sonesson et al., 2005; Holtsmark et al., 2008). However, in situations where natural mating in groups is used to obtain offspring, assumptions of controlled parental contributions and mating do not hold. In addition, the number of parents may be fixed for practical and management reasons. Furthermore, rates of inbreeding are difficult to predict and

response to selection can be highly fluctuating depending on superiority of the highest contributing families. Although OCS under such circumstances is less effective when compared to situations with controlled mating, the method may be able to suppress rates of inbreeding.

Until now it remained unclear how many animals should be mass selected and genotyped in 2-stage breeding schemes where natural mating in groups is used to obtain offspring. Further, for such situations it is not known how big the number of selected parents should be to prevent rates of inbreeding reaching levels above the generally accepted level of 1% (Bijma, 2000). The objective of this paper is to determine effects of mass selected and genotyped fractions with OCS on response to selection and rates of inbreeding in aquaculture populations with natural mating of parents. Additionally, we evaluated effects of heritability of the selected trait and the required number of selected parents.

Materials and methods

Using stochastic simulation, response to selection (ΔG) and rates of inbreeding (ΔF) were determined in various breeding schemes with natural mating of parents. Two single trait breeding schemes were simulated: 1) mass selection and 2) 2-stage selection using OCS on a pre-selected and genotyped fraction (figure 6.1) which was obtained by mass selection from the population. In addition, we simulated 2-stage selection on a *random* mating population with *equal* full sib family sizes.

Two different heritabilities (h^2 = 0.2 and 0.5) and numbers of selected parents (N = 150, 200 and 300 with $n_{sires} = n_{dams} = ½\ N$) were simulated with N_{tot} = 8,000 offspring per generation. Selected parents were distributed over broodstock groups and the number of animals was kept at 50 per broodstock.

Natural mating of parents

Natural mating of parents was simulated using skewed parental contributions (c) to offspring generations. To construct skewed parental contributions, n_{sires} and n_{dams} random numbers were drawn from a gamma distribution with shape 0.75 and scale 1/3 [$x \sim \Gamma(0.75, 1/3)$] and transformed to relative contributions. This gamma distribution was chosen as it resembles the observed natural mating pattern of common sole (Blonk et al., 2009). Only sires and dams with $c_{sire} > 0$ or $c_{dam} > 0$ were used to contribute to the next generation and are further denoted as n_{csires} and n_{cdams}. Parental contributions were not

correlated to the selected trait. A full factorial mating design with dimensions n_{csires} by n_{cdams} was set up with the contribution of full sib families (c_{FS}) as the Kronecker product of the relative parental contributions.

To reflect that some full fib families are not contributing, a maximum number of possible families (n_{maxfam} = n_{csires} + n_{cdams}) was randomly drawn from the full factorial mating designs. The final relative contribution per full sib family (c_{FS}) was calculated as $c_{FS} \cdot 100 / \Sigma c_{FS}$.

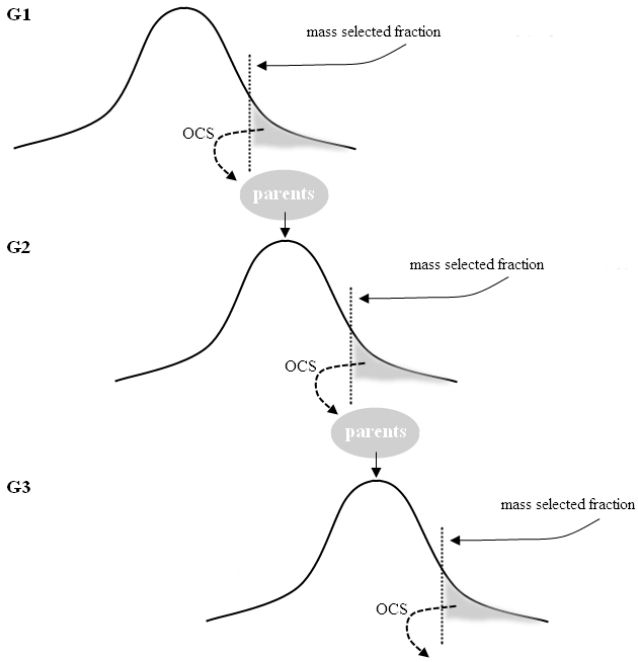


Figure 6.1. *Schematic overview of 2-stage selection including BLUP and OCS (Optimal contribution selection) on a mass selected fraction of the total population.*

Phenotypes

Sire and dam breeding values of founder populations were simulated as $z\sqrt{\left(h^2 \cdot \sigma_p^2\right)}$ with *z* as a normal deviate generated from a simulated normal distribution with *N~(0,1)*) and phenotypic variance σ_p^2 of 1. Founders were set unrelated (*F* = 0).

The total number of offspring generated in each subsequent generation was denoted as N_{tot}. For each full sib family, phenotypes of $N_{tot} \cdot c_{FS}$ offspring were calculated using the parental breeding values, a Mendelian sampling term including inbreeding coefficients of the parents and a residual as follows:

$$P_k = \frac{1}{2}A_i + \frac{1}{2}A_j + z_1 \cdot \sqrt{\frac{1}{2} \cdot h^2 \cdot \sigma_p^2 \cdot \left(1 - \frac{1}{2}\left(F_i + F_j\right)\right)} + z_2 \cdot \sqrt{\left(1 - h^2\right) \cdot \sigma_p^2}$$

where P_k is the phenotype of offspring k, A_i and A_j breeding values for sire i and dam j, F_i and F_j inbreeding coefficients of sire i and dam j and z_1 and z_2 normal deviates generated from a simulated normal distribution $N\sim(0,1)$.

Response and Rate of inbreeding

Offspring of all full sib families were pooled for each generation. ΔG per generation was calculated as $\bar{P}_{10} / 10$ where $\bar{P}_{10}$ is the mean phenotype of the offspring population at the tenth generation (ΔG was not calculated from the Bulmer equilibrium). Mean ΔF of the offspring population was calculated using $F_t = 1 - (1 - \Delta F)^t$ (Falconer and Mackay, 1996). Because ΔF was calculated relative to the level of inbreeding in the generation where Bulmer equilibrium was reached, the formula was rearranged to:

$$\Delta F = 1 - \left(\sqrt[(t - t_B)]{\frac{1 - F_t}{1 - F_{t_B}}} \right) \qquad 6.1$$

where t_B is the generation where Bulmer equilibrium was reached. Bulmer equilibrium was assumed to be in the fifth generation.

Mass selection breeding schemes

To compare results from 2-stage selection schemes, we simulated a breeding scheme with natural mating of parents and mass selection of offspring. In this scheme, a fixed number of animals (N) with highest phenotypes were selected each generation from N_{tot} = 8,000 pooled offspring and subsequently used as parents for the new generation. Results of ten generations were averaged over 30 replicates.

2-stage selection

With 2-stage selection, a fraction from all offspring was mass selected and genotyped (see figure 6.1 for a graphical presentation). Tested mass selected and genotyped fractions ranged from 10% to the smallest possible fraction i.e. N/N_{tot}.
ASReml (Gilmour et al., 2006) was used to obtain BLUP breeding values in the selected fraction with linear model $y_i = \mu + a_i + \varepsilon_i$ where a_i is the random animal effect and ε_i is the residual error. For BLUP, only phenotype and pedigree information from the selected fraction was used. Phenotypes and pedigree information of the other selection candidates were ignored, as in reality, only animals in the selected fractions would be genotyped.

With estimated breeding values and pedigree data, a fixed number of parents N was selected using OCS software package GENCONT (Meuwissen, 1997; 2002) with ΔF restricted to 1% per generation. With OCS, response to selection is maximised while restricting ΔF. The program is designed to optimise selection by setting optimal contributions of selected parents and by varying the number of selected parents each selection round.
In this paper OCS is used to suppress ΔF as the algorithm strives for suppression of ΔF through selection of less related animals when set ΔF restrictions are not reached. The optimal design of the breeding scheme with 2-stage selection may then be found by adjusting the fixed number of parents. For consistency with literature we maintained the term OCS in this document. To account for violation of contributions which were proposed by GENCONT due to natural mating, constraints of relatedness ($\bar{r}_{t+1}$) in the next generation were calculated as $\bar{r}_t + \Delta F(2 - \bar{r}_t)$ with $\bar{r}_t$ as the mean relatedness from the previous *parental* population and not from the actual *offspring* population (using the *previous* option in GENCONT).
Effects of 2-stage selection on ΔF and ΔG in situations with natural mating were analysed using simulation of ten generations and 30 replicates.

Effect of natural mating contributions

To determine effects of 2-stage selection and natural mating on mean relatedness in parental and offspring populations, one founder population G_0 (N = 200) with unrelated individuals was generated and reproduced once with simulated natural mating. From the resulting G_1 population (N_{tot} = 8,000), 200 parents were selected using 2-stage selection

with OCS from different mass selected and genotyped fractions (2.5 to 10%), and reproduced with natural mating to obtain G_2. Mean relatedness was calculated for parents selected from G_1 and for all offspring in G_2. Mean relatedness was obtained from 50 replications in each fraction.

The effect of skewness of parental contributions during natural mating was determined by simulating ten generations of 2-stage selection schemes with *random* mating parents and *equal* full sib family sizes. Random mating of parents was simulated by drawing a maximum number of possible families ($n_{maxfam} = n_{csires} + n_{cdams}$) from a full factorial mating design including all parents. Full sib family contribution c_{FS} was then calculated as $1 / n_{maxfam}$. Results are shown for heritabilities 0.2 and 0.5 in populations with 8,000 offspring and 200 selected animals. Results were averaged over 30 replicates.

Results

Mass selection

With stochastic simulation of ten generations of mass selection in populations with skewed contribution of parents, 200 selected parents and heritability 0.2, mean ΔF per generation is 1.5% (see figure 6.2). In this scenario, mean ΔG per generation is $0.352\sigma_P$. With 300 selected parents and heritability 0.2, ΔF is 1% and ΔG is 0.338 σ_P (figure 6.4). With 150 selected animals, ΔF increases to 2.1%. However, ΔG is also higher with 0.362 σ_P (figure 6.5). Compared to simulations with heritability 0.2, values of ΔF and ΔG increase when heritability is 0.5. For example, with a heritability of 0.5 and 200 selected parents, ΔF increases to 1.8% and ΔG increases to 0.805 σ_P (figure 6.3). Similar results are obtained for scenarios with 150 or 300 selected parents (data not shown).

2-stage selection

ΔF and ΔG of 2-stage selection breeding schemes with natural mating of parents are shown in figure 6.2 to 6.5 for different mass selected and genotyped fractions, number of selected parents and heritabilities. In figure 6.2 and 6.3, results are shown for 200 selected parents and heritabilities of respectively 0.2 and 0.5. In figure 6.4, results are shown for a heritability of 0.2 and 300 selected parents. In figure 6.5, ΔF and ΔG are presented for populations with 150 selected parents and the same heritability.

Observed patterns of *ΔF* and *ΔG* with increasing genotyped fractions are similar across different heritabilities and numbers of selected parents. With relatively small fractions, *ΔF* is initially high and sometimes above the set restriction of 1 % (e.g. fractions < 4%). This is particularly clear when heritability is 0.5 and 200 parents are selected. Here, *ΔF* is 1.3 with a fraction of 2.75% (figure 6.3).

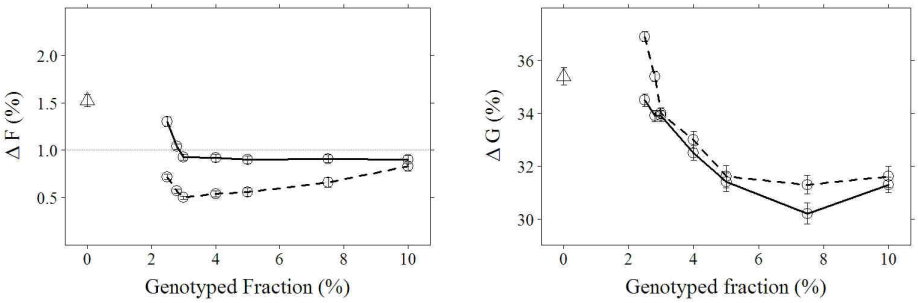


Figure 6.2. *Mean ΔF (in %) and ΔG (in % of σ_P) with standard error bars for a 2-stage breeding scheme including OCS on different genotyped fractions (population size: 8,000; 200 selected parents; heritability: 0.2; ΔF restriction: 1%; replicates: 30). Results are shown for natural mating populations with skewed parental contributions (solid line) and populations with random mating and equal full sib family sizes (dashed line). Results for mass selection on a natural mating population are depicted at genotyped fraction 0 (triangle).*

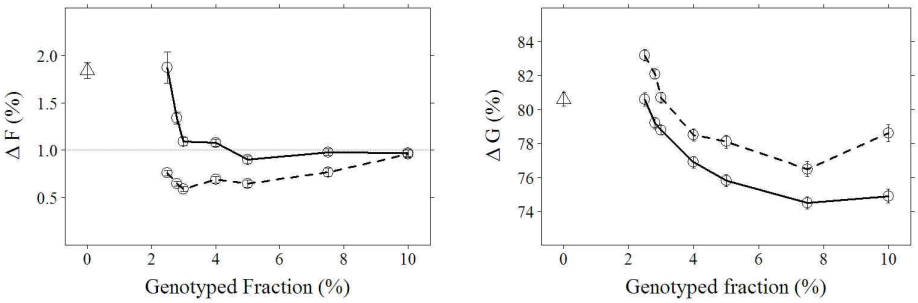


Figure 6.3. *Mean ΔF (in %) and ΔG (in % of σ_P) with standard error bars for a 2-stage breeding scheme including OCS on different genotyped fractions. As in figure 6.2 but with heritability 0.5.*

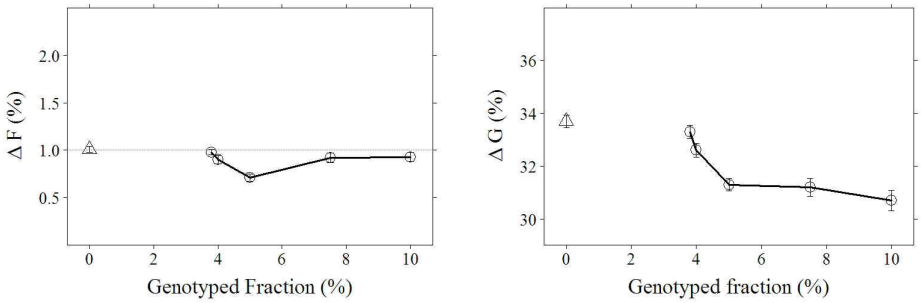


Figure 6.4. *Mean ΔF (in %) and ΔG (in % of σ_P) with standard error bars for a 2-stage breeding scheme including OCS on different genotyped fractions. As in figure 6.2 but with 300 selected parents, heritability 0.2 and without populations with random mating and equal full sib family sizes.*

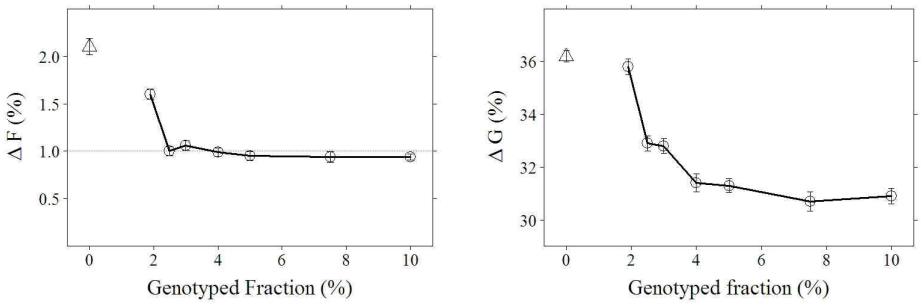


Figure 6.5. *Mean ΔF (in %) and ΔG (in % of σ_P) with standard error bars for a 2-stage breeding scheme including OCS on different genotyped fractions. As in figure 6.2 but with 150 selected parents, heritability 0.2 and without populations with random mating and equal full sib family sizes.*

With lower heritability, effects are less distinct but a similar pattern is present (figure 6.2). When fractions increase, *ΔF* decreases and stabilizes to 1% in all scenarios. The fact that *ΔF* is above 1% in small fractions is most likely due to the relative small number of families presented in these fractions. Consequently, on the long term, OCS can not suppress high *ΔF* due to lack of genetic diversity and OCS nears mass selection. As expected, *ΔF* and *ΔG* are close to the values obtained with mass selection at the smallest

possible fractions for each scenario. With increasing fractions, genetic diversity in fractions increases, leading towards more flexibility for OCS to select more diverse broodstocks and restrict ΔF.

In accordance with results for ΔF, with small fractions ΔG is initially high and close to results of mass selection. For example, with heritability 0.2 and 150 selected parents, response of 2-stage selection is between 0.36 and 0.33 σ_P at genotyped fractions 2 to 2.5% (figure 6.5). Selecting parents from larger fractions reduces ΔG as OCS is able to suppress ΔF; with fractions of approximately 5% and higher, overall responses stabilize to approximately 0.31 σ_P.

In agreement with results for mass selection, ΔF with 2-stage selection is lower for heritability 0.2 than for heritability 0.5 at smaller and equal fractions. Results are only shown for situations with 200 selected parents (figures 6.2 and 6.3). For example, with 200 selected parents and a genotyped fraction of 2.75%, ΔF is 1.3% when heritability is 0.5 (figure 6.2). Under the same conditions but with a heritability of 0.2, mean obtained ΔF is approximately 1% (figure 6.3). Similarly, ΔG is lower for heritability 0.2 than for heritability 0.5 at small fractions. With 200 selected parents and a genotyped fraction of 2.75%, ΔG is 0.338 σ_P with heritability 0.2 and 0.790 σ_P with heritability 0.5.

Effect of natural mating

The effect of skewness of contributions with natural mating of parents is reflected by the difference of mean relatedness in the selected parents and their offspring after natural mating. The mean relatedness in the offspring is always higher than in the parents. This is caused by many null or small parental contributions and few large parental contributions.

Effects of skewed contributions on ΔF are also illustrated when comparing selection on natural mating populations with populations where mating is random but where full sib family sizes are equal (e.g. figure 6.2). With random mating and equal full sib family sizes, parental contributions are less skewed than with natural mating. Overall ΔF in these populations is lower than in populations with natural mating (figures 6.2 and 6.3). This supports the idea that ΔF is increased due to skewness of contributions.

Effect of number of selected parents

At the smallest possible fractions (1.9% with 150 parents, 2.5% with 200 parents and 3.8% with 300 parents), the number of selected parents corresponds to the selection intensity. Consequently, 2-stage selection with 150 parents (figure 6.5) results in the highest ΔF (1.6%) and ΔG (0.360 σ_P). At this point, OCS is not able to restrict *ΔF* and 2-stage selection should roughly equal mass selection. In our results this is not always the case because at these small fractions, OCS can not always select sufficient new animals of both sexes.

Moreover, at small but equal genotyped fractions, breeding schemes with 2-stage selection of fewer parents may show lower *ΔF* than schemes where more parents are selected. For example, with genotyped fractions of 2.5% (200 animals), breeding schemes with selection of 200 parents yield *ΔF* of 1.3% (figure 6.2) whereas with selection of 150 parents a *ΔF* of 1% is reached (figure 6.5). The reason for this is that with selection of fewer parents from the same fraction, more room is left to suppress *ΔF* with OCS. In contrast, with more parents relatively many and possibly related animals need to be selected in order to gather the required number of parents. This may then result in higher *ΔF*. This implies that when the selected number of parents is increased, also the genotyped fraction should be increased to obtain the same *ΔF*.

At larger fractions, the number of selected parents has less effect on the achieved *ΔF* in 2-stage selection. In general, lower *ΔF* is realised when more parents are selected at larger fractions. For example, with fractions larger than 4%, *ΔF* is approximately 0.8 - 0.9 with 300 selected parents whereas *ΔF* is approximately 1% with 150 selected parents. This is probably caused by the fact that the number of selected parents was fixed. Consequently, with larger numbers of selected parents, OCS algorithms need to select more parents -and thus families- than required to achieve the restriction. Obtained *ΔF* is then lower than the set restriction. Note that when the number of parents to select is set *variable*, OCS algorithms will decrease the number of selected parents to increase *ΔF* (and *ΔG*) in such cases.

In contrast to results of mass selection, *ΔG* of 2-stage selection is not much affected by the selected number of parents at relatively large fractions. For example, at heritability 0.2, *ΔG* is approximately 0.310 σ_P for all scenarios with fractions larger than 4% (figures

6.2, 6.4 and 6.5). The small difference between scenarios is explained by the fact that OCS methods roughly maintain equal ΔF at larger fractions. Therefore, OCS more intensely suppresses ΔF when a relatively small number of parents is selected. This leads to (lower) values of ΔG which then are possibly equal to values obtained when larger numbers of parents are selected.

Mass selection vs. 2-stage selection

Stochastic simulation shows that ΔF under 2-stage selection (figures 6.2, 6.3, 6.4 and 6.5) is comparatively lower than values obtained with mass selection when natural mating populations are used to obtain offspring. ΔG of mass selection exceeded ΔG of 2-stage selection.

To compare response and efficiency, breeding schemes may be best compared under dimensions where ΔF is equal. Under 2-stage selection, ΔF of 1% is realised with 150 selected parents and a minimum genotyped fraction of 2.5% (figure 6.5) while mass selection requires selection of 300 parents (figure 6.4). Here ΔG is slightly higher for mass selection (0.336 σ_P) than for 2-stage selection (0.330 σ_P) but differences are small. With selection of 200 parents and a minimum genotyped fraction of 3%, ΔG of 2-stage selection is virtually equal to results of mass selection.

Discussion

In this paper we used stochastic simulation to assess ΔF and ΔG in breeding schemes where natural mating of parents is used to obtain offspring. We compared breeding schemes with mass selection and with 2-stage selection including mass selected and genotyped fractions. We determined effects of genotyped fractions in 2-stage selection schemes, effects of heritability of the selected trait and the number of selected parents. Additionally, we compared breeding schemes with skewed parental contributions due to natural mating of parents and schemes with random mating of parents with equal full sib family sizes.

Our results show that when natural mating of parents is used (i.e. when parental contributions are skewed), ΔG was always higher with mass selection than with 2-stage selection when the same number of parents is selected. However, observed ΔF confirms expectations from literature that mass selection schemes yield excessive ΔF in

populations with skewed parental contributions unless the number of selected parents is large (Meuwissen and Woolliams, 1994; Gjerde et al., 1996; Bentsen and Olesen, 2002). With equal numbers of selected parents, ΔF from 2-stage selection was generally lower than ΔF from mass selection (e.g. see figure 6.*2*). It was found that with skewed parental contributions, 2-stage selection schemes require minimum genotyped fractions of 2.5% and less parents (N = 150) than mass selection schemes (N = 300) to obtain ΔF of 1% when heritability is 0.2 (figure 6.2). However, when the number of selected parents increases, also larger genotyped fractions are required. When heritabilities increase, larger fractions are required to meet the set restriction of ΔF. It can be concluded that 2-stage selection is more effective when restricting ΔF than mass selection.

Effect of parental contributions

ΔF and ΔG of populations under selection are largely affected by type of selection, heritability of traits and the number of selected parents (Verrier et al., 1993; Caballero et al., 1996; Gjerde et al., 1996) as well as parental contributions and their skewness (Wray and Thompson, 1990; Wray et al., 1994; Bijma, 2000; Fernandez et al., 2003). Performance of OCS is optimal once parental contributions are fully controlled and the number of animals to select is free (Meuwissen, 1997; Hinrichs et al., 2006). In cases where parental contributions are controlled e.g. with artificial insemination methods in cattle, ΔF of offspring populations exactly follows ΔF based on parents with optimal contributions (Meuwissen, 1997; Hinrichs et al., 2006). However, with natural mating of parents in groups, control of contribution is limited as animals are free in choosing their partners. This has been shown in animals including e.g. sheep (Preston et al., 2005), fishes (Forsgren, 1997; Bekkevold, 2006) and birds (Wiley, 1973). Additionally, in fish, numbers of offspring vary considerably per mating (Brown et al., 2005; Blonk et al., 2009). Such variations of family size are also observed in situations with family specific reproductive problems or mortality (Sonesson, 2005; Vehvilainen et al., 2008). Under such circumstances, actual parental contributions often are not conform values imposed by OCS algorithms. Consequently, mean relatedness of offspring populations are likely to exceed mean relatedness based on selected parents.

It must be noted that in this paper: 1) optimal contributions as proposed by GENCONT were not followed because with natural mating of parents this is not possible and 2) the number of selected parents was not set variable, as this is often not done in practical fish

breeding. This implies that optimising capabilities of OCS in GENCONT are severely restricted in this setting and that OCS does not work in an optimal way. However, even though optimal contribution theory is not optimized for natural mating systems, this method can lower average relationships between selected parents and thus constrain ΔF. This is in line with our results where we showed that 2-stage selection including OCS generally yields lower ΔF than mass selection (e.g. figure 6.*2*).

Effect of natural mating

The impact of skewness of parental contributions on OCS methods is illustrated by the fact that selection on populations with random mating and equal full sib family sizes invariably yielded lower ΔF than selection on skewed parental contributions due to natural mating (e.g. see figure 6.4). The reason for this is that with random mating and equal full sib family sizes, it is likely that more families are included in mass selected and genotyped fractions. It follows that OCS routines then have more possibilities to select parents with lower average relationships.

Skewness of parental contributions due to natural mating of parents in groups may be limited by decreasing broodstock size further (i.e. dividing the number of selected parents into multiple sub groups) or by using single pair spawning tanks. However, this requires large infrastructure and is labour intensive. Moreover, broodstock size may also have an optimum size from a production point of view. In case of e.g. common sole, small broodstocks (< 6 animals) or single pair spawning were not successful in reproduction (Blonk, unpublished data) but this may be different for other species.

Effect of genotyped fractions

With OCS in breeding schemes where offspring is obtained with natural mating of parents, genotyping of selection candidates is required. However, genotyping is expensive and should be minimised. With 2-stage selection and natural mating of parents, our simulations show that genotyped fractions can be minimised to effectively restrict ΔF. Although genotyped fractions can be minimised with natural mating and skewed contribution of parents, a relatively large number of genotyped selection candidates is required to restrict ΔF to 1%. For example, it is shown that in a population of 8,000 selection candidates at least 200 (2.5%) animals need to be genotyped, whereas in a

population with controlled contributions this may be around 100 animals (Sonesson, 2005).

Effect of the number of selected parents

Performance of optimal contribution selection methods is optimal once contributions can be controlled and when the number of parents to select is free (Meuwissen, 1997). For example, when the expected ΔF with a certain selection of parents is above 1%, new animals are added to the selection to decrease ΔF. Vice versa, when the expected ΔF is below 1% and ΔG is not maximised, the number of selected parents is decreased to increase ΔG.

However, under practical circumstances the number of parents is often fixed, especially when the selected animals (nucleus) are the same as the parents for reproduction. One reason is that the number of broodstocks is limited to a maximum for economical reasons. Also, a minimum number of parents is often set to ensure reproductive output and profitability of farms. Although reproductive capacity of fishes is extremely high with animals often capable to produce tens of thousands of offspring (Gjerde et al., 1996; Komen et al., 2006), reproductive success can be fluctuating, especially under natural mating circumstances (Brown et al., 2005; Blonk et al., 2009).

Implications

Our results imply that both mass selection and 2-stage selection schemes can be used to genetically improve populations where natural mating of parents is used to obtain offspring. With ΔF of 1% and a heritability of 0.2, 2-stage selection schemes require 150 parents with a minimum genotyped fraction 2.5% whereas mass selection schemes need selection of 300 parents. Here, expected ΔG under mass selection is somewhat higher than with 2-stage selection, but differences are small. It is important to realize that with higher heritabilities and/or when more parents are required (e.g. for reproductive output or safety reasons), larger genotyped fractions are needed as well.

In general, costs for broodstock maintenance are high for breeding schemes using mass selection, but breeding schemes with 2-stage selection require costs for genotyping. It depends on the economic importance of the trait whether increase in ΔG due to mass selection outweighs increasing costs for broodstock maintenance.

With rapid developments in the field of genomics (e.g. Glover et al., 2010), application of OCS methods may become more accessible through reduction of costs. Also, accuracy of selection can be improved by estimation of molecular relatedness and Mendelian sampling terms (Hayes et al., 2009; Blonk et al., 2010). This means that there are good prospects for genetic improvement aquaculture industry when natural mating in groups is used to obtain offspring.

Acknowledgements

The authors thank Piter Bijma and Theo Meuwissen for their useful comments.

References

Bekkevold, D. X., 2006. Male size composition affects male reproductive variance in Atlantic cod *Gadus morhua* L. spawning aggregations. Journal of Fish Biology 69, 945-950.

Bentsen, H. B. and Olesen, I., 2002. Designing aquaculture mass selection programs to avoid high inbreeding rates. Aquaculture 204, 349-359.

Bijma, P., 2000. Long-term genetic contributions: prediction of rates of inbreeding and genetic gain in selected populations. PhD thesis. Wageningen University, Wageningen.

Bijma, P., Van Arendonk, J. A. and Woolliams, J. A., 2001. Predicting rates of inbreeding for livestock improvement schemes. J. Anim Sci. 79, 840-853.

Blonk, R. J. W., Komen, H., Kamstra, A. and Van Arendonk, J. A. M., 2010. Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations of Common sole, *Solea solea*. Genetics 184, 1-7.

Blonk, R. J. W., Komen, J., Kamstra, A., Crooijmans, R. P. M. A. and 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.

Brown, C. R., Woolliams, J. A. and McAndrew, B. J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247, 219-225.

Caballero, A., Santiago, E. and Toro, M. A., 1996. Systems of mating to reduce inbreeding in selected populations. Animal Science 62, 431-442.

Duchesne, P., Godbout, M. H. and Bernatchez, L., 2002. PAPA (Package for the Analysis of Parental Allocation) : A computer program for simulated and real parental allocation. Molecular Ecology Notes 2, 191-194.

Falconer, D. S. and Mackay, T. F. C., 1996. Introduction to quantitative genetics. Pearson Prentice Hall Harlow, Essex.

Fernandez, J. J., Toro, M. M. A. and Caballero, A. A., 2003. Fixed contributions designs vs. minimization of global coancestry to control inbreeding in small populations. Genetics 165, 885-894.

Fessehaye, Y., El-Bialy, Z., Rezk, M. A., Crooijmans, R., Bovenhuis, H. and Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. Aquaculture 256, 148-158.

Forsgren, E., 1997. Female sand gobies prefer good fathers over dominant males. Proceedings: Biological Sciences 264, 1283-1286.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. and Thompson, R., 2006. ASReml User Guide Release 2.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Gjerde, B., Gjøen, H. M. and Villanueva, B., 1996. Optimum designs for fish breeding programmes with constrained inbreeding mass selection for a normally distributed trait. Livestock Production Science 47, 59-72.

Glover, K., Hansen, M., Lien, S., Als, T., Hoyheim, B. and Skaala, O., 2010. A comparison of SNP and STR loci for delineating population structure and performing individual genetic assignment. BMC Genetics 11, 2.

Hayes, B. J., Visscher, P. M. and Goddard, M. E., 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91, 47-60.

Hinrichs, D., Wetten, M. and Meuwissen, T. H. E., 2006. An algorithm to compute optimal genetic contributions in selection programs with large numbers of candidates. J. Anim Sci. 84, 3212-3218.

Holtsmark, M., Klemetsdal, G., Sonesson, A. K. and Woolliams, J. A., 2008. Establishing a base population for a breeding program in aquaculture, from multiple subpopulations, differentiated by genetic drift: I. Effects of the number of subpopulations, heritability and mating strategies using optimum contribution selection. Aquaculture 274, 232-240.

Komen, J., Bovenhuis, H. and Van Arendonk, J. A. M., 2006. Consequences of reproductive characteristics for fish breeding schemes. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production.

Meuwissen, T. H. E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. Journal of Animal Science 75, 934-940.

Meuwissen, T. H. E., 2002. GENCONT: An operational tool fo controlling inbreeding in selection and conservation schemes. Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. 33, 769-770.

Meuwissen, T. H. E. and Woolliams, 1994. Effective sizes of livestock populations to prevent a decline in fitness. Theoretical and Applied Genetics 89, 1019-1026.

Preston, B. T., Stevenson, I. R., Pemberton, J. M., Coltman, D. W. and Wilson, K., 2005. Male Mate Choice Influences Female Promiscuity in Soay Sheep. Proceedings: Biological Sciences 272, 365-373.

Sonesson, A. A. K., 2005. A combination of walk-back and optimum contribution selection in fish: a simulation study. Genetics, Selection, Evolution 37, 587-599.

Sonesson, A. A. K., Gjerde, B. B. and Meuwissen, T. T. H. E., 2005. Truncation selection for BLUP-EBV and phenotypic values in fish breeding schemes. Aquaculture 243, 61-68.

Vehvilainen, H., Kause, A., Quinton, C., Koskinen, H. and Paananen, T., 2008. Survival of the Currently Fittest: Genetics of Rainbow Trout Survival Across Time and Space. Genetics 180, 507-516.

Verrier, E., Colleau, J. J. and Foulley, J. L., 1993. Long-term effects of selection based on the animal-model blup in a finite population. Theoretical and Applied Genetics 87, 446-454.

Wiley, R. H., 1973. Territoriality and non-random mating in sage grouse, *Centrocercus urophasianus*. Animal Behaviour 6, 85-169.

Wray, N. R. and Thompson, R., 1990. Prediction of rates of inbreeding in selected populations. Genetical Research 55, 41-54.

Wray, N. R., Woolliams, J. A. and Thompson, R., 1994. Prediction of rates of inbreeding in populations undergoing index selection. Theoretical and Applied Genetics 87, 878-892.

7

General Discussion

Introduction

The aim of this thesis was to design a breeding program for increased productivity of farmed common sole 1) using natural mating in groups to obtain offspring and 2) using present farm infrastructures as much as possible. The most important aspects to consider were pointed out in the general introduction (chapter 1) and individually analysed and discussed in chapters 2 to 6.

In the general discussion, findings of chapters 2 to 6 are used to present blue prints for a breeding program for common sole. Effects of natural mating of parents on rates of inbreeding and response to selection are put into context. Also, profitability of different breeding programs are elaborated and discussed. Further, the use of natural mating and its potentially positive effects are discussed with relation to genetic improvement of natural mating stocks. Finally, the effect of husbandry practices as grading on estimated breeding values is assessed.

Inbreeding and response to selection under natural mating

Contributions

In idealised populations with random and non-assortative mating systems, parental contributions are equal on average and variation of family size is random. Here, Poisson distributions are valid to represent family size (Wright, 1969; Falconer and Mackay, 1996). However, with natural mating of parents in groups, family size (and parental contribution) is generally not a matter of random deviation where some contributions happen to be larger than others. Here, parental contributions result from e.g. hierarchy within genders and assortative mating systems (Penn and Potts, 1999; Avise et al., 2002; Chenoweth and Blows, 2006). Such systems may be found across a wide range of species including fish (Reynolds and Gross, 1992; Forsgren, 1997), birds (Wiley, 1973) and mammals (Goldsworthy et al., 1999). For example, in natural mating systems with captive Atlantic cod, relatively large males had higher contribution than smaller males (Rowe, 2007). Further, male reproductive success depended on size differences between males and females: smaller size differences yielded more offspring (Bekkevold et al., 2002).

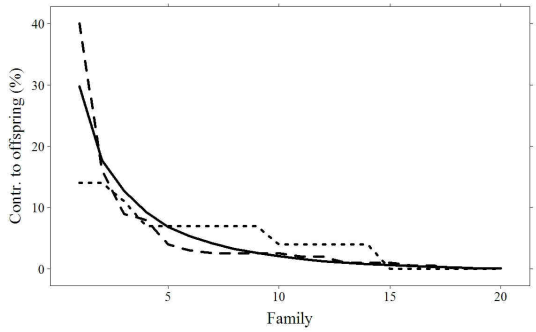


Figure 7.1 Simulated (continuous), observed (dashed) and poisson distributed (dotted) contribution of families to the total number of offspring obtained by natural mating of parents.

As a result of these mating systems, contribution of parents to the next generation is highly skewed, with only few parents producing majority of the offspring (Hutchings, 1999; Bekkevold et al., 2002; Brown et al., 2005; Bekkevold, 2006; Fessehaye et al., 2006b; Rowe, 2007). For common sole, parental contributions to the next generation are described in chapter 2 and results for one broodstock are shown in figure 7.1. From this figure it can be seen that poisson distribution of family size does not reflect patterns found with natural mating of parents. Hence, distributions may be better represented by Gamma distributions (e.g. with $\Gamma\sim(0.75,3)$ as shown in figure 7.1).

With highly skewed parental contribution and unbalanced family sizes, the chance of selecting a large number of animals from the same family increases dramatically. Simple mass selection programs are then expected to yield high rates of inbreeding (Falconer and Mackay, 1996). Chapter 6 showed that rates of inbreeding indeed reach 1.5% per generation when performing mass selection (figure 7.2) on offspring from natural mating parents with a selection proportion of 2.5% (200 parents selected from 8,000 animals). In ideal populations without selection, expected rates of inbreeding equal $1/2N_e$ where N_e is the effective population size. With 200 parents in such populations, the expected rate of inbreeding is 0.25%. With mass selection on a population with Poisson distributions and a selection percentage of 2.5%, predicted rates of inbreeding increase to 0.6% (SelAction software: Rutten et al., 2002). However, rates of inbreeding are still lower than under populations with skewed contributions due to natural mating of parents.

Selection under natural mating

The found rates of inbreeding under selection of populations that were obtained with natural mating of parents where above the maximum acceptable rate of inbreeding of 1% (Bijma, 2000). This implies important consequences for breeding programs which use natural mating. In chapter 6, the use of optimal contribution selection and best linear unbiased prediction (BLUP) of breeding values on a mass selected and genotyped fraction (“2-stage selection”) was proposed to overcome high rates of inbreeding under natural mating of parents (figure 7.2).

Optimal contribution selection with BLUP of breeding values (BLUP-OC) requires genetic relationships between selection candidates. To infer relationships, offspring from

natural mating of parents needs to be genotyped and in chapter 6, optimal numbers of mass selected animals to genotype were assessed for 2-stage selection. Results showed that it is possible to maintain rates of inbreeding below 1% when genotyping a mass selected fraction (i.e. selection based on phenotype alone) of 2.5% of the population at least, with 150 selected parents.

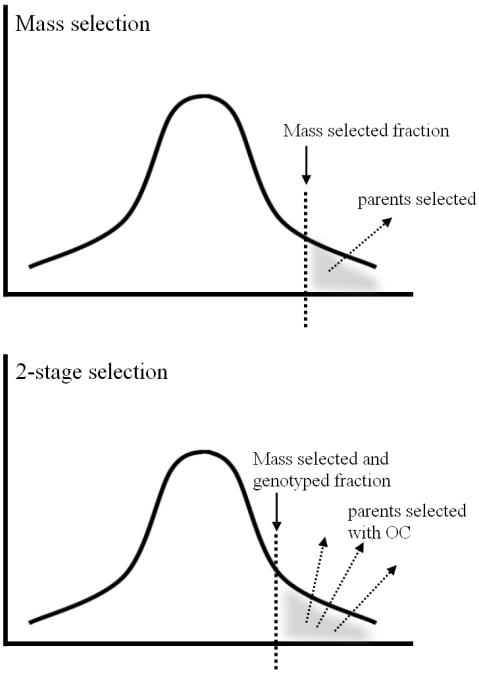


Figure 7.2 Schematic overview of mass selection and 2-stage selection with optimal contribution selection (OC) and BLUP of breeding values on a mass selected and genotyped fraction.

In chapter 6 it was also shown that with increasing nucleus sizes (i.e. decreasing selection intensities with equal population size), overall rates of inbreeding drop. For example, with a nucleus of 300 animals, ΔF nears 1% per generation when using only mass selection. Consequently, mass selection is possible under natural mating in groups when broodstock sizes are large enough. The use of mass selection seems an attractive idea as these schemes are fairly simple and cheap. However, there is little control on rates of

inbreeding (Meuwissen, 1997). In contrast, selection schemes with BLUP-OC monitor and control rates of inbreeding constantly.

The choice between a breeding program based on mass selection and a breeding program with more controlled rates of inbreeding is highly dependent on costs, returns and differences in rates of inbreeding (Belonsky and Kennedy, 1988). In chapter 6 it was shown that simulated mass selection with nucleus sizes of 300 parents gave a response of 0.336 σ_P and rates of inbreeding of 1% per generation. At the same level of inbreeding, 2-stage selection (including a "safety zone") with a genotyped fraction of 5% and a nucleus with 200 parents yielded a response of approximately 0.320 σ_P. It will depend on economic importance of the selected trait(s) whether an increase in response under mass selection outweighs genotyping costs of 2-stage selection or not.

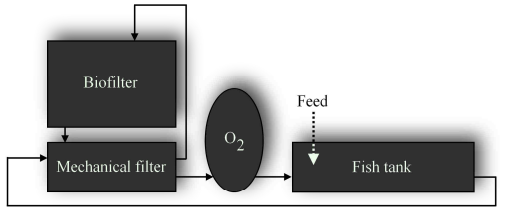
Economic value of growth rate in common sole

To evaluate costs and returns for breeding programs, economic values of selected traits should be known as well as costs for maintaining broodstocks, tagging and, if applicable, for testing environments and genotyping. The economic value of a trait is defined as the change in profitability as a consequence of one unit change of the trait, while keeping other traits constant (Hazel, 1943). Economic values can be estimated using partial derivatives from a profit function in which costs and returns are described as a function of physical, biological and economic parameters (e.g. in: van Arendonk, 1991; Jiang et al., 1998).

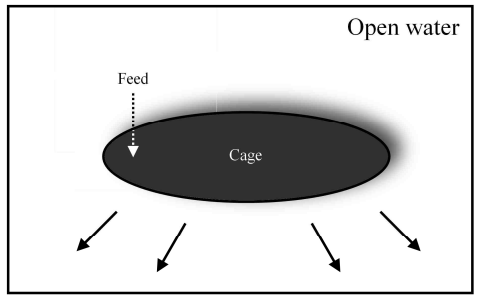
For calculation of economic values in common sole, parameters were obtained from a 90 MT recirculation production system with shallow raceways in the Netherlands (Solea B.V. IJmuiden). The economic value was calculated assuming unlimited filter capacity. Further assumptions were: feed conversion ratio: 1.2 $kg \cdot kg^{-1}$ (set fixed); feed price: 1.5 $€ \cdot kg^{-1}$; market price: 9 $€ \cdot kg^{-1}$. The overall density of the system was fixed on 900,000 animals. The obtained economic value per individual was 2.63 € / $g \cdot d^{-1}$, i.e. improvement of one unit growth rate (1 $g \cdot d^{-1}$) yields € 2.63 extra profit per animal [adapted from Hoon (2009)].

However, in aquaculture recirculation systems, costs as well as returns per kg produced are fixed, and thus independent of growth rates. This is because recirculation systems are

closed systems where filtration capacity of waste (determined by the dimensions of the system) is limited. As filtration capacity relates directly to maximum feeding level, (yearly) production is fixed (figure 7.3). Also, assuming that feed conversion rates remain constant, production can not increase with higher growth rates. Improved cost-return ratios of recirculation systems may then result from lower feed conversion rates, smaller standing stocks or lower density and thus lower costs for interests or insurance (personal communication A. Kamstra).



Recirculation system - "restricted" waste production



Cage system - "unrestricted" waste production

Figure 7.3 Schematic overview of aquaculture recirculation systems with restricted waste production and cage systems with virtually unrestricted waste production.

In contrast, with cage culture or flow through systems, production capacities are virtually unrestricted whereas densities are fixed. Higher growth rates then result in more production rounds per year. Fixed costs per kg produced then drop with increasing growth rates.

As bio economics of recirculation systems are complex (Kazmierczak and Caffey, 1995; De Ionno et al., 2006), accurately determining economic values in such an approach becomes very elaborate. However, to compare profitability of different breeding

programs, a simplified approach with restriction of density rather than production capacity may be used.

Comparison of breeding programs

Costs and predicted returns of two breeding programs for common sole with natural mating of parents were compared. Both a program with mass selection and a program including 2-stage selection with mass selected and genotyped fractions (figure 7.2) were analysed using simulation. Optimal designs were obtained using stochastic simulation procedures presented in chapter 6. In both programs, selection was from a population of 8,000 selection candidates. It was assumed that all selection candidates were randomly sampled from production populations directly and that no additional costs for test environments were required.

Parameters for simulation were obtained from a 90 MT production farm in the Netherlands. Mean bodyweight at harvest was 187 g, phenotypic variation was 8556 g^2, heritability for body weight at harvest was 0.23 and the generation interval was three years. Response to selection was expressed as linear growth (g/d) and was derived from body weight at harvest / production period. Growth rate in the founder population thus was 0.26 g/d with a production period of two years. The number of animals in one broodstock was set to 50.

Both programs were designed to obtain a mean ΔF of approximately 1% per generation. For mass selection programs, a nucleus with 350 animals was required. This equals 7 broodstocks. For the breeding program with 2-stage selection, 5% of the population (400 animals) was needed for mass selected and genotyped fractions whereas the required nucleus size was 250 animals (5 broodstocks). Response to selection in growth rate after simulating ten generations was 0.0559 g/d (40.8 g) per generation for mass selection and 0.0518 g/d (37.8 g) per generation for optimal contribution selection.

To calculate costs of breeding programs, maintenance (including electricity and depreciation), labour and feed for the broodstocks as well as tagging and genotyping were taken into account. Yearly costs for tagging and genotyping were calculated as one third of the total costs, as these were made only once per generation interval (3 years). Expected returns due to genetic improvement of production populations were calculated

as the product of the economic value for the total population (fixed density of 900,000 animals) and the response to selection (g/d) per year.

Table 7.1 Costs, predicted returns and increase of yearly profit in breeding programs with mass selection and breeding programs with 2-stage selection (optimal contribution selection on mass selected and genotyped fractions). For each breeding program, the required number and costs (€) of tanks and animals to genotype or tag are shown per component. Selection costs were averaged over the generation interval (GI). Costs for genotyping were set to € 20 per animal and were based on 10 markers per sample (personal communication W. van Haeringen).

		Mass selection			2-Stage selection		
Broodstock costs	**€/Tank**	**Tanks**	**€**	**Total**	**Tanks**	**€**	**Total**
Maintenance	3500	7	24,500		5	17,500	
Labour	1700	7	11,900		5	8,500	
Feed	450	7	3,150		5	2,250	
Subtotal				**39,550**			**28,250**
Selection costs	**€/Animal**	**Animals**	**€**		**Animals**	**€**	
Genotyping	20				400	8,000	
Tagging	2	350	700		400	800	
Subtotal per GI				700			8,800
Subtotal per year				**233**			**2,933**
Total costs				**39,783**			**31,183**
Expected returns				**132,192**			**122,472**
Expected profit				**92,409**			**91,289**

Compared to 2-stage selection, mass selection has higher costs for broodstock maintenance as more broodstocks are required (table 7.1). Adversely, for 2-stage selection costs are increased due to genotyping. Returns from mass selection programs were higher than with 2-stage selection. Due to this, expected yearly profit with mass selection was approximately € 1100 higher than with 2-stage selection. This small difference suggests that profit obtained by both programs can be considered as equal.

In the current breeding programs, returns were calculated for situations were the aim is on sale of animals for slaughter. However, profit of breeding programs strongly depends on the commercial focus of the enterprise. When companies aim at sale of fingerlings to grow out production farms, market prices are generally based on number rather than on body weight at slaughter. At this point, expected returns are equal for both breeding programs, provided that the number of fingerlings produced remains equal. Given this, profit of breeding programs largely depends on costs for broodstock maintenance and genotyping, thus also favouring breeding programs with optimal contribution selection and smaller nucleus size.

Regardless differences in expected profit, the most optimal breeding program in long term perspectives may be the one with a better control on rates of inbreeding, i.e. a breeding program including genotyping and optimal contribution selection. When natural mating in groups is used to obtain offspring, one must realise that mass selection runs the risk of accidentally high rates of inbreeding as there is no control on relatedness between animals (Gjerde et al., 1996b; Meuwissen, 1997; Bentsen and Olesen, 2002). With optimal contribution selection, pedigrees of selected animals are always monitored. This offers considerable advantages in emergency situations with unexpected high levels of relatedness due to e.g. extreme skewness of parental contributions in a previous generation. In such cases, one could decide to specifically search for additional parents by expanding the genotyped fraction to avoid increase of rates of inbreeding.

Another advantage of genotyping selection candidates is the possibility of including information on traits in the selection index that can not be recorded on live animals. For example, when fillet weight or degutted body weight is in the breeding goal, breeding programs may use information obtained from sibs of selection candidates in the index. With inclusion of sib information, relationships between animals are required and breeding programs using mass selection do not suffice. However, under the proposed breeding programs, use of extra sib information requires genotyping of additional offspring which subsequently increases costs. This has to be taken into account when assessing costs and revenues of incorporating correlated traits that can not be measured on live animals.

Accuracy of estimated breeding values from mass selection is lower and responses deviate more strongly from predicted values when compared to regular BLUP schemes where the full or a random sample of the population is pedigreed (Falconer and Mackay, 1996). However, with 2-stage selection including BLUP-OC on fractions as used in this thesis, only a fraction of the population is used for BLUP evaluation. Consequently, BLUP breeding values are biased and accuracy may be lower than expected. With regard to accuracy of selection alone, 2-stage selection may not be preferred above mass selection.

However, with currently rapid developments in the field of genomics, the use of genotyping in breeding programs is likely to lead to more efficient and accurate selection methods, e.g. the possibility of estimating Mendelian sampling terms and the use of genomic selection (Hayes and Goddard, 2008; Hayes et al., 2009). The increase in efficiency was also illustrated by chapter 4, where breeding values were more accurately estimated with molecular relatedness than with relationships from reconstructed pedigrees, especially in situations with skewed contributions of parents.

Consequences of natural mating

Natural mating in groups can be used to obtain offspring in many fish species. For example, artificial reproduction of soles has not proved successful at the time of this writing (Guzman et al., 2009). Other species can be reproduced artificially whereas in practice, natural mating is often used to obtain offspring, e.g. gilthead seabream (Brown et al., 2005), sea bass (Massault et al., 2009) and Nile tilapia (Fessehaye et al., 2006b). In these species, use of natural mating in groups is often a matter of efficiency as artificial reproduction can be very labour intensive.

Artificial reproduction of fishes often includes hormonal therapies and complex treatments to achieve efficient production of fertilised eggs. For example, stages of gonadal development of animals should carefully be monitored, stress should be minimised or avoided and one should consider the appropriate hormonal treatments, hormone-delivery-systems and latency periods. If such factors are not taken into account, treatments are likely to fail and for this reason, artificial reproduction is often associated with decreasing gamete quality (Guzman et al., 2009; Mylonas et al., 2010).

However, as opposed to artificial methods, natural mating systems are relatively free of stress, beneficial for animal welfare, free of hormone treatments and less laborious ways to obtain offspring in large amounts. Further benefits of the use of natural mating of species are 1) possible presence of mate choice and 2) indirect selection for reproductivity. These will be discussed below.

Mate choice

One effect of natural mating is the possible presence of mate choice (Wiley and Poston, 1996). Evidence that individuals choose partners based on specific characteristics such as colour, posture, behaviour or chemical compounds has been found for several species including birds (Wiley, 1973), insects (Lihoreau and Zimmer, 2007) but also fish (Reynolds and Gross, 1992; Landry et al., 2001). Often, this mechanism is based upon major histocompatibility complex (MHC) compatibility between partners to produce more heterozygous and more immunological robust offspring (Penn and Potts, 1999; Milinski, 2006; Forsberg et al., 2007). Hence, breeding programs with natural mating of parents are likely to select for more heterozygosity and disease resistance. This will increase efficiency of production, benefit animal welfare and reduce use of medicaments and antibiotics. In chapter 2, we showed that heterozygosity in offspring populations was higher than in parental populations of common sole. Although observed patterns may also result from accidental large contributions from relatively homozygous parents, this suggests presence of mate choice mechanisms in common sole. This implies that with the possible beneficial effects of mate choice, an increase in production and efficiency may be expected from incorporating natural mating into breeding programs.

Selection for reproductivity

Apart from being simple and labour extensive, the use of natural reproduction implies indirect selection for reproductivity of both sexes. Selection on reproductive traits might be feasible as heritability of reproductive output like e.g. fecundity or gonad weight, proved moderate in several species (Gjerde, 1986; Gjedrem, 2000; Charo-Karisa et al., 2007). Consequently, the reproductive capacity of natural mating populations may indirectly be maintained or improved since the chance of selecting families with genetically low reproductive capabilities (and thus small contributions) is relatively small.

In contrast, selection for improved reproduction is limited in artificial reproductive methods where quantities of eggs are often standardised and when reproductivity is not in the breeding goal.

With selection on growth, presence of correlated response of reproductive output should be determined. In other species, phenotypic and genetic correlations between production traits and reproductive traits were low to moderate but positive (Gjerde, 1986; Gall and Neira, 2004; Charo-Karisa et al., 2007). This suggests that with selection for production traits, reproductive traits can be affected negatively. If negative correlated responses of reproductive output are to be expected in sole, natural mating systems need additional monitoring of this trait. To prevent problems with reproduction on the long term, reproductive output needs to be both monitored (e.g. using gonad somatic indices) and included into the breeding goal. This probably will lead to lowered response to selection of growth and additional costs for monitoring of reproductive output.

Breeding goals – Correlated responses to selection for growth

The primary goal of most commercial selection programs is improvement of economical benefit. For most cultured species, the breeding goal should be (increased) body weight at harvest, or increased growth. In chapters 3, 4 and 5, attention was paid to estimation of heritabilities of the most important –especially for slow growing species such as common sole– production traits body weight, body height and body length.

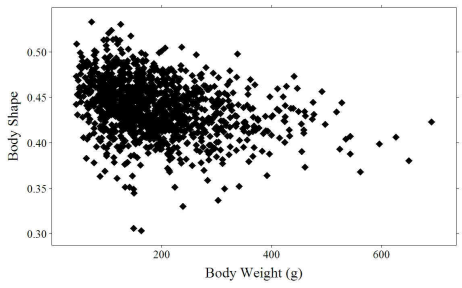


Fig 7.4 Relationship between body weight (g) and shape. Values of shape vary between 0 and 1. Smaller values represent more circular shapes.

When fish are sold un-filleted, shape of the animal may also play a role in consumer acceptance. In chapter 5 a negative correlation between body shape and body weight or body height was shown; increasing body weight at harvest leads to more circular shaped animals (figure 7.4). It is clear that shape should be included in the selection index, and selection should preferably be on maintenance of the present average shape.

Selection for early puberty

For many aquaculture species, early puberty is considered as a major problem due to its negative effects on production. During puberty, the production of gametes is initiated and development of gonads is at the cost of somatic growth (Zanuy et al., 2001; Taranger et al., 2010). Among many other factors, fast growth is thought to induce early puberty (figure 7.5). Here, the onset of puberty is induced by high adiposity and energy stores as a result of favourable farmed conditions but also due to genetic selection (Gjerde et al., 1994; Taranger et al., 2010). It follows that with a positive correlation between early puberty and growth (i.e. young puberty connects to fast growth), selection on growth ultimately decreases productivity of stocks.

Examples of positive correlations between early puberty and growth are found in rainbow trout and Atlantic salmon (Gjerde et al., 1994; Kause et al., 2003b), and can be considerable (up to approximately 0.4), e.g. in Atlantic cod or in Coho salmon, *Oncorhynchus kisutch* (Gjerde, 1986; Gall and Neira, 2004; Kolstad et al., 2006; Neira et al., 2006). From this, one can conclude that undesired negative responses of early puberty due to selection for growth can occur in common sole and needs to be quantified and monitored.

Although selection against early puberty can be relevant for some species, increasing mean age at puberty through selection will take several generations and thus may take long. In contrast, faster and less elaborate means to reduce early puberty are available and may be preferred above genetic alteration. For example, non-natural photoperiod regimes were associated with suppressed puberty and increased growth in Atlantic cod, turbot, Atlantic halibut and sea bass (Simensen et al., 2000; Hansen et al., 2001; Norberg et al., 2001; Rodríguez et al., 2001; Imsland et al., 2003a; Begtashi et al., 2004). A benefit of non-genetic alteration of age at maturation is that generation intervals for selection of new breeders may not be prolonged. Furthermore, selection methods for increased

growth and against early puberty while both traits are positively correlated will decrease response to selection of growth (compare effects of weighed selection indices with shape and bodyweight in chapter 5), which is not desired.

It depends on the marketed product if correlated response of early puberty negatively affects production volumes. If market size is higher than weight at puberty, negative effects can be expected (figure 7.5). However, in common sole, current age at maturation is minimal three years and the optimal (Dutch) market weight is generally before weight at time of puberty. At this point, no direct negative effect of early puberty due to fast growth is expected. On the contrary, if time to puberty is shortened, so are generation intervals and this may even increase response to selection of the first generations of breeding programs.

Selection from production stocks

In practical situations, grading of stocks is used to separate small and large size fractions to solve problems caused by social interactions. In chapter 3 it was shown that effects of grading on heritability estimates are predictable and can be corrected. However, effects on breeding values were not shown but might be relevant if correlations breeding values between graded and ungraded situations are low. To show effect of grading on estimated breeding values, the same approach as chapter 3 was used here. A simulated population was sorted to tanks using different size ranges to represent different time points after grading. Here, it was assumed that with time, size ranges of tanks increase. Consequently, phenotypic overlap of tanks increases. This reduces confounding of tanks and genetic effects. Equal to chapter 2, a full factorial mating design was simulated with 30 sires and 30 dams and with ten offspring per mating. Heritability was set at 0.3. No competition or tank effects were simulated.

Figure 7.6 shows correlation coefficients between breeding values of the non-graded population and breeding values of the same population but at different time points after grading. Shortly after grading, correlations of estimated breeding values between non graded populations and graded populations are approximately 0.7. However, from figure 7.6 it can be concluded that when more time is allowed between grading and breeding

value estimation, correlations of estimated breeding values between non-graded and graded populations can increase up to 0.99. At this point, ranking of estimated breeding values under graded situations equals that of estimated breeding values under non-graded situations.

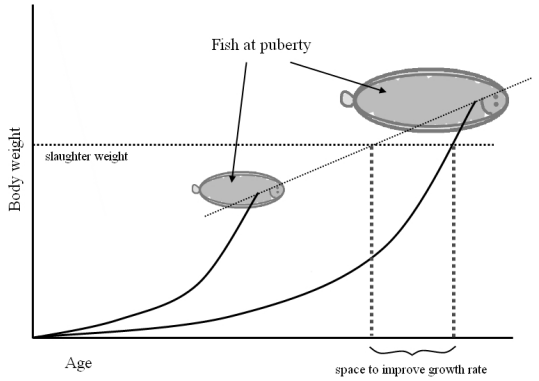


Figure 7.5 Relation (schematic) between growth (body weight / age) and onset of puberty (adapted from Taranger et al., 2010). The dotted line on the y-axis indicates market size. It is shown that market (slaughter) weight determines the space to improve growth through e.g. genetic selection.

The ability to use information from populations with grading on selection candidates has at least three advantages. First, selection may be from production populations directly, and selection candidates need not to be stocked in separate systems. Since no specialised infrastructure for testing environments is required, this significantly saves costs, especially for slow growing species such as common sole.

Secondly, selection is performed under situations where effects of social interactions are minimised. Social interactions as agonistic behaviour and competition for access to food is notorious for increasing phenotypic variation and mortality of stocks (Fessehaye et al., 2006a; Ellen et al., 2008). Analysis of social genetic variance of production traits has revealed “hidden” genetic variance in pigs and chicken (Bergsma et al., 2008; Ellen et al., 2008). This implies that population social behaviour can be altered through selection. However, genetic improvement of stocks may also focus on circumstances where social

interactions are reduced and where both aggressive and shy animals receive an equal chance to be selected. This is in same line with the breeding program “PROSPER” (Chevassus et al., 2004).

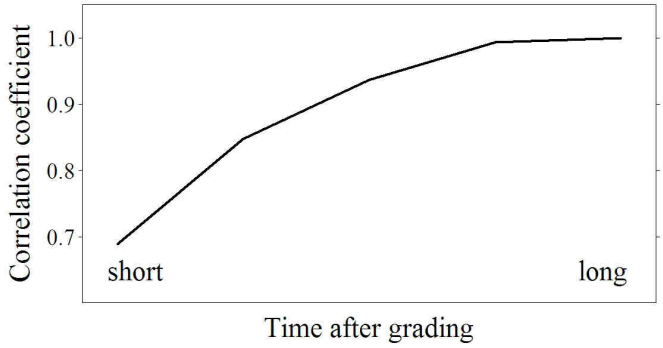


Figure 7.6 correlation coefficients between breeding values of non-graded populations and breeding values of the same populations at different time points after grading.

The third advantage is due to possible presence of “genotype by environment (GxE) interactions” which cause re-ranking of breeding values and selection candidates across environments. Here, testing environments without grading and with a high level of negative social interactions may differ greatly from “real life” production environments. Consequently, offspring performance may be optimal once parents are selected from similar environmental conditions.

GxE interactions can subsequently be estimated from the correlation between genetic values for a trait in different environments. For many species and traits, GxE interactions have been analysed (Mulder, 2007). For example, in pigs, moderate genetic correlations of production traits were shown between different levels of the pig production chain, i.e. test stations and production farms (Merks, 1988). In cattle, small effects of GxE interactions were found for production traits under different milking systems (Mulder et al., 2004), whereas larger effects were found for more different (e.g. tropical v.s. moderate) environments (Cienfuegos-Rivas et al., 1999). In two studies on production traits of Nile tilapia and Atlantic cod, no significant GxE interactions were found between

different farming conditions (Kolstad et al., 2006; Khaw et al., 2009) whereas in European sea bass, low GxE interactions were observed (Saillant et al., 2006; Dupont-Nivet et al., 2008). This shows that depending on the type of environment and species, GxE interactions can be relevant and should therefore be considered.

Recommendations for aquaculture breeding programs with natural mating of parents

In this thesis it is shown that natural mating of parents in groups can be used effectively once genotyping of selection candidates is used. Due to smaller required nucleus sizes, costs of breeding programs for a natural mating species including genotyping, BLUP analysis and optimal contribution selection (2-stage selection) are lower than costs for breeding programs based on mass selection. Response for breeding programs with mass selection was higher but differences between yearly profit of both breeding programs was small. Although accuracy of estimated breeding values may be biased under 2-stage selection, rates of inbreeding are far better controlled under 2-stage selection than under mass selection. To increase long term profit, it is therefore recommended to use 2-stage selection rather that mass selection under situations with natural mating parents.
Analysis showed that selection for shape of common sole is needed to compensate for negatively correlated response of shape due to selection for growth. However, correlations of production traits with reproductive output or early puberty in common sole are unknown and should be estimated. If undesirable correlations are present, reproductive output and early puberty should be monitored to prevent long term negative effects.

Profit of breeding programs may further increase when selection is from production stocks directly. Bias of estimated breeding values when using information from populations with grading is small but costs are lower due to absence of testing systems. Also, chances on unexpected low responses caused by re-ranking of genotypes due to GxE interactions between testing systems and “real life” production decrease.

Summary

Natural mating of parents in groups to obtain offspring for commercial aquaculture is used in many fish species. Sometimes this type of mating if preferred above artificial methods for its efficient way to obtain larvae. Examples are found in culture of species such as European sea bass, gilthead seabream and Atlantic cod. In many other fish species, such as common sole, artificial reproduction is impossible and stocks rely fully on natural mating. With natural mating of parents in groups, parental contributions are not a matter of random deviation where some contributions happen to be larger than others. Due to hierarchy within genders and assortative mating systems, natural mating of parents typically yields highly skewed parental contributions with few parents producing majority of the offspring. This type of reproduction implies that batches of produced eggs commonly include mixed families of highly variable sizes. Consequently, parental contributions as well as relationships between individual offspring are unknown. Incorporation of natural mating into breeding programs for genetic improvement of common sole involves several problems. First, simple breeding programs with mass selection are expected to yield high rates of inbreeding as the chance to select many members of one family is relatively large. Second, estimation of heritabilities or breeding values with relationships between animals (e.g. BLUP) requires genotyping. Genotyping is rather costly, especially when the number of candidates is large.

The aims of this thesis were to evaluate possibilities for a breeding program to increase productivity of farmed common sole using 1) natural mating in groups to obtain offspring and using 2) present farm infrastructures. Using this approach, a breeding program with natural mating parents is designed while costs are minimised.

In **chapter 2**, parental contributions and levels of coancestry between offspring were determined for one entire reproductive season of wild common sole kept in two broodstocks (28 animals in broodstock A; 20 animals in broodstock B). To estimate parental contributions, we performed parentage analysis on 48 parents and 2100 offspring with 10 highly polymorphic microsatellite markers. As expected, contribution of parents to offspring was highly skewed: in both broodstocks, five or less parental pairs produced more than half of the progeny. Compared to parental populations, few alleles were lost (9 of 134) and levels of heterozygosity (F_{IS}) increased in offspring populations. Next, coefficients of coancestry in offspring were calculated using parental contributions and parental relatedness from a relatedness estimator. Levels of coancestry in progeny were substantially high (2.0 and 4.9%), suggesting high rates of inbreeding. This shows that

natural mating in groups can result in significant inbreeding in future generations. This chapter also shows that commonly used population genetic parameters as loss of alleles and change in level of heterozygosity should not be used as indicators of rates of inbreeding in small captive populations. A quantitative method with contributions and relatedness ($c'Ac = 2\bar{f}$) should rather be used.

In commercial circumstances, grading of animals is used to reduce negative social interactions as competition for food. However, grading potentially introduces bias in estimates of genetic parameters. In **chapter 3**, effects of grading untagged fish on estimation of genetic parameters of body weight and body length was analysed using data from simulated populations with different heritabilities and a real dataset with 1336 animals. Simulated datasets demonstrated that heritability of graded traits (body length) was underestimated shortly after grading when overlap of phenotypic ranges in tanks is small. At later stages after grading, i.e. when phenotypic overlap of tanks increases, underestimation of genetic parameters decreased. Underestimation of genetic parameters over different phenotypic distributions of fish over tanks were predictable across heritabilities and correlations. Consequently, upper estimates of heritability can be determined once size distributions of fish over tanks are known. In the real dataset, estimated heritability for body length was 0.28 (±0.11) and extrapolation with results from simulation suggested that the true heritability was 0.40. Grading had a minor effect on estimated heritabilities of the correlated trait body weight. In the real dataset heritability for body weight was estimated at 0.21 (±0.09). These results show that that there is enough genetic variation in production parameters for common sole to genetically improve populations.

With missing parents and natural mating in groups, commonly used methods that reconstruct explicit pedigrees lead to loss of data as not all offspring can be allocated to their parents. Relatedness estimators however, infer relationships between all animals sampled. **Chapter 4** compares accuracy of breeding values estimated with relationships from a reconstructed pedigree and from a relatedness estimator (“molecular relatedness”) when parental contributions are skewed and when few (10) but highly polymorphic microsatellite markers are used. Due to missing parents, a pedigree could be reconstructed for 1336 animals (reduced dataset) out of a dataset with 1953 harvest size offspring (full dataset) and 51 known putative parents. In the reduced dataset, cross validation showed that accuracy of estimated breeding values was 0.54 with pedigree

reconstruction and 0.55 with molecular relatedness. In the full dataset, accuracy of estimated breeding values from molecular relatedness increased to 0.60. This indicates that pedigree reconstruction and molecular relatedness predict breeding values equally well in a population with skewed contributions to families and with few markers. An advantage of molecular relatedness is that loss of genotyping data is minimised while accuracy increases. However, compared to pedigree reconstruction methods, molecular relatedness as used in this study is less accurate for estimating coefficients of inbreeding and should therefore not be used in breeding programs on the long term.

From a marketing point of view it may be desired to maintain "natural" shape of fish which are sold un-filleted. In **chapter 5**, the need to include shape into a breeding goal for common sole was discussed using data on 1222 pedigreed animals at harvest size. Comparison of manual measurements and digital image analysis of shape showed that estimated genetic parameters were similar across methods. Also, correlations of measurements were high between methods. However, digital image analysis had significantly lower standard errors and was much faster. Heritability for shape was 0.34 (±0.11). Body weight correlated more strongly with body height than with body length. This indicates that body height is a better predictor for body weight. The genetic correlation between shape and body weight was –0.44. The data suggested that selecting for increased body weight at harvest yields more circular shaped fish. However, by using a selection index for shape and body weight, it is possible to maintain shape while improving body weight at harvest.

The skewness of parental contributions and heritabilities as obtained in the previous chapters were used in **chapter 6** to evaluate two breeding programs with natural mating of parents. First, a breeding program with mass selection was simulated. Second, a breeding program with best linear unbiased prediction of breeding values and optimal contribution selection of fixed nucleus sizes from a mass selected and genotyped fraction of a population (n=8,000) was simulated (2-stage selection). With breeding programs including 2-stage selection, a restriction of 1% was set on the rate of inbreeding. In this study, excessive high rates of inbreeding (>1.5%) were observed after 10 generations of mass selection unless a large nucleus (300 parents) was used. With a nucleus of 300 animals and mass selection, rates of inbreeding neared 1% and response per generation was 0.336 σ_P. With 2-stage selection, rates of inbreeding and response were restricted with mass selected and genotyped fractions 2.5 % of the population. At this point, the rate

of inbreeding was 1% and response was 0.330 σ_p when nucleus size was 150 and heritability 0.2. This study indicates that mass selection schemes can be used in breeding programs where natural mating of parents in groups is used. However, a large nucleus size is required. With 2-stage selection, smaller nucleus sizes are required but costs for genotyping (although minimised) are to be made and responses were lower. It depends on the economic importance of the trait whether the increase in response of mass selection outweighs the increasing costs for broodstock maintenance.

In the general discussion (**chapter 7**), several aspects are discussed with regard to development of a breeding program including natural mating of parents. First, in chapter 6 proposed breeding programs were evaluated using overall costs and economical values for growth of common sole in a 90 tonnes recirculation system in the Netherlands. Response and yearly profit of mass selection were higher than with 2-stage selection but differences were small. Benefits of 2-stage selection are a better control on rates of inbreeding and the possibility to include sib information on traits which can not be recorded on live animals.

Consequences of including natural mating into breeding programs were discussed. It was suggested that possible presence of mate choice systems may increase variability and immunological resistance of offspring. Further, with natural mating of parents, breeding programs automatically select for reproductivity whereas programs with artificial reproduction do not. However, possible negative correlations between production parameters and reproductivity imply that reproductivity should be monitored.

Correlated responses due to selection on body weight were discussed. It was shown that inclusion of body shape into breeding goals for genetic improvement of common sole is required as body weight negatively correlates with shape. In addition, means to prevent early puberty and concomitantly decreasing growth rates were discussed. It was suggested that selection to prevent early puberty might be possible, but that husbandry practices as changing photoperiod regimes are likely to yield faster response. Moreover, selection on late puberty prolongs generation intervals and therefore slows down yearly response of breeding programs whereas non-genetic means do not. However, it was suggested that for slow growing species and where weight at harvest is lower than weight at puberty, decreasing time to puberty also decreases generation interval. This ultimately increases response to selection of breeding programs.

Selection from production stocks which undergo grading is likely to increase profitability of breeding programs, since no separate test environments are required. This is especially relevant for slow growing species. Moreover, selection from production circumstances is more optimal if genotype by environment interactions are present. Effects of grading on estimated breeding values were discussed using the same approach as in chapter 3. It was shown that estimated breeding values from non graded populations and populations shortly after grading correlated with 0.7. When more time is allowed between grading and breeding value estimation, correlations increase to 0.99. This implies that breeding values can be estimated from graded fish once phenotypic distributions of tanks have some overlap.

Conclusions

From this PhD thesis, the following conclusions can be made:

- There is enough genetic variation in production parameters for common sole to genetically improve populations.
- Body shape should be included into the breeding goal when selecting in increased growth of common sole.
- Optimal contribution selection with best linear unbiased prediction of breeding values is recommended to control rates of inbreeding in selected populations with natural mating of parents.
- The use of a molecular relatedness estimator is recommended to increase efficiency of genotyping once parental genotypes are (partly) lacking and to increase accuracy of breeding value estimation in populations with natural mating of parents.
- Animals can be selected from production circumstances with grading.

Samenvatting

Natuurlijke voortplanting van ouderdieren in groepen om nakomelingen te verkrijgen is een veel gebruikte methode binnen de commerciële aquacultuur. Doordat vislarven bij natuurlijke voortplanting van ouders relatief efficiënt gewonnen kunnen worden verkrijgt deze methode van voortplanting vaak de voorkeur boven kunstmatige methoden. Voorbeelden hiervan zijn kweek van (Europese) zeebaars, goudbrasem, en Atlantische kabeljauw. In veel andere soorten zoals bijvoorbeeld Noordzeetong is kunstmatige voortplanting niet mogelijk. Hier is kweek volledig afhankelijk van natuurlijke voortplanting. Wanneer groepen ouderdieren op natuurlijke wijze voortplanten zijn contributies van ouders sterk scheef verdeeld door hiërarchische systemen binnen geslachten en door niet-willekeurige paringsystemen. Hierdoor produceert een klein aantal ouders het overgrote deel van de nakomelingen. Als gevolg van deze wijze van reproductie bevatten partijen geproduceerde eieren meerdere families van sterk verschillende groottes en zijn zowel individuele ouderlijke contributies als verwantschappen tussen nakomelingen onbekend.

Het inpassen van natuurlijke voortplantingsystemen in fokprogramma's ter verbetering van productiviteit van gekweekte Noordzeetong stuit op verschillende problemen. Ten eerste hebben fokprogramma's die enkel selecteren op de beste fenotypes een grote kans op hoge inteelttoename omdat de kans op selectie van veel leden van één families groot is. Ten tweede is genotyperen van dieren nodig voor het schatten van erfelijkheidsgraden en op verwantschap gebaseerde fokwaarden. Genotyperen is kostbaar, vooral wanneer het aantal kandidaten groot is.

Het doel van dit proefschrift is het evalueren van mogelijkheden voor een fokprogramma ter verbetering van productiviteit van gekweekte Noordzeetong, daarbij gebruik makend van 1) natuurlijke voortplanting in groepen en 2) de aanwezige infrastructuur van kwekerijen zodat respons maximaal is en kosten worden geminimaliseerd.

In **hoofdstuk 2** worden ouderlijke contributies en niveaus van verwantschap (coancestry) bepaald voor twee groepen nakomelingen en hun ouders. Om ouderlijke contributies te schatten wordt een ouderschapsanalyse uitgevoerd met 10 DNA merkers (microsatellieten).

Zoals verwacht is de contributie van ouders aan de populatie nakomelingen scheef verdeeld: in beide oudergroepen wordt meer dan de helft van de nakomelingen geproduceerd door minder dan zes ouders. Verlies van allelen (gekoppeld aan genetische

variatie) in de populatie nakomelingen is beperkt: slechts 9 van de 134 allelen worden niet teruggevonden en niveaus van heterozygotie nemen toe ten opzichte van de ouders. Deze parameters uit de populatiegenetica wekken de suggestie dat het verlies van genetische variatie beperkt is. Echter, niveaus van verwantschap in de nakomelingen, berekend aan de hand van ouderlijke contributies en hun geschatte verwantschappen, tonen aan dat de eigenlijke toename van verwantschap hoog is (2.0 and 4.9%). Dit resultaat geeft aan dat natuurlijke voortplanting van ouders in groepen kan leiden tot aanzienlijke inteelttoename in toekomstige generaties.

In commerciële kweekomstandigheden wordt sorteren van dieren gebruikt om negatieve sociale interacties zoals competitie om voer te verminderen. Bij sorteren worden verschillende groottes (lengtes) van dieren gescheiden en in aparte tanks geplaatst. Zodat kleinere dieren een betere kans tot voeropname krijgen. Echter, sorteren van ongemerkte dieren veroorzaakt een systematische fout bij het schatten van genetische parameters omdat genetische effecten niet gescheiden kunnen worden van bakeffecten. De best presterende dieren zitten namelijk in dezelfde tank en het is moeilijk om te bepalen of prestatie bepaald werd door het genotype of door de bak. In **hoofdstuk 3** wordt het effect van sorteren op het schatten van genetische parameters bepaald. Hiertoe worden de erfelijkheidsgraad en correlaties tussen lichaamsgewicht en lichaamslengte bepaald op verschillende tijdstippen na sorteren. Hierbij is aangenomen dat de overlap tussen fenotypes in tanks na sorteren toeneemt waardoor de systematische fout van de geschatte parameters kleiner wordt. Gesimuleerde datasets tonen aan dat erfelijkheid onderschat wordt door sorteren maar dat onderschatting van genetische parameters plaatsvindt volgens voorspelbare patronen over de tijd. Sorteren heeft minder effect op gecorreleerde kenmerken, vooral wanneer de daadwerkelijke correlatie 0.5 is. Hierdoor is het mogelijk om een geschatte erfelijkheid te extrapoleren naar de maximale daadwerkelijke waarden, mits de verdeling van dieren over tanks bekend is.
In een echte dataset wordt de erfelijkheidsgraad voor lichaamslengte geschat op 0.28 (±0.11). Extrapolatie van resultaten wijst op een daadwerkelijke erfelijkheidsgraad van 0.40 voor lichaamslengte en een erfelijkheid van 0.36 voor het gecorreleerde kenmerk lichaamsgewicht. Deze resultaten tonen aan dat er in gekweekte Noordzeetong voldoende genetische variatie is voor effectieve selectie op groei.

Naast phenotypische kenmerken zijn verwantschappen tussen dieren nodig om genetische parameters te schatten. In populaties die geproduceerd zijn middels natuurlijke voortplanting van ouders in groepen kunnen verwantschappen worden bepaald door reconstructie van stambomen met DNA merkerinformatie van ouders en nakomelingen. Echter, wanneer DNA informatie van (een aantal) ouders mist, leiden methoden voor stamboomreconstructie tot verlies van data. Als alternatief kunnen verwantschappen geschat worden door enkel DNA informatie van nakomelingen te gebruiken. **Hoofdstuk 4** vergelijkt nauwkeurigheid van fokwaarden die geschat zijn met gereconstrueerde stambomen of met geschatte verwantschappen ("moleculaire verwantschappen") in een situatie met scheef verdeelde ouderlijke contributies en tien DNA merkers. Een stamboom wordt gereconstrueerd voor 1336 dieren (verkleinde dataset) van een populatie met 1953 dieren (volledige dataset) en met 51 bekende, mogelijke ouders. De resultaten van de kleine dataset geven aan dat fokwaarden even goed kunnen worden voorspeld met beide methoden. Een voordeel van moleculaire verwantschap is dat alle data gebruikt kan worden wat een hogere nauwkeurigheid oplevert.

Vanuit een marketingperspectief is een natuurlijke visvorm gewenst wanneer vis als geheel (gestript) en niet gefileerd wordt verkocht. De noodzaak om visvorm op te nemen in het fokdoel voor Noordzeetong wordt bediscussieerd in **hoofdstuk 5**. Handmatige metingen en digital image analysis worden vergeleken als methodes om visvorm te bepalen. Geschatte genetische parameters verschillen nauwelijks tussen methodes en correlaties tussen methodes zijn hoog. Echter, standaardfouten zijn significant lager met digitale image analyse en ook verkort deze methode de tijd benodigd voor data verzamelen.
De geschatte erfelijkheidsgraad van visvorm is 0.34 (±0.11). De geschatte genetische correlatie tussen visvorm en gewicht is –0.44. Dit suggereert dat selectie op lichaamsgewicht leidt tot rondere visvormen. Door een selectie-index te gebruiken met zowel gewicht als vorm kan de huidige "natuurlijke" vorm van Noordzeetong behouden worden terwijl gewicht op leeftijd verbetert.

In **hoofdstuk 6** worden twee fokprogramma's met natuurlijke voortplanting van ouders geëvalueerd. Als eerste wordt een fokprogramma met selectie op de beste fenotypes gesimuleerd. Daarnaast wordt een 2-staps fokprogramma gesimuleerd dat gebruik maakt

van fokwaarden van een op fenotype geselecteerde fractie van de populatie (stap 1) en dat dieren selecteert op basis van optimale contributies om inteelttoename te beperken en respons te maximaliseren (stap 2). Fokwaarden worden geschat met behulp van verwantschappen. Hierbij word aangenomen dat verwantschappen werden bepaald door middel van met DNA merkers gereconstrueerde stambomen.
Dit hoofdstuk toont aan dat selectie op fenotype gebruikt kan worden in fokprogramma's met natuurlijke voortplanting in groepen. Echter, hier is een relatief grote groep ouderdieren nodig. Met 2-staps selectie volstaan kleinere groepen maar moeten kosten gemaakt worden voor genotyperen. Daarnaast is de respons over het algemeen lager. Afhankelijk van de economisch waarde van het geselecteerde kenmerk zal afhangen of de selectierespons van selectie op fenotype compenseert voor kosten voor grotere groepen ouderdieren.

In de algemene discussie (**hoofdstuk 7**) worden meerdere aspecten met betrekking tot de ontwikkeling van een fokprogramma met natuurlijke voortplanting van ouders besproken. De in hoofdstuk 6 besproken fokprogramma's worden geëvalueerd op basis van kosten en economische waarden voor groei van Noordzeetong. De winstgevendheid is het hoogst voor een fokprogramma met selectie op fenotype, maar de verschillen zijn klein. Voordelen van 2-staps selectie zijn onder andere dat controle op inteelttoename beter is en dat het mogelijkheid is te selecteren op kenmerken die niet op levende dieren gemeten kunnen worden (bijvoorbeeld filetpercentage).
Consequenties van het gebruik van natuurlijke voortplanting van groepen ouders worden besproken in het licht van mogelijke aanwezigheid van (niet-willekeurige) partnerkeuze systemen. Dit leidt mogelijk tot verhoogde variatie en immunologische weerstand van nakomelingen. Daarnaast wordt besproken dat fokprogramma's moeten letten op mogelijke negatieve correlaties tussen productie- en reproductiekenmerken. Vervroegde puberteit en maturatie verlagen vaak groeisnelheden en vormen daarom een groot probleem in de hedendaagse aquacultuur. Besproken wordt dat aanpassing van kweekomstandigheden, zoals bijvoorbeeld lichtregimes, beter en sneller resultaat zullen leveren dan selectiemethoden. In het geval van langzaam groeiende soorten zoals Noordzeetong zal een vervroegde puberteit in eerste instantie leiden to verkorte generatie-intervallen en dus verhoogde respons omdat slachtgewicht meestal lager is dan gewicht bij puberteit.

In de algemene discussie wordt tevens bediscussieerd dat selectie onder productieomstandigheden met sorteren de winstgevendheid van fokprogramma's verhoogt doordat geen aparte testomgeving nodig is. Ook wordt aangegeven dat selectie onder productieomstandigheden voordelen heeft wanneer genotype-milieu interacties aanwezig zijn. Genotype-milieu interacties houden in dat dieren (genotypen) verschillend presteren onder verschillende milieus. In andere woorden: de ranglijst van fokwaarden is afhankelijk van het milieu. Het effect van sorteren op geschatte fokwaarden is berekend met de methode van hoofdstuk 3. Resultaten geven aan dat fokwaarden van gesorteerde dieren met redelijke zekerheid geschat kunnen worden zodra de fenotypische spreiding van tanks overlapt.

Conclusies

De belangrijkste conclusies uit dit proefschrift zijn:

- Er is voldoende genetische variatie in productieparameters voor effectieve verbetering van productie van Noordzeetong door selectie.
- Visvorm moet opgenomen worden in het fokdoel bij selectie op groei van Noordzeetong.
- Optimale contributie selectie met "best linear unbiased prediction of breeding values" is aanbevolen om inteelttoename in geselecteerde populaties met natuurlijke voortplanting in groepen te controleren.
- Het is aanbevolen moleculaire verwantschappen te gebruiken in populaties met natuurlijke voortplanting in groepen wanneer genotypen van ouders gedeeltelijk missen en om nauwkeurigheid van fokwaardeschatting te verbeteren.
- Dieren kunnen effectief geselecteerd worden vanuit productieomstandigheden met sorteren.

Dankwoord

En dan nu het hoofdstuk waar het minst over gedacht, gewikt en gewogen is. Niet uit onachtzaamheid, maar omdat het zo vanzelfsprekend was dit te schrijven. En eigenlijk is vriendschap en hulp is niet altijd vanzelfsprekend. Daarom is dit ook wat mij betreft het belangrijkste hoofdstuk. Een flink aantal mensen hebben een bijdrage gehad aan het schrijven van dit proefschrift. Ik wil hen dan ook op gepaste wijze bedanken want dat hebben zij verdiend.

Allereerst natuurlijk Hans Komen, mijn begeleider. Al flink wat jaar werken wij samen en we zijn na alle algemeen bekende student - begeleider perikelen inmiddels goed op elkaar ingespeeld en verworden tot een team. Hans, ik ben je bovenal ontzettend dankbaar voor alle goede raad en natuurlijk alle goede, stevige, kritische en vooral geanimeerde gesprekken. Ik vond en vind het een genot met jou te werken.
Natuurlijk gaat een groot woord van dank uit naar Andries Kamstra. We waren elkaar al eens eerder tegengekomen in het kader van een van mijn afstudeervakken en je was direct enthousiast toen ik je vroeg of je interesse had in een samenwerking met ABG en mij in de vorm van een PhD project. Sindsdien is er ongelofelijk veel ontwikkeld, gebeurd en gedaan bij Solea BV. Ik ben je dankbaar dat ik mee mocht kijken en mijn steentje bij kon dragen. Dit is voor mij een echte verrijking geweest.
Natuurlijk ook een woord voor Johan van Arendonk, begeleider en promotor. Jouw kritische en verhelderende commentaar en tips hebben ontzettend geholpen en waren onmisbaar. Ik wil je bedanken voor de mogelijkheid en de vrijheid die mij gegeven is voor het uitvoeren van dit onderzoek.

Dan wil ik ook graag de mensen van Solea B.V. bedanken: Allereerst Ewout Blom voor de leuke tijd in IJmuiden, je tijd, hulp, de goede gesprekken en het samenwerken. Natuurlijk ook een woord van dank aan Jan van der Heul, want wát zonder Jan? Al je hulp rondom experimenten en meedenken was onmisbaar. De heren Ton en Daan allereerst voor jullie hulp bij het grote experiment maar natuurlijk ook voor alle geweldige koffiepauzes over “vroeger” en vis. Corina en Leroy bedankt voor de hulp en gezelligheid!
Uiteraard dank aan al de mensen van ABG met in het bijzonder Richard, Bert, Sylvia en Tineke voor jullie eindeloze geduld met mijn labwerk. Door jullie werd mijn tijd in het lab super...helaas kan de centrifuge dat dan niet navertellen... Piter Bijma voor je advies

en Bart Ducro voor de hulp met ASReml en Fortran maar vooral de gezellige buurman – buurman gesprekken. Albart “R” Coster, need I say more? Natuurlijk ook voor je gesprekken! Ilse van Grevenhof voor de lol in onze gezamenlijke fotohobby. De omslag van dit boekje is ook daar een product van. Off course, my students Panya Sae-Lim, Brian Babigumira, Amabel Tenghe, Celine Hoon and Chontida Phuthaworn, you guys helped me a lot and it was great fun to help you! Zonder Henry Ford, Bløf, Chiel, Coen en Sander, Bas van Werven en Kanis & Gunnink was dit proefschrift niet to stand gekomen. You made my day!
Mijn goede vrienden Sander Huurman en Job Munten wil ik bedanken omdat zij voor mij in zo’n apepakkie op een podium willen gaan zitten...mannen, het staat jullie goed!

Ik dank mijn ouders voor de stimulans om altijd een stapje verder te gaan.

Maar boven alles en iedereen: Lieve Gerdine, jij zei ja, en dát gaan we vieren. Met jou is het leven een feest en dat zal altijd zo blijven!

DANK JULLIE ALLEMAAL!

ROBBERT

Curriculum Vitae

Robbert Johannes Willem Blonk werd geboren op 15 juli 1981 te Rotterdam en groeide op in Capelle aan den IJssel waar hij in 1999 aan het Comenius College het Gymnasiumdiploma behaalde. Datzelfde jaar begon hij de studie Zoötechniek aan Wageningen Universiteit met als specialisatie Aquaculture and Fisheries. Een stage van zes maanden werd uitgevoerd bij Selonda S.A., Griekenland. In de doctoraalfase werden drie afstudeervakken voltooid. Het eerste afstudeervak werd uitgevoerd bij Zon Aquafarming B.V. (NL) onder begeleiding van de leerstoelgroep Aquaculture and Fisheries en betrof analyse van reproductiekenmerken in vier verschillende lijnen Nijl-tilapia. Het tweede afstudeervak behandelde de spermatogenese van Noordzeetong bij wilde en gekweekte dieren. Dit onderzoek werd uitgevoerd voor Solea B.V. (NL) onder begeleiding van de leerstoelgroep Aquaculture and Fisheries te Wageningen universiteit en de leerstoelgroep Endocrinologie aan de universiteit van Utrecht. Het derde afstudeervak werd uitgevoerd bij de leerstoelgroep Animal Breeding and Genetics te Wageningen universiteit. In dit project werd het gebruik van YY mannetjes in een fokprogramma voor genetische “all male” populaties van Nijl-tilapia besproken. Voor dit onderzoek ontving Robbert de tweede NZV scriptieprijs 2005 (Nederlandse Zoötechnische Vereniging). Robbert studeerde af in 2005 en eind dit jaar ontving hij een persoonlijke subsidie (NWO Casimir) voor een aanstelling als promovendus bij de leerstoelgroep Animal Breeding and Genetics aan Wageningen universiteit. Het resultaat van dit traject is dit proefschrift. Momenteel is Robbert werkzaam als post-doc bij de leerstoelgroep Animal Breeding and Genetics te Wageningen universiteit.

Robbert Johannes Willem Blonk was born on the 15th of July 1981 in Rotterdam, the Netherlands. He was raised in Capelle aan den IJssel and in 1999 he obtained his high school diploma at the Comenius College Gymnasium. In 1999, Robbert started a bachelor Animal sciences and in 2002 he started the MSc Aquaculture and Fisheries at Wageningen University. An internship of six months was performed at Selonda S.A., Greece. For his master degree, three theses were carried out. The first thesis was carried out at Zon Aquafarming B.V. (NL) under supervision of the Aquaculture and Fisheries group and analysed reproduction of different strains of Nile tilapia. The second thesis concerned spermatogenesis of wild and cultured common sole and was conducted for Solea B.V. (NL) under supervision of the Aquaculture and Fisheries group at Wageningen university and the endocrinology group at Utrecht university. The third thesis was conducted in 2005 at the Animal Breeding and Genetics group at Wageningen university. In this study, the use of YY males in breeding programs for production of all male populations of Nile tilapia was analysed. For this work, he received the second NZV thesis award (Dutch Zootechical Society). Robbert graduated in 2005 and in the end of 2005 he received a personal grant (NWO Casimir) for a PhD studentship at the Animal Breeding and Genetics group at Wageningen university. The result of this project is this thesis. Currently, Robbert is working as post-doc at the Animal Breeding and Genetics group at Wageningen university.

Publications

Peer reviewed articles

Blonk, R. J. W., Komen, J., Kamstra, A., Crooijmans, R. P. M. A. and 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.

Blonk, R. J. W., Komen, H., Kamstra, A. and Van Arendonk, J. A. M., 2010. Estimating breeding values with molecular relatedness and reconstructed pedigrees in natural mating populations of common sole, *Solea solea*. Genetics 184, 1-7.

Blonk, R. J. W., Komen, H., Kamstra, A. and Van Arendonk, J. A. M., 2010. Effects of grading on heritability estimates under commercial conditions: A case study with common sole, *Solea solea*. Aquaculture 300, 43-49.

Blonk, R. J. W., Komen, H., Tenghe, A., Kamstra, A. and Van Arendonk, J. A. M., 2010. Heritability of shape in common sole, *Solea solea*, estimated from image analysis data. Aquaculture 307, 6-11.

Conference proceedings

Blonk R.J.W., Eding, E., Schülz, R.W. and Kamstra, A., 2006. Spermatogenesis of sole, *Solea solea* in natural and culture conditions. In Cultivation of Soles III, Cadiz, Spain on March 22-23 2006.

Blonk, R.J.W. & Komen, H., 2008. Consequences of natural mating for genetic variation and inbreeding in captive broodstock populations of common sole, *Solea solea*. In European Aquaculture Society conference, Krakow, Poland on September 15-18 2008.

Blonk R.J.W., Komen J., Kamstra A. and Van Arendonk J.A.M., 2008. Genetic improvement of common sole, *Solea solea*. In Cultivation of Soles IV, Faro, Portugal on November 11-14 2008.

Blonk, R.J.W., Komen, J., Kamstra, A. and van Arendonk, J.A.M., 2009. Estimating heritability and breeding values using molecular relatedness in common sole, *Solea solea*. In International Symposium of Aquaculture Genetics 2009, Bangkok, Thailand.

Blonk R.J.W., Komen J., and van Arendonk J.A.M., 2010. Minimizing Genotyping In Breeding Programs With Natural Mating. In 9th World Congress on Genetics Applied to Livestock Production, Leipzig, Germany.

Student Theses

Blonk, R.J.W., 2003. Reproduction parameters of four different Nile Tilapia (*Oreochromis niloticus*) strains. MSc thesis Aquaculture and Fisheries group, Wageningen University.

Blonk, R.J.W., 2005. Spermatogenesis of Sole, *S. solea* (L.), in natural and culture conditions. MSc thesis Aquaculture and Fisheries group, Wageningen University.

Blonk R.J.W., 2005. Integration of YY male tilapia in breeding schemes for production of Genetically Male Tilapia (GMT). MSc thesis Animal Breeding and Genomics group, Wageningen University.

Training and Supervision Plan

Training and Supervision Plan

The basic package	
Course on philosophy of science and/or ethics	2006
WIAS Introduction Course	2007
Scientific exposure	
Congresses and seminars	
Sole workshop III, Cadiz, Spain	2006
Sole workshop IV, Faro, Portugal	2008
Aquaculture Europe congress 2008, Krakow, Poland	2008
Fokkerij en Genetica connectiondagen, Vught, the Netherlands	2008
Aquaculture genetics 2009, Bangkok, Thailand	2009
Wias science day, Wageningen, the Netherlands	2006-2009
World Congress on Genetics Applied to Livestock Production, Leipzig, Germany	2010
In-depth studies	
Courses	
Modern statistics of Life Sciences	2006
Linear models in animal breeding	2007
Understanding GxE Interactions	2007
Experimental designs	2009
ASReml course	2009
Quantitative Genetics Discussion Group	2006-2010
Statutory courses	
Use of Laboratory Animals	2006
Professional skills support courses	
Project- and Time Management	2006
Course Supervising MSc thesis	2006
Techniques for Writing and Presenting a Scientific Paper	2008
Introduction to R for statistical analysis	2008
Research skills training	
Preparing PhD research proposal	2006
Didactic skills training	
Supervising practical Genetic improvement of livestock	2008
Supervising 4 MSc theses	2007-2010
Supervising 1 BSc theses	2008

The research described in this thesis was financially supported by the Netherlands organisation for scientific research (NWO).

Artwork on cover: x-pressions.nl by Robbert Blonk and Ilse van Grevenhof

Printed by: Wöhrmann Print Service, Zutphen, the Netherlands