Comparison of Selection Methods for Natural Mating Mass-Spawning Species

Cécile Massault[1,2], Robbert Blonk[2], Hans, Komen[2], Dirk Jan de Koning[1], Chris Haley[3], van Arendonk Johan[2], Henk Bovenhuis[2]

Introduction

For a number of mass-spawning species in aquaculture, especially those recently domesticated, natural mating is still used for the production of the next generation, where a number of males and females are placed together in a tank and results in a large number of offspring with unknown parents. The offspring generation is a mix of full-sib families and half-sib families (both paternal and maternal) thus creating a complex structure with few large families and many small families. Variation in family size may increase the rate of inbreeding per generation (Falconer, D. and Mackay, T. (1996)).

Three main methods are used for selective breeding: mass selection is currently the most-used method of selection in mass-spawning species (Vandeputte, M., Dupont-Nivet, M., Haffray, P., et al. (2009)), where only records of animals' own performance are used as criteria of selection, best unbiased linear prediction (BLUP), which uses information on relatives to increase the accuracy of selection and genomic selection (Meuwisen et al. 2001) which includes genomic information to increase further accuracy. Optimum contribution with constrained rate of inbreeding is used to limit inbreeding (Sonesson, A., and Meuwissen, T. (2000)). However, inbreeding has been rarely evaluated in mass-spawning and no study on restricting inbreeding for such population structure exists at the moment.

In this study, we use simulation to determine the rate of genetic gain and rate of inbreeding in a mass spawning population with skewed parental distribution. We compared consequences of three breeding schemes, i.e. mass selection, BLUP selection and genomic selection, during 10 generations of selection.

Material and methods

We simulated a genome size of 10 Morgans, divided into 10 chromosomes (1 Morgan length each) under a finite locus genetic model. Mutation-drift equilibrium for a given population is reached when the loss of genetic variance due to drift equalled the gain of genetic variance due to mutation (Meuwissen, T., Hayes, B. And Goddard, M., (2001)). We simulated 1 700 generations with random mating, starting with 10 000 polymorphic markers (4 400 polymorphic markers at drift-mutation equilibrium) with an effective population size of 100 (50 males and 50 females) and a mutation rate of 10^{-5}. We kept record of the pedigree for the last 5 generations of random mating in order to account for existing inbreeding prior selection. The base population was the last generation of random mating after reaching mutation-drift equilibrium. The base population G_1 was identical for three selection methods: mass selection, BLUP selection and genomic selection (detailed below). The next generation was either produced from natural mating or controlled mating, depending on the scenario. Pedigrees were stored for each method of selection to calculate inbreeding level, although pedigree records were only used for BLUP selection. We simulated 3 different scenarios with 512 selection candidates (with no missing phenotypes or genotypes) and for the three different heritabilities h^2 (0.2, 05 and 0.8) over 10 generations of selection. Scenario 1 uses natural mating with selection of 20 males and 20 females, scenario 2 uses controlled mating with selection 16 males and 32 females and scenario 3 uses natural mating selecting 50 males and 50 females. Each scenario was simulated 50 times.

After each generation of selection, selected fish reproduced in two different ways: natural mating or controlled mating. During controlled mating, 16 males were mated to two females (32 in total). Females were mated to only

[1] Division of Genetics and Genomics, Roslin Institute and Royal (Dick) School of Veterinary Sciences, University of Edinburgh, Roslin, Midlothian, EH25 9PS, United Kingdom

[2] Animal Breeding and Genetics Group, Wageningen University, postbox 338, NL-6700AH, Wageningen, The Netherlands

[3] MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU

one male. Sixteen offspring were kept per full-sib family, 8 males and 8 females and the total of offspring was therefore 512. Natural mating was simulated by assigning parental contribution, that were drawn from a gamma distribution of parameter $\alpha = 1$ and $\beta = 0.75$ using 20 males and 20 females to reproduce the sea bream population structure observed by Antonello et al. (2009)

We decided to include 100 known QTLs in the genome (10 QTLs per chromosome) at generation G_1 (base population). The QTL effects were sample from a gamma distribution of shape parameter 0.4 (Meuwissen, T., Hayes, B. And Goddard, M., (2001)) and scale parameter 0.12 in order to obtain a genetic variance σ^2_A of 15. QTL effects had a probability 0.5 to be either positive or negative. The genetic value is the sum of QTL values per locus:

$$g_i = \sum_{j=1}^{n} (p_{ij} + q_{ij})$$

where g_i is the genetic value of animal i, n the total number of QTLs, p_{ij} the paternal QTL allele value of animal i and at locus j and q_{ij} the maternal QTL allele value of animal i and at locus j (Hayes and Goddard 2009). With a heritability h^2, the environmental values e_i were drawn from a normal distribution $N(0,(\sigma^2_A/h^2 - \sigma^2_A))$. The phenotypic value for animal i was the sum of genetic value g_i and environmental value e_i.

Parents for the next generation are selected using three different method of selection: mass selection, GBLUP and BLUP. No restriction on inbreeding was applied for any of the methods. Among the selection candidates, sex ratio was equal. We will refer the number of selected males as *nselmales* and the number of selected females as *nselfemales*. For all three methods, all animals were phenotyped and for genomic selection, all animals were genotyped with certainty. No constrain on inbreeding was applied.

For mass selection, the parents for the future generations were selected according to their own performance. Selection candidates were ranked from the highest phenotypes to the lowest. n*selmales* and *nselfemales* with the highest phenotypes were therefore kept for reproduction.

Estimated breeding values for the selection candidates using genomic selection (GBLUP) and BLUP were calculated using the following model:

$$y = \mu + Zu + e$$

where y was the vector of phenotypes, μ the mean phenotype, Z the incidence matrix relating phenotypes to breeding values, u vector of estimated breeding values, and e vector of environmental effects, with var(e) = I σ^2_e. with I the identity matrix. In our case, no fixed effects other than the mean were added to the model. The new brood stock was therefore composed of the *nselmales* males and *nselfemales* females with the highest EBVs. In BLUP selection, var(u) equals A σ^2_A, with A the genetic relationship matrix, while var(u) equals G σ^2_A, with G the genomic relationship matrix, calculated as described by Hayes & Goddard (2008), by calculated a similarity value between two individuals at each locus, which can be either 0 (no common alleles), 0.5 (one common allele) or 1 (2 common alleles). The genomic relationship value between 2 individuals is the sum of loci similarity values. The new brood stock was therefore composed of *nselmales* males and *nselfemales* females with the highest GEBVs.

Results and discussion

Figure 1 a) shows the genetic level with natural mating, 40 animals selected and low heritability (0.2). After 10 generations of selection, GBLUP method performs the best in term of genetic level, while we obtained the worst results with mass selection. Genetic level obtained using BLUP method falls in between the one from GBLUP and the one from mass selection. The difference of genetic level between GBLUP and BLUP is more important in the long term (G_6 to G_{11}). Rate of inbreeding (ΔF) are very high in general (up to 14% increase for one generation with BLUP). For the three methods of selection, inbreeding rate increases a lot between G_2 and G_3 (first selection). On average over 10 generations, ΔF was 3.7%, 6.7% and 9.4 % for mass, GBLUP and BLUP selection respectively. Average of accuracy (r_{gg}) of selection over 10 generations was 0.34, 0.43 and 0.54 for

mass, BLUP and GBLUP selection respectively. Therefore, GBLUP selection performs better with the highest accuracy, but inbreeding rate remains large compared to mass selection. Genetic level is similar for GBLUP, BLUP and mass selection in the long term (Figure 1 b) with high heritability. Inbreeding rate decreases as the heritability increases for BLUP and GBLUP and remains the same for mass selection (ΔF of 7.2%, 5% and 3.8% respectively). Accuracy increases as heritability increases for the three methods (r_{gg} of 0.64, 0.69 and 0.77 for mass, BLUP and GBLUP respectively for $h^2 = 0.8$).

Figure 1 c) represents genetic level of GBLUP selection respectively, between a natural mating population with 40 selected animals, a controlled mating population with 48 selected animals and a natural mating population with 100 animals selected. No major difference are observed between genetic level when selecting 40 animals with natural mating and 48 animals with controlled mating are very similar, while genetic level obtained with 100 animals selected using natural mating is not as high as the two other scenarios. In the specific conditions of this simulation, controlled mating is not an advantage over natural mating. r_{gg} is similar for the three cases (between 0.52 and 0.54). ΔF is the highest when using natural mating with 40 selected animals (6.7%) and the lowest when using natural mating with 100 selected animals (3%). ΔF of the controlled mating schemes (4.7%) is lower than natural mating with 40 animals selected, due to the population structure and the slight higher number of selected animals. While the use of genomic relationship matrix provides more accurate information to select the best animals, it acts as well as a natural constrain against inbreeding, when compared to BLUP selection. Genomic relationship matrix based on markers is more accurate than genetic relationship matrix based on pedigree and become more accurate by increasing the number of markers (VanRaden, P. (2007)).

As the results showed, rate of inbreeding with GBLUP selection is very high (6.7% in scenario 1 with h^2 of 0.2), which is similar to Sonesson, A. and Meuwissen, T. (2009). However, inbreeding rate using BLUP selection method without restriction on inbreeding was larger (9.4 %). The high rate of inbreeding observed can be explained by several factors. Inbreeding was recorded and taking into account from the last 5 generations before base generation, therefore additional inbreeding was accounted prior selection. The second factor is the actual effect of using variable family size in the population. Brood stocks were kept fairly small in our simulation, because in natural mating of mass-spawning, not all parents contribute to the offspring. This will not be the case for species, where reproduction can be controlled. The low number of candidates although can have an impact on inbreeding, because few families will be represented in the sample, due to uneven parental contributions. Finally, we did not apply any restriction on inbreeding to assess the actual increase of inbreeding using natural mating in mass-spawning. It confirms the reported high inbreeding rate after one generation of 3-6% (Fesshaye, Y., El-Bialy, Z, Rezk, M. *et al.* (2006)). Those characteristics are specific from mass-spawning species, where natural mating is still used to produce the next generation. Our results show that, while the effect on gain of genetic level is rather minor compare to breeding programs with control over reproduction, it is essential to restrain inbreeding and genetic gain of different breeding schemes should be compared at equal inbreeding rate.

We did not use a training data set to perform GBLUP selection method. With the low number of selection candidates and the uneven family size using natural mating, it was not possible to have a large number of families with a large number of offspring in each family. Therefore all animals were selection candidates and all animals were phenotyped. The uneven parental contribution to offspring is a limitation in genomic selection for natural mating mass-spawning species. Sonesson, A. and Meuwissen, T. (2009) use 100 full-sib families of at least 60 individuals (30 as candidate selection and 30 as test) – i.e. controlled mating. However, with the sea bream reference population of 1257 individuals, only 3 families have the required amount of progeny. A natural mating population of roughly 40 000 individuals will meet the criteria of 100 full-sib families of at least 60 offspring each, which is a very large sample size for parentage assignment. Although, genetic gain is similar in controlled mating and natural mating, accuracy of selection is higher with controlled mating. Effort in controlling reproduction in mass-spawning species (undergoing in sea bass, Vandeputte, M., Dupont-Nivet, M., Haffray, P., et al. (2009)) will increase the efficiency of genomic selection for those species.

Conclusion

We demonstrated that GBLUP is an efficient method of selection for natural mating mass-spawning species compared to BLUP and mass selection. However high inbreeding rate is a limitation and need to be constrained.

References

Antonello J., Massault C., Franch R. (2009) *Aquaculture* 298, 29-35

Falconer D.S., Mackay T.F.C (1996) Longman. New York. 464

pp. Fessehaye Y., El-Bialy Z., Rezk M.A. *Aquaculture* 256, 148-

158

Hayes B.J, Goddard M.E. (2008) *J. Anim. Science* 86, 2089-2092

Meuwissen T.H.E., Hayes B.J., Goddard M.E. (2001) *Genetics* 157, 1819-1829

Sonesson A.K., Meuwissen, T.H.E. (2000) *Genet. Sel. Evol* 32, 231-248

Sonesson A.K., Meuwissen T.H.E (2009*) Genet. Sel. Evol.* 41, 37-46

Vandeputte, M., Dupont-Nivet, M., Haffray, P., *et al.* (2009) *Aquaculture* 286, 20-27

VanRaden, P. M. (2007) *Interbull Bull* 37, 33-36.

Figure 1 - Comparison of genetic level: a) between GBLUP, BLUP and mass with natural mating, h2 = 0.2 and 40 selected animals, b) between three heritability (0.2, 0.5 and 0.8) of GBLUP with natural mating, 40 animals selected, c) between GBLUP with natural mating (40 animals selecetd) , controlled mating (48 animals selected) and natural mating (100 animals selected)