One size fits all?

Optimization of rainbow trout breeding program under diverse producer preferences and genotype-by-environment interaction.

Panya Sae-Lim

Thesis committee

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

Co-promotors
Dr. ir. J. Komen
Associate professor, Animal Breeding and Genetics Centre
Wageningen University

Dr. A. Kause
Principal Research Scientist
MTT Agrifood Research Finland, Jokioinen, Finland

Other members
Prof. dr. ir. J.A.J. Verreth, Wageningen University
Dr. M. Vandeputte, INRA, Jouy-en-Josas, France
Dr. A.K. Sonesson, Nofima, Ås, Norway
Prof. dr. ir. J.C.M. van Trijp, Wageningen University

This research was conducted under the auspices of the Graduate School of Wageningen Institute of Animal Sciences (WIAS).

One size fits all?

Optimization of rainbow trout breeding program under diverse producer preferences and genotype-by-environment interaction.

Panya Sae-Lim

Thesis

submitted in fulfillment 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 8 March 2013
at 1:30 p.m. in the Aula.

P. (Panya) Sae-Lim
One size fits all? Optimization of rainbow trout breeding program under diverse producer preferences and genotype-by-environment interaction,

PhD thesis, Wageningen University, the Netherlands (2013)
With references, with summaries in English and Dutch

ISBN 978-94-6173-464-8

Abstract

Global fish breeders distribute improved animal material to several continents to be farmed under diverse environments, and for very different market conditions. When establishing a global breeding program, there is a need to assess whether or not a single breeding objective satisfies the markets across different countries. It may be challenging to develop a single fish stock that performs well across all environments due to genotype-by-environment interaction (GxE). GxE is a phenomenon describing the possibility that different genotypes have a different sensitivity to changes in an environment. The objective of this thesis was to develop an optimized global breeding program for rainbow trout (*Oncorhynchus mykiss*) in terms of a balanced breeding goal that satisfies preferences of trout producers and maximized genetic gains across environments in the presence of GxE in production traits. Analytic hierarchy process (AHP) was used to estimate preferences, which can be aggregated to consensus preference values using weighted goal programming (WGP). The analysis revealed that the 6 most important traits were thermal growth coefficient (TGC), survival (Surv), feed conversion ratio (FCR), condition factor (CF), fillet percentage (FIL%), and late maturation (LMat). Individual trait preferences are different for farmers having different farming environments and producing different end-products. Calculating consensus preference values resulted in consensus desired genetic gains. To satisfy most farmers, consensus desired genetic gains can be taken into account in a global breeding strategy. Strong genotype re-ranking was found for all growth traits across environments. Based on simulation, re-location of breeding program led to highest total genetic gain for body weight at harvest. Alternatively, including sib performance into selection index increased genetic gain in all environments. Finally, environment-specific program can be used, but this is costly. There is a possibility of a conflict between 2 profits: from a breeding company and fish farmers and an optimum solution for that conflict can be found by using macroeconomics and cost-benefit analysis.

To my family and all my teachers...

Contents

1

General introduction

1.1 Introduction

1.1.1 Global rainbow trout breeding programs

Global animal breeding programs distribute improved animal material to several continents. The same genetic material is hence farmed under diverse environments and commercialized for very different market conditions. Such a set-up provides a unique opportunity for a study combining both animal genetics and global market development. Sustainable selective breeding programs improve multiple traits. A selection index typically includes production, fertility, product quality and animal health traits. The inclusion of traits into a selection index depends on a breeding objective and which traits can be measured. The breeding objective, i.e. aggregate genotype, summarizes the traits that are to be improved and their economic value for the food supply chain. To obtain economically optimized selection, breeding values of traits are weighted by their economic value, i.e., a euro change per unit of a change in a trait (Hazel, 1943). In this way, the emphasis in genetic changes occurs in the traits with the highest economic importance, and maximum economic gain is obtained.

Derivation of economic values from profit equations has been widely applied for terrestrial animals (Lindholm and Stonaker, 1957; Rothschild et al., 1981; Smith, 1983; Brascamp et al., 1985; Smith et al., 1986; Hermesch et al., 2003). The special characteristic of rainbow trout is that it is marketed in diverse ways depending on a country. Even within Europe, the market body weight varies from 400g to large-trout production at body weight of up to 4kg. The small portion-sized fish are typically marketed as whole, whereas large trout are processed into diverse products (cold or hot smoked fillets, fresh gutted fish, fresh fillets, gravlax fillets and convenient foods). It can be hypothesized that the breeding objective of a breeding company differs from individual farmer's breeding objectives, which depend on their local markets and farming conditions.

Fish breeding programs are at the initial phases of globalization, and there are no publications on global variation in market requirements or economic values. In fact, there are only two previous studies on economic values for aquaculture species (Henryon et al., 1999; Ponzoni et al., 2007). When animal production expands to novel areas or when a novel trait to be selected does not have direct economic impact (e.g. quality or animal welfare traits), profit functions need to be replaced by a desired genetic gains method to obtain selection index weights. Using this approach, it is possible to first derive genetic gains that the market requires ("desired" gains), and then back-calculate the index weights to be used to get these

gains. With the desired gains approach, it is important to involve all stakeholders in the process. Most participatory approach studies on defining breeding objectives have been limited to identifying traits that should be included in the breeding-objective (e.g., Tozer and Stokes, 2001; Tano et al., 2003; Nielsen et al., 2005; Gizaw et al., 2009). These "traits" are usually defined in general composite terms such as adaptation, growth, or reproduction. Little emphasis has been placed on using information from participatory studies to derive relative weights for such traits.

The potential weakness of the participatory approach is that the method is prone to biases due to personal opinions and preferences that may deviate from an economical optimum. However, for a company, satisfying the market requirements is the first step for a successful commercial activity. Moreover, for many traits, such as animal health and product quality, deriving euro based index weights is difficult or impossible in practice (Kanis et al., 2005).

1.1.2 Genotype-by-environment interaction

Even if the breeding objective would be shared across countries, it may be challenging to develop a single fish stock that performs well across all environments. This is because there may be genetic constraints in terms of genotype-by-environment interactions. Genotype-by-environment (GxE) interaction is a phenomenon that different genotypes have a different sensitivity to changes in an environment. GxE interaction has different forms: re-ranking across environments and heterogeneity of genetic variances (Falconer and Mackay, 1996; Lynch and Walsh, 1998).

Re-ranking is more serious than heterogeneity of genetic variance because re-ranking means that a single genotype is not superior across all environments (Calus, 2006; Mulder, 2007). GxE re-ranking is estimated by the strength of a genetic correlation of a trait measured from different environments (Falconer, 1952). It has been suggested that GxE interaction is unimportant from an agricultural point of view when the genetic correlation between environments is higher than 0.8 (Robertson, 1959). However, there is no formal proof for this general value and one can argue that any correlation lower than 1 can be significant provided the economic relevance of the trait is high enough. When there is no re-ranking (r_g = 1), selection in one environment (e.g., in nucleus) leads to the same genetic changes in all environments (e.g. nucleus and production environments), allowing an easy development of a single superior population.

However, with increasing levels of re-ranking, it becomes more difficult to develop a population that is superior across all environments. In practice, this means that when GxE interaction exists, either families need to be evaluated for their performance in several environments for breeding value estimation (e.g. sib testing stations) or environment-specific breeding programs need to be developed (de Jong and Bijma, 2002; Mulder and Bijma, 2005; Martinez et al., 2006; Mulder et al., 2006). In dairy cattle, Mulder et al. (2006) studied an optimization of breeding programs for different environments when G x E interaction was present. They designed four different breeding strategies accounting for breeding objectives in different environments and determined when a breeding program should be subdivided to meet the different breeding objectives. They concluded that the highest average genetic gain from a single breeding strategy in two equally important environments was accomplished when genetic correlation was higher than 0.6. When genetic correlation was lower than 0.6, it was best to have two environment-specific breeding programs. Dominik and Kinghorn (2008) evaluated the effect of neglecting G x E interaction on the efficiency of index selection, genetic gain and dollar response calculations. The efficiency of selection index was reduced by 1 to 25% and the total US dollar response per year was on average 33% overestimated. Consequently, the selection index should account for GxE interaction to avoid biased or less accurate index values that lead to lower genetic gains.

Breeding values obtained for fish reared in a closed breeding nucleus may not necessarily predict performance at every farm that these fish are sold to. Including many rearing conditions from various countries into a breeding program may solve the problem, but it also makes the running of a breeding program complicated and costly (Kolstad et al., 2006; Maluwa et al., 2006). Besides GxE interaction, there is also the problem that with so many different aquaculture farms, demands for specific traits in fish can be different. Previous studies in fish have studied GxE interaction across diets, production environments, and temperature treatments, e.g., in European sea bass (Saillant et al., 2006; Dupont-Nivet et al., 2008), tilapia (Khaw et al., 2009 and references therein), common carp (Wang and Li, 2007), Atlantic cod (Kolstad et al., 2006) and Rainbow trout (Fishback et al., 2002; Kause et al., 2003, 2007; Pierce et al., 2008; Vehviläinen et al., 2008 references therein). The general observation is that GxE interaction is often present but not important enough to be accounted for in a breeding program. The hypothesis is that GxE interaction will be most strong for traits which are determined by several environmental factors that each vary from one environment to another. Traits most probably prone to GxE interaction are survival, growth, age at maturation and

animal health, for which even zero genetic correlations between environments can occur (Kause et al., 2007; Vehvilainen et al., 2008).
The previously mentioned GxE studies on aquaculture species focus only on genetics and innately assume that a trait needs to be improved in the same amount and in the same direction in all environments (breeding objective is constant). Here it is argued that on a global scale, GxE interaction needs to be interpreted in the context of potentially different breeding objectives in different countries (breeding objective can vary). For instance, GxE interaction is not an issue if a trait has high importance only in the country where the nucleus is located, whereas the trait is not important elsewhere.

1.1.3 Bias and precision in estimation of genotype-by-environment interaction

Any experiment aiming at estimating genetic parameters should be designed to minimize bias (the degree to which the mean estimate is biased downwards or upwards from the true value) and precision (the degree of variation around the mean estimate) given fixed resources for the experiment. Such an optimization can be done using a simulation.
In this project a simulation study on bias and precision in GxE interaction studies was conducted prior to the design and execution of a GxE experiment. This simulation study had two functions. First, an optimal experimental design given a fixed number of animals to be used (a cost factor) was derived for the global GxE experiment conducted by the Troutlodge company. Second, it was possible to assess the degree to which suboptimal experimental designs lead to a bias in the estimated GxE correlation. The latter is of especial interest because of the fundamental role of the GxE correlation for the theories of evolutionary biology and quantitative genetics, and on practical animal breeding (Falconer 1952, de Jong and Bijma 2002, Mulder and Bijma 2005). There are previous studies on the statistical methods for estimation and significance testing of genetic correlations across environments (Windig, 1997; Astles et al., 2006). However, to our knowledge, there are no previous studies on the effects of population structure on the GxE estimation. Restricted maximum likelihood (REML) is known to be a robust genetic parameter estimation method even for unbalanced designs or when low number of individuals is used (Patterson and Thompson, 1971; Harville, 1977). This allows effective estimation of unbiased genetic parameters for diverse population structures. However, at extreme population structures (e.g., low family size or low

family number), biases can occur because parameter estimation is difficult, and standard errors become high. Given the low number of individuals used in many experiments in fish breeding, bias and precision in the estimates is thought to be a real issue.

1.2 Objectives

The overall objective of this project was to develop an optimized global rainbow trout breeding program in terms of 1) a balanced breeding goal that satisfies preferences of trout producers, and 2) maximized genetic gains across environments in the presence of GxE interaction of production traits. To achieve the overall objective, the project was divided into 4 stages:

1) To execute a survey to obtain information on preferred market traits and local production characteristics from trout farmers located on 3 continents.

2) To perform a simulation study to determine the optimal design of the GxE interaction experiment.

3) To execute an experiment using 3 major production environments (USA, Germany, Peru) to determine the amount of GxE interaction in fish growth.

4) To optimize rainbow trout breeding program accounting for GxE interaction.

The breeding program and customer base maintained by Troutlodge Inc., USA was used in the study. Troutlodge has the breeding nucleus in USA and distributes the genetically improved fish material to more than 50 countries across continents.

1.3 Outline of this thesis

This thesis consists of 3 parts. Firstly, it is examined whether or not trout farmers in the USA, South America and Europe prefer different animal characteristics. The market survey sets the breeding objective for the breeding program with respect to the relative importance of different animal traits in different parts of the world, i.e., which traits are the most important ones to be improved through selective breeding. Secondly, the genetic part of the study examines the genetic potential to satisfy the global market and the extent of GXE interaction. GxE interaction determines whether or not there are certain genotypes or families that are superior across all markets and all environmental conditions. Finally, optimization of the breeding program was conducted by means of comparing alternative breeding strategies.

In **Chapter 2**, market survey methods were combined with animal breeding methodology to assess the possibility to develop animal material that would satisfy global markets. A rainbow trout breeding program providing fish to Europe, South America, and USA was used as a study system. In **Chapter 3**, a stochastic simulation was performed to determine an optimal design of the GxE interaction experiment for unbiased estimation of a genetic correlation of a trait measured in different environments. In addition, it was investigated whether or not REML is robust against unbalanced designs. In **Chapter 4**, re-ranking of genotypes during growth and two-stage selection was studied. Two-stage selection is sometimes used in rainbow trout breeding programs. This chapter shows the general implication of two-stage selection to enhance multiple traits selection. In **Chapter 5**, the degree of GxE interaction was quantified in terms of heterogeneity of additive genetic variance and re-ranking of genotypes across 4 environments, located in 3 different continents. The genetic analysis was performed using multi-trait multi-environment model in multivariate analysis. In **Chapter 6**, the environmental factors explaining genotype-by-environment interaction for body weight across continents were identified. Changing the identified environmental parameters to be similar across environments would reduce GxE interaction and hence enhance breeding, leading to higher genetic gains in all environments. Two alternative methods were used to identify environmental parameters causing GxE: reaction norm model and two-step factor analytic model. In **Chapter 7**, optimization of a breeding program for GxE interaction was investigated using deterministic simulation. First, the consequence of GxE interaction to genetic gain in multiple environments was simulated when the selection methods were varied: one- and two- stage selection. Second, alternative breeding strategies were evaluated: re-location of the breeding

program, exploitation of sib information from production environments, the potential to compete against competing breeding programs, and establishment of separate environment-specific breeding programs. The developed approaches can be used by any breeding programs to direct genetic gains into desired directions that reflect market conditions.

References

Astles, P. A., Moore, A. J., Preziosi, R. F., 2006. A comparison of methods to estimate cross-environment genetic correlations. J. Evol. Biol. 19: 114-122.

Brascamp, E. W., Smith, C., Guy, D. R., 1985. Derivation of economic weights from profit equations. Anim. Prod. 40: 175-180.

Calus, P. M., 2006. Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle. Ph.D. Diss., Animal Breeding and Genetics, Wageningen Univ., The Netherlands.

De Jong, G., Bijma, P., 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. Livest. Prod. Sci. 78: 195-214.

Dominik, S., Kinghorn, B. P., 2008. Neglecting genotype × environment interaction results in biased predictions from selection index calculations. Livest. Sci. 114: 233-240.

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., 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., 1952. The Problem of environment and selection. Am. Nat. 86: 293-298.

Falconer, D. S., Mackay, T. F. C., 1996. Introduction to quantitative genetics. p 464 pp. Longman, Essex, UK.

Fishback, A. G., Danzmann, R. G., Ferguson, M. M., Gibson, J. P., 2002. Estimates of genetic parameters and genotype by environment interactions for growth traits of rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. Aquaculture 206: 137-150.

Gizaw, S., Komen, H., van Arendonk, J. A. M., 2009. Participatory definition of breeding objectives and selection indexes for sheep breeding in traditional systems. Livest. Sci. 128: 67-74.

Harville, D. A., 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72: 320-338.

Hazel, L. N., 1943. The genetic basis for constructing selection indexes. Genetics 28: 476-490.

Henryon, M., Purvis, I. W., Berg, P., 1999. Definition of a breeding objective for commercial production of the freshwater crayfish, marron (*Cherax tenuimanus*). Aquaculture 173: 179-195.

Hermesch, S., Kanis, E., Eissen, J. J., 2003. Economic weights for feed intake in the growing pig derived from a growth model and an economic model. J. Anim Sci. 81: 895-903.

Kanis, E., K. H. De Greef, A. Hiemstra, and J. A. M. van Arendonk., 2005. Breeding for societally important traits in pigs. J. Anim Sci. 83: 948-957.

Kause, A., O. Ritola, T. Paananen, E. Mäntysaari, and U. Eskelinen. 2003. Selection against early maturity in large rainbow trout *Oncorhynchus mykiss*: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228: 53-68.

Kause, A., Tobin, D., Mäntysaari, E. A., Martin, S. A. M., Houlihan, D. F., Kiessling, A., Rungruangsak-Torrissen, K., Ritola, O., Ruohonen, K., 2007. Genetic potential for simultaneous selection of growth and body composition in rainbow trout (*Oncorhynchus mykiss*) depends on the dietary protein and lipid content: Phenotypic and genetic correlations on two diets. Aquaculture 271: 162-172.

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

Kolstad, K., Thorland, I., Refstie, T., Gjerde, B., 2006. Genetic variation and genotype by location interaction in body weight, spinal deformity and sexual maturity in Atlantic cod (*Gadus morhua*) reared at different locations off Norway. Aquaculture 259: 66-73.

Lindholm, H. B., Stonaker, H. H., 1957. Economic importance of traits and selection indexes for beef cattle. J. Anim Sci. 16: 998-1006.

Lynch, M., Walsh, B. 1998. Genetics and analysis of quantitative traits. p 980 pp. Sinauer Associates, Inc., U.S.A.

Maluwa, A. O., Gjerde, B., Ponzoni, R. W., 2006. Genetic parameters and genotype by environment interaction for body weight of *Oreochromis shiranus*. Aquaculture 259: 47-55.

Martinez, V., Kause, A., Mäntysaari, E., Mäki-Tanila, A., 2006. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: II. Two-stage selection. Aquaculture 254: 195-202.

Mulder, H. A., Bijma, P., 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83: 49-61.

Mulder, H. A., Veerkamp, R. F., Ducro, B. J., van Arendonk, J. A. M., Bijma, P., 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment Interaction. J. Dairy Sci. 89: 1740-1752.

Mulder, H.A., 2007. Methods to optimize livestock breeding programs with genotype by environment interaction and genetic heterogeneity of environmental variance. Ph.D. Diss., Animal Breeding and Genetics group, Wageningen Univ., The Netherlands.

Nielsen, H. M., Christensen, L. G., Groen, A. F., 2005. Derivation of sustainable breeding goals for dairy cattle using selection index theory. J. Dairy Sci. 88: 1882-1890.

Patterson, H. D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58: 545-554.

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

Ponzoni, R. W., Nguyen, N. H., Khaw, H. L., 2007. Investment appraisal of genetic improvement programs in Nile tilapia (*Oreochromis niloticus*). Aquaculture 269: 187-199.

Robertson, A., 1959. The Sampling variance of the genetic correlation coefficient. Biometrics 15: 469-485.

Rothschild, M. F., Carlson, J. P., Christian L. L., 1981. Comparison of selection index economic weights and prices paid for performance-tested boars. J. Anim Sci. 53: 575-579.

Saillant, E., Dupont-Nivet, M., Haffray, P., 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.

Smith, C., 1983. Effects of changes in economic weights on the efficiency of index selection. J. Anim Sci. 56: 1057-1064.

Smith, C., James, J. W., Brascamp, E. W., 1986. On the derivation of the economic weights in livestock improvement. Anim. Prod. 43: 545-551.

Tano, K., Kamuanga, M., Faminow, M. D., Swallow, B., 2003. Using conjoint analysis to estimate farmer's preferences for cattle traits in West Africa. Ecol. Econ. 45: 393-407.

Tozer, P. R., Stokes, J. R., 2001. Using multiple objective programming in a dairy cow breeding program. J. Dairy Sci. 84: 2782-2788.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H., Paananen, T., 2008. Survival of the currently fittest: Genetics of rainbow trout survival across time and space. Genetics 180: 507-516.

Wang, C. H., Li, S. F., 2007. Genetic effects and genotype × environment interactions for growth-related traits in common carp, *Cyprinus carpio* L. Aquaculture 272: 267-272.

Windig, J. J., 1997. The calculation and significance testing of genetic correlations across environments. J. Evol. Biol. 10: 853-874.

2

Defining desired genetic gains for rainbow trout breeding objective using analytic hierarchy process

P. Sae-Lim[1], H. Komen[1], A. Kause[1], J. A. M. van Arendonk[1], A. J. Barfoot[2], K. E. Martin[2], J. E. Parsons[2]

[1] Animal Breeding and Genomics Centre, Wageningen University, P. O. Box 338, 6700 AH, Wageningen, the Netherlands; [2] Troutlodge, Inc., 12000 McCutcheon Rd., Sumner, WA 98390

Journal of Animal Science (2012) 90:1766-1776

Abstract

Distributing animals from a single breeding program to a global market may not satisfy all producers, as they may differ in market objectives and farming environments. Analytic hierarchy process (AHP) is used to estimate preferences, which can be aggregated to consensus preference values using weighted goal programming (WGP). The aim of this study was to use an AHP-WGP based approach to derive desired genetic gains for rainbow trout breeding and to study whether breeding trait preferences vary depending on commercial products and farming environments. Two questionnaires were sent out. Questionnaire-A (Q-A) was distributed to 178 farmers from 5 continents and used to collect information on commercial products and farming environments. In this questionnaire, farmers were asked to rank the 6 most important traits for genetic improvement from a list of 13 traits. Questionnaire B (Q-B) was sent to all farmers who responded to Q-A (53 in total). For Q-B, preferences of the 6 traits were obtained using pairwise comparison. Preference intensity was given to quantify (in % of a trait mean; G%) the degree to which 1 trait is preferred over the other. Individual preferences, social preferences, and consensus preferences (Con-P) were estimated using AHP and WGP. Desired gains were constructed by multiplying Con-P by G%. The analysis revealed that the 6 most important traits were thermal growth coefficient (TGC), survival (Surv), feed conversion ratio (FCR), condition factor (CF), fillet percentage (FIL%), and late maturation (LMat). Ranking of traits based on average Con-P values were Surv (0.271), FCR (0.246), TGC (0.246), LMat (0.090), FIL% (0.081), and CF (0.067). Corresponding desired genetic gains (in % of trait mean) were 1.63, 1.87, 1.67, 1.29, 0.06, and 0.33%, respectively. The results from Con-P values show that trait preferences may vary for different types of commercial production or farming environments. This study demonstrated that combination of AHP and WGP can be used to derive desired gains for a breeding program and to quantify differences due to variations market demand or production environment.

Key words: breeding goals, desired genetic gain, participatory approach, rainbow trout, relative weights

2.1 Introduction

A breeding objective describes the desired rate and direction of genetic change for a specified group of traits in a breeding program for a particular species. The benefits of these changes are commonly expressed in economic terms from the perspective of commercial farmers. However, in practice, many large breeding organizations determine the direction of genetic change through trial-and-error adjustments of their selection index based on perceived market demands and preferences (Kanis et al., 2005; Amer, 2006). When a breeding company distributes animals from a single breeding program to a global market, the breeding objective may not satisfy all customers, as they differ in products and market objectives. In such situations, market research techniques can be used to derive information on trait preferences of producer groups with different objectives. The analytic hierarchy process (AHP; Saaty, 1980) is often used in multicriteria decision-making processes to estimate preference values for a given set of traits. First, pairwise comparison (Thurstone, 1927) is used to assess relative importance of different alternatives for a set of criteria or attributes. Next, individual preferences are estimated using the AHP and aggregated to group preferences using weighted goal programming (WGP; Linares and Romero, 2002). When groups differed in opinion, WGP-based models can be used to construct consensus preference values.

Rainbow trout (*Oncorhynchus mykiss*) are marketed as eyed eggs distributed from a single breeding program. Rainbow trout producers differ in terms of commercial products (e.g., fry, pan-sized fish, or fillet) or environments. The aim of this study was to use AHP and WGP to define desired genetic gains for a rainbow trout breeding objective, based on consensus preference values and to investigate potential variation in preferences for breeding traits due to differences in commercial products or environments.

2.2 Materials and methods

2.2.1 Questionnaire distribution

Troutlodge, Inc. (Sumner, WA) supplies eyed rainbow trout eggs to fi sh farmers in more than 60 countries. Troutlodge and its customer database were used to conduct a worldwide survey. The questionnaires were originally designed in English and translated into Spanish, German, and Polish. The farmers with Internet connections received questionnaires by e-mail; otherwise, by regular mail. The

questionnaire (Q) consisted of 2 parts, Q-A and Q-B. Questionnaire A was designed to collect information that was needed for the design of the pairwise comparison (Q-B), as detailed below. Questionnaire A was distributed to 178 farmers on 5 different continents: North America, South America, Europe, Africa, and Asia. In total, 53 responses (29.8%) were received. Questionnaire B was sent to the respondents of Q-A. Thirty-nine responses (73.6%) were returned and used in the analysis of Q-B.

Questionnaire A

The objectives of Q-A were to obtain information on each participant's farming environment and to identify and rank the traits they would most like to see improved. Questions were grouped into 2 sections: 1) commercial products and farming environments, and 2) traits to be included in the breeding goal. Questions about commercial products and farming environments were grouped into 4 categories: commercial products, altitude, water circulation, and water temperature. In the second part, farmers were asked to select and rank the 6 most preferred traits out of 13 alternative traits to be included in the breeding objective, by using percentages (from 0 to 100%). The rankings summed to 100%. A detailed definition of each trait was provided (Table 2.1) to ensure consistent meaning between participants. Following the general guidelines of Saaty (2003), participants were restricted to choosing 6 traits to limit the number of pairwise comparisons in Q-B to 15.

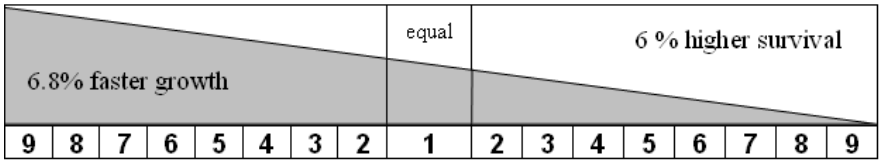

Figure 2.1. Pairwise comparison. The scale expresses both the intensity of trait preference and the percentages of genetic improvement to be obtained in the traits. The diagonal line indicates the decrease in the improvement of growth, when survival is preferred (right hand side). For example, if number 9 is selected from the right hand side, this means a respondent would like to have 6% higher survival and no genetic improvement in growth.

Table 2.1 Trait definitions used in the questionnaire A

Trait	Definition
Growth	Growth is defined as a thermal growth coefficient (TGC), a measurement of the growth rate corrected for different water temperatures.
Condition factor	Condition factor reflects the appearance in terms of body shape, taking into account fish weight and length.
Survival at harvest	Percentage of fish survived until harvest (market-sized fish), i.e., percentage of fish live at harvest out of the number of fish initially stocked in a raceway.
Fillet percentage	Fillet percentage is the ratio of fillet weight and ungutted body weight.
Uniformity	Consistency of growth and body size within a given population. Higher uniformity indicates that harvest fish are of similar size, and that the population of fish reaches harvest size within a short window of time.
Late maturation	This means the age at first maturation. Maturation may reduce flesh quality and the appearance of harvested fish, and may also make fish aggressive towards other fish.
Flesh color (redness)	Coloration is normally measured using the Roche Scale. The scale ranges from 20 (least red) to 34 (most red).
Skin color and spottiness	For spottiness, fish can vary from not spotted to fully spotted. For skin color, base skin color can vary from silvery to brown, blue or green. Besides, maturation can further induce more dense spottiness, darkened skin color, and a pink stripe on the side.
Feed Conversion Ratio (FCR)	A fish's efficiency in terms of converting feed mass into increased body weight. Ordinary, the lower FCR, the better efficiency.
High water temperature tolerance	Upper thermal tolerance or heat stress tolerance. Low tolerance may reduce production traits, e.g. growth performance, and survival, when fish are exposed to the higher than optimum level of water temperature (exceed 20^{o}C or 68^{o}F).
Low water temperature tolerance	Low water temperature (1-4^{o}C) may reduce growth performance, osmotic regulation, and survival. Rearing rainbow trout in lower water temperature (1-4^{o}C) than usual may be required in some geographical areas.
Deformity	Deformities commonly result from disorders of bone structures, which can be caused by either genetic defects or a wide range of environmental factors or both.
Disease resistance	Fish with higher disease resistance have lower chance of getting an infection or have higher survival when fish are exposed to an outbreak.

Questionnaire B

From Q-A, the 8 highest ranked traits were growth (thermal growth coefficient: TGC), disease resistance, survival (Surv), feed conversion ratio (FCR), uniformity, condition factor (CF), fillet percentage (FIL%), and late maturation (LMat). Disease resistance was not included in the follow-up as there were many different diseases mentioned. Uniformity was removed from further analysis due to the lack of genetic parameters. The objective of Q-B was to determine the relative importance of the remaining 6 traits using multiple pairwise comparisons. To do this, the 6 traits were assigned to 15 pairwise comparison questions [6 × (6 − 1)/2 = 15 pairs]. Each pairwise comparison expressed both the intensity of trait preference and the percentage of genetic improvement of both traits being compared. Saaty's scale was used for scoring the intensity of trait preference (Saaty, 1980), with scores ranging from 1 to 9. A score of 1 meant that both traits were equally preferred. A score of 9 meant that 1 trait was absolutely more preferred than the alternative trait. In that case, the percentage of improvement in the preferred trait was at its maximum potential whereas the other trait was not improved at all (see Figure 2.1 for detailed explanation). Expressing scores only in terms of "higher" or "faster" might have resulted in biased responses based on exaggerated expectations that could not be matched with actual realized genetic gains. Therefore, the percentage of genetic improvement of a trait was estimated as the genetic gain in % of trait mean obtained after 1 generation of phenotypic selection on that trait only.

The mean and variance of traits reported in literature usually differ in scale due to differences in age and environments. To eliminate the scaling effect, σp was substituted with CV (CV = σ_p/μ_p) in calculating the response to phenotypic selection (Falconer and Mackay, 1996):

$$G = ih^2 CV\mu_p \quad [1]$$

where G = genetic gain, i = selection intensity, $h2$ = heritability, σp = phenotypic standard deviation, and μp= phenotypic mean. A percentage of maximum potential genetic improvement (G%; in % of original trait mean) was defined as a ratio of G to the trait mean of the previous generation:

$$G\% = ih^2 CV \cdot 100\% \quad [2]$$

Equation 2 was used for the continuous traits (TGC, FCR, CF, and FIL%). For the binary traits (Surv and LMat), an expected genetic change in observed trait value was derived by assuming that selection was changing the underlying liability scale of a trait. Selection intensity was fixed to 1 (38% selected animals). Heritability and CV for TGC, Surv, CF, FIL%, and LMat of rainbow trout and other salmonid fish were obtained from literature. The parameter estimates and their references are summarized in Table 2.2.

2.2.2 Trait preference estimation

Preference values were estimated on 3 different levels. First, the individual preference values of each respondent were calculated. Next, individual preferences were aggregated in social group preference values. Finally, consensus preference values across all respondents were calculated and used for calculating the desired genetic gains.

Individual preference value (Ind-P)

The AHP (Saaty, 1980) was used to calculate the individual preference value (Ind-P).

Table 2.2 Average h^2 and CV derived from literature and the expected percentage of genetic improvement of the 6 most important traits used in questionnaire part B

Trait	h^2	Ref.[1] (h^2)	CV	Ref. (CV)	Percentage of improvement
TGC	0.32	17	21.23	13,17	6.8% faster
Surv[2]	0.17	16	-	-	6.0% higher
FCR[3]	0.17	12	45.69	12	7.6% lower
CF	0.44	2,3,5,6,7, 9	11.22	3,5,6,9,10	4.9% higher
FIL%	0.19	6,9,14,15	3.82	6,9,14,15	0.7% higher
LMat[2]	0.18	1,4,8,11	-	-	14.3% later

[1]Ref. = Reference: 1 = McKay et al. (1986), 2 = Gjerde and Schaeffer (1989), 3 = Elvingson and Johansson (1993), 4 = Wild et al. (1994), 5 = Rye and Refstie (1995), 6 = Kause et al. (2002), 7 = Kause et al. (2003a), 8 = Kause et al. (2003b), 9 = Neira et al. (2004), 10 = Gjedrem (2005), 11 = Kause et al. (2005), 12 = Kause et al. (2006), 13 = Dumas et al. (2007), 14 = Kause et al. (2007), 15 = Powell et al. (2008), 16 = Vehviläinen (2008), 17 = Silverstein et al. (2009).
[2]These traits were modelled as threshold characters.
[3]The parameters for FCR were derived using the Delta method (Lynch and Walsh, 1998) from the parameters of daily feed intake and daily weight gain reported by Kause et al. (2006).

The intensity of preferences was expressed as relative importance (a_{ij}), score 1 to 9, between the *i*th and *j*th traits (*i*, *j* =1, 2, 3…, 6) respectively. A 6 × 6 pairwise comparison matrix (P),

$$P = \begin{bmatrix} 1 & a_{12} & a_{13} & . & . & a_{16} \\ 1/a_{12} & 1 & a_{23} & . & . & a_{26} \\ 1/a_{13} & . & 1 & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & 1 & a_{56} \\ 1/a_{16} & 1/a_{26} & . & . & 1/a_{56} & 1 \end{bmatrix}$$

was constructed in which all a_{ij} were elements of the upper diagonal. The lower diagonal contains the reciprocal relative importance: $a_{ji} = 1/a_{ij}$. In this matrix, the diagonal elements represent the comparison of each trait with itself, that is, $a_{i=j} = 1$. For each set of responses from an individual, the eigenvector of the P matrix, corresponding with maximum eigenvalue, was calculated. Individual preference values were obtained by normalizing the eigenvector (Saaty, 1980, 1990).

Consistency of the responses of farmers in pairwise comparisons was checked by calculating the consistency ratio (CR) of the pairwise comparison matrix. A judgment is commonly reliable when CR is less than 0.10 (Saaty, 1980). The vector of Ind-P values and corresponding CR were estimated using Super Decisions software (Saaty, 2003). When CR was greater than 0.1, a new P matrix with maximum eigenvalue less than the one from the original P matrix was reconstructed using an iterative algorithm (Zeshui and Cuiping, 1999). This resulted in CR less than 0.1.

Social group preference value

To assess whether or not farmers with different commercial products or farming environments have different preferences for traits, the social group preference (Soc-P) value was calculated separately for 4 categories: commercial products, altitude, water circulation, and water temperature. Each category contained different social groups. For commercial products, 3 social groups were distinguished: producers of both pan-sized and large fish, producers of only pan-sized fish, and producers of fry. Farm altitude and water temperature may affect fish performance, e.g. by changing the amount of dissolved oxygen in water. Thus they may impact trait preferences. For altitude, 2 social groups were compared:

producers at high altitude, and low altitude. Water circulation was separated into 2 social groups: farmers that use flow-through system, and farmers that use recirculating aquaculture system (RAS). Fish farms having maximum water temperature in the range of 9-14^{o}C and minimum temperature below 9^{o}C were assigned to "low" water temperature social group. Fish farms having minimum water temperature in the range of 9-14^{o}C and maximum temperature exceeding 14^{o}C were assigned to "high" water temperature social group. Finally, fish farms having minimum water temperature below 9^{o}C and maximum water temperatures exceeding 14^{o}C were assigned to "high and low" water temperature social group. The cut-off values for water temperature were obtained from FAO (2011). Weighted goal programming (WGP; Linares and Romero, 2002) was applied for aggregation of individual preference values (González-Pachón and Romero, 1999) into social groups. This approach minimizes the sum of negative and positive deviations ($n_{ik} + p_{ik}$) between the social group preference weight (W_i^j) and Ind-P values (A_i^{kj}) of each member in the social group:

Achievement function:

$$Min \sum_{i=1}^{6} \sum_{k=1}^{Nj} \left(n_{ik} + p_{ik}\right)^{\pi}$$

subject to

Goals:

$$W_i^j + n_{ik} - p_{ik} = A_i^{kj} \qquad [3]$$

where *i* = 1 to 6 traits to be judged by *j* = 1, 2,…, *m* social groups within each category. A_i^{kj} is Ind-P derived from AHP of the *i*th trait judged by the *k*th member of the *j*th social group. W_i^j = aggregated preference weight attached to the *i*th trait by the *j*th social group. n_{ik} and p_{ik} are negative and positive deviations, respectively of A_i^{kj} from W_i^j. The weight attached to the sum of deviation variables π, was set to 1. When π equals 1, the solution is statistically defined by the median weight (Linares and Romero, 2002). Model [3] was solved by using LINGO computer software (LINDO system Inc., 1999). Social group preference values were derived from normalized W_i^j and summed up to 1 for each category.

Consensus preference value (Con-P)

Consensus or "society" preference values across all respondents were calculated using extended WGP (Linares and Romero, 2002). Extended WGP is defined as follows:

Achievement function:

$$Min\left[(1-\lambda)D+\lambda Z\right] \quad \text{with} \quad Z=\sum_{i=1}^{6}\sum_{j=1}^{m}(\bar{n}_{ij}+\bar{p}_{ij})$$

Here, $\bar{n}_{ij}+\bar{p}_{ij}$ = the sum of the negative and positive deviations of the *j*th social group preference value W_i^j from the consensus preference value W_i^s for the *i*th trait (i =1, 2,.., 6). Summing these values for all 6 traits combined for social group *j* produces the disagreement D_j of social group *j* with the consensus preference values. *Z* is calculated by summing all values D_j with *j* =1 to m. *D* represents the maximum value observed for D_j:

The goals are:

$$\sum_{i=1}^{6}(\bar{n}_{i1}+\bar{p}_{i1})-D\le 0,$$

$$\cdots$$

$$\sum_{i=1}^{6}(\bar{n}_{im}+\bar{p}_{im})-D\le 0,$$

$$W_i^s+\bar{n}_{ij}-\bar{p}_{ij}=W_i^j, \quad i\in\{1,\ldots,6\}, j\in\{1,\ldots,m\} \qquad [4]$$

The coefficient value λ determines the emphasis on the minority groups. When λ = 0, model (4) equals WGP with $\pi=\infty$ and the disagreement of the most displaced social group is minimized. When λ =1, model (4) equals WGP with π=1 and the consensus values represent the median weights (Linares and Romero, 2002). Model [4] with λ varied from 0 to 1 was solved in LINGO.

In our dataset, each Ind-P value results from combination of commercial products and farming environments. Consensus preference values were first estimated for commercial products and water temperature, as these categories showed large disagreement between social groups. To use the full dataset and to avoid sampling bias, we classified all observations that did not belong to a particular social group

as “unassigned” when estimating Soc-P values. These values were then used to calculate Con-P values using model [4].

To obtain Con-P values across all different commercial products and farming environments, Ind-P values were randomly assigned to 5 social groups (R_1,…,R_5). Then Soc-P values were first estimated and normalized using model (3). Finally, Con-P values were estimated using extended WGP. This procedure was replicated 6 times to obtain the average Con-P values.

2.2.3 Derivation of desired genetic gains

Desired genetic gains (desiredG%) of each trait was derived by multiplying percentage of improvement (G%, Table 2.2) with the average Con-P values from random groups, calculated with model [4].

2.3 Results

2.3.1 Preference value estimation

Individual preference value (Ind-P) obtained from AHP were plotted as a boxplot shown in Figure 2.2. Median of the preferences for TGC (0.237), Surv (0.216), FCR (0.209), CF (0.048), FIL% (0.080), and LMat (0.092) showed that improvement in TGC, Surv, and FCR were more preferred than improvement in CF, FIL%, and LMat. There were, however, large differences between individual respondents. There were a few extreme values found for CF, FIL% and LMat, indicating that a few farmers heavily preferred these processing traits.

There were 9 respondents that had original CR lower than 0.1 (0.024 to 0.081) which means that these respondents were consistent in their answers. Thirty respondents had original CR ranging 0.118 to 1.506. After modification of the *P* matrix of those 30 observations, the final CR ranged from 0.002 to 0.108.

Social preference value (Soc-P) on traits differed between commercial products. Sample size of each social group is shown in Figure 2.3. Not all respondents could be included in 1 of the catedories due to missing or ambiguous information in the questionnaire (see Figure 2.3 for numbers). The preference for FIL% was greater when farmers were producing pan-sized and large trout. LMat was more important for producers of large fish. Fry producers regarded growth as the most preferred trait for improvement.

There were large differences between preference values among water temperature groups. The high water temperature group expressed a stronger preference for TGC and Surv, at the expense of CF and LMat. In contrast, the low water temperature group emphasized LMat at the expense of preferences on TGC and Surv. Farmers experiencing wide fluctuations in water temperature regarded Surv as the most important trait to be improved.

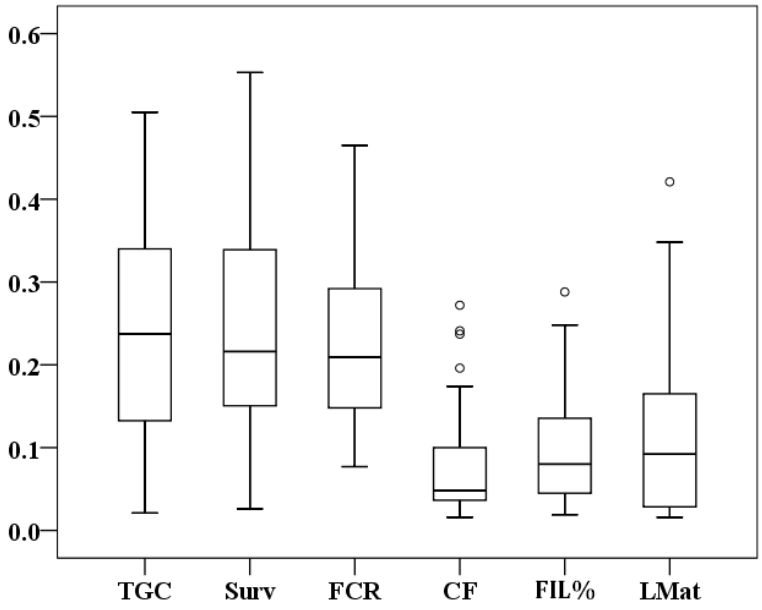

Figure 2.2. Boxplot of individual preference values (Y-axis) obtained from AHP on the 6 most important traits (X-axis). Boxplot shows median (horizontal line within the box), interquartile range, most extreme values of individual preferences (Y-bar error), and outliers (circle dots). TGC = thermal growth coefficient; Surv = overall survival; FCR = feed conversion ratio; CF = condition factor; FIL% = fillet percentage; LMat = late maturation.

There were small differences in Soc-P values for altitude and water circulation categories (Figure 2.3). Farmer's preferences for TGC and Surv differed slightly, depending on altitude. At high altitude, TGC was considered less important than Surv whereas the reverse was true for low-altitude farmers. Preference for FIL% was higher for high altitude farmers compared to low altitude farmers. Farmers using recirculating aquaculture system expressed a higher preference for TGC and Surv and a lower preference for FCR, FIL%, and LMat compared to farmers using flow through system.

As the most variation among Soc-P values was observed in commercial products and water temperature categories, Con-P values were calculated for these categories (Table 2.3). A set of Con-P values with a combination of the lowest *Z* and D_j represents the best consensus among groups because such Con-P values are the solution from the highest agreement of each social group and overall. For commercial products, the best Con-P values, obtained with λ[1.0, 0.3), were FCR (0.272), Surv (0.258), TGC (0.231), FIL% (0.093), CF (0.080), and LMat (0.066). For water temperature, the best consensus values, corresponding with λ(1.0, 0.3), were Surv (0.293), TGC (0.253), FCR (0.220), LMat (0.105), FIL% (0.074), and CF (0.055).

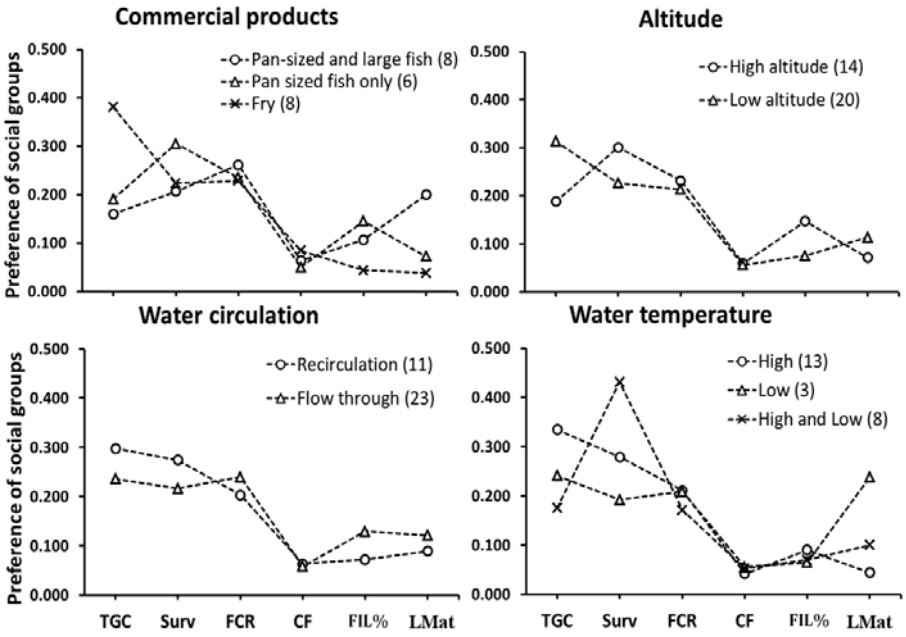

Figure 2.3. Social group preference values based on different commercial products and farming environments. TGC= thermal growth coefficient; Surv = overall survival; FCR = feed conversion ratio; CF = condition factor; FIL% = fillet percentage; LMat = late maturation. Numbers in parenthesis are the number of respondents assigned into each social group.

2.3.2 Desired genetic gain of rainbow trout

Average Con-P values from all replicates of random groups were Surv (0.271), TGC=FCR (0.246), LMat (0.090), FIL% (0.081), and CF (0.067) (Table 2.4). Desired gains (desiredG%) calculated using the average of Con-P values multiplied with G%

are shown in Table 2.4. DesiredG% was 1.87%, 1.67%, 1.63%, 1.29%, 0.33%, and 0.06% for FCR, TGC, Surv, LMat, CF, and FIL%, respectively.

2.4 Discussion

In this study, we propose a novel participatory approach with a use of pairwise comparison, analytical hierarchy process (AHP), and weighted goal programming (WGP), together with expected maximum genetic improvement (G%, model 2) to derive desired genetic gains for a rainbow trout breeding objective.

2.4.1 Participatory approach

Aggregation of different criteria to consensus preference values is normally used in multiple criteria decision-making. Animal breeders have developed and applied methods to estimate preference values for defining a breeding objective, for example, using a conjoint analysis (Tano et al., 2003) and choice experiment (Wurzinger et al., 2006). The general steps of these participatory approaches are firstly, to define the most important breeding traits, and secondly, to compare the traits, e.g. fast growth, coat color, and high milk yield. The final step is to estimate preference values of traits, however, most participatory approaches do not provide desired genetic gains or weighting factors for traits under selection.

Here, we developed an alternative participatory approach to estimate preference values and to determine desired genetic gains for a breeding objective. When comparing preferences for different traits, most participatory studies use qualitative attributes for both qualitative traits (e.g. coat color in sheep), and quantitative traits (e.g. slow or fast growth). Not all respondents may interpret such qualitative attributes in the same way. In this study, G% was used as a quantitative (genetic) attribute to compare traits. As a result, more consistent and precise (pairwise) comparison of the traits can be made by all respondents, which provided quantitative information for deriving desired gains. In our study, all individual respondents were given equal weight when calculating consensus preference values. Our method can be further modified by weighting preferences of individual responders, e.g. according to the amount of eggs they buy.

Table 2.3 Consensus preference values (Con-P) estimated using extended WGP model.

Method	λ[1]	Consensus preference value[2]						Social group disagreement				
		TGC	Surv	FCR	CF	FIL%	LMat	D_1[3]	D_2	D_3	D_4	Z[4]
Commercial products	[1.0, 0.3)	**0.231**	**0.258**	**0.272**	**0.080**	**0.093**	**0.066**	0.258	0.192	0.258	0.136	0.844
	[0.3, 0.0)	0.236	0.254	0.276	0.079	0.091	0.065	0.258	0.206	0.258	0.122	0.844
	[0.0]	0.226	0.264	0.278	0.082	0.084	0.067	0.258	0.258	0.258	0.154	0.928
Temperature	[1]	0.281	0.228	0.234	0.062	0.078	0.117	0.165	0.250	0.352	0.100	0.867
	(1.0, 0.3)	**0.253**	**0.293**	**0.22**	**0.055**	**0.074**	**0.105**	0.234	0.181	0.255	0.197	0.867
	[0.3, 0.0)	0.245	0.297	0.222	0.055	0.074	0.106	0.245	0.192	0.245	0.208	0.888
	[0.0]	0.251	0.301	0.217	0.054	0.073	0.104	0.245	0.245	0.245	0.208	0.941
Random groups[5]		0.246	0.271	0.246	0.067	0.081	0.090					
S.D.[6]		0.030	0.023	0.027	0.009	0.013	0.022					

[1] λ : range of λ producing equal solutions, where "[" and "]" mean equal to and "(" and ")" mean higher or lower than boundary value, respectively.

[2] TGC = thermal growth coefficient; Surv = overall survival; FCR = feed conversion ratio; CF = condition factor; FIL% = fillet percentage; LMat = late maturation.

[3] Underlined number indicates maximum social group disagreement (*D*) on the Con-P obtained. *D*1, *D*2, *D*3, and *D*4 are pan-sized and large fish, pan-sized fish, fry, unassigned group for commercial products, and are high, low, high and low water temperature, and unassigned group for water temperature. Bolding indicates the best consensus preference values.

[4] *Z* = overall disagreement (summation of *D*1 to *D*4).

[5] Average of Con-P values from all replicates.

[6] SD = standard deviation of all replicates.

Table 2.4 Percentage of improvement (G%) and desired genetic gain (desiredG%), obtained using extended WGP and random social groups.

Trait[1]	G%[2]	Con-P[3]	desiredG%
TGC	6.80%	0.246	1.67%
Surv	6.00%	0.271	1.63%
FCR	7.60%	0.246	1.87%
CF	4.90%	0.067	0.33%
F%	0.70%	0.081	0.06%
LMat	14.30%	0.090	1.29%

[1]TGC = thermal growth coefficient; Surv = overall survival; FCR = feed conversion ratio; CF = condition factor; FIL% fillet percentage; LMat = late maturation.
[2]Expected maximum genetic improvement (in % of a trait mean) with selection intensity of 1.
[3]Average of consensus preference values (Con-P) from all replicates.

The participatory approach presented here, and the calculation of economic values are two complementary ways of defining the weight of traits for a selection index. The participatory approach is based on satisfying farmer preferences which is essential for a successful breeding company. Furthermore, sustainable breeding objectives are increasingly important in modern breeding programs (Amer, 2006). Societal or non-market value traits such as animal health or behavior are included into a sustainable objective, but calculation of economic values for these traits is difficult (Kanis, 2005; Monsen et al., 2010). By including societal, or non-market value traits into pairwise comparison (Olesen et al., 1999), it is possible to define desired gains by using our approach. However, this approach still requires that traits, such as welfare traits are first translated into traits for which G% can be estimated.
The merit of economic values, in turn, is that they do not depend on potentially subjective opinions of farmers, like questionnaires do. However, in study systems like ours, it is unrealistic to calculate economic values for extensive number of traits using economic information on very diverse production systems and environments originating from different continents.

2.4.2 Methods

In our questionnaire we used 6 traits to construct a 6x6 pairwise comparison matrix. This number is generally recommended as do-able; it may be a mental challenge for respondents to be consistent when the number of traits is more than 7, leading to 21 pairwise comparisons (Saaty, 2003). However, most respondents gave inconsistent responses (CR> 0.1) in this study. Based on literature, there are 2

other factors that could lead to high consistency ratio. Firstly, the use of 9-point scale may not necessarily lead to a consistent pairwise comparison matrix (Murphy, 1993). Assume there are 3 traits being compared, T_1, T_2, and T_3. Trait T_1 is 7 times more important than T_2, and T_2 is 9 times more important than T_3. Thus, T_1 should be 63 times more important than T_3 but the maximum score that can be given is 9. Secondly, the traits in the comparison should be of similar magnitude (Saaty, 2003; Tano et al., 2003). In our questionnaire, most traits were similar in terms of G%. However, late maturation (G% = 14.3%) and fillet percentage (G% = 0.7%) were clearly different and this might have contributed to high CR. When constructing a questionnaire, all these factors should be considered.

A number of computational methods have been developed to reduce consistency ratio, for example, fuzzy AHP (Leung and Cao, 2000), fuzzy linguistic preference (Wang and Chen, 2008), modification of pairwise comparison matrix (Zeshui and Cuiping, 1999), and heuristic approach (Cao et al., 2008). Here we used Zeshui and Cuiping's method which reduces consistency ratio but the obtained individual preference values remain close to the original version.

In this study, we used extended WGP to obtain Con-P values. The benefits of using WGP are 1) possibility to study of variation in preferences of traits due to different social groups and 2) estimation of Con-P values which account for disagreement among social groups. Fitting WGP can easily be done using, for example, LINGO. Moreover, Linares and Romero (2002) demonstrated how to fit both WGP and extended WGP models. Consequently, our approach is a suitable alternative for breeders to estimate desired gains and thus to define a breeding objective.

2.4.3 Breeding traits and preference values for rainbow trout

The 6 most important traits selected out of the thirteen offered were 3 production traits: TGC, Surv, and FCR, and 3 quality/processing traits: CF, FIL%, and LMat. In general, TGC and Surv are economically very important because they are the main criteria for a farmer to get return. Moreover, feed is one of the major costs of fish production. Thus improved FCR is important to reduce cost and gain more farm profit. Also FIL% is economically important, especially for processors because the fillet is the edible portion of the fish. Late maturation is preferred because early maturation during grow-out period spoils flesh quality. In addition, mature males are aggressive against other fish resulting in wounds, fin erosion and poor appearance. Condition factor reflects fish shape which is important for both producers and processors. High CF is preferred by processors willing to have wide fillets, but excessively high CF is disliked because of the non-natural trout

appearance (Kause et al. 2003a) or because it may reflect an underlying skeletal deformation (Sullivan et al. 2007). Hence the breeding objective for rainbow trout needs to cover both production and quality/processing traits.

Flesh color, skin color and spottiness, high and low water temperature tolerance, and deformity were not selected. Farmers may consider these traits as less economically important, or only a minority of farmers preferred them. Moreover, some of these traits overlap with the 6 most important traits. For example, high and low water temperature tolerance may negatively affect production traits, e.g. growth and survival. Selecting for growth and survival in cold (or warm) water most likely increases tolerance to cold (or warm) water. Assuming that there is only weak genotype-by-environment interaction for growth or survival with rearing temperature (as shown by Fishback et al. (2002) for growth), selecting for these traits at low water temperature may lead to improvement of fish performance at high water temperature, and vice versa. Flesh color is in fact economically important for farmers because feed with color additives is costly and consumers are willing to pay for redder fillet (Alfnes et al., 2006). However, the importance of fillet color is non-existent or small when producing fry and portion-sized fish. In our study, only a minor part of the responders were producers of large trout. This may explain the low preference for flesh color here.

Uniformity and disease resistance were removed from the original list of 6 most important traits. We wanted to quantify the importance of uniformity relative to other traits in the questionnaire A. However, it was removed from questionnaire B because there are no genetic parameters for uniformity yet. A desire for increased resistance against over 25 diseases was requested by the responders of the questionnaire A (data not shown). Including them all into breeding objective is practically impossible. Even though disease resistance had high ranking score, overall survival until harvest was alternatively chosen. Selection for overall survival improves resistance against multiple mortality factors (Vehviläinen et al., 2008).

Our results showed that variation of preferences for traits occurs with respect to different commercial products and farming environments. For different commercial products, fry producers focus on growth as the first priority, followed by FCR and survival. Producers of pan-sized fish consider Surv and F % more important than fry producers. FIL% is not important at fry stage because under commercial production, FIL% is expressed at a later age. Late maturation becomes more preferable trait when farmers produce large-sized trout. This is logical because of the negative impact of maturation on flesh quality.

Recirculating aquaculture system (RAS) is commonly used in the areas where water supply is limited or environmental load need to be reduced. The water circulation used by farmers influenced preference values for growth, Surv, and FIL%. High fish densities and feeding rates as well as high CO_2 content and low water exchange rate in RAS reduce fish health, increase mortality, and reduced growth (Danley et al., 2005; Good et al., 2009). Consequently, growth and survival are more preferable to be improved by RAS users. Rainbow trout farming can be practiced at high altitudes. Atmospheric partial pressure of oxygen decreases when altitude is high and results in low oxygen solubility. Low dissolved oxygen affects growth and ultimately survival. Therefore, fish farmers at higher altitude may have higher preference for fish survival.

Late maturation is more preferred by farmers with low water temperature compared to farmers with high temperature. Photoperiod is the major factor determining age at maturation while low temperature leads to slow growth rate. Together this could result in matured fish before they reach harvest size (Davies and Bromage 2002). When the water temperature is high, production traits, especially growth, become much more important. As expected, farmers that rear rainbow trout at extreme water temperatures consider survival as the most important trait.

In this study it was not possible to distinguish each social group category *a priori*, as we did not know which categories were relevant and how farmers were distributed over these categories. This is also shown by the numbers of unassigned observations. For these, it was not possible to determine with certainty to which category they should be assigned. In fact, each observation is a complex combination of commercial products and farming environments, e.g. pan-sized fish produced at high altitude using RAS with low water temperature. Consequently, the trends observed here for Soc-P values should be considered as hypotheses for further testing.

The differences in Soc-P values within the categories “water circulation” and “altitude” are smaller than those from “commercial products” and “water temperature”. It indicates that consensus is more important for fish farmers who have different product objectives or water temperature. Con-P values from commercial products were very different from Con-P values from water temperature, especially for late maturation. For commercial products conflicting interests were seen between fry producers and large fish producers. For water temperature, there were three groups with conflicting interests: low, high, and low and high temperature farmers. In these situations, extended WGP is a valuable tool to try and minimize the disagreement between groups. Extended WGP takes both

minority opinion and overall disagreement into account. This results in a desired genetic gain beneficial to most of customers.

Average Con-P values from random groups can be seen as a compromise between Con-P values for both categories, and account for all different commercial products and farming environments. We conclude that random social groups can be practically used in situations where stakeholder groups cannot be explicitly identified.

2.4.4 Desired genetic gains for rainbow trout

Desired genetic gains derived in this study can be used to estimate selection indexes and weighting factors of the 6 most important traits of the breeding objective when the trait means and phenotypic and genotypic (co)variances are known. Brascamp (1984) developed formulas to obtain selection index weights producing desired genetic gains. Alternatively, available software, e.g. Sel*A*ction (Rutten et al., 2002) can be used to calculate weighting factors on an iterative way (Gizaw et al., 2009). Estimates of genetic parameters for a number of traits, including the 6 most important traits found here, are available in literature for rainbow trout (e.g. Kause et al., 2003ab, 2005, 2006, 2007; Vehviläinen et al., 2008, 2010). Yet, complete information on phenotypic and genotypic correlations for all 6 important traits are still not available.

Based on the current literature, development of breeding objectives for aquaculture breeding programs is at its infancy. There are only 2 previous studies on economic values for aquaculture species (Henryon et al., 1999; Ponzoni et al., 2007). There are no previous participatory studies to define a breeding objective in aquaculture species.

Fish breeding programs that serve global markets and take farmer preferences and local conditions into account contribute to increasing the profitability and sustainability of global fish farming. Most trout farms are typically located in rural areas and they can contribute to local income, nature conservation, and rural development through specialized production methods. Including farmer preferences in the breeding strategy will therefore indirectly improve livelihood.

In conclusion, variation in preference values of breeding traits exists. Estimation of consensus preference values and percentage of genetic improvement can be used to derive desired genetic gains for a rainbow trout breeding objective. Our participatory approach can be used to define a breeding objective for any breeding program which serves large and potentially diverse markets.

2.5 Acknowledgements

This study is a part of the PhD project: Optimization of rainbow trout breeding program for multi-trait selection with GxE interaction. This project is funded by Ananda Mahidol Foundation (The King Scholarship, Thailand) and Troutlodge, Inc., USA. We thank translators from Troutlodge, Inc. for multilingual questionnaires in this study.

References

Alfnes, F., Guttormsen, A. G., Steine, G., Kolstad, K., 2006. Consumer's willingness to pay for the color of salmon: A choice experiment with real economic incentives. Amer. J. Agric. Econ. 88: 1050-1061.

Amer, P. R., 2006. Approaches to formulating breeding objectives. 8th World Congr. on Genet. Appl. Livest. Prod., Belo Horizonte, MG, Brazil: PP31-01.

Brascamp, E. W., 1984. Selection indices with constraints. Anim. Breed. Abstr. 52:645-654.

Cao, D., Leung, L. C., Law, J. S., 2008. Modifying inconsistent comparison matrix in analytic hierarchy process: A heuristic approach. Decis. Support Syst. 44:944-953.

Danley, M. L., Kenney, P. B., Mazik, P. M., Kiser, R., Hankins, J. A., 2005. Effects of carbon dioxide exposure on intensively cultured rainbow trout, *Oncorhynchus mykiss*: Physiological responses and fillet attributes. J. World Aquacult. Soc. 36:249-261.

Davies, B., Bromage, N., 2002. The effects of fluctuating seasonal and constant water temperatures on the photoperiodic advancement of reproduction in female rainbow trout, *Oncorhynchus mykiss*. Aquaculture 205:183-200.

Dumas, A., France, J., Bureau, D. P., 2007. Evidence of three growth stanzas in rainbow trout (*Oncorhynchus mykiss*) across life stages and adaptation of the thermal-unit growth coefficient. Aquaculture 267:139-146.

Elvingson, P., Johansson, K., 1993. Genetic and environmental components of variation in body traits of rainbow-trout (*Oncorhynchus mykiss*) in relation to age. Aquaculture 118:191-204.

Falconer, D. S., Mackay, T. F. C., 1996. Introduction to quantitative genetics. 4th ed. Longman Group Ltd., Essex, UK.

FAO, 2011. Cultured aquatic species information programme *Oncorhynchus mykiss* (Walbaum, 1792): Fisheries and aquaculture department. Accessed May 18, 2011. http://www.fao.org/fishery/culturedspecies/Oncorhynchus_mykiss/en.

Fishback, A. G., Danzmann, R. G., Ferguson, M. M., Gibson, J. P., 2002. Estimates of genetic parameters and genotype by environment interactions for growth traits

of rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. Aquaculture 206:137-150.

Gizaw, S., Komen H., van Arendonk, J. A. M., 2009. Participatory definition of breeding objectives and selection indexes for sheep breeding in traditional systems. Livest. Sci. 128:67-74.

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

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

González-Pachón, J., Romero, C., 1999. Distance-based consensus methods: A goal programming approach. Omega 27:341-347.

Good, C., Davidson, J., Welsh, C., Brazil, B., Snekvik, K., Summerfelt, S., 2009. The impact of water exchange rate on the health and performance of rainbow trout *Oncorhynchus mykiss* in water recirculation aquaculture systems. Aquaculture 294:80-85.

Henryon, M., Purvis, I. W., Berg, P., 1999. Definition of a breeding objective for commercial production of the freshwater crayfish, marron (*Cherax tenuimanus*). Aquaculture 173:179-195.

Kanis, E., De Greef, K. H., Hiemstra, A., van Arendonk, J. A. M., 2005. Breeding for societally important traits in pigs. J. Anim Sci. 83:948-957.

Kause, A., Ritola O., Paananen, T., Mantysaari, E., Eskelinen, U., 2002. Coupling body weight and its composition: A quantitative genetic analysis in rainbow trout. Aquaculture 211:65-79.

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

Kause, A., Ritola, O., Paananen, T., Mantysaari, E., Eskelinen, U., 2003b. Selection against early maturity in large rainbow trout *Oncorhynchus mykiss*: The quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228:53-68.

Kause, A., Ritola, O., Paananen, T., Wahlroos, H., Mantysaari, E. A., 2005. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). Aquaculture 247:177-187.

Kause, A., Tobin, D., Houlihan, D. F., Martin, S. A. M., Mantysaari, E. A., Ritola, O., Ruohonen, K., 2006. Feed efficiency of rainbow trout can be improved through

selection: Different genetic potential on alternative diets. J. Anim. Sci. 84:807-817.

Kause, A., Paananen, T., Ritola, O., Koskinen, H., 2007. Direct and indirect selection of visceral lipid weight, fillet weight, and fillet percentage in a rainbow trout breeding program. J. Anim. Sci. 85:3218-3227.

Leung, L. C., Cao, D., 2000. On consistency and ranking of alternatives in fuzzy AHP. Eur. J. Oper Res. 124:102-113.

Linares, P., Romero, C., 2002. Aggregation of preferences in an environmental economics context: A goal-programming approach. Omega 30:89-95.

LINDO Systems Inc., 1999. LINGO User's Guide, LINDO Systems Inc., Chicago, USA.

Lynch, M., and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc., Massachusetts, USA.

McKay, L. R., Ihssen, P. E., Friars, G. W., 1986. Genetic parameters of growth in rainbow trout, *Salmo gairdneri*, as a function of age and maturity. Aquaculture 58:241-254.

Monsen, B. B., Ødegård, J., Arnesen, K. R., Toften, H., Nielsen, H. M., Damsgård, B., Bijma, P., Olesen, I., 2010. Genetics of social interactions in Atlantic cod (*Gadus morhua*). 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany:100.

Murphy, C. K., 1993. Limits on the analytic hierarchy process from its consistency index. Eur. J. Oper. Res. 65:138-139.

Neira, R., Lhorente, J. P., Araneda, C., Diaz, N., Bustos, E., Alert, A., 2004. Studies on carcass quality traits in two populations of Coho salmon (*Oncorhynchus kisutch*): Phenotypic and genetic parameters. Aquaculture 241:117-131.

Olesen, I., Gjerde, B., Groen, A. F., 1999. Methodology for deriving non-market trait values in animal breeding goals for sustainable production systems. Interbull Bull. 23:13-21.

Ponzoni, R. W., Nguyen, N. H., Khaw, H. L., 2007. Investment appraisal of genetic improvement programs in Nile tilapia (*Oreochromis niloticus*). Aquaculture 269:187-199.

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

Rutten, M. J. M., Bijma, P., Woolliams, J. A., van Arendonk, J. A. M., 2002. Sel*A*ction: Software to predict selection response and rate of inbreeding in livestock breeding programs. J. Hered. 93:456-458.

Rye, M., Refstie, T., 1995. Phenotypic and genetic parameters of body size traits in Atlantic salmon *Salmo salar* L. Aquac. Res. 26:875-885.

Saaty, R. W., 2003. Decision making in complex environments; The analytical hierarchy process (AHP) for decision making and the analytic network process

(ANP) for decision making with dependence and feedback. Super decisions. Creative decisions foundation in Pittsburgh, Pennsylvania, USA.

Saaty, T. L., 1980. The analytic hierarchy process: Planning, priority setting, resource allocation. McGraw-Hill, New York, USA.

Saaty, T. L., 1990. How to make a decision: The analytic hierarchy process. Eur. J. Oper. Res. 48:9-26.

Silverstein, J. T., Vallejo, R. L., Palti, Y., Leeds, T. D., Rexroad III, C. E., Welch, T. J., Wiens, G. D., Ducrocq, V., 2009. Rainbow trout resistance to bacterial cold-water disease is moderately heritable and is not adversely correlated with growth. J. Anim. Sci. 87:860–867.

Sullivan, M., Hammond, G., Roberts, R. J., Manchester, N. J., 2007. Spinal deformation in commercially cultured, *Salmo salar* L.: a clinical and radiological study. J. Fish Dis. 30:745-752.

Tano, K., Kamuanga, M., Faminow, M. D., Swallow, B., 2003. Using conjoint analysis to estimate farmer's preferences for cattle traits in West Africa. Ecol. Econ. 45:393-407.

Thurstone, L. L., 1927. A law of comparative judgement. Psychol. Rev. 34:273-286.

Vehviläinen, H., A. Kause, C. Quinton, H. Koskinen, and T. Paananen. 2008. Survival of the currently fittest: Genetics of rainbow trout survival across time and space. Genetics 180:507-516.

Vehviläinen, H., Kause, A., Kettunen-Praebel, A., Koskinen, H., Paananen, T., 2010. Comparison of direct and indirect selection for rainbow trout sea grow-out survival. 9th World Congr. Genet. Appl. Livest. Prod., Leipzig, Germany:PP4-124.

Wang, T. C., Chen, Y. H., 2008. Applying fuzzy linguistic preference relations to the improvement of consistency of fuzzy AHP. Inform. Sci. 178:3755-3765.

Wild, V., Simianer H., Gjoen, H. M., Gjerde, B., 1994. Genetic-parameters and genotype x environment interaction for early sexual maturity in Atlantic salmon (*Salmo-salar*). Aquaculture 128:51-65.

Wurzinger, M., Ndumu, D. B., Drucker, A. G., Okeyo, A. M., Semambo, D. K., Solkner, J., 2006. Assessing stated preferences through the use of choice experiments: Valuing (re) production versus aesthetics in the breeding goals of Uganda Ankole cattle breeders. 8th World Congr. Genet. Appl. Livest. Prod., Belo Horizonte, MG, Brazil:PP31-09.

Zeshui, X., Cuiping, W., 1999. A consistency improving method in the analytic hierarchy process. Eur. J. Oper. Res. 116:443-449.

3

Bias and precision of estimates of genotype-by-environment interaction: A simulation study

P. Sae-Lim[1], H. Komen[1], A. Kause[1]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands

Aquaculture (2010) 310:66-73

Abstract

Re-ranking of genotypes across environments is a form of genotype-by-environment (GxE) interaction with serious consequences for breeding programmes. The degree of such GxE interaction can be estimated using the genetic correlation (r_g) between measurements in two environments for a given trait. When r_g is lower than 0.8, GxE interaction is commonly considered to be biologically significant. Here a stochastic simulation was used to study the impact of population structure on bias and precision of genetic correlation estimates between two environments. Simulated populations resulted from a nested mating design (1 sire to 2 dams). Simulated r_g was 0.0, 0.5, or 0.8. A trait with heritability (h^2) of either 0.3 or 0.1 in both environments was simulated. Simulation results show that genetic correlation estimates are biased downwards especially when the simulated r_g is 0.8, heritability is 0.1, and family size is less than 10. A downward biased genetic correlation estimate incorrectly suggests the existence of GxE interaction. This can lead to the erroneous conclusion that a multi-environment breeding programme is needed. The optimal design with the lowest mean square error for r_g for a trait with low h^2 requires a large family size (20-25) and a low number of families (100-80 or 50-40 for population size fixed to 2000 and 1000 animals, respectively). For traits with moderate h^2, the optimal family size is 10 with 200 or 100 families for population size fixed to 2000 and 1000, respectively. We also studied the effect of selective mortality on GxE estimates. However, schemes with unequal family sizes due to differences between families in survival produced similar results for the optimum design as schemes with equal family sizes. Equal-family-size design can thus be used to determine the optimal design for estimating GxE interaction. Our study can be used as a guideline for estimating a genetic correlation for practical breeding programmes.

Key words: breeding programme, genetic correlation, genotype by environment interaction, optimal design, population structure, simulation

3.1 Introduction

Many breeding programmes distribute animal material across diverse production environments, sometimes even at a global scale. Selection within a nucleus broodstock may lead to lower-than-expected genetic gains in other production environments when genotype-by-environment (GxE) interaction exists but it is not introduced in the selection criteria.

GxE interaction is defined as a phenomenon that genotypes respond differently to an environment gradient (Falconer and Mackay, 1996). There are two main types of GxE interaction: scaling effects and re-ranking. A scaling effect means that the amount of genetic variation in two environments differs. Re-ranking means that ranking of genotypes changes across different environments (Lynch and Walsh, 1998). Re-ranking in particular is a challenge for breeding because genotypes in one environment are not necessarily the best ones in other environments. Re-ranking across environments can be estimated using a genetic correlation between measurements in two environments for a given trait (Falconer, 1952). GxE interaction is commonly considered to be biologically significant when genetic correlation is lower than 0.8 (Robertson, 1959b).

In aquaculture, a number of studies on GxE interaction have been conducted under diverse management practices. The published studies on genetic correlations between environments have used family sizes and family numbers ranging from tens to several hundred (e.g. Sylvén et al. 1991; Fishback et al., 2002; Kause et al., 2003, 2004; Saillant et al., 2006; Quinton et al., 2007; Dupont-Nivet et al., 2008; Pierce et al., 2008; Vehviläinen et al., 2008; Khaw et al., 2009).

To accurately estimate a genetic correlation between environments, an optimal design needs to be established; an experimental design which produces a precise and unbiased result while using minimum testing capacity. Enlarging population size typically increases the power of a design but simultaneously increases costs. In contrast, too small population size or suboptimal population structure (number of families, family size, and mating design) may potentially result in biased and inaccurate estimates. Furthermore, differences in family size caused by differential survival or differences in parental contributions to the whole population size will result in unequal family sizes. The resulting population structure is unbalanced which may influence the bias and precision of genetic correlation estimates.

To our knowledge, no study has been conducted to assess bias and precision of estimates of GxE interaction. The present study describes the use of a stochastic simulation to construct an optimal population structure promoting precise and unbiased estimation of a genetic correlation between environments. The

simulations employed here are divided into three scenarios. Firstly, various population sizes were simulated. Secondly, varying combinations of family size and family number were used to find an optimal population structure under a fixed population size. Thirdly, in practice, an experimental design is unintentionally challenged with between-family variation in survival leading to an unbalanced design. Unequal family sizes may result in larger sampling variance compared to equal family sizes (Hammersley, 1949; Tallis, 1959). Therefore this scenario was used to study the influence of unequal family sizes on the bias and precision of the estimation.

3.2 Materials and methods

In the simulation, three different population structures were constructed, and (co)variance components were estimated.

3.2.1. Population construction

The simulated population structure was a split-family design with two environments, where the offspring generation had trait records and their parents only contributed to the pedigree. In each environment, phenotype of an individual was calculated as $y = 0.5a_s + 0.5a_d + m + e$, where a_s and a_d are additive genetic values of sire and dam, respectively, m is Mendelian sampling term, and e is environmental effect. Additive genetic values were sampled from a bivariate normal distribution of environments A and B: $Var(a) = \begin{bmatrix} \sigma^2_{a,A} & \sigma_{a,AB} \\ \sigma_{a,AB} & \sigma^2_{a,B} \end{bmatrix}$, Mendelian sampling terms from $Var(m) = \begin{bmatrix} 1/2\,\sigma^2_{a,A} & 1/2\,\sigma_{a,AB} \\ 1/2\,\sigma_{a,AB} & 1/2\,\sigma^2_{a,B} \end{bmatrix}$, and environmental effects from $Var(e) = \begin{bmatrix} \sigma^2_{e,A} & 0 \\ 0 & \sigma^2_{e,B} \end{bmatrix}$. Each of these effects had a mean of zero.

Phenotypic variance (σ^2_P) was set to 1. Additive genetic variance (σ^2_a) was calculated as $\sigma^2_P h^2$ and environmental variance (σ^2_e) was calculated as $\sigma^2_P(1 - h^2)$. Genetic covariance between measurements of a trait in two environments ($\sigma_{a,AB}$) determined the degree of family re-ranking, and was sampled from a simulated value (described below). No environmental covariance was simulated between the

two environments because each animal inhabited only one environment. The population construction was done in *R* (R Development Core Team, 2008).

3.2.2 Simulated scenarios

A population was simulated with a genetic correlation (r_g) of 0.8, 0.5, or 0.0 between environments. A value of 0.8 is often considered a threshold value for GxE interaction to be significant for a breeding programme (Robertson, 1959b), whereas genetic correlations of 0.5 and 0.0 mean that strong re-ranking occurs. A trait with heritability of either 0.3 or 0.1 in both environments was used in all scenarios. Number of sires, dams, and offspring were constructed following three population design scenarios. For all scenarios, the mating design was one sire mated to two different dams (paternal nested design). The paternal nested mating designs are used e.g. in GIFT and Troutlodge breeding programmes.

Varied population size (scenario A)

Family size is one important factor that determines the amount of bias, standard error and mean square error. Therefore this scenario was to evaluate the impact of family size on precisions and bias. The simulated population had a fixed family number of 100 but family size ranged from 3 to 75 within each environment. Note that with the increase of family size, also the population size increases, e.g. family size of 3 x 100 families = 300. The range of family sizes is given in Table 3.1.

Fixed population size (scenario B)

An experiment typically has a limit for the maximum number of fish reared, tagged or genotyped. The results from scenario A showed that estimates of r_g were unbiased for traits with both low (0.1) and moderate (0.3) heritabilities when population size was larger than 2000 (100 families x 20 individuals). Therefore, the starting point for this simulation was a fixed population size of 2000 in both environments. In this scenario, both family size and family number were varied. Given a fixed population size, this means that increasing family size results in decreasing family number. The results from population size of 2000 were compared to the bias and precision of the r_g estimates when simulating a fixed population size of 1000, i.e. when the number of animals was decreased to 50%. Table 3.1 summarizes the used family sizes and number of families for population sizes of 2000 and 1000.

Unequal family size (scenario C)

In this scenario, the effect of unequal family size on the bias and precision of the estimate of r_g was studied. The initial population size was 2000 and survival was 50%, meaning that the population size at harvest trait recording was reduced to 1000. To generate differences between families in size, each individual was assigned a trait record for survival (0 = alive, 1 = died). Survival was not correlated with the traits recorded in two environments, and was not analysed as a correlated trait in the genetic analyses. Survival was modelled as a binary threshold trait with the following underlying liability scale phenotypic and genetic parameters. Phenotypic variance for survival was assumed to be one, and thus additive genetic variance is equal to heritability for survival (h^2_{surv}). To generate different degrees of between-family variation in survival for the population construction, three alternative sets of parameters were used for survival: h^2_{surv} = 0.00 and c^2_{surv} = 0.3; h^2_{surv} = 0.15 and c^2_{surv} = 0.1; h^2_{surv} = 0.30 and c^2_{surv} = 0.0, where c^2_{surv} is the ratio of variance for common environment of full-sibs to phenotypic variance. These represent realistic estimates for rainbow trout (Kanis et al., 1976; Vehviläinen et al., 2008, 2010).

Scenario C was performed for a trait with h^2 of 0.1 and 0.3 with a simulated genetic correlation of 0.8 between two environments. Results from scenario A show that this is the most difficult scenario to estimate genetic correlation correctly.

3.2.3 Estimation of (co)variance components

The simulated data were analysed using a bivariate animal model in which the same trait in two environments was treated as two different traits. The model fitted was:

$$y_{ij} = \mu_i + a_{ij} + e_{ij} \qquad [1]$$

where y_{ij} represents a trait measured in one of two environments (i =1, 2) for an individual j (j = number of individuals); μ_i is the overall mean of the trait i; a_{ij} is the random additive genetic effect of individual j; and e_{ij} is the random residual effect. Due to only one observation for each individual, residual covariance was fixed to zero. Estimated genetic correlation between two environments ($\hat{r}_g$), its standard error, and heritabilities with their standard errors were estimated using restricted maximum likelihood (REML) in ASReml software (Gilmour et al., 2006). The (co)variance matrix was constrained to be positive definite.

Table 3.1 Simulated population structures in three scenarios: varied population size, fixed population size (1000 and 2000), and fixed population size with unequal family sizes.

Scenario	Population structure													
Varied population size	Family size	3	5	10	15	20	25	30	40	50	60	75		
	Family no.	100	100	100	100	100	100	100	100	100	100	100		
	Population size	300	500	1000	1500	2000	2500	3000	4000	5000	6000	7500		
Fixed population size of 1000	Family size	2	3	4	5	6	10	20	25	42	50	63	83	100	125
	Family no.	500	334	250	200	166	100	50	40	24	20	16	12	10	8
	Population size	1000	1002	1000	1000	996	1000	1000	1000	1008	1000	1008	996	1000	1000
Fixed population size of 2000	Family size	2	3	4	5	6	10	20	27	40	50	59	77	100	125
	Family no.	1000	668	500	400	334	200	100	74	50	40	34	26	20	16
	Population size	2000	2004	2000	2000	2004	2000	2000	1998	2000	2000	2006	2002	2000	2000
Fixed population size of 1000; Unequal family size	Average family size	2	3	4	5	6	10	20	25	50	100	125			
	Family no.	500	334	250	200	166	100	50	40	20	10	8			
	Population size	1000	1002	1000	1000	996	1000	1000	1000	1000	1000	1000			

3.2.4 Summarising output from the simulation

Each population structure alternative was simulated 500 times. For each alternative, the results were summarized using: (*i*) average, median and mode of estimated genetic correlations ($\hat{r}_g$) from all replicates, (*ii*) standard deviation (SD) of the distribution of replicates of $\hat{r}_g$, (*iii*) average of the estimated standard errors (SE) of $\hat{r}_g$ given by ASReml, and (*iv*) average mean square error (MSE) of the estimated genetic correlations. Runs that did not converge because of a lack of genetic variation in the sample data were removed from the analysis.

3.3 Results

3.3.1 Varied population size

When population size was allowed to vary, the designs with small family size and non-zero simulated r_g resulted in downward biased genetic correlation estimates (Figure 3.1b-c, top panel). For h^2 of 0.1, the estimated genetic correlations were biased when family size was 3-10 for r_g of 0.5, or 3-20 for r_g of 0.8. For h^2 of 0.3, only a slight downward bias was observed with family size of 3-10 for r_g of 0.8. The downward bias was also observed for the median and mode of $\hat{r}_g$ values. For instance, for r_g of 0.5, h^2 of 0.1 and family size of 3, $\hat{r}_g \pm$ SD was 0.315 ± 0.44 (mode = 0.525, median = 0.416). For r_g of 0.8, h^2 of 0.1 and family size of 3, $\hat{r}_g \pm$ SD was 0.482 ± 0.40 (mode = 0.675, median = 0.584) (Figure 3.1). In contrast, the mean genetic correlation estimate was unbiased when r_g was 0.0 (Figure 3.1a, top panel).

These results indicate that genetic correlation estimates become more biased with lower h^2 (0.1) and higher simulated r_g (0.8). Consequently, more animals are needed to obtain reliable estimates when the simulated genetic correlation is higher (Figure 3.1a top vs. c top).

Standard errors of the estimates were drastically reduced when family size was increased to 10-15 for h^2 = 0.3, and to 20-25 for h^2 = 0.1 (Figure 3.1a-c, middle panel). Similarly, MSEs were strongly reduced when family size was increased to 10 and 15-20 for h^2 = 0.3 and h^2 = 0.1, respectively (Figure 3.1a-c, bottom panel).

3.3.2 Population size fixed to 2000

For r_g of 0.5, mean estimates of r_g were weakly downward biased when family size was either 2-10 or 77-125, corresponding to family number of 1000-200 or 26-16, respectively (Figure 3.2b, top panel).

For r_g of 0.8, biased $\hat{r}_g$ was obtained when family size was either 2-10 or 40-125, corresponding to family number of 1000-200 or 50-16, respectively (Figure 3.2c, top panel). Yet for both r_g of 0.5 and 0.8, there was no bias at low family size when h^2 was 0.3. Moreover, when the mean $\hat{r}_g$ was biased at high family sizes (i.e. low family numbers, the mode $\hat{r}_g$ was always close to the simulated value (either 0.5 or 0.8). For instance, when family size was 77 (family number of 26), h^2 was 0.1 and r_g was 0.8, $\hat{r}_g \pm$ SD was 0.631 ± 0.29 but mode was 0.800 and median was 0.750. This indicates that the bias at high family sizes was not that strong, but that the distribution became skewed to the left.

3.3.3 Population size fixed to 1000

For r_g of 0.5, estimates of genetic correlation were downward biased when family size was either 2-10 or 50-125, corresponding to family number of 500-100 or 20-8, respectively (Figure 3.3b, top panel). For r_g of 0.8, biased $\hat{r}_g$ were obtained when family size was either 2-20 or 25-125, corresponding to family number of 500-50 or 40-8, respectively (Figure 3.3c, top panel). Yet for both r_g of 0.5 and 0.8, there was no bias at low family size when h^2 was 0.3. Similar to population size of 2000, only with very low family sizes, mean, mode and median of the estimates were all biased downward, whereas at high family sizes mainly mean and median were biased. For example, with family size of 4 (family number = 250), simulated h^2 of 0.1 and r_g of 0.8, estimated genetic correlation was 0.644 ± 0.42 (mode = 0.650, median = 0.681). In contrast, family size of 100 (family number = 10), simulated h^2 of 0.1 and r_g of 0.8, estimated genetic correlation was 0.533 ± 0.34 (mode = 0.720, median =0.651).

It is noteworthy that for population size fixed to 1000 with simulated genetic correlation of 0.8 and h^2 of 0.1, the mean of estimated genetic correlation never reached the simulated genetic correlation of 0.8. The reason is that the distribution of correlation estimates is pushed against the limit of unity, preventing normal distribution to occur. Instead, the proportion of correlations that should have been above unity was in fact between 0.8 and 1.0. This is indicated by the fact that

although the mean was lower than the simulated 0.8, at family sizes of 20-25 both the median and mode of the estimated correlations were 0.801 and 0.840.
The results for the optimal designs, i.e. designs with the lowest MSE, are shown in Table 3.2. Also the designs with the lowest SEs are listed.

Table 3.2 Ranges of family sizes[a] and family numbers that produce the lowest mean square error (MSE) and standard error (SE) for the estimated genetic correlation (optimal designs).

			Lowest MSE			Lowest SE		
Population size	r_g	h^2	Family size	Family number	MSE	Family size	Family number	SE
2000		0.1	20-27	100-74	0.045	20-27	100-74	0.207
	0	0.3	4-6	500-334	0.012	4-6	500-334	0.111
		0.1	20-40	100-50	0.036	20-40	100-50	0.185
	0.5	0.3	6-10	334-200	0.009	6-10	334-200	0.098
		0.1	20-27	100-74	0.015	40-59	50-34	0.137
	0.8	0.3	10-40	200-50	0.005	10-40	200-50	0.074
1000		0.1	20	50	0.091	20-25	50-40	0.304
	0	0.3	3-10	334-100	0.027	4-6	250-166	0.159
		0.1	10-63	100-16	0.078	20-42	50-42	0.28
	0.5	0.3	6-10	166-100	0.018	5-10	500-100	0.144
		0.1	20-25	50-40	0.024	20-63	50-16	0.242
	0.8	0.3	5-25	200-40	0.01	10-25	100-40	0.104

Population sizes were fixed to 1000 or 2000; simulated genetic correlations (r_g) were 0.0, 0.5 and 0.8, and simulated trait heritabilities (h^2) were 0.1 and 0.3.
[a] All families are of equal size.

For both population sizes, the lowest SE from h^2 of 0.1 was roughly two times higher than the lowest SE from h^2 of 0.3. The designs providing the lowest MSE and SE tended to be different only when r_g was high and h^2 was low. When comparing the same simulated h^2 and r_g, the lowest MSE from population size of 1000 was approximately two times higher than MSE from population size of 2000.

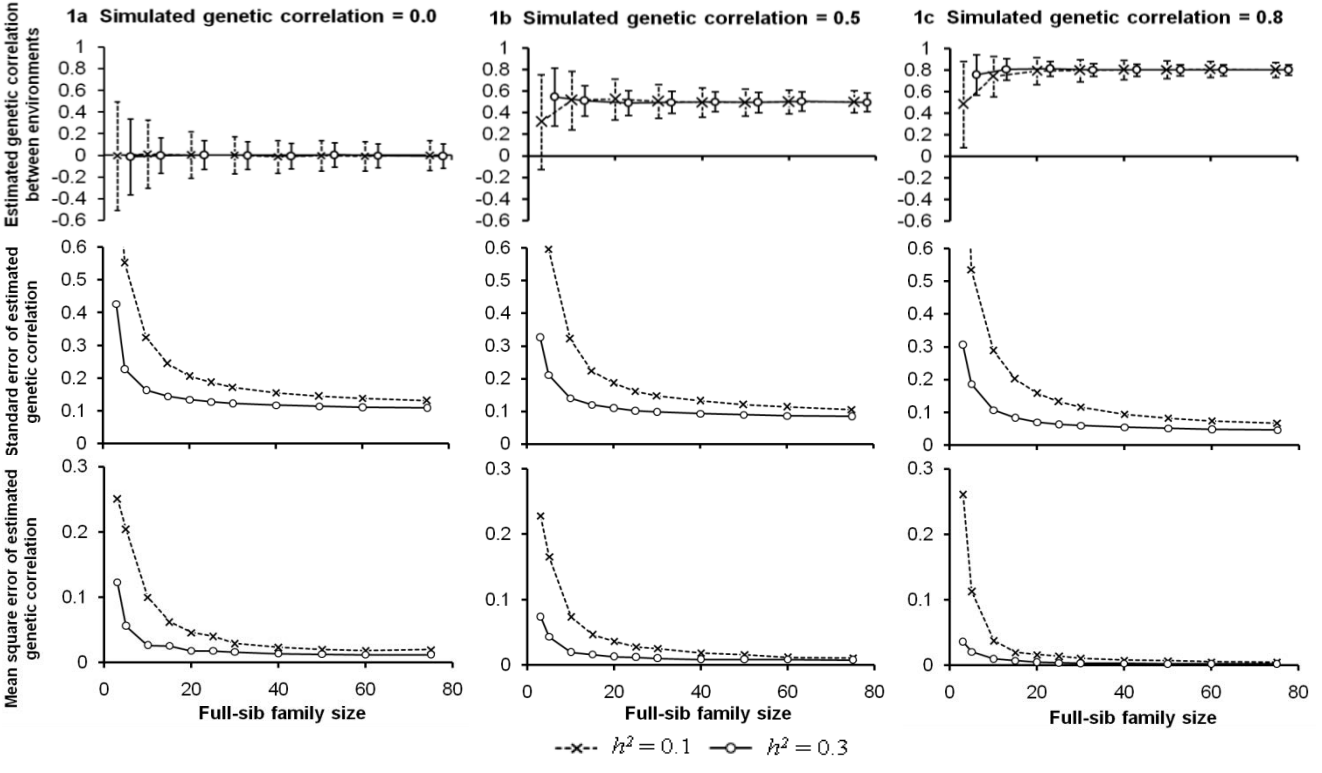

Figure 3.1 Average estimated genetic correlation ($\hat{r}_g$) between two environments, average standard error, and average mean square error under varied family size scenario. In the $\hat{r}_g$ graphs, symbols for h^2 of 0.3 have been moved 3 points forward on the x-axis to distinguish the lines of h^2 = 0.1 and 0.3. Some family size points were removed from the graphs to avoid excessively dense data points. The error bar of the top panel is standard deviation of estimated genetic correlations from all simulated replicates.

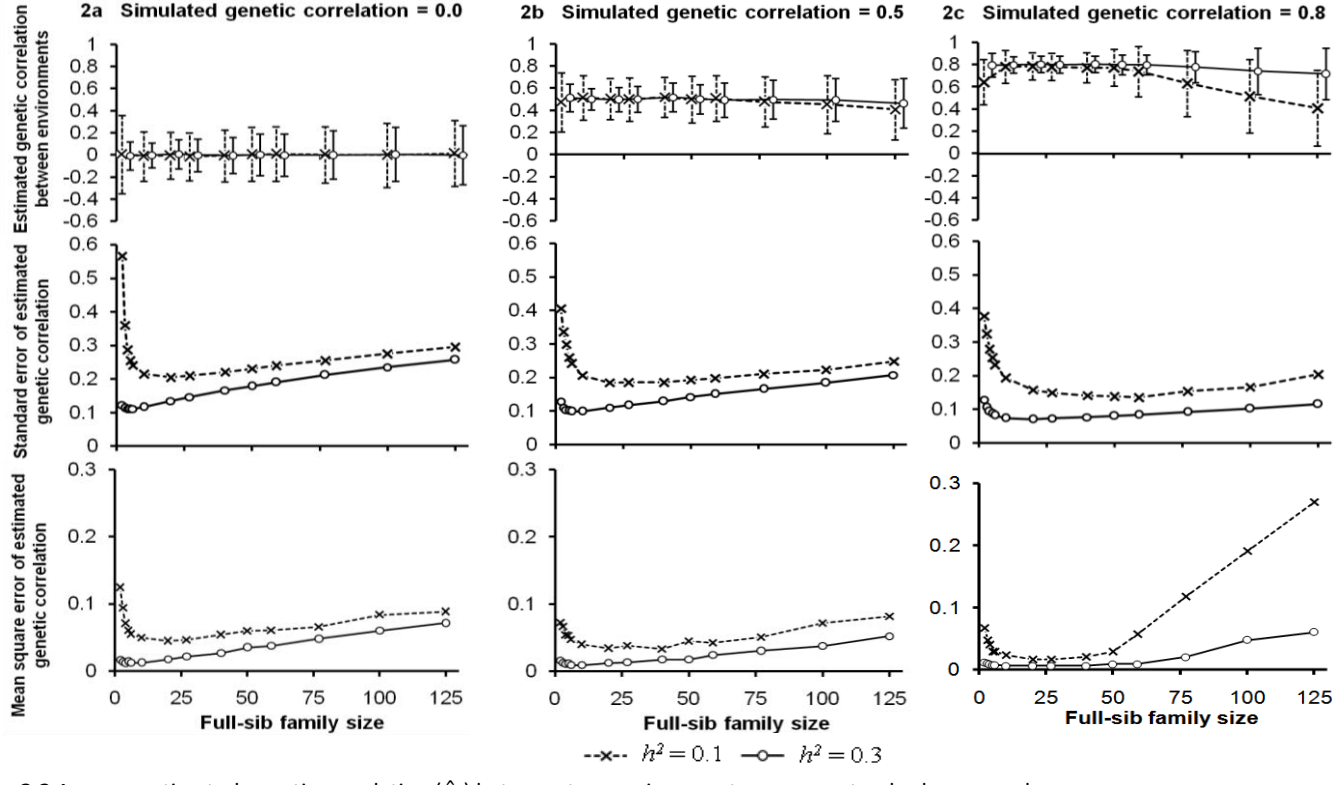

Figure 3.2 Average estimated genetic correlation ($\hat{r}_g$) between two environments, average standard error, and average mean square error under population size fixed to 2000 scenario. In the$\hat{r}_g$ graphs, symbols h^2 of 0.3 have been moved 3 points forward on the x-axis to distinguish the line of h^2 of 0.1 and 0.3. Some family size points were removed from the graphs to avoid excessively dense data points. The error bar of the top panel is standard deviation of estimated genetic correlations from all simulated replicates.

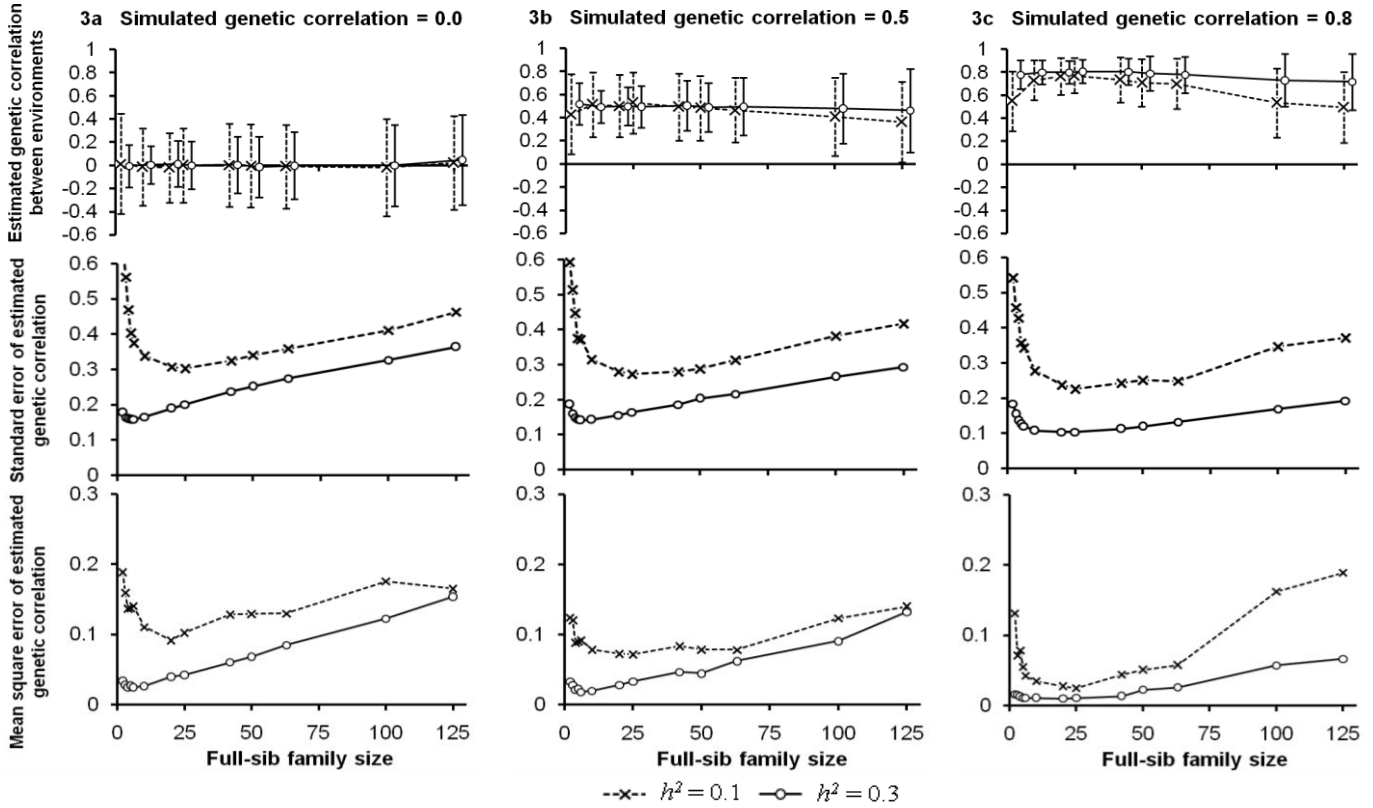

Figure 3.3 Average estimated genetic correlation ($\hat{r}_g$) between two environments, average standard error, and average mean square error under population size fixed to 1000 scenario. In the $\hat{r}_g$ graphs, symbols h^2 of 0.3 have been moved 3 points forward on the x-axis to distinguish the line of h^2 of 0.1 and 0.3. Some family size points were removed from the graphs to avoid excessively dense data points. The error bar of the top panel is the standard deviation of estimated genetic correlations from all simulated replicates.

3.3.4 Unequal family size

Figure 3.4 shows the comparison between equal and unequal family sizes for a population size of 1000. All combinations of parameter values for survival (h^2_{surv} and c^2_{surv}) for generating unequal family sizes gave similar results, especially for the design with the lowest SD, SE and bias (Figure 3.4a-b).

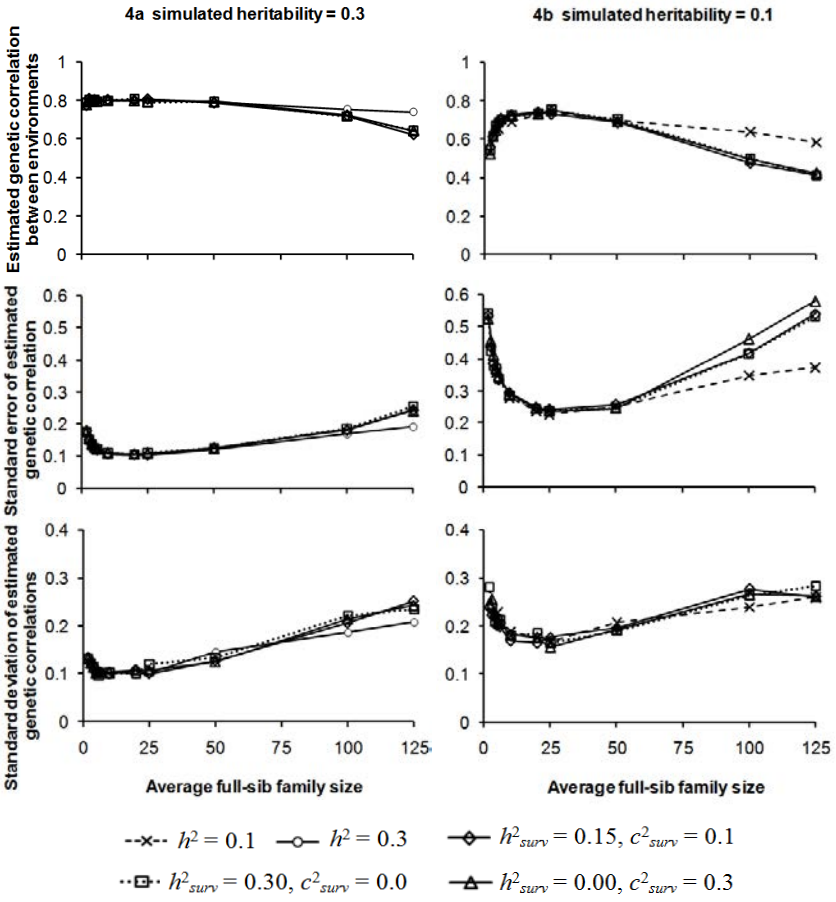

Figure 3.4 Average estimated genetic correlation ($\hat{r}_g$) between two environments, average standard error, and average standard deviation of the estimates under equal and unequal family size scenarios. The simulated genetic correlation was 0.8. Plotted lines are for equal family size scenarios (h^2 = 0.1, h^2 = 0.3) and for unequal family size with alternative parameters used for selective mortality (h^2_{surv}, c^2_{surv} scenarios).

However, at these family sizes median and mode ranged between 0.771-0.808 and 0.775-0.875, respectively.

Consequently, the optimal designs (in terms of bias and precision) for the equal and unequal family size scenarios did not differ. The estimated genetic correlation at the optimal design was slightly less than the expected value of 0.8 when h^2 was 0.1 (Figure 3.4b).

For h^2 of 0.3, there was no difference in $\hat{r}_g$, SE, and SD between balanced and unbalanced designs, except when average family size was larger than 75 (Figure 3.4a). For h^2 of 0.1, the unbalanced design produced more bias of $\hat{r}_g$ and higher corresponding SE at larger average family size (>50) than balance design (Figure 3.4b). The lowest bias and lowest SE in estimating r_g from unbalanced population structures was observed when the average family size was around 10-20 for h^2 of 0.3, and 20-25 for h^2 of 0.1.

3.4. Discussion

3.4.1. Bias

The results show that population structures characterised by very small family sizes (<10) can result in downward biased genetic correlation estimates between environments when a correlation of 0.8 was simulated. Such poor designs may indicate the presence of (strong) GxE interaction, when it in fact does not exist. In the worst case, these results may then be used to decide that a multi-environment breeding programme is needed. GxE interaction can be accounted for in two ways. Firstly, sibs of breeding candidates can be tested in alternative production environments and breeding values are estimated for each environment. Secondly, separate breeding programmes specific to each production environment may be established. Both solutions result in increased costs and complexity of breeding activities.

Small family size may also lead to a dramatic increase in standard deviation of genetic correlation estimates. Thus, there is a chance that in a single experiment, estimated genetic correlation can be either over or underestimated.

Given a fixed population size of 1000-2000, the optimal design to estimate if r_g < 0.8 (i.e. to test for significant GxE interaction) is 40-200 families with a family size of 10-40 individuals (Table 3.2). If the genetic correlation is less than 0.8, the GxE interaction will typically be moderate enough to be accounted for in a breeding programme (Robertson, 1959b). In aquaculture, experimental designs to detect GxE use highly varied population structures. For example, in rainbow trout

(*Oncorhynchus mykiss*), population structure was 10-42 x 35 (family size per environment x family number) (McKay et al., 1984). In sea bass (*Dicentrarchus labrax* L.), population structure was 27-30 families, corresponding with family size of 124-114 (a total of 3440 fish sampled; Saillant et al., 2006). In European whitefish (*Coregonus lavaretus* L.), population structure was 12 x 70 (Quinton et al., 2007). In Arctic charr (*Salvelinus alpinus* L.), two population structures of 44x29 and 20x26 were used in two different environments (Nilsson, 1990). In general, these studies tend to have sufficient family size but, based on our simulations, the small number of families in some of listed studies may result in low precision and bias of genetic correlation estimates.

Our results showed downward biased genetic correlation estimates, especially when family size was small, family number was low, simulated genetic correlation was high and heritability was low (h^2=0.1). When family size and heritability are reduced, it is increasingly difficult to estimate the true family mean in each environment and thus also breeding values of individuals. This artificially increases re-ranking and lowers the estimate for a genetic correlation. To minimize bias, it can be concluded that a less heritable trait therefore requires a larger sample size per family compared to a more highly heritable trait.

The results showed no biased estimation for simulated genetic correlation of 0.0. The reason is that with simulated genetic correlation of zero, family ranking in one environment is independent on family ranking in the second environment. Thus, mixing of the ranking due to poor population structure has no influence on the degree of bias. Accordingly, bias was more influenced when the simulated genetic correlation increased (Figure 3.2b vs. 3.2c, and Figure 3.3b vs. 3.3c).

In the fixed population size scenarios, low family number with large family size also caused downward biased estimates. The mode of the estimates, however, tended to remain close to the simulated value but the mean was shifted downward, implying that bias was not very severe. The left-hand skewness perhaps results from the fact that the estimation of genetic (co)variance components becomes increasingly difficult when the family number is lower.

In addition to population structure, the method of estimating genetic correlation also influences the degree of bias. A comparison of different methods revealed that using the family-means method produces a downward biased genetic correlation estimate between environments when in fact the genetic correlation is high (Windig, 1997; Astles et al., 2006). The family-means method refers to the practice of calculating Pearson correlation coefficient between environment-specific family means.

Restricted maximum likelihood (REML) is known to be a robust method allowing effective estimation of unbiased genetic parameters for diverse population structures and unbalanced designs (Patterson and Thompson, 1971; Harville, 1977). However, at extreme population structures, bias can occur because parameter estimation is difficult, as shown here.

3.4.2 Optimal designs

An optimal design plays an important role in accurately detecting GxE interaction. In this study, mean of MSE and SE generally indicated the same optimal design. However, when the simulated genetic correlation increased, optimal designs based on MSE or SE were different. For instance, with a fixed population size of 1000, a trait with h^2 of 0.1 and a simulated r_g of 0.8, the lowest mean MSE was obtained with a family size of about 20-25, whereas for the lowest SE the optimal full-sib family size was 20-63 (Figure 3.2c; Table 3.2). MSE includes both bias and sampling variance of estimated genetic correlation, and therefore MSE should be the first priority to indicate an optimal design.

When population size was fixed, both family size and family number were varied. Consequently, optimal design tended to be the balance of the two to reach the lowest MSE (and SE). The results from all fixed population scenarios and all simulated genetic correlations showed that for h^2 of 0.1, the lowest MSE and SE were found when family size was about 20-25 individuals (family number of 100-80 and 50-40 for population fixed to 2000 and 1000, respectively). For h^2 of 0.3, the lowest MSE and SE were found for family size of 10 (family number of 200 and 100 for population fixed to 2000 and 1000, respectively). Similar to our results, Martinez et al.(2006) found increased accuracy of estimated breeding value of sea body weight for rainbow trout breeding candidates located at a freshwater nucleus when the number of individuals tested at the sea increased from 7 to 20 per family.

For a population with 1000 individuals and family size of 20, family number of 50, r_g of 0.8 and h^2 of 0.1, the 95% confidence interval of estimated genetic correlation is 0.300-1.232 (0.766±1.96x0.238). For a population size of 2000, the same design (family size of 20 and family number of 100) gives a confidence interval of 0.477-1.095 (0.786±1.96x0.158). In both cases the design is at the optimum. As a result, population size of 2000 is almost two times more powerful than population size of 1000. Increasing population size to increase the power of detecting GxE interaction is essential but will make a study more costly.

Two alternative h^2 values were used here. Heritability of 0.1 is often seen for health, survival and some quality traits (Fjalestad et al., 1996; Gitterle et al., 2006; Tobin et al., 2006; Vehviläinen et al., 2008; Navaro et al., 2009) whereas heritability of 0.3 is typical for growth-related traits (Gjedrem, 1983; Kause et al., 2003; Charo-Karisa et al., 2007; Dupont-Nivet et al., 2008). It can be concluded that when studying GxE interaction for multiple traits, optimizing a design simultaneously for traits with low h^2 and moderate h^2 is challenging. The exception to this is the optimisation of designs for survival and growth traits at harvest. When initially having large family size optimised for survival, family size at harvest is smaller for harvest growth traits (h^2=0.3), but still the design may remain close to the optimal for both traits. Moreover, the expected low and moderate h^2 traits can be recorded from different number of animals using a predefined animal list at trait recording.

In this study, parental animals were assumed non-related and a paternal nested mating design was used. In multigenerational nucleus breeding programmes, parents are expected to be related creating more complex family structures within an offspring generation. Moreover, with full control over reproduction, factorial or partial factorial designs can also be used in practice (Dupont-Nivet et al., 2006). Partial factorial designs will create large half-sib structures which could potentially influence the optimal design. Both increased genetic links between parents and large half-sib families will potentially make the design more robust against bias when using small full-sib families. However, simultaneously the number of effective parents is reduced which may increase sampling variance.

Our study focused on the optimal design for genetic correlation estimates. However, there is a relationship between the optimal design for heritability and genetic correlation estimates. Osborne and Paterson (1952) and Robertson (1959a) derived the sampling variance of intraclass correlation equation for balanced analysis of variance (ANOVA). This expression can be used to estimate standard error of heritability for a half-sib design, and thus optimal population structure for h^2 estimation (Lynch and Walsh, 1998). Furthermore, Robertson (1959b) and Tallis (1959) studied optimal designs for estimating genetic correlations using nested analysis of covariance (ANCOVA). They showed that the optimal design for minimizing sampling variance of estimated heritability is also the optimal design for estimating genetic correlation (Lynch and Walsh, 1998). However, this depends on the value of the genetic correlation (Robertson, 1959b).

When r_g approaches unity, results from this study show that the optimal family size to estimate h^2 becomes approximately one quarter that of the optimal design for estimating r_g (results not shown). For population size fixed at 2000, our results did

show that when the simulated genetic correlation was 0.0, the design with the smallest SE for estimating h^2 of 0.1 and for estimating a r_g of 0.0 were the same. When genetic correlation was increased to 0.8, the optimal design for estimating h^2 was less than twice (family size of 10-15 individuals, corresponding to family number of 200-134) that of the optimal design for estimating r_g (family size of 40-60 individuals, corresponding to family number of 50-34). The reason is that when two traits are highly genetically correlated, estimation of h^2 in one environment is supported by sib information from the other environment.
However, for a trait with h^2 of 0.3, the difference between the optimal design for estimating genetic correlation (family size of 15-20 individuals, corresponding to family number of 134-100) and h^2 (family size of 10 individuals, corresponding to family number of 200) was smaller than for a trait with h^2 of 0.1. This suggests that the relationship between the optimal design for genetic correlation and heritability also depends on the heritability of a trait studied.

3.4.3 Unequal family size

Even when a breeding programme is started with equal family sizes, e.g. number of eggs or number of fish stocked, there is typically family-specific mortality during all life stages (Kanis et al., 1976; Vehviläinen et al., 2008, 2010). For mass spawning species, unequal contribution of parents to the offspring population can occur (Blonk et al., 2009). Both mechanisms lead to unequal family sizes at harvest. Using ANCOVA, it has been shown that unequal family size leads to higher sampling variance for a genetic correlation compared to equal family size designs (Hammersley, 1949; Tallis, 1959). In contrast to the papers by Hammersley (1949) and Tallis (1959), we found estimated SE of genetic correlation from data with unequal and equal family size to be the same at the optimal design. The likely explanation is that the REML approach used in our simulation is robust against unbalanced designs (Patterson and Thompson, 1971; Lynch and Walsh, 1998). Thus, SE was not strongly influenced by unequal family size at the optimal design.
The equal and unequal family size scenarios had the same population size at GxE trait recording, so family size variation was the only factor varied. A range of realistic h^2 and c^2 parameters for survival (Kanis et al., 1976; Vehviläinen et al., 2008, 2010) was used to generate variation in family size, and to make realistic designs. However, our result shows that an optimal design for estimating genetic correlation is not influenced by family differences in survival (or parental contribution). Nevertheless, in practice, it is difficult to obtain the exact planned

final population size for trait recording because average survival level of cohorts can vary significantly (Vehviläinen et al., 2008).

In conclusion, our simulation study specifically focused on the within population GxE interaction where the estimated genetic correlation between two environments is a measurement of GxE interaction. There are four factors leading to downward biased estimate; low h^2, high (simulated) r_g, small family size, and low family number. An optimal design is not affected by unequal family sizes caused by differential mortality among families.

3.5 Acknowledgements

This study is part of the PhD project: Optimization of rainbow trout breeding program for multi-trait selection with GxE interaction. This project is funded by Ananda Mahidol Foundation (The King Scholarship, Thailand) and Troutlodge, Inc., USA.

References

Astles, P.A., Moore, A.J., Preziosi, R.F., 2006. A comparison of methods to estimate cross-environment genetic correlations. J. Evol. Biol. 19, 114-122.

Blonk, R.J.W., Komen, J., Kamstra, A., Crooijmans, R., 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.

Charo-Karisa, H., Bovenhuis, H., Rezk, M.A., Ponzoni, R.W., van Arendonk, J.A.M., Komen, H., 2007. Phenotypic and genetic parameters for body measurements, reproductive traits and gut length of Nile tilapia (*Oreochromis niloticus*) selected for growth in low-input earthen ponds. Aquaculture 273, 15-23.

Dupont-Nivet, M., Vandeputte, M., Haffray, P., 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.

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., 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., 1952. The problem of environment and selection. Am. Nat. 86, 293-298.

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

Fishback, A.G., Danzmann, R.G., Ferguson, M.M., Gibson, J.P., 2002. Estimates of genetic parameters and genotype by environment interactions for growth traits of rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. Aquaculture 206, 137-150.

Fjalestad, K.T., Larsen, H.J.S., Røed, K.H., 1996. Antibody response in Atlantic salmon (*Salmo salar*) against *Vibrio anguillarum* and *Vibrio salmonicida* O-antigens: Heritabilities, genetic correlations and correlations with survival. Aquaculture 145, 77-89.

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

Gitterle, T., Odegard, J., Gjerde, B., Rye, M., Salte, R., 2006. Genetic parameters and accuracy of selection for resistance to white spot syndrome virus (WSSV) in *Penaeus* (*Litopenaeus*) *vannamei* using different statistical models. Aquaculture 251, 210-218.

Gjedrem, T., 1983. Genetic variation in quantitative traits and selective breeding in fish and shellfish. Aquaculture 33, 51-72.

Hammersley, J.M., 1949. The unbiased estimate and standard error of the interclass variance. Metron 15, 189-205.

Harville, D.A., 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Am. Stat. Assoc. 72, 320-338.

Kanis, E., Refstie, T., Gjedrem, T., 1976. Genetic analysis of egg, alevin and fry mortality in salmon (*Salmo salar*), sea trout (*Salmo trutta*) and rainbow trout (*Salmo gairdneri*). Aquaculture 8, 259-268.

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

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E., Eskelinen, U., 2003. Selection against early maturity in large rainbow trout *Oncorhynchus mykiss*: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228, 53-68.

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

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

Martinez, V., Kause, A., Mäntysaari, E., Mäki-Tanila, A., 2006. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout (*Oncorhynchus mykiss*): I. One-stage selection. Aquaculture 254, 182-194.

McKay, L.R., Friars, G.W., Ihssen, P.E., 1984. Genotype x temperature interactions for growth of rainbow-trout. Aquaculture 41, 131-140.

Navarro, A., Zamorano, M.J., Hildebrandt, S., Ginés, R., Aguilera, C., Afonso, J.M., 2009. Estimates of heritabilities and genetic correlations for body composition traits and G × E interactions in gilthead seabream (*Sparus auratus* L.). Aquaculture 295, 183-187.

Nilsson, J., 1990. Heritability estimates of growth-related traits in Arctic charr (*Salvelinus alpinus*). Aquaculture 84, 211-217.

Osborne, R., Paterson, W.S.B., 1952. On the sampling variance of heritability estimates derived from variance analyses. Proc. R. Soc. Edind. 64, 456-461.

Patterson, H.D., Thompson, R., 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58, 545-554.

Pierce, L.R., Palti, Y., Silverstein, J.T., Barrows, F.T., Hallerman, E.M., 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.

Quinton, C.D., Kause, A., Ruohonen, K., Koskela, J., 2007. Genetic relationships of body composition and feed utilization traits in European whitefish (*Coregonus lavaretus* L.) and implications for selective breeding in fishmeal- and soybean meal-based diet environments. J. Anim. Sci. 85, 3198-3208.

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.

Robertson, A., 1959a. Experimental design in the evaluation of genetic parameters. Biometrics 15, 219-226.

Robertson, A., 1959b. The sampling variance of the genetic correlation coefficient. Biometrics 15, 469-485.

Saillant, E., Dupont-Nivet, M., Haffray, P., 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.

Sylvén, S., Rye, M., Simianer, H., 1991. Interaction of genotype with production system for slaughter weight in rainbow trout (*Oncorhynchus mykiss*). Livest. Prod. Sci. 28, 253-263.

Tallis, G.M., 1959. Sampling errors of genetic correlation-coefficients calculated from analyses of variance and covariance. Aust. NZ. J. Stat. 1, 35-43.

Tobin, D., Kause, A., Mäntysaari, E.A., Martin, S.A.M., Houlihan, D.F., Dobly, A., Kiessling, A., Rungruangsak-Torrissen, K., Ritola, O., Ruohonen, K., 2006. Fat or lean? The quantitative genetic basis for selection of lipid and protein composition traits in breeding schemes of rainbow trout. Aquaculture 261, 510-521.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H., Paananen, T., 2008. Survival of the currently fittest: Genetics of rainbow trout survival across time and space. Genetics 180, 507-516.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H., Paananen, T., 2010. Genetic architecture of rainbow trout survival from egg to adult. Genet. Res. 92, 1-11.

Windig, J.J., 1997. The calculation and significance testing of genetic correlations across environments. J. Evol. Biol. 10, 853-874.

4

Enhancing selective breeding for growth, slaughter traits and overall survival in rainbow trout (*Oncorhynchus mykiss*)

P.Sae-Lim[1], H. Komen[1], A. Kause[2], K. E. Martin[3], R. Crooijmans[1], J. A. M. van Arendonk[1], J. E. Parsons[3]

[1] Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2] MTT Agrifood Research Finland, Biotechnology and Food Research, Biometrical Genetics, FI-31600 Jokioinen. Finland; [3]Troutlodge, Inc., 12000 McCutcheon Rd., Sumner, WA 98390, USA

Aquaculture (2013) 372-375:89-96

Abstract

Enhancing selection using two-stage selection is normally implemented by pre-selection for tagging weight (BWT) and by final selection for ungutted harvest weight (BWH) and thermal growth coefficient from tagging to harvest (TGC_{TH}). However, selection on harvest traits, i.e., gutted weight (GBWH), visceral percentage (VISW%), condition factor (CFH), and overall survival (SURV) can be enhanced by exploiting correlated traits. It can be hypothesized that the efficiency of two-stage selection on genetic response in BWH and TGC_{TH} is dependent on their genetic (r_g) and phenotypic (r_p) correlations with BWT and therefore dependent on the time point of pre-selection. The aims of this study were first, to estimate genetic parameters (heritability: h^2, r_p, and r_g) for BWT (7 months), BWS (weight at sorting, 9 months), BWH (14 months), TGC_{TH}, GBWH, VISW%, CFH, and SURV. Second, these genetic parameters were used in two deterministic simulation studies; *i*) one- and two-stage selection to compare genetic responses in BWH and TGC_{TH}, and *ii*) alternative selection indices using correlated traits to compare corresponding accuracy of selection (r_{IH}) for slaughter traits, CFH, and SURV. Genetic parameters were estimated using an animal mixed model in ASReml on 2,041 fish records. The main results showed that first; r_g of BWT was 0.35 with BWH but -0.25 with TGC_{TH} whereas the r_g of BWS was 0.72 with BWH but 0.39 with TGC_{TH}. Pre-selection for BWS led to genetic response of 54.15 g in BWH which was higher than the genetic response from pre-selection for BWT (51.90 g). Similarly, pre-selection on BWS enhanced correlated genetic response in TGC_{TH} to 0.30 $g^{(1/3)}/^oC*day$. In contrast, pre-selection for BWT resulted in lower correlated genetic response in TGC_{TH} of 0.20 $g^{(1/3)}/^oC*day$. It can be concluded that genetic improvement of BWH and TGC_{TH} can be enhanced by postponing pre-selection to a later age. However, an optimal time point for tagging and pre-selection should be found to minimize common environmental effects and rearing costs during communal rearing of full-sibs. Second, including GBWH in a selection index can reduce unfavourable selection responses in VISW%. The GBWH is highly genetically correlated with BWH and can be easily indirectly selected. TGC_{TH} is a good predictor for selection for lower VISW%, and higher SURV, but not for higher CFH. To control genetic changes in the condition factor, it should be included to the selection index.

Key words: body weight, condition factor, genetic correlation, response to selection, thermal growth coefficient, two-stage selection

4.1 Introduction

Rainbow trout (*Oncorhynchus mykiss* Walbaum) is a globally important fish species for aquaculture. In 2009, world aquaculture production of rainbow trout was 732,432 tons with an economic value of 3.4 billion dollars (FAO, 2012). Growth is the main breeding goal in rainbow trout breeding programmes. In addition, production traits such as maturity age, gutted weight, and overall survival are also targeted (Gjedrem, 1985; Crandell and Gall, 1993; Kause et al., 2007a; Vehviläinen et al., 2008).

Growth in trout is usually measured as body weight at a given harvest age (BWH). In Salmonids, average heritability for body weight across ages is around 0.24 (range: 0.10-0.52) (Gunnes and Gjedrem, 1981; Linder et al., 1983; Gall and Huang, 1988; Elvingson and Johansson, 1993; Kause et al., 2002, 2007a; Quinton et al., 2005). Instead of body weight at a fixed age, growth rate corrected for water temperature can be selected for, i.e., thermal growth coefficient (TGC; Cho, 1992). Faster growth shortens the rearing period, lowers feed cost, and reduces water usage. Heritability for TGC from 9 to 12 month post-hatch (Silverstein et al., 2009) was 0.32. Fish breeders may implement at fingerling stage pre-selection for tagging body weight (BWT), aiming at enhancing the genetic improvement of harvest body weight (Martinez et al., 2006). Because the positive genetic relationships between age-specific body weights become weaker as time between measurements increases (Su et al., 2002; Rutten et al., 2005a), it can be hypothesized that the efficiency of pre-selection on correlated genetic response in harvest body weight and TGC is influenced by the time point of pre-selection.

Condition factor and slaughter traits such as gutted weight and visceral percentage are important traits for farmers, processors and customers. Farmers are typically paid for gutted weight, and offal from gutting is waste. Condition factor is a highly valued trait by both farmers and consumers as rounded body shape is not preferred and it is also related to fish health because fish with skeletal deformations tend to have a high condition factor (McKay and Gjerde, 1986; Kause et al., 2003; Sullivan et al., 2007; Sae-Lim et al., 2012). Fish breeders typically select for wet body weight, however this may result in undesired changes in body composition and appearance. Lipid deposition, visceral percentage and condition factor are known to be unfavourable genetically correlated with ungutted body weight (Gjerde and Schaeffer, 1989; Kause et al., 2002, 2003, 2007a). Accordingly, selection strategies for rapid growth need to improve gutted weight with high economic value and to avoid detrimental correlated genetic responses in body composition and appearance. This is challenging because traits such as gutted

weight, visceral percentage and fillet percentage cannot be directly recorded from live breeding candidates. High overall survival across a rearing period yields high production volume at harvest and increases fish welfare. Overall survival is a binary trait, and thus, in a univariate breeding value evaluation, all surviving full-sibs get the same estimated breeding value (EBV). When a normally distributed trait genetically correlated with survival, e.g. body weight recorded from breeding candidates, is analysed together with survival in a multi-trait analysis, the accuracy of selection for survival is increased. This is because body weight contains information on the genetic potential to survive, and hence surviving full-sibs with different body weight will get different EBVs for survival (Thompson and Meyer, 1986). Moreover, the heritability of overall survival is low (Vehviläinen et al., 2008, 2010), and thus means to enhance selection for survival are needed. The first aim of the present study was to investigate whether pre-selection on BWT can enhance genetic improvement of BWH and TGC. Second, we assessed the possibility of including correlated traits into alternative multi-trait selection indices to enhance genetic responses in gutted body weight, visceral percentage, and condition factor. Finally, we studied whether growth or slaughter traits are genetically correlated with survival, and assessed the degree to which the accuracy of selection can be increased by selection on the correlated traits. Selection accuracy is a measure of increase in efficiency of selection when using additional traits in selection.

4.2 Materials and methods

4.2.1 Experiment

The fish used in this study were all-female progeny, from selected parents of the Kamloop strain, obtained from Troutlodge, Inc., USA. In August 2009, a total of 58 sires and 100 dams were mated to produce 100 families. The sires were sex-reversed by using sex hormonal manipulation which causes females to produce male gonads. Each sire was mated to one to three dams (average = 1.7) and each dam was mated to one sire. Production of families took place over a period of 4 weeks. Fertilized eggs were incubated in 100 incubators (one for each family). Eggs from each of the 4 spawning weeks were incubated at different temperatures to ensure that all groups hatched at approximately the same time. In September 2009, approximately two weeks prior to hatch, a total of 25 eyed eggs from each family were randomly sampled, pooled and shipped to Forellenzucht Trostadt, Germany. The eyed eggs were incubated at 10^{o}C in a single incubator until hatching

(October 2009). Hatched fry were stocked in a 0.7 m^3 indoor flow-through tank for yolk-sac absorption and at mean weight of 5.7 g moved to a 2.5 m^3 circular indoor tank. Fish were daily fed 2.5% of average body weight with powdered feed (Aller Futura), and then 1.5 mm and 2 mm size pelleted feed, by automated feeder (Aller Aqua, Germany). Dissolved oxygen was maintained at 10 ppm and water temperature was monitored daily using an automatic water quality controller (HACH LANGE sc 1000, HACH LANGE GmbH, Germany). The average water temperature was 9.9 ^oC (range: 5.3 to 14.5 ^oC).

4.2.2 Traits measurement

In April 2010, the fish reached tagging size (mean = 27.1 g; n = 2235 fish). They were anesthetized by using clove oil (10 ppm), and individually tagged in the body cavity using passive integrated transponders (DORSET Identification b.v., the Netherlands). At tagging, body weight (BWT) and fork length (FLT) were measured and a fin-clip was collected from the caudal fin tip of each fish and stored at -20 ^oC in 96-well plates filled with 100% ethanol. Tagging occurred over a period of 3 days. All tagged experimental fish were stocked in a single 5 m^3 indoor raceway, mixed with approximately 4000 non-tagged fish of similar size to obtain fish density used in commercial production. Feeding was 1.5 to 1.7% of average body weight.

In June 2010 at an average body weight of 64.1 g (n = 2091 fish), fish were moved from the indoor raceway to a 400m^3 outdoor pond. The tagged fish were separated from the non-tagged fish using the tag reader (GR250: DORSET Identification b.v., the Netherlands). During sorting, body weight and fork length was measured of each tagged fish (BWS, in grams and FLS, in mm). In order to keep the experimental fish separated from the commercial fish, all tagged fish were stocked in a 16m^3 net-cage, suspended in the pond. The fish were fed with 3-4.5 (mix) mm to 6 mm pellets (Aller Aqua, Germany) until harvest. The feeding was reduced to 0.53 to 0.87 % of average body weight. The feed contained 42 to 64% protein and 11 to 30% lipid, the protein content decreasing and lipid increasing with fish age.

In December 2010 at an average body weight of 376.4 g (n = 1992 fish), the tagged fish were measured for ungutted body weight at harvest (BWH, in grams) and fork length (FLH, in mm). Then the fish were gutted, and gutted body weight (GBWH) was measured. The condition factor at tagging and harvest (CFT and CFH)) was calculated as $(10^5*BW)/FL^3$. Visceral weight percentage at harvest (VISW%) was calculated as (BWH - GBWH)*100/BWH. Thermal growth coefficient (TGC) was calculated as

$$\left[\frac{(\sqrt[3]{Wt}-\sqrt[3]{Wo})}{T\times t}\right]\times 1000, \quad [1]$$

where W_t = body weight at time t (sorting or harvest), W_o = initial body weight (tagging or sorting), T = average daily water temperature (^oC), and t = rearing period in days. To correct for the non-linear relationship between growth and water temperature (Jobling, 2003), TGC formula was modified by substituting T with k calculated from the model used by Mallet et al. (1999):

$$k=\frac{T_{opt}(T-T_{min})(T-T_{max})}{(T-T_{min})(T-T_{max})-(T-T_{opt})^2}, \quad [2]$$

where k = new temperature (modified from equation 2 in Mallet et al. (1999)) corrected for the concave relationship between growth and temperature. Optimum water temperature: T_{opt} = 14.83 ^oC, was taken from the average for salmonid growth (Hokanson et al., 1977; Austreng et al., 1987; FAO, 2011). Daily water temperature: T was from the daily measurement at the farm. The limits for the lower and upper thermal tolerance: T_{min} = 0 ^oC, and T_{max} = 23 ^oC, respectively, were taken from the literature (Hokanson et al., 1977; Ojolick et al., 1995; Matthews and Berg, 1997). The modification makes sense biologically because growth rate is reduced at temperature above the optimum. The switch from the standard to the new modified TGC had no influence on (co) variance component estimates because all fish were recorded for body weight during the same period and thus the standard TGC was modified with the same constant value for all fish. TGC was calculated for 3 periods, i.e. tagging to sorting (TGC_{TS}), sorting to harvest (TGC_{SH}), and tagging to harvest (TGC_{TH}). Overall survival (SURV) was coded as "1" for fish surviving from tagging to harvest and as "0" for fish whose tag was not recorded at harvest.

4.2.3 Pedigree reconstruction

To re-construct the pedigree, DNA was isolated from the fin-clips of the tagged fingerlings and from their 158 parents. The DNA isolation was done using Nucleospin® 96 Tissue Core Kit. The following 9 microsatellite markers were used for PCR: OMM1008, OMM1051, OMM1088, OMM1097 (Rexroad et al., 2002), OMM5007, OMM5047 (Rexroad et al., 2005), OMM5233, OMM5177 (Coulibaly et al., 2005), and OMM1325 (Palti et al., 2002). Multiplex PCR amplification, i.e.

quadroplex and pentaplex, was done as follows (Johnson et al., 2007): an initial 5 min denaturation at 95 ^{o}C, followed by 35 cycles of 30 s denaturation at 95 ^{o}C, 45 s annealing at 55 ^{o}C, and 90 s extension at 72 ^{o}C, and a final 10 min extension at 72 ^{o}C. Fragment analysis of the PCR products was done by relatively setting the fragment sizes to Genescan LIZ 500 size standard (Applied Biosystem). Output data were analysed using Genemapper software version 4 (Applied Biosystem).
Parental allocation was performed using PAPA software (Duchesne et al., 2002). In PAPA, likelihood is estimated for each potential parental pair, and the offspring are assigned to the pair with the highest likelihood. The known mating data were used to increase the accuracy of parental assignments. Afterwards, the results were manually checked for Mendelian inheritance. In total, 2104 out of 2235 sampled offspring were successfully allocated to the 100 full-sib families.

4.2.4 Genetic analysis

The number of records for each trait is shown in Table 4.1. After parental allocation, the number of observations was less than 2104 due to missing data (Table 4.1). The average of fish per sire ranged from 30.52 to 35.19 and the average of fish per dam ranged from 17.70 to 20.41. Heritabilities, phenotypic (r_p) and genetic (r_g) correlations were estimated using restricted maximum likelihood in a multivariate animal model (ASReml v. 3.0; Gilmour et al., 2009). Each trait was modelled as:

$$y_{ij} = \mu + Anim_i + FS_j + e_{ij,} \quad [3]$$

where y_{ij} is the observation of the *i*th individual from the *j*th full-sib family, μ is the overall mean, *Anim* is the additive genetic effect of the *i*th animal, *FS* is the full-sib common environmental effect, and *e* is the random error term. The full-sib effect was modelled without a pedigree, accounting for common effects to full-sibs, e.g. incubator effects and environmental maternal effects and non-additive genetic effects. Models with and without full-sib effect were compared for heritability estimates. To estimate genetic correlations, the full-sib effect was excluded from the model because in many cases the full-sib effect captured all the (co)variance of the traits (Maluwa et al., 2006). Potential selection bias due to selective mortality was accounted for by always including BWT as a reference trait in a multi-trait analysis (Pollak et al., 1984; Kause et al., 2011).
Heritability from the model without full-sib effect was calculated as $h^2 = V_A/(V_A + V_e)$, and heritability from the model with full-sib effect was quantified as h^2_{FS} =

$V_A/(V_A + V_{FS} + V_e)$. The common environmental effect due to full-sib families was calculated as $c^2 = V_{FS}/(V_A + V_{FS} + V_e)$. Heritability for SURV was first estimated on the observed binary scale and subsequently transformed to the underlying normally distributed liability scale following the formula of Dempster and Lerner (1950). Residual (co)variance between overall survival and traits measured at harvest was set to zero. Phenotypic correlation between survival and traits at harvest was, therefore, not applicable. Phenotypic and genetic correlation matrices were bended to be positive definite (Hayes and Hill, 1981). The bending had minor effect on phenotypic (range: 0 to 0.0009) and genetic (range: 0 to 0.0027) correlation estimates.

4.2.5 Deterministic simulation

To evaluate the efficiency of alternative direct and indirect selection strategies, accuracy of selection index (r_{IH}) and in some cases, expected genetic responses to selection, were predicted using Sel*A*ction computer software (Rutten et al., 2002). Phenotypic and genetic parameters estimated in this study were used as input. In general, c^2 is high in early life stages, but decreases over time. Consequently, for the traits related to tagging and sorting, h^2_{FS} was used while for harvest traits, h^2 was used.

One-stage and two-stage selection

First, we compared one- and two-stage selection schemes on total genetic response in BWH and total genetic correlated response in TGC_{TH}. Breeding goal (H) was body weight at harvest (H = 1*BWH). Two-stage selection included pre-selection for either BWT or BWS, followed by multi-trait EBV index selection for BWH and TGC_{TH}. These were compared to one-stage multi-trait EBV index selection for BWH and TGC_{TH} only.

Initial population size was 50,000 fish and was reduced to 200 fish (selection candidates) at final selection. The number of full-sib families was 100 (1:1 mating design), producing 500 (250 males and 250 females) offspring per dam. At the first stage of selection, the population of 50,000 fish was reduced to 5,000 by either random selection (one-stage selection) or by own performance selection (two-stage selection). In practical selection schemes, the first-stage selection is typically practiced within families at tagging. However, it was not possible to do within-family selection for pre-selection in Sel*A*ction. Therefore own performance selection was used. In the second stage, selection was based on indices I_1 = BWT +

BWH + TGC_{TH}, or I_2 = BWS+ BWH + TGC_{TH}, in which BWT and BWS have own performance observations, the other traits have EBVs. The population structure is summarized in Table 4.2. Two-stage selection had lower proportion of selected animals (P_{total} = 0.004) than one-stage selection (P_{total} = 0.04).

The response to selection for BWH, and TGC_{TH} and the ratio r_{IH1}/r_{IH2} of the accuracy of selection index at the first stage (r_{IH1}) and the second stage (r_{IH2}) were used to evaluate the effect of pre-selection on total genetic response for BWH and TGC_{TH}. The higher ratio of r_{IH1}/r_{IH2} indicates better pre-selected trait as a predictor for BWH and TGC_{TH}.

Enhancing selection response using correlated traits (one stage selection only).

In this scenario, the effect of including correlated traits in the selection index on the accuracy of selection against VISW%, or selection against CFH, or for SURV was studied. The breeding goals were H = -1*VISW%, H = -1*CFH, and H = 1*SURV, respectively. First, single trait selection for the target trait (VISW%, CFH, SURV) was simulated to estimate r_{IH} (direct selection). Second, indices combining the target trait and correlated traits (BWT, BWS, BWH, GBWH, TGC_{TH}, and CFT) were simulated (combined direct and indirect selection). Finally, the effect on the target trait due to selection on correlated traits only (indirect selection) was studied. All simulations in this scenario were based on BLUP-EBVs. The population structure was the same as one-stage selection described above (Table 4.2). Additionally, for the slaughter traits (GBWH and VISW%), we assumed that 20 fish per family were gutted to provide the information for selection indices.

4.3 Results

4.3.1. Heritability of traits

Estimates for variance components are given in Table 4.1. For the model with common environmental effect, h^2_{FS} for BWT, BWS, BWH, and GBWH was low to moderate (0.11-0.28). Estimates for CFT, CFH and VISW% were moderate to high h^2_{FS} (0.24±0.15, 0.16±0.10, and 0.33±0.05 respectively). For overall survival (SURV), h^2_{FS} was 0.07±0.06 on observed scale and 0.19±0.17 on the underlying liability scale. Estimates for TGC were inconsistent, depending on the body weight-recording interval. From tagging to sorting (TGC_{TS}), h^2_{FS} was 0.25±0.11 while from

sorting to harvest (TGC_{SH}), h^2_{FS} was close to zero (0.01±0.00). For the whole experimental period (TGC_{TH}), h^2_{FS} was 0.12±0.10.

Estimates for common environmental effect (c^2) varied from 0.00±0.00 to 0.10±0.07 (Table 4.1). When c^2 was excluded from the model, heritability estimates were two to three fold higher compared to the estimates obtained from the model with the full-sib effect. The most extreme case was TGC_{SH} which showed moderate h^2 of 0.18±0.04, whereas h^2_{FS} was 0.01 and c^2 was 0.08. This might indicate that c^2 sometimes captures all genetic variance, leading to underestimated $h^2_{FS.}$

4.3.2 Correlations

Phenotypic and genetic correlations are given in Table 4.3. Correlations of BWH with BWT were lower (r_p = 0.35, r_g = 0.35) than with BWS (r_p = 0.63, r_g =0.72). The BWS was positively correlated to TGC_{TS} (r_p = 0.72±0.01, r_g = 0.66±0.07), TGC_{SH} (r_p = 0.02±0.03, r_g = 0.04±0.15) and TGC_{TH} (r_p = 0.35±0.02, r_g = 0.39±0.12).

In contrast, BWT showed negative or close-to-zero correlation with TGC_{TS} (r_p = 0.01±0.03, r_g = -0.10±0.13), TGC_{SH} (r_p = -0.11±0.03, r_g = -0.21±0.14) and TGC_{TH} (r_p = -0.10±0.03, r_g = -0.25±0.13). Similar to correlations between BWT and BWH, correlation between TGC_{TS} and TGC_{SH} was weak (r_p = 0.17±0.03, r_g = 0.26±0.14). This suggests that growth rate of one period is not a good predictor for the next period.

Correlations of BWH with VISW% were positive (r_p = 0.27, r_g = 0.31), indicating that selection for BWH was expected to indirectly increase percentage of visceral weight (Table 4.3). However, the correlations of GBWH with VISW% were weaker (r_p = 0.18, r_g = 0.21), compared to those of BWH.

Correlations between CFT and CFH showed moderate phenotypic relationship (r_p=0.50), and strong genetic correlation (r_g=0.80). The BWT and CFT were moderately correlated (r_p=0.35, and r_g=0.50), and reduced over time (BWH and CFH: r_p=0.08, and r_g=0.23); this was supported by the observation that TGC_{TH} had a negative or close-to-zero correlation with CFT and CFH.

Overall survival (SURV) showed stronger positive correlations with BWS (r_p = 0.34, r_g = 0.50) than with BWT (r_p = 0.06, r_g = 0.07), indicating that pre-selection for BWS may increase overall survival until harvest. Moderate genetic correlations were also equally found between SURV and both BWH and GBWH (r_g = 0.57). However, TGC_{TH} showed the highest genetic correlation with SURV (r_g = 0.62).

Table 4.1 Number of observation (N), trait mean, phenotypic variance (V_p), heritability from model excluding full-sib effect (h^2), heritability from model including full-sib effect (h^2_{FS}), full-sib common environmental effect ratio (c^2) and their standard error (SE).

Traits	N	Mean	V_p	h^2	SE(h^2)	h^2_{FS}	SE(h^2_{FS})	c^2	SE(c^2)
BWT	2041	27.06	43.75	0.41	0.06	0.28	0.13	0.05	0.06
CFT	2010	1.20	0.01	0.46	0.06	0.24	0.15	0.10	0.07
BWS	1900	64.06	280.35	0.33	0.05	0.11	0.11	0.10	0.06
BWH	1819	376.39	6677.50	0.23	0.05	0.13	0.09	0.04	0.04
GBWH	1770	316.20	4532.70	0.24	0.05	0.13	0.09	0.05	0.04
CFH	1818	1.34	0.01	0.30	0.05	0.16	0.10	0.07	0.05
VISW%	1770	16.08	3.24	0.33	0.05	0.32	0.05	0.00	0.00
SURV (observed scale)	1993	0.89	0.10	0.14	0.03	0.07	0.06	0.04	0.03
SURV (liability scale)	1993	N.A.	N.A.	0.39	0.09	0.19	0.17	N.A.	N.A.
TGC_{TS}	1900	1.42	0.15	0.34	0.05	0.25	0.11	0.04	0.05
TGC_{SH}	1815	1.42	0.04	0.18	0.04	0.01	0.00	0.08	0.02
TGC_{TH}	1818	1.43	0.04	0.23	0.05	0.12	0.10	0.05	0.05

BWT=body weight at tagging, CFT = condition factor at tagging, BWS = body weight at sorting, BWH = ungutted body weight at harvest, GBWH = gutted body weight at harvest, CFH = condition factor at harvest, VISW% = visceral weight percentage, SURV = overall survival at harvest, TGC_{TS} = thermal growth coefficient from tagging to sorting, TGC_{SH} = thermal growth coefficient from sorting to harvest, and TGC_{TH} = thermal growth coefficient from tagging to harvest.

N.A. = Not applicable.

4.3.3 Deterministic simulations

One and two-stage selection

The deterministic simulation for one-stage selection showed that total response to direct selection for BWH was 46 g. Including pre-selection for BWS gave a total response of 54.15 g in BWH. This was 2.25 g higher compared to pre-selection for BWT (Table 4.4). After pre-selection for BWT, correlated genetic response in TGC_{TH} was negative (-0.09 $g^{(1/3)}$/day*^{o}C).

In contrast, pre-selection for BWS resulted in positive correlated genetic response TGC_{TH} (0.04 $g^{(1/3)}$/day*^{o}C). Consequently, total correlated genetic response in TGC_{TH} from pre-selection for BWT was 0.20 $g^{(1/3)}$/day*^{o}C, which was even lower than total correlated genetic response in TGC_{TH} from one-stage selection (0.27 $g^{(1/3)}$/day*^{o}C).

Two-stage selection with pre-selection for BWS enhanced total genetic response in TGC_{TH} to 0.30 $g^{(1/3)}$/day*^{o}C. The ratio of r_{IH1}/r_{IH2} was higher for BWS (0.33) than BWT (0.25), indicating that BWS was a better pre-selected trait as a predictor for BWH and TGC_{TH} than BWT.

Table 4.2 Number of sires, dams, and full-sib families used in one- and two-stage selection.

Selection	[1]P_1	Sire	Dam	Fish/family	[2]P_2	Sire	Dam	[3]P_{total}
One-stage	1.0	2500	2500	50	0.040	100	100	0.040
Two-stage	0.1	2500	2500	50	0.040	100	100	0.004

[1] Proportion of selected animals at pre-selection,

[2] Proportion of selected animals at final selection, and

[3] Total proportion of selected animals ($P_{total} = P_1 * P_2$).

Enhancing selection response using correlated traits

The accuracy of direct selection against VISW% was 0.63 (Table 4.5). Because growth traits are unfavourably genetically correlated with VISW%, low r_{IH} to predict VISW% are favourable for selection against VISW%. Among the growth traits, selection for GBWH produced the lowest accuracy (r_{IH} = 0.11), followed by BWH and TGC_{TH} (r_{IH} = 0.17). Using all body weight measurements (BWT, BWS, BWH, and GBWH) in the index produced unfavourably high r_{IH} of 0.39. These results suggested that the most effective way to reduce the correlated response in VISW% was by selection for GBWH.

Table 4.3 Genetic correlations and their standard errors (below diagonal), and phenotypic correlations (above diagonal) between traits.

	BWT	CFT	BWS	BWH	CFH	GBWH	VISW%	SURV	TGC_{TH}
BWT	1	0.35±0.03	0.68±0.02	0.35±0.02	0.06±0.03	0.36±0.03	-0.01±0.03	0.06±0.03	-0.10±0.03
CFT	0.50±0.09	1	0.21±0.03	0.08±0.03	0.50±0.02	0.07±0.03	0.06±0.03	0.04±0.03	-0.08±0.03
BWS	0.66±0.07	0.33±0.12	1	0.64±0.02	0.03±0.03	0.64±0.02	0.10±0.03	0.34±0.02	0.35±0.02
BWH	0.35±0.12	0.15±0.13	0.73±0.07	1	0.08±0.03	0.99±0.00	0.27±0.03	N.A.	0.88±0.01
CFH	0.24±0.13	0.80±0.05	0.17±0.13	0.23±0.14	1	0.06±0.03	0.11±0.03	N.A.	0.02±0.03
GBWH	0.32±0.12	0.12±0.14	0.71±0.07	0.99±0.00	0.20±0.14	1	0.18±0.03	N.A.	0.87±0.01
VISW%	0.01±0.13	0.02±0.13	0.26±0.13	0.31±0.13	0.15±0.13	0.21±0.14	1	N.A.	0.31±0.03
SURV	0.07±0.15	0.08±0.15	0.50±0.12	0.57±0.11	0.08±0.14	0.57±0.11	0.22±0.13	1	N.A.
TGC_{TH}	-0.25±0.13	-0.15±0.13	0.39±0.12	0.80±0.05	0.08±0.15	0.79±0.05	0.30±0.13	0.62±0.11	1

BWT=body weight at tagging, CFT = condition factor at tagging, BWS = body weight at sorting, BWH = ungutted body weight at harvest, GBWH = gutted body weight at harvest, CFH = condition factor at harvest, VISW% = visceral weight percentage, SURV = overall survival at harvest, and TGC_{TH} = thermal growth coefficient from tagging to harvest.

N.A. = Not applicable.

For CFH, accuracy of direct selection was 0.66. Similar to VISW%, growth traits are genetically correlated with CFH. Indirect selection on TGC_{TH} produced r_{IH} of 0.04, a value lower than indirect selection on each body weight separately (0.07-0.13) or on all body weights together (0.19).
To improve overall survival, combining direct selection for SURV with indirect selection for either TGC_{TH} (r_{IH} = 0.58) or BWH (r_{IH} = 0.57) produced higher accuracy than direct selection for SURV only (r_{IH} = 0.54). Thus, genetic improvement of SURV could be enhanced by including one of these traits into the selection index. In addition, indirect selection on TGC_{TH} resulted in the highest r_{IH} (0.36), compared to indirect selection on BWT (0.04), BWS (0.19), or BWH (0.33). Including all body weight measurements for combined direct and indirect selection did not improve r_{IH} considerably, compared to r_{IH} from TGC_{TH} (Table 4.5).

4.4 Discussion

4.4.1 Heritability of traits

Heritabilities were estimated using two models: either with (h^2_{FS}) or without the common environmental effect (h^2). Heritability estimates were dramatically reduced when including the full-sib effect. This infers that without the full-sib effect, heritability may be overestimated. Martinez et al. (1999) showed that a genetic model ignoring full-sib effect inflated additive genetic variance. On the other hand, it is also possible that the full-sib effect artificially captures part of the additive genetic (co)variance, leading to downward biased heritability estimates (Maluwa et al., 2006). In our analysis, additive genetic variance for BWS and TGC_{SH} may have been artificially captured by the common environmental effect. These traits had moderate heritabilities when full-sib effect was excluded from the model (h^2 = 0.33 and 0.18, respectively), but low or close-to-zero when included (h^2_{FS} = 0.11 and 0.01, respectively). In general, full-sib effect is high for the early life stage, but it is low in the later stage. To avoid overestimated heritability, we therefore use h^2_{FS} for the traits measured at tagging and sorting, while h^2 for the traits at harvest is used for the further discussion.
For growth traits, the h^2_{FS} for BWT was 0.28 (7 months after hatching), which is lower than h^2 for nursery weight (3 months after hatching) of 0.37 and 0.52 (Gall and Huang, 1988). The h^2_{FS} for BWS (9 months after hatching) was 0.11, which is at the lower end of the range of h^2 reported for yearling weight from literature (0.10

to 0.52) (Linder et al., 1983; Gall and Huang, 1988; Elvingson and Johansson, 1993). Similar to both BWT and BWS, the h^2 for BWH (15 months after hatching) was 0.23, which is lower than h^2 for body weight (0.28 to 0.51) at age of 1.0 to 2.5 years old (Elvingson and Johansson 1993). Heritability for GBWH was 0.24; this value is difficult to compare to literature because most h^2 estimates are for large size fish (1 to 3 kilograms) or older than 15 months (age at harvest for portion-sized fish). However, h^2 for gutted weight ranges from 0.22 to 0.45 (Gjerde and Schaeffer, 1989; Elvingson and Johansson, 1993; Kause et al., 2002; 2007a). The h^2 for TGC_{SH} (between 9 and 15 months post-hatch) was 0.18, which is lower than h^2 for TGC (between 9 and 12 months post-hatch) of 0.32 estimated by Silverstein et al. (2009). The h^2 for TGC_{TH} was 0.23, which is lower than h^2 for overall growth rate (TGC_{14}: h^2(M)=0.46 and h^2(PB) =0.65) (Le Boucher et al., 2011). The h^2(M) and h^2(PB) are heritabilities estimated from rainbow trout fed with marine (M) and plant-based (PB) diets, respectively. The finding of Le Boucher et al. (2011) showed that h^2 for TGC may depend on the diet.

In our analysis visceral percentage (VISW%) showed lower h^2 (0.33) than estimated by Kause et al. (2007a) (0.58). Heritability estimates for condition factor range from 0.19 to 0.59 for various ages (Gjerde and Schaeffer, 1989; Elvingson and Johansson, 1993; Rye and Refstie, 1995; Kause et al., 2002, 2003; Neira et al., 2004). Our h^2_{FS} estimate for CFT (0.24) and h^2 for CFH (0.30) is in this range. For survival, h^2_{FS} is similar (observed scale = 0.07, liability scale =0.19) to the h^2 estimated previously for overall survival, ranging from 0.08 to 0.17 (Vehviläinen et al., 2008).

The magnitude of most heritability estimates is moderate to high, showing the possibilities for selective breeding. Yet, our heritability estimates are on average still lower than reported in previous studies. This may be due to lower temperature (average water temperature = 9.97^{o}C and in winter = 6^{o}C) which can lead to lower heritability estimate because fish have less opportunity to show their genetic potential. Moreover, different statistical models for genetic parameter estimates may lead to the variation of the heritability estimates.

4.4.2 Enhancing selection response for growth by two-stage selection

Pre-selection for body weight at an early age can be used to reduce the number of individuals to be reared until harvest, and to indirectly select for body weight at harvest. Pre-selection can be practiced within families at tagging, i.e., by tagging the biggest fish within families, and the final selection is done based on estimated breeding values of tagging and harvest body weights. Two-stage selection had

lower proportion of selected animals than one-stage selection (Martinez et al., 2006). This is called two-stage selection (Cunningham, 1975). For other livestock, two-stage selection has been successfully implemented, for example, in swine to increase litter size (Ruiz-Flores and Johnson, 2001).

Phenotypic and genetic correlations among body weights decrease when the time period between the measurements increases (Su et al., 2002; Rutten et al., 2005a). Re-ranking across life stages has also been found for other traits, e.g. overall survival for rainbow trout (Vehviläinen et al., 2010). The likely reasons for the re-ranking are 1) growth at different stages in life is affected by different sets of genes leading to different growth curves, 2) differences in individual environmental sensitivity in response to differences in environment before tagging, and between tagging and harvesting, and 3) interaction of genes or epistasis affecting body weight changing over time (Le Rouzic et al., 2008).

In this study, we quantified the effect of timing of pre-selection on genetic improvement of BWH and $TGC_{TH.}$ The genetic correlation between BWT and BWH was 0.35, while it was 0.72 between BWS and BWH. This implies that pre-selection on BWT improves BWH, but given the negative genetic correlation between BWT and TGC_{TH}, pre-selection for BWT does not result in fish that grow faster from tagging onwards. Pre-selection on BWT may lead to discarding smaller fish that may have higher potential for faster future growth, measured as TGC, compared to fish that are initially bigger. This consequently leads to a negative genetic gain in TGC_{TH}. Thus, the total genetic response of TGC_{TH} after two-stage selection can be lower than the response from one-stage selection for TGC_{TH}.

Alternatively, pre-selection for BWS resulted in 17.4% higher total genetic response in BWH compared to the response from one-stage selection for BWH, and a 4.3% higher response in BWH compared to the response in BWH with pre-selection for BWT. Moreover, pre-selection for BWS resulted in a higher genetic gain in TGC_{TH} than pre-selection for BWT (0.20 vs. 0.30 $g^{(1/3)}$/day*^{o}C). In addition, pre-selection for BWS led to higher ratio of r_{IH1}/r_{IH2} (0.33) than ratio of r_{IH1}/r_{IH2} (0.25) from pre-selection for BWT, indicating that BWS is a better predictive trait for BWH and TGC_{TH} than BWT. It can be concluded that selection for BWH and TGC_{TH} can be enhanced by postponing pre-selection from tagging body weight to a later age.

However, for a breeding program, the drawback of postponing pre-selection is that full-sibs are held in the family-tanks longer before tagging, and consequently, common tank effects are increasingly introduced to the full-sibs. To be able to accurately separate genetic and environmental effects, it would be beneficial to offer the same rearing environment for all fish as early as possible. In addition, postponing tagging and keeping all families separated will increase rearing cost.

Table 4.4 Response to selection and accuracy of selection index (r_{IH}) from one- and two-stage selection simulation.

	Trait selected		Genetic response				r_{IH1}/r_{IH2}
	[1]Pre-selection	[2]Final selection	Pre-selection	[4]r_{IH1}	Final selection	[4]r_{IH2}	
One-stage	N.A.	[3]*TGC*$_{TH}$	N.A.	N.A.	0.27	N.A.	N.A.
		BWH			46.13		
Two-stage	BWT	*TGC*$_{TH}$	-0.09	0.16	0.20	0.63	0.25
		BWH	9.23		51.90		
	BWS	*TGC*$_{TH}$	0.04	0.21	0.30	0.63	0.33
		BWH	12.16		54.15		

BWT=body weight at tagging, BWS = body weight at sorting, BWH = ungutted body weight at harvest, and TGC_{TH} = thermal growth coefficient from tagging to harvest.

[1] = Pre-selection for either BWT or BWH,

[2] = Final selection for BWH,

[3] = Italic letter indicated correlated genetic responses, and

[4] = Accuracy of pre-selection (r_{IH1}) and final selection (r_{IH2}).

N.A. = not applicable. Responses are per generation.

Furthermore, the optimal time point for pre-selection will depend on the harvest weight. The bigger the targeted harvest weight, the smaller the genetic correlation between tagging and harvest body weight. Consequently, early pre-selection will become less efficient. Thus, an optimal timing for tagging and pre-selection should be found. Harvest body weight and TGC are two complementary traits for improving growth. Body weight is a measure of cumulative growth over lifetime, whereas TGC measures growth rate during a specific growth period. For instance, when TGC is recorded at the end of the growing period, its improvement enhances growth when feeding costs are high. To be able to select for TGC, initial BW needs to be recorded which may increase labour cost compared to single stage selection.

4.4.3 Enhancing selection response in slaughter traits and condition factor

In some countries, gutted body weight (GBWH) is of commercial interest because farmers are paid for weight of gutted fish.

Additionally, for farmed fish used mainly for human consumption, gutted weight can be regarded as more sensible breeding goal trait compared to harvest weight because viscera are not included in gutted weight. The current and previous studies (Gjerde and Gjedrem, 1984; Kause et al., 2007a) show that GBWH can be genetically improved by selecting for BWH or TGC_{TH} because these are highly genetically correlated traits. Similarly, indirect improvement of fillet weight is effective when selecting for either gutted or ungutted body weight (Rutten et al., 2005b; Kause et al., 2007a; Nguyen et al., 2010). In addition, selection for low VISW% also indirectly improves fillet percentage (Kause et al., 2007a). Selection for BWH can lead to unfavourable correlated genetic response in VISW% (present study) and in proximate body composition such as lipid deposition (Kause et al., 2007ab). When fish are depositing high quantities of lipid, visceral waste consisting mostly of lipids, will increase. We observed that the selection for gutted harvest weight leads to a weaker unfavourable genetic change in VISW%, compared to the selection for BWH, even though GBWH and BWH weights are very highly positively correlated (r_g = 0.99). This is because BWH includes visceral weight, which is excluded from GBWH. Our study performed with portion-sized 376 g fish supports the findings by Kause et al. (2007a) who used 1.15 kg rainbow trout. Thus, it seems to be a general feature across ages in rainbow trout. In contrast, including all body weight measurements in the selection index resulted in a more unfavourable

correlated genetic response in VISW%. Gutted weight can be recorded during sib testing, and then selection index can include both GBWH and VISW% to select breeding candidates.

Condition factor measured at tagging and harvest showed strong positive genetic correlations, implying low re-ranking of condition factor over time. Selection for CFH can be easily enhanced by selection for CFT. However, phenotypic and genetic correlations between body weight and condition factor were reduced over time, i.e., between BWT and CFT (r_g = 0.50) vs. BWH and CFH (r_g = 0.23). This result is inconsistent with a study by Elvingson and Johansson (1993) who found an increasing genetic correlation between body weight and condition factor in trout from 1.5 to 2.5 years old. Genetic correlation between length and CFH tends to be lower than between BWH and CFH but length was not included in the genetic analysis. Fishback et al. (2002) found lower genetic correlation between length and CF (r_g = 0.22 at T = 8.5 ^{o}C, and r_g = -0.09 at T = 15 ^{o}C) than between BW and CF (r_g = 0.31 at T = 8.5 ^{o}C, and r_g = 0.03 at T = 15 ^{o}C).

4.4.4 Enhancing direct selection response in survival

Including traits genetically correlated with survival into multi-trait breeding value evaluations will improve selection accuracy for overall survival (SURV). Our results show that for indirect selection for SURV, the BWS (r_{IH} = 0.19) was a more informative trait than BWT (r_{IH} = 0.04). However, for the combined direct and indirect selection, BWT and BWS were not, but BWH (r_{IH} = 0.57) was more informative, compared to direct selection for SURV (r_{IH} = 0.54) (Table 4.5). The TGC_{TH} is also an informative trait whose inclusion in a selection index improves accuracy of direct selection for SURV from 0.54 to 0.58. When including all body weight measurements for combining direct and indirect selection, accuracy of selection for SURV is slightly higher (0.61) than accuracy from including TGC_{TH}.

In our study, mortality occurred gradually over time and the causes of mortality were not identified. Thus, survival reflects resistance and/or tolerance against potentially multiple unknown factors. It is well established that when mortality factors vary in time and space, also the heritability and genetic correlations of survival are not constant across time, space and species (Vehviläinen et al., 2008, 2010, 2012).

Table 4.5 Alternative selection indices and their accuracy of selection index (r_{IH}) for visceral percentage (VISW%), condition factor at harvest (CFH) and overall survival (SURV).

Scenario	r_{IH}		
	[1]Direct	[2]Combined	[3]Indirect
[4]H= -1*VISW%			
[5]I = VISW% (*Hypothetical direct selection*)	**0.63**		
I = BWH		0.63	0.17
I = GBWH		0.63	0.11
I = TGC_{TH}		0.63	0.17
I = BWT + BWS + BWH + GBWH		0.66	0.39
H = -1*CFH			
I = CFH *(Hypothetical direct selection)*	**0.66**		
I = CFT		0.66	0.40
I = BWT		0.66	0.12
I = BWS		0.66	0.07
I = BWH		0.66	0.13
I = TGC_{TH}		0.66	0.04
I = BWT + BWS + BWH + GBWH		0.66	0.19
H = 1*SURV			
I = SURV *(Hypothetical direct selection)*	**0.54**		
I = BWT		0.54	0.04
I = BWS		0.54	0.19
I = BWH		0.57	0.33
I = TGC_{TH}		0.58	0.36
I = BWT + BWS + BWH + GBWH		0.61	0.35

BWT=body weight at tagging, CFT = condition factor at tagging, BWS = body weight at sorting, BWH = ungutted body weight at harvest, GBWH = gutted body weight at harvest, CFH = condition factor at harvest, VISW% = visceral weight percentage, SURV = overall survival at harvest, and TGC_{TH} = thermal growth coefficient from tagging to harvest.

[1] = Direct selection, [2] = Combined direct and indirect selection, [3] = Indirect selection, [4] = Breeding objective, and [5] = Selection index.

In conclusion, selection for harvest body weight and overall thermal growth coefficient (TGC_{TH}) can be enhanced by postponing pre-selection to a later age. However, an optimal time point for tagging and pre-selection should be found to minimize common environmental effect and rearing cost. Including gutted weight in a selection index can reduce unfavourable selection responses in visceral percentage. TGC_{TH} and harvest body weight can be used to enhance selection for overall survival. To control genetic changes in condition factor, condition factor should be included to the selection index.

4.5 Acknowledgements

This study is part of the PhD project: Optimization of rainbow trout breeding program for multi-trait selection with GxE interaction. This project is funded by Ananda Mahidol Foundation (The King Scholarship, Thailand) and Troutlodge, Inc., USA. We thank Alexander and Johannes Tautenhahn at the Forellenzucht, Trostadt, Germany for running the experiment and collecting data on daily water temperature and dissolved oxygen. We are thankful to Robbert J. W. Blonk, Gyula Kovács, Katrijn Peeters, Chontida Phutaworn and Shripathi Bhat for data collection. We appreciate Gyula Kovács and Bert Dibbits for genotyping work and significant time they spent in the laboratory. We would like to thank Dr. Piter Bijma for his guidance on Sel*A*ction and Dr. Herman A. Mulder for his useful comments.

References

Austreng, E., Storebakken, T., Åsgård, T., 1987. Growth rate estimates for cultured Atlantic salmon and rainbow trout. Aquaculture 60, 157-160.

Cho, C.Y., 1992. Feeding systems for rainbow trout and other salmonids with reference to current estimates of energy and protein requirements. Aquaculture 100, 107-123.

Coulibaly, I., Gharbi, K., Danzmann, R.G., Yao, J., Rexroad, C.E., 2005. Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. Anim. Genet. 36, 309-315.

Crandell, P.A., Gall, G.A.E., 1993. The genetics of age and weight at sexual maturity based on individually tagged rainbow trout (*Oncorhynchus mykiss*). Aquaculture 117, 95-105.

Cunningham, E.P., 1975. Multi-stage index selection. Theor. Appl. Genet. 46, 55-61.

Dempster, E.R., Lerner, I.M., 1950. Heritability of threshold characters. Genetics 35, 212-236.

Duchesne, P., Godbout, M.-H., Bernatchez, L., 2002. PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. Mol. Ecol. Notes. 2, 191-193.

Elvingson, P., Johansson, K., 1993. Genetic and environmental components of variation in body traits of rainbow trout (*Oncorhynchus mykiss*) in relation to age. Aquaculture 118, 191-204.

FAO, 2011. Cultured aquatic species information programme *Oncorhynchus mykiss* (Walbaum, 1792): Fisheries and aquaculture department. Accessed May 18, 2011. http://www.fao.org/fishery/culturedspecies/Oncorhynchus_mykiss/en.

FAO, 2012. World aquaculture production of fish, crustaceans, molluscs, etc., by principal species in 2009: Fisheries and Aquaculture Statistics. Accessed April 10, 2012. ftp://ftp.fao.org/FI/CDrom/CD_yearbook_2009/navigation/index_content_aquaculture_e.htm.

Fishback, A.G., Danzmann, R.G., Ferguson, M.M., Gibson, J.P., 2002. Estimates of genetic parameters and genotype by environment interactions for growth traits of rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. Aquaculture. 206, 137-150.

Gall, G.A.E., Huang, N., 1988. Heritability and selection schemes for rainbow trout: body weight. Aquaculture 73, 43-56.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2009. ASReml User Guide Release 3.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Gjerde, B., Gjedrem, T., 1984. Estimates of phenotypic and genetic parameters for carcass traits in Atlantic salmon and rainbow trout. Aquaculture 36, 97-110.

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

Gjedrem, T., 1985. Improvement of productivity through breeding schemes. GeoJournal. 10, 233-241.

Gunnes, K., Gjedrem, T., 1981. A genetic analysis of body weight and length in rainbow trout reared in seawater for 18 months. Aquaculture 24, 161-174.

Hayes, J.F., Hill, W.G., 1981. Modification of Estimates of Parameters in the Construction of Genetic Selection Indices ('Bending'). Biometrics 37, 483-493.

Hokanson, K.E.F., Kleiner, C.F., Thorslund, T.W., 1977. Effects of Constant Temperatures and Diel Temperature Fluctuations on Specific Growth and Mortality Rates and Yield of Juvenile Rainbow Trout, *Salmo gairdneri*. J. Fish. Res. Board. Can. 34, 639-648.

Jobling, M., 2003. The thermal growth coefficient (TGC) model of fish growth: a cautionary note. Aquac. Res. 34, 581-584.Johnson, N.A., Rexroad Iii, C.E., Hallerman, E.M., Vallejo, R.L., Palti, Y., 2007. Development and evaluation of a new microsatellite multiplex system for parental allocation and management of rainbow trout (*Oncorhynchus mykiss*) broodstocks. Aquaculture 266, 53-62.

Kause, A., 2011. Genetic analysis of tolerance to infections using random regressions: A simulation study. Genet. Res. 93, 291-302.

Kause, A., Paananen, T., Ritola, O., Koskinen, H., 2007a. Direct and indirect selection of visceral lipid weight, fillet weight, and fillet percentage in a rainbow trout breeding program. J. Anim. Sci. 85, 3218-3227.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E., Eskelinen, U., 2002. Coupling body weight and its composition: a quantitative genetic analysis in rainbow trout. Aquaculture 211, 65-79.

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

Kause, A., Quinton, C., Airaksinen, S., Ruohonen, K., Koskela, J., 2011. Quality and production trait genetics of farmed European whitefish, *Coregonus lavaretus*. J. Anim. Sci. 89, 959-971.

Kause, A., Tobin, D., Mäntysaari, E.A., Martin, S.A.M., Houlihan, D.F., Kiessling, A., Rungruangsak-Torrissen, K., Ritola, O., Ruohonen, K., 2007b. Genetic potential for simultaneous selection of growth and body composition in rainbow trout (*Oncorhynchus mykiss*) depends on the dietary protein and lipid content: Phenotypic and genetic correlations on two diets. Aquaculture 271, 162-172.

Le Boucher, R., Quillet, E., Vandeputte, M., Lecalvez, J.M., Goardon, L., Chatain, B., Médale, F., Dupont-Nivet, M., 2011. Plant-based diet in rainbow trout (*Oncorhynchus mykiss* Walbaum): Are there genotype-diet interactions for main production traits when fish are fed marine vs. plant-based diets from the first meal? Aquaculture 321, 41-48.

Le Rouzic, A., Álvarez-Castro, J.M., Carlborg, Ö., 2008. Dissection of the genetic architecture of body weight in chicken reveals the impact of epistasis on domestication traits. Genetics. 179, 1591-1599.

Linder, D., Sumari, O., Nyholm, K., Sirkkomaa, S., 1983. Genetic and phenotypic variation in production traits in rainbow trout strains and strain crosses in Finland. Aquaculture 33, 129-134.

Mallet, J.P., Charles, S., Persat, H., Auger, P., 1999. Growth modelling in accordance with daily water temperature in European grayling (*Thymallus thymallus* L.). Can. J. Fish. Aquat. Sci. 56, 994-1000.

Maluwa, A.O., Gjerde, B., Ponzoni, R.W., 2006. Genetic parameters and genotype by environment interaction for body weight of *Oreochromis shiranus*. Aquaculture 259, 47-55.

Martinez, V., Neira, R., Gall, G.A.E., 1999. Estimation of genetic parameters from pedigreed populations: lessons from analysis of alevin weight in Coho salmon (*Oncorhynchus kisutch*). Aquaculture 180, 223-236.

Martinez, V., Kause, A., Mäntysaari, E., Mäki-Tanila, A., 2006. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: II. Two-stage selection. Aquaculture 254, 195-202.

Matthews, K.R., Berg, N.H., 1997. Rainbow trout responses to water temperature and dissolved oxygen stress in two southern California stream pools. J. Fish. Biol. 50, 50-67.

McKay, L.R., Gjerde, B., 1986. Genetic variation for a spinal deformity in Atlantic salmon. Aquaculture 52, 263–272.

Neira, R., Lhorente, J.P., Araneda, C., Diaz, N., Bustos, E., Alert, A., 2004. Studies on carcass quality traits in two populations of Coho salmon (*Oncorhynchus kisutch*): phenotypic and genetic parameters. Aquaculture 241, 117-131.

Nguyen, N.H., Ponzoni, R.W., Abu-Bakar, K.R., Hamzah, A., Khaw, H.L., Yee, H.Y., 2010. Correlated response in fillet weight and yield to selection for increased harvest weight in genetically improved farmed tilapia (GIFT strain), *Oreochromis niloticus*. Aquaculture 305, 1-5.

Ojolick, E.J., Cusack, R., Benfey, T.J., Kerr, S.R., 1995. Survival and growth of all-female diploid and triploid rainbow trout (*Oncorhynchus mykiss*) reared at chronic high temperature. Aquaculture 131, 177-187.

Palti, Y., Fincham, M.R., Rexroad, C.E., 2002. Characterization of 38 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Mol. Ecol. Notes. 2, 449-452.

Pollak, E.J., van der Werf, J., Quaas, R.L., 1984. Selection Bias and Multiple Trait Evaluation. J. Dairy. Sci. 67, 1590-1595.

Quinton, C.D., McMillan, I., Glebe, B.D., 2005. Development of an Atlantic salmon (*Salmo salar*) genetic improvement program: Genetic parameters of harvest body weight and carcass quality traits estimated with animal models. Aquaculture 247, 211-217.

Rexroad, C.E., Rodriguez, M., Coulibaly, I., Gharbi, K., Danzmann, R., DeKoning, J., Phillips, R., Palti, Y., 2005. Comparative mapping of expressed sequence tags containing microsatellites in rainbow trout (*Oncorhynchus mykiss*). BMC Genomics 6, 54.

Rexroad, C.E., Coleman, R.L., Hershberger, W.K., Killefer, J., 2002. Eighteen polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Anim. Genet. 33, 76-78.

Ruiz-Flores, A., Johnson, R.K., 2001. Direct and correlated responses to two-stage selection for ovulation rate and number of fully formed pigs at birth in swine. J. Anim. Sci. 79, 2286-2297.

Rutten, M.J.M., Komen, H., Bovenhuis, H., 2005a. Longitudinal genetic analysis of Nile tilapia (*Oreochromis niloticus* L.) body weight using a random regression model. Aquaculture 246, 101-113.

Rutten, M.J.M., Bovenhuis, H., Komen, H., 2005b. 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., van Arendonk, J.A.M., 2002. Sel*A*ction: Software to Predict Selection Response and Rate of Inbreeding in Livestock Breeding Programs. J. Hered. 93, 456-458.

Rye, M., Refstie, T., 1995. Phenotypic and genetic parameters of body size traits in Atlantic salmon (*Salmo Salar* L). Aquac. Res. 26, 875-885.

Sae-Lim, P., Komen, H., Kause, A., van Arendonk, J.A.M., Barfoot, A.J., Martin, K.E., Parsons, J.E., 2012. Defining desired genetic gains for rainbow trout breeding objective using analytic hierarchy process. J. Anim. Sci. 90, 1766-1776.

Silverstein, J.T., Vallejo, R.L., Palti, Y., Leeds, T.D., Rexroad, C.E., III, Welch, T.J., Wiens, G.D., Ducrocq, V., 2009. Rainbow trout resistance to bacterial cold-water disease is moderately heritable and is not adversely correlated with growth. J. Anim. Sci. 87, 860-867.

Su, G.S., Liljedahl, L.E., Gall, G.A.E., 2002. Genetic correlations between body weight at different ages and with reproductive traits in rainbow trout. Aquaculture 213, 85-94.

Sullivan, M., G. Hammond, R. J. Roberts and N. J. Manchester. 2007. Spinal deformation in commercially cultured, *Salmo salar* L.: a clinical and radiological study. J. Fish. Dis. 30, 745-752.

Thompson, R., Meyer, K., 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest. Prod. Sci. 15, 299-313.

Vehviläinen, H., Kause, A., Kuukka-Anttila, H., Koskinen, H., Paananen, T., 2012. Untangling the positive genetic correlation between rainbow trout growth and survival. Evol. Appl., in press.

Vehviläinen, H., Kause, A., Quiton, C., Kuukka-Anttila, H., Koskinen, H., Paananen, T., 2010. Genetic architecture of rainbow trout survival from egg to adult. Genet. Res. 92, 1-11.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H., Paananen, T., 2008. Survival of the currently fittest: Genetics of rainbow trout survival across time and space. Genetics 180, 507-516.

5

Genotype-by-environment interaction of growth traits in rainbow trout (*Oncorhynchus mykiss*): A continental scale study

P. Sae-Lim[1], H. Komen[1], A. Kause[2], H. A. Mulder[1], J. A.M. van Arendonk[1], K. E. Martin[3], A. J. Barfoot[3], J. E. Parsons[3], J. Davidson[4], C. E. Rexroad III[5]

[1]Animal Breeding and Genomics Centre, Wageningen University, P. O. Box 338, 6700 AH, Wageningen, The Netherlands; [2]Agrifood Research Finland, Biotechnology and Food Research, Biometrical Genetics, FI-31600 Jokioinen, Finland; [3]Troutlodge, Inc., 12000 McCutcheon Rd., Sumner, WA 98391, USA; [4]The Conservation Fund's Freshwater Institute, 1098 Turner Road, Shepherdstown, WV 25443, USA; [5]National Center for Cool and Cold Water Aquaculture, ARS-USDA, 11861 Leetown Road, Kearneysville, WV 25430, USA

Submitted to Journal of Animal Science

Abstract

Rainbow trout is a globally important fish species for aquaculture. However, fish for most farms worldwide are produced by only a few breeding companies. Selection based solely on fish performance recorded at a nucleus may lead to lower-than-expected genetic gains in other production environments when genotype-by-environment (GxE) interaction exists. GxE interaction has 2 forms: genotype re-ranking measured by the genetic correlation between environments and heterogeneity of genetic variances across environments. The aim was to quantify the magnitude of GxE interaction of growth traits (tagging weight; BWT, harvest weight; BWH, and growth rate; TGC) measured across 4 environments, located in 3 different continents. A total of 100 families, of size 25, were produced from the mating of 58 sires and 100 dams. In total, 13,806 offspring were reared at the nucleus in Washington State (NUC) and in 3 other environments: a recirculating aquaculture system in West Virginia, USA (FI), a high-altitude farm in Peru (PE), and a cold-water farm in Germany (GE). To account for selection bias due to selective mortality, a multi-trait multi-environment animal mixed model was applied to the performance data. Genetic correlation (r_g) of a trait measured in different environments and r_g of different traits measured in different environments were estimated. The results show that heterogeneity of additive genetic variances was mainly found for BWH measured in FI and PE. Additive genetic coefficient of variation for BWH in NUC, FI, PE and GE was 7.63, 8.36, 8.64, and 9.75, respectively. Strong genotype re-ranking was found for all traits (BWT: r_g = 0.15 to 0.37, BWH: r_g = 0.19 to 0.48, TGC: r_g = 0.31 to 0.36) across environments. The r_g between BWT in NUC and BWH in both FI (0.31) and GE (0.36) were positive, which was also found between BWT in NUC and TGC in both FI (0.10) and GE (0.20). However, r_g were negative between BWT in NUC and both BWH (-0.06) and TGC (-0.20) in PE. Correction for selection bias resulted in higher additive genetic variances. In conclusion, strong GxE interaction was found for BWT, BWH, and TGC. Accounting for GxE interaction in the breeding program either by using sib-information from testing stations or environment-specific breeding programs would increase genetic gains for environments that differ significantly from NUC.

Key words: heterogeneous genetic variance, multi-trait multi-environment, re-ranking, scaling effect, selection bias, thermal growth coefficient

5.1 Introduction

Rainbow trout is a globally important species for aquaculture. It is produced under very diverse production conditions, such as different altitudes, water qualities, and farming managements. In addition, the market size differs across production systems, e.g., from 300 g portion-sized fish to 2-3 kg large trout. However, a breeding company may distribute trout from a single breeding program to these diverse production and market conditions. Trait recording and selection practiced at a single nucleus station may lead to lower-than-expected genetic gains in other production environments when genotype-by-environment (GxE) interaction exists (Mulder and Bijma, 2005). GxE interaction has different forms: re-ranking across environments and heterogeneity of genetic variances (Falconer and Mackay, 1996; Lynch and Walsh, 1998). In addition, genetic correlations between pairs of traits may differ between environments (Calus, 2006; Mulder, 2007).

A survey among rainbow trout farmers revealed that growth was the most preferred trait among 13 traits (Sae-Lim et al., 2012a). Body weight and thermal growth coefficient (TGC) are two complementary traits for improving growth rate. Body weight is a measure of cumulative growth over lifetime whereas TGC measures growth rate corrected for water temperature during a specific growth period. In rainbow trout, weak to moderate GxE interaction has been found for body weight and TGC (Fishback et al., 2002; Kause et al., 2003; Kause et al., 2006; Le Boucher et al., 2011; Pierce et al., 2008). However, production systems located in different continents can differ greatly in temperature, altitude, photoperiod, water used and feeding, which may result in stronger GxE interaction. In this study, the aim was to quantify the magnitude of GxE interaction of growth traits in the forms of re-ranking, heterogeneity of genetic variances, heritabilities, and correlations between traits across 4 different production environments, located in 3 continents.

5.2 Materials and methods

5.2.1 GxE experiment

The fish used in this study were all-female offspring, obtained from Troutlodge, Inc., Washington State, USA. Troutlodge, Inc. is a global breeding company, shipping salmonid eggs to more than 60 countries around the world. In August 2009, a total of 58 sex-reversed XX sires and 100 dams were mated to produce 100

Table 5.1 Environmental parameters measured during the genotype-by-environment interaction experiment.

Environmental parameter	NUC[a]	FI[b]	PE[c]	GE[d]
Hatching	Sept, 2009	Sept to Oct, 2009	Sept, 2009	Oct, 2009
Tagging	Jan, 2010	Jan, 2010	Mar, 2010	Apr, 10
Harvest	Jun& Jul, 2010	Jul, 2010	Sept, 2010	Dec, 2010
No. of fish at harvest	2372	2243	2890	1992
Age at tagging (dph[e)]	118 to 120	118 to 120	162 to 166	193 to 195
Age at harvest (dph)	280 to 295	294 to 296	357 to 359	445 to 446
Avg[f]. dissolved oxygen (mg/l)	7.30	10.20	6.63	10.0
Avg. water temperature (^{o}C)	13.4	11.8	13.4	9.9
Feeding (% BW)	7.3 to 1.2%	11.2 to 1.2%	3.61 to 0.5%	2.5 to 0.5%
Protein%	44 to 53%	42 to 55%	42 to 50%	42 to 64%
Fat%	16 to 25%	15 to 16%	13 to 15%	11 to 30%
Photoperiod* (minute)	223.1	163.3	-53.1	292.9
Altitude** (Above sea level; m)	25	129	3812	361
Recirculation*** (%)	0%	85%	0%	65%
Captive environment	2 flow-through raceways	2 partial reused circular dual drain culture tanks	net pen submerged in Titicaca Lake	outside pond

[a]NUC = nucleus, [b]FI = recirculating aquaculture system, [c]PE = high elevated farm, and [d]GE = low temperature farming, [e]day post hatch, [f]average. *Photoperiod was calculated from the difference between the highest day length (minute) and average day length from overall rearing period. Day length was calculated from the difference between sunrise and sunset in minute. The sunrise and sunset data (option: actual time) was assessed from http://www.wunderground.com/history/. The negative sign indicates different directions of the change in day length. **Altitude of each location was obtained from; http://www.daftlogic.com/sandbox-google-maps-find-altitude.htm. *** Recirculation aquaculture system in FI and re-sued water system in GE.

full-sib families. Each sire was mated to 1 to 3 dams (average =1.7), and each dam was mated to 1 sire. Production of families took place over a period of 4 weeks. Fertilized eggs from each of the 4 spawning weeks were incubated using different water temperatures, resulting in all groups hatching at approximately the same time. Fertilized eggs were incubated in 100 incubators (1 for each family) until the eyed-egg stage.

In September 2009, groups of 25 eyed-eggs from each family were randomly sampled and pooled into a batch. In total, 5 batches with 100 families of family size 25 were generated (a total of 12,500 eyed-eggs). The number of families and the family size were based on the guidelines of a simulation study (Sae-Lim et al., 2010). Batch no. 1 was shipped to The Freshwater Institute, West Virginia, USA (FI) and grown in a recirculating aquaculture system. Batches no. 2 to 3 were shipped to Parsiri and Huancayo farms, a high-altitude farm in Lake Titicaca in Peru (PE). Batch no. 4 was split by approximately half in case of poor survival. In total, Pasiri received a total of 3,743 eggs and Huancayo received a total of 3,757 eggs. Due to flooding in November 2009, all fish in Huancayo farm were lost. Batch no. 5 was shipped to Forellenzucht Trostadt in Germany (GE), a farm characterized by year-round low water temperatures. Finally, 600 randomly sampled eyed-eggs from each family were shipped to Troutlodge's Eastern Washington facility (NUC). In NUC, the 600 eggs were hatched and from the surviving sack fry, 50 were randomly chosen for hormone-treated to obtain sex-reversed XX-males for the breeding program. The remaining all female groups was grown to 25g at which point 32 fish per family were pre-selected as selection candidates for the breeding program. The pre-selection was within family selection, based on the phenotype of body weight at tagging (BWT). Subsequently, 25 fish from each family were randomly sampled from the approximately remaining 520 fish for inclusion in the study.

The environmental conditions of the 4 farms are given in Table 5.1. The farms had been selected to represent extremes in rearing conditions. In brief, the German farm was chosen as an example of a low water temperature farm; the Peru farm was chosen for its location at 3812 m above sea-level, and the Freshwater institute was chosen as being representative for a recirculating aquaculture system (RAS).

5.2.2 Pedigree reconstruction

The fish were tagged using passive integrated transponders (PIT tag; Allflex USA, Inc. for NUC, FI and PE, and DORSET Identification b.v., the Netherlands, for GE) and the PIT tag scanned (scanner SF2001ISO: Destron Fearing, USA for NUC, FI and PE, and GR250: DORSET Identification b.v., the Netherlands, for GE) at the average

size of 26.3 to 33.2g (5 to 7 months of age) (Table 5.2). Before tagging, fish were anesthetized using MS222 (150 mg/l) in NUC, FI, and PE and using clove oil (10 mg/l) in GE. Fin clips were collected from all 158 broodstock fish and from fish at tagging from FI, PE and GE for DNA extraction. In NUC, fish were kept in separated family tanks until tagging; therefore fin clips were not collected in NUC.

DNA was isolated from the fin-clips to reconstruct the pedigree. Genotyping of the DNA samples was done in 3 laboratories: National Center for Cool and Cold Water Aquaculture, United States Department of Agriculture; Troutlodge, Inc.; and Animal Breeding and Genomics Centre, Wageningen University. The protocols for DNA isolation and genotyping were synchronized across the labs. In brief, the DNA isolation was done using Nucleospin® 96 Tissue Core Kit. Multiplex PCR amplification was done as described by Johnson et al. (2007). Fragment analysis of the PCR products was done by setting the fragment sizes relatively to Genescan LIZ 500 size standard (Applied Biosystem). Output data were analysed using Genemapper software version 4 (Applied Biosystem) (Sae-Lim et al., 2012b).

Parental allocation was performed using PAPA software (Duchesne et al., 2002). The known mating data were used to increase the accuracy of parental assignments (Sae-Lim et al., 2012b). In total, 2,142 out of 2,243 fish sampled in FI, 3,106 out of 3,236 fish sampled in PE, and 2,104 out of 2,235 fish sampled in GE were successfully allocated to the 100 full-sib families. Fish that were not successfully allocated to the families were removed from the dataset.

In total, 6 generations of pedigree information used in the genetic analysis were from the DNA reconstructed pedigree and from the 5 previous generations of pedigree information.

5.2.3 Trait measurement

During tagging, fish in all environments were measured for body weight (BWT, in units of g). All surviving fish were measured for body weight at harvest (BWH, in g), which is the round weight prior to any processing. The age at harvest ranged from 9 months in NUC to 14 months in GE (Table 5.1).

Thermal growth coefficient from tagging to harvest (TGC) was calculated as

$$\left[\frac{(\sqrt[3]{BWH}-\sqrt[3]{BWT})}{T\times t}\right]\times 1000, \qquad [1]$$

where T = average water temperature (°C), and t = rearing period in days. To correct for the non-linear relationship between growth rate and water temperature (Jobling, 2003), formula TGC was modified to

$$\left[\frac{(\sqrt[3]{BWH}-\sqrt[3]{BWT})}{k\times t}\right]\times 1000, \quad [2]$$

by substituting T with k calculated from the model used by Mallet et al. (1999):

$$k=\frac{T_{opt}(T-T_{min})(T-T_{max})}{(T-T_{min})(T-T_{max})-(T-T_{opt})^2}, \quad [3]$$

where k = new temperature, corrected for the concave relationship between growth rate and temperature. The optimum water temperature (T_{opt}) was set to 14.8 °C, which was calculated as the average optimal water temperature for salmonid growth (Austreng et al., 1987; Hokanson et al., 1977; FAO, 2011). Daily water temperature: T was from the daily measurement at a farm. The limits for the lower and upper thermal tolerance: T_{min} = 0 °C, and T_{max} = 23 °C, respectively, were taken from the literature (Hokanson et al., 1977; Matthews and Berg, 1997; Ojolick et al., 1995).

5.2.4 Genetic analysis

Heritability (h^2), common environmental effect for full sibs (c^2), phenotypic (r_p) and genetic (r_g) correlations were estimated using restricted maximum likelihood in an animal mixed model in ASReml v. 3.0 (Gilmour et al., 2009).

Heritability

Significant fixed effects were tested in SAS v 9.2 using PROC GLM. The fixed effects tested were different across environments due to different data structure. Thus the final model for different environments varied and included only the significant effects.

In NUC, each trait was modelled as:

$$y_{ijklm} = \mu + Sex_i + \beta^{*}AGE_j + FERT_k + a_l + FS_m + e_{ijklm,} \quad [4]$$

In FI, each trait was modelled as:

$$y_{jklmn} = \mu + \beta * AGE_j + FERT_k + Tank_n + a_l + FS_m + e_{jklmn,} \quad [5]$$

In PE and GE, each trait was modelled as:

$$y_{jklm} = \mu + \beta * AGE_j + FERT_k + a_l + FS_m + e_{jklm,} \quad [6],$$

where y is the observation of the lth individual from the mth full-sib family, μ is the overall mean, Sex is the fixed effect corrected for gender of observation (i =1: male, 2: female, 9: unknown). The Sex effect was only modelled in the NUC for body weight at tagging, as we included BWT of the selection candidates in the dataset. Otherwise this effect was omitted. The β is a regression coefficient of fixed effect AGE_j. The fixed effect AGE_j was included as a covariate in the model to correct for different measurement dates within environment, and corrected for rearing periods from hatching to the day of trait measurement (Table 5.2). For TGC, AGE was not included in the model because TGC is already corrected for the rearing period. $Tank$ is the fixed effect for BWT due to the 2 circular tanks used in FI for stocking fish from fingerling up to tagging (j =1, 2). $FERT$ is the fixed effect corrected for fertilization period of 4 weeks (k = 1, 2, 3, or 4) due to different groups of available fertile dams. The a_l is the random additive genetic effect, a ~N(0, $\boldsymbol{A}\sigma_a^2$) of the lth animal, where $\boldsymbol{A}$ is the additive genetic relationship matrix and σ_a^2 is the additive genetic variance. FS_m is the random full-sib common environmental effect, FS ~ N(0, $\boldsymbol{I}\sigma_{FS}^2$), and e is the random error term, e ~N(0, $\boldsymbol{I}\sigma_e^2$), where $\boldsymbol{I}$ is the identity matrix, σ_{FS}^2 is the common environmental variance and σ_e^2 is the residual variance. The full-sib effect was included into the model to account for effects common to full-sibs, e.g. incubator effects, environmental maternal effects, and a quarter of the dominance variance.

Univariate analysis was performed for each trait to test for the significance of common environmental effect. The models with and without the full-sib effect were compared using likelihood ratio test (LRT). The LRT = $-2[\ln(L)_r - \ln(L)_f]$, where $\ln(L)_r$ and $\ln(L)_f$ are natural logarithm of likelihood from the reduced model (without full-sib effect) and the full model (with full-sib effect), respectively (Lynch and Walsh, 1998). The asymptotic distribution of likelihood ratio follows Chi-square (χ^2) distribution with a mixture (50:50) of degrees of freedom between 0 and 1 (Stram and Lee, 1994). The 5% significance level was therefore χ^2 = 2.706.

After the LRT, h^2 and c^2 were estimated using a bivariate model. Selection bias (Henderson, 1984; Pollak et al., 1984; Ouweltjes et al., 1988) due to selective mortality was accounted for by always including BWT of each environment as a reference trait in the bivariate model (Kause et al., 2011). Full-sib effect was always included in the bivariate model to avoid overestimated h^2. Heritability from the model with full-sib effect was quantified as $h^2 = V_A/(V_A + V_{FS} + V_R)$, where V_A, $V_{FS,}$ and V_R are estimated additive genetic, estimated full-sib and estimated residual variances. The common environmental effect was calculated as $c^2 = V_{FS}/(V_A + V_{FS} + V_R)$. In addition, variation across environments was compared by estimating phenotypic $(CV_P = (SD_P/\bar{X}) \times 100)$, genetic $(CV_A = (SD_A/\bar{X}) \times 100)$ and residual $(CV_R = (SD_R/\bar{X}) \times 100)$ coefficients of variation. The SD_P, SD_A, and SD_R are phenotypic, genetic, and residual standard deviations, respectively. Values were obtained from the models 1 to 3: $\bar{X}$ is phenotypic trait mean. The V_A and CV_A were used to quantify the degree of heterogeneous genetic variation across environments.

Phenotypic and genetic correlations

Three types of genetic correlations were estimated: *a*) genetic correlations of different traits within an environment, *b*) genetic correlation of a trait measured in different environments (measure of genotype re-ranking), and *c*) genetic correlations of different traits in different environments.

To estimate all types of genetic correlations simultaneously, we performed a multi-trait multi-environment (MTME) analysis using a multivariate animal mixed model. The first MTME model contained 3 traits measured in 4 environments, but ASREML had difficulty in estimating the parameters. Therefore, the size of a single MTME model was reduced to 2 traits and 4 environments (a total of 8 traits). The full-sib effect was excluded from the model because in many cases, the full-sib effect captured all the (co)variance of the traits (Maluwa et al., 2006). Residual (co)variances of the same trait and different traits, measured in different environments were set to zero:

$$\text{VAR}(e) = \begin{bmatrix} R_{T1,E1} & & \cdots & \\ R_{T12,E1} & R_{T2,E1} & & Symmetry \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & R_{T1,E4} \\ 0 & 0 & & R_{T12,E4} \quad R_{T2,E4} \end{bmatrix},$$

where $R_{T1,E1}$ is the residual variance of trait: T_1 measured in environment: E_1. $R_{T12,E1}$ is the residual covariance between T_1 and T_2 measured in E_1. Therefore, phenotypic correlations (r_p) were only calculated between the traits measured within the same environment.

After estimating all variance components, phenotypic and genetic correlation matrices were bended to be positive definite (Hayes and Hill, 1981) in Octave computer software (A. Kause, MTT, Finland: personal communication). The bending induced only minor changes in phenotypic (range: 0 to 0.005) and genetic (range: -0.016 to 0.068) correlation estimates. The bended estimates were presented.

Effect of selection bias

To study the effect of selection bias on V_A and V_R estimates, a comparison of V_A and V_R from 2 models was made. These models were *i*) multivariate model for BWH measured in 4 environments, and *ii*) MTME model for BWT and BWH measured in 4 environments. These models did not include the full-sib effect, to enhance the comparison of the models. Simultaneous estimation of V_A and V_{FS} typically creates discrepancy in V_A across the compared models because of the difficulty of accurately estimating the two at the same time.

5.3 Results

5.3.1 Genotype-by-environment interaction

Heterogeneity of genetic variation

The V_A in BWT ranged from 11.98 to 17.63 (Table 5.2). In contrast, V_A of BWH in PE (2054.06) was twice as high as V_A of BWH in FI (1092.29). However, the CV_A in BWH was very similar in PE (8.64) and FI (8.36), suggesting that the variances differed because of the differences in trait means. Similarly, V_A and CV_A of BWH in NUC was 1742.67 and 7.63 whereas V_A and CV_A of BWH in GE was 1345.45 and 9.75. For TGC, CV_A varied between environments from 2.79 in PE to 6.35 in GE.

Table 5.2 Mean and its standard deviation (SD), phenotypic (V_p), genetic (V_A) and residual (V_R) variance estimates, phenotypic (CV_p), genetic (CV_A) and residual (CV_R) coefficients of variance, heritability (h^2), common environmental effect (c^2), and their standard error (SE) for growth traits in each production environment. The estimates were from bivariate analysis.

Trait [a]	Environment [b]	N	Mean	SD	V_P	V_A	V_R	CV_P	CV_A	CV_R	h^2	SE (h^2)	c^{2c}	SE (c^2)
BWT	NUC	6448	33.15	5.96	36.70	14.65	19.23	18.27	11.55	13.23	0.40	0.15	**0.08**	0.05
	FI	2138	26.26	6.20	40.51	17.63	20.62	24.24	15.99	17.29	0.44	0.15	0.06	0.05
	PE	3179	29.15	5.81	33.88	13.56	19.72	19.97	12.63	15.24	0.40	0.13	0.02	0.04
	GE	2041	27.06	6.62	44.57	11.98	29.27	24.67	12.79	20.00	0.27	0.13	**0.07**	0.05
BWH	NUC	2364	546.82	94.70	9035.40	1742.67	6749.19	17.38	7.63	15.02	0.19	0.10	**0.06**	0.03
	FI	1893	395.15	75.84	6127.50	1092.29	4589.94	19.81	8.36	17.15	0.18	0.11	**0.07**	0.04
	PE	2795	524.28	105.17	11212.00	2054.06	8682.51	20.20	8.64	17.77	0.18	0.09	0.04	0.03
	GE	1819	376.39	81.72	6148.90	1345.45	4715.37	20.83	9.75	18.24	0.22	0.09	0.01	0.03
TGC	NUC	2364	2.07	0.18	0.03	0.009	0.022	8.37	4.47	7.14	0.27	0.12	0.04	0.04
	FI	1891	1.73	0.16	0.03	0.003	0.022	10.01	2.96	8.48	0.10	0.08	**0.07**	0.03
	PE	2790	1.75	0.20	0.04	0.002	0.034	11.43	2.79	10.58	0.06	0.06	**0.05**	0.03
	GE	1818	1.43	0.19	0.04	0.008	0.029	13.99	6.35	11.84	0.22	0.09	0.01	0.03

[a]BWT = body weight at tagging, BWH = harvest body weight, TGC = thermal growth coefficient with mallet correction. [b]NUC = nucleus, FI = recirculating aquaculture system, PE = high elevated farm, and GE = low temperature farming. [c]Bold letter indicates significant effect when using Likelihood Ratio Test (LRT) ~χ^2 with mixture of degrees of freedom (50:50) between 0 and 1, $\alpha = 0.05$.

Table 5.3 Phenotypic (r_p) and genetic (r_g) correlations and their standard error (SE) between different traits measured within environment.

Trait [a]	Environment [b]	r_p ± SE	r_g ±SE
BWT-BWH	NUC	0.56 ± 0.02	0.47 ± 0.09
	FI	0.64 ± 0.02	0.65 ± 0.07
	PE	0.50 ± 0.02	0.58 ± 0.08
	GE	0.36 ± 0.03	0.41 ± 0.12
BWT-TGC	NUC	0.13 ± 0.03	-0.15 ± 0.12
	FI	0.22 ± 0.03	0.13 ± 0.13
	PE	0.16 ± 0.02	0.20 ± 0.13
	GE	-0.09 ± 0.03	-0.14 ± 0.14
BWH-TGC	NUC	0.90 ± 0.01	0.71 ± 0.05
	FI	0.88 ± 0.01	0.72 ± 0.05
	PE	0.92 ± 0.00	0.81 ± 0.03
	GE	0.88 ± 0.01	0.75 ± 0.05

[a]BWT = body weight at tagging, BWH = harvest body weight, TGC = thermal growth coefficient with Mallet correction. [b]NUC = nucleus, FI = recirculating aquaculture system, PE = high elevated farm, and GE = low temperature farming.

Heterogeneity of heritabilities

Heritability for BWT was similar in NUC (0.40), FI (0.44), and PE (0.40) but lower in GE (0.27) (Table 5.2). The lower estimate of h^2 in GE was due to lower V_A (11.98) and higher V_R (29.27) compared to the other environments. For BWH, h^2 ranged from 0.18 to 0.22. For TGC, h^2 was heterogeneous across environments (0.06 to 0.27). In addition, c^2 for TGC was significant in FI (0.07) and PE (0.05), indicating some effects common to full-sibs beyond additive genetic effects.

Heterogeneity of within environment correlations

The r_g between BWT and BWH were heterogeneous, especially between FI (r_g = 0.65) and GE (r_g = 0.41) but less so between NUC (r_g = 0.47) and PE (r_g = 0.58) (Table 5.3). Similarly, r_g between BWT and TGC showed heterogeneity between on one hand FI (r_g = 0.13) and PE (r_g = 0.20) and on the other hand GE (r_g = -0.14) and NUC (r_g = -0.15). In contrast, r_g between BWH and TGC tended to be more homogeneous across environments, r_g ranged from 0.71 to 0.81.

Genetic correlation for same trait across environments

Genetic correlation of BWT measured in NUC and the 3 production environments ranged from 0.15 (PE) and 0.37 (GE) (Table 5.4). Genetic correlation of BWH measured in NUC and the 3 production environments ranged from 0.19 (PE) to 0.48 (GE). Genetic correlation for TGC measured in NUC and the 3 production environments ranged from 0.31 (PE) to 0.36 (GE).

Moderate r_g of all traits was found among the production environments in FI, PE and GE. The r_g of BWT ranged from 0.55 to 0.65. Lower r_g were found for BWH (0.40 to 0.51) and TGC (0.32 to 0.42). Overall, the results indicated strong re-ranking across environments for BWT, BWH, and TGC.

Genetic correlation between different traits measured in different environments

Genetic correlations between BWH in NUC and TGC in FI (0.44), GE (0.36) and PE (0.19) were all positive, showing that single trait selection for BWH in NUC will lead to favourably correlated response for TGC across environments (Table 5.5).

Table 5.4 Genetic correlation and its standard error (± SE) for genotype-by-environment (GxE) interaction for growth traits.

Trait[a]	Environment[b]	Environment		
		FI	PE	GE
BWT	NUC	0.34 ± 0.10	0.15 ± 0.11	0.37 ± 0.10
	FI		0.58 ± 0.08	0.65 ± 0.07
	PE			0.55 ± 0.09
BWH	NUC	0.41 ± 0.11	0.19 ± 0.13	0.48 ± 0.12
	FI		0.40 ± 0.12	0.51 ± 0.12
	PE			0.43 ± 0.12
TGC	NUC	0.35 ± 0.13	0.31 ± 0.13	0.36 ± 0.13
	FI		0.32 ± 0.14	0.42 ± 0.14
	PE			0.34 ± 0.14

[a]BWT = body weight at tagging, BWH = harvest body weight, TGC = thermal growth coefficient with mallet correction. [b]NUC = breeding environment, FI = recirculating aquaculture system, PE = high elevated farm, and GE = low water temperature farming.

On the other hand, r_g between BWT in NUC and BWH in FI (0.31) and in GE (0.36) were positive, but weakly negative with BWH in PE (-0.06). Similarly, r_g between BWT in NUC and TGC in FI (0.10) and in GE (0.20) were positive but negative with

TGC in PE (-0.20), indicating that selection on BWT in NUC may lead to different directions and size of correlated responses for BWH and TGC across environments.

Effect of selection bias

Overall, including BWT in the multi trait analysis resulted in higher estimates of V_A and V_R for BWH than that from multivariate model without BWT (Table 5.6). This suggested it is important to include BWT in multiple trait analysis to avoid selection bias in estimates for BWH.

Table 5.5 Genetic correlation and its standard error (± SE) between different traits measured in different environments.

Environment[a]	Trait[b]	Environment		
		FI		
		BWT	BWH	TGC
NUC	BWT	-	0.31 ± 0.11	0.10 ± 0.13
	BWH	0.21 ± 0.12	-	0.44 ± 0.13
	TGC	-0.04 ± 0.13	0.32 ± 0.13	-
Environment	Trait	PE		
		BWT	BWH	TGC
NUC	BWT	-	-0.06 ± 0.12	-0.20 ± 0.13
	BWH	0.14 ± 0.12	-	0.19 ± 0.14
	TGC	0.05 ± 0.13	0.27 ± 0.13	-
Environment	Trait	GE		
		BWT	BWH	TGC
NUC	BWT	-	0.36 ± 0.12	0.20 ± 0.13
	BWH	0.21 ± 0.12	-	0.36 ± 0.14
	TGC	-0.03 ± 0.13	0.33 ± 0.13	-

[a]NUC = nucleus, FI = recirculating aquaculture system, PE = high elevated farm, and GE = low temperature farming. [b] BWT = body weight at tagging, BWH = harvest body weight, TGC = thermal growth coefficient with mallet correction.

5.4 Discussion

5.4.1 Genotype-by-environment interaction

GxE interaction can have different consequences: re-ranking of breeding values or genotypes across environments, heterogeneous genetic variation across environments (also known as scaling effect), heterogeneous heritabilities, and heterogeneous correlations between traits (measured within environment) across environments (Lynch and Walsh, 1998; Calus, 2006; Mulder, 2007). Statistically, re-ranking is absent when genetic correlation (r_g) of a trait measured in different environments does not differ from 1. However, in practice, the presence of re-ranking is commonly considered unimportant when $r_g \geq 0.8$ (Robertson, 1959). Re-ranking is more serious than heterogeneity of genetic variance because re-ranking means that a single genotype is not superior across all environments (Calus, 2006; Mulder, 2007). For example, in dairy cattle genetic evaluation, it is important to account for heterogeneity of genetic variance between farms to accurately estimate breeding values (EBV) when farms differ in variance (e.g. Hill, 1984; Meuwissen et al., 1996). However, in fish, heterogeneity of genetic variation is less important because selection candidates are located in a single environment and multiple environments are treated as genetically different traits.

Table 5.6 Additive genetic (V_A) and residual (V_R) variances of BWH* from two different models: multivariate model with 4 traits (BWH measured in 4 environments) and multi-trait multi-environment (MTME) model with 8 traits (BWT** and BWH measured in 4 environments).

Environment[a]	Multivariate		MTME	
	V_A	V_R	V_A	V_R
NUC	3304.17	5250.71	3552.10	5869.99
FI	2404.87	3687.04	2525.57	3896.43
PE	3558.27	7758.52	3812.78	7857.38
GE	1637.63	4537.64	1654.06	4563.28

[a]NUC = nucleus, FI = recirculating aquaculture system, PE = high elevated farm, and GE = low temperature farming. *body weight at harvest. **body weight at tagging.

In our study, heterogeneity of additive genetic variation, heritabilities and correlations across environments was also found. High re-ranking between NUC and other environments was found for all traits, but re-ranking was stronger for

TGC than for BWH, especially between NUC and FI, and between NUC and GE. The BWH differed between environments due to variation in age at harvest which resulted from the differences in local market objectives. These differences in age at harvest have influenced r_g estimates between environments. The BWH is the cumulative result of growth from hatching to harvest and there is a common period between hatching to harvest across environments. In contrast, TGC is a more dynamic trait than BWH because TGC was calculated for specific grow-out period, i.e., between BWT and BWH measured at different ages. Consequently, it is expected that re-ranking in time (Rutten et al., 2005; Sae-Lim et al., 2012b) and between environments is higher in TGC than in BWH. Higher re-ranking across environments in TGC is in the agreement with a previous study in European seabass (*Dicentrachus labrax*) which daily gain coefficient (DGC: r_g = 0.21 to 0.61) had higher re-ranking than harvest body weight (r_g>0.80) (Dupont-Nivet et al., 2010).

Re-ranking has been studied in different livestock species for multiple environments. For different locations, in Atlantic cod, weak re-ranking (r_g = 0.82 to 0.94) for 2-year body weight measured in 3 different locations off the coast of Norway was found (Kolstad et al., 2006). In rainbow trout, moderate GxE exists (r_g = 0.61) between fresh and brackish water environments in body weight measured at 2 years of age (Kause et al., 2003). In tilapia (Oreochromis shiranuis) grown at different altitudes, (Maluwa et al., 2006) weak re-ranking (r_g = 0.74) for body weight measured between high and low altitudes has been reported. In contrast to tilapia, Colorado Angus cattle weaning weight, for example, measured at high, medium, and low altitude showed moderate to weak re-ranking (r_g = 0.47 to 0.83) (Williams et al., 2012). Under partially controlled environment, weak re-ranking was found in slow-growing chickens for 8 week body weight (r_g = 0.74 to 0.98) and body weight at slaughter (r_g = 0.76 to 0.97), and initial specific growth rate in chicken (r_g = 0.83 to 0.99) was found when measuring in different husbandry systems; cages, floor pens, and outdoor (N'Dri et al., 2007). Similarly, moderate to weak re-ranking (r_g = 0.67 to 0.96) was reported in body weight of rainbow trout measured between 2 different diets (Kause et al., 2006; Le Boucher et al., 2011). The previous studies above do not show consistent pattern of GxE interaction across livestock kept in different environments. However, most studies tend to show weak re-ranking across regions or locations or countries.

The high re-ranking in this study may be due to the large diversity of commercial environments combined with differences in age at harvest. Differences in various macro-environmental parameters, such as altitude, dissolved oxygen, water temperature, photoperiod, water sources, and feeding, may have contributed to

the strong GxE interaction observed. GxE interaction may be reduced by changing environmental parameters to be similar to the breeding environment. Identifying the environmental parameter explaining the GxE interaction will help in finding a breeding scheme to meet the different environments.

In our study, the r_g among the 3 production environments are more similar and it does not explicitly indicate which environment is the most different from the others. However, there is a tendency that r_g between PE and other environments are slightly lower for all traits. This suggests that PE is a slightly different environment than FI, and GE.

In all production environments, we collected information at a single location. Caution should, therefore, be taken in generalizing our results. To confirm the current result, we recommend that more experiments, using multiple farms, should be conducted.

5.4.2 Trait selection at nucleus

In trout breeding, two-stage selection (Cunningham, 1975) is sometimes used to enhance genetic gain, for instance in Finland (Kause et al., 2005; Martinez et al., 2006) and by Troutlodge. Two-stage selection can be implemented by tagging only the biggest fingerlings typically within families in the first stage, and the final selection among tagged individuals based on EBVs for all traits of interest (Martinez et al., 2006). To implement two-stage selection efficiently in trout across multiple environments, a positive r_g between trait used in the first stage (BWT) and the traits in the breeding goal (BWH and/or TGC across environments) are needed. Our study revealed that pre-selection for BWT in NUC will yield favourable correlated genetic responses in both BWH and TGC in FI and GE. Therefore, two-stage selection can be efficiently implemented for FI and GE. In contrast, pre-selection for higher BWT in NUC will indirectly contribute to lower-than-expected genetic gain (due to GxE interaction) of BWH and TGC in PE. Postponing pre-selection may improve the efficiency of two-stage selection and enhance genetic gain in PE (Sae-Lim et al., 2012b).

5.4.3 Method and selection bias

In this study, a multi-trait multi-environment (MTME) model was used and by including BWT, we accounted for selection bias due to selective mortality (Henderson, 1984; Pollak et al., 1984; Ouweltjes et al., 1988). This resulted in higher additive genetic and residual variances for BWH. The explanation could be

that mortality related to low body weight resulted in reduced variance among the surviving fish. Tagging weight is recorded on all fish, and hence BWH of culled fish can be estimated when both BWT and BWH are included into the model, returning the variance closer to its original value. In European whitefish, the impact of selection bias due to pre-selection was accounted for by using multi-trait analysis (Kause et al., 2011). In dairy cattle, an approximate multi-trait model was used to account for selection bias which resulted in higher accuracy of estimated breeding values (Lassen et al., 2007). In Dutch Warmblood horse, bivariate model accounting for selection bias due to pre-selection increased h^2 for dressage competition from 0.15 to 0.21 (Ducro, 2010).

5.4.4 Implication for breeding

Where GxE interaction is present, optimization of a breeding program allows genetic gains in all environments to be maximized. There are several strategies of optimization to be used, as described by Mulder et al. (2006). First, adjusting farming management to be similar to breeding environment may reduce GxE interaction. This certainly holds for market weight. Differences in market weight between environments can be accommodated better by collecting multiple weights in the nucleus. However, not all environmental parameters can be controlled. Second, sibs' performance information collected in different production environments can be incorporated in estimating breeding values for selection candidates in the nucleus. Using the sib information, it is possible to select breeding candidates in the nucleus that have high EBVs for performance in another environment (Mulder and Bijma, 2005). Third, environment-specific breeding programs can be implemented. To make a decision from a genetic point of view whether or not a single breeding program should be divided into 2 environment-specific breeding programs, "break-even correlation" can be used as a criterion (Mulder and Bijma, 2005; Mulder et al., 2006). The break-even correlation is defined as the intersection of genetic correlations when the genetic gain of different breeding strategies is equal. When the genetic correlation across environments is lower than the break-even correlation, separated breeding programs are recommended. The estimated break-even correlation in a dairy cattle breeding program ranges from 0.61 (Mulder, 2007) to 0.70 (James, 1961). In fish breeding, the break-even correlation is expected to be higher, i.e., ≥ 0.70, due to sib testing and higher selection intensity compared to cattle.

In our study, we found that r_g of a trait measured in different environments is lower than 0.7. This suggests that from a strictly genetic point of view separated breeding programs for the different environments seem to lead to a higher genetic gain than a single breeding program. However, it is very costly to organize environment-specific breeding programs. Opportunities to exploit sib information to overcome the disadvantage of GxE interaction needed to be further explored in combination with recording weight over different periods in the nucleus. Moreover, for example in dairy cattle, a single breeding program with progeny testing all bulls in 2 environments (OJ-2 strategy; Mulder et al., 2006) resulted in lower genetic gain in an overall objective than in 2 separate breeding programs (TE-1 strategy). But overall genetic gain from OJ-2 is not severely lower than TE-1 even though the r_g is lower than the break-even correlation of 0.61.

In conclusion, strong GxE interaction was found in tagging body weight, harvest body weight, and even stronger GxE interaction in growth rate. Pre-selection in nucleus may indirectly contribute to lower-than-expected genetic gain in Peru, due to GxE interaction. This study calls for a further research on optimization of breeding schemes that meets the different environments. A better understanding of the causes of the GxE interaction will help to design the most optimal breeding scheme from not only a genetic but also an economic point of view.

5.4 Acknowledgement

This study is a part of the PhD project: Optimization of rainbow trout breeding program for multi-trait selection with GxE interaction. This project is funded by The King Scholarship, Thailand (Ananda Mahidol Foundation) and Troutlodge, Inc., USA. We thank Sean M. Nepper, Doug P. Dixon, Mary L. Elliott, Lenore Brown, Brent A. Hinners, Shaun B. Elliott, Paul B. Goodmanson, Forellenzucht Trostadt (Germany), The Conservation Fund's Freshwater Institute, and Piscifactorías de los Andes S.A (Peru) for their considerable contributions to and significant time spent in the GxE interaction experiments.

References

Austreng, E., T. Storebakken, and T. Åsgård. 1987. Growth rate estimates for cultured Atlantic salmon and rainbow trout. Aquaculture 60: 157-160.

Banos, G., Sigurdsson, A., 1996. Application of contemporary methods for the use of international data in national genetic evaluations. J. Dairy. Sci. 79: 1117-1125.

Calus, P. M., 2006. Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle. Ph.D. Diss., Animal Breeding and Genetics, Wageningen Univ., The Netherlands.

Cunningham, E. P., 1975. Multi-stage index selection. Theor. Appl. Genet. 46: 55-61.

Duchesne, P., Godbout, M. H., Bernatchez, L., 2002. PAPA (package for the analysis of parental allocation): A computer program for simulated and real parental allocation. Mol. Ecol. Notes. 2: 191-193.

Ducro, B. J. 2010., Accounting for preselection in genetic evaluation of dressage performance of Dutch Warmblood horses, the 9th World Congress on Genetics Applied to Livestock Production (WCGALP), Leipzig, Germany.

Dupont-Nivet, M., Karahan-Nomm, B., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., Chatain, B., Vandeputte, M., 2010. Genotype by environment interactions for growth in European seabass (*Dicentrarchus labrax*) are large when growth rate rather than weight is considered. Aquaculture 306: 365-368.

Falconer, D. S., Mackay, T. F. C., 1996. Introduction to quantitative genetics. p 464 pp. Longman, Essex, UK.

FAO., 2011. Cultured aquatic species information programme *Oncorhynchus mykiss* (Walbaum, 1792): Fisheries and aquaculture department. Accessed May 18, 2011. http://www.fao.org/fishery/culturedspecies/Oncorhynchus_mykiss/en.

Fishback, A. G., Danzmann, R. G., Ferguson, M. M., Gibson, J. P., 2002. Estimates of genetic parameters and genotype by environment interactions for growth traits of rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. Aquaculture 206: 137-150.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., 2009. ASReml User Guide Release 3.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Hayes, J. F., Hill, W. G., 1981. Modification of estimates of parameters in the construction of genetic selection indices (bending). Biometrics 37: 483-493.

Henderson, C. R., 1984. Estimation of variances and covariances under multiple trait models. J. Dairy. Sci. 67: 1581-1589.

Hill, W. G., 1984. On selection among groups with heterogeneous variance. Anim. Sci. 39: 473-477.

Hokanson, K. E. F., Kleiner C. F., Thorslund, T. W., 1977. Effects of constant temperatures and diel temperature fluctuations on specific growth and mortality rates and yield of juvenile rainbow trout, *Salmo gairdneri*. J. Fish. Res. Board. Can. 34: 639-648.

James, J. W., 1961. Selection in two environments. Heredity 16: 145-152.

Jobling, M., 2003. The thermal growth coefficient (TGC) model of fish growth: A cautionary note. Aquac. Res. 34: 581-584.

Johnson, N. A., Rexroad III, C. E., Hallerman, E. M., Vallejo, R. L., Palti, Y., 2007. Development and evaluation of a new microsatellite multiplex system for parental allocation and management of rainbow trout (*Oncorhynchus mykiss*) broodstocks. Aquaculture 266: 53-62.

Kause, A., Quinton, C., Airaksinen, S., Ruohonen, K., Koskela, J., 2011. Quality and production trait genetics of farmed European whitefish, *Coregonus lavaretus*. J. Anim. Sci. 89: 959-971.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E., Eskelinen, U., 2003. Selection against early maturity in large rainbow trout *Oncorhynchus mykiss*: The quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228: 53-68.

Kause, A., Ritola, O., Paananen, T., Wahlroos, H., Mantysaari, E. A., 2005. Genetic trends in growth, sexual maturity and skeletal deformations, and rate of inbreeding in a breeding programme for rainbow trout (*Oncorhynchus mykiss*). Aquaculture 247: 177-187.

Kause, A., Tobin, D., Houlihan, D. F., Martin, S. A. M., Mäntysaari, E. A., Ritola, O., Ruohonen, K., 2006. Feed efficiency of rainbow trout can be improved through selection: Different genetic potential on alternative diets. J. Anim. Sci. 84: 807-817.

Kolstad, K., Thorland, I., Refstie, T., Gjerde, B., 2006. Genetic variation and genotype by location interaction in body weight, spinal deformity and sexual maturity in Atlantic cod (*Gadus morhua*) reared at different locations off Norway. Aquaculture 259: 66-73.

Lassen, J., Sorensen, M. K., Madsen, P., Ducrocq, V., 2007. An approximate multitrait model for genetic evaluation in dairy cattle with a robust estimation of genetic trends. Genet. Sel. Evol. 39: 353-367.

Le Boucher, R., Quillet, E., Vandeputte, M., Lecalvez, J. M., Goardon, L., Chatain, B., Médale F., Dupont-Nivet, M., 2011. Plant-based diet in rainbow trout (*Oncorhynchus mykiss* Walbaum): Are there genotype-diet interactions for main production traits when fish are fed marine vs. Plant-based diets from the first meal? Aquaculture 321: 41-48.

Lynch, M., Walsh, B., 1998. Genetics and analysis of quantitative traits. p 980 pp. Sinauer Associates, Inc., U.S.A.

Mallet, J. P., Charles, S., Persat, H., Auger, P., 1999. Can. J. Fsih. Aquat. Sci. 56: 994-1000.

Maluwa, A. O., Gjerde, B., Ponzoni, R. W., 2006. Genetic parameters and genotype by environment interaction for body weight of *Oreochromis shiranus*. Aquaculture 259: 47-55.

Martinez, V., Kause, A., Mäntysaari, E., Mäki-Tanila, A., 2006. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: II. Two-stage selection. Aquaculture 254: 195-202.

Matthews, K. R., Berg, N. H., 1997. Rainbow trout responses to water temperature and dissolved oxygen stress in two southern California stream pools. J. Fish. Biol. 50: 50-67.

Meuwissen, T. H. E., De Jong, G., Engel, B., 1996. Joint estimation of breeding values and heterogeneous variances of large data files. J. Dairy. Sci. 79: 310-316.

Mulder, H.A., 2007. Methods to optimize livestock breeding programs with genotype by environment interaction and genetic heterogeneity of environmental variance. PhD Diss., Animal Breeding and Genetics group, Wageningen Univ., The Netherlands.

Mulder, H. A., Bijma, P., 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83: 49-61.

Mulder, H. A., Veerkamp, R. F., Ducro, B. J., van Arendonk, J. A. M., Bijma, P., 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. J. Dairy. Sci. 89: 1740-1752.

N'Dri, L., Sellier, N., Tixier-Boichard, M., Beaumont, C., Mignon-Grasteau, S., 2007. Genotype by environment interactions in relation to growth traits in slow growing chickens. Genet. Sel. Evol. 39: 1-16.

Ojolick, E. J., Cusack, R., Benfey, T. J., Kerr, S. R., 1995. Survival and growth of all-female diploid and triploid rainbow trout (*Oncorhynchus mykiss*) reared at chronic high temperature. Aquaculture 131: 177-187.

Ouweltjes, W., Schaeffer, L. R., Kennedy, B. W., 1988. Sensitivity of methods of variance component estimation to culling type of selection. J. Dairy. Sci. 71: 773-779.

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

Pollak, E. J., van der Werf, J., Quaas, R. L., 1984. Selection bias and multiple trait evaluation. J. Dairy. Sci. 67: 1590-1595.

Robertson, A., 1959. The sampling variance of the genetic correlation coefficient. Biometrics 15: 469-485.

Rutten, M. J. M., Komen, H., Bovenhuis, H., 2005. Longitudinal genetic analysis of Nile tilapia (*Oreochromis niloticus* L.) body weight using a random regression model. Aquaculture 246: 101-113.

Sae-Lim, P., Komen, H., Kause, A., 2010. Bias and precision of estimates of genotype-by-environment interaction: A simulation study. Aquaculture 310: 66-73.

Sae-Lim, P., Komen, H., Kause, A., van Arendonk, J. A. M., Barfoot, A. J., Martin, K. E., Parsons, J. E., 2012a. Defining desired genetic gains for rainbow trout breeding objective using analytic hierarchy process. J. Anim. Sci. 90: 1766-1776.

Sae-Lim, P., Komen, H., Kause, A., Martin, K. E., Crooijmans, R., van Arendonk, J. A. M., Parsons, J. E., 2012b. Enhancing selective breeding for growth, slaughter traits and overall survival in rainbow trout (*Oncorhynchus mykiss*). Accepted.

Stram, D. O., Lee, J. W., 1994. Variance components testing in the longitudinal mixed effects model. Biometrics 50: 1171-1177.

Williams, J. L., Bertrand, J. K., Misztal, I., Łukaszewicz, M., 2012. Genotype by environment interaction for growth due to altitude in United States Angus cattle. J. Anim. Sci. 90: 2152-2158.

6

Identifying environmental parameters explaining genotype-by-environment interaction in body weight of rainbow trout (*Onchorynchus mykiss*): reaction norm and factor analytic models

P. Sae-Lim[1], H. A. Mulder[1], H. Komen[1], A. Kause[2]

[1]Animal Breeding and Genomics Centre, Wageningen University, P.O. Box 338, 6700 AH, Wageningen, The Netherlands; [2]MTT Agrifood Research Finland, Biotechnology and Food Research, Biometrical Genetics, FI-31600 Jokioinen. Finland.

Submitted to Journal of Animal Science

Abstract

Understanding the biological causes for genotype-by-environment interaction (GxE) is important for breeders. Finding the relevant environmental parameters (EP) that cause GxE is often difficult when environmental factors cannot be experimentally manipulated. Two statistical approaches can be taken to address this question. When data on candidate EP are available, GxE can be quantified along the specific EPs using a reaction norm. Alternatively, a factor analytical (FA) model can be used to identify the unknown common factor that explains GxE. This factor can then be correlated with known EPs to identify the relevant EP. In a previous study, we reported significant GxE for body weight at harvest (BWH) in rainbow trout, grown on 3 different continents. Here we explore the possible causes for this GxE. Reaction norm and factor analytic (FA) model were both used to identify which of the following EP: age at harvest, water temperature, oxygen, and photoperiod, potentially caused the observed GxE. Data on BWH was recorded from 8,976 offspring reared in one of the following locations: 1) the breeding environment in Washington State, USA (NUC), 2) a recirculating aquaculture system in West Virginia, USA (FI), 3) a high-altitude farm in Peru (PE), and 4) a cold-water farm in Germany (GE). Akaike information criteria (AIC) and Bayesian information criteria (BIC) were used for model comparison. Photoperiod and the combination of days to harvest multiplied with daily temperature (Day*Degree) were identified by the reaction norm model as the EPs causing GxE. The unknown factor, identified by the FA model showed the highest correlation with Day*Degree. Photoperiod and Day*Degree were the EPs differing the most between PE and other environments. The AIC and BIC indicated that FA model was more parsimonious than the reaction norm model. The low variation in EPs reduced the power of the reaction norm model to identify EPs, and future studies should use multiple farms per production environment. A factor analytical model is preferred over a reaction norm model when only limited information on the variation of EP between farms is available.

Key Words: environmental sensitivity, factor analytic model, genotype by environment interaction, rainbow trout, reaction norm model

6.1 Introduction

Harvest body weight is an economically important trait in rainbow trout (*Onchorynchus mykiss*) and in other farmed fish species. Rainbow trout can be produced in a wide range of farming environments. When genotype-by-environment interaction (GxE) is present, selection practiced solely in a breeding environment may lead to lower-than-expected genetic gains in the other production environments. Optimization of a breeding program to account for GxE can increase genetic gain across environments (Mulder and Bijma, 2005; Martinez et al. 2006; Mulder et al., 2006). Optimization may be expensive, for instance when environment-specific breeding programs need to be established. Alternatively, changing environmental parameters (EP) to be similar across production environments may reduce GxE. To do so, the EPs causing GxE should be identified.

To identify EPs causing GxE, a reaction norm model can be used to quantify GxE as the function of specific environmental parameters (Fikse et al., 2003; Zwald et al., 2003; Schaeffer, 2004). Alternatively, in a two-step factor analysis, an unknown common factor causing GxE is first identified, and subsequently correlations between the common factor and EPs are calculated to identify the significant EPs (Van Eeuwijk et al., 2001). In this study, the aim was to identify the environmental parameters causing strong GxE in harvest body weight of rainbow trout using a reaction norm model and a factor analytic model.

6.2 Materials and methods

6.2.1 Data

The data used in this study were from a GxE experiment conducted in 4 different environments located in 3 continents (North America, South America, and Europe as described by Sae-Lim et al. 2012b). In August 2009, 100 full-sib families were produced from 58 sires and 100 dams (1 to 1.7 mating ratio) at Troutlodge breeding company in Washington State (NUC). The fertilization took place during a period of 4 weeks. Different water temperature was used to synchronize embryonic development and hatching. At eyed-egg stage, each family was split into 5 groups to be tested at 4 different production environments. At least 25 eyed eggs per family were shipped to each of the following 3 locations: 1) recirculating aquaculture system located at the Freshwater Institute, Virginia, USA (FI); 2) a high altitude farm with low oxygen dissolved in the water (Titicaca Lake) in Peru (PE);

and 3) a low water temperature farm in Germany (GE). A random sample of 25 eyed eggs per family was held at NUC as a control. All fish were measured for harvest body weight (BWH, in grams), in June 2010 (NUC), in July 2010 (FI), in August 2010 (PE) and in December 2010 (GE) (Table 6.1).

Table 6.1 Descriptive statistics for body weight at harvest (BWH, unit g) in 4 different environments.

Environment	No.	Mean	SD
NUC	2367.0	546.7	94.7
FI	1893.0	395.2	75.8
PE	2897.0	524.1	105.0
GE	1819.0	376.4	81.7

No. = number of observations, NUC = breeding environment, FI = Freshwater institute, PE = Peru, GE = Germany; SD =standard deviation

6.2.2 Pedigree reconstruction

The fish were tagged using passive integrated transponders (PIT tag; Allflex USA, Inc. for NUC, FI and PE, and DORSET Identification b.v., the Netherlands, for GE) and the PIT tag was scanned (scanner SF2001ISO: Destron Fearing, USA for NUC, FI and PE, and GR250: DORSET Identification b.v., the Netherlands, for GE) at the average size of 26.3 to 33.2g (5 to 7 months age). Before tagging, fish were anesthetized using MS222 (150 mg/l) in NUC, FI, and PE and using clove oil (10 mg/l) in GE. Fin clips were collected from all 158 parents and from the fish at tagging from FI, PE and GE for DNA extraction. In NUC, fish were kept in separated family tanks until tagging, allowing pedigree construction; therefore fin clips were not collected in NUC.

DNA was isolated from the fin-clips to reconstruct the pedigree. Genotyping of the DNA samples was done in 3 laboratories: National Center for Cool and Cold Water Aquaculture, USDA; Troutlodge, Inc.; and Animal Breeding and Genomics Centre, Wageningen University. The protocols for DNA isolation and genotyping were synchronized across the labs. The DNA isolation was done using Nucleospin® 96 Tissue Core Kit. Multiplex PCR amplification was done as described by Johnson et al. (2007). The following 9 microsatellite markers were used for PCR: OMM1008, OMM1051, OMM1088, OMM1097 (Rexroad et al., 2002), OMM5007, OMM5047 (Rexroad et al., 2005), OMM5233, OMM5177 (Coulibaly et al., 2005), and OMM1325 (Palti et al., 2002). Multiplex PCR amplification, i.e. quadroplex and

pentaplex, was done as follows (Johnson et al., 2007): an initial 5 min denaturation at 95 ^{o}C, followed by 35 cycles of 30 s denaturation at 95 ^{o}C, 45 s annealing at 55 ^{o}C, and 90 s extension at 72 ^{o}C, and a final 10 min extension at 72 ^{o}C. Fragment analysis of the PCR products was done by setting the fragment sizes to Genescan LIZ 500 size standard (Applied Biosystem). The output data were analysed using Genemapper software version 4 (Applied Biosystem) (Sae-Lim et al., 2012a).
Parental allocation was performed using PAPA software (Duchesne et al., 2002). The known mating data were used to increase the accuracy of parental assignments (Sae-Lim et al., 2012b). In total, 2142 out of 2243 fish sampled in FI, 3106 out of 3236 fish sampled in PE, and 2104 out of 2235 fish sampled in GE were successfully allocated to the 100 full-sib families. The 362 fish that were not successfully allocated to the families were removed from the dataset. In total, 6 generations of pedigree information, from the DNA reconstructed pedigree and from the 5 previous generations of pedigree information were used in the genetic analysis.

6.2.3 Environmental parameters

Summary statistics of EPs are given in Table 6.2. Data on the following EPs were available:

Temperature

Average water temperature (^{o}C) measured in the tank (NUC, FI), raceway (GE) or lake (PE) in the farm during the rearing period of the experiment. In NUC, the average ambient temperature was between 13 to 14 ^{o}C throughout the growing season. In FI, PE, and GE the water temperature followed the natural (daily and seasonal) fluctuations. Water temperature was recorded for every 15 minutes using data logging Transmitter SC100 (Hach Lange, Germany) in NUC and GE. For FI, temperature was measured once a day using either a Hach HQ40d hand held meter or a SC100 Universal Controller (Hach Company, Loveland, CO). For PE, measurement using a standard mercury thermometer was done at Titicaca Lake once a day for only a short period (September 3 to 16, 2010). However, the water temperature in Titicaca Lake is not fluctuating through the whole year and season and variation in water temperature is small (12 to 14 ^{o}C).

Age

Average age at harvest (in days) was calculated from hatching up to day of harvest. Differences in age at harvest were caused by differences in preferred market sizes across environments. In NUC, it was not possible to harvest and measure all fish simultaneously. Consequently, harvest was done twice (2 weeks interval).

*Day*Degree*

In salmonids, growth rate is dependent on temperature. The product of days to harvest and daily temperature is therefore commonly used in salmonid farming to compare days to harvest across temperature regimes. Day*degree was calculated as: average water temperature multiplied with average age at harvest.

Oxygen

The amount of oxygen dissolved in the water during the rearing period, recorded in mg/l or ppm was calculated as the average of daily measurements, as follows: In NUC, oxygen was measured daily in the morning (7:30 to 9:00 am) using YSI model 550 (YSI, Yellow Springs, OH) at the inlet and the outlet of the rearing tanks. In FI, oxygen was measured at a single position in the circular tanks once a day between 8:00 to 9:30 am using a Hach HQ40d with a Hach LDO probe attachment, or a SC100 Universal Controller (Hach Company, Loveland, CO). In PE, dissolved oxygen was measured at in the net pens of Titicaca Lake in the morning (9:00 to 10:00 am) for a short period of time (same as temperature) using Hach dissolved oxygen test kit (Hach Company, Loveland, CO). In GE, oxygen level was controlled to be above 10 mg/l. When the oxygen decreased, supplement oxygen was automatically released until the oxygen was above 10 mg/l. The oxygen measurement was done for every 15 minutes using data logging Transmitter SC100 (Hach Lange, Germany). Average dissolved oxygen for each environment was calculated as the total amount of dissolved oxygen divided by the number of observations.

Photoperiod

The experiment was conducted across continents. Consequently, changes in day length are different. “Photoperiod” was defined as the difference between maximum day length observed during the rearing period and the average day length of overall rearing period. The locations that were used for calculating

photoperiod were: Seattle in Washington State (NUC), Martinsburg in West Virginia (FI), Juliaca in Peru (PE), and Leipzig Schkeuditz in Germany (GE). The data of weekly sunrise and sun set in 2009 and 2010 were obtained from http://www.wunderground.com/history/. The average day length was calculated from the difference between sunset and sunrise in minutes, for each week in the rearing period (Figure 1). To account for differences between northern and southern hemisphere (NUC, FI and GE versus PE), we used negative and positive signs to indicate the directions of change in the photoperiod.

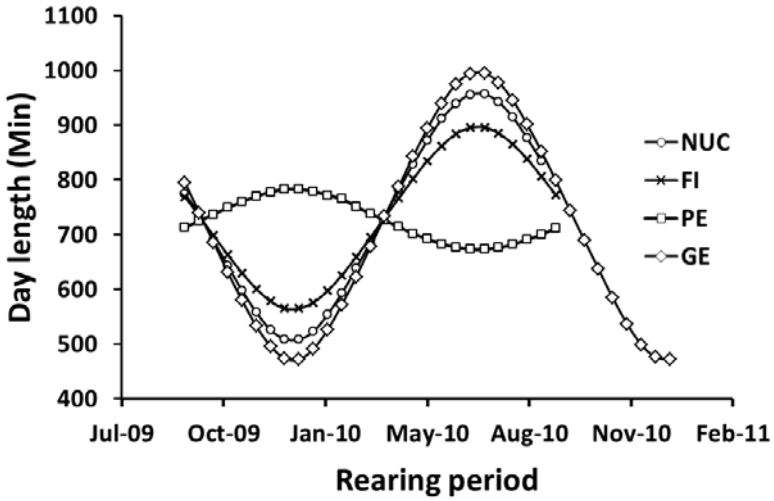

Figure 6.1 Day length profiles in the 4 experiment environments. The x-axis represents the rearing period in two-month intervals (month-year). Each observation represents the average day length during a two-week interval. The rearing period differed across environments: NUC =breeding environment, FI = Freshwater Institute, PE = Peru and GE= Germany.

6.2.3 Genetic analysis

In a previous study, we reported significant GxE for body weight at harvest in rainbow trout, grown on 3 different continents. Genetic correlations between the locations ranged from 0.19 to 0.48 (Sae-Lim et al., submitted). In the present study, the same data were used to identify EP.

Table 6.2 Means of environmental parameters during the rearing period (unit in parenthesis).

Environment	Age (day)	Temperature (°C)	Day*Degree (day*°C)	Oxygen (mg/l)	Photoperiod (minute)
NUC	287.5	13.8	3940.3	7.3	223.1
FI	294.0	12.5	3685.6	10.5	163.3
PE	357.0	13.4	4805.3	6.6	-53.1
GE	444.0	9.9	4439.2	12.0	292.9

NUC = breeding environment, FI = Freshwater Institute, PE = Peru, GE = Germany.

Reaction norm model

The EP causing GxE can be identified by fitting each EP in the reaction norm model. Random regression was used to estimate (co)variance components. The random animal effect was modelled as a function of EP. The fixed covariate *AGE* was included in the model to account for differences in body weight at harvest due to differences in age at harvest across environments. The fixed *FERT* effect was included in the model to correct for different fertilization periods. The random regression model was:

$$Y_{hijkl} = \mu + \eta_h + \beta_0 AGE_{hi} + FERT_{hj} + \sum_{k=0}^{m} \alpha_k P_{kl} + E_{hijkl}, \qquad [1]$$

where Y_{hijkl} is the BWH of each animal sorted by environment; μ is the average BWH of all animals; η_h is fixed environmental effect (h =1: NUC, 2: FI, 3: PE, and 4:GE), accounting for different levels of environment; β_0 is the coefficient of linear fixed regression on age at harvest (AGE_{hi}) within the *h*th environment; $FERT_{hj}$ is the *j*th fertilization period within the *h*th environment; α_k ($\begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_m \end{bmatrix}$ ~ MVN[0, $\mathbf{A} \otimes \mathbf{G_{RN}}$], where MVN is multivariate normal distribution, $\mathbf{A}$ is additive genetic relationship matrix, and $\mathbf{G_{RN}}$ is genetic (co)variance matrix from reaction norm model), is the coefficient *k* of the random regression on element *k* of the orthogonal polynomial P_{kl} resembling an EP of environment *l* and *m* is the maximum order of the polynomial. The E_{ijkl} is the random residual effect (E~ N(0, $\mathbf{I}\sigma_E^2$)), where $\mathbf{I}$ is the identity matrix. Common environmental effect was excluded from the model due to difficulty in estimating genetic parameters (Maluwa et al., 2006).

The orthogonal polynomial of *AGE* (fixed regression effect) was tested for the significance up to the third order; however the quadratic and the cubic orders were

not significant (Wald test). The order of orthogonal polynomial (*m*) in the random term was obtained from the test using Akaike's information content (AIC: Akaike, 1973) and Bayesian's information content (BIC: Schwarz, 1978). The lowest AIC and BIC indicates the most parsimonious polynomial order for the polynomial random term. The third order polynomial is the most parsimonious reaction norm model, but it was not used in this study. The third order polynomial results in 4x4 **G** matrix, which is the same dimension as the original multivariate model with 4 environments. Consequently, r_g from 3rd order model are similar to the estimates from the multivariate model.

Additive genetic variance (V_A) for each level of EP was calculated by $\mathbf{\Phi G_{RN} \Phi'}$, where $\mathbf{\Phi}$ is a vector of polynomial coefficients for each level of EP with size of 1**m*. The *m* is the highest order of polynomial + 1; $\mathbf{G_{RN}}$ is a genetic covariance matrix with size of *m***m*; and $\mathbf{\Phi'}$ is the transposed vector of $\mathbf{\Phi}$. The covariance (COV) between different levels of EP, i.e., level *i*th and *j*th, was calculated by $\mathbf{\Phi_i G_{RN} \Phi_j'}, i \neq j$. The genetic correlation (r_g) between different levels of EP, representing the average EP of each environment was calculated as $\mathrm{COV}(EP_i, EP_j)/(\sqrt{V_{A,EP_i} \times V_{A,EP_j}})$. The sire BLUP-estimated breeding value (EBV) for each level of EP was calculated as $\mathbf{H\Phi'}$, where **H** is a vector of sire BLUP-EBV for α_k polynomial coefficients with size of 1**m.* The sire BLUP-EBV was plotted against EPs to show the degree of heterogeneity of variance and re-ranking. Only 10 sires were randomly selected in order to avoid excessively dense information in the plot.

Factor-analytic model

The factor analytic (FA) model is a model to identify unknown common factors explaining the variation of the data. The FA model can be used to estimate GxE (Meyer, 2009) using unknown common effects or loadings.The FA animal mixed model was:

$$Y_{hijk} = \mu + \eta_h + \beta_0 AGE_{hi} + FERT_{hj} + Ac_k + As_k + E_{ijk}, \quad [2]$$

where *Ac* and *As* are the random animal effects due to a common factor and specific effects, respectively, $(A_c + A_s) \sim \mathrm{MVN}[\mathbf{0}, \ \mathbf{A} \otimes \mathbf{G_{FA}}]$, where MVN is multivariate normal distribution, $\mathbf{A}$ is the additive genetic relationship matrix, and $\mathbf{G_{FA}}$ is the genetic variance-covariance matrix for common and specific animal effects. The genetic variance-covariance matrix $\mathbf{G_{FA}} = \mathbf{\Gamma\acute{\Gamma}} + \mathbf{\Psi}$, where **Γ** is the

matrix of factor loadings; $\mathbf{\Gamma}'$ is the transposed matrix of loadings; and $\mathbf{\Psi}$ is the diagonal matrix of specific variances (ψ_l), accounting for additional variance, i.e., the variation that is not explained by unknown common factors, of the *l*th environment (Meyer, 2009). The total number of parameters fitted in FA model is $n(k+1)-k(k-1)/2$ and may not exceed $n(n+1)/2$, where n is the size of **G** matrix, and k is the number of unknown common factors. When k is 1, the number of parameter fitted in FA was $4(1+1) - 1(1-1)/2 < 4(4+1)/2$, or $8 < 10$ (Gilmour et al., 2009). The 8 parameters were 4 for loading parameters and 4 specific variances. The number of factors cannot be higher than 1 in this study.

In ASReml, different types of FA models can be implemented (Gilmour et al., 2009). In this study, we used the extended FA model which provides $\mathbf{G_{FA}}$ matrix, loading parameters, and correlation between genetic effects in 4 environments, and unknown common factor. Additive genetic variance (V_A) for a certain environment was calculated as: $V_A = \hat{\gamma}\hat{\gamma}' + \hat{\psi}$, where $\hat{\gamma}$ and $\hat{\psi}$ are the estimated loading vector and estimated specific variance, respectively. The square of the loading parameter indicates the amount of additive genetic effect explained by the unknown common factor. The high loading in an environment indicates higher additive genetic variance in such environment explained by unknown common factor. The percentage of additive genetic variance explained by unknown common factor was calculated as: $\%\text{Expl} = \frac{\hat{\gamma}^2}{V_A} \times 100$. The COV of BWH between environment *i*th and *j*th was calculated as $\hat{\gamma}_i\hat{\gamma}_j'$. The r_g between BWH measured in different environments was calculated as: $r_g = \frac{\text{COV}_{BWH_i, BWH_j}}{\sqrt{V_{A,BWH_i} \times V_{A,BWH_j}}}$, where *i* and *j* represent different environments. Pearson ($\rho_{EP,L}$) and Kendall rank ($\tau_{EP,L}$) correlations between EP and loading parameters of the unknown common factor were calculated to identify EP, causing GxE.

Model comparison

Reaction norm and FA models were compared with AIC and BIC to determine the most parsimonious model. All the models were kept the same with respect to fixed effects to make all the models comparable in terms of REML log likelihood.

6.2.4 Identification of EP

With a reaction norm model, the best fitted EP will have the highest Log likelihood. In addition, mean square deviation (MSD) was calculated from the difference

between estimated genetic correlation from reaction norm and factor analytic model. The $MSD = \frac{\sum_{i=1}^{n}\left(r_{g_{RN,i}} - r_{g_{FA,i}}\right)^2}{n}$, where $r_{g_{RN,i}}$ and $r_{g_{FA,i}}$ are estimated genetic correlation of BWH between different environments from reaction norm and factor analytic models, respectively. The *i*th genetic correlation is from the same pair of environments for both models and *n* is equal to 6, because with 4 environments there are 6 genetic correlations. Genetic correlations from factor analytic model were very similar to genetic correlations estimated from multivariate model. Therefore, the reaction norm model with the lowest MSD is the model that deviates least from the multivariate model indicating that the EP used in the reaction norm model is able to capture GxE.

FA model was used as the first step in a two-step approach. The second step was to estimate correlations between loadings and EPs. The EP that is highly correlated with the unknown common effect is most likely the significant EP causing GxE.

6.3 Results

6.3.1 Reaction norm model

For reaction norm model, r_g of BWH between different means of EP, representing each environment are shown in Table 3. For age at harvest, r_g varied from 0.572 to 0.998 (MSD = 0.17). For water temperature, r_g varied from 0.606 to 0.998 (MSD = 0.17). Genetic correlations for Day*Degree were lower (0.349 to 0.973: MSD = 0.10). For dissolved oxygen, the range of r_g (0.603 to 0.996: MSD = 0.14) was similar to r_g for water temperature. For photoperiod, r_g ranged from 0.368 to 0.969 (MSD = 0.10). Reaction norm models with day-degree and photoperiod as EP resulted in genetic correlations closest to the FA-model or multivariate model (results not shown), indicating that day-degree and photoperiod are the most important EP explaining GxE.

The plot of sire BLUP-EBVs against age at harvest showed GxE: both heterogeneity of variance and re-ranking was observed (Figure 2). Sire BLUP-EBVs tended to be sensitive to the change of water temperature in the same direction as age at harvest. Sire BLUP-EBVs were even more sensitive to changes in Day*Degree, dissolved oxygen, and photoperiod; their plots showed more heterogeneity and stronger re-ranking than age at harvest and water temperature.

Table 6.3 Genetic correlation of body weight at harvest measured in different environments and mean square deviation (MSD), estimated by reaction norm and factor analytic models.

Model	Environmental parameter	Environment	FI	PE	GE	MSD
Reaction norm	Age	NUC	0.99	0.91	0.57	0.17
		FI		0.94	0.63	
		PE			0.86	
	Temperature	NUC	0.97	1.00	0.61	0.17
		FI		0.98	0.79	
		PE			0.65	
	Day*Degree	NUC	0.97	0.56	0.82	0.10
		FI		0.35	0.66	
		PE			0.93	
	Oxygen	NUC	0.85	1.00	0.67	0.14
		FI		0.80	0.96	
		PE			0.60	
	Photoperiod	NUC	0.97	0.60	0.96	0.10
		FI		0.78	0.87	
		PE			0.37	
Factor analytic model	Unknown common factor	NUC	0.56 ± 0.06	0.36 ± 0.04	0.54 ± 0.06	N.A.
		FI		0.41 ± 0.05	0.61 ± 0.07	
		PE			0.39 ± 0.05	

NUC = breeding environment, FI = Freshwater Institute, PE = Peru, GE = Germany, N.A. = Not applicable.

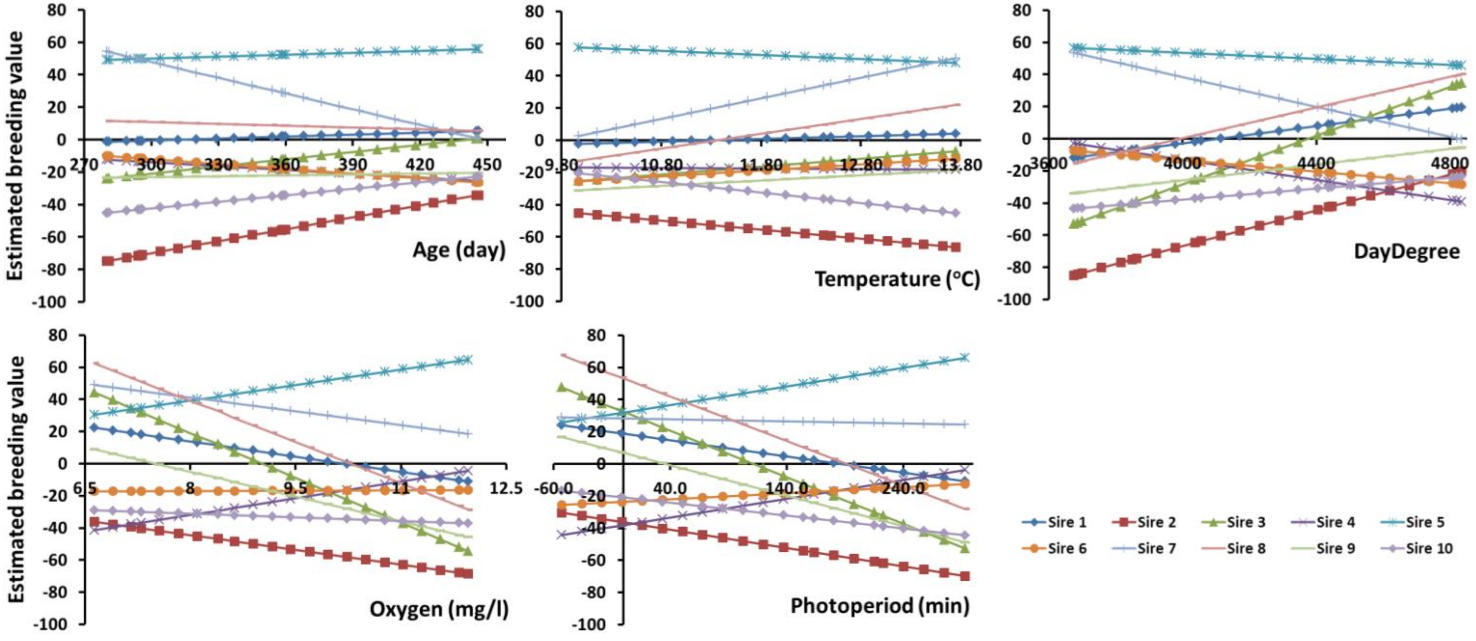

Figure 6.2 Estimated breeding values of sires for body weight (y-axis: in g) across the age at harvest (day), water temperature (^{o}C), Day*Degree (day*^{o}C), dissolved oxygen (mg/l), and photoperiod (min) using reaction norm model. Only 10 randomly chosen sires are plotted in this graph to illustrate the degree of re-ranking.

6.3.2 Factor analytic model

For FA model, r_g of BWH between PE and NUC (0.36), between PE and FI (r_g= 0.41), and between PE and GE (0.39) were low, indicating moderate to strong re-ranking. The $\hat{\gamma}$ was 40.34 in NUC, 38.57 in FI, 30.41 in PE, and 30.70 in GE, indicating that the unknown common factor mainly explained the majority of V_A in NUC and FI (Table 4). The proportion of genetic variance explained by the common factor was only 26.20% in PE. The $\hat{\psi}$ was high in PE (2606.73), showing high additive genetic variance not accounted for by the unknown common factor.

Table 6.4 Total genetic variance (V_A), estimated loading ($\hat{\gamma}$), specific genetic variance ($\hat{\psi}$), genetic variance explained in percentage by unknown factor (%Expl).

Environment	V_A	$\hat{\gamma}$	$\hat{\psi}$	%Expl
NUC	3283.1	40.3	1656.1	49.6
FI	2361.7	38.6	874.2	63.0
PE	3531.3	30.4	2606.7	26.2
GE	1613.4	30.7	670.9	58.4

NUC = breeding environment, FI = Freshwater Institute, PE = Peru, GE = Germany

To correlate the common unknown factor to the known EPs, Pearson correlation ($\rho_{EP,L}$) was negative and high between loadings and day*degree (-0.91), and loadings and age at harvest (-0.86). The Kendall rank correlation ($\tau_{EP,L}$) was in agreement with $\rho_{EP,L}$ but lower for both day*degree ($\tau_{EP,L}$ = -0.67) and age at harvest ($\tau_{EP,L}$ = -0.67) (Table 5). Water temperature was moderately correlated with loadings ($\rho_{EP,L}$ = 0.50). Dissolved oxygen was weakly correlated ($\rho_{EP,L} = -0.14$) or not correlated ($\tau_{EP,L} = 0.00$) with the loadings. Photoperiod was positively correlated with loadings ($\rho_{EP,L}$ =0.32, $\tau_{EP,L}$ = 0.33). The results indicated that Day*Degree was the most likely EP causing GxE for BWH.

Table 6.5 Correlation between loadings from factor analytic model and environmental parameters.

Environmental parameter	Pearson	Kendall rank
Age	-0.86	-0.67
Temperature	0.50	0.33
Day*Degree	-0.91	-0.67
Oxygen	-0.14	0.00
Photoperiod	0.32	0.33

6.3.3 Model comparison

For the reaction norm model, the lowest AIC (87645.7) and BIC (87695.3) indicated photoperiod as the best fitted EP, compared to the other EPs (Table 6). However, Day*Degree (AIC=87656.5, BIC =87706.2) fitted the model similarly well. The best fit was concordant with lower average r_g for either photoperiod or Day*Degree. The AIC (87513.0) and BIC (87528.6) from FA model were lower than those from reaction norm models, indicating that the FA-model is more parsimonious than the reaction norm model.

Table 6.6 Model comparison between random regression and factor analytic models.

Model	EP	LogL	NPar	dfR	AIC	BIC
	Age	-43853.7	7	8854	87721.5	87771.1
	Temperature	-43850.6	7	8854	87715.3	87764.9
Random regression	Day*Degree	-43821.3	7	8854	87656.5	87706.2
	Oxygen	-43840.2	7	8854	87694.4	87744.1
	Photoperiod	-43815.9	7	8854	**87645.7**	**87695.3**
Factor analytic	Unknown	-43748.5	8	8854	**87513.0**	**87528.6**

LogL = natural logarithm of likelihood, NPar = number of parameters, dfR = residual degree of freedom, AIC = Akaike's information content, BIC = Bayesian's information content, bold letter indicates the lowest AIC and BIC from both random regression and factor analytic models.

6.4 Discussion

The aim of this study was to identify the environmental parameters (EP) explaining GxE in harvest body weight (BWH) of rainbow trout using a reaction norm and a factor analytic model.

6.4.1 Identifying environmental parameters

To our knowledge, this is the first study that implemented reaction norm model and factor analytic (FA) model to identify significant EPs causing GxE in aquaculture. Our findings show that both methods can be used to identify significant EPs. However, reaction norm model identified different significant EP than the two-step FA model. Based on AIC and BIC, photoperiod gave a slightly better fit to reaction norm model than Day*Degree, indicating that photoperiod is the significant EP. However, in FA model Day*Degree was highly negatively correlated (Pearson correlation: $\rho_{EP,L}$ = -0.91) with loadings of the unknown common factor, suggesting that Day*Degree was the most significant EP. On the

contrary, both the reaction norm model and the FA model indicate that Day*Degree is an important EP, whereas temperature seems not to be the most likely EP responsible for GxE. Power to identify EP is, however, limited due to having 4 environments. Identifying environmental parameters explaining GxE has been studied using different methods. In Guernsey cows from 4 different countries, 15 environmental parameters were studied using random regression model and it was found that 9 of them indicated the presence of GxE (r_g= 0.85 to 0.98) (Fikse et al., 2003). By calculating genetic correlations between animals being at the opposite ends of environmental gradients, Zwald et al. (2003) found that 7 out of 13 EPs caused genetic correlations deviating from unity (r_g = 0.79 to 0.90).

Identifying significant EP causing GxE is valuable because this information may be used to reduce GxE before optimization of a breeding program. Optimization of a breeding program may be more expensive due to the possible need to establish multiple sib testing stations or environment-specific breeding programs, than changing the significant EP to be similar across environments reducing GxE. However, changing EP to be similar across environments may be expensive for the farmers/producers or impossible, e.g. sea water temperature. The reduction in genotype re-ranking across environments would lead to an increase of genetic gain of BWH in the production environments (FI, PE, and GE) but the extra profit that this generates may be offset by the extra costs of EP manipulation. Finding the significant EP is also of biological interest, because it provides evidence for environmental sensitivity of growth in rainbow trout. Artificial selection may change the gene pool of the population in the direction that fish perform best in the controlled environment where selection is practiced, leading to increased environmental sensitivity across multiple environments (Kolmodin et al., 2002; van der Waaij, 2004). Higher sensitivity to environments may lead to negative consequences, such as reduced fitness and animal health or negative economic impact (Ashley, 2007).

Previous studies have shown that photoperiod is one of the major factors influencing growth in rainbow trout (Sumpter, 1992; Taylor et al., 2005, 2006). In general, longer day length tends to increase growth rate. Taylor et al. (2005) found that rainbow trout exposed to L:D=18:6 rhythm, where L =light hours, and D =dark hours, grew significantly faster than rainbow trout exposed to L:D=8:16, and expressed significantly higher circulating levels of insulin-like growth factor-I (IGF-I) hormone. This hormone is positively correlated with growth rate in rainbow trout (Taylor et al., 2005). These observations support the idea that photoperiod may cause significant GxE in growth if genetic variation in sensitivity to photoperiod exists. The direction of change in day length in Peru is the opposite from that in the

other locations. The light rhythm can be manipulated in aquaculture production and manipulation of photoperiod by placing lamps under or above the water, is becoming common practice to enhance growth and delay sexual maturation in Atlantic salmon and in rainbow trout (Taylor et al., 2006). Therefore, it may be possible to reduce GxE due to different photoperiods.

Day*Degree is a combination of 2 individual factors: days to harvest determining the length of the rearing period, and average water temperature in ^{o}C. The differences in Day*Degree across environments may result from differences in age, differences in temperature, or both. Age at harvest is easy to adjust to be the same across environments, to reduce the observed re-ranking. However, the commercial market weight differs across countries, and thus the age differences need to be maintained. A bulk of the rainbow trout production occurs in fresh and sea water net pens, pool or raceways in which temperature control is difficult.

6.4.2 Model comparison

In this study, the most significant EP was identified by the reaction norm model using the following criteria: the best fitted EP to the model according AIC and BIC, resulting in the lowest mean square deviation (MSD) between r_g from reaction norm and factor analytic models. Due to the lack of replicates within environments, reaction norm model resembled a model with categorical EP. Reaction norm model would pinpoint the EPs more efficiently if the EPs were measured on a more continuous scale (e.g. more environments or treatments). Factor analytic models are frequently used in plant breeding, for example in multi-environment trials to analyse variety testing (Kelly et al., 2007), which is the random version of the additive main effect and multiplicative interaction model (AMMI) (Gauch, 1988; van Eeuwijk, 1995; van Eeuwijk et al., 1995). Recently, factor analytic model was suggested to be useful in GxE investigations in animal breeding (Meyer, 2009). The factor analytic model was used in international sire evaluation to reduce the number of parameters to estimate in comparison to estimating the full genetic variance-covariance matrix across countries (Tyrisevä et al., 2011). Our study used two-step factor analytic model to identify environmental parameter causing GxE. The advantage of using factor analytic model is the ability to analyse unknown common factors, which can be correlated to known EPs (van Eeuwijk et al., 2001) as shown in our study. The unknown common factor can be regarded as either a single factor or a composite of environmental factors, because several environmental factors may contribute to GxE between environments.

The unknown common factor in this study explained genetic variance in harvest body weight differently across environments. For instance, the unknown common factor explained only 26.2 % of the total additive genetic variance in BWH recorded in PE but 63% in FI. The variation in percentage of explained additive genetic variance indicates the presence of GxE. The percentage of additive genetic variance was in all environments less than 100% indicating that there is more than 1 unknown common factor explaining GxE. Due to a limitation of the number of unknown factors, the second unknown common factor could not be studied unless dimension of the **G** matrix is larger, e.g. 5x5 matrix. This requires an experiment in at least 5 farms or locations. The second unknown common factor is expected to mainly explain additive genetic variance in PE because common factors are orthogonal and V_A in the other environments was mainly explained by the first unknown common factor. Moreover, with a limited number of environments, the correlation between the unknown common factor and the EP may not be accurate and therefore no solid conclusions can be made about EP explaining GxE. Therefore, it is recommended that a higher number of environments should be used in a future GxE research.

Based on AIC and BIC, the factor analytic model was more parsimonious than the reaction norm model, indicating that factor analytic model is the most suitable in our data set. The factor analytic model is suitable when the experiment is not designed for multiple farms per environment, and to study unknown factors common across environments. With more than 5 environments, multiple common factors can be studied (van Eeuwijk et al., 1995).

In conclusion, photoperiod and Day*Degree were identified as environmental parameters causing the strong GxE of BWH in rainbow trout across 4 different environments. Both the reaction norm model and the factor analytic model can help in revealing the environmental parameters responsible for GxE. A factor analytical model is preferred over a reaction norm model when limited information on the variation of EP between farms is available.

6.5 Acknowledgements

This study is a part of the PhD project: Optimization of rainbow trout breeding program for multi-trait selection with GxE interaction. This project is funded by Ananda Mahidol Foundation (The King Scholarship, Thailand) and Troutlodge, Inc., USA. We thank professor dr. ir. Johan A. M. van Arendonk for his useful comments.

References

Akaike, H., (1974). A new look at the statistical model identification. IEEE. Trans. Autom. Control. 19: 716-723.

Ashley, P. J., 2007. Fish welfare: Current issues in aquaculture. Appl. Anim. Behav. Sci. 104: 199-235.

Coulibaly, I., Gharbi, K., Danzmann, R.G., Yao, J., Rexroad III, C. E., 2005. Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. Anim. Genet. 36, 309-315.

Duchesne, P., Godbout, M. H., Bernatchez, L.. 2002. PAPA (package for the analysis of parental allocation): a computer program for simulated and real parental allocation. Mol. Ecol. Notes. 2: 191-193.

Fikse, W. F., Rekaya, R., Weigel, K. A., 2003. Assessment of environmental descriptors for studying genotype by environment interaction. Livest. Prod. Sci. 82: 223-231.

Gauch, H. G., Jr., 1988. Model selection and validation for yield trials with interaction. Biometrics 44: 705-715.

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Thompson, R., 2009. ASReml user guide release 3.0. VSN International Ltd. Hemel Hempstead, HP1 1ES, UK.

Johnson, N. A., Rexroad III, C. E., Hallerman, E. M., Vallejo, R. L., Palti, Y., 2007. Development and evaluation of a new microsatellite multiplex system for parental allocation and management of rainbow trout (*Oncorhynchus mykiss*) broodstocks. Aquaculture 266: 53-62.

Kelly, A. M., Smith, A. B., Eccleston, J. A., Cullis, B. R., 2007. The Accuracy of Varietal Selection Using Factor Analytic Models for Multi-Environment Plant Breeding Trials. Crop Sci. 47: 1063-1070.

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, Jorjani, H., 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. Acta. Agric. Scand. 52: 11-24.

Maluwa, A. O., Gjerde, B., Ponzoni, R. W., 2006. Genetic parameters and genotype by environment interaction for body weight of *Oreochromis shiranus*. Aquaculture 259: 47-55.

Martinez V, Kause, A., Mäntysaari, E. A., Mäki-Tanila, A., 2006. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: I. One-stage selection. Aquaculture 254: 182-194.

Meyer, K., 2009. Factor-analytic models for genotype × environment type problems and structured covariance matrices. Genet. Sel. Evol. 41: 1-11.

Mulder, H. A., Bijma, P., 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83: 49-61.

Mulder, H. A., Veerkamp, R. F., Ducro, B. J., van Arendonk, J. A. M., Bijma, P., 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. J. Dairy. Sci. 89: 1740-1752.

Palti, Y., Fincham, M.R., Rexroad III, C.E., 2002. Characterization of 38 polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Mol. Ecol. Notes. 2, 449-452.

Price, E. O., 1999. Behavioral development in animals undergoing domestication. Appl. Anim. Behav. Sci. 65: 245-271.

Rexroad, C.E., Rodriguez, M., Coulibaly, I., Gharbi, K., Danzmann, R., De Koning, J., Phillips, R., Palti, Y., 2005. Comparative mapping of expressed sequence tags containing microsatellites in rainbow trout (*Oncorhynchus mykiss*). BMC Genomics 6, 54.

Rexroad, C.E., Coleman, R.L., Hershberger, W.K., Killefer, J., 2002. Eighteen polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Anim. Genet. 33, 76-78.

Sae-Lim, P., Komen, H., Kause, A., Martin, K. E., Crooijmans, R., van Arendonk, J. A. M., Parsons, J. E., 2012a. Enhancing selective breeding for growth, slaughter traits and overall survival in rainbow trout (*Oncorhynchus mykiss*). Accepted.

Sae-Lim, P., Komen, H., Kause, A., Mulder, H. A., van Arendonk, J. A.M., Martin, K. E., Barfoot, A. J., Parsons, J. E., Davidson, J., Rexroad III, C. E.. 2012b. Genotype-by-environment interaction of growth traits in rainbow trout (*Oncorhynchus mykiss*): A continental scale study. Submitted.

Schaeffer, L. R., 2004. Application of random regression models in animal breeding. Livest. Prod. Sci. 86: 35-45.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statistics. 6:461–464.

Sumpter, J. P. 1992. Control of growth of rainbow trout (*Oncorhynchus mykiss*). Aquaculture 100: 299-320.

Taylor, J. F., Migaud, H., Porter, M. J. R., Bromage, N. R., 2005. Photoperiod influences growth rate and plasma insulin-like growth factor-I levels in juvenile rainbow trout, *Oncorhynchus mykiss*. Gen. Comp. Endocr. 142: 169-185.

Taylor, J. F., North, B. P., Porter, M. J. R., Bromage, N. R., Migaud, H., 2006. Photoperiod can be used to enhance growth and improve feeding efficiency in farmed rainbow trout, *Oncorhynchus mykiss*. Aquaculture 256: 216-234.

Tyrisevä, A. M., Mayer, K., Fikse, F. W., Ducrocq, V., Jakobsen, J., Lidauer, M., Mäntysaari, E., 2011. Principal component and factor analytic models in international sire evaluation. Genet. Sel. Evol. 43: 1-10.

van der Waaij, E. H., 2004. A resource allocation model describing consequences of artificial selection under metabolic stress. J. Anim. Sci. 82: 973-981.

van Eeuwijk, F. A., 1995. Linear and bilinear models for the analysis of multi-environment trials: I. An inventory of models. Euphytica 84: 1-7.

van Eeuwijk, F. A., Keizer, L. C. P., Bakker, J. J., 1995. Linear and bilinear models for the analysis of multi-environment trials: II. An application to data from the Dutch maize variety trials. Euphytica 84: 9-22.

van Eeuwijk, F. A., Cooper, M., DeLacy, I. H., Ceccarelli, S., Grando, S., 2001. Some vocabulary and grammar for the analysis of multi-environment trials, as applied to the analysis of FPB and PPB trials. Euphytica 477-490.

Zwald, N. R., Weigel, K. A., Fikse, W. F., Rekaya, R., 2003. Identification of factors that cause genotype by environment interaction between herds of Holstein cattle in seventeen countries. J. Dairy. Sci. 86: 1009-1018.

7

General discussion

7.1 General introduction

The aims of this thesis were 1) to quantify GxE interaction using an optimal design, 2) to define breeding objective that serves diverse local markets and farming environments, and 3) to optimize breeding program when GxE interaction exists.

The consensus of trait preferences was combined with the breeder's equation, which resulted in a method to quantify desired genetic gains of the most important traits in rainbow trout. The desired genetic gains can be used to derive breeding goal weights when the genetic parameters of all traits are available (Chapter 2). Genetic parameter estimates for GxE interaction should be unbiased and precise. The experimental design for GxE interaction was simulated and found that when the h^2 is moderate (h^2 = 0.3), fewer fish will be needed for each environment (100 family with 10 per each). In contrast, when the h^2 of a studied trait is low (h^2 = 0.1), around 20-25 fish per family are required to avoid biased estimate of genetic correlation of a trait measured in 2 different environments (Chapter 3). In rainbow trout, two-stage selection is sometimes implemented to enhance the genetic response to selective breeding, especially for growth. We found that two-stage selection can be used to increase genetic gain of body weight at harvest (BWH) when fish are pre-selected for body weight at tagging (BWT). However, this is not true for thermal growth coefficient (TGC) from tagging to harvest due to a negative genetic correlation between BWT and TGC (Chapter 4). Re-ranking on breeding values of body weight measured in different time points was found in Chapter 4, i.e. BWT and BWH. Postponing pre-selection for BWT by using a later body weight measurement that has a positive correlation with TGC will increase genetic gains for both BWH and TGC. However, a balance needs to be found between the increased gain obtained by this method and the increased common environmental effects, which reduce the accuracy of selection. Based on the simulation study on the optimal GxE design in Chapter 3, an experiment was performed to investigate GxE interaction for growth traits (BWT, BWH and TGC) of rainbow trout reared in 4 different environments located in 3 continents. Results of this study were reported in Chapter 5 and 6. Significant GxE interaction was found for all growth traits. Subsequently, the environmental parameters (EP) responsible for causing GxE interaction for BWH were studied using 2 models: linear reaction norm model, and factor analytic model. Photoperiod and Day*Degree, i.e., the multiplication between water temperature (^{o}C) and rearing period (day) were identified as the most likely responsible EP for GxE interaction for BWH.

In this chapter, findings in Chapters 2 to 6 are used to study how to optimize the breeding program for GxE interaction. First, the consequence of GxE interaction on

genetic gain is demonstrated through deterministic simulation. Second, the consequence of re-location of the single breeding program on resulting overall genetic gain across environments was studied. Third, sib information from production environments was used to optimize a single breeding program under multiple simulated situations. In the final paragraph, the overall genetic gain from different strategies is discussed.

7.2 Consequence of genotype-by-environment interaction

Most previous GxE interaction studies showed that GxE interaction for production traits in various fish species is moderate-to-weak. In rainbow trout, moderate GxE is present (r_g = 0.61) between fresh and brackish water environments in body weight measured at 2 years of age (Kause et al., 2003). Similarly, moderate to weak GxE interaction (r_g = 0.67 to 0.96) was found in body weight of rainbow trout measured between 2 different diets (Kause et al., 2006; Le Boucher et al., 2011). In Atlantic cod (*Gadus morhua*), weak GxE interaction (r_g = 0.82 to 0.94) was found for 2-year body weight measured in 3 different locations off the coast of Norway (Kolstad et al., 2006). In European sea bass (*Dicentrarchus labrax* L), weak re-ranking (r_g =0.70 to 0.99) was reported in body weight measured from 4 farms in different locations: France, Israel, Italy, and Portugal (Dupont-Nivet et al., 2008). In tilapia (*Oreochromis shiranuis*) weak re-ranking (r_g = 0.74) was found for body weight measured between fish grown at high and low altitudes (Maluwa et al., 2006). Our results for rainbow trout, grown in different production systems across different continents, show heterogeneity of genetic variances, heritability and correlations of multiple traits across environments, and strong genotype re-ranking across environments. It may be difficult to draw a general conclusion for magnitude of GxE interaction in fish breeding, as the genetic parameters are dynamic by generations, populations, traits, and environments (Gjedrem, 2005; Vehviläinen et al., 2008; Khaw et al., 2012). Nevertheless, previous studies have clearly indicated the existence of GxE interaction in body weight of different fish species across environments as shown in previous studies and ours.

GxE interaction leads to different consequences: heterogeneity of additive genetic variances, and re-ranking of genotypes across environments. Heterogeneity is less important in genetic evaluation in fish than for example in dairy cattle breeding program where selection candidates are selected across multiple environments (Hill, 1984; Meuwissen et al., 1996). In fish breeding, selection candidates are evaluated in a single environment, but heterogeneity can lead to bias in prediction

of genetic responses. Re-ranking is more the concern of fish breeders due to reduction of genetic gains in production environments (Mulder and Bijma, 2005). In this section, the consequence of GxE interaction in term of strong re-ranking is evaluated by predicting genetic gain. For simplicity, body weight at harvest (BWH) measured in 4 different environments: breeding environment (NUC), a recirculating aquaculture system in Freshwater Institute (FI), a high altitude farm in Peru (PE), and a cold water temperature environment (German farm, GE), was used under the situation of using performance information from the nucleus only (i.e. no optimization for GxE interaction). The genetic parameters from Chapter 5 (Summarized in Table 7.1 and 7.2) were used in the deterministic simulation in SelAction (Rutten et al., 2002).

Table 7.1 Phenotypic variance (V_P), heritability (h^2) common environmental effect (c^2) and their standard error (SE) for body weight at harvest (BWH) measured in breeding environment (NUC), recirculating aquaculture system (FI), high altitude farm (PE), and cold water farm (GE).

Environment	V_P	h^2	SE (h^2)	c^2	SE (c^2)
NUC	9035	0.19	0.10	0.06	0.03
FI	6127	0.18	0.11	0.07	0.04
PE	11212	0.18	0.09	0.04	0.03
GE	6149	0.22	0.09	0.01	0.03

The consequence of re-ranking has been mainly considered in the context of single trait selection or one-stage selection Mulder and Bijma, 2006; Mulder et al., 2006; Bijma and van Arendonk 1998; Bijma et al., 2001). However, in rainbow trout, two-stage selection is sometimes implemented in a breeding program (Martinez et al., 2006b; Sae-Lim et al., 2012).

Table 7.2 Genetic correlation (r_g) of BWH measured in NUC, FI, PE, and GE.

Environment	Environment		
	FI	PE	GE
NUC	0.41	0.19	0.48
FI		0.40	0.51
PE			0.43

The classical multi-stage selection is to keep the initial number of animals and number of selection candidates fixed. Strong pre-selection leads to lower genetic gain than one-stage selection due to insufficient information source, e.g. pre-

selection (mass selection) for bull before progeny testing is available (Schrooten et al., 2005).

Box 1: Deterministic simulation

Deterministic simulation was used to predict the genetic gain of BWH. The breeding objective (*H*) was defined as: $H = BWH_{NUC}$, where BWH_{NUC} is body weight at harvest measured at NUC. Breeding program is in the nucleus station, where selection candidates are evaluated and selected. Initial population was 50,000 fish and was reduced to 150 selection candidates (50 males and 100 females) at final selection. The number of full-sib families was 100 (1 male: 2 females mating design), producing 500 (250 males and 250 females) offspring per dam. For two-stage selection, pre-selection for BWT aims at enhancing genetic gain for BWH_{NUC}. At the first stage of selection, the population of 50,000 fish was reduced to 4,000 based on BWT (selection index of the first stage: $I_1 = b_1 \times OP_{BWT,NUC}$). Proportion of selection in the first stage: P_1 was 8%. The first-stage selection is typically practiced within families at BWT. However, it was not possible to implement within-family selection for pre-selection in SelAction. Therefore mass selection (own performance) was used. In the second stage, selection was based on index:

$$I_2 = b_1 \times OP_{(BWT,NUC)} + b_2 \times OP_{(BWH,NUC)} + b_3 \times FS_{(BWH,NUC)} + b_4 \times HS_{(BWH,NUC)} + b_5 \times EBV_{sire} + b_6 \times EBV_{dam},$$

where *OP* is own performance, *FS* is information from full-sibs, *HS* is information from half-sibs, *EBV* is estimated breeding values from sire and dam. Proportion of selected animals at the final stage was 2.5 % in male ($P_{2,M}$) and 5% in female ($P_{2,F}$). The total proportion of selected animals was 0.2% in male ($P_{total,M}$) and 0.4% in female ($P_{total,F}$). For one-stage selection, 40 fish per family (20 males and 20 females) from 100 families was randomly selected (P_1 = 1) at tagging. The final stage of selection was the same as two-stage selection. The total proportion of selected animals in male: $P_{total,M}$ was 2.5% and in female: $P_{total,F}$ was 5%.

Multi-stage selection can enhance genetic gain when initial population size is enlarged (Schrooten et al., 2005) because the proportion of selected animals after

pre-selection is lower. Two-stage selection in trout is slightly different from other livestock species. In trout, pre-selection typically takes place at tagging, and consists of within family selection based on phenotype, i.e., body weight at tagging (BWT) and equal number of fish per family selected. This pre-selection lowers rearing costs because fewer fish are kept and tagged using passive integrated transponder (PIT) tags. Alternatively, the number of fish before tagging can be increased while the number of fish reared until harvest is fixed. This will result in higher selection intensity at pre-selection and higher genetic gain with the same costs. Pre-selection within family only exploits the variation of Mendelian sampling terms and therefore leads to minimum inbreeding rate (Falconer and Mackay, 1996). Final selection is based on multi-trait BLUP-EBV index. Two-stage selection has higher selection intensity than one-stage selection and should result in enhanced genetic gain (Martinez et al., 2006b; Sae-Lim et al., 2012).

Consequence of re-ranking of genotypes to genetic gain may differ according to different selection methods, i.e., one-stage or two-stage selection when the genetic correlation between BWT measured in NUC and BWH measured in the production environments is not consistent across environments. To study this, the response to selection for BWH in NUC and correlated responses of BWH measured in FI, PE, and GE were compared. Simulation (Box 1) was run under 4 different situations: 1) no re-ranking (r_g fixed to be 1) and 2) re-ranking (r_g equal to the estimates in Chapter 5). For two-stage selection, the correlations between BWT in NUC and BWH in all environments were obtained from Chapter 5. The results are shown in Table 7.3.

When GxE interaction is absent, genetic gains from one-stage selection range from 33.3g (FI) to 45.1g (PE) whereas two-stage selection increased genetic gains in NUC, FI, and GE. This indicates that two-stage selection enhances selective breeding. The difference in absolute gains is different due to different genetic parameters, especially genetic variance is different. However, in PE, two-stage selection reduced genetic gain (from 45.1 to 41.2 g). This is because of a weak negative genetic correlation between BWT measured in NUC and BWH measured in PE (r_g = -0.06: Chapter 5). When GxE interaction is present, genetic gains from one-stage selection ranged from 8.6 (PE) to 41.6 (NUC) g. As expected, a large reduction of genetic gain occurred in PE.

Table 7.3 Genetic gain (in gram) of BWH measured in 4 different environments after one stage and two-stage selection, in the presence or absence of GxE interaction.

Selection	Environment	Genetic gain (g)	
		No re-ranking*	Re-ranking
One-stage	NUC	41.6	41.6
	FI	33.3	13.7
	PE	45.1	8.6
	GE	36.9	17.7
Two-stage	NUC	53.1	53.1
	FI	38.9	20.2
	PE	41.2	6.3
	GE	44.4	26.1

*r_g was fixed to 1. If there is re-ranking, r_g in Table 7.2 was used. The correlations between tagging body weight (BWT) measured in breeding environments (NUC) and BWH in all environments (NUC, recirculating aquaculture system in Freshwater Institute: FI, high altitude farm in Peru: PE, and cold-water farm in Germany: GE) were taken from Chapter 5. The genetic gain is per generation.

Two-stage selection however led to higher genetic gain in NUC, FI and GE in comparison to one-stage selection when GxE interaction existed. In contrast, two-stage selection led to lower genetic gain (6.3 g) than one-stage selection (8.6 g) in PE.

It can be concluded that the genetic gain in the other environments are lower than the expected genetic gain in NUC, due to GxE interaction, and that two-stage selection does not always increase genetic gain across environments. It is possible to exploit postponing pre-selection (Chapter 4) for BWT in NUC because r_g between later body weight measurement and BWH in PE may increase and become positive (Rutten et al., 2005; Su et al., 2002). However, postponing pre-selection may increase common environmental effect because fish are kept in the family tanks for a longer period. Genetic correlation between multiple body weight measurements in NUC and BWH measured in production environments should be investigated. To reduce the impact of GxE interaction on genetic gain for BWH in 3 production environments (FI, PE, and GE), the breeding program should be optimized. In the next paragraphs, several strategies to do this are compared. For simplicity, one-stage selection will be used to evaluate solely the effect of G x E on genetic gain in BWH.

7.3 Optimization of breeding program in the presence of GxE interaction

With large numbers of full-sibs per family, it is possible for a single trout breeding program to distribute eggs across continents, with high diversity in farming management, local markets, and rearing environments. GxE interaction may be weakly detected, depending on magnitude of differences among macro-environments (Calus, 2006; Mulder, 2007). Only when there is strong re-ranking of animals, i.e. r_g < 0.61 (Mulder et al., 2006), re-ranking needs to be taken into account for the breeding plans (Gjedrem, 2005). The threshold value of GxE interaction for decision-making whether or not a single breeding program should be separated is discussed in more detail in Chapter 5. The expectation of the threshold r_g in fish breeding is higher (≥ 0.70) than in dairy cattle breeding program, i.e., ranging from 0.61 (Mulder, 2007) to 0.70 (James, 1961). It is a challenge to improve fish genetically for very different production environments across continents by using a single breeding program, as in the case of trout in the present study.

One way to increase genetic gain in all environments before optimizing breeding program is to manipulate environmental parameters (EP) causing GxE interaction for growth traits. In Chapter 6, photoperiod and Day*Degree were identified as they may cause GxE interaction for BWH of rainbow trout in 4 environments. The EP in production environments may differ from breeding environment or even be suboptimal. When the animals are in the suboptimal environment, most energy is allocated to fitness to be able to adapt and to survive. More energy will be allocated to growth when the rearing environment is optimal (van der Waaij, 2004). It is possible to manipulate the photoperiod, in a commercial environment in order to enhance growth in rainbow trout kept in freshwater (Taylor et al., 2006). This approach may be used to reduce GxE interaction caused by a photoperiod that deviates from breeding environment. Day*Degree is assumed to be the same when the water temperature and feeding regime are the same across environments. Previous study has shown that it is possible to change EP, e.g., photoperiod, but it may be costly to implement. In addition, it depends on willingness of the fish farmers to cooperate with the breeding company. Nevertheless, fish farmers may have limited facilities and financial support to devote to manipulation of EP. The fish farmers' investment may depend on the increased profit after changing responsible EPs.

In the next section, different ways to optimize the breeding program for BWH by using alternative approaches are discussed, i.e., re-location of breeding program,

incorporating sib's performance in the selection, and separated breeding programs. The alternative approaches can be used under different situations and they will be useful for a future breeding plan for any fish breeding programs.

7.3.1 Re-location of breeding program

With a single breeding program, relocation of the breeding program may increase the overall genetic gain in all environments. The overall genetic gain ($\Delta \boldsymbol{G}$) is the weighted genetic gain in each environment. The environment that yields the maximum ΔG is the environment where the breeding program should be located. The breeding objective was the same as in section 7.2. The ΔG can be calculated as:

$$\Delta G = w_1 \times \Delta G_{NUC} + w_2 \times \Delta G_{FI} + w_3 \times \Delta G_{PE} + w_4 \times \Delta G_{GE} \quad [1]$$

where w_i is standardized relative weight (v_i): ($\sum_i^4 v_i = 100$). This relative weight can refer to economic importance, economic values, proportion of egg sold per environment, etc. However, in this section, equally important weight ($v_i = 25\%$) was used as a demonstration. The w_i was standardized using additive genetic standard deviation of each environments (Chapter 5): $w_i = 25/\sigma_{A,i}$, where $\sigma_{A,i}$ is genetic standard deviation in the *i*th environment to account for heterogeneity of additive genetic variances among environments (Chapter 5). When substituting w_i into this equation:

$$\frac{w_i \sigma_{A,i}}{\sum_{i=1}^{4} w_i \sigma_{A,i}} \quad , \quad [2]$$

the relative weight v_i will always be equal to 25. There was no sib information from other environments, except from the environment where the selection took place. Genetic parameters in Table 7.1 and 7.2 were used to predict the genetic gains. Table 7.4 shows the overall genetic gain of BWH when the breeding program was re-located.

The total genetic gain for BWH is highest when the breeding program is located in Germany (67.14 g). The average genetic gain is also highest in GE (25.80) but it is difficult to express, as the unit of relative importance is unknown. However, the

ranking of ΔG and average genetic gain is the same. The low c^2 estimate in GE (c^2 = 0.01: Table 7.1), may partly contribute to higher ΔG in GE compared to the other environments. Lower c^2 and slightly higher h^2 for BWH in GE lead to higher accuracy of estimated breeding value and thus higher accuracy of selection.

Table 7.4 Genetic gain of BWH when the breeding program is re-located in FI, PE, and GE.

Environment	Location of breeding program			
	NUC	FI	PE	GE
NUC	41.64	16.68	8.04	22.09
FI	13.69	32.60	13.57	18.81
PE	8.58	17.64	45.89	21.46
GE	17.74	18.42	16.15	40.85
ΔG*	52.12	56.93	51.40	67.14
Average#	20.41	21.33	20.91	25.80

#calculated based on $\Delta G = w_1\Delta G_{NUC} + w_2\Delta G_{FI} + w_3\Delta G_{PE} + w_4\Delta G_{GE}$ and $w_i = 25/\sigma_{A,i}$, therefore the unit is not in g* trait unit of g.

Both $\Delta \boldsymbol{G}$ and average genetic gain may be deviated from the predicted numbers shown in Table 7.3 when the environments are not equally weighted, e.g. when one environment is considered more important than the others. This approach can be used without sib testing in the other production environments. However, moving a breeding program is costly. Therefore, cost-benefit analysis should be conducted prior to decision-making.

7.3.2 Sib's performance

Sib's performance information can be used to account for performance in different environments by combining information from selection candidates, measured in NUC, and sib information from other environments into the selection index (Mulder and Bijma, 2005). In this section, BWH measured in different environments was treated as different traits. The relative weight in NUC (v_1) was investigated to understand how much v_1 should be in NUC compared to the other environments to increase genetic gain in all environments. The breeding objective was defined as:

$$H = w_1 \times g_{BWH,NUC} + w_2 \times g_{BWH,FI} + w_3 \times g_{BWH,PE} + w_4 \times g_{BWH,GE}, \quad [3]$$

where g_{BWH} is the true breeding value of BWH in different environments, and $w_i = v_i/\sigma_{A,i}$. In NUC, the v_1 was varied from 0 to 100%. The v_i in the other production environments (FI, PE, and GE) was treated as of equal economic importance: $\frac{(100-v_1)}{3}$. The genetic gain of BWH in all environments was plotted against v_1 as shown in Figure 7.1.

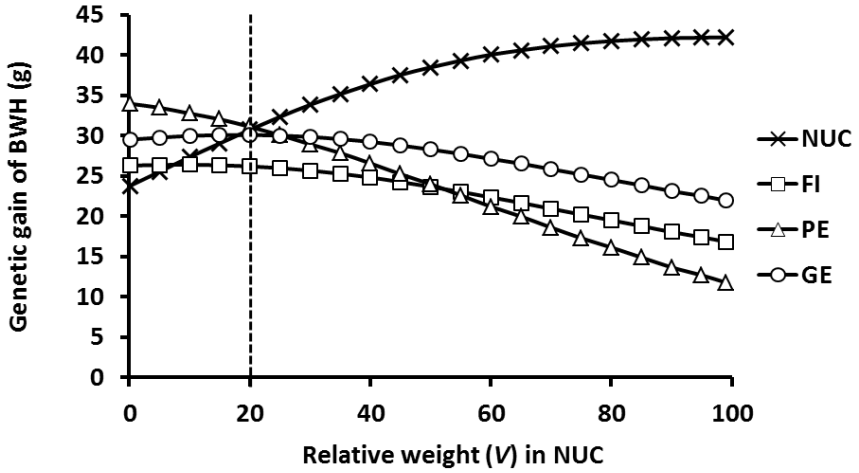

Figure 7.1 Genetic gain from the breeding program located in NUC optimized using sib's performance from all production environments (FI=recirculating aquaculture system, PE=high altitude farm, and GE = low water temperature farm). The y-axis is the genetic gain of BWH in the trait unit (g). The x-axis is the relative weight ($\boldsymbol{v_1}$) in NUC prior to standardizing. The relative weight in the other environment will be equal and ranging from 33.3 to 0 %. The vertical broken line indicates the relative weight in the NUC that results in the most similar genetic gain across environments.

The result shows a trade-off between genetic gain in NUC and genetic gain in the other environments. When lower relative weight was given to BWH measured in NUC, the genetic gain in the other environment increased at different rate, i.e., PE versus GE and FI. The lower c^2 in GE (Chapter 5) and higher h^2 contributed to overall higher genetic gain in GE than in FI. When the relative weight in NUC was at 25%, the genetic gain in NUC (32.3 g) is still slightly higher than the gain in the other environments (FI =26.0 g, PE = 30.1 g, and GE = 30.0 g). When the relative weight was 20% in NUC (vertical broken bar: Figure 7.1), the gain in PE was higher (31.1 g) than gain in NUC (30.7 g). Therefore, to maintain genetic gain in NUC equal

to or greater than genetic gain in any other environments, relative weight of 20 to 25% in NUC is the best compromise, leading to almost equal genetic gain in all environments. However, if the breeding company needs to increase gain significantly in other environments, genetic gain in NUC will be sacrificed (relative weight in NUC < 20%).

The genetic gains from the breeding program optimized by using sib's performances were compared to the gain from the program without any optimization. To measure the change in genetic gain in all environments, the genetic gain of BWH from the optimized breeding program with sib information was compared with the genetic gain from the breeding program located in NUC without sib information , i.e. as shown in Table 7.3 for one-stage selection: Percentage of genetic gain (for each environment) = $\frac{\Delta G\ (optimization)}{\Delta G\ (no\ optimization)} \times 100$ (Figure 7.2).

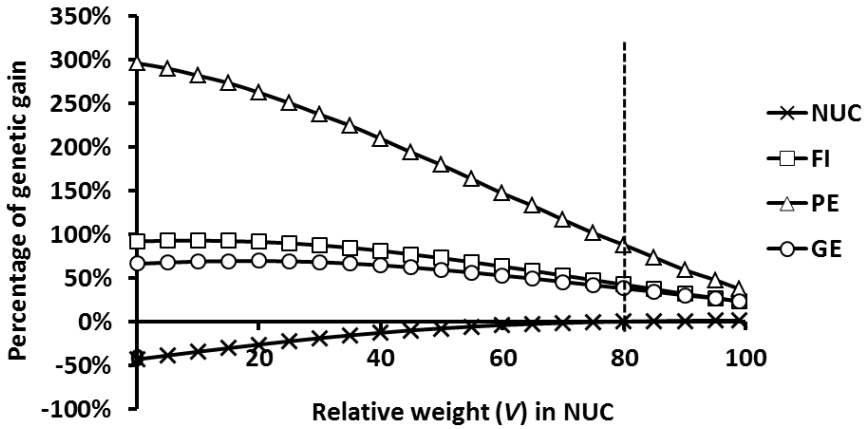

Figure 7.2 percentage of change in genetic gain for BWH (%) between breeding program with and without sibs' performance (y-axis). The x-axis is the relative economic weight ($\boldsymbol{v_1}$) in NUC. The broken vertical line indicates the relative weight given approximately zero change in NUC (0.2%).

When the relative weight in NUC was 80% (vertical broken line: Figure 7.2), the change of genetic gain in NUC was close to zero (0.2%) comparing to the breeding program without optimization). This indicates that it is possible to maintain genetic gain in NUC equal to the program without sib information while the genetic gain in the other environments is increased (FI = 42.5%, PE= 87.9%, and GE = 38.1%). The percentage of genetic gain in NUC was negative when the relative weight was

below 80% (Figure 7.2). As expected, the rate of change in gain in PE was faster than any other environments due to strong GxE interaction of BWH between NUC and PE.

Based on these results, it can be concluded that when other environments are equally important (ranging from 0 to 33.3%) and sib information is available from production environments, a single breeding program can be optimized, leading to increased genetic gain across environments. The relative weight should be between 20 to 80% to optimally increase all gain without too much reduction of genetic gain in NUC. Optimization of Finnish breeding programs for body weight in freshwater and seawater using sib information has been studied in rainbow trout using different population structures, i.e., different mating designs (Martinez et al., 2006a). The situation is similar to the cases described in Mulder and Bijma (2005) with a nucleus and one production environment. In this work, the considered situations reflected mostly chicken or pig breeding programs. In dairy cattle, studies have been carried out to optimize breeding programs in the presence of GxE (Meuwissen and Woolliams, 1993; Mulder and Bijma, 2006; Mulder et al., 2006) and in the combined crossbred and purebred selection for different species, e.g. broiler and pig (Bijma and van Arendonk, 1998; Jiang and Groen, 1999). When GxE interaction exists in the breeding program, the reduction of genetic gain is mainly due to lower accuracy of selection (Mulder and Bijma, 2005). Incorporating more information into the selection index results in higher accuracy of selection for fish that perform well in multiple environments.

The optimization method can be used in the cases where no other breeding companies are selling eyed eggs into a particular market. However, in reality market competition is typically present. It is likely that there is market competition in each production environment which means that genetic gain for BWH of rainbow trout from the competitor may be higher than the genetic gain in the own breeding company (de Vries, 1989). The relative weight ($\boldsymbol{v_i}$) can be increased when the genetic gain needs to be increased for a particular environment because of strong competition for sales. This will result in more emphasis of sibs' performance in production environments where competition is strong.

Box 2 Pseudo-BLUP selection index

The selection index in this section had information of *OP* of BWH measured in NUC, *FS* and *HF* in NUC and 3 production environments, EBV of parents, which equivalent to pseudo-BLUP selection index (Campos-Montes et al., 2009; Valdez-Nava et al., 2011). The selection index in SelAction (Rutten et al., 2002) is therefore equivalent to:

$$I = w_1 \times EBV_{BWH,NUC} + w_2 \times EBV_{BWH,FI} + w_3 \times EBV_{BWH,PE} + w_4 \times EBV_{BWH,GE,}$$

where w_i is the same as described in equation [1]. Thus, the more weight (w_i) given into BWH in a certain environment (equation [3]) will emphasize the EBV of BWH in that environment, leading to higher genetic response.

Three scenarios were simulated: **Scenario A** means that 1 of the 4 environments requires much higher relative weight (94%) than the others (2% per each environment). This scenario is for a single environment that has a competitor selling higher performance eggs and thus the breeding company needs to outcompete the performance of competitor's fish stock. **Scenario B** shows a situation in which NUC and 1 of the 3 production environments (FI, PE, or GE) have equally moderate relative weight (48%) and the others have relative weight of only 2% each. This scenario shows competitive position in multiple environments, but NUC is always one of the environments that are emphasized. **Scenario C** was the same as B but with emphasis on 2 production environments (48% per each) not on the NUC (2%).

In scenario A, the genetic gain for BWH in NUC is 42.1 g whereas genetic gain in PE is 12.7 g when the relative weight in NUC is 98%. In contrast, the genetic gain for BWH in NUC is 12.8 g whereas genetic gain in PE is 41.6 g when the relative weight in PE is 98%. This shows a possibility to increase genetic gain for BWH either in NUC or PE to out compete the performance of the fish from the competitors (Figre 7.3: A). In scenario B (NUC & PE), the simulation indicated that it is also possible to push genetic gain simultaneously (NUC = 34.3 g, and PE =32.5 g) for competing in both environments. In scenario B, the genetic gain in NUC tended to increase (37.3 to 38.2 g) when FI (26.6 g) or GE (30.4 g) received equally relative weight as NUC (48%). In contrast, when relative weight is low in NUC (2%) but high in PE (48%) with any other environments in scenario C, the genetic gain in PE was high (36.1 to 36.7 g) and higher than the gain of PE in scenario B.

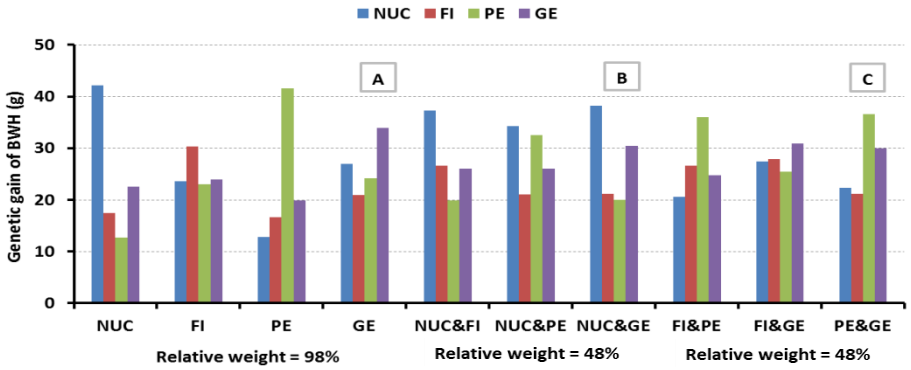

Figure 7.3 The genetic gain of BWH (g) across environments when the relative weight is varied. NUC = breeding environment, FI = recirculating aquaculture system, PE= high altitude farm, and GE = low water temperature farm. There are 3 scenarios: A = a single environment has extremely high economically relative weight than the others (94%), B = NUC and 1 of the other environments have equal moderate relative weight (48%), and C = B but other production environments not in NUC.

This section demonstrates the possiblity to optimize breeding program based on sibs' performance to increase genetic gain in certain environments where the performance of the fish from the competitor is higher. Incorporating sibs' performance in the index and vaying the relative weights in the breeding goal leads to a dynamic tool that allows a breeding company to change strategies over the years as the position of the company to its competitors changes. This example is for a single breeding objective for BWH. It is possible to use the desired genetic gain approach e.g. Chapter 2 to derive the relative weight as shown in literature (Cunningham et al., 1970; Yamada et al., 1975; Brascamp, 1984; Gibson and Kennedy, 1990). The desired genetic gain may be to have higher gain for BWH than the competitors in a certain environment.

In Chapter 2, variation of breeding trait preferences existed among rainbow trout fish farmers. This suggests that BWH may not be the most important in the environment with high GxE interaction with NUC. For example, in PE, condition factor, fillet percentage and FCR are more important than growth , based on individual preference values of condition factor = 0.272, fillet percentage = 0.248, FCR =0.178, and growth = 0.114 (results not shown in Chapter 2). Thus GxE interaction of BWH between NUC and PE seems to be less important. However, trait preferences can be dynamic values and questionairres to adjust the preference values should be routinely carried out to monitor changes.

Optimization of breeding program for GxE interaction by including sib's performance in the selection index may have some limitations. Firstly, the increase of genetic gains is a trade-off, especially where there is very strong GxE interaction of BWH (NUC and PE: Figure 7.3, Scenario A). If the fish from a competitor can grow faster and the genetic gain in BWH is not enough to outcompete this competitor, this method are no longer relevant as the economic value becomes zero. In such situations, the emphasis in the breeding program should shift to traits for which the company does have a market advantage (de Vries, 1989). Secondly, it is more serious than the situation above when the competitors have selective breeding programs in the environments where there is very strong GxE interaction of BWH with PE. It may not be possible anymore to increase genetic gain for BWH based on sibs' performance to outcompete local egg suppliers. The situation is even worse when all environments have local competitors. However, based on Chapter 5, selective breeding for BWH from competitor located in PE may lose market in NUC due to strong GxE interaction for BWH. The alternative breeding stategy is environment-specific breeding programs which are normally very costly.

7.3.3 Environment-specific breeding program and a comparison of breeding strategies

Environment-specific breeding program that give emphasis only in a single environment is the strategy to be used when the genetic gain cannot be higher than the genetic gain in fish from competitors. In fish breeding, fish normally produce a number of offspring per family which is enough to have the same number of families for each environment-specific breeding program. However, the cost may be approximately *n* times due to *n* number of separated breeding programs. Facilities from NUC may be shared with established environment-specific program to reduce fixed costs. Alternatively, it is also possible to divide fish families into 2, e.g. 50:50 families per each environment. However, rate of inbreeding may increase when the low number of families used per environments (Bijma et al., 2001; Kincaid, 1976, 1983; Su et al., 1996). In this section, 4 environment-specific breeding programs were located in NUC, FI, PE, and GE. Selection was as before, i.e. no sib information and one-stage selection only. The breeding objective in each environment was to improve BWH. Genetic gain from separated breeding programs (Strategy 4) for BWH is predicted based on the population structure described in Box 1 (one-stage selection) and is compared to

the baseline strategy (no optimization, breeding in the NUC) and the 2 other strategies previously discussed (Table 7.5).

Table 7.5 A comparison of 4 different breeding strategies with respect to genetic gain ($\Delta \boldsymbol{G}$) of body weight at harvest (BWH) and different economically relative weights ($\boldsymbol{v_i}$). All breeding strategies are based on one-stage selection.

Environment*	Strategy 1		Strategy 2		Strategy 3		Strategy 4	
	$\boldsymbol{v_i}$	$\Delta \boldsymbol{G}$	$\boldsymbol{v_i}$	$\Delta \boldsymbol{G}$	$\boldsymbol{v_i}$	$\Delta \boldsymbol{G}$	$\boldsymbol{v_i}$	$\Delta \boldsymbol{G}$
NUC	100	41.6	20	30.7	80	41.7	100	41.6
FI	0	13.7	26.7	26.2	6.7	19.5	100	32.6
PE	0	8.6	26.7	31.1	6.7	16.1	100	45.9
GE	0	17.7	26.7	30.1	6.7	24.5	100	40.9

* NUC = breeding environment, FI = recirculating aquaculture system, PE = high altitude farm, and GE = low water temperature farm. Strategy 1 = no optimization in the presence of GxE interaction and a single breeding program is located in NUC. Strategy 2&3 = optimized breeding program based on sib's performance. Strategy 2 is a trade-off of genetic progress in NUC to other production environments and gain in NUC is restricted to be similar to the other environments. Strategy 3 is to maintain gain in NUC to be the same as in Strategy 1. Strategy 4 is 4 separated breeding programs.

Strategy 1 is the breeding program without any optimization leading to the lowest cost, compared to the other strategies, but it results in lower genetic gain of BWH in all production environments. Strategy 2&3 are the optimized breeding program based on sibs' performance from 3 sib testing stations located in FI, PE, and GE. Strategy 3 results in higher genetic progress in all environments and it is possible to increase genetic gain in PE even more. For the sib testing, Strategy 2 has higher total genetic gain (118.1g) than Strategy 3 (101.8g) but it is more costly than Strategy 1. Strategy 4, providing the highest genetic gain of BWH across environments is the most costly for the breeding company.

In general, it is assumed that the breeding goal leading to the maximum profit for the fish farmers will also give the maximum profit to the breeding company. However, this assumption is arguable, for example, it depends on the salability of the eggs/fingerlings and the competitive position of the breeding company (de Vries, 1989). In the current situation of a global rainbow trout breeding program, there are two different profits: profit for the company and profit for farmers. Rainbow trout farmers across environments may prefer as fast growing fish as possible, e.g., Strategy 4. However, the maximum profit for the breeding company is from salability of high performance eggs/fingerlings, produced at a minimum

cost. When the trout farmers prefer the most expensive strategy, the profit at the company will reduce due to the higher cost of the breeding program. If the breeding company increases the price per eyed egg unit, the farmers may not be able to buy, depending on the revenue at farms. This shows a conflict of both profits from the trout farmers and the breeding company. The optimal solution for both profits is to find the most optimal point that satisfies both farmers and the company. A cost-benefit analysis based on macroeconomics can be used to find the optimal solution (e.g., Boardman et al., 2011; Mankiw, 2009). An example of a macroeconomic model for making optimal decision is shown in Figure 7.4.

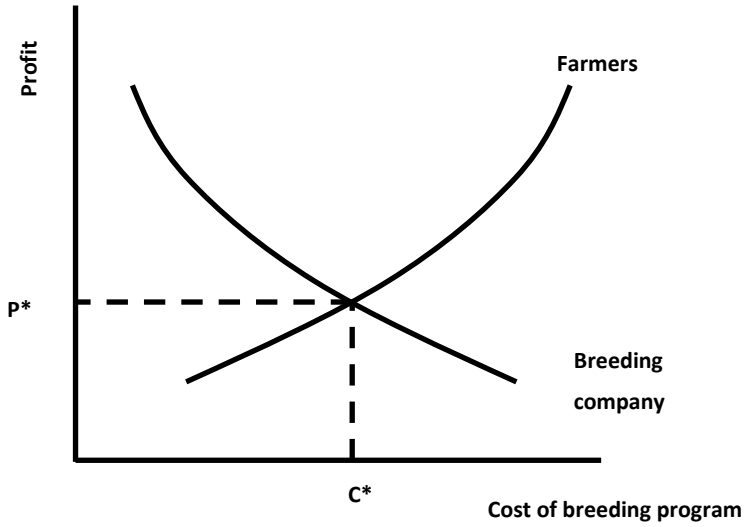

Figure 7.4 An example of macroeconomic model explains profit of farmers and breeding company (modified from Mankiw, 2009 and Andrew J. Barfoot, personal communication, October 2012). Y-axis is the profit in US dollars and X-axis is the cost of breeding company. The P* is an optimal profit for both farmers and breeding company with the optimal cost (C*) for the breeding program. The profits from farmers and breeding company are equalized in unit.

Assuming that Figure 7.4 is the equilibrium of both profits and there is not a shift of the curves, as the cost of the breeding program increases, the profitability of the breeding company will decrease. Thus, the slope of the line is negative in Figure 7.4. At the same time, the more the company spends on the breeding program to increase genetic gain, the more advantages it will be to the farmers. These advantages will result in increased profits for the farmer, thus the positive upward

slope of the line in the graph. The optimal expenditure on the breeding program is going to be the point at which these two curves intersect (as shown by the dashed black lines). This is because the farmers will only pay a premium for the eggs of the breeding program up to and until the point at which the increased profitability equals the added expense of obtaining eggs from the breeding program. For example, if the farmer must pay $1 more for the eggs from the breeding program, he must expect to receive at least $1 of added benefit. If he only receives $0.99 of added benefit (profit), he will no longer buy the more expensive eggs (Andrew J. Barfoot, personal communication, October 2012). The cost-profit can be analyzed more sophisticatedly using real data. Strategies may be combined to suit the best actual situation, leading to the most optimal decision for both the breeding company and the fish farmers.

Trout farms are typically located in rural areas and can contribute to local income, nature conservation and rural development through specialized production methods (e.g. organic, smoked/processed). Fish breeding programs that serve global markets, like those for rainbow trout, should take variation of customer preferences, local conditions, and the possible occurrence of G x E interaction into account. Ignoring local market conditions and requirements will not result in the production of "one size fits all" fish, and less opportunity to increase the sustainability (Olesen et al., 2000) and profitability of all farms combined.

7.4 Conclusions

Our study shows that for a global trout breeding company, genotype-by-environment interaction for growth of rainbow trout in different continents should be considered and investigated/monitored. It is possible to optimize a single breeding program for genotype-by-environment interaction using alternative strategies as presented in this thesis. The optimal strategy depends on the competitive position and optimum point of cost-benefit for both farmers and the global breeding company. Farmers' preferences on breeding traits should be accounted for in the breeding strategy to satisfy a wide array of trout farmers. Optimization of the breeding program for both genotype-by-environment interaction and farmers' preferences will contribute to a sustainable aquaculture and increased well-being for both trout farmers and farmed rainbow trout in the world.

7.5. Acknowledgements

Antti Kause, Han A. Mulder, Hans Komen, Johan A. M. van Arendonk, Andrew J. Barfoot, and Kyle E. Martins are acknowledged for their helpful comments, and suggestions in optimization of breeding program. I would like to thank Piter Bijma for his helpful theoretical explanation on SelAction.

References

Ashley, P. J., 2007. Fish welfare: Current issues in aquaculture. Appl. Anim. Behav. Sci. 104: 199-235.

Bijma, P., van Arendonk, J. A. M., 1998. Maximizing genetic gain for the sire line of a crossbreeding scheme utilizing both purebred and crossbred information. Anim. Sci. 66: 529-542.

Bijma, P., Woolliams, I. A., van Arendonk, J. A. M., 2001. Genetic gain of pure line selection and combined crossbred purebred selection with constrained inbreeding. Anim. Sci. 72: 225-232.

Boardman, A. E., Greenberg, D. H., Vining, A. R., Weimer, D. L.., 2011. Cost-benefit analysis: concepts and practice. p 560 pp. Prentice Hall, Boston, M.A.

Brascamp, E. W., 1984. Selection indices with constraints. Anim. Breed. Abstr. 52: 645-654.

Calus, M. P. L., 2006. Estimation of genotype x environment interaction for yield, health and fertility in dairy cattle. Ph.D. Diss. Wageningen Univ., Wageningen.

Cunningham, E.P., Moen, R. A., Gjerdrem, T., 1970. Restriction of selection indexes. Biometrics 26:67-74.

de Vries, A. G., 1989. A method to incorporate competitive position in the breeding goal. Animal Science 48: 221-227.

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., 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., Mackay, T. F. C., 1996. Introduction to Quantitative Genetics. p 464 pp. Longman, Essex, UK.

Gibson, J. P., Kennedy, B. W., 1990. The use of constrained selection indexes in breeding for economic merit. Theor. Appl. Genet. 80: 801-805.

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

Hill, W. G., 1984. On selection among groups with heterogeneous variance. Anim. Sci. 39: 473-477.

James, J. W. 1961., Selection in two environments. Heredity 16: 145-152.

Jiang, X. S., Groen, A. F., 1999. Combined crossbred and purebred selection for reproduction traits in a broiler dam line. J. Anim. Breed. Genet. 116: 111-125.

Kause, A., Ritola, O., Paananen, T., Mantysaari, E., Eskelinen, U., 2003. Selection against early maturity in large rainbow trout *Oncorhynchus mykiss*: the quantitative genetics of sexual dimorphism and genotype-by-environment interactions. Aquaculture 228: 53-68.

Kause, A., Tobin, D., Houlihan, D. F., Martin, S. A. M., Mäntysaari, E. A., Ritola, O., and Ruohonen, K., 2006. Feed efficiency of rainbow trout can be improved through selection: Different genetic potential on alternative diets. J. Anim. Sci. 84: 807-817.

Khaw, H. L., Ponzoni, R. W., Hamzah, A., Abu-Bakar, K. R., Bijma, P., 2012. Genotype by production environment interaction in the GIFT strain of Nile tilapia (*Oreochromis niloticus*). Aquaculture 326-329: 53-60.

Kincaid, H. L., 1976. Effects of inbreeding on rainbow trout populations. T. Am. Fish. Soc. 105: 273-280.

Kincaid, H. L., 1983. Inbreeding in fish populations used for aquaculture. Aquaculture 33: 215-227.

Kolstad, K., Thorland, I., Refstie, T., Gjerde, B., 2006. Genetic variation and genotype by location interaction in body weight, spinal deformity and sexual maturity in Atlantic cod (*Gadus morhua*) reared at different locations off Norway. Aquaculture 259: 66-73.

Le Boucher, R., Quillet, E., Vandeputte, M., Lecalvez, J. M., Goardon, L., Chatain, B., Médale, F., Dupont-Nivet, M., 2011. Plant-based diet in rainbow trout (*Oncorhynchus mykiss* Walbaum): Are there genotype-diet interactions for main production traits when fish are fed marine vs. Plant-based diets from the first meal? Aquaculture 321: 41-48.

Maluwa, A. O., Gjerde, B., Ponzoni, R. W., 2006. Genetic parameters and genotype by environment interaction for body weight of *Oreochromis shiranus*. Aquaculture 259: 47-55.

Mankiw, N. G., 2009. Macroeconomics. Worth, New York.

Martinez, V., Kause, A., Mäntysaari, E. A., Mäki-Tanila, A., 2006a. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: I. One-stage selection. Aquaculture 254: 182-194.

Martinez, V., Kause, A., Mäntysaari, E. A., Mäki-Tanila, A., 2006b. The use of alternative breeding schemes to enhance genetic improvement in rainbow trout: II. Two-stage selection. Aquaculture 254: 195-202.

Meuwissen, T. H. E., De Jong, G., Engel, B., 1996. Joint estimation of breeding values and heterogeneous variances of large data files. J. Dairy. Sci. 79: 310-316.

Meuwissen, T. H. E., Woolliams, J. A., 1993. Responses of multi-trait selection in open nucleus schemes for dairy cattle breeding. Anim. Sci. 56: 293-299.

Mulder, H. A., 2007. Methods to optimize livestock breeding programs with genotype by environment interaction and genetic heterogeneity of environmental variance. Ph.D. Diss. Wageningen Univ., Wageningen.

Mulder, H. A., Bijma, P., 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83: 49-61.

Mulder, H. A., Bijma, P., 2006. Benefits of cooperation between breeding programs in the presence of genotype by environment interaction. J. Dairy. Sci. 89: 1727-1739.

Mulder, H. A., Veerkamp, R. F., Ducro, B. J., van Arendonk, J. A. M., Bijma, P., 2006. Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. J. Dairy Sci. 89: 1740-1752.

Olesen, I., Groen, A.F., Gjerde, B., 2000. Definition of animal breeding goals for sustainable production systems. J. Anim. Sci. 78: 570-582.

Rutten, M. J. M., Bijma, P., Woolliams, J. A., 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.

Rutten, M. J. M., Komen, H., Bovenhuis, H., 2005. Longitudinal genetic analysis of Nile tilapia (*Oreochromis niloticus* L.) body weight using a random regression model. Aquaculture 246: 101-113.

Sae-Lim, P., Komen, H., Kause, A., Martin, K. E., Crooijmans, R., van Arendonk, J. A. M., Parsons, J. E., 2012. Enhancing selective breeding for growth, slaughter traits and overall survival in rainbow trout (*Oncorhynchus mykiss*). Accepted.

Schrooten, C., Bovenhuis, H., van Arendonk, J. A. M., Bijma, P., 2005. Genetic progress in multistage dairy cattle breeding schemes using genetic markers. J. Dairy. Sci. 88: 1569-1581.

Su, G. S., Liljedahl, L. E., Gall, G. A. E., 1996. Effects of inbreeding on growth and reproductive traits in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 142: 139-148.

Su, G. S., Liljedahl, L. E., Gall, G. A. E., 2002. Genetic correlations between body weight at different ages and with reproductive traits in rainbow trout. Aquaculture 213: 85-94.

Taylor, J. F., North, B. P., Porter, M. J. R., Bromage, N. R., Migaud, H., 2006. Photoperiod can be used to enhance growth and improve feeding efficiency in farmed rainbow trout, *Oncorhynchus mykiss*. Aquaculture 256: 216-234.

van der Waaij, E. H., 2004. A resource allocation model describing consequences of artificial selection under metabolic stress. J. Anim. Sci. 82: 973-981.

Vehviläinen, H., Kause, A., Quinton, C., Koskinen, H., Paananen, T., 2008. Survival of the currently fittest: Genetics of rainbow trout survival across time and space. Genetic 180: 507-516.

Villanueva, B., Wray, N. R., Thompson, R., 1993. Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. Anim. Sci. 57: 1-13.

Wray, N. R., Hill, W. G., 1989. Asymptotic rates of response from index selection. Anim. Sci. 49: 217-227.

Yamada, Y., Yokouchi, K., Nishida, A., 1975. Selection index when genetic gains of individual traits are of primary concern. Jpn. J. Genet. 50: 33-41.

S

Summary

Summary

Global fish breeding distribute improved animal material to several continents to be farmed under diverse environments and for very different market conditions. Yet, fish breeding programs are at the initial phases of globalization.

When establishing a global breeding program, there is a need to assess whether or not a single common breeding objective satisfies the markets across different countries and production systems. To direct genetic changes in multiple traits, selection index weights have been commonly calculated using profit functions. When animal production expands to novel areas or when a novel trait to be selected which does not have a direct economic impact, profit functions need to be replaced by the desired genetic gains method to obtain relative index weights. Using this approach, it is possible to first derive genetic gains that the market requires and then back-calculate the relative weights to be used to get these gains.

Even if the breeding objective would be shared across countries, it may be challenging to develop a single fish stock that performs well across all environments. This is because there may be genotype-by-environment (GxE) interactions. GxE interaction is a phenomenon describing the possibility that different genotypes have a different sensitivity to changes in an environment. GxE interaction has two different forms: genotype re-ranking across environments and heterogeneity of genetic variances. Re-ranking is more serious than heterogeneity of genetic variance because re-ranking means that a single genotype is not superior across all environments. The degree of re-ranking is quantified by the strength of a genetic correlation (r_g) of a trait measured from different environments. When there is no re-ranking ($r_g = 1$), selection in one environment leads to parallel genetic response in all environments, allowing an easy development of a single superior population. However, with increasing levels of re-ranking ($r_g < 1$), it becomes more difficult to develop a population that is superior across all environments. The overall objective of this project was to develop an optimized global breeding program for rainbow trout (*Oncorhynchus mykiss*) in terms of 1) a balanced breeding goal that satisfies preferences of trout producers (Chapters 2 and 7), and 2) maximized genetic gains across environments (Chapter 7) in the presence of GxE interaction in production traits (Chapters 3, 4, 5, and 6). In **chapter 2**, distributing animals from a single breeding program to a global market may not satisfy all producers, as they may differ in market objectives and farming environments. Analytic hierarchy process (AHP) can be used to estimate preferences, which can be aggregated to consensus preference values using weighted goal programming (WGP). The aim of this study was to use an AHP-WGP based approach to derive desired genetic gains for rainbow trout breeding and to study whether breeding trait preferences vary depending on production environments and end-products produced by farmers. The analysis revealed that the 6 most important traits were thermal growth coefficient (TGC), survival (Surv), feed conversion ratio (FCR), condition factor (CF), fillet percentage (FIL%), and late maturation (LMat). Individual trait preferences are different for farmers having different farming

environments and producing different end-products. Calculating consensus preference values resulted in the following ranking of traits: Surv (0.271), FCR (0.246), TGC (0.246), LMat (0.090), FIL% (0.081), and CF (0.067). Corresponding desired genetic gains (in % of a trait mean) were 1.63, 1.87, 1.67, 1.29, 0.06, and 0.33%, respectively. In **chapter 3**, a stochastic simulation was used to study the effect of population structure on bias and precision of genetic correlation estimates (r_g) between two environments. The simulated r_g was 0.0, 0.5, or 0.8 for a trait with heritability (h^2) of either 0.3 or 0.1 in both environments. Simulation results showed that r_g is biased downwards especially when the simulated r_g is 0.8, heritability is 0.1, and family size is less than 10. A downward biased r_g incorrectly suggests the existence of G×E interaction. The optimal design with the lowest mean square error for r_g for a trait with low h^2 requires a large family size (20–25) and a low number of families (100–80 or 50– 40 for population size fixed to 2000 and 1000 animals, respectively). For traits with moderate h^2, the optimal family size is 10 with 200 or 100 families for population size fixed to 2000 and 1000, respectively. We also studied the effect of selective mortality on G×E estimates. Schemes with unequal family sizes due to differences between families in survival produced similar results for the optimum design compared to schemes with equal family sizes. Simulations using equal family sizes can thus be used to determine the optimal design for estimating G×E interaction. In **chapter 4**, enhancing selection using two-stage selection is normally implemented by pre-selection for tagging weight (BWT) and by final selection for harvest weight (BWH) and thermal growth coefficient from tagging to harvest (TGC_{TH}). However, the efficiency of two-stage selection on genetic response in BWH and TGC_{TH} is dependent on their genetic (r_g) and phenotypic (r_p) correlations with BWT and therefore dependent on the time point of pre-selection. The aim of this chapter was first to estimate h^2, r_p, and r_g for BWT (at 7 months of age), BWS (weight at sorting, at 9 months), BWH (at 14 months), and TGC_{TH}. Second, the estimates of h^2, r_p, and r_g were used in deterministic simulation to compare genetic responses in BWH and TGC_{TH} in one- and two-stage selection schemes. Genetic correlation of BWT was 0.35 with BWH but -0.25 with TGC_{TH}, whereas the r_g of BWS was 0.72 with BWH but 0.39 with TGC_{TH}. Pre-selection for BWS led to genetic response of 54.1 g in BWH which was higher than the genetic response from pre-selection for BWT (51.9 g). Similarly, pre-selection on BWS enhanced correlated genetic response in TGC_{TH} to 0.30 $g^{1/3}/^oC*day$. In contrast, pre-selection for BWT resulted in lower correlated genetic response in TGC_{TH} of 0.20 $g^{1/3}/^oC*day$. Genetic improvement of BWH and TGC_{TH} can be enhanced by postponing pre-selection to a later age. However, an optimal time point for tagging and pre-selection should be found to minimize common environmental effects and rearing costs during communal rearing of full-sibs. In **chapter 5**, the aim of this chapter was to quantify the magnitude of GxE interaction in growth traits (BWT, BWH and TGC) of rainbow trout, grown in 4 different environments: nucleus in Washington State, USA (NUC), a recirculating aquaculture system in West Virginia, USA (FI), a high-altitude farm in Peru (PE), and a cold-water

farm in Germany (GE). The results show that heterogeneity of additive genetic variances was mainly found for BWH measured in FI and PE. Strong genotype re-ranking was found for all traits (BWT: r_g = 0.15 to 0.37, BWH: r_g = 0.19 to 0.48, TGC: r_g = 0.31 to 0.36) across environments. This implies that selection based on performance information from the nucleus only is inefficient in increasing fish performance in the other environments. Moreover, the r_g of BWT in NUC with BWH and TGC in both FI (0.31 and 0.10) and GE (0.36 and 0.20) were positive. However, r_g was negative between BWT in NUC and both BWH (-0.06) and TGC (-0.20) in PE. This implies that pre-selection for BWT may not increase harvest traits in Peru. Accounting for GxE interaction in the breeding program either by using sib-information from test stations or by establishing environment-specific breeding programs would increase genetic gains for environments that differ significantly from the nucleus environment. In **chapter 6**, reaction norm (RN) and factor analytic (FA) models were both used to identify which of the following environmental parameters (EP), age at harvest, water temperature, oxygen and photoperiod, correlated with the observed GxE for BWH reported in Chapter 5. Reaction norm quantifies GxE along the specific EPs. In FA model, an unknown common factor causing GxE is first identified, and then correlations between the common factor and EPs are calculated. Photoperiod and the combination of days to harvest multiplied with daily temperature (Day*Degree) were identified by the RN model as the EPs causing most GxE. With the two-step FA model Day*Degree was the most likely EP responsible for GxE. Results showed that the degree of GxE was the highest between Peru and the other environments, and that photoperiod and Day*Degree were the most likely EP's explaining the differences between environments. The Akaike information criteria and Bayesian information criteria indicated that FA model was more parsimonious than the RN model. The low variation in EPs reduced the power of RN model to identify EPs, and future studies should use multiple farms per production environment. Factor analytic model is preferred over a RN model when only limited information on the variation of EP between farms is available. In **chapter 7**, a deterministic simulation was used to investigate alternative breeding program designs in the presence of GxE across environments Results from chapter 5 showed that selection for BWH in a nucleus breeding program with information from the nucleus only would result in reduced genetic gains of BWH in FI, PE and GE due to GxE interaction. Therefore, alternative approaches to optimize a single breeding program for GxE interaction were proposed: changing EPs, re-location of a breeding program, incorporating sib's performance in the selection, and separated breeding programs. This part showed diverse possibilities to increase genetic gain of BWH in all environments. Manipulating photoperiod to reduce GxE is commercially possible; however, this depends on the extra profit at the production farm. Re-location of breeding program to GE led to highest total genetic gain for BWH. Alternatively, including sib performance into selection index increased genetic gain in all environments. The relative weight should be between 20 to 80% to optimally increase all gain without

too much reduction of genetic gain in nucleus. When the competitive position of the breeding program in one environment is inferior, more weight can be given to emphasize desired genetic response in that environment. Desired gain (higher than that from a competitor) can be used to obtain relative index weights to outcompete the competitor. If the genetic gain is still lower than that of the competitor, environment-specific program can be used, but this is costly. There is a possibility of a conflict between 2 profits: from a breeding company and fish farmers and an optimum solution for that conflict can be found by using macroeconomics and cost-benefit analysis.

Main conclusions of this thesis

1. Farmers' preferences on breeding traits vary across farming environments and local markets. To satisfy most farmers, consensus desired genetic gains can be derived using a combination of analytical hierarchy process and weighted goal programming which can be taken into account in a global breeding strategy.
2. Strong GxE interaction across continents in growth traits of rainbow trout was found; however, alternative breeding scheme designs can be used to account for GxE to increase genetic gain in all environments.

S

Samenvatting

Samenvatting

Wereldwijde visteelt transporteert verbeterd dierlijk materiaal naar verschillende continenten, waar het in verschillende omgevingen en voor verschillende markten geproduceerd wordt. Echter staan fokprogramma's voor wereldwijde visteelt nog in de kinderschoenen. Bij het opzetten van een wereldwijd fokprogramma is het nodig om vast te stellen of één fokdoel volstaat voor de verschillende markten, over landen en teeltsystemen heen. Om genetische verandering in meerdere eigenschappen te sturen, werden selectie index gewichten veelal berekend aan de hand van winst functies. In geval dat dierlijke productie uitbreidt naar nieuwe gebieden of wanneer een nieuw kenmerk zonder directe economische waarde moet worden geselecteerd, moet de winst functie vervangen worden door een gewenste genetische winst methode om gewenste relatieve index gewichten te verkrijgen. Met deze methode wordt de door de markt gewenste genetische vooruitgang bepaald, op basis waarvan de relatieve gewichten terug gerekend worden om deze vooruitgang te verkrijgen. Zelfs wanneer het fokdoel hetzelfde is voor verschillende landen, kan het door aanwezigheid van genotype-bij-omgeving (GxE) interacties een uitdaging zijn om één enkele vispopulatie te fokken welke goed presteert in alle omgevingen. GxE omvat de mogelijkheid dat verschillende genotypen een verschillende gevoeligheid voor omgevingsveranderingen kunnen hebben. GxE interactie kent twee verschillende vormen: genotype herschikking tussen omgevingen and heterogeniteit van genetische varianties. Herschikking is problematischer dan heterogeniteit van varianties omdat herschikking impliceert dat één enkel genotype niet het beste is in alle omgevingen. De mate van herschikking wordt gekwantificeerd door de sterkte van de genetische correlatie (r_g) van een kenmerk gemeten in verschillende omgevingen. Wanneer er geen herschikking is (r_g=1), leidt selectie in de ene omgeving tot parallelle genetische respons in alle omgevingen, wat de ontwikkeling van één enkele superieure populatie gemakkelijk toelaat. Echter, met toenemende mate van herschikking (r_g<1) wordt het moeilijker om een populatie te ontwikkelen welke superieur is in alle omgevingen. Het algemene doel van dit project was om een geoptimaliseerd wereldwijd fokprogramma voor regenboog forel (*Oncorhynchus mykiss*) te ontwikkelen met 1) een gebalanceerd fokdoel dat voldoet aan de voorkeuren van forel telers (hoofdstukken 2 en 7) en 2) gemaximaliseerde genetische vooruitgang in productiekenmerken over omgevingen heen (hoofdstuk 7) in de aanwezigheid van GxE interactie (Hoofdstukken 3, 4, 5 en 6). In **hoofdstuk 2** staat beschreven dat niet alle telers noodzakelijkerwijs tevreden zijn als dieren afkomstig uit één enkel fokprogramma beschikbaar worden gesteld voor de wereldwijde markt, gezien mogelijke verschillen in markten en teeltomgevingen. Een analytisch hiërarchisch proces (AHP) kan gebruikt worden om voorkeuren van telers te kwantificeren, die kunnen samengevoegd worden tot een consensus voorkeurswaarde doormiddel van gewogen doel programmering (WGP). Het doel

van deze studie was om een AHP-WGP gebaseerde methode te gebruiken om de gewenste genetische vooruitgang voor regenboog forel fokkerij te verkrijgen en om te bestuderen of voorkeuren met betrekking tot fok-kenmerken afhankelijk zijn van teeltomgevingen en de eindproducten. De analyse onthulde de zes belangrijkste kenmerken: thermale groei coëfficiënt (TGC), overleving (Surv), voederconversie ratio (FCR), conditie factor (CF), filet percentage (FIL%) en late maturatie (LMat). Individuele fok-kenmerken zijn afhankelijk van teeltomgevingen en de eindproducten. De berekening van de consensus voorkeurswaarden resulteerden in de volgende rangschikking: Surv (0,271), FCR (0,246), TGC (0,246), LMat (0,090), FIL% (0,081) en CF (0,067). De corresponderende gewenste genetisch vooruitgang (in % van het gemiddelde van het kenmerk) was respectievelijk 1,63, 1,87, 1,67, 1,29, 0,06 en 0,33%. In **hoofdstuk 3** werd een stochastische simulatie gebruikt om het effect te bepalen die populatie structuur heeft op de afwijking en de precisie van de geschatte genetische correlatie (r_g) tussen twee omgevingen. De gesimuleerde r_g was 0, 0,5 of 0,8 voor een eigenschap die in beide omgevingen een erfelijkheidsgraad (h^2) had van 0,3 of 0,1. De gesimuleerde resultaten toonden aan dat r_g onderschat wordt, vooral wanneer de gesimuleerde r_g 0,8 is, de h^2 0,1 is en de familiegrootte minde dan tien is. Een onderschatte r_g kan ten onrechte het bestaan van GxE interactie suggereren. De optimale opzet voor een kenmerk met een lage h^2, die resulteert in de laagste mean square error voor r_g, vereist een grote familie (20-25) en een gering aantal families (50-40 en 100-80 voor een gefixeerde populatieomvang van respectievelijk 1000 en 2000 dieren). De opzet voor een kenmerk met een middelmatige h^2 is optimaal wanneer, bij een gefixeerde populatieomvang van 1000 of 2000 dieren, de familiegrootte 10 is met respectievelijk 100 en 200 families. We bestudeerden ook het effect van selectieve mortaliteit op GxE schattingen. Scenario's met ongelijke familiegroottes, veroorzaakt door verschillen in overleving tussen families, en scenario's met gelijke familiegroottes gaven gelijkaardige resultaten voor de optimale opzet. Simulaties waarin een gelijke familiegrootte wordt toegepast, kunnen dus gebruikt worden om de optimale opzet te bepalen met betrekking tot het schatten van GxE interactie. In **hoofdstuk 4** staat beschreven dat verbeterde selectie door middel van een twee-stap's selectie gewoonlijk wordt toegepast door een pre-selectie op gewicht bij merken (BWT) en een tweede selectie op zowel slachtgewicht (BWH) als de thermale groeicoëfficiënt tussen merken en slachten (TGC_{TH}). Echter, de efficiëntie van de twee-stap's selectie op de genetische respons in BWH en TGC_{TH} is afhankelijk van de genetische (r_g) en fenotypische correlatie (r_p) met BWT en is daardoor afhankelijk van het tijdstip van pre-selectie. Het doel van dit hoofdstuk was allereerst om h^2, r_g en r_p te schatten voor BWT (op de leeftijd van zeven maanden), BWS (gewicht bij sorteren, op de leeftijd van negen maanden), BWH (op de leeftijd van 14 maanden) en TGC_{TH}. Vervolgens werden de schattingen van h^2, r_p en r_g gebruikt in een deterministische simulatie om de genetische respons in

BWH en TGC_{TH} in één- en twee-stap's selectie scenario's te vergelijken. De genetische correlatie van BWT was 0,35 met BWH, maar -0,25 met TGC_{TH}, terwijl de r_g van BWS 0,72 was met BWH, maar 0,39 met TGC_{TH}. Pre-selectie voor BWS leidde tot een genetische respons van 54,1 g in BWH, wat hoger was dan de genetische respons bij pre-selectie voor BWT (51,9 g). Gelijkaardig verbeterde de pre-selectie op BWS de gecorreleerde genetische respons in TGC_{TH} tot 0,30 $g^{1/3}/^oC*dag$. Tegengesteld resulteerde pre-selectie voor BWT in lagere gecorreleerde genetische respons in TGC_{TH} van 0,20 $g^{1/3}/^oC*dag$. Genetische vooruitgang van BWH en TGC_{TH} kan verbeterd worden door het uitstellen van pre-selectie. Een optimaal moment voor merken en pre-selectie zou gevonden moeten worden om de gemeenschappelijke omgevingseffecten en de opfokkosten tijdens de gemeenschappelijke opfok van full-sibs te minimaliseren. Het doel van **hoofdstuk 5** was om de grootte van GxE interacties te kwantificeren in groeikenmerken (BWT, BWH en TGC) van regenboog forel geteeld in vier verschillende omgevingen: de nucleus in Washington State, USA (NUC); een recirculatie aquacultuur systeem in West Virginia, USA (FI); een op een grote hoogte gelegen bedrijf in Peru (PE) en een koud-water bedrijf in Duitsland (GE). De resultaten toonden dat de heterogeniteit van additief genetische varianties vooral gevonden werd voor BWH gemeten in FI en PE. Sterke genotype herschikking werd gevonden voor alle kenmerken, over alle omgevingen heen (BWT: r_g = 0,15 tot 0,37, BWH: r_g = 0,19 tot 0,48, TGC: r_g = 0,31 tot 0,36). Dit suggereert dat wanneer selectie enkel gebaseerd is op prestatie informatie van de nucleus, dit inefficiënt is in het verbeteren van de prestatie in andere omgevingen. Verder was de r_g van BWT in NUC met BWH en TGC zowel in FI (0,31 en 0,10) en GE (0,36 en 0,20) positief. Echter was de r_g tussen BWT in NUC en zowel BWH (-0,06) en TGC (-0,20) in PE negatief. Dit suggereert dat pre-selectie voor BWT wellicht niet de slachtkenmerken in Peru verbetert. Rekening houdend met GxE interactie in het fokprogramma, enerzijds door gebruik te maken van sib-informatie in teststations en anderzijds door het opzetten van omgevings-specifieke fokprogramma's, zou de genetische vooruitgang in omgevingen die significant verschillen van de nucleus omgeving kunnen verbeteren. In **hoofdstuk 6** werden zowel een reactie norm (RN) model en factor analytisch (FA) model gebruikt om te identificeren welke van de volgende omgevingsparameters (EP), slachtleeftijd, watertemperatuur, zuurstof en voederperiode, gecorreleerd waren met de geobserveerde GxE voor BWH beschreven in hoofdstuk 5. Het RN model kwantificeert de GxE interactie voor de specifieke EPs. Met behulp van het FA model werd eerst een onbekende gemeenschappelijk factor, verantwoordelijk voor GxE, geïdentificeerd en vervolgens werden correlaties tussen de gemeenschappelijke factor en EPs berekend. Het RN model identificeerde de fotoperiode en de combinatie van dagen tot slacht vermenigvuldigd met dagelijkse temperatuur (Dag*Graden) als de EPs verantwoordelijk voor de meeste GxE. Met het twee-stap's FA model was

Dag*Graden de meest waarschijnlijke EP verantwoordelijk voor GxE. De mate van GxE was het hoogste tussen Peru en de andere omgevingen en de fotoperiode en Dag*Graden waren de meest waarschijnlijke EPs welke de verschillen tussen omgevingen verklaren. Het Akaike informatie criterium en het Bayesiaanse informatie criterium gaven aan dat het FA model meer parsimonieus was dan het RN model. De lage variatie in EPs reduceerde de power van het RN model bij het identificeren van EPs en toekomstig onderzoek zou meerdere bedrijven per omgeving moeten gebruiken. Wanneer enkel gelimiteerde informatie over de variatie van EP tussen bedrijven beschikbaar is, heeft een factor analytisch model de voorkeur over een RN model. In **hoofdstuk 7** werd een deterministische simulatie gebruikt om alternatieve fokprogramma scenario's te bestuderen in aanwezigheid van GxE. Resultaten uit hoofdstuk 5 toonden dat selectie op BWH in een fokprogramma die enkel gebruik maakt van informatie uit de nucleus zou resulteren in verminderde genetische vooruitgang van BWH in FI, PE en GE door GxE interactie. Daarom werden de volgende alternatieven voorgesteld om het fokprogramma te optimaliseren voor GxE.: EPs veranderen, het verplaatsen van het fokprogramma, sib-informatie meenemen in de selectie en gescheiden fokprogramma's. Het aanpassen van de fotoperiode om GxE te reduceren voor BWH is commercieel mogelijk. Echter is dit afhankelijk van de extra winst die het teeltbedrijf maakt. Verplaatsing van het fokprogramma naar GE leidde tot de hoogste totale genetische vooruitgang voor BWH. Het meenemen van sib-informatie in de selectie-index verbeterde de genetische vooruitgang voor BWH in alle omgevingen. Het relatieve gewicht zou tussen 20 en 80% moeten zijn om optimale vooruitgang te boeken zonder een te sterke afname van de genetische vooruitgang in de nucleus. Wanneer de competitieve positie van een fokprogramma inferieur is in één omgeving, kan meer gewicht worden toegekend om de gewenste genetische respons in die omgeving te benadrukken. De gewenste vooruitgang (hoger dan die van een concurrent) kan gebruikt worden om relatieve indexgewichten te verkrijgen om de concurrent te overstijgen. Als de genetische vooruitgang nog steeds lager is dan die van de concurrent kan een omgevings-specifiek programma gebruikt worden, maar dit is duur. Er is een mogelijk conflict tussen de winst van het fokbedrijf enerzijds en de teler anderzijds. Een optimale oplossing voor dat conflict kan worden gevonden middels macro-economie en kosten-baten analyse.

Belangrijkste conclusies van deze thesis:

1. De voorkeuren van telers met betrekking tot fok-kenmerken variëren tussen teeltomgevingen en lokale markten. Om aan de wensen van de meeste telers tegemoet te komen kan een consensus over de gewenste genetische vooruitgang verkregen worden middels een combinatie van een analytisch hiërarchisch proces

en gewogen doel programmering welke opgenomen kan worden in een wereldwijde fokstrategie.

2. Sterke GxE interactie werd gevonden over continenten heen in groeikenmerken van regenboog forel. Echter kunnen alternatieve fokprogramma scenario's gebruikt worden om rekening te houden met GxE om genetische vooruitgang in alle omgevingen te verbeteren.

P

Publications

Peer-reviewed publications

Sae-Lim, P., H. Komen, and A. Kause. 2010. Bias and precision of estimates of genotype-by-environment interaction: A simulation study. Aquaculture 310: 66-73.

Sae-Lim, P. et al. 2012. Defining desired genetic gains for rainbow trout breeding objective using analytic hierarchy process. Journal of Animal Science 90: 1766-1776.

Sae-Lim, P. et al. 2013. Enhancing selective breeding for growth, slaughter traits and overall survival in rainbow trout (*Oncorhynchus mykiss*). Aquaculture 372–375: 89-96.

Conference proceedings

Sae-Lim, P., Komen, H., Kause, A., 2010.Effect Of Different Population Structures On Bias And Precision Of Genotype-By-Environment (GxE) Interaction Estimates: A Simulation Study. 9th world congress genetics applied to livestock production (WCGALP), Leipzig, Germany.

Sae-Lim, P., Komen, H., Kause, A., Barfoot, A.J., Martin, K.E., Parsons, J.E., 2011. Desired genetic gains for a breeding objective: a novel participatory approach. Annual Meeting of the European Association for Animal Production (EAAP), Norway.

Sae-Lim, P., Komen, H., Kause, A., van Arendonk, J.A.M., Martin, K.E., Parsons, J.E., 2012. Effect of pre-selection on genetic responses for growth in rainbow trout (*Oncorhynchus mykiss*). The International Symposium on Genetics in Aquaculture (ISGA) XI, Auburn University, Alabama, USA.

Sae-Lim, P., Mulder H.A., Komen, H., Kause, A., van Arendonk, J.A.M., Martin, K.E., Parsons, J.E., 2012. Genotype-by-environment interaction for growth traits in rainbow trout: A continental scale study. The International Symposium on Genetics in Aquaculture (ISGA) XI, Auburn University, Alabama, USA.

A

About the author

Panya Sae-Lim was born on the 21st of January 1982 in Bangkok, Thailand. In 2000, he graduated from Phrakhanong Pittayalai high school and he started a bachelor program in Aquaculture at Faculty of Fisheries, Kasetsart University (Kasetsart = Agriculture). He also attended the Reserved Affairs Center for five years. In 2004, he graduated from Kasetsart University with first class of honor, and was ranked to acting second lieutenant. He decided continuing his higher education aboard. In 2005, he was awarded the prestigious King Scholarship (Anadamahidol Foundation) from the King of Thailand. In 2006, he started his master program in Animal Breeding and Genetics at Wageningen University. Simultaneously, he joined Thai students association in the Netherlands and actively worked for Thai students. In 2007, he did two master theses. In 2008, he graduated his master degree in Animal Breeding and Genetics, and he continued his Ph.D. which was the collaboration between Troutlodge Inc., USA and Animal Breeding and Genomics Centre, Wageningen University. In 2011, he was elected as the president of Thai Students Association in the Netherlands (TSAN). He organized a successful Thai student academic conference (TSAC) 2012 at Volendam, the Netherlands with the participants from 11 different counties around Europe, the UK, and Japan. The results of the Ph.D. project are shown in this thesis.

Training and Education

Training and Education

The Basic Package (3.0 ECTS)

WIAS Introduction Course	2009
Course on philosophy of science and/or ethics	2009

Scientific Exposure (15.2 ECTS)

International conferences

World Congress on Genetics Applied to Livestock Production (WCGALP),Leipzig, Germany	2010
62nd EAAP Annual Meeting , Stavanger, Norway	2011
The 11th International Symposium on Genetics in Aquaculture (ISGA 2012)	2012

Seminars and workshops

WIAS Sciences Day	2010-2012
Healthy as a (sport)horse, Wageningen, the Netherlands	2011
Friends or Fiends? Consequences of social interactions for artificial breeding programs and evolution in natural populations, Wageningen, the Netherlands	2009
Mini symposium on Advanced Genetics, Wageningen, the Netherlands	2012

Presentations

Troutlodge Inc., USA (3x oral)	2009/2012
WIAS Science day, Wgaeningen, theNetherlands (poster)	2010-2011
WCGALP, Leipzig, Germany (oral)	2010
EAAP, Starvenger, Norway (poster)	2011
ISGA XI (2x oral)	2012
Trout seminar by COPPENS, Hotel Leuther Muhle, Germany (invited speaker)	2012

In-Depth Studies (21.4 ECTS)

Disciplinary and interdisciplinary courses

Design and optimization of animal breeding strategies, Denmark	2009
Getting started in AS-Reml, Wageningen, the Netherlands	2009
Quantitative Genetics with a focus on selection theory , Wageningen, the Netherlands	2010

Survival analysis, Wageningen, the Netherlands	2011
Animal Breeding Schemes, NOVA course, Denmark	2011
Genomic selection in Livestock, Wageningen, the Netherlands	2011
Advanced methods and algorithms in animal breeding with focus on genomic selection, Wageningen, the Netherlands	2012
Advanced statistics courses	
Design of Experiment	2011
Statistic for Life Sciences	2012
PhD students' discussion groups	
QDG (1hour/week)	2008-2009
MSc level courses	
Questionnaire Construction	2009

Professional Skills Support Courses (9.3 ECTS)	year
Research Master Cluster	2008
Course Techniques for Scientific Writing	2011
Project & Time Management	2011
Course Supervising MSc thesis work	2011

Research Skills Training (8.0 ECTS)	year
Preparing own PhD research proposal	2008
External training period, Troulodge Inc., USA	2009/2010

Didactic Skills Training (4.5 ECTS)	
Supervising theses	
MSc major thesis: Gyula Kovács	2011
MSc major thesis: Doreen Lamuno	2012
Tutorship	
reviewer in Research Master Cluster	2011

Training and education total	**61.4 ECTS**

A

Acknowledgements

First of all, I would like to express my deepest gratitude to prestigious scholarship from Anandamahidol Foundation under the Royal Patronage of His Majesty the King of Thailand to pursue my PhD in Animal Breeding and Genetics.

There were many things happened during my PhD (3 years in ABGC-WUR, the Netherlands, 1 year in Troutlodge, Inc., USA, 1 week in Forellenzuncht Trostadt, Germany, and 1 week in MTT, Finland).

Dear people at ABGC-WUR, MTT, and Forellenzucht Trostadt

I would like to thank my promoter: Prof. dr. ir. Johan van Arendonk for the opportunity to continue my PhD at Wageningen University. Also thanks for your contribution and your full support during my PhD, including your teaching in NOVA course, and your critical comments during the PhD meetings and R&O. I appreciated your understanding my difficult situations at the beginning of my Ph.D. and your trust that I can do it.

I would like to express my deepest gratitude to my co-promoters: Dr. ir. Hans Komen and Dr. Antti Kause (MTT, Finland).

Having me as your Ph.D. was a big challenge, wasn't it, Hans? It is always fun when we looked back to the first day we met at the Schiphol airport in August, 2006. It was even fun to see your face when you were saying: "Oh boy". When was the last time I heard that? ☺ Hans, you are a very smart, helpful, and practical supervisor (what a cool combination ☺). I would like to deeply thank for your trust in me. It had to be difficult for you to believe in me whether or not I was going to make it when there was a big difference between me and other students. You always smile (like this ☺) and led me step by step. You saw all my weak points and planned to improve me just like having a selection index on me. I could say that without you, I may not have come this far. You have watched and patiently significantly improved me day by day. Thus this has proven how great in supervision you are. You motivated me and helped me making a lot of novel works in this project and taught me focusing on the quality of the papers. We were working together from a single computer at ABGC to the farm in the middle of nowhere. During the working hours, we worked hard but during the road trip, we learnt each other from our own life experiences. Hans, it is very difficult to find a proper word to explain how great you are. It was a great experience working with, and learning from you. I hope that I will have a chance to work with you more in the future.

Antti  , I am always grateful to have you as my co-promoter. I really learnt a lot from you since the first paper that I wrote. I really appreciate the way you supervised me and the way you gave your responses. You are very good at explaining theories. Personally, you always think thoroughly around me and you even planed things ahead for me. You are a very

kind supervisor and a great teacher simultaneously, not only for the theory but also for the life experience. I always feel lucky to have you as my supervisor and have your full supports when you were in WUR. However, when you went to MTT, Finland, you even worked harder on me to considerably improve me in many aspects. You are a very initiative person and very creative as you came up with many new ideas in this projects. I would like to thanks for the great hospitality when I worked on the final part of my PhD thesis at MTT, Finland. The trip in the forest was fun and the sausages were great☺. I hope I will have a chance to visit you again, especially not for work but for exploring Finland and watching birds.

I have a special thank to Dr. ir. Han A. Mulder who significantly contributed to the GxE interaction parts. Han, you are very smart and patiently explained theories to me when I had any difficulty in understanding complex problems. You were even kind enough to spend significant time on my project and even had many lunch meetings with me (therefore, you could not have lunch/coffee peacefully :P). You even helped me in the questions that I was just curious to know and even arranged a meeting with Dr. Marcos Malosetti from the Plant Breeding group. You are the one who inspired me to do factor analytical mixed model for Animal Breeding.

I would like to thank Forellenzucht Trostadt (Alexander and Johannes Tautenhahn) for their excellent collaboration and hospitality when we were working in your farm. It was a very enjoyable and great experience to work with the fish you farm. We miss your dishes in your restaurant☺. Thanks Dr. ir. Robbert J. W. Blonk, Gyula Kovács, Katrijn Peeters, Chontida Phutaworn and Shripathi Bhat for data collection. I particularly appreciate Gyula Kovács and Bert Dibbits for genotyping work.

Thanks all the ABGC stuffs: Dr. ir. Henk Bovenhuis, Dr. ir. Piter Bijma, Alex Hulzebosch for their basic (but very crucial) knowledge. Thanks QDG initiative (Dr. ir. John Bastiaansen and Dr. ir. Esther Ellen) that I had helpful discussion. Thanks to the paper discussion group, that I had Katrijn Peeters, Dianne van der Spek, Yogesh Paudel, André Marubayashi Hidalgo, Mahlet Anche, and Kim van Hulzen, helped me improve my papers ☺, and TLM (Dr. ir. Marleen Visker) and all the ABGC secretaries, Ada, Monique, and Lisette for their helps.

Of course, my PhD beloved friends! First of all, Dr. ir. Anouk Schurink and Susan Wijga, thanks for all your time and your direct and indirect contributions. It had been nice to have both of you as my friends and my office mates. We shared a lot of stress and helped each other (but you mostly helped me :P). Peng Xue and Trong, we shared the office for a short time but it was fun to talk to you. Thanks Aniek, Sebastian, Mahlet, Yogesh, Yan Yan, Albart, Ewa, Katrijn, Naomi, Irene, Mirte, Hooi-Ling, Ane Marie, Danica, Cécile, Gus, Laurent, Dianne, Sandrine, Mahlet, Nancy, Marzieh, Juan, Ester, André, Hamed, Gabriel, Marcos, Margherita, Ilse, Marcin, Luqman, and the names that I may not be able to cover. It was fun to play badminton with you and on top of that, we did a good job together on WE-DAY.

My beloved people in the US

This project could not have finished without the great collaboration from Troulodge, Inc., USA. I remember the wonderful blessing that they do in every monthly meeting: "*we bless that we can do better so that we can feed more hungers in the world*", said Mr. Jim Barfoot, the President. I would like to express my gratitude to Jim Barfoot and Jim Parsons for giving us (ABGC-WUR) the permission to do the global scale experiment in your company and for your financial supports. I would like to thank my supervisors: Mr. Andrew J. Barfoot and Mr. Kyle E. Martin. They are extremely great supervisors form Troutlodge, Inc., USA.

Kyle, we mostly worked alongside on the field. We had been together in Washington State, Idaho, West Virginia, and (even) Washington, D.C. ☺. We talked and we learnt each other a lot. I really enjoyed working with you all the time. You worked hard and simultaneously planned everything for me and took excellent cares of me. You were always there to support me and made this project moving forward. You are a very polite, kind, and very honest person. It was fun to hang with you and your dad and to watch Seattle Mariners at Seattle, the Puyallup Fair ☺ and even going for fishing in the Pacific Ocean. Thanks to Jim Parsons for the ticket ☺.

Andrew, you are an amazing supervisor! You are a fantastic person I ever met. Why is that? ☺ Because I see your life and your life is wonderful. You are extremely a well organizing person. You are a very active person who spends every second to improve yourself and to contribute to the (world) society. Side-by-side camp (the camp that gives at least one happy moment to the children who have cancer before they (may) pass away) has shown this truth to me. I was very curious whether or not you have ever been tired. You have become my motivation to work hard and well organized things around me. You are always clam and very smart and very keen to the solutions for all the situations. I learnt a lot from you about marketing and economics part of this thesis. Thanks for the great hospitality while working in the US and all the activities you invited me to ☺.

Thanks to Sean M. Nepper, Doug P. Dixon, Mary L. Elliott, Lenore Brown, Brent A. Hinners, Shaun B. Elliott, Paul B. Goodmanson, The Conservation Fund's Freshwater Institute (John Davidson), National Center for Cool and Cold Water Aquaculture (Caird E. Rexroad III), and Piscifactorías de los Andes S.A (Peru) for their considerable contributions to this project.

Apart from work, the Barfoot is a very warm family. They made me home away from my home. I celebrated my (first) 4th of July in your house where I met Jacob, David, Melissa, and Daniel. Jim and Leah, thanks for your amazing love and your deep kindness that you have given to me. Thanks for adopting me as your son and for having shown me the US. I remember all the activities we did. Do you still remember the heart disease menu before I left the US? ☺. Thanks for letting me see Jacob and David getting bigger every year I meet them. Thanks Mellissa and Daniel for all excellent hospitalities every time I got your home.

Thanks Mary Kay Wasserstrom and mom for the haircut. I am always happy to be a part of your family wherever I am.

My Thai community

I would like to thank Thai communities in Wageningen for their direct and indirect supports and encouragements since my master program. I would like to express my gratitude to P Veow and Kess for their hospitality and their supports.
I am certainly happy that I could contribute to the Thai student society in the Netherlands and in Europe in many aspects since my master program. Being the President of Thai Students Association in the Netherlands (TSAN) during the last year of my PhD was a big challenge. However, thanks to all the TSAN staffs for their full support while we were working together.

I would like to express my gratefulness to Nidhiutai family, Atipatya family, Miss Chana Sindhvananda who are being friendly and always take a good care of me.

Finally, I would like to express my deepest gratitude to my family in Thailand for their infinite encouragement and all of my teachers that I could not mention all their names.

Dank U!

Colophon

This research project was financially supoorted by Anandamahidol Foundation under the Royal Patronage of His Majesty the King of Thailand and Troutlodge, Inc., USA.

The photograph of rainbow trout on the cover was kindly provided by Troutlodge, Inc., USA, and was edited by Panya Sae-Lim.

This thesis was printed by GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede, the Netherlands.