The Open Access Israeli Journal of Aquaculture – Bamidgeh

As from **January 2010** The Israeli Journal of Aquaculture - Bamidgeh (IJA) will be published exclusively as **an on-line Open Access (OA)** quarterly accessible by all AquacultureHub (http://www.aquaculturehub.org) members and registered individuals and institutions. Please visit our website (http://siamb.org.il) for free registration form, further information and instructions.

This transformation from a subscription printed version to an on-line OA journal, aims at supporting the concept that scientific peer-reviewed publications should be made available to all, including those with limited resources. The OA IJA does not enforce author or subscription fees and will endeavor to obtain alternative sources of income to support this policy for as long as possible.

Published under auspices of
The Society of Israeli Aquaculture and Marine Biotechnology (SIAMB),
University of Hawaii at Manoa Library
and
University of Hawaii Aquaculture Program in association with
AquacultureHub
http://www.aquaculturehub.org

ISSN 0792 - 156X

PUBLISHER:
Israeli Journal of Aquaculture - BAMIGDEH -
Kibbutz Ein Hamifratz, Mobile Post 25210,
ISRAEL
Phone: + 972 52 3965809
http://siamb.org.il

Genetics and Genomics - Integration of Molecular Genetics into a Breeding Program for Rainbow Trout

Jeffrey T. Silverstein*, Gregory M. Weber, Caird E. Rexroad III and Roger L. Vallejo

USDA/ARS National Center for Cool and Cold Water Aquaculture, Leetown, West Virginia 25430, USA

Key words: rainbow trout, selective breeding, marker assisted selection

Abstract

At the National Center for Cool and Cold Water Aquaculture (US Department of Agriculture, Ag. Research Service) in Leetown, WV, we have a broodstock development program now entering the second generation of family-based selective breeding using expected breeding values (EBVs). Our major breeding objectives are faster growth and resistance to *Flavobacterium psychrophilum*, the causative agent of bacterial coldwater disease. For these traits we have developed assays to evaluate phenotypic performance. In addition to our breeding program, we are participating in an international collaboration to develop a microsatellite marker linkage map for the rainbow trout with the intent of identifying quantitative trait loci (QTL) and using them in marker (MAS) or gene-assisted (GAS) selection. There are several possible approaches with regard to the types and numbers of markers to develop and the strategies and methods for implementing the markers in a selective breeding program. This paper describes the choices we made concerning QTL identification for traits of high, low, and unknown degrees of heritability. These traits are plasma cortisol response to a crowding stress ($h^2 = 0.4$), feed intake ($h^2 = 0.1$) and resistance to *F. psychrophilum* ($0.3<h^2<0.4$). In order to identify QTL in a relevant commercially important rainbow trout line, we are making crosses from within our resource population. The development of research family crosses, choice of markers for genome scanning, and planned steps to implementation of these results are described.

Genetic improvement for agriculture has long been a goal of quantitative genetic studies and is increasingly a goal common to molecular genetic studies. For decades quantitative genetics has provided the theoretical and applied basis for genetic improvement of agricultural species. More recently, molecular genetics has brought the development of a large variety and number of genetic markers and the promise of more efficient discovery of the identity and location of genes with relatively large effects on important traits (quantitative trait loci or QTL). If this promise is upheld then progress toward developing animals with superior genetic value will be enhanced by the combined use of quantitative and molecular genetic techniques.

At the National Center for Cool and Cold Water Aquaculture (NCCCWA) we are developing improved germplasm for the rainbow trout industry in the USA. The goals of the improvement program were fashioned with considerable input from various domestic pro-

* Corresponding author. Tel.: +1-304-7248340, fax: +1-304-7250351, e-mail: Jeff.Silverstein@ARS.USDA.GOV

ducer groups and the priorities are to improve growth, feed efficiency, disease resistance and general hardiness. The improvement is to be achieved by selective breeding based on family performance, and we are initially focused on improving growth and resistance to *Flavobacterium psychrophilum*, the causative agent of bacterial cold water disease**.** Whereas breeding programs typically use economic values to weight selection intensity for various traits, because we do not use a selection index weighted by economic value our breeding program objectives should be considered to be informal. This paper briefly reviews the status and goals of the NCCCWA breeding program. The limitations to genetic improvement and the potential for QTL identification and marker or gene assisted selection (MAS or GAS) implementation to alleviate these limitations are outlined. Finally, the initial steps being taken to implement MAS at NCCCWA are described.

Quantitative genetics has formed the basis of genetic improvement in plants and animals for over seven decades. The genetic improvement success in farm animals of the 1930s stemmed from application of statistical genetic models (Fisher, 1918; Wright, 1921ab) to important production traits that generally had high heritability and were measured in both sexes, such as daily weight gain and feed conversion (Merks, 2000). Carp improvement in Israel took advantage of quantitative theory and crossbreeding experiments were conducted to exploit heterosis (Hulata et al., 1985; Moav and Wohlfarth, 1976). With the advent of animal models for variance component estimation in the 1970s and 80s (Henderson, 1984), the calculation of best linear unbiased predictors (BLUPs) became routine and have been in common use since the late 1980s (Belonsky and Kennedy, 1988; Lo et al., 1997). These evaluations of performance for individuals enabled progress in traits that were sex limited, had low heritability and required selection of individuals based on sib performance traits such as milk production, or litter size. The Norwegian Atlantic salmon breeding program took off during this period and while initially focused on growth traits with high heritability, it has subsequently included multiple traits including those with lower heritability such as age at sexual maturity and disease resistance (Gjedrem, 2000). Gains in performance were fantastic. In spite of the tremendous gains achieved, the actual genes involved in the traits being selected and their location in the genome remained unknown and outside the realm of these investigations.

The recent explosion of interest in genetics has been driven by molecular information, identifying the genes and the location of the genes controlling specific traits. For example, molecular genetics has had incredible success identifying the genes causing unique phenotypes ranging from cystic fibrosis in humans (Kerem et al., 1989) to leanness of flesh linked with halothane susceptibility in pigs (Fujii et al., 1991) and more recently a mutation in the insulin like growth factor 2 (IGF2) gene affecting muscle growth, fat deposition and heart size in pigs (Van Laere et al., 2003). These are traits which are due to mutations in single genes. While genetic variation in some important traits may be traced to mutations in single genes, molecular genetics also aims to dissect multifactorial quantitative traits of complex inheritance by identifying multiple regions of the genome with sequence variation that explains significant amounts of phenotypic variation observed for traits such as disease resistance, and quality traits such as fillet yield.

Among quantitative and molecular geneticists, perhaps there has been a reciprocal lack of appreciation of the strengths and limitations of these complementary approaches to achieving genetic improvement. Reports from molecular specialists announcing that to apply marker assisted selection effectively more molecular markers are necessary (and coming soon) to identify allele variants in strong linkage disequilibrium due to physical proximity with the trait of interest, may not fully resonate with quantitative geneticists. Quantitative geneticists wonder how to employ the markers we already have, whether more markers are needed, and when molecular approaches will fulfill their promise.

The limit to genetic improvement through selective breeding is the availability of genetic

variation in the target population. Genetic variation indicates that there is a heritable component to performance and that selection of parents with higher breeding values will lead to production of superior offspring. Theoretically, selection on a trait depletes genetic variation and eventually genetic variation can be exhausted (Falconer and Mackay, 1996). Accurate determination of genetic value is also important, and limitations in determining accuracy can limit the rate of genetic improvement (Tribout et al., 1998). Progress in selective breeding will be proportional to the accuracy in determining an animals' genetic value. More accurate determination of genetic value is the focus of marker assisted selection (Meuwissen et al., 2001).

The systematic integration of quantitative and molecular genetics is relatively new although the ideas for mapping QTL and genes have been around for decades (Weller, 2001) and examples of MAS based on biochemical polymorphisms, such as the link between hemoglobin type and reproductive performance in Merino sheep (Evans and Turner, 1965), go back over 40 years. The recent advent of nearly limitless genetic markers has dramatically changed the practicality of identifying genetic regions of interest. Further, the possibility of sequencing whole genomes, no longer limited to humans and a few model species, has opened the door to clearly identifying the sequence polymorphisms responsible for trait genetic variation (Andersson, 2001).

In terms of integrating marker information into a breeding program, Dekkers (1999) identified four necessary steps. The steps are (a) map development: identification and mapping of genes and genetic polymorphisms, (b) QTL detection: detection and estimation of associations between genetic markers and performance traits, (c) genetic evaluation: integration of phenotypic and genotypic data in statistical methods to estimate breeding values of individual animals in a breeding population, and (d) breeding strategy development: development of breeding strategies and programs for the use of molecular genetic information in selection and mating programs. At NCCCWA we are currently between the first two steps.

The NCCCWA linkage map for rainbow trout (Rexroad, unpubl. data) is based on microsatellite loci and represents a large effort with marker contributions from researchers throughout the world. There are 568 markers ordered with a cutoff of LOD score 10, the markers were distributed across 29 linkage groups with a total map distance of 2042 cM. The markers were developed with several purposes in mind, including parental verification, inbreeding monitoring and MAS. For the purpose of QTL discovery, highly polymorphic loci are important in an outbred population, thus microsatellite markers which are more polymorphic than the typically diallelic restriction fragment length polymorphism (RFLP) markers were chosen as the main marker class for this map.

To progress into the second step of the above program, we must identify associations between genetic markers or gene allelic variants and traits of interest. There are two broad approaches to identifying markers associated with specific traits, the candidate gene approach and the genome scan approach.

The candidate gene approach is based on some background knowledge of the physiology of a trait. Work on the estrogen receptor (ESR) gene in relation to litter size in pigs by Rothschild et al. (1996) is a good example of the candidate gene approach. The estrogen receptor gene was investigated because of the importance of this hormonal pathway in reproduction. A polymorphism within the gene provided a marker that could be mapped. This polymorphism was tracked in Meishan and Large White breeds of pig and two common alleles (A and B) yielding three genotypes (AA, AB and BB) were identified. Next, an association between the ESR genotype and phenotype was confirmed. The AA genotype resulted in an average of 2.3 pig increase in first parity litter size for Meishan pigs. A similar trend in increased litter size was seen in other breeds, but the effect was not as large. In general, the candidate gene approach is an educated guess based on physiological knowledge of the underlying biology of a trait.

Alternate phenotypic classes are genotyped at specific markers testing for association. The drawback is that there are a large number of potential candidate genes for testing and limited biological knowledge. There has been some success using this approach in fish. Palti et al. (1999) identified markers associated with disease resistance segregating with major histocompatibility genes in rainbow trout. As comparative mapping tools are developed, applications of findings from other species like zebrafish or pufferfish with fully sequenced genomes may improve the likelihood of success using the candidate gene approach in rainbow trout.

The genome scan approach depends less on prior information. It is a brute force method that relies on testing many markers for non-random association between genotype and phenotype, rather than on knowledge of the genes underlying phenotypic variation. In other words, it identifies chromosomal regions that are co-segregating with the trait of interest. A non-random association of alleles at a marker and trait locus is known as linkage disequilibrium (LD) and it can play a role in the search for genes affecting the trait (Weir, 1996). A number of markers spanning the genome at reasonable map densities would be tested for LD between marker genotype and trait phenotype.

The number of markers needed depends on several factors. Generally 100-150 markers are needed for complete genome coverage (Haley and Visscher, 2000). The higher the heritability of the trait the fewer the number of markers needed (or the greater the spacing allowed between markers) to identify a QTL of moderate effect, meaning that the locus explains a substantial fraction of the phenotypic variance. Another consideration is whether the marker/QTL association is in LD throughout the population of interest (LD marker) or in LD in specific families but in linkage equilibrium over the entire population (LE markers; Dekkers, 2004). LE markers require only sparse marker maps (15-30 cM spacing) whereas LD markers must be close to the causative mutation, within 5 cM.

Another question that arises when considering experimental design to identify markers linked to traits is which animals should be genotyped. If the candidate gene approach is used, then the genotypes for specific markers known to be in proximity to the candidate genes will be tracked in animals with divergent phenotypes. Each marker will be tested for a significant association between genotype and phenotype. If the genome scan approach is used, animals of divergent phenotypes can still be used but the marker loci examined are not chosen based on previous knowledge; they are selected to be distributed throughout the genome with no prior expectation of association with a particular gene.

Within the genome scan approach, there are several breeding strategies that can be used in the search for QTL: inbred lines, crossbred populations, and outbred populations. Each of these strategies has been used to some extent in QTL identification in rainbow trout. QTL detection from inbred lines is the most straightforward statistically; however, QTL variants detected by inbred line crosses typically represent differences between lines, and whether the QTL will be of general use within other populations remains unclear (Lynch and Walsh, 1998). In rainbow trout, a number of QTL have been identified using doubled haploid lines as the inbred parents (Robison et al., 2001). Use of crossbred populations has been used to great advantage in pigs. Andersson et al. (1994) made crosses between the European wild boar and Large White domestic pigs which revealed several QTL in the pig with large effects on body composition. This strategy was employed by Sakamoto et al. (1999) in identifying QTL associated with spawning time in rainbow trout by crossing spring and fall spawning strains. The outbred population approach is common in animal agriculture studies because it is often not practical to create inbred lines, and breeds with different phenotypes of interest may not exist for crossbreeding. The outbreeding approach is more complicated statistically, yet often this approach is most relevant for detecting markers associated with QTLs to be used in improvement for aquaculture because populations under

improvement are typically not inbred and populations have not generally been characterized for crossbreeding type studies (however see Sakamoto et al., 1999). QTL detection studies using outbred populations will tend to identify loci with significant effects in particular families, not loci with significant effects across the population. For example, the study on upper thermal tolerance in rainbow trout by (Perry et al., 2001) was one of the first examples of QTL detection in an outbreeding population of rainbow trout. In a study including 44 families from eleven 2 x 2 factorial crosses, one family was identified in which one locus accounted for 7.5% of the phenotypic variation. To find marker loci with significant effects across many families, increased resolution through identifying additional markers closer to the causative mutation would be required.

At NCCCWA there are three separate QTL identification projects at various stages of planning and implementation. The three traits for which QTL markers are being sought are resistance to *F. psychrophilum*, feed intake, and cortisol response to a standardized stressor. Disease resistance is a typical candidate trait for QTL markers to conduct MAS. In the absence of QTL resistance, susceptibility can only be determined by exposure, and all siblings from a full-sib family would have equal value. With markers, resistance may be evaluated through presence or absence of particular alleles and siblings with particular alleles would have a higher value based on QTL genotype. Feed intake can be measured on breeders, however it is expensive and difficult to measure on individuals. Further, whether measurements of feed intake are consistent over time has not been determined. The expense and difficulty of measuring feed intake on individuals make it a good trait for MAS. Cortisol response to stress is straightforward, but time consuming to measure in individuals. Also, the relationship between cortisol responsiveness and production traits of interest is not well understood. Through identification of QTL, fine mapping, and eventual identification of genes responsible for variation in cortisol response, associations and mechanisms underlying associations between production traits and stress physiology might be elucidated.

Before discussing the details of the QTL studies, a brief discussion of the design of our breeding program is necessary. Based on a two-year cycle of sexual maturity, there are two year class groups that form separate broodstocks, an even and an odd year class group. Each year class is a synthetic line produced by crossing fish from six different lines: University of Washington, Shasta, Troutlodge Inc, House Creek, Arlee, and Kamloops. Each year class is made up of approximately 100 full and half-sib families.

The two year class populations are being selected for different traits. The even year line is being selected for improved growth. The odd year line will be selected for resistance to bacterial cold water disease (also known as rainbow trout fry syndrome) caused by the pathogen, *F. psychrophilum*, the bacterial disease of largest impact to the rainbow trout industry. The odd year line was selected on growth characters in 2005, however, a challenge for *F. psychrophilum* was also conducted in that year and the fish which will spawn in 2007 will be selected based entirely on their resistance to *F. psychrophilum*.

Feed efficiency and fish hardiness are being monitored, though not yet being selected. Feed efficiency is being examined by measurement of feed intake and growth, however methods to measure individual feed intake over time in group reared fish are still being developed. Hardiness is being approached by measuring stress response. Following protocols developed by Pottinger and Carrick (1999; see Lankford and Weber, 2006) plasma cortisol levels in rainbow trout following a standardized crowding stressor have been quantified. It is not clear whether a robust elevation in cortisol is desirable or less responsiveness is better. At this time we are monitoring whether selection for growth and disease resistance will cause change in stress response. We have found increased cortisol response is correlated with increased growth in our lines (Lankford and Weber, 2006).

The cortisol response QTL identification study is the furthest along. Parental fish phe-

notypes were measured in 2004 and F_1 crosses were produced in January 2006. Heritability of cortisol response to a standardized handling/crowding stressor was estimated at approximately 0.4 (unpubl. data). The crosses made are an effort to combine outbreeding and inbreeding strategies. Parental fish from an outbred population with highly divergent phenotypes were crossed to produce several F_1 families. The use of highly divergent parental phenotypes (P_0) was an attempt to maximize the allelic difference in loci involved with cortisol response; this should increase the likelihood that the parents are informative and consequently the prospects of detecting QTL. A sample of F_1 animals will be phenotyped to determine their cortisol response to stress phenotype. A response intermediate to the parental phenotypes, for example, would suggest that divergent parental phenotypes had alternate codominant alleles. Animals in the subsequent F_2 generation, to be produced in 2008, will be challenged with the standard stressor and then will be genotyped at approximately 150 loci. A genome scan will follow to evaluate the association between phenotype and genotype at each informative marker.

Resistance to *F. psychrophilum* appears to have moderate heritability of between 0.3 and 0.4 estimated from the longitudinal data of days to death in our population (unpubl. data), slightly higher than estimates of heritability for resistance to *F. psychrophilum* shown by Henryon et al. (2006). In 2005, seventy-five full-sib families were challenged with *F. psychrophilum*. Breeders from the full-sibs of the families at the extremes for survival will be chosen as the parents for an F_1 generation to be crossed in 2007. The F_1 fish will be raised to sexual maturity and bred to form F_2s in 2009, which will be evaluated for disease resistance phenotype and marker genotypes at approximately 150 loci. Association between phenotypic variation and markers will be evaluated via a genome-wide linkage scan.

Feed intake which has been estimated to have a low heritability of about 0.1 (unpubl. data) is another trait being examined. Here too, parental fish identified as having highly divergent phenotypes among the 2005 families will be crossed to produce the F_1s. The parental fish are to be identified from feed intake records taken in four feeding bouts over a 3-month time frame on 200 individual fish. The F_1s will be evaluated for variance and mean performance. The F_2 crosses will be made in 2009 and the phenotypic evaluation completed in 2010. A genome scan to test for associations between marker genotype and phenotype will follow.

The hoped-for outcome for each of these studies is to identify QTL that explain enough trait variation to be useful in our breeding program through application of MAS. To identify markers closely linked to trait causal mutations that can be used via MAS will require additional fine mapping efforts. As the types and numbers of markers increases and the cost of scoring markers decreases, the barrier represented by the need for QTL fine mapping will become less daunting.

References

Andersson L., 2001. Genetic dissection of phenotypic diversity in farm animals. *Nat. Rev. Genet.,* 2:130-138.

Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M., Hansson I., Håkansson J. and K. Lundström, 1994. Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science*, 263:1771-1774.

Belonsky G.M. and B.W. Kennedy, 1988. Selection on individual phenotype and best linear unbiased predictor of breeding value in a closed swine herd. *J. Anim. Sci.,* 66:1124-1131.

Dekkers J., 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *J. Anim. Sci.,* 82:E313-328.

Evans J. and H. Turner, 1965. Haemoglobin type and reproductive performance in Australian Merino sheep. *Nature,* 207:1396-1397.

Falconer D. and T. Mackay, 1996. *Introduction to Quantitative Genetics* (4th ed.). Pearson Education Ltd., Essex, UK. 464 pp.

Fisher R.A., 1918. The correlation between

relatives on the supposition of Mendelian inheritance. *Trans. Royal Soc. Edinburgh*, 52:399-433.

Fujii J., Otsu K., Zorzato F., de Leon S., Khanna V.K., Weiler J.E., O'Brien P.J. and D.H. MacLennan, 1991. Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science*, 253:448-451.

Haley C. and P. Visscher, 2000. DNA markers and genetic testing in farm animal improvement: current applications and future prospects. *Roslin Inst. Ann. Rep.,* 1998-1999:28-39.

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

Hulata G., Wohlfarth G. and R. Moav, 1985. Genetic differences between the Chinese and European races of the common carp *Cyprinus carpio* L. IV. Effects of sexual maturation on growth patterns. *J. Fish Biol.,* 26:95-103.

Kerem B., Rommens J.M., Buchanan J.A., Markiewicz D., Cox T.K., Chakravarti A., Buchwald M. and L.C. Tsui, 1989. Identification of the cystic fibrosis gene: genetic analysis. *Science*, 245:1073-1080.

Lankford S. and G. Weber, 2006. Associations between plasma growth hormone, insulin-like growth factor-1 and cortisol with stress responsiveness and growth performance in a selective breeding program for rainbow trout. *North Am. J. Aquacult.,* 68:151-159.

Lo L.L., Fernando R.L. and M. Grossman, 1997. Genetic evaluation by BLUP in two-breed terminal crossbreeding systems under dominance. *J. Anim. Sci.,* 75:2877-2884.

Merks J.W.M., 2000. *One Century of Genetic Changes in Pigs and Future Needs*. Brit. Soc. Anim. Sci., Edinburgh.

Meuwissen T.H., Hayes B.J. and M.E. Goddard, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157:1819-1829.

Moav R. and G. Wohlfarth, 1976. Two-way selection for growth rate in the common carp (*Cyprinus carpio* L.). *Genetics*, 82:83-101.

Perry G.M.L., Danzmann R.G., Ferguson M.M. and J.P. Gibson, 2001. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). *Heredity*, 86:333-341.

Pottinger T.G. and T.R. Carrick, 1999. Modification of the plasma cortisol response to stress in rainbow trout by selective breeding. *Gen. Comp. Endocrinol.,* 116:122-132.

Robison B.D., Wheeler P.A., Sundin K., Sikka P. and G.H. Thorgaard, 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *J. Hered.,* 92:16-22.

Rothschild M., Jacobson C., Vaske D., Tuggle C., Wang L., Short T., Eckardt G., Sasaki S., Vincent A., McLaren D., Southwood O., van der Steen H., Mileham A. and G. Plastow, 1996. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc. Nat. Acad. Sci. (USA),* 93:201-205.

Sakamoto T., Danzmann R.G., Ferguson M.M., Okamoto N. and P.E. Ihssen, 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture,* 173:33-43.

Tribout T., Bidanel J.P., Ducos A. and H. Garreau, 1998. Continuous genetic evaluation of farm and station tested pigs for production and reproduction traits in France. *6th WCGALP*, pp. 491-494.

Van Laere A.S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A.L., Haley C.S., Buys N., Tally M., Andersson G., Georges M. and L. Andersson, 2003. A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature*, 425:832-836.

Weir B., 1996. *Genetic Data Analysis II.* Sinauer Assoc., Inc., Sunderland, MA, USA, 445 pp.

Weller J., 2001. *Quantitative Trait Loci Analysis in Animals*. CABI Publ., New York, 287 pp.

Wright S., 1921a. Systems of mating. I. The biometric relations between parents and offspring. *Genetics,* 6:111-123.

Wright S., 1921b. Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics,* 6:144-161.