Manual: 2013-12

# Strain comparisons in aquaculture species: a manual

R.W. Ponzoni, J.W. James, N.H. Nguyen, W. Mekkawy and H.L. Khaw.

CGIAR
RESEARCH PROGRAM ON
Livestock and Fish

Manual: 2013-12

# Strain comparisons in aquaculture species: a manual

**Authors**
R.W. Ponzoni[a], J.W. James[b], N.H. Nguyen[c], W. Mekkawy[a] and H.L. Khaw[a]

**Affiliation**
[a] WorldFish, PO Box 500 GPO, 10670 Penang, Malaysia
[b] Faculty of Veterinary Science, University of Sydney, Australia
[c] University of Sunshine Coast, Locked Bag 4, Maroochydore DC, Queensland, 4558 Australia

# Table of contents

## Preface

When different strains or breeds of a particular species are available, the best choice is seldom immediately obvious for producers. Scientists are also interested in the relative performance of different strains because it provides a basis for recommendations to producers and it often stimulates the conduct of work aimed at unraveling the underlying biological mechanisms involved in the expression of such differences. Hence, strain or breed comparisons of some sort are frequently conducted. This manual is designed to provide general guidelines for the design of strain comparison trials in aquaculture species. Example analyzes are provided using SAS and SPSS.

The manual is intended to serve a wide range of readers from developing countries with limited access to information. The users, however, are expected to have a basic knowledge of quantitative genetics and experience in statistical methods and data analysis as well as familiarity with computer software. The manual mainly focuses on the practical aspects of design and data analysis, and interpretation of results. The statistical theory can be reviewed in several standard textbooks.

The authors would like to acknowledge the interest of Drs Modadugu V. Gupta, Shrinivas S. Jahageerdar and Eskandar Safari during some phases of the preparation of the present manual. An early version of the manual was used in 'Quantitative Genetics Applied to Fish Improvement' training courses imparted by some of the authors (NHN, HLK, RWP) to staff from partner countries. We are grateful for the input made by the staff from partner countries at the time. In addition, the authors would especially like to thank Dr Malcolm Beveridge for his valuable input editing several versions of the manuscript. He greatly contributed to greater overall clarity in the manual.

WorldFish, Penang
16 December 2012

# 1. Introduction

Aquaculture production has been expanding rapidly in recent decades, especially in developing countries, due to growing demand for high quality protein from aquatic sources. This expansion has contributed in some instances to local food security. However, production in many aquaculture species still depends on stock collected from wild populations. The continued reliance on wild stocks is unwise as the performance of their undomesticated offspring is unpredictable and no long-term gains will accrue from this practice. Thus there is an immediate need to replace wild stock with domesticated and selected stock if productivity is to increase.

Selective breeding is a highly effective and sustainable means of improving productivity. Several successful genetic improvement programs of cultured aquatic animal species have been developed over the last four decades achieving genetic gains between 10% and 20% per generation for economically important traits. Before applying genetic principles to the improvement of any particular species, it is necessary to evaluate the available strains to identify those with superior performance under prevailing production system(s). Presently, formal treatments of experimental designs and statistical methods for the analysis and interpretation of strain comparison trials are limited. This has often led to poorly designed and analyzed strain comparison trials. This manual is, therefore, designed with the intention of remedying that situation. A short account of the main principles has recently been presented by Ponzoni et al. (2011a).

# 2. Features of aquatic animal strain comparison trials

A brief summary of the features of past strain comparison trials is presented in Table 1. Whereas a number of species, including finfish, shellfish, mollusks and crustaceans were used in the trials, the major ones were Atlantic salmon (*Salmo salar*) and other salmonids, as well as tilapia (*Oreochromis niloticus*). Harvest weight or growth-related traits were commonly examined. A number of studies also evaluated reproduction traits, survival rate, sexual maturity, smoltification rate, and immunological parameters.

The experimental scale has differed among studies, with the number of strains ranging from 2 to 37 for pure strain and up to 64 in diallel crossing trials. A similar pattern was found with respect to the total number of families or number of families per strain. There were as many as 62 families per strain or as few as 4 families per strain. The nested mating design was used in most of the studies, with a ratio of one male to 2 or 3 females. Other mating types, such as single pair mating, rectangular factorial mating or incomplete factorial by set have also been employed.

A number of studies tested strains under different production systems, including laboratory conditions with two types of rearing regimes (separate or communal) using tanks, ponds or cages. The majority of the trials used identification at the strain and family levels, but there were some trials with identification at the strain, family and individual levels and some with identification at the strain level only.

Despite the advantage of high prolificacy in aquaculture species and the opportunity for in *vitro* fertilization that would enable application of different mating designs, strain comparison trials are generally limited by the size of the available rearing facilities (Blanc 2003) and lack of efficient and affordable methods of pedigree recording. As a consequence, most often only a fixed effects model can be fitted. In about 65 % of the trials a fixed effects model was fitted for the analysis (Table 1). This has implications for potentially misleading inferences about strain differences, due to either overestimation of precision or loss in statistical power. When identification is not fully recorded, and analysis of variance is carried out to detect between strain differences, the error variance is underestimated (James 1975). It is, therefore, more appropriate to use the genetic residual (e.g. between sire mean square) than the residual (between individuals variance) when testing for the significance of between strains differences. A more valid comparison can be made if the complete pedigree of the individuals is known. Mixed model methodology accounts for environmental effects (Sorensen and Kennedy 1984) and changes in additive genetic variances due to inbreeding, assortative mating and gametic disequilibrium using the numerator relationship matrix (Kennedy 1990), though these latter complications are not likely to be important in most strain comparisons.

**Table 1.** Characteristics of strain comparison trials in aquaculture species.

| Species | No. of Strains | Traits † | No. of families ‡ | Mating Ratio £ | Model | ID § | Source |
|---|---|---|---|---|---|---|---|
| Salmonids | 3 | S | 25♂:75♀ | 1:3 | Mixed | FS | Kanis et al. (1976) |
| Atlantic Salmon | 37 | SR | 306 | 1:3 | Fixed | FS | Refstie et al. (1977) |
| Atlantic Salmon | 37 | BW, L | 308 | 1:3 | Fixed | FS | Refstie and Steine (1978) |
| Atlantic Salmon | 37 | GR, L | | 1:3or 3×3 | Fixed | FS | Gunnes and Gjedrem (1978) |
| Atlantic Salmon | 3 | GR | | 1:3 | Mixed | FS | Bailey and Loudenslager (1986) |
| Atlantic Salmon | 6 | R, GR | 376 | 1:3 | Mixed | FS | Jonasson (1996) |
| Chinook Salmon | 2 | S, GR | 36 | 1:1 | Fixed | FS | Cheng et al. (1987) |
| Chinook Salmon | 3 | BW, L | 144 | 1:3 | Mixed | IFS | Winkelman and Peterson(1994) |
| Chinook Salmon | 2 | S | 72 | 1:2 | Mixed | FS | Unwin et al. (2003) |
| Rainbow Trout | 3 | RP | 27-33/ yr | 1:1 | Fixed | Str | Gall and Gross (1978) |
| Rainbow Trout | 4 | BW, L | 39♂:113♀ | 1:(2-9) | Mixed | FS | Sylvén and Elvingson (1992) |
| Rainbow Trout | 5 | IP | 30 RT | 1:3 | Fixed | FS | Overturf et al. (2003) |
| Rainbow Trout | 3 | S | | | Fixed | IFS | Quinton et al. (2004) |
| Striped bass | 5 | GR | 21 | 1:1 | Mixed | FS | Jacobs et al. (1999) |
| Tilapia | 5 | M | 95 | 1:1 | Fixed | FS | Oldorf et al. (1989) |
| Tilapia | 3 | GR | | 1:1 | Fixed | Str | Romana-Eguia and Doyle (1992) |
| Tilapia | 8 | BW | 200 | 1:1 | Fixed | FS | Eknath et al. (1993) |
| Tilapia | 4 | GR, S | 50♂:100♀ | 1:2 | Fixed | Str | Macaranas et al. (1997) |
| Red Tilapia | 5 | GR | 20 RT | 1:3 | Fixed | FS | Romana-Eguia and Eguia (1999) |
| Common carp | 3 | GR | 72 | (1:3)×3 | Fixed | FS | Vandeputte et al. (2002a) |
| Rohu | 6 | S, GR | 57 | 1:(2-3) | Mixed | FS | Reddy et al. (2002) |
| Scallop | 2 | BW, L,S | 4 | | Fixed | Str | Cruz et al. (1998) |
| Yabby Crayfish | 3 | BW | 35 | 1:4 | Fixed | FS | Jerry et al. (2002) |

† Survival (S), smolt rate (SR), body weight (BW), growth rate (GR), length (L), reproduction traits (RP), return rate (R), maturity (M), immunological parameters (IP).
‡ If total no. of families of all strains is not available, the ratio of total males and females, or the number of replicate tanks (RT) is given in the table.
£ (1:3) × 3 means incomplete factorial mating by three sets each with one male mated to 3 females.
§ Identification (ID) for individual, family and strain (IFS), identification for sib-family (mostly full-sib) (FS) and strain, and identification for only strain (Str).

## 3. General considerations in strain Comparison trials

We need to bear in mind the population whose mean we are trying to estimate. Suppose we compare two stocks, each formed by sampling males and females from two different rivers and mating the males and females from the same river to produce fish for the trial. Then if the stocks to be compared are the hatchery stocks set up from these samples, the parents sampled comprise the whole population, whereas if the aim is to compare the wild stocks which were sampled, the parents involved in the trial are only a sample. The difference can be seen by asking the following question. On the results of this trial, would I choose one of these hatchery stocks, but make no conclusion about the relative merits of the wild stocks? If we would speak of the relative merits of the wild stocks, then the parents are only a sample, and the sampling variance from this source must be taken into account. The same principle applies if only a few animals are taken from a stock to breed the experimental animals.

Hill (1974) argued that using statistical significance testing to decide on the design of a strain comparison was misconceived. He claimed that the purpose of a strain comparison is to select, on the basis of the trial results, the best strain of those under test for further use. Therefore, the appropriate criterion for optimizing the design is the difference between the expected gain from using the strain which performs best in the test and the cost of the experiment. Thus, if we expect to find only small differences among the strains compared, the expected gain will be small and there will be little justification for spending much money on the experiment. This is in contrast with the significance testing approach, which would lead to a requirement of a large experiment to be confident of detecting a small difference as significant.

To use Hill's (1974) approach we need to have some idea of the current usage of the strains, the likely magnitude of their differences in economic value, and, for a given precision of the trial, the likelihood that usage would change in a particular way. It then would be possible to calculate the expected gain in productivity. The final result will depend on how well the real differences between strains are predicted before the trial is run, as well as on industry structures. For example, if the trial is conducted by a breeding organization in order to choose which of various stocks it will use in an improvement program it may decide to choose the best performer in the trial and discard the others. But if the trial compares stocks produced by different breeding organizations, the change in use of these stocks will depend on how the results are perceived by customers of the organization. Whereas formal analysis is likely to be very difficult, an informal consideration

comparing the costs of the trial with possible outcomes would usually be conducted at an early stage of the planning.

It may be, of course, that we do not know even very roughly how large the strain differences are, and in this case it is not possible to design an optimal strain comparison which is intended to lead to a choice between strains. Before this can be done it would be necessary to conduct a pilot trial on a relatively small scale in order to check whether it seems likely that a large well designed strain comparison would be justified. Such a pilot study should be well designed so that no biases are involved, and that the accuracy is as great as possible for the money spent on running it, but the aim would be simply to get estimates of strain differences in order to see whether a larger scale trial should be conducted.

Before the consolidation of the poultry industry into the control of a few large firms, it was common for Government Departments of Agriculture or similar bodies to run random sample tests of egg laying hens. This was done by taking a random sample of eggs from the hatchery of the poultry breeder. The eggs were then incubated at the test site, and, after hatching, hens were run as a group whose eggs were collected, counted and graded, so that the value of the eggs produced during the test period could be determined, as could the sale value of the surviving hens at the end of the trial. All feed consumed by the group was measured so that the cost of feed for the group could be calculated. It was then possible to calculate the difference between the income produced by each group and the cost of feed consumed, and report the result as "income over feed cost". Other costs, such as providing housing, labour, purchase price of chicks and so on, were taken as being the same for each group and thus the income over feed cost value could be taken as a measure of the desirability for potential buyers of the stock from which the sample was taken. For breeders there would be other criteria to consider, such as fertility, but for the buyer of day-old chicks these would matter only if they affected the purchase price. In these trials the comparison was of the stocks commercially available at the time, so a simple random sample of hatchery stock was appropriate.

Many strain comparisons in large domestic animals have been made by crossing males of several strains to a common strain of females. This is often convenient, as one set of females may be readily available, and sires can be introduced by artificial insemination or by introduction of a comparatively small number of animals. Though this may be very convenient, it is inherently inferior to using both males and females of each strain. There are two reasons for this, one statistical and one biological. For an experiment of a given size, the standard error of the strain difference will be the same in both cases, but the actual difference will be only half as large with the cross design, so that the power will be much reduced. To achieve equal power would require an experiment four times as large. Also, with strain crosses the progeny may show heterosis and the degree of heterosis may not be the same for all crosses, causing some bias, and if there are maternal effect differences between strains these will not be detected with the crossing design. Therefore, if it is possible, strain comparisons should be based on full strain differences. In the significance testing approach, one considers the probability of detecting as statistically significant a difference regarded as practically important, so that such an approach is only undertaken if it is judged that the strains are likely to differ in important ways. Small differences will not be of interest.

In fact, it is not really the statistical significance in itself that is important. What is important is to estimate mean differences with the required accuracy. If the accuracy of an estimated difference is such that we cannot be confident that it will have the right sign when the true difference is large enough to be of practical importance, then the experiment is unsatisfactory. This is why the significance testing approach is useful. It will also be helpful, if it is intended to publish the results, to have a significant result in order to convince journal editors and referees.

Conventionally, the significance testing approach to design is as follows. We need to specify four quantities: the size of the mean difference we want to detect as significant; the standard deviation of the measurement; the significance level; the power or probability of obtaining a significant result. However, it is common for experimenters not to have clear ideas on all of these quantities. It is therefore useful to consider a range of possibilities. This will also be helpful because a trial often involves measurements with different relations between mean difference and standard deviation.

The standard deviation is usually fixed. But it is open to the researcher to ask such questions as: If I test at the 5% level, what chance would I have of getting a significant result if the true mean difference is (say) one of four values? Or, what mean differences would be needed to give me powers of 50%, 80%, 90% and 95%? Answering these questions will help determine how to choose an appropriate experiment size. Although there are four variables, the mean difference (D) and the standard deviation ($\sigma$) enter only as their ratio, so there are really only three variables determining experiment size. In many cases there will be several traits each with its own $D/\sigma$ value and it will be necessary to consider size calculations for all of them in order to decide on the design.

Strain trials in plants are usually referred to as plant variety trials and normally include large numbers of varieties. Very often these are inbred lines or crosses so that sampling the base population does not raise problems of adequately representing the genotypes present in the strain. A major concern in plant variety trials is accounting for patterns of soil quality in the trial area. While there are now methods of analysis which can make *post hoc* allowance for such trends there is still advantage in trying to set up a trial with a minimum of confounding between variety and soil quality. In aquaculture the analogous problem is to allow for effects of rearing conditions and similar environmental factors.

Not all strain comparisons are concerned with identifying the best strain for commercial exploitation. Some are conducted in order to find just what differences there are between strains in a number of traits.

An example of a strain trial in which it was clear that there were very large strain differences comes from poultry breeding. Havenstein et al. (1994) studied a well-known commercial strain of broiler chickens sampled in 1991 and a control strain which had been unselected since 1957. The two strains were each fed on two diets: one typical of broiler diets used in 1957 and the other typical of 1991 practice. The object in this work was not to test whether the strains differed, because it was clear that there were very substantial differences between them. Rather, the aim was to partition the gain in productivity over 34 years into components due to genetic change and to improved nutrition, and to test the extent to which the strain differences were dependent on the diet provided. The superiority in six-week weight of the improved strain was 1266 g on the 1957 diet but 1506 g on the 1991 diet, a substantial difference. But the ratio of weights was 3.40 on the 1957 diet and 3.05 on the 1991 diet, so that on a logarithmic scale the difference was much less marked. This example illustrates that it may be important to make comparisons in two or more environments, and that the presence or absence of interaction may depend on the scale of measurement.

Another example of a strain trial which was not concerned with finding a superior genotype comes from psychological studies in mice. A common way of checking for genetic influences on traits in mice is to compare inbred lines, of which there are a large number. Sanford et al. (2003) compared three strains for the effect on sleep patterns of training mice to associate a particular sound with an electric shock. This was done on a small scale (7 to 10 mice per strain) for technical reasons, but the strains were chosen as being very different in anxiety levels so that large

differences were expected, and indeed were found to be statistically significant. Of course, inbred strains have the advantage of being genetically uniform, and strains for comparison can be chosen freely for experimental convenience, not being governed by industry relevance as in species of economic importance.

Despite the fact that strain comparisons may have other purposes, this manual will be primarily concerned with experiments intended to provide information on strains of commercial relevance.

# 4. Design considerations in relation to analysis and interpretation of results

## 4.1 Development of statistical models for analysis: fixed effects, random effects, covariates

Understanding the meaning and the reason for the inclusion of certain factors in the analysis of strain comparisons is important. In this section we define, in simple terms, the main features of model development.

One of the important concepts is that of 'fixed effects' fitted in the model. It is necessary to be able to distinguish them from 'random effects'. Effects are either fixed or random depending on the peculiarities of the effect or on how the levels of the factors that appear in the study were chosen. An effect is called fixed if the levels in the study represent all possible levels of the factor, or at least the levels of interest in the study. In a strain comparison of aquatic animals, sex, batch in which the fish were produced, and environment in which they were reared, would typically be fitted as fixed effects. With fixed effects we are interested in differences between the levels of the effect. For example, regarding sex, most often we have only two levels, and we are generally interested in the weight of males and females and in the magnitude of the difference between the sexes. Strain would also be treated as a fixed effect when we are interested in the relative performance of two or more strains among which farmers may be able to choose for production purposes.

By contrast, an effect is called random when the levels of the factor that are used in the study represent only a random sample of a larger set of potential levels. In such cases the factor effects corresponding to the larger set of levels constitute a population of effects with a probability distribution. In strain comparisons, we may have information about the sires (male parents), the dams (female parents) or, more generally, about the family structure of the individuals that we are trying to compare. Then, the families constitute a sample of all possible families in that strain, and the 'family effect' is treated as random. Also, we may have 'replicates' in an experiment; that is, the experimental units may be replicated in a number of ponds or tanks. In such cases we are not particularly interested in the magnitude of the difference between replicates. Our replicates are a sample of a whole universe of potential replicates, and in consequence, 'replicates' should be treated as a random effect.

We may also need to fit some effects as covariates. For instance, the fish we are dealing with may have been born over a period of six weeks. If we harvest and weigh all of them on the same day or within a few days, there will be variation in age that could be affecting harvest weight. To account for that variation in age in the analysis we can fit age (harvest date minus spawning date) of the fish as a linear or curvilinear covariate.

When there are several fixed factors in the data and there is doubt about whether all of them actually influence the outcome, it is often valuable to run preliminary analyzes to check whether all should be included in the final model for analysis. The reason for this is that if factors with genuine effects are omitted, bias may occur and affect the estimates. On the other hand, if factors with no real effect are included, the adjustments made to account for such effects are simply random errors and reduce the precision of the analyzes. Whereas it is often not possible to find an unambiguous answer, it is useful to analyze the data using a fixed effects model to see which factors appear likely to be important. The preliminary model should include all possible effects such as batch, rearing systems (pond or cage), sex and age. For continuous variables such as age, further analyzes should be performed to investigate their relationships with the dependent variable. For example, the relationships might be linear or non-linear. Perhaps splines will be the best way to fit effects of covariates. If important non-linearity is found, consultation with an experienced statistician is recommended.

After running the complete model, often only the fixed effects that are statistically significant ($P<0.05$) are retained. Retaining effects in the model which in reality do not influence the trait being analyzed simply adds random error to the estimates of the fixed effects, whereas removing non-significant effects which in reality do influence the trait will possibly add bias as well as random error. Judgment must be applied, and subject-matter knowledge may be an important contributor to such decisions.

## 4.2 Definition of strain

There is no consensus on the definition of 'strain' in aquaculture species. Often the terms line, strain, stock and population are used interchangeably. In this manual, we use the term strain. Gunnes and Gjedrem (1978) define strain as 'a discrete breeding population from a river, river system, or a fjord leading to a river.' Based on this definition, any discrete breeding population from a hatchery may also be termed a strain for evaluation purposes. It is essential to clearly define the criteria used in the choice of strains to be included in trials as well as their origin. The choice of strains is usually based on information about geographical location, agro-ecological regions, suitability for aquaculture and production performance collected either through literature surveys or personal contacts.

Several methods have been used to assess the degree of differentiation among discrete populations, including linear and meristic morphometrics, as well as biochemical and molecular (DNA) polymorphisms. Characterization of individual populations can help to determine which ones are to be included in a strain comparison trial. However, experience with electrophoresis analysis in both aquatic and terrestrial animals has shown that despite a high level of homogeneity, typical of many crustaceans, there may still be marked differences in production characteristics between the strains (e.g. Jones et al. 2000). Marker analysis can help to identify genetically distinct populations, but there is a weak correlation between molecular information and phenotypic performance in many species (Reed and Frankham 2001). It is therefore at the discretion of the researchers to utilize these techniques or not, to differentiate populations into different strains.

Note that among aquatic animals the term 'strain' has a different meaning from that in livestock, where it is used to define a sub-division within a breed (Hall 2004, Chapter 2). Note also that the term 'breed' is only rarely applied to aquatic animals, but it is sometimes used as synonymous with 'strain'.

## 4.3 Aim and sampling

Strain comparisons have two main aims: First, to compare specific strains for production under a specific set(s) of environments. For instance, we may wish to compare local strains with exotic or genetically improved strains produced in another production system. In this context strain is considered to be a fixed effect and given that part of the variation among animals for any trait is due to genetic differences, it is essential to include both environmental (systematic) and genetic (information on the relatedness of animals) sources of variation in the model. Here genetic information is being used to allow correct estimation of the error variance.

The second aim would be to estimate how much variation there is among a population of strains, in which case strain would be regarded as a random effect. This would require a different set of design criteria depending on the purpose of estimating between strain variation. If the purpose were to establish a foundation for a breeding program using a synthetic strain, an approach such as that of Hill (1974) could be attempted, although that could be difficult in the absence of between strain information. In this manual we assume that the strains are chosen at will by the experimenter(s) and that the strain effect is treated as fixed.

In any strain comparison it is crucial to select a random representative sample from each strain. However, individuals within a strain are not unrelated and generally there is a family structure within each strain (e.g. half sibs, full sibs). This means that sampling needs to be applied at two levels; sampling of families within each strain, and sampling of individuals within families. Given that at least some of the traits of interest are heritable, it is likely that there will be genetic differences between families within each strain. If the family structure is ignored in the analysis, the error variance may be seriously underestimated, especially when heritability is high and the number of families is small. This has been a problem in livestock breed comparisons, where the frequency of 'false' significant differences between two breeds has been estimated to be as high as 50% (Sellier 1980; Komender and Hoeschele 1989).

In order to clarify this issue further, following James (1975), assume a strain comparison trial with *s* strains, *f* families per strain and *n* individuals per family. Analysis of variance is the statistical procedure which is generally used to estimate differences between strains after accounting for all identifiable fixed and random effects in a linear model. The analysis of variance for this scenario is presented in Table 2, where $\sigma^2$, $\sigma^2_f$, and $\sigma^2_s$ are the variances for individuals, families and strains, respectively. In this table it is assumed that strain differences are random, but if they are fixed, $\sigma^2_s$ can be interpreted as a measure of the magnitude of strain differences.

**Table 2.** Analysis of variance including family structure.

| Source of variation | Degrees of freedom | Mean Square (MS) | Expected Value of MS |
|---|---|---|---|
| Strains | s-1 | MSS | $\sigma^2 + n\,\sigma^2_f + nf\sigma^2_s$ |
| Families within strain | s(f-1) | MSF | $\sigma^2 + n\,\sigma^2_f$ |
| Individuals within strain and family | sf(n-1) | MSI | $\sigma^2$ |

It is often impossible to mark larvae or fry soon after hatching due to their small size, making separate rearing of families mandatory until individuals are large enough to be physically tagged. This means that generally there will be an extra component of variance due to common environmental effects ($\sigma^2_c$) which will increase the between family variance. If data are then analyzed without consideration of family structure the analysis of variance will be as shown in Table 3.

**Table 3.** Analysis of variance ignoring family structure.

| Source of variation | Degrees of freedom | Mean Square (MS) | Expected Value of MS |
|---|---|---|---|
| Between strains | s-1 | MSB | $\sigma^2 + n\,\sigma^2_f + nf\sigma^2_s$ |
| Within strains | s(fn-1) | MSW | $\sigma^2 + n\sigma^2_f\,(f-1)/(nf-1)$ |

It is clear that the F test would be misleading and standard errors of strain means would be underestimated. The bias will be worse when the between family variance and the family size are large. For instance, if the between family variance is one quarter of the within family variance, and the experiment has 20 families of 50 individuals per family, the apparent error variance will be less than a tenth of the true value. If due to lack of facilities families cannot be identified, then it is essential to have replication. The analysis of variance for a replicated trial is shown in Table 4. In this scenario we will be able to account for variance due to common environmental effects ($\sigma^2_c$) by taking the between replicate variance as the error term. The variation between fish within replicates is irrelevant to the strain comparison. For the use of replicates to be valid, the replicates must be truly independent samples. It would not be correct to choose a small number of breeding animals to produce fish for the trial, and then split these progeny into replicates. We may note that if individuals are identified and family sizes are equal the family means may be regarded as replicates, which shows that the correct error degrees of freedom in Table 3 should be s(f-1) so that not only the mean square for error is wrong. If family identification is unavailable, the between family variance cannot be estimated, but it may be possible from earlier information to make a reasonable guess. In this case putting θ as the ratio of the between family to within family variance, an estimate of the correct mean square for testing strains may be calculated as:

$$MSW[1 + n\theta]/[1 + \theta(T - s)/(T - 1)]$$

where T = sfn is the total number of animals in the trial. It may be worth trying several plausible values of θ to see for which of them the strain difference would be significant, with s(f – 1) taken as the error degrees of freedom.

**Table 4.** Analysis of variance with replicates (r) within strains.

| Source of variation | Degrees of freedom | Mean Square (MS) | Expected Value of MS |
|---|---|---|---|
| Between strains | s-1 | MSB | $\sigma^2 + n\,\sigma^2_r + nr\sigma^2_s$ |
| Replicates within strains | s(r-1) | MSR | $\sigma^2 + n\,\sigma^2_r$ |
| Fish within replicates | sr(n-1) | MSW | $\sigma^2$ |

The precision of the estimate of between strain differences will be improved by the inclusion of more families from each strain. The recent advent of DNA technologies, especially microsatellites, allows the posterior assignment of parents and tracking identification of family origin of animals. Using this approach all families can be mixed and communal testing can be carried out from birth onward (e.g. Herbinger et al. 1999; Vandeputte et al. 2002b). The posterior assignment of parents enables communal early rearing (CER) after hatching and eliminates the common environmental effect. Early communal rearing reduces the need for using hapas which are not a favorable growing environment for fish in general, especially in carp species. Consequently, fish under the CER scheme grow faster than under separate rearing of families until tagging and have lower generation interval (Ninh et al. 2011). The advantages of CER demonstrate the effectiveness of molecular parentage assignment as a useful tool in selective breeding programs for aquaculture species if it is affordable.

## 4.4 Synchronization of spawning and accounting for known systematic effects

The accuracy of strain comparison trials may be affected by variation in age of individuals within and between strains due to differences in when they were spawned. One way of improving the accuracy is through synchronization of spawning. An ideal situation would be to spawn all required families within 24 to 48 hours in order to minimize differences due to age. This condition is hardly ever met in practice and it is more likely that the required number of families will be bred in a time frame of 4 to 6 weeks. Whereas protracted spawning periods should be avoided, they are sometimes unavoidable, and it is necessary to adjust for variation due to age differences introduced by different spawning times. Bailey and Loudenslager (1986) considered day of spawning as a fixed effect, whereas others have grouped the number of days (usually 3 to 4) into batches and used batch as a fixed factor in the model to account for age differences (e.g. Eknath et al. 1993). The age effect can also be adjusted for by the inclusion of age as a covariate in the model during the analysis of data.

In spite of spawning synchronization, there will still be other environmental sources of variation influencing performance. It is a common practice to obtain a sample of offspring of different strains from various sources, with or without pedigree information, in order to reduce the cost of strain comparison trials. Under this scenario management and environment carry-over effects due to different sample origin will be confounded with the strain effect. This may severely undermine the accuracy of strain comparison and cannot be overcome by increasing the sample size. In order to avoid previous management and environmental carry-over effects, it is necessary to breed the offspring of the different strains at one single location. Here it is important to take a random sample of parents from each strain. In order to avoid dam age effect Unwin et al. (2003) selected 3 year old dams based on their scale pattern, whereas Bailey and Loudenslager (1986) pre-selected dams by visual inspection before tagging. However, the data can be adjusted for dam age by inclusion of dam age as a fixed effect or covariate when information on age is available. Furthermore, adjustments for other environmental effects, such as management and sex, should be applied by their inclusion as fixed effects in the model.

## 4.5 Test environment

To produce meaningful results, strain comparisons need to be conducted in an environment(s) similar to that where they are eventually expected to be used for culture purposes because genotype by environment interactions (G × E) can occur for economically important traits. G × E are said to occur when the observed differences among strains vary depending on the environment. Two types of interactions may be defined. A significant rank-type interaction will lead to dissimilar ranking of strains in different environments (Figure 1), whereas a scale-type interaction will not affect the ranking of strains, but it will affect the magnitude of differences between them. The results from studies in one particular environment should not be extrapolated to another environment unless there is good evidence of absence of rank-type G × E.

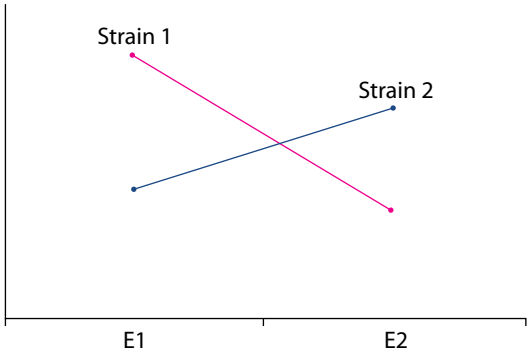


**Figure 1.** Rank type genotype by environment (G × E) interaction.

## 4.6 Number of strains

Although in principle one would like to test as many strains as possible, physical limitations such as hatchery capacity, number of available ponds and hapas for breeding and separate rearing of families, and variation in response to synchronization treatments, impose restrictions on the number of strains that can be included in one trial. Furthermore, in order to detect small differences between strains, a large number of animals from each group is often required, and therefore only a limited number of strains are evaluated in most experiments. As pointed out above, Hill (1974) states that if differences are small the conduct of a strain comparison may not be justified. Taylor (1976a, b) suggests testing several breeds in one experiment, but this design requires some prior information on the breed performance that is seldom available with strains of aquaculture species.

## 4.7 Number of animals per strain

In general, the sample size is dictated by the availability of resources and the biological constraints imposed by the species in question. Ideally, the appropriate sample size needs to be determined prior to the conduct of an experiment to ensure conclusions can be drawn confidently. Power analysis is an important aspect of experimental design for strain comparison. It will help to find the required sample size for detecting a given effect size with a given confidence. The required sample size per strain can be calculated under the assumption of independent samples, using standard statistical theory (Snedecor and Cochran 1971, p 111; Sokal and Rohlf 1969, p 246).

Appendix A describes the classical approach to determine the number of animals for a comparison of any type of 'treatment', in our case, strains. The number of individuals per strain depends upon the magnitude of the difference to be detected, the variance of the trait, desired statistical power and the level of statistical significance applied in the test. The inter-relationship among these parameters can be visualized by varying the individual values. The sample size required to detect a significant difference increases as the difference becomes smaller, and also when the power of the test increases. For a given magnitude of difference and power, an increase in the variation of the trait of interest is accompanied by an increase in the required sample size.

The calculation of sample size based on the approach described in Appendix A assumes that the design is balanced (i.e. the same number of individuals measured in each strain). It has limitations because it ignores between family variability, resulting in a downward bias of the treatment mean standard errors (James

1975). This may cause the apparent detection of a greater number of significant differences than would be the case if the proper model were fitted (Sellier 1980; Komender and Hoeschele 1989). The chance of assessing the strains correctly is therefore reduced. Appendix B presents the equation to determine sample size taking family structure into consideration. Of course, if families are identified, the expected variance between families can be used as the error variance, and the number of families then takes the place of the number of individuals in the calculations. The between family variance will depend on family size, being $\sigma^2_f + \sigma^2/n$ where n is the family size. Alternatively, mixed model procedures to assess power, precision and sample size in the design of experiments may be used but are not presented here. A detailed discussion of the method is provided in Stroup (1999, 2001).

## 4.8 Accuracy

The accuracy of strain evaluations with repeated measurements can be calculated following the method given in SCA (1982). Suppose there are s sires of each strain used to produce n progeny each, whose performance is recorded on *k* occasions. If the performance character recorded has heritability equal to $h^2$ and repeatability equal to *r*, the variance of the mean of the progeny produced by sires of a strain is:

$$V\left(\frac{h^2}{4s}+\frac{r-\frac{1}{4}h^2}{sn}+\frac{1-r}{snk}\right) \qquad [4.1]$$

where *V* is the phenotypic variance. The variance of a mean difference is twice this value.

Let the two strains differ in mean breeding value by *D* standard deviations; their progeny will differ by *½D* standard deviations if the sires of each strain are mated to females of a common strain, so the coefficient of variation of progeny mean difference is:

$$CV=\frac{2\sqrt{2}}{D}\sqrt{\frac{h^2}{4s}+\frac{r-\frac{1}{4}h^2}{sn}+\frac{1-r}{snk}} \qquad [4.2]$$

Now consider a case where 5, 10, or 20 sires from different strains each produced 10, 50, or 100 offspring measured on three occasions. It is also assumed that the heritabilities of the examined traits were at three magnitudes: low (0.05), moderate (0.30) and high (0.70) and the repeatability either equaled or was 0.1 greater than the heritability. Appendix C (Table C1) shows that sire number is the dominant factor affecting accuracy of the trial, as measured by the product of coefficient of variation (CV) and strain difference (D) which is the standard error of the mean difference. There is almost no influence of the repeatability on the CV×D value. The value of CV×D increases with the increase in the magnitude of heritability. Furthermore, over 50, the number of offspring per sire has little impact on the accuracy. The situation discussed here is one that is common in (say) sheep or dairy strain trials, where individuals produce for several years, but may be of limited interest in aquaculture. It may perhaps be relevant to repeated reproduction rates. If sires are mated to females of their own strain, their progeny will differ by D rather than ½D as given above so the standard error of D will be half as large.

If a strain comparison is to be replicated over years it will obviously be better to use a different set of parents in the second year, as this will double the size of the parental sample as well as the size of the progeny sample, whereas if the same parents were used to produce a second crop of offspring, only the progeny group size would be doubled. For example, if we have *f* full-sib families of size *n* and a heritability of $h^2$, the variance of the mean difference would be halved by repeating the experiment with new families in a second year. However, if the same *f* pairs of parents were used in the second year, the variance would be reduced to a fraction which may be much greater than 0.5. If the heritability is 0.4 and family size is 50, this ratio is 0.96, and is hardly smaller than the variance based on a single year's data. The higher the heritability and the larger the family size, the less is gained by using the same parents a second time. This is obvious since, if there is little genetic variance, increasing the size of families will considerably increase the precision of the family mean. If family size is very large, the families used will be accurately evaluated in the first year and little is to be gained by getting more progeny from them in a second year.

In practice, the number of offspring per family is often high but the numbers of families per strain are low in most trials. More offspring per sire or dam can help to increase the precision of the estimated breeding values of those sires or dams, whereas increasing the number of families will increase the accuracy of estimation of strain performance. Several studies reported a correlation between apparent bias in estimation of the genetic trends and the number of sires available for evaluation (e.g. Biffani et al. 2001). It is therefore desirable to optimize the number of families and the offspring per family at the beginning of the experiment. For instance, in European sea bass, Vandeputte et al. (2001) suggested that the greatest precision is obtained if 20 to 50 sires are bred in a full factorial mating design, generating 10 to 25 progeny per sire. However, in most cases, extra family numbers or individuals per family are needed due to unexpected mortalities or losses during the course of experiments. There is no general rule as such to decide how many more families or individuals should be selected, but as a rule of thumb, 10% more than the required sample size should be maintained. Other prevailing environmental conditions such as disease prevalence, severe fluctuations in climate, testing at farmers' ponds, should also be taken into account and if challenge tests are involved for evaluating disease resistance among strains, then the sample size may need to be increased.

## 4.9 Traits measured

It is desirable to measure all economically important traits in strain comparison trials, but in practice only a limited number of traits are measured. This is due to cost and lack of efficient and simple methods for the measurement of some of the traits. In general, aquaculture production aims to achieve greater growth and survival rate, higher fillet yield and better fillet quality, greater disease resistance and better adaptation to the environment. Traits, therefore, can be grouped in the following categories: production performance, carcass composition, meat (flesh) quality, and fitness traits (Table 5). Economic importance, genetic correlations between traits, cost of measurement, precision and feasibility for commercial settings are the criteria usually considered for the inclusion of traits in strain comparison trials. Growth rate measured as harvest weight has been the main selection objective for the majority of selective breeding programs and it has been improved by 10 to 20% per generation.

**Table 5.** Traits of economic importance in aquaculture species.

| Traits | Measurement |
|---|---|
| **Production performance** | |
| Body weight | Weight at harvest |
| Body length | Length at harvest |
| Body width and depth | Width and depth at harvest |
| Feed conversion efficiency | Feed intake/weight gain |
| **Carcass composition** | |
| Carcass weight | Live weight minus total viscera, head and skin weight |
| Dressing percentage (%) | (Carcass weight/live weight)*100 |
| Fillet yield (%) | (Fillet weight/live weight)*100 |
| Head weight percentage (%) | (Head weight/live weight)*100 |
| Abdominal fat weight | Dissected internal fat weight |
| **Flesh quality** | |
| pH | pH meter |
| Color | Color coding system (objective or subjective) |
| Tenderness | Panel assessment; shear force |
| Taste | Panel assessment |
| Texture | Texture measured by Near Infrared Spectrophotometry |
| Fillet fat content | Chemical analysis (Soxhlet Extractor) |
| Fatty acid composition | Specialized laboratory analyzes |
| **Fitness** | |
| Age at maturity | Age at first spawning |
| Number of eggs | Total number of eggs per spawning per female |
| Egg weight | Total egg weight per spawning per female |
| Egg diameter | Egg diameter measured by image analysis techniques |
| Survival rate (%) | Survival at harvest |
| Disease resistance | Challenge test |
| Immunological parameters | e.g. cortisol, lysozyme, total antibody activity (IgM) |
| Salinity tolerance | Median lethal salinity index |
| Thermal tolerance | Challenge tests |

### 4.10 Cost-benefit considerations

Strain comparison trials are often very expensive. Hill (1974) developed an economic approach to derive the optimum size for the trials which depends upon the return for each unit of improvement, variable or fixed cost per animal and between strain variation. The cost-benefit analysis which has been discussed in detail in SCA (1980) in the context of livestock species is presented in Appendix D.

## 5. Design

A range of designs have been used in strain comparison trials (Table 1). The choice of a suitable design generally depends on three main factors:

i) Mating methods (single pair or mass spawning)
ii) Identification at strain, family or individual level
iii) Resources (e.g. facilities, labor and other costs). In practice, there are two main experimental possibilities:
   - Strain only identified, family or individual not identified (the case of mass spawning with only strain identification is illustrated in Appendix E).
   - Pair mating with strain, family and individual identification. With this option, a sib or a factorial mating design can be employed. A step-wise example is given in Appendix F.

In the following sections we briefly describe some designs that require the identification of both family and individuals with the strains being compared.

### 5.1 Sib design

The sib design (full and half sibs) is the simplest and most common form of experimental design used in aquaculture (Figure 2). In a strain comparison there would be males, females and progeny from each of the strains represented in the trial, with a family structure along the lines of that described in Figure 2. In this example each male from each strain is mated to three females, and each female produces progeny of the $y_{ij}$ combination of males and females. If only one female were mated to each male we would only produce full sib families, in what is commonly known in aquaculture as a pair mating design.

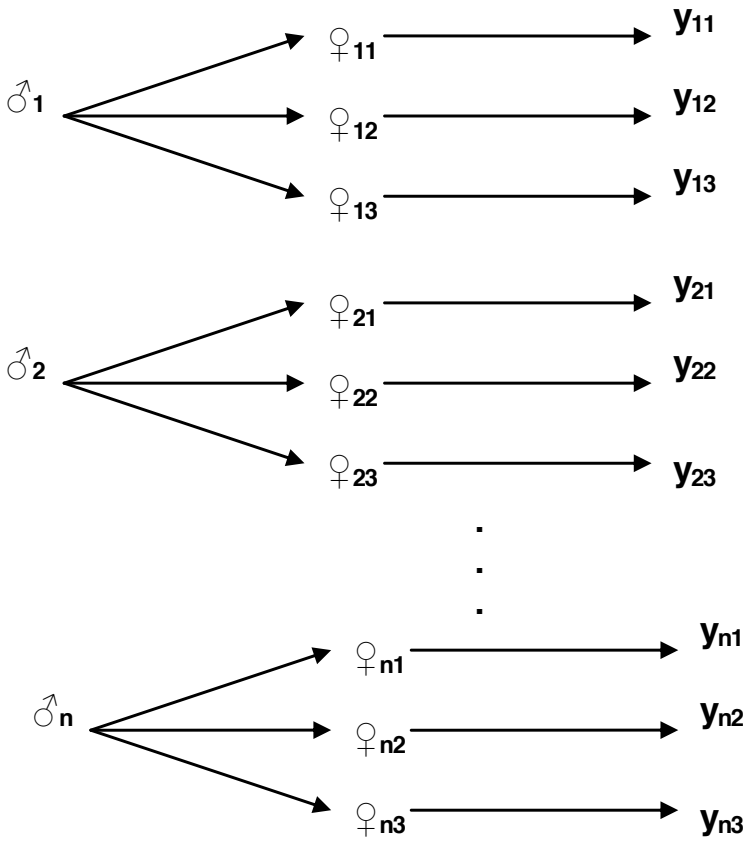


**Figure 2.** Half-sib family design.

## 5.2 Diallel design

A diallel cross is an experimental breeding design used to test all possible combinations of distinct strains or lines. Diallel crosses are used for estimation of genetic effects and evaluating quantitative traits of economic or biological importance. In some instances a diallel cross has been the starting point in the establishment of a base population for a genetic improvement program (e.g. Bentsen et al. 1998; Maluwa and Gjerde 2006; Thanh et al. 2010).

Figure 3 represents a diallel cross of four strains. It is assumed that 10 pairs would be mated within each cross. It is easy to visualize that the diallel design is very informative. In the leading diagonal we have the purebreds, whereas in the off-diagonal we have all possible crosses among the strains involved. The presence of reciprocal crosses enables the estimation of maternal effects.

| Sex of parent stock | Females | | | |
|---|---|---|---|---|
| Males | A | B | C | D |
| A | A × A<br>10 families | A × B<br>10 families | A × C<br>10 families | A × D<br>10 families |
| B | B × A<br>10 families | B × B<br>10 families | B × C<br>10 families | B × D<br>10 families |
| C | C × A<br>10 families | C × B<br>10 families | C × C<br>10 families | C × D<br>10 families |
| D | D × A<br>10 families | D × B<br>10 families | D × C<br>10 families | D × D<br>10 families |

**Figure 3.** Mating scheme (four sources of the species involved: A, B, C and D).

Despite the fact that a diallel design enables the conduct of analyzes with the potential to elucidate genetic influences over important traits, they have not been widely used due to the fact that they require extensive breeding facilities and to the complexities in the computations and genetic interpretation. Furthermore, with aquatic animals, synchronization of spawning is not always feasible, especially when the number of animals increases, and this can result in an excessively extended mating season. We do not deal with the detail of analyzes of diallel designs. Anyone embarking upon the conduct of a diallel cross among strains is advised to consult a statistician with knowledge of the subject matter.

## 6. Analysis

A well run strain comparison involves much time, effort and cost, so it is important that the results are analyzed properly. It is wrong to waste the resources expended in conducting the trial by not extracting the information present in the gathered data. Hence, as much as possible, we should endeavor to use appropriate statistical methods. Often it is well worth consulting a skilled statistician with experience in the field. However, it must be remembered that no matter how clever the analysis may be, it cannot extract information which is not present in the data. Only if the experiment is well designed will the maximum amount of information be derived from the effort.

A range of statistical software is available for the analysis of trial data. Here we present the analysis of a Nile tilapia data set, using two commercial software packages (SAS and SPSS). A simple model is fitted to the data set to show the general process of the analysis and familiarize readers with the application of these packages and interpretation of the results. The analyzes are for illustration, and should not be taken as general guides. For details about SAS and SPSS their respective documentation should be consulted. Every trial should receive careful consideration of its particular characteristics in order to determine a suitable statistical analysis. It is also important that the results of the analysis be presented clearly and convincingly so that workers in the field can understand what has been done and accept the conclusions given.

The data used in this section are a sub-set of a much larger data set collected in the context of the GIFT (Genetic Improvement of Farmed Tilapia) project (see Eknath et al. 1993 for full details). The fish from which the data are used in this section belong to four strains of tilapia. Four different sires were used in each strain (a total of 16 sires). Other information about the fish includes sex (female or male), environment in which the fish were reared (three different locations), batch in which they were produced (five batches), and age at harvest (calculated from the difference between spawning and harvest dates, and used as a linear covariate). The characters recorded at harvest were live weight (g) and length (cm), but only the former trait was analyzed here. The data set used in the following analyses can be obtained by contacting WorldFish [worldfishcenter@cgiar.org].

### 6.1 SAS

#### 6.1.1 Fixed effects model

The analysis can be carried out using the PROC MIXED procedure in SAS (SAS 2008). This analysis ignores the family structure, which is included in a later analysis, in order to show the consequences. The SAS code for the analysis is presented below. We assume that the data are in a text file called 'datastr.prn' in the 'c' drive. The INPUT statement reads the data. The CLASS statement in PROC MIXED includes the effects of strain (str), sire, batch, environment (env) and sex. The MODEL statement specifies the dependent variable harvest body weight (wt) and the independent variables defined in the class statement that we include in the model, as well as age as a linear covariate. The LSMEANS statement produces least squares means for the fitted fixed effects.

```
TITLE 'ANALYSIS OF FISH DATA';
filename fish 'c:\datastr.prn'
   ;
data raul
   ;
infile fish lrecl=140 missover
   ;
*
*   THIS INPUT READS THE PEDIGREE AND THE FISH RECORDS
*;
   INPUT
       ind 1-10 sire 13-21  sex $ 36 age 42-43 env $ 51-55
       str $ 61-62  batch 67 wt 75-78 lgth 83-86
                          ;
 *
 *   THESE STATEMENTS FIT A FIXED EFFECTS MODEL AND age AS COVARIATE
 *;
PROC MIXED
        ;
   CLASS  str  sire batch env sex
        ;
   MODEL  wt = str batch env sex age
        ;
   LSMEANS str batch env sex
        ;
   run
        ;
```

Below we show an edited version of the output corresponding to the SAS code above. The first section of the PROC MIXED output provides information on the number of levels for each class variable and total number of observations used in the analysis. This is followed by the analysis of variance table showing the significance of the fixed effects and of the covariate fitted. Note that in this analysis strain is statistically significant. The next section of the output provides least squares means for the different levels of the fixed effects, with their standard errors.

The Mixed Procedure

Class Level Information

| Class | Levels | Values |
|---|---|---|
| str | 4 | A1 A2 P2 P4 |
| sire | 16 | 198921649 198921681 198921696 198921910 198922210 198922323 198922944 198922963 198923510 198923513 198923529 198923896 198924263 198924455 198924625 198924665 |
| batch | 5 | 1 2 3 4 5 |
| env | 3 | BFAR1 BFAR2 FAC2 |
| sex | 2 | F M |

Number of Observations Used 2081

Residual variance 368.12

Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| str | 3 | 2069 | 9.85 | <.0001 |
| batch | 4 | 2069 | 8.75 | <.0001 |
| env | 2 | 2069 | 408.19 | <.0001 |
| sex | 1 | 2069 | 882.46 | <.0001 |
| age_st | 1 | 2069 | 2.25 | 0.1334 |

Least Squares Means

| Effect | str | env | sex | batch | Estimate | Standard Error |
|---|---|---|---|---|---|---|
| str | A1 | | | | 61.3434 | 1.2023 |
| str | A2 | | | | 55.9216 | 1.2569 |
| str | P2 | | | | 56.9056 | 1.0882 |
| str | P4 | | | | 60.5418 | 1.2250 |
| batch | | | | 1 | 54.1753 | 3.9149 |
| batch | | | | 2 | 61.9511 | 0.8873 |
| batch | | | | 3 | 63.2224 | 1.8226 |
| batch | | | | 4 | 55.8275 | 0.8140 |
| batch | | | | 5 | 58.2141 | 0.7087 |
| env | | BFAR1 | | | 72.7054 | 1.1845 |
| env | | BFAR2 | | | 59.9976 | 1.1746 |
| env | | FAC2 | | | 43.3313 | 1.0037 |
| sex | | | F | | 45.8645 | 1.0567 |
| sex | | | M | | 71.4917 | 0.9791 |

**6.1.2 Mixed model analysis with sire as a random effect**

The analysis presented above is appropriate when there is no relationship among the animals in each strain. In the data we are analysing the fish were of four families per strain, so that the analysis given above is incorrect. We now show how the data should be treated.

The SAS code and output from the model with sire as a random effect are presented below. The SAS code is the same as in the case of fitting only fixed effects, except that now, sire nested within strain is fitted as a random effect.

```
TITLE 'ANALYSIS OF FISH DATA';
filename fish 'c:\datastr.prn'
   ;
data raul
   ;
infile fish lrecl=140 missover
   ;
*
*   THIS INPUT READS THE PEDIGREE AND THE FISH RECORDS
*;
   INPUT
       ind 1-10 sire 13-21  sex $ 36 age 42-43 env $ 51-55
       str $ 61-62  batch 67 wt 75-78 lgth 83-86
                          ;
 *
 *   THESE STATEMENTS FIT A MIXED MODEL WITH sire AS RANDOM AND age AS COVARIATE
 *;

PROC MIXED
        ;
   CLASS  str  sire batch env sex
        ;
   MODEL  wt = str batch env sex age
        ;
   RANDOM sire(str)
        ;
   LSMEANS str batch env sex
        ;
   run
        ;
```

As in the case of the fixed effects model, we now present an edited version of the SAS output. The first part of the output is identical to that when we fitted the fixed effects model. It provides information on the number of levels for each class variable and total number of observations used in the analysis. This is followed by the analysis of variance table showing the significance of the fixed effects and of the covariate fitted, and illustrating important differences relative to the fixed effects model in the previous section. Note that in this analysis strain is not statistically significant (P~0.28), in contrast to what happened in the fixed effects model. Note also that the denominator degrees of freedom for strain are smaller than in the fixed effects model (12 vs 2069). This is due to the fact that in this analysis, rightly accounting for the family structure in the data, the strain mean square is tested against the mean square for sires within strains, not against the residual. The conclusion now is that there are no significant differences between strains. The results show that if family structure (e.g. sires within strains) is not included in the model we may end up with a false conclusion that there are significant differences between strains, whereas a correct analysis shows that there are no significant differences between them.

The Mixed Procedure

Class Level Information

| Class | Levels | Values |
|---|---|---|
| str | 4 | A1 A2 P2 P4 |
| sire | 16 | 198921649 198921681 198921696 198921910 198922210 198922323 198922944 198922963 198923510 198923513 198923529 198923896 198924263 198924455 198924625 198924665 |
| batch | 5 | 1 2 3 4 5 |
| env | 3 | BFAR1 BFAR2 FAC2 |
| sex | 2 | F M |

Number of Observations Used 2081

Variance Estimates

| Parameter | Estimate |
|---|---|
| sire(str) | 16.9453 |
| Residual | 357.80 |

Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| str | 3 | 12 | 1.44 | 0.2796 |
| batch | 4 | 2057 | 4.40 | 0.0015 |
| env | 2 | 2057 | 419.83 | <.0001 |
| sex | 1 | 2057 | 859.20 | <.0001 |
| age_st | 1 | 2057 | 0.81 | 0.3669 |

Least Squares Means

| Effect | str | env | sex | batch | Estimate | Standard Error |
|---|---|---|---|---|---|---|
| str | A1 | | | | 60.9492 | 2.3956 |
| str | A2 | | | | 55.5934 | 2.4271 |
| str | P2 | | | | 56.3143 | 2.3332 |
| str | P4 | | | | 60.0620 | 2.4095 |
| batch | | | | 1 | 52.6904 | 4.3060 |
| batch | | | | 2 | 61.4140 | 1.3939 |
| batch | | | | 3 | 61.8474 | 2.1705 |
| batch | | | | 4 | 56.8866 | 1.3243 |
| batch | | | | 5 | 58.3102 | 1.2561 |
| env | | BFAR1 | | | 72.4976 | 1.5839 |
| env | | BFAR2 | | | 59.6906 | 1.5765 |
| env | | FAC2 | | | 42.5010 | 1.4567 |
| sex | | | F | | 45.6654 | 1.4873 |
| sex | | | M | | 70.7941 | 1.4422 |

## 6.2 SPSS

In this section we show that the data can be analyzed with the statistical package SPSS, and the results are the same as those obtained using SAS.

### 6.2.1. Fixed effects model

We reanalyzed the data processed with SAS in section 6.1.1 using the MIXED procedure in SPSS (SPSS 2011). The MIXED procedure enables the use of linear mixed models for the analysis of the data. However, as mentioned in section 6.1.1, when only fixed effects are fitted, the analysis ignores the family structure. We assume that the data are in an EXCEL file called 'datastr.xls' in the 'c' drive, given that EXCEL files are easier to read in SPSS than text files. To open your data file in SPSS, from the menu bar choose: File ----> Open ----> Data, then go to the 'c' drive and choose your file then click on open. To run the analysis we also require a syntax file. This is a file in which you write the code for analysis in SPSS. To open a new syntax file, from the menu bar choose: File ----> New ----> Syntax. The SPSS code for analysis is:

```
MIXED wt BY str sire batch env sex WITH age
/FIXED= str batch env sex age | SSTYPE(3)
/EMMEANS=TABLES(str)
/EMMEANS=TABLES(batch)
/EMMEANS=TABLES(env)
/EMMEANS=TABLES(sex) .
```

The code for the analysis starts with the keyword MIXED, followed by the dependent variable harvest body weight (wt). Next, the keyword BY separates the dependent variable (wt) from the independent variables strain (str), sire, batch, environment (env) and sex. Any covariates in the model are listed at the end of the statement after the keyword WITH. The covariate fitted in this case is age. The FIXED statement includes the fixed effects that are in the model of the analysis, strain (str), batch, environment (env) and sex. The keyword | SSTYPE(3) specifies the calculation of type III sum of squares. EMMEANS statements produce the least squares means (called Estimated Marginal Means in SPSS) for the fitted fixed effects.

To run the analysis and get the results, from the window of the syntax file, choose from the menu bar Run -- --> All.

Below we show an edited version of the output corresponding to the SPSS code above. The first section of the output provides information on the number of levels for each class variable and the total number of observations used in the analysis. This is followed by the analysis of variance table showing the statistical significance of the fixed effects and of the covariate fitted. Note that the result is identical to the result using SAS in section 6.1.1.

## Mixed Model Analysis

**Dependent Variable: wt. Model Dimension**

| | | Number of Levels | Number of Parameters |
|---|---|---|---|
| Fixed Effects | Intercept | 1 | 1 |
| | Str | 4 | 3 |
| | Batch | 5 | 4 |
| | Env | 3 | 2 |
| | Sex | 2 | 1 |
| | Age | 1 | 1 |
| Residual | | | 1 |
| Total | | 16 | 13 |

## Fixed Effects

**Type III Tests of Fixed Effects**

| Source | Numerator df | Denominator df | F | Significance P>F |
|---|---|---|---|---|
| Intercept | 1 | 2069 | 2088.827 | .000 |
| Str | 3 | 2069 | 9.845 | .000 |
| Batch | 4 | 2069 | 8.754 | .000 |
| Env | 2 | 2069 | 408.194 | .000 |
| Sex | 1 | 2069 | 882.456 | .000 |
| Age | 1 | 2069 | 2.254 | .133 |

## Estimated Marginal Means

1. str

| str | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| A1 | 61.343 | 1.202 | 2069 | 58.986 | 63.701 |
| A2 | 55.922 | 1.257 | 2069 | 53.457 | 58.387 |
| P2 | 56.906 | 1.088 | 2069 | 54.772 | 59.040 |
| P4 | 60.542 | 1.225 | 2069 | 58.139 | 62.944 |

2. batch

| batch | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| 1 | 54.175 | 3.915 | 2069 | 46.498 | 61.853 |
| 2 | 61.951 | .887 | 2069 | 60.211 | 63.691 |
| 3 | 63.222 | 1.823 | 2069 | 59.648 | 66.797 |
| 4 | 55.828 | .814 | 2069 | 54.231 | 57.424 |
| 5 | 58.214 | .709 | 2069 | 56.824 | 59.604 |

3. env

| env | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| BFAR1 | 72.705 | 1.184 | 2069 | 70.382 | 75.028 |
| BFAR2 | 59.998 | 1.175 | 2069 | 57.694 | 62.301 |
| FAC2 | 43.331 | 1.004 | 2069 | 41.363 | 45.300 |

4. sex

| sex | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| F | 45.865 | 1.057 | 2069 | 43.792 | 47.937 |
| M | 71.492 | .979 | 2069 | 69.571 | 73.412 |

### 6.2.2 Mixed model analysis with sire as a random effect

In this section we repeat the analysis of section 6.1.2 where sire is included in the model as a random effect. The SPSS code is the same as in the case of fitting only fixed effects (section 6.2.1), except that now sire nested within strain is fitted as a random effect using statement RANDOM.

```
MIXED wt BY str sire batch env sex WITH age
/FIXED= str batch env sex age | SSTYPE(3)
/RANDOM=sire(str)
/EMMEANS=TABLES(str)
/EMMEANS=TABLES(batch)
/EMMEANS=TABLES(env)
/EMMEANS=TABLES(sex) .
```

Below we show an edited version of the SPSS output. The first part provides information on the number of levels for each class variable and total number of observations used in the analysis. This is followed by the analysis of variance table showing the significance of the fixed effects and of the covariate. The result is identical to the result of SAS in section 6.1.2, except for the denominator degrees of freedom for the fixed effects. The reason is that SPSS MIXED uses an approximation called Satterthwaite for the calculation of the degrees of freedom, whereas SAS, by default, uses a method called CONTAINMENT. However, the Satterthwaite method is an option in PROC MIXED of SAS.

## Mixed Model Analysis

Dependent Variable: wt. Model Dimension

| | | Number of Levels | Covariance Structure | Number of Parameters |
|---|---|---|---|---|
| Fixed Effects | Intercept | 1 | Variance Components | 1 |
| | Str | 4 | | 3 |
| | Batch | 5 | | 4 |
| | Env | 3 | | 2 |
| | Sex | 2 | | 1 |
| | Age | 1 | | 1 |
| Random Effects | sire(str) | 16 | | 1 |
| Residual | | | | 1 |
| Total | | 32 | | 14 |

## Fixed Effects

Type III Tests of Fixed Effects

| Source | Numerator df | Denominator df | F | Significance P>F |
|---|---|---|---|---|
| Intercept | 1 | 53.773 | 1014.304 | .000 |
| Str | 3 | 10.979 | 1.436 | .285 |
| Batch | 4 | 1904.273 | 4.397 | .002 |
| Env | 2 | 2028.196 | 419.836 | .000 |
| Sex | 1 | 2065.936 | 859.194 | .000 |
| Age | 1 | 265.820 | .818 | .366 |

## Covariance Parameters

Estimates of Covariance Parameters

| Parameter | | Estimate | Std. Error |
|---|---|---|---|
| Residual | | 357.792555 | 11.158762 |
| sire(str) | Variance | 17.011302 | 8.563131 |

## Estimated Marginal Means

1. str

| str | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| A1 | 60.949 | 2.399 | 14.337 | 55.815 | 66.083 |
| A2 | 55.594 | 2.430 | 15.059 | 50.415 | 60.772 |
| P2 | 56.314 | 2.337 | 12.937 | 51.263 | 61.364 |
| P4 | 60.062 | 2.413 | 14.662 | 54.908 | 65.215 |

2. batch

| batch | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| 1 | 52.690 | 4.307 | 797.133 | 44.236 | 61.143 |
| 2 | 61.414 | 1.395 | 25.472 | 58.542 | 64.285 |
| 3 | 61.847 | 2.172 | 131.715 | 57.551 | 66.142 |
| 4 | 56.887 | 1.326 | 21.189 | 54.132 | 59.643 |
| 5 | 58.310 | 1.258 | 17.316 | 55.660 | 60.960 |

3. env

| env | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| BFAR1 | 72.498 | 1.585 | 42.623 | 69.300 | 75.695 |
| BFAR2 | 59.691 | 1.578 | 41.820 | 56.506 | 62.875 |
| FAC2 | 42.500 | 1.458 | 30.580 | 39.525 | 45.476 |

4. sex

| sex | Mean | Std. Error | Df | 95% Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower Bound | Upper Bound |
| F | 45.665 | 1.489 | 33.469 | 42.638 | 48.693 |
| M | 70.794 | 1.444 | 29.419 | 67.843 | 73.744 |

## 6.3 Assumptions of analysis of variance

Sokal and Rohlf (1969, Chapter 13) provide a detailed account of underlying assumptions in analysis of variance. These include independence of the error terms, homogeneity of variances (homoscedasticity), additivity of main effects, and normality. There is generally little awareness about these assumptions among aquaculturists planning or analyzing strain comparison data, except perhaps regarding the last mentioned one. Note, however, that the consequences of non-normality are not too serious. Only a very skewed distribution will have a marked effect on the significance level of effects or on the efficiency of the experimental design. If lack of normality is a concern, a good way of addressing the issue is by carrying out a transformation (e.g. logarithmic, square root, arcsine). Figure 4 shows the distribution of fish weights for the example we have been analyzing, before and after transformation to square root. It is clear that the distribution of square root transformed data looks more like 'normal' than it does in actual units. Note however, that when the transformed data were analyzed in the same way that we earlier analyzed the raw data, essentially the same results were obtained.

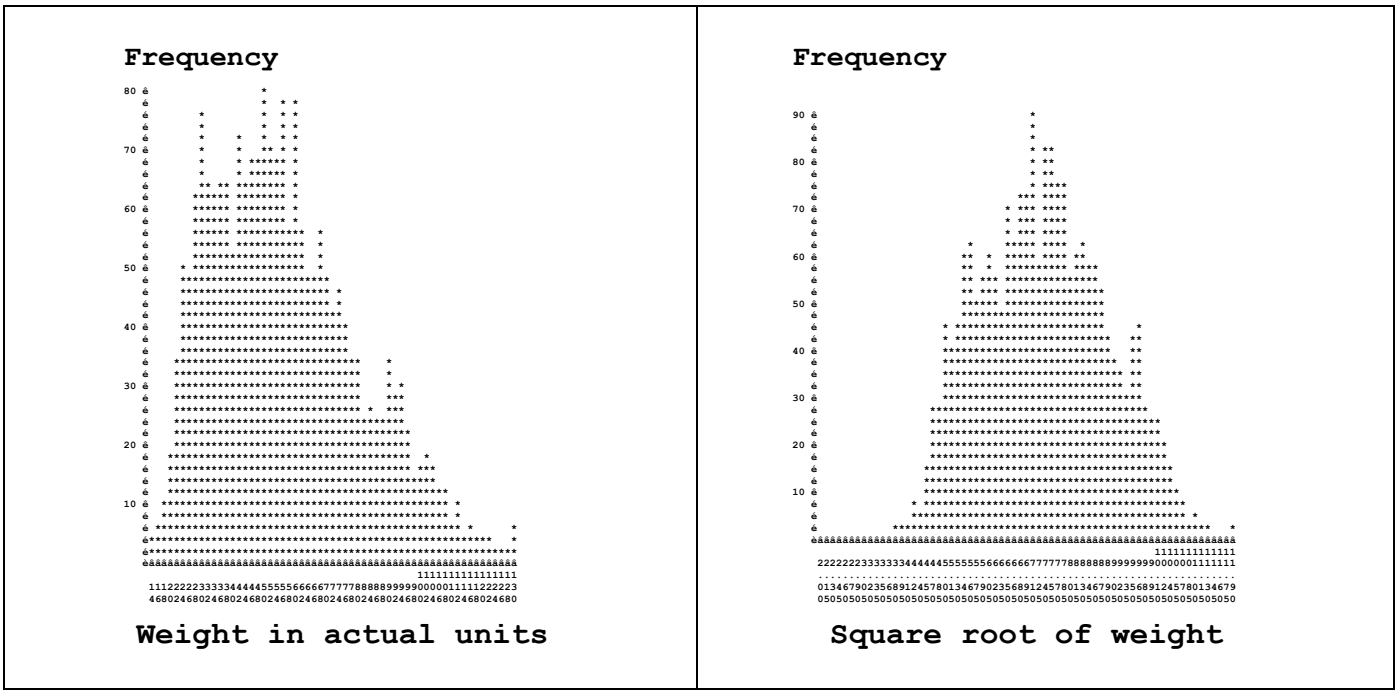


**Figure 4.** Distribution of fish weights, before and after transformation to square root.

Sokal and Rohlf (1969, p. 381) comment that a fortunate fact about transformations is that often several departures from the assumptions of the analysis of variance are simultaneously cured by the same transformation to a new scale. For instance, if the data are made homoscedastic they are usually also made to approach normality and to approximate additivity of treatment effects. A visual inspection of data distribution such as that displayed in Figure 4 may point to the need for a transformation. Consulting with a statistician is recommended.

If there is evidence that the assumptions of the analysis of variance are violated, it is wise to transform the variable before analysis. Power transformation (e.g. logarithm, square root) and angular or arcsine are among the options. The logarithmic transformation is commonly used when the standard deviations of samples are proportional to the means. The square root is appropriate when dealing with counts of rare events and data tend to follow a Poisson distribution. The arcsine or angular transformation is required for data expressed as percentages or proportions (e.g. survival rate in fish). The power transformation may be applied when strain comparison trials are conducted over a broad range of locations or environments and the means and residual variances differ markedly among them. Before analysis of transformed data one should test that the transformation has produced an improvement in the approximation to normality and equality of variances. Note that the tests of significance are performed on the transformed data, but the means and standard errors need to be back-transformed to the original units in reports for ease of interpretation by readers.

## 7. CASE STUDIES

In this section we give a brief account of a sample of strain comparison studies. The purpose is to illustrate how strain comparisons are designed and conducted by researchers with different aims. Note that we provide only a sketchy description of the trials. Full details are available in the original papers. The hope is that readers interested in conducting strain comparisons will consult these references and gain further insight on how to go about their work.

### 7.1 Nile tilapia (*Oreochromis niloticus*) in Egypt

Ibrahim et al. (2013) report the relative performance of two Nile tilapia strains in Egypt: the Abbassa selection line and the Kafr El Sheikh commercial strain. The Abbassa selection line (developed by selective breeding) and the Kafr El Sheikh commercial strain (widely used in the tilapia industry in Egypt) were evaluated at two stocking densities (two and four fish per $m^2$). Harvest weight, length, depth, width and head length were recorded.

The experiment was repeated with the same design in 2008 and 2010. In both years there were 10 ponds of 100 $m^2$ each available. Five were randomly assigned to a lower stocking density (two fish per $m^2$), whereas the other five were assigned to a higher stocking density (four fish per $m^2$). Both strains were present in each pond. Figure 5 is a schematic representation of the experimental design used in this trial, where pond is nested within year and stocking density, but pond and strain are cross-classified. The statistical model fitted to harvest weight included the fixed effects of Year, Strain, Density, Sex, as well as all the two way interactions; Pond nested within Year and Density, and Strain by Pond were fitted as random effects. Mathematically the model may be written as:

$$Y_{ijklmn} = \mu + Yr_i + Str_j + Dens_k + Sx_l + (Yr\ Str)_{ij} + (Yr\ Dens)_{ik} + (Yr\ Sx)_{il} + (Str\ Dens)_{jk} + (Str\ Sx)_{jl} + (Dens\ Sx)_{kl} + P_{ikm} + (Str\ P)_{jm} + e_{ijklmn}$$

where: $Y_{ijklmn}$ is the nth observation in the mth pond, the lth sex, the kth density, the jth strain and the ith year, $\mu$ is the overall mean,
$Yr_i$ is the effect of the ith year (i = 2008 or 2010),
$Str_j$ is the effect of the jth strain (j = Abbassa selection line or Kafr El Sheikh strain),
$Dens_k$ is the effect of the kth stocking density (k = two or four fish per $m^2$),
$Sx_l$ is the effect of the lth sex (l = female or male),
$(Yr\ Str)_{ij}$, $(Yr\ Dens)_{ik}$, $(Yr\ Sx)_{il}$ + $(Str\ Dens)_{jk}$, $(Str\ Sx)_{jl}$ and $(Dens\ Sx)_{kl}$ are self-explanatory two way interaction terms,
$P_{ikm}$ is the random effect of the mth pond nested within the kth density and the ith year (m=1,..., 20), $(Str\ P)_{jm}$ is the random effect of the interaction between strain and pond, and $e_{ijklmn}$ is the random error term.

Males were heavier than females but the between sex difference was greater in the commercial than in the Abbassa selection line (39 and 31 per cent, respectively). Females in the Abbassa selection line grew almost as fast as males in the commercial line. Both

strains grew faster at the lower density of two fish per $m^2$, and the percentage reduction in harvest weight at the higher density was about the same for both strains (27 per cent). The advantage of the Abbassa selection line over the commercial line was about 28 per cent at both densities. Both strains had a similar (and good) survival rate (approx. 80 per cent) during the grow-out period. The authors concluded that the Abbassa selection line was ready for release to the tilapia industry in Egypt.

The aim of the comparison in this case was clear: there was a strain widely used commercially (the Kafr El Sheikh) and a new one (Abbassa line) developed by WorldFish scientists by selective breeding. A comparison was essential in order to make a rational decision about the use of the Abbassa line. The paper by Ibrahim et al. (2012) provides a detailed account and discussion of all the results.

2008

| | | | | |
|---|---|---|---|---|
| Pond no. 1<br>Strains A and K<br>Density 2 | Pond no. 2<br>Strains A and K<br>Density 4 | Pond no. 3<br>Strains A and K<br>Density 2 | Pond no. 4<br>Strains A and K<br>Density 4 | Pond no. 5<br>Strains A and K<br>Density 4 |
| Pond no. 6<br>Strains A and K<br>Density 2 | Pond no. 7<br>Strains A and K<br>Density 4 | Pond no. 8<br>Strains A and K<br>Density 2 | Pond no. 9<br>Strains A and K<br>Density 4 | Pond no. 10<br>Strains A and K<br>Density 2 |

2010

| | | | | |
|---|---|---|---|---|
| Pond no. 11<br>Strains A and K<br>Density 4 | Pond no. 12<br>Strains A and K<br>Density 4 | Pond no. 13<br>Strains A and K<br>Density 2 | Pond no. 14<br>Strains A and K<br>Density 2 | Pond no. 15<br>Strains A and K<br>Density 4 |
| Pond no. 16<br>Strains A and K<br>Density 2 | Pond no. 17<br>Strains A and K<br>Density 4 | Pond no. 18<br>Strains A and K<br>Density 4 | Pond no. 19<br>Strains A and K<br>Density 2 | Pond no. 20<br>Strains A and K<br>Density 2 |

**Figure 5.** Schematic representation of the experimental design of the strain comparison (A=Abbassa selection line; K=Kafr El Shiekh strain; Density 2=stocking density of 2 fish per $m^2$, 100 A and 100 K fish; Density 4=stocking density of 4 fish per $m^2$, 200 A and 200 K fish).

## 7.2 Rohu carp (*Labeo rohita*) in India

Reddy et al. (2002) compared growth and survival of six stocks of rohu carps (one domesticated and five wild stocks), in earthen ponds, under monoculture and polyculture in India. The mating design used to produce full and half sib families was as follows. In the first year (1993) of the experiment, eggs for each female were fertilized with milt from three different males whereas in the second year (1994) each male was mated with two or three different females. After hatching, each full sib group was reared in separate nursery ponds or tanks until fingerlings when they were randomly sampled and individually tagged at an average body weight between 12 and 31 g. The same number of fingerlings from each full sib group was stocked in either monoculture or polyculture ponds. At harvest, individual body weight and survival were recorded. Coding for the latter trait was based on fish present at harvest time, relative to those that were present at tagging time.

Body weight and survival were analyzed using a linear model with the following effects: a fixed effect of production system (monoculture or polyculture), a random effect of pond nested within production system, a fixed effect of stock, a fixed interaction effect between production system and stock, a random effect of full sib family nested within stock, and a random error term. Survival data were coded as '0' or '1' corresponding to dead or alive fish, respectively. The stock effect was tested against the full sib family effect nested within stock as the error term.

The results from the analysis indicated that there is a statistically significant difference in harvest weight (but not survival) between strains in 1994. The effects of full sib group on harvest weight and survival were highly significant, implying that there is substantial genetic variation in both harvest weight and survival within strains. It is suggested that these two traits can be effectively improved by selective breeding. Overall, both harvest weight and survival were not different between monoculture and polyculture systems but the interaction between production systems and strains was significant though of low magnitude. Given this interaction, one could conclude that careful consideration should be given to the decision on which strain is to be used as base stock for breeding programs in the different production systems. However, the authors comment that there was no re-ranking of the stocks for harvest weight, and that the development of specialized strains for each of the two production systems is not required.

This is one of the well designed trials for strain comparisons. The analysis used the linear mixed model, taking into account the random effect of families and ponds within production systems. Therefore, the estimates of strain differences are expected to be unbiased. The analysis of survival could be improved by using statistical procedures that are specific for the treatment of discrete data, although generally one would not expect greatly different results.

The main objective of the study conducted by Reddy et al. (2002) was to compare the performance of one farmed and five wild rohu stocks for growth and survival rate in two production environments, with a view to establishing a base population for a selective breeding program in India.

## 7.3 Striped bass (*Morone saxatilis*) in the USA

The paper by Jacobs et al. (1999) illustrates the inclusion of covariates in mixed model analyzes. The main aim of this study was to evaluate the performance of five different striped bass strains in the US. Families were produced by pair mating from randomly captured wild brood stock from different rivers in different locations. Experimental fish were grown in two facilities, mainly differing in water source, tank size, feeding regime and photoperiod. Families from each strain were randomly stocked in at least three replicate tanks. If fish were not initially tagged, each strain was kept in separate tanks. Otherwise, three to eight families of the same strain were held in one tank. Body weight of either all or a sample of the fish (depending on the grow-out facility) per tank was taken every month or every 6 weeks for a period of five months.

The authors carried out a mixed model analysis whereby facility, strain, and facility by strain interactions were treated as fixed effects, whereas family nested within strain and tank nested within strain were treated as random. As all weights were not

the same at the start of the experiment across facilities, starting weights and starting weights by strains were included as covariates in the model.

Overall, significant differences in growth rate existed among striped bass strains. The effect of facility was substantial and the correlation of family growth between facilities was high (0.82). There was no evidence of genotype by environment interaction. The variance due to family nested within strain accounted for as much as 60% of the variance, but genetic variability in these populations could not be fully described due to the limited number of families in the study. The effect of tank (i.e. common environmental effect) was not given in the paper although it was included in the model for analysis. The common environmental effect may result from the similarity between full sibs, which are often reared in the same tanks or ponds. In a number of aquaculture species, the magnitude of this effect has been estimated as ranging between 2 and 15% (e.g. Rye and Mao 1998; Pante et al. 2002; Ponzoni et al. 2005). It is recommended that whether or not the effect of tank or pond is significant, it should be included in the model to avoid bias due to the presence of non-additive genetic effects. In this way, some parts of the maternal genetic effect, that is, the effect of dam's genotype on performance of her progeny, are also accounted for. It is always likely that environmental effects of tank and pond are present, regardless of non-additive genetic effects, and it is usually unwise to omit them from the model.

The researchers conducting this study comment that perhaps one of the most effective means of starting a selective breeding program is through strain selection. Strain evaluations that have been conducted with a number of species indicate that by examining a wide range of strains under common culture conditions, improvements in body weight (at a fixed age) up to 50% are possible by appropriate choice of strain before selective breeding efforts are initiated. We concur with authors that a correct choice of strain(s) may result in productivity increases equivalent to those achieved by several generations of selection.

## 7.4 Yabby crayfish (*Cherax destructor*) in Australia

Jerry et al. (2002) conducted a trial to evaluate performance characteristics of five wild populations of freshwater yabby crayfish (*Cherax destructor*). The populations were randomly sampled from rivers in Southern-eastern Australia. A nested mating design was employed in which each male was mated with four females in an aquarium. Once berried, females were moved at random to individual aquaria until juveniles were released approximately 28 days later. Juveniles randomly chosen from each dam were then pooled into half and full sib family groups based on sire lines and stocked in the grow-out system. Measurements were made every 3 months and they included individual body weight, orbit carapace length, abdomen length, abdomen width and total length. Analyzes were carried out fitting a general linear model. In preliminary runs the significance of initial stocking weight and density at measurement as covariates was tested. They were, however, dropped from the model due to their insignificant effects. The final model included the fixed effects of sire, population and sex. There were significant differences in body weights between populations; for instance the population coded as number 2 was 40% heavier than the population coded as number 1. However, both abdomen width and length were not statistically different between strains. In the present study, the model included sire as a fixed effect; however, it could be of interest to fit sire as a random term, which probably would have been more appropriate.

The authors reporting this study had the aim of evaluating wild populations and generating information to identify foundation stocks on which to base a future genetic improvement program for this species. The results indicate that there is large genetic variation in growth between the five populations and that the genetic base for a selective breeding program can be formed from population 2. If genetic gain in harvest weight generally reported for aquatic animals is a conservative average of 10% per generation, selection of population 2 can advance four generations ahead of other populations.

## 7.5 Tilapia (*Oreochromis niloticus*) in the Philippines

The aim of strain comparison trials sometimes includes the evaluation of the strains in a range of environments with the purpose of finding out whether there is an important genotype by environment interaction (G × E). Eknath et al. (1993) reported growth performance of eight different strains of Nile tilapia under different farm environments in the Philippines. Four strains were collected from the wild in Africa, as well as another four domesticated strains farmed in Asia at the time. Single pair matings (25 pairs per strain) were carried out in hapas in breeding ponds. Fry were collected in batches (3 to 7 days) and reared separately for each strain and batch until they reached 3 to 5 g, when they were individually tagged. Fingerlings from different strains were communally stocked in different testing environments, ranging from fertilized ponds (with and without supplementary feeding), ponds fertilized with on farm agricultural residues, rice-fish systems, cages, and hatcheries. Harvest weight was recorded after an average rearing period of 90 days. The general linear model shown below was fitted to the data:

$$y_{ijklm} = \mu + E_i + G_j + S_k + B_l + (GE)_{ij} + (SE)_{ik} + (BE)_{il} + e_{ijklm}$$

where:
$y_{ijklm}$ is the phenotypic performance of individual m of batch l and sex k from strain j in the environments i (m=1,2, ..., 3420)
$E_i$ is the fixed effect of the test environments (i=1, 2,...8)
$G_j$ is the fixed effect of strains (j=1,...8)
$S_k$ is the fixed effect of sex k (k=1, 2)
$B_l$ is the fixed effect of batch (l=1,2)
$(GE)_{ij}$ is the fixed interaction effect between strains and test environments
$(SE)_{ik}$ is the fixed interaction effect between sexes and test environments
$(BE)_{il}$ is the fixed interaction effect between batches and test environments and $e_{ijklm}$ is residual random error with mean 0 and variance $\sigma^2$

The main finding from this study was that there were significant differences in final body weight among strains, with the fastest growing strain being that from Egypt and the slowest one from Ghana. Overall, the growth performance of strains was relatively consistent across the testing environments. The interactions between strains and environments, although significant, were low and explained only 0.3% of the total variation in the model.

The authors suggested that, given the relative unimportance of G × E, there was no need to develop specialized tilapia strains for specific environments, and that selection in a single composite population was the appropriate course of action to follow. Note that the work reported in this paper provided the foundation stock for the well known and highly productive GIFT (Genetically Improved Farmed Tilapia) strain (Ponzoni et al. 2011b).

## 8. CONCLUSION

Strain comparison trials involve several steps including aim and choice of strains, sampling populations, preparation of testing environments, determination of sample sizes associated with pre-defined statistical power, implementation of the experiment, collection of data, statistical analysis and interpretation of the results. The accuracy of the strain comparison trial is generally determined by the number of families used. If the assumption of normal distribution is violated, the data should be transformed before carrying out statistical analyzes. The development of DNA and reproductive technologies enable posterior parentage assignment, minimizing early testing environment effects and providing an opportunity for the application of a range of experimental designs. Financial limitations are most often a major constraint, but nevertheless, within the limits imposed by available resources, experiments should be carried out in a manner that maximizes the chance of detecting between strain differences without bias.

# APPENDICES

## A. Sample size using standard method

An important step in the design of strain comparison trails is the calculation of sample size. The least significant difference (*d*) between two treatments with equal size and a common estimate of error variance (e.g. Morris, 1999) is:

$$d = t_{\alpha} \times \sqrt{2} \times \sqrt{[\sigma^2/n]} \quad [A1]$$

where *t* is Student's *t* value for a chosen probability with error degrees of freedom (*df*), $\sigma^2$ is the variance of the trait in question, and n is sample size for each treatment. Sample size can be obtained after rearrangement of [A1] from the following equation:

$$n = 2 \times t_{\alpha}^{\ 2} \times \sigma^2/d^2 \quad [A2]$$

The coefficient of variation of a variable is a measure of its standard deviation expressed as a fraction of the mean. Dividing both numerator and denominator of equation [A2] by the square of the mean and expressing as a percentage gives:

$$n = 2 \times t_{\alpha}^{\ 2} \times (CV)^2/ (d\%)^2 \quad [A3]$$

The $t_{\alpha}$ value can be obtained from a table in a statistical textbook (or elsewhere) for a chosen level of probability, conventionally taken as 0.05 or 0.01. There are several sources to obtain CV (%) such as: 1) results of previous trials, 2) average value from literature, or 3) make a guess. The value of d depends on how big a difference is regarded as important. Further details and explanations are given in Morris (1999).

**Example A1**: Assume we wish to design an experiment to detect a difference of 2% (d=2) between harvest weight of two strains of tilapia at a probability level of 0.05 (P=0.05). A survey of the literature shows that the CV of harvest weight at 6 months age is 25% (CV=25). Given P=0.05, the $t_{\alpha}$ value is approximately 2.0 (Morris, 1999). The required number of animals from equation A3 is 1250.

$$n = 2 \times (2.0)^2 \times (25)^2/ (2)^2 = 1250$$

This is the quickest way to determine the required sample size. However, since the observed value, *d*, is different from the true mean value we need to specify a probability of successfully detecting this difference by including a second *t* value in equation [A3], which then becomes:

$$n = 2 \times (t_{\alpha} + t_{\beta})^2 \times (CV)^2/ (d\%)^2 \quad [A4]$$

where $t_{\beta}$ is Student's *t* value with *df* of the error variance and a probability of 2(1-*p*), with *p* = the probability of success or power of the test.

**Example A2**: With the above example, we now wish to have a 90% chance of success to detect the difference of 2% between the two strains, then *p* = 0.9 and 2(1-*p*) = 0.2. The value $t_{\beta}$ is approximately 1.3, so the number of animals according to equation [A4] will be

$$n = 2 \times (2.0 + 1.3)^2 \times (25)^2/ (2)^2 = 3403$$

Note that equation [A4] becomes [A3] when *p* = 0.5, then 2(1-*p*) = 1 and $t_{\beta} = 0$. This indicates that there is a dramatic increase in the number of replicates with increasing probability of success.

In conclusion, applying the equation [A4] it is therefore possible to determine the number of replicates for any given values of CV and d% and desired power. However, there are also several computer software programs for calculating the number of replications. Some are available on the internet and free to download, such as: http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/ (Erdfelder et al. 1996), and http://www.stat.uiowa.edu/~rlenth/Power/ (Lenth 2001).

In addition, users can write scripts to perform power analyzes with regular general-purpose statistical software. There are three main advantages: independence of software, customization of the program for the analysis, and easy implementation of new advances in statistical theory. For example using some basic information from Ponzoni et al. (2005), the sample size required to have a significant difference in harvesting weight between two groups of Tilapia (means: 166 and 192 g, standard deviations: 80 and 116 g, respectively), with a statistical power of 0.8 and level of significance of 0.05, can be calculated using the following SAS script:

```
Data size;
DO n=2 to 1000;
alpha=0.05;
mi1=166; /* the mean of population 1 */
mi2=192; /* the mean of population 2 */
std1=80; /* the standard deviation of population 1 */
std2=116; /* the standard deviation of population 2 */
df=2*n-2; /* degrees of freedom */
pstd=sqrt(((n-1)*std1+(n-1)*std2*std2)/(n+n-2)); /* pooled SD */
lambda=(abs(mi2-mi1)/pstd)/sqrt(1/n+1/n); /* the noncentral
parameter*/
tcrit_low=TINV(alpha/2,df); /* lower limit of the critical value */
tcrit_up=TINV(1-alpha/2,df); /* upper limit of the critical value */
tcrit_onetail=TINV(1-alpha,df); /* critical value for one-side test */
power_onetail=1-CDF('t',tcrit_onetail, df, lambda); /* the power */
power_twotail=CDF('t',tcrit_low, df, lambda)+1-CDF('t',tcrit_up, df,
lambda);
output;
end;
PROC print data=size (obs=1);
where power_onetail>0.8;
var alpha n df power_onetail;
run;
PROC print data=size (obs=1);
where power_twotail>0.8;
var alpha n df power_twotail;
run;
```

The DO statement directs calculation of power for the sample sizes from 2 to 1000. The following lines declare population means, standard deviations and express formula to calculate the degrees of freedom, pooled standard deviations and noncentral parameter (lambda). The critical values are computed using the TINV function and the statistical power with the CDF function. The output is printed with the first observation having the power >0.80.

The SAS output is:

| Obs | alpha | n | df | power_ one tail |
|---|---|---|---|---|
| 124 | 0.05 | 125 | 248 | 0.80146 |

| Obs | alpha | n | df | power_ two tail |
|---|---|---|---|---|
| 158 | 0.05 | 159 | 316 | 0.80216 |

It is concluded that the required sample sizes are at least 125 and 159 tilapia fish in each group for one and two-sided tests, respectively, to detect a difference between two groups at the significance level of 0.05 and with a statistical power of 0.80. The sample sizes increase to 218 and 262 for one and two-sided tests, respectively, if the power of the test is 0.95.

## B. Sample size including family structure

By ignoring family structure while determining sample size, the test for significance between strains may not be correct. The equation to determine sample size allowing for family structure is as follows (Hill 1980):

$$N_f = N_0 [1+(n-1)t] \quad [B1]$$

Here, $N_0$ is the number of animals per strain with no family structure; n is size of families, t is the intra-class correlation ($t = \frac{\sigma^2_B}{\sigma^2_B + \sigma^2_w}$) where $\sigma^2_B$ is the between-family and $\sigma^2_w$ the within family variance.

For instance, the number of animals per strain determined in the example A2 is 3403. Now assume there are 20 progeny per family (n= 20) and the intra-class correlation is 0.1 (t = 0.1). The required number of animals according to equation B1 is:

$$N_f = N_0 [1+ (n-1)t] = 3403[1+(20-1)0.1] = 9869$$

It is clear that the required number of animals greatly increases when taking family structure into consideration.

An alternative approach is to assume f families of size n will be used, so that the variance of a family mean is $\sigma^2[1 + (n-1)t]/n$ or $\sigma^2_{FM}$. Then equation [A4] can be used with f replacing n and CV being calculated with $\sigma^2_{FM}$ instead of $\sigma^2$. This could be done for a range of family sizes to find the most practical combination of number of families and family size. This seems a more useful procedure than fixing the family size in advance, unless this is necessary for practical reasons.

## C. Accuracy of strain comparison trials

The results from changes in different parameters and their impact on the accuracy are presented in Table C1, calculated using equation [4.2]. Table C1 shows that sire is the dominant factor affecting accuracy of the trial, as measured by the standard error of the strain difference (D). The values in the table are for trials in which sires from different strains are mated to dams from a common source. If sires are mated to dams from their own strain, the standard error is halved, because the difference between progeny means does not need to be doubled to estimate the strain difference.

**Table C1.** Accuracy (standard error of strain difference) expressed as percentage of phenotypic standard deviation. Smaller values show greater accuracy.

| $h^2$ | r | n | | | | |
|---|---|---|---|---|---|---|
| | | s | k | 10 | 50 | 100 |
| 0.05 | 0.05 | 5 | 1 | 42 | 23 | 19 |
| | | 5 | 2 | 32 | 19 | 17 |
| | | 5 | 3 | 28 | 18 | 16 |
| | | 10 | 1 | 30 | 16 | 13 |
| | | 10 | 2 | 23 | 13 | 12 |
| | | 10 | 3 | 20 | 13 | 11 |
| | | 20 | 1 | 21 | 11 | 9 |
| | | 20 | 2 | 16 | 10 | 8 |
| | | 20 | 3 | 14 | 9 | 8 |
| 0.05 | 0.15 | 5 | 1 | 42 | 23 | 19 |
| | | 5 | 2 | 33 | 19 | 17 |
| | | 5 | 3 | 30 | 18 | 16 |
| | | 10 | 1 | 30 | 16 | 13 |
| | | 10 | 2 | 23 | 14 | 12 |
| | | 10 | 3 | 21 | 13 | 12 |
| | | 20 | 1 | 21 | 11 | 9 |
| | | 20 | 2 | 17 | 10 | 9 |
| | | 20 | 3 | 15 | 9 | 8 |
| 0.30 | 0.30 | 5 | 1 | 52 | 39 | 37 |
| | | 5 | 2 | 47 | 37 | 36 |
| | | 5 | 3 | 45 | 37 | 36 |
| | | 10 | 1 | 37 | 27 | 26 |
| | | 10 | 2 | 33 | 26 | 25 |
| | | 10 | 3 | 32 | 26 | 25 |
| | | 20 | 1 | 26 | 19 | 18 |
| | | 20 | 2 | 23 | 19 | 18 |
| | | 20 | 3 | 23 | 18 | 18 |
| 0.30 | 0.40 | 5 | 1 | 52 | 39 | 37 |
| | | 5 | 2 | 47 | 37 | 36 |
| | | 5 | 3 | 45 | 37 | 36 |
| | | 10 | 1 | 37 | 27 | 26 |
| | | 10 | 2 | 33 | 26 | 25 |
| | | 10 | 3 | 32 | 26 | 25 |
| | | 20 | 1 | 26 | 19 | 18 |
| | | 20 | 2 | 23 | 19 | 18 |
| | | 20 | 3 | 23 | 18 | 18 |
| 0.70 | 0.70 | 5 | 1 | 64 | 55 | 54 |
| | | 5 | 2 | 62 | 55 | 54 |
| | | 5 | 3 | 62 | 55 | 54 |
| | | 10 | 1 | 45 | 39 | 39 |
| | | 10 | 2 | 44 | 39 | 38 |
| | | 10 | 3 | 44 | 39 | 38 |
| | | 20 | 1 | 32 | 28 | 27 |
| | | 20 | 2 | 31 | 27 | 27 |
| | | 20 | 3 | 31 | 27 | 27 |
| 0.70 | 0.80 | 5 | 1 | 64 | 55 | 54 |
| | | 5 | 2 | 63 | 55 | 54 |
| | | 5 | 3 | 62 | 55 | 54 |
| | | 10 | 1 | 45 | 39 | 38 |
| | | 10 | 2 | 44 | 39 | 38 |
| | | 10 | 3 | 44 | 39 | 38 |
| | | 20 | 1 | 32 | 28 | 27 |
| | | 20 | 2 | 31 | 28 | 27 |
| | | 20 | 3 | 31 | 27 | 27 |

$h^2$= heritability; r = repeatability; s=number of sires; k= number of repeated measurements; n = number of progeny.

## D. Cost-benefit considerations in strain comparisons

Details of the economic approach are given in Hill (1974) and SCA (1980). The equation for the expected return (R) from the test is given by:

$$R=Wi\,\sigma_b^2(\sigma_b^2+\sigma^2/n)^{-1/2}-nkC-F+DW \quad [D1]$$

where:

W = value of one unit of improvement for the characteristic in question.
i = standardized selection differential of the observed superiority of the best strains.
$\sigma_b^2$ = variance between strains.
$\sigma^2$ = variance within strain.
n = number of animals per strain.
k = number of strains.
C = variable cost of test per animal.
F = fixed cost of the test.
D = mean of all trial strains relative to mean of commonly used strains.

In the equation the response is calculated as the standardized selection differential (i) times the strain standard deviation ($\sigma_b$) times the correlation between true and observed strain means $\sigma_b/\sqrt{(\sigma_b^2+\sigma^2/n)}$. The selection response in [D1] is given in terms of selection for a single trait, which logically would need to be overall economic value, but might in practice be a trait of dominant importance. In practice also the value of D is unlikely to be known, unless some of the widely used strains are included in the trial. It is assumed that animals within a strain are unrelated in this equation. Another assumption is that there are no per-strain costs, only fixed and per-animal costs, in the trial, but this does not affect the choice of n, only the profitability.

If other parameters in [D1] are known, differentiation with respect to *n* and setting the derivative to zero gives the optimum size of the test by the following solution:

$$n^2(\sigma_b^2+\sigma^2/n)^{3/2}=\sigma_b^2\sigma^2Wi/2kC \quad [D2]$$

The upper limit of *n* is approximately

$$n<(\sigma^2\,Wi/2kC\,\sigma_b^2)^{3/2} \quad [D3]$$

When the optimum value of n is determined it can be entered into equation [D1] to calculate the economic return from the trial. If R turns out to be negative, then the trial should not be conducted.

Hill (1974) was primarily concerned with the general point that the possible gains from strain trials should be balanced against the trial costs and so did not consider more general designs. It is possible to modify his approach to consider designs with (say) f families of m progeny, and to calculate the correlation between the observed and true strain means with this design. The resulting correlation would replace $\sigma_b/\sqrt{(\sigma_b^2+\sigma^2/n)}$ in [D1]. In the simplest case the appropriate correlation would be $\sigma_b/\sqrt{(\sigma_b^2+\sigma_f^2/f+\sigma_w^2/fm)}$, where $\sigma_b^2$ and $\sigma_w^2$ are the between and within family variances within a strain. Evaluation of designs would be done numerically. The absence of family costs would mean that the best design would have m = 1, thus reducing to the case considered by Hill (1974), but it would be straightforward to add a family cost and use numerical optimization.

## E. Example of experimental design with strain identification only: performance comparison of GIFT and other Nile tilapia strains in Bangladesh

Brooders were selected from: (1) the latest generation of GIFT tilapia at Bangladesh Fisheries Research Institute (BFRI) or from the nucleus in Malaysia (120 females and 40 males), and (2) from other strains of Nile tilapia available in Bangladesh (120 females and 40 males for each strain).

Simultaneous spawning of all strains was carried out by separate stocking of chosen brooders in breeding hapas (three replicate mating hapas per strain).

Fry collection is made after 2 to 3 weeks of mating.

Collected fry of each strain are stocked in separate cages (at a density of approx. 6000 fry per cage, may vary with cage size).

Fingerlings are randomly sampled from cages for fin-clipping. The number of fingerlings to be fin-clipped depends on the number of testing environments and number of replicates per environment.

Fin clipping is conducted when the fish reach 3 to 5 g.

Fin clipped fingerlings are kept separately in conditioning tanks (or cages) after potassium permanganate ($KMNO_4$) treatment.

Transfer fin clipped fingerlings to testing environments for communal grow-out, three replicates per environment.

Testing environments could include, for example, ponds at BFRI, farmers' ponds and cages.

Records:

- Body weight at fin-clipping.
- Weight and length of 100 to 200 fish of each strain sampled in each location monthly or once every two months.
- At final harvest, body weight and length of all experimental fish.
- Sex of individual fish at harvest.
- Spawning date, nursing date, stocking date, harvesting date, initial weight and water parameters.

Water quality parameters (pH, temperature, alkalinity, NH3, visibility, conductivity, dissolved oxygen (DO)) are taken twice a month from each culture system, preferably at the same time of day.

A general format for data collection and data entry in spreadsheets for statistical analyzes is as follows:

| Strain | Spawning date | Finclip date | Stocking date | Stocking weight | Location | Harvest date | Harvest weight | Length | Notes |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

The analysis of variance for an experimental design such as that described in this section would proceed by fitting the effects indicated in the table below:

Suppose there are s strains tested in k environments with r replicates per strain (r = 3 in example) in each environment giving a total of skr replicates. Sex is fitted as a fixed effect with two levels. Preliminary analyzes would be conducted to test or adjust for covariates such as age at harvest and stocking weight. Covariates would be included as necessary in the final analyzes. The factors of interest in the analysis would be included in an analysis of variance in the following form, assuming the same number of fish measured in each replicate, with analysis of replicate means. The format given assumes data from both sexes are analyzed together, but on occasion each sex might be treated separately:

| SOURCE | DF | Expected Mean Square |
|---|---|---|
| Strains | s-1 | $V_R + rkV_S$ |
| Environments | k-1 | $V_R + rsV_E$ |
| Str × Env | (s-1)(k-1) | $V_R + rV_{SE}$ |
| Sex | 1 | $V_I + rksV_X$ |
| Sex × Str | s-1 | $V_I + rk\,V_{XS}$ |
| Sex × Env | k-1 | $V_I + rsV_{XE}$ |
| Sex × Str × Env | (s-1)(k-1) | $V_I + rV_{XSE}$ |
| Sex × Rep | sk(r-1) | $V_I$ |
| Replicates | sk(r-1) | $V_R$ |

This analysis of variance table is based on a split plot form of analysis, in which replicates are treated as main plots and sex is a sub-plot treatment, whereas strain and environment are main-plot treatments. $V_I$ is the sub-plot error mean square and $V_R$ is the main-plot mean square. The main purpose is to show that in this set-up the replicate means are the appropriate basic elements for testing significance and allocating standard errors. The fixed effect contributions are shown as variance components ('V' symbol). Although this is not strictly correct, it does not affect the above reasoning regarding how the different fixed effects ought to be tested.

## F. Example of experimental design with strain, family and animal identification

- Strains: four strains (GIFT from BFRI or from the nucleus in Malaysia and three other available strains of Nile tilapia currently cultured in Bangladesh)[1].
- Testing environment: one (standard grow-out earthen pond)[2].
- Sample 150 females (F) and 150 males (M) of similar weight per strain[3]. Bring all experimental populations to one station.
- Stock these four strains in four separate ponds for conditioning.
- Synchronize fry production for all strains, following single pair mating (1 F x 1 M). Target to produce 30 families (pair mating of 30 F x 30 M) per strain within two to three weeks.
- Collect 400 fry from each family and from all strains[4].
- Rear fry of each family in separate hapas within the same pond (two or three replicate nursing hapas per family).
- When the fish reach a body size between 5 and 15 g, randomly sample 50 fry of each family for physical tagging (using PIT tag)[5].
- A total of 6000 tagged fish (50 fish tagged per family x 30 families per strain x 4 strains = 6000) will be communally grown out in two ponds (a random sample of 25 fish per family in one pond and another 25 fish of each family in another pond)[6].
- Records:
  - Growth data
    - Body weight of individual fish at tagging.
    - Measure a sample of 100 to 200 fish of each strain every 2 months (or every month).
    - At final harvest (after a grow-out period of 4 to 6 months), measure body weight and length on individual fish.
    - Record sex of individual fish at harvest.
    - Record 'tag lost' fish at final harvest in order to take this into consideration when calculating survival rate.
  - Reproduction data (breeders).
    - Total number of eggs per spawning per female.
    - Total number of fry per spawning per female.
    - Total fry weight.
    - Body weight of females prior to mating.
  - Additional measurements for consideration.
    - Fillet weight.
    - Chemical composition (protein, fat, moisture and ash content).
    - Flesh quality attributes (pH, color, texture, water holding capacity).
  - Other data recording: spawning date, nursing date, stocking date, harvesting date, initial weight, testing location, water parameters at each location.

With 4 strains, 2 ponds, 30 families of 50 fish per strain, assuming for simplicity of presentation that all fish survive, and that sex and other effects can be ignored, an analysis of variance of the following form would be carried out:

| SOURCE | DF | Expected Mean Square |
|---|---|---|
| Strains | 3 | $V_W + 50V_F + 1500V_S$ |
| Ponds | 1 | $V_W + 25V_{PF} + 3000V_P$ |
| Strains × Ponds | 3 | $V_W + 25V_{PF} + 750V_{SP}$ |
| Families(Strains) | 116 | $V_W + 50\,V_F$ |
| Ponds× Families(Strains) | 116 | $V_W + 25V_{PF}$ |
| Fish(Families × Ponds) | 5760 | $V_W$ |

If sex cannot be ignored, there should be a term for sex, and terms for interactions involving sex. In the table it is assumed that pond is a fixed effect, so there is no component of variance for Strains by Ponds interaction included in the Strains expected mean square. In practice the coefficients of the variance components would differ because of unequal survival, but the main ideas would be as shown.

Notes:

[1] Evaluation of more than four strains may be required in some instances.

[2] Test the fish in more than one environments if resources are available, e.g. in cages, semi-intensive systems.

[3] Alternatively, sample about one thousand fry or fingerlings per population and rear them to sexual maturity at the same station (also under the same culture environment).

[4] Number of fry nursed per family × *n* testing environments. For example, if there are 3 test environments, so we would need to nurse 400 × 3 = 1,200 fry.

[5] Number of fish per family tagged × number of test environments. For example, if there are 3 test environments, so we would need to tag 50 × 3 = 150 fish per family.

[6] Number of communally grown fish can be adjusted to fit pond size or rearing areas. In all cases, fish from each family and strain must be represented in all ponds (and environments).

## References


Bailey, J.K. and Loudenslager, E.J., 1986. Genetic and environmental components of variation for growth of juvenile Atlantic salmon (*Salmo salar*). *Aquaculture* 57: 125-132.

Biffani, S., Samore, A., Canavesi, F. and Boettcher, P., 2001. Impact of data structure on validation of genetic trend in a small population. Proceedings of the 2001 INTERBULL meeting, Budapest. INTERBULL Bulletin No. 27:143.

Bentsen, H.B., Eknath, A.E., Palada-de Vera, M.S., Danting, J.C., Bolivar, H.L., Reyes, R.A., Dionisio, E.E., Longalong, F.M., Circa, A.V., Tayamen, M.M. and Gjerde, B., 1998. Genetic improvement of farmed tilapias: Growth performance in a complete diallel cross experiment with eight strains of *Orechromis niloticus. Aquaculture* 160: 145-173.

Blanc, J.M., 2003. Comparison of experimental designs for estimating quantitative genetic parameters in fish. *Aquaculture Research* 34: 1099-1105.

Cheng, K.M., McCallum, I.A., McKay, R.I. and March, B.E., 1987. A comparison of survival and growth of two strains of Chinook salmon (*Oncorhynchus tshawytscha*) and their crosses reared in confinement. *Aquaculture* 67: 301-311.

Cruz, P., Ramirez, J.L., Garcia, G.A. and Ibarra, A.M.,1998. Genetic differences between two populations of catarina scallop (*Argopecten ventricosus*) for adaptations for growth and survival in a stressful environment. *Aquaculture* 166: 321-335.

Eknath, A.E., Tayamen, M.M., Palada-de Vera, M.S., Danting, J.C., Reyes, R.A., Dionisio, E.E., Capili, J.B., Bolivar, H.L., Abella, T.A., Circa, A.V., Bentsen, H.B., Gjerde, B., Gjedrem, T. and Pullin, R.S.V., 1993. Genetic improvement of farmed tilapias: the growth performance of eight strains of *Orechromis niloticus* tested in different farm environments. *Aquaculture* 111:171-188.

Erdfelder, E., Faul, F. and Buchner, A., 1996. GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, and Computers* 28: 1-11.

Gall, G.A.E. and Gross, S.J., 1978. A genetic analysis of the performance of three rainbow trout broodstocks. *Aquaculture* 15: 113-127.

Gunnes, K. and Gjedrem, T., 1978. Selection experiments with salmon IV. Growth of Atlantic salmon during two years in the sea. *Aquaculture* 15: 19-33.

Hall, S.J.G., 2004. Livestock Biodiversity: Genetic Resources for the Farming of the Future. Blackwell Science Ltd., Oxford, UK, 269 pp.

Havenstein, G.B., Ferket, P.R., Scheideler, S.E. and Larson, B.T., 1994. Growth, livability and feed conversion of 1991 versus 1957 type broilers when fed "typical" 1957 and 1991 broiler diets. *Poultry Science* 73: 1785–1794.

Herbinger, C.M., O'Reilly, P.T., Doyle, R.W., Wright, J.M. and O'Flynn, F. 1999. Early growth performance of Atlantic salmon full-sib families reared in single family tanks versus in mixed family tanks. *Aquaculture* 173: 105–116.

Hill, W.G., 1974. Size of experiment for breed or strain comparison. Proceeding of Working Symposium on breed evaluation and crossing experiments, Zeist, Germany, p 43-54.

Hill, W.G., 1980. Experimental design in quantitative genetics and animal breeding. Lecture Notes, Gottingen University, Germany, 70 pp.

Ibrahim, N.A., Zead, M.Y.A., Khaw, H.L., El-Naggar, G.O. and Ponzoni, R.W., 2012. Relative performance of two Nile tilapia (*Oreochromis niloticus*) strains in Egypt: The Abbassa selection line and the Kafr El Sheikh commercial strain. *Aquaculture Research* 44: 508-517.

Jacobs, J.M., Lindell, S., Heukelem, W.V., Hallerman, E.M. and Harrell, E.M., 1999. Strain evaluation of striped bass (*Morone saxatilis*) under controlled conditions. *Aquaculture* 173: 171-177.

James, J.W., 1975. Genetical considerations in large field experiments. In Developments in field experiment design and analysis. *Comm. Agric.Bureau.*: 155-167

Jerry, D.R., Purvis, I.W. and Piper, L.R., 2002. Genetic differences in growth among wild populations of the yabby, *Cherax destructor* (Clark). *Aquaculture Research* 33: 917–923.

Jonasson, J., 1996. Selection experiments on Atlantic salmon ranching. II. Variation among release sites and strains for return rate, body weight and ratio of grilse to total return. *Aquaculture* 144: 277-294.

Jones, C.M., McPhee, C.P. and Ruscoe, I.M., 2000. A review of genetic improvement in growth rate in red claw crayfish *Cherax quadricarinatus* (von Martens) (Decapoda: Parastacidae). *Aquaculture Research* 31: 61–67.

Kanis, E., Refstie, T. and Gjedrem, T., 1976. A 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.

Kennedy, B.W., 1990. Use of mixed model methodology in analysis of designed experiments. In Gianola, D. and Hammond, K. (eds). Advances in statistical methods for genetic improvement of livestock. Berlin, Springer-Verlag.

Komender, P. and Hoeschele I., 1989. Use of mixed-model methodology to improve estimation of crossbreeding parameters. *Livest. Prod.Sci.* 21: 101-113.

Lenth, R.V., 2001. Some Practical Guidelines for Effective Sample Size Determination. *The American Statistician* 55: 187-193.

Macaranas, J.M., Mather, P.B., Lal, S.N., Vereivalu, T., Lagibalavu, M. and Capra, M.F., 1997. Genotype and environment : A comparative evaluation of four tilapia stocks in Fiji. *Aquaculture* 150: 11-24.

Maluwa, A.O. and Gjerde, B., 2006. Genetic evaluation of four strains of *Oreochromis shiranus* for growth in a diallel cross. *Aquacultur*e 259: 28-37.

Morris, T.R. 1999. Experimental design and analysis in animal sciences. CABI, Wallingford, oxon, UK.

Ninh, H.N., Ponzoni, R.W., Nguyen, N.H., Woolliams, J.A., McAndrew, B.J. and Penman, D.J., 2011. Communal or separate rearing of families in selective breeding of common carp (*Cyprinus carpio*). *Aquaculture* 322-323: 39-46.

Oldorf, W., Kronert, U., Balarin, J., Haller, R., Hörstgen-Schwark, G. and Langholz, H.J., 1989. Prospects of selecting for late maturity in Tilapia (*Oreochromis niloticus*): II. Strain comparisons under laboratory and field conditions. *Aquaculture* 77: 123-133.

Overturf, K., Casten, M.T., LaPatra, S.L., Rexroad III, C. and Hardy, R.W., 2003. Comparison of growth performance, immunological response and genetic diversity of five strains of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 217: 93-176.

Pante, Ma. J.R., Gjerder, B., McMillan, I. and Misztal, I., 2002. Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* 204: 383-392.

Ponzoni, R.W., Nguyen, N.H. and Khaw, H.L., 2011a. Fundamental considerations about design and sample size in strain comparisons and their implications. *Aquaculture Research* 42: 1855 – 1858.

Ponzoni, R.W, Nguyen, N.H., Khaw, H.L., Hamzah, A., Abu-Bakar, K.R. and Yee, H.Y., 2011b. Genetic improvement of Nile tilapia (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish Center with the GIFT strain. *Reviews in Aquaculture* 3: 27-41.

Ponzoni, R.W., Hamzah, A., Tan, S. and Kamaruzzaman, N., 2005. Genetic parameters and response to selection for live weight in the GIFT strain of Nile Tilapia (*Oreochromis niloticus*). *Aquaculture* 247: 203-210.

Quinton, C.D., McKay, L.R. and McMillan, I., 2004. Strain and maturation effects on female spawning time in diallel crosses of three strains of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* 234: 99-110.

Reddy, P.V.G.K., Gjerde, B., Tripathi,S.D., Jana, R.K., Mahapatra, K.D., Gupta, S.D., Saha, J.N., Sahoo, M., Lenka, S., Govindassamy, P., Rye, M. and Gjedrem, T., 2002. Growth and survival of six stocks of rohu (*Labeo rohita, Hamilton*) in mono and polyculture production systems. *Aquaculture* 203: 239-250.

Reed, D.H. and Frankham, R., 2001. How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution* 55: 1095-1103.

Refstie, T. and Steine, T.E., 1978. Selection experiments with salmon III. Genetic and environmental sources of variation in length and weight of Atlantic salmon in the freshwater phase. *Aquaculture* 14: 221-234.

Refstie, T., Steine, T.E. and Gjedrem, T., 1977. Selection experiments with salmon. II. Proportion of Atlantic salmon smoltifying at 1 year of age. *Aquaculture* 10: 231-242.

Romana-Eguia, M.R.R. and Doyle, R.W., 1992. Genotype-environment interaction in the response of three strains of Nile tilapia to poor nutrition. *Aquaculture* 108: 1-12.

Romana-Eguia, M.R.R. and Eguia, R.V.,1999. Growth of five Asian red tilapia strains in saline environments. *Aquaculture* 173:161-170.

Rye, M. and Mao, I.L. 1998. Nonadditive genetic effects and inbreeding depression for body weight in Atlantic salmon (*Salmo salar L.*). *Livestock Production Science* 57: 15-22.

Sanford, L.D., Tang, X., Ross, R.J. and Morrison, A.R., 2003. Influence of shock training and explicit fear- conditioned cues on sleep architecture in mice: strain comparison. *Behavior Genetics* 33: 43–58.

SAS (Statistical Analysis System) Institute Inc., 2008. SAS/STAT® 9.2 User's Guide. Cary, NC: SAS Institute Inc., USA.

Standing Committee on Agriculture (SCA), 1980. The design and conduct of production competitions for comparison of genotypes. Report No.6, Technical Report Series, Canberra, Australia, 26 pp.

Standing committee on Agriculture (SCA), 1982. Recommended procedures for the sampling, introduction, evaluation and dissemination of genotypes imported into Australia. Technical Report Series, Canberra, Australia, 38 pp.

Sellier, P., 1980. Design and analysis of quantitative genetics experiments: a review with particular references to pigs. *Livest. Prod. Sci.* 7: 539-544.

Snedecor, G.W. and Cochran, W.G., 1971. Statistical Methods. Sixth Edition. Iowa State University Press, Ames, USA.

Sokal, R.R. and Rohlf, F.J., 1969. Biometry. The principles and practice of statistics in biological research. W. H. Freeman and Company, San Francisco, USA.

Sorensen, D.A. and Kennedy, B.W., 1984. Estimation of response to selection using least-squares and mixed model methodology. *J. Anim. Sci. 58*: 1097-1106.

SPSS (Statistical Package for the Social Sciences) Inc., 2011. IBM SPSS Statistics for Windows, Version 19.0. Armonk, NY: IBM Corp.

Stroup, W.W., 1999. Mixed model procedure to assess power, precision and sample size in the design of experiments. Proceedings of the Biopharmaceutical Section, Alexandria, VA: American Statistical Association, p 15-24.

Stroup, W.W., 2001. Power analysis based on spatial effect mixed models: A tool for comparing design and analysis strategies in the presence of spatial variability. *Journal of Agricultural Biological and Environmental Statistics* 7: 491-511.

Sylvén, S. and Elvingson, P., 1992. Comparison of rainbow trout (*Oncorhynchus mykiss*) strains for body weight, length and age at maturity in different Swedish production systems. *Aquaculture* 104: 37-50.

Taylor, St. C.S., 1976a. Multibreed Designs 1. Variation between breeds. *Anim. Prod.* 23: 133-144.

Taylor, St. C.S., 1976b. Multibreed Designs 2. Genetic variation within and between breeds. *Anim. Prod.* 23: 145-154.

Thanh, M.N., Nguyen, N.H., Ponzoni, R.W., Vu, N.T., Barnes, A. and Mather, P.B., 2010. Estimates of strain additive and non-additive genetic effects for growth traits in a diallel cross of three strains of giant freshwater prawn (*Macrobrachium rosenbergii*) in Vietnam. *Aquaculture* 299: 30-36.

Unwin, M.J., Kinnison, M.T., Boustead, N.C. and Quinn, T.P., 2003. Genetic control over survival in Pacific salmon (*Oncorhynchus spp.*): experimental evidence between and within populations of New Zealand Chinook salmon (*O. tshawytscha*). *Can. J. Fish. Aquat. Sci.* 60: 1-11.

Vandeputte, M., Dupont-Nivet, M., Chatain, B. and Chevassus, B., 2001. Setting up a strain-testing design for the seabas, *Dicentrarchus labrax*: a simulation study. *Aquaculture* 202: 329-342.

Vandeputte, M., Peignon, E., Vallod, D., Haffray, P., Komen, J. and Chevassus, B., 2002a. Comparison of growth performance of three French strains of common carp (*Cyprinus carpio*) using hemi-isogenic scaly carp as internal control. *Aquaculture* 205: 19-36.

Vandeputte, M., Quillet, E. and Chevassus, B., 2002b. Early development and survival in brown trout (*Salmo trutta fario L.*): indirect effects of selection for growth rate and estimation of genetic parameters. *Aquaculture* 204: 435-445.

Winkelman, A.M. and Peterson, R.G., 1994. Heritabilities, dominance variation, common environmental effects and genotype by environment interactions for weight and length in chinook salmon. *Aquaculture* 125: 17-30.

## Other publications of interest related to strain comparisons


Cameron, N.D., 1997. Selection indices and prediction of genetic merit in animal breeding. CAB International. Wallingford, Oxon OX10 8DE, UK.

Dickerson, G.E., 1969. Experimental approaches in utilising breed resources. *Anim. Breed. Abstr.* 37: 191- 202.

Dupont-Nivet, M., Vandeputte, M. and Chevassus, B., 2002. Optimization of factorial mating designs for inference on heritability in fish species. *Aquaculture* 204: 361-370.

Falconer, D.S., 1952. The problem of environment and selection. *Amer. Nat.* 86: 293-298.

Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R., 1999. Asreml reference manual. NSW Agriculture Biometric Bulletin No.3. Orange Agricultural Institute, Forest Road, Orange 2800 NSW Australia.

Gitterle, T., Rye, M., Salte, R., Cock, J., Johansen, H., Lozano, C., Suárez, J.A. and Gjerde, B., 2005. Genetic (co)variation in harvest body weight and survival in *Penaeus* (*Litopenaeus*) *vannamei* under standard commercial conditions. *Aquaculture* 243: 83-92.

Henderson, C.R., 1984. Applications of linear models in animal breeding. University of Guelph, Guelph.

Horstgen-Schwark, G. and Langholz, H.-J., 1998. Prospects of selecting for late maturity in tilapia (*Oreoshromis niloticus*) III. A selection experiment under laboratory conditions. *Aquaculture* 167: 123- 133.

Kause, A., Ritola, O., Paananen, T., Mäntysaari, E., and 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.

Kirpichnikov, V.S., 1966. Methods of progeny testing spawners in carp fish farms. Selective breeding in carp. *Israel Program for Scientific transaction* 5737: 36-55.

Lemarié, G., Baroiller, G.F., Clota, F., Lazard, J. and Dosdat, A., 2004. A simple test to estimate the salinity resistance of fish with specific application to *O. niloticus* and *S. melanotheron*. *Aquaculture* 240: 575- 587.

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

Meyer, K., 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. *Livestock Production Science* 86: 69-83.

Moav, R. and Wolfarth, G.W., 1974. Magnification through competition of genetic differences in yield capacity in carp. *Heredity* 33: 181-202.

Mrode, R.A., 1996. Linear models for the prediction of animal breeding values. CAB International. Wallingford, Oxon OX10 8DE, UK.

Rutten, M.J.M., Komen, H. and Bovenhuis, H., 2005. Longitudinal genetic analysis of Nile tilapia (*Oreochromis niloticus L.*) body weight using a random regression model. *Aquaculture* 206: 101-113.

Saxton, A., 2004. Genetic selection. In Genetic analysis of complex traits using SAS. SAS Institute Inc.

Schaeffer, L.R., 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86: 35-45.

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

Withler, R.E. and Beacham, T.D., 1994. Genetic variation in body weight and flesh colour of the coho salmon (*Oncorhynchus kisutch*) in British Columbia. *Aquaculture* 119: 135-148.

Wohlfarth, G.W., Moav, R., Hulata, G. and Beiles, A., 1975. Genetic variation in seine escapability of the common carp. *Aquaculture* 5: 375-387.

This publication should be cited as: R.W. Ponzoni, J.W. James, N.H. Nguyen, W. Mekkawy and H.L. Khaw. (2013) Strain comparisons in aquaculture species: a manual. WorldFish, Penang, Malaysia. Manual: 2013-12.

The CGIAR Research Program on Livestock and Fish aims to increase the productivity of small-scale livestock and fish systems in sustainable ways, making meat, milk and fish more available and affordable to poor consumers across the developing world. For more information. http://livestockfish.cgiar.org/.

Design and layout: Eight Seconds Sdn Bhd.

Photo credits: Front cover, Khaw Hooi Ling; back cover, Azhar Hamzah (Department of Fisheries Malaysia). Printed on 100% recycled paper.



Contact Details:
WorldFish, PO Box 500 GPO,
10670 Penang, MALAYSIA
Web: www.worldfishcenter.org

WorldFish

CGIAR RESEARCH PROGRAM ON Livestock and Fish