A NEW ERA: THE MERGING OF QUANTITATIVE AND MOLECULAR GENETICS – PROSPECTS FOR TILAPIA BREEDING PROGRAMS

Hans Magnus Gjøen

Department of Animal and Aquacultural Sciences, Agricultural University of Norway, P.O.Box 5025, 1432 Aas, Norway

Abstract

Traditionally, selective breeding has been the method of choice for improving broodstock, both in terrestrial and aquatic animal production. Considerable improvements have been achieved for most species when sound base populations have been established and appropriate selection methods were applied. Results are shown for a major breeding program in tilapia. This approach, however, is based more or less on a theoretical "black box" in terms of understanding the concrete bridge between single genes and phenotypic expression. In parallel, there is an ongoing revolution within experimental biology, and enormous amounts of data are produced. The sequencing of whole genomes has given new possibilities, and e.g. micro-array technology makes it possible to measure genome wide expression levels simultaneously. In addition, genetic markers are widely developed and used. Marker-data in itself can be analysed with the traditional models to find so called QTLs. However, the traditional model of gene action, which is based on additivity, is not well suited for handling data on the levels of mRNA or proteins, as these are expressions of dynamic system of interacting genes, rather than independent effects. Thus, new gene-regulatory models have to be developed.

Traditional selective breeding will remain a "main engine" in tilapia breeding program in the foreseeable future, but the enormous amount of new experimental data and new gene technological tools requires a whole new modelling concept for further refinement of our genetic improvement efforts. This will also include specific physiological, biological and hands-on knowledge and multidisciplinary research networks. Tilapia is likely to be a forerunner for other aquaculture species in this respect.

Introduction

Status of breeding technology: Traditional selection works

It is well documented that conventional selective breeding works well for all species tested. If inbreeding is avoided and the population has sufficient genetic variation, a proper selection scheme will normally create around 10% genetic gain per generation for a trait like growth rate. Selection response is accumulated each generation, so considerable accumulated response can be generated if we work patiently.

The main principle in traditional quantitative genetic theory is the infinitesimal genetic model and the concept of additivity of genetic effects. The infinitesimal genetic model assumes that there are infinite number of genes behind a quantitative trait, whereas the principle of additivity assumes that all genetic effects, also dominance and epitasis, can be modelled independent of each other. We will come back to the shortcomings of these assumptions later, but fact is that despite of its simplicity, these models have proven to be very powerful and useful when describing, simulating and planning any breeding program.

When simulating breeding programs, it is customary to generate about 15 generations in order to clearly display the long-term effect of selection. The most important parameter to monitor carefully is the rate of inbreeding, which should be kept at an acceptable level. A rule of thumb that is often applied and sited is to at least keep the effective population size (N_e) above 50, which corresponds to a rate of inbreeding of 1% per generation. By using these tools, it is possible to tailor make breeding schemes to fit biology, resources and facilities at hand.

The final goal for any breeding program would be to maximize the genetic gain for the trait of interest. Genetic gain is generally expressed by:

$$\Delta G = i\, \sigma_G\, r_{GI}$$

where:

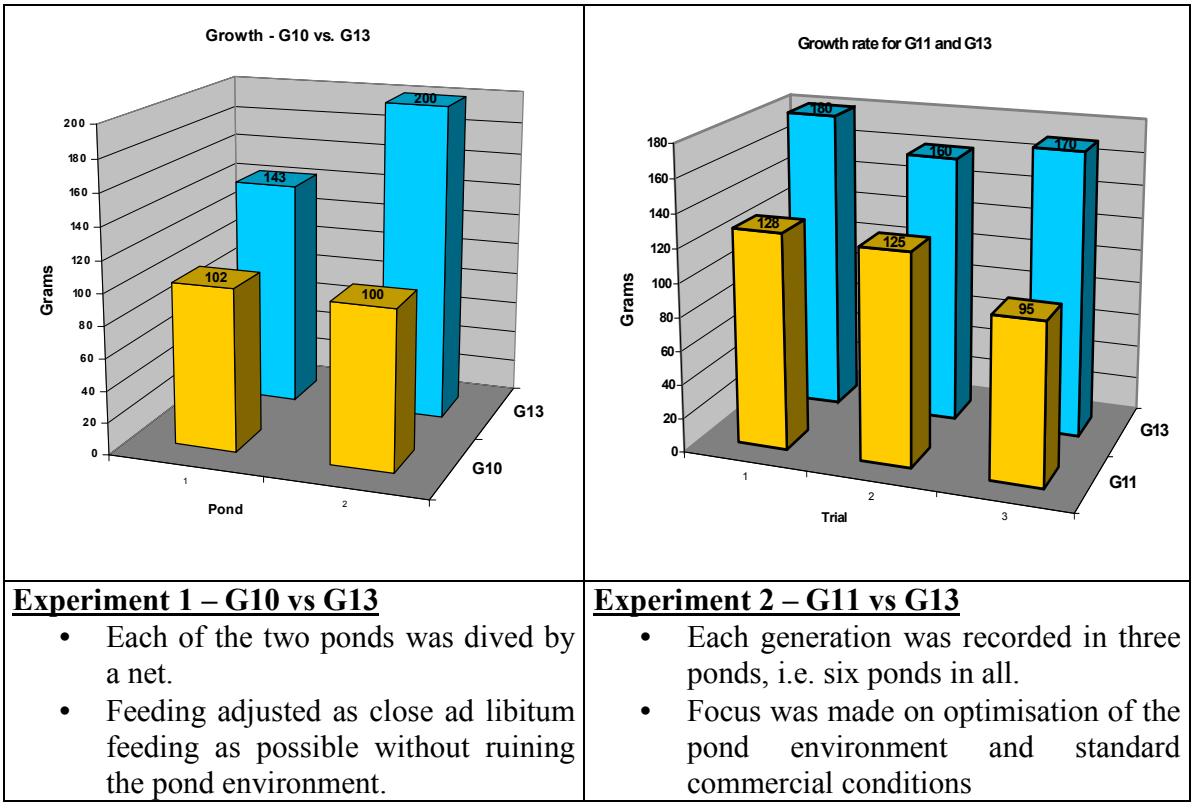
- i is the selection intensity
- σ_G is the true genetic variation
- r_{GI} is the correlation between the true genetic value and the estimated breeding value, often called the accuracy of the breeding value estimate.

The difficult task of any breeder is to optimise these parameters, since none of them can be easily enlarged without negatively influencing one or both of the others. This can be adequately done through simulation studies as described above, and this will be an important tool for breeders in the future too.

An example from tilapia

An example of a breeding program that has tried to apply the principles described above is the breeding program of GenoMar. The GST^{TM} strain that they have developed is the continuation of the GIFT breeding program run over a 10-year period, ending in 1998. The GIFT Project was concluded after 10 generations of testing and selection. Since then, GenoMar has done some changes in the program based on scientific consideration of new technologies and results from extensive simulation studies. The two major changes made were to apply DNA-fingerprinting as an ID-tool and to change to a revolving mating scheme, which means that a generation is completed after nine monthly batches. DNA-typing gave several advantages in form of increased selection intensity, shorter generation interval and operational benefits. The strain has now been selected for 4-5 more generations by these means.

In order to estimate actual genetic gain in the breeding program, comparison trials have been conducted. Generation 10 and 11 (G10/G11) as a base, were tested against generation 13 (G13). The results are shown in Figure 1, and they demonstrate that selection has been successful with an average genetic gain of almost 20% per generation. This is somewhat higher than expected, but clearly indicates that the applied breeding scheme has been adequate. Even though the selection intensity has been high, the change in inbreeding rate per generation in these generations has been less than 0.3% per generation, which is well within recommended limits. The overall survival rate also increased with in average 11% per generation, giving a total average survival of more than 80% in G13. The overall FCR was satisfactory in all ponds, with a total average less than 1.1, except for one G10 pond that experienced an uncontrolled algae bloom, and thus was excluded from the data. In experiment 1, G10 and G13 had the same FCR, whereas in experiment 2, FCR was reduced with 6% per generation.



Experiment 1 – G10 vs G13	**Experiment 2 – G11 vs G13**
• Each of the two ponds was dived by a net. • Feeding adjusted as close ad libitum feeding as possible without ruining the pond environment.	• Each generation was recorded in three ponds, i.e. six ponds in all. • Focus was made on optimisation of the pond environment and standard commercial conditions

Figure 1. Growth performance in two comparison trials of GST^{TM}.

Immediate potentials for new gene technology

High expectations have been heralded concerning the results that new gene technology can achieve for the benefit of animal and fish production. This has caused public and private bodies to spend millions of dollars on finding vital genes for economical important traits. Sad to say, many of these projects have given few results, and this might

lead some to the conclusion that gene technology is still a futuristic adventure. However, there are some immediate outputs of the technology that should be harvested. These may be categorised as 1) improvement of traditional systems and 2) surpass of "difficult" traits.

Improvement of traditional methods

An efficient breeding program for fish requires the possibility of family identification. The main drawbacks with conventional breeding schemes and tagging methods are the introduction of common environmental effects prior to tagging, high costs related to establishing and operating facilities for separate family rearing and limited number of individuals tagged within each family.

These drawbacks can all be solved by using DNA-typing as a tagging system. This allows for:

- Communal rearing of individuals from many families in the very early life stage
- Minimal need for special facilities due to separate family rearing
- Utilisation of the enormous display of Mendelian sampling that is represented within the large full-sib groups produced by most fish species

Towards the end of the growth period, the required number of the largest fish are DNA-typed and marked with a physical tag. In addition, other traits can be efficiently recorded by using smaller parallel lines with all families represented in fewer and standardised numbers. Here all fish are recorded and typed. With appropriate software, pedigrees can be established, and the breeder is enabled to estimate high accuracy breeding values through use of BLUP.

Simulation studies and preliminary growth results from tilapia has shown that this technology can improve the efficiency of selective breeding by at least 10-20% compared to conventional schemes.

Surpass "difficult traits"

Many traits important for economical fish farming, like disease resistance, fillet colour and fillet quality, are not measurable at the live individual and thus not obtainable for the breeding candidate itself. For such traits, the breeder is dependent on other sources of information, the full-sib records being the most important. Even though this is a valuable breeding method, unique to fish, it only allows us to utilise half of the genetic variation present for these trait. Individual records would allow full benefit of this variation. One way to obtain observation associated with an individual is to obtain QTL-information. This is why emphasis in QTL-scans has been put on these traits and is more likely to give important contributions to multiple trait breeding programs. Most of these traits have also never been subject to selection before, and selection based on QTL-information is thus more likely to be more efficient.

A new era - opening the black box between phenotype and genotype

There is an ongoing revolution within experimental biology as far as enormous amounts of data are produced from various new technologies. The sequencing of whole genomes has given new possibilities for collection of system wide biological data. Genes

can be identified within a sequence and micro array technology makes it possible to measure genome wide expression levels simultaneously. Furthermore, differences at single nucleotides between individual sequences, SNPs, are a new class of genetic markers which gives us even more powerful ways to fine-map the genomes of our farmed species. At present, nearly 1.8 million SNPs have been detected in the human genome. Analysing the amounts of information that can be produced from such experiments represents a huge challenge for modern biology. SNP data in itself can be analysed with the methods of genetics. As genetic markers, the SNPs can be arranged linearly in linkage groups, and recombination fractions can be estimated to make dense genetic maps. These genetic maps can then be used in traditional genetic studies, for example in a search for QTLs.

However, the traditional model of gene action, which is based on additivity, is not well suited for handling data on the levels of mRNA or proteins, as these are expressions of dynamic system of interacting genes rather than independent effects. Several studies show that the effectiveness of marker assisted selection (MAS) is not as efficient as previously thought (e.g. 4), and that there is a need for new models that can enable us to utilise this vast amount of data in a better and sustainable way. New models are also needed to mimic the true nature of genetic interaction effects, like dominance and epistasis, where conventional additive models have shown to be inadequate.

New technology and disciplines often require new models and methods to communicate the new concept in a proper way. This new area is now being developed in the union of at least three major sciences: genomics, quantitative genetics and physiology/biology (Figure 2).

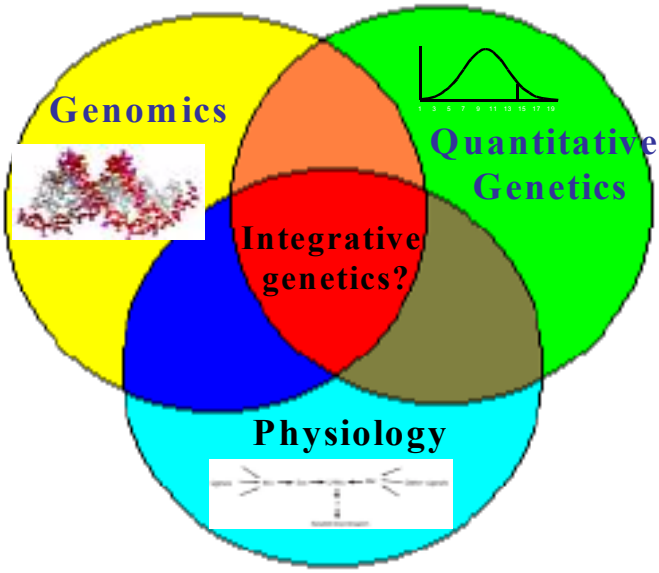


Figure 2. New technology and knowledge allows the merging or linking of sciences that previously did not interact efficiently.

Many terms have been used in relation to this new area (e.g. systems biology, bioinformatics or computational biology), but the term *Integrative genetics* is trying to encompass the fact that we now are on the edge of integrating all biological observations and systems down to the genetic building blocks. We are about to bridge the gap between genotypes and phenotypes by integrating experimental and theoretical approaches; integrating processes and mechanisms connecting genotypic data with phenotypic data in a coherent mechanistic explanatory structure, and by applying the explanatory frameworks of non-linear system dynamics and statistics. This will typically include the use of positive and negative feedback-loops to describe the way that the expression of genes interact, in contrast to simple linear manners in conventional models. The new approach has shown to be very efficient also in describing genetic effects like dominance and epistasis.

One other important aspect of this approach is that it will allow us in a new way to link phenotype (P), genotype (G) and environment (E) in a much more realistic way. Conventionally, we are used to limit our models to treat these factors in a simple additive way P = G + E. But genetic and environmental aspects has to be considered jointly, not only as an interaction element in the conventional equation, e.g. P = G + E + GxE, but as a dynamic system allowing for the non-linearity described above.

Describing such systems completely is likely to be extremely complicated for composite traits, like growth rate. However, by using adequate methodology, it is doable; especially for more explicit traits, like colouration. For instance, recently a larger project to accurately describe and model the filet colouration in salmon has thus been initiated in Norway at the Centre for Integrative Genetics (CIGENE). Generally, the new tools require and enable multi-disciplinary approaches that will lead to an increased need for networking among scientists.

Conclusion

Traditional selective breeding will remain a ”main engine” in breeding program in the foreseeable future. Some immediate benefits of using new gene technology should be harvested.

The enormous amount of new experimental data and new gene technological tools requires a completely new modelling concept. Utilisation of this new gene-toolbox gives great prospective for integrating increasingly larger amount of data and insight, both from biology, physiology and genetics to the benefit of a growing aquaculture industry. Since tilapia is the fastest growing aquaculture species, has a short generation interval and has many research projects allocated, it is believed that it will be leading the development of modern aquaculture breeding technology.

References

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. Livest. Prod. Sci. 47: 59-72.

Gjøen, H. M. and Gjerde, B. 1998. Comparing breeding schemes using individual phenotypic values and BLUP breeding values as selection criteria. Proc. 6th World Congress on Genetics applied to Livestock Production 27, 111-114. Armidale, NSW, Australia.

Melamed, P., Gong, Z., Fletcher, G., and Hew, C. L. 2002. The potential impact of modern biotechnology on fish aquaculture. Aquaculture 204: 255-269.

Hayes, B. and Goddard, M.E. 2003. Evaluation of marker assisted selection in pig enterprises. Livestock Production Science 81: 197-211.

Omholt, S. W., Plahte, E., Øyehaug, L., and Xiang, K. 2000. Gene regulatory networks generating the phenomena of additivity, dominance and epistasis. Genetics 155: 969-981.

Plahte, E., Mestl, T., and Omholt, S. W. 1998. A methodological basis for description and analysis of systems with complex switch-like interactions. Journal of Mathematical Biology 36: 321-349.

Omholt S.W., Kefang X., Andersen Ø., and Plahte E., 1998. Description and Analysis of Switchlike Regulatory Networks Exemplified by a Model of Cellular Iron Homeostasis. Journal of Theoretical Biology 00195: 339-351.

Øyehaug, L., Plahte, E., Våge, D.I. and Omholt, S.W. 2002. The Regulatory Basis of Melanogenic Switching. Journal of Theoretical Biology 215: 449-469.