Critical assessment of the “internal reference” method to eliminate non-genetic effects within a Combined Family Selection program on the abalone species (*Haliotis midae*)

by

Gareth Frank Difford

Thesis presented in partial fulfilment of the requirements for the degree Master of Science (Agriculture) in Genetics and Animal breeding at the University of Stellenbosch

Supervisor: Prof Daniel Brink
Co-supervisors:
Faculty of Agricultural Science
Department of Genetics
Division of Aquaculture

December 2013

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2013

ABSTRACT

The aim of this study was to critically assess the implementation of the internal reference method within the most recent 173 full-sibling growth trial of the Innovation Fund Abalone Breeding Project. The trial was conducted over two locations for a period of five years, with minimal replication for the majority of test families and a single full-sibling family was entered into each experimental unit (basket) as an internal reference group. The primary focus was firstly, to validate the performance of the internal reference group as a control for comparisons and correction of environmental variation in test family performances. Secondly, to identify areas of weakness and either make recommendations to remedy areas of weakness or justify devoting resources to alternative methods of reducing extraneous environmental variance with limitations on replication. The efficiency and statistical power associated with utilising internal reference information as a covariate and for manual correction respectively were examined for the 6 full-sibling test families that were replicated.

This study reports on the evaluation of factors which are potential sources of bias in the internal reference method, the first of which, tag loss, was found to be significant after 6-12 months. However, it was not found to bias internal reference group performances as the factors which contribute to tag loss were found to act randomly. Variability in size ratio of internal reference to test family at co-stocking proved a significant source of bias, as reference groups smaller than their test family counterparts had reduced performances. Testing for genotype by environment interactions was precluded due to the inherent lack of replication and the subsequent confounding of genotype effects with inter-rearing structure effects at one of the locations. However, significant differences were detected for both traits of interest of the internal reference group over the two locations. Significant antagonistic interactions were detected and identified as a source of bias for average daily weight gain of replicate test families.

The evaluation of average daily length gain for the efficiency of adjustment when the internal reference is a covariate and the change in statistical power when the internal reference is used for a manual correction, yielded conflicting results. The latter shows a decrease in statistical power and the former shows an increase in efficiency, both resulting in poor goodness of fit in the respective models. There was however evidence that when no antagonistic interactions occurred “between replicate variance” decreased and therefore the internal reference method has statistical merit provided all critical success factors are satisfied.

Recommendations were made for future implementation of the internal reference method to facilitate adequate statistical testing for sources of bias and the prevention thereof. Additionally, an alternative method which may have merit in decreasing environmental variance and the need for replication, is discussed.

OPSOMMING

Die doel van die studie was om die gebruik van ’n interne verwysingsgroep te ontleed, soos toegepas tydens die evaluering van 173 volsib families as deel van die Innovasiefonds Perlemoen Teelprogram. Die evaluering is gedoen op twee lokaliteite oor 'n tydperk van vyf jaar, met minimale replikasie van die toets families en die gebruik van ‘n enkele volsib familie as 'n interne verwysingsgroep in elke eksperimentele eenheid (mandjie). Die primêre fokus was eerstens om die gebruik van die interne verwysingsgroep vir die korreksie van omgewingsvariasie in die toets familie optredes te evalueer. Tweedens, om spesifieke gebreke te identifiseer ten opsigte van die gebruik van die interne verwysingsgroep en aanbevelings maak dit reg te stel en om die meriete van alternatiewe metodes te oorweeg. Die doeltreffendheid en statistiese onderskeidingsvermoë van die gebruik van interne verwysingsgroep as 'n kovariaat is ondersoek met betrekking tot die 6 volsib groepe wat oor voldoende replikasies beskik het.

Die studie doen voorts verslag oor die evaluering van potensiële oorsake van sydigheid ten opsigte van die gebruik van die interne verwysingsgroep, insluitend die beduidende verlies van identifikasie vanaf 6 tot 12 maande. Geen aanduiding van sydigheid is egter gevind en die aanleidende oorsake van verlies van identifikasie blyk van ’n ewekansige aard te wees. Verskille in die grootte tussen die interne verwysingsgroep en toets-families met aanvang van evaluering blyk 'n belangrike bron van sydigheid te wees, waar die kleiner groepering aan verminderde prestasie gekoppel word. Bepaling van genotipe-omgewing-interaksies kon nie uitgevoer word nie as gevolg van die inherente gebrek van replisering oor lokaliteite. Beduidende verskille is egter waargeneem tussen interne verwysingsgroepe oor die twee lokasies ten opsigte van die beide groei eienskappe. Beduidende antagonistiese interaksies is waargeneem en geïdentifiseer as 'n bron van sydigheid ten opsigte van die gemiddelde daaglikse gewigstoename van replikaat toetsfamilies.

Die evaluering van gemiddelde daaglikse lengtetoename met die interne verwysingsgroep as is 'n kovariaat en die verandering in statistiese ontledingsvermoë tydens die gebruik van die interne verwysingsgroep het teenstrydige resultate opgelewer. Laasgenoemde toon 'n afname in statistiese ontledingsvermoë en die eersgenoemde toon 'n toename in doeltreffendheid, met beide swak passing op die onderskeie modelle. In die afwesigheid van antagonistiese interaksies tussen replikasies het variansie afgeneem en beskik die interne verwysingsgroep oor die nodige statistiese meriete indien daar aan al die kritiese vereistes voldoen word.

Aanbevelings is gemaak ten opsigte van die toekomstige implementering van die interne verwysingsmetode met verwysing na voldoende statistiese toetsing vir bronne van sydigheid en die voorkoming daarvan. 'n Verdere metode wat oor die nodige meriete beskik om die omgewingsvariasie en die noodsaaklikheid vir replikasie te verminder, word bespreek.

ACKNOWLEDGEMENTS

I would like to acknowledge and express gratitude to the following individuals and institutions for their support and contributions, without which this thesis would not have been possible:

Professor Danie Brink, Department of Genetics, Stellenbosch University, whose constant guidance and input proved invaluable.

Mev Analine Sadie, Department of Genetics, Stellenbosch University, for guidance and encouragement through statistical analyses.

Mev Gail Jordaan, Department of Animal Sciences, Stellenbosch University, for never giving up even when SAS did

Mnr Theo Pepler, Department of Genetics, Stellenbosch University, for supplying relevant statistical literature.

Dr Steven Jansens and *Prof Nadine Buys* contributing to my genetics knowledge.

Mnr Arnold Vlok and *Mej Mariette Gerber* for the many hours spent acquiring the data.

Mnr Michael Joubert for completing the 2012 measurements.

The staff at Abagold (Pty.) Ltd. and HIK Abalone (Pty.) Ltd. for assistance in managing and measuring the abalone.

My friends and family for their support and understanding

My parents for contributing the genes and the environment which got me where I am today

Crystal for trying to be patient…

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS & SYMBOLS

%	Percentage
<	Less than
(Pty) Ltd.	Property Limited
ABA	Abagold (Pty) Ltd.
ADLG	Average Daily Length Gain (5 years)
ADWG	Average Daily Weight Gain (5 years)
AFSA	Abalone Farmers Association of South Africa
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
BLUP	Best Linear Unbiased Predictor
cm	Centimetres
CO_2	Carbon Dioxide
CV	Coefficient of Variation
FAO	Food Agriculture Organisation of United Nations
FAOSTAT	Food Agriculture Organisation: Statistics Division
g	Grams
GLM	Generalized Linear Models
H	Height (Dimension)
Ha	Alternative Hypothesis
Ho	Null Hypothesis
HIK	HIK (Pty) Ltd.
L	Length (Dimension)
LS-means	Least-Square Means
Max.	Maximum
Min.	Minimum
mm	millimetres
mt	Metric tons
N	Count
p	Statistical Probability
Std Dev	Standard Deviation
TM	Trade Mark
USD	United States of America Dollar
W	Width (Dimension)

Chapter 1

RATIONAL

1.1 Background

The abalone is a marine gastropod species which is consumed globally and is of great economic importance to commercial fisheries and aquaculture in tropical and temperate oceans around the world. Abalone belongs to the family Haliotidae and the genus *Haliotis*, of which there are 56 recognised species (Geiger & Owen, 2012). One of the species of greatest commercial importance is the Perlemoen abalone (*Haliotis midae*), endemic to South Africa and noted for favourable organoleptic properties of meat and large size (Van der Merwe, 2009). *H. midae* production accounted for approximately 8.6% of the value of the total 2011 global abalone production of 676 million USD (FAOSTAT, 2012).

The South African abalone farming industry began in the early 1990's, following the first successful spawning of the indigenous species (*Haliotis midae*) in captivity (Genade *et al*., 1988). Development of the industry was driven by high international prices, declining fishery and capture quotas and the subsequent closing of the commercial abalone fishery in 2008 (Cook, 1998; Raemaekers *et al*., 2011). The industry rapidly expanded due to successful transfer of parallel technologies from abalone species cultured internationally as well as availability of favourable coastal habitats with associated infrastructure and accessibility to an agreeable work force in South Africa (Troel *et al*., 2006). The abalone industry in South Africa is not surprisingly the most affluent component of the South African aquaculture industry, with production reaching 1052 mt and a value of 41.53 million USD in 2011, comprising more than 80% of the rand value of the aquaculture outputs (Sales & Brits, 2001; Kirkendale *et al*., 2010; FAOSTAT, 2012).

The South African abalone industry is distinguished as a global leader in land-based culture systems, utilising intensive pump ashore technologies (Kirkendale, *et al*., 2010), towards the production of a premium value product (Cook & Gordon, 2010; AFASA, 2010). This has led to South Africa being one of the largest producers of abalone outside of Asia (Sales & Brits, 2001). Over the period of 2002 until 2010 the total production of abalone increased by approximately 235%, reaching a production level of 2500 tons per annum (FAO, 2010; AFASA, 2010). This increase in production, although significant, is dwarfed when contrasted by that of China, which achieved a 575% increase in production over the 2002 – 2010 period, culminating with a total production of 56 511 tons (Nie & Wang, 2004; FAO, 2010). This increasing trend is mirrored in Chile, the second largest producer of abalone outside of Asia, which achieved an approximate increase of 700% and total production of 794 tons over the same period of time (Flore-Aguilar *et al*., 2007; FAO, 2010).

Closer examination of the South African abalone industry shows specific limiting factors such as lack of single governmental aquaculture policy, strengthening Rand value, an escalation in cost of electricity, the downward economic turn and increasing costs of land near suitable coastline (AFASA, 2010). These limiting factors appear to negatively influence the profitability of the industry, the growth capabilities and international competitiveness. The industry has little control over these limiting factors, so strategic planning towards increased efficiency and productivity are required to restore profitability and competitiveness.

All cultured abalone in South Africa are derived from wild, genetically undomesticated brood stocks, juxtaposed to other traditional animal production industries globally and locally, for example dairy, swine and poultry, which make use of highly superior genetic material (Eknath *et al*., 1991; Gjedrem, 2005). The benefit cost ratio of genetic improvement in aquaculture species has been estimated between 5:1 and 50:1 (Gjedrem, 2005). With the broad aims of improving the profitability and international competitiveness of the industry, a genetic improvement programme for the species *Haliotis midae* was established in 2005. The programme comprised of a core selective breeding element and multiple related fields of molecular genetics, biotechnology and assisted reproduction. The specific aim of the selective breeding element is to increase biological productivity of the species through improvement of production traits namely, growth rate and yield within a combined family selection structure.

1.2 Internal Reference Method

The largest comparative growth trial within the selective breeding program of the Perlemoen abalone has made use of the 'internal reference' technique as a tool to reduce random environmental effects. Although this technique has been implemented since 2006, the method has not been assessed in terms of necessity, the optimal implementation procedure and the efficiency. The risks associated with the incorrect implementation of the 'internal reference' method within a selective breeding program can result in costs of foregone productivity as well as decreased genetic response (Basiao *et al*., 1996). The families spawned for the growth trial were almost entirely genetically unrelated full sibling families, of 173 families spawned only 6 families had successful enough settlement numbers to be replicated. In addition, limitations on resources and biosecurity concerns resulted in the use of two locations, as opposed to the five locations conventionally utilised for the comparative growth trials. One of the full sib families recorded extraordinary success in terms of settlement that allowed for this family to be replicated across locations as well as to be used as an internal reference group amongst the other full sib families.

While the lack of replication greatly diminishes the statistical value of the trial, an opportunity was realised to utilise the extremely successful family in an attempt to increase

accuracy of growth comparisons in the absence of replication. A *posteriori* analysis of the few replicated families included in the trial and the performance of the internal reference may reveal useful information for future trials and the selection program. The benefit of investigating the technique critically being obtaining the necessary information required to make educated decisions on the further use and optimisation of the method, or justify devoting resources to investigating alternative methods.

The appeal of the 'internal reference' method only becomes apparent when viewed specifically within the context of aquaculture. Statistical analyses conducted within aquaculture environments in particular, are sensitive to the multiple random environmental variables which are too costly or impractical to control directly, for example water temperature, diet and water quality (Moav *et al*., 1976; Basiao & Doyle, 1990; Kocour *et al*., 2005). Resultantly, inter-rearing structure variability is high, which subsequently inflates the error variance within experiments and lowers the sensitivity of the statistical analyses (Buck *et al*., 1970; Uraiwan & Doyle, 1986; Ling & Cotter, 2003). In the cases where individual families or strains must be reared separately throughout the selection experiment, the inter-rearing structure differences are confounded with the genetic differences of the individual families or strains in question (Moav *et al*., 1976).

Increasing replicate numbers will decrease the error variance and increase the precision of estimation of treatments (Cochran & Cox, 1957). However, this too is very costly and often impractical due to limitations in testing space, resources and reproductive capacity of the species (Gjedrem, 2005). In some instances *posteriori* statistical power analyses of aquaculture experiments, revealed that the adequate number of replications required to achieve acceptable statistical sensitivity far exceeds the capacity of the production unit (Ling & Cotter, 2003; Seary-Bernal, 1994). This poses a problem for breeding programmes in aquaculture environments, particularly in developing countries, as testing too few families with adequate replication can result in decreased selection intensity, genetic response and increases the chances of inbreeding and genetic drift. On the other hand testing many families without replication can result in type ii errors, in which a null hypothesis is incorrectly 'accepted' and this too will result in decreased genetic response to selection (Fairweather, 1991). The downstream economic implications of both situations are opportunity costs of foregone productivity, potentially reaching millions of USD (Basiao *et al*., 1996).

A possible solution to this hazardous situation is the 'internal reference' method (Kirpichnikov, 1966; Moav & Wohlfarth, 1974; Doyle *et al*., 1990), in which a distinguishable control group is included within each experimental unit, the control and treatment being subjected to the same random environmental variables which contribute to differences between replicates. This enables the partitioning and removal of environmental variance common to all observations in the experimental unit and subsequently, increases precision

of estimation of treatments with relatively little cost (Cochran, & Cox. 1957). Depending on the magnitude of environmental variation, it is possible to use the 'internal reference' method to supplement replication and achieve acceptable statistical sensitivity, reduce costs and facilitate the evaluation of a larger number of families (Moav & Wohlfarth, 1974; Basiao *et al*., 1996).

The use of the internal reference method information can be utilised in two manners as a "base line" to evaluate test group performances free of inter-rearing structure environmental effects (Doyle *et al*., 1990). The first method is to simply subtract the reference performance from that of the respective test group performance and proceed with an analysis of variance or rank performances (Hill, 1972a; Muir, 1986). The second method is to utilise the the internal reference group as a concomitant covariate in an analysis of covariance (Fisher, 1932; Doyle *et al*., 1990). According to Kirpichnikov (1966) and Basiao et al, (1990), the internal reference method is heavily reliant on the following factors and statistical considerations to precluded unwanted bias:

- Reference group individuals must be absolutely distinguishable from test individuals
- Conditions prior to stocking must be the same for all individuals
- When testing in multiple environments, no genotype by environment interaction must occur
- The weights (size) of reference group individuals must be the same as test individuals at co-stocking
- The test or reference group must not benefit from suppressing the other (antagonistic interactions)

1.3 Conclusion

This study focuses on a dataset obtained from a comparative growth trial of full sibling families as part of the South African abalone genetic improvement program obtained from the 2007-2008 spawning season. The aim of this is study is to critically analyse the current implementation of the internal reference method, in terms of the factors and statistical considerations listed above to determine if the following:

1. The justification for the inclusion of an internal reference group.
2. The use of the internal reference group to facilitate selection over locations.
3. The need for the assessment of the efficiency/accuracy of the method.
4. Identification/recommendation of improvements to existing method or alternative methods.

1.4 References

AFASA. 2010. Abalone Sector Report. Available from: http://www.aasa-aqua.co.za/site/wp-content/uploads/2010/12/Abalone-Sector-Report-August-2010-_1_1.pdf. [Accessed 24/10/12].

Basiao, Z.U., Doyle, R.W., & Arago, A.L. 1996. A statistical power analysis of the 'internal reference' technique for comparing growth and growth dispensation of tilapia strains. *Journal of Fish Biology*. **49**: 277-286.

Basiao, Z.U. & Doyle, R.W. 1990. Use of internal reference population for growth rate comparison of tilapia strains. I. In a crowded environment. The Second Asian Fisheries Forum (Hirano, R. & Hanyu, I., eds) Manila. pp. 503-506.

Cook, P. 1998. The current status of abalone farming in South Africa. *Journal of Shellfish Research*. **17**(3): 601-602.

Cook, P.A., & Gordon, H.R. 2010. World abalone supply, markets, and pricing. *Journal of Shellfish Research*. **29**(3): 569-571.

Cochran, W.G., & Cox, G.M. 1957. Experimental Designs. New York: Wiley.

Doyle, R.W., Field, C.A., & Basiao, Z.U. 1990. A Statistical procedure for using 'reference fish' to compare the growth of genetic strains in aquaculture. The Second Asian Fisheries Forum (Hirano, R. & Hanyu, I., eds), 499-502. Manila: Asian Fisheries Society.

Eknath, A.E., Bentsen, H.B., Gjerde, B., Tayamen, M.M., Abella, T.A., Gjerdem T., & Pullin, R.S.V. 1991. Approaches to national fish breeding programs: pointers from a tilapia pilot study. NAGA, The ICLARM Quarterly, April: 10-12.

FAO, 2010. The State of World Fisheries and Aquaculture 2010. FAO Fisheries and Aquaculture Department, Food and Agriculture Organization of the United Nations. Rome.

FAOSTAT, 2012. Fisheries and Aquaculture Department: *Global Aquaculture Production* (Online Query). http://www.fao.org/fishery/statistics/en. [Accessed January 2013].

Fairweather, P.G. 1991. Statistical power and design for environmental monitoring. *Australian Journal of Marine and Freshwater Research*. **42**(5): 555-567.

Fisher, R.A. 1925. Statistical methods for research workers. Oliver and Boyd Ltd., Edinburgh. 1st edition.

Flore-Aguilar, R. A., Gutierrez, A., Ellwagner, A., & Searcy-Bernal, R. 2007. Development and current status of abalone aquaculture in Chile. *Journal of Shellfish Research*. **26**: 705-711.

Geiger, D.L.,& Owen, B. 2012. *Abalone - Worldwide Haliotidae.* Hackenheim: ConchBooks. pp 361.

Genade, A.B., Hirst, A.L. & Smit, C.J. 1988. Observations on the spawning, development and rearing of the South African abalone *Haliotis midae* Linn. *South African Journal of Marine Science*. **6**: 3-12.

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

Hill, W.G.1972a. Estimation of genetic change, I. General theory and design of control populations. *Animal Breeding Abstracts*. **40**: 1-15.

Kirkendale, L., Robertson-Anderson, D.V. & Winberg, P.C. 2010. Review on the use and production of algae and manufactured diets as feed for sea-based abalone aquaculture in Victoria. University of Wollongong, Shoalhaven Marine & Freshwater Centre, Nowra, for the Department of Primary Industries, Fisheries Victoria, 198.

Kirpichnikov, V.S., 1966. Methods of progeny testing spawners in carp fish farms. *Selective Breeding of Carp. Israel Program for Scientific Translations*. **5737**: 36– 55.

Ling, E.N.,& Cotter, D. 2003. Statistical power in comparative aquaculture studies. *Aquaculture*. **224**: 159-168.

Moav, R. & Wohlfarth, G. 1974. Magnification through competition of genetic differences in yield in carp. *Heredity*. **33**: 181-202.

Moav, R., Brody, T., Wohlfarth, G., & Hulata, G. 1976. Applications of electrophoretic genetic markers to fish breeding. I. advantages and methods. *Aquaculture.* **9**: 217-228.

Muir, W.M. 1986. Estimation of response to selection and utilization of control populations for additional information and accuracy. *Biometrics*. **42**: 381-391.

Nie, G.H., & Wang, S. 2004. The status of abalone culture in China. *Journal of Shellfish Research*. **23**: 941-945.

Raemaekers, S., Hauck, M., Bürgener, M., Mackenzie, A., Maharaj, G., Plagányi, É.E. & Britz, P.J. 2011. Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. *Ocean & Coastal Management*. **54**: 413-445.

Sales, J., & Britz, P.J. 2001. Research on abalone (Haliotis midae L.) cultivation in South Africa. *Aquaculture Research*. **32**: 863-874.

Seary-Bernal, R. 1994. Statistical power and aquaculture research. *Aquaculture*. **127**: 371-388.

Troell, M., Robertson-Anderson, D., Anderson, R.J., Bolton, J.J., Maneveldt, G., Halling, C. and Probyn, T., 2006. Abalone farming in South Africa: An overview with perspectives on kelp resources, abalone feed, potential for on-farm seaweed production and socio-economic importance. *Aquaculture*, **257**: 266-281.

Van der Merwe, A.E. 2009. Population genetic structure and demographical history of South African abalone, *Haliotis midae*, in a conservation context. Ph.D Thesis. Stellenbosch University, South Africa.

Chapter 2

LITERATURE REVIEW

From the inception of agricultural sciences, once the need for replication due to variability in response of experiments was recognised and later the expense of adequately doing so (Fisher, 1925), reference groups have been implemented to supplement replication and the cost thereof (Yates, 1936). Over the last century the development of experimental designs in agricultural experiments, particularly those of selection and breeding experiments, has gone hand in hand with that of reference group methodology. The first experimental designs which implemented reference group methodologies where developed in the early 20th century for use in field trials of crops (Yates, 1936; Cochran, 1937), shortly thereafter it was adapted to suit the needs of conventional livestock breeding such as poultry and dairy (Gowe *et al*., 1959; Dickerson, 1969) and later to that of the relatively recent aquaculture breeding (Kirpichnikov, 1966; Doyle *et al*., 1990).

The term “reference group” is somewhat ambiguous, in the context of selection experiments it encompasses all forms of control population, strain, variety, line or families as well as test groups replicated for the purpose of estimating environmental effects between locations, as in check varieties, reference sires and internal reference groups as discussed below. Reference groups form an integral part of statistical control methodology, wherein they act as a known standard or zero treatment against which to evaluate treatment effects (Fredeen, 1986).

In the context of selection and breeding programmes the phenotypic performance of an individual (response), is a combination of the genetic effects and environmental effects. However the term environmental effects can be misleading, as it is all “non-genetic” effects contributing to experimental error, not necessarily being climatic in nature (Falconer & MacKay, 1996). As phenotypic performance is the response variable and genetic effects are underlying latent variable of interest in breeding, some form of statistical control is required to estimate the genetic effects free of the environmental effects (Gall *et al.,* 1993). A reference group replicated spatially at all locations (between location effects), throughout a location (within location effects) or temporally through generations can be utilised for the purpose of estimating and reducing environmental effects (Hill, 1972a).

An important distinction is made between measuring relative genetic differences and absolute genetic differences. The evaluation of relative genetic differences is carried out by comparing the phenotypic performance criteria of two or more populations to determine their relative ranking (Fredeen, 1986). In these instances specialised control populations are not required as any of the populations under evaluation can serve as a reference group against

which to evaluate all other populations (Fredeen, 1986). Resources are typically then used to maintain uniform environmental conditions for all populations to better facilitate contemporaneous comparisons. In the evaluation of absolute genetic differences some means of statistical control is required to separate the environmental and genetic components of the differences in phenotypic performances between test populations (Gall *et al.,* 1993). In essence comparisons must be made with a genetically stable control population, structured to minimise inbreeding and genetic drift over generations (Hill, 1972a). The trend in performance of the control population can be taken as a measure of environmental trend through time, the deviation thereof from the selected populations is a measure of the absolute genetic component (Fredeen, 1986). In effect, the same genetic material is repeated in successive generations.

According to Fredeen (1986) the choice of a control or reference method in selection experiments is dependent on:

1. The specific objective of the experiment in terms of evaluating absolute genetic differences or relative genetic differences between test groups.
2. The feasibility of controlling all non-genetic factors through housing, temperature, lighting, nutrition etc.
3. The cost of using various control methods vs. the relative efficiency of control methods, this is dependent on aspects such as scale of operations, limitations on resources, availability of replicates etc.
4. Biological attributes of the species under selection such as fecundity, generation interval, controlled reproduction etc.

2.1 Reference Groups in Plant Breeding

The need for reference groups in plant breeding stems from the restraints imposed by the high environmental variability in plant performances between plots due to factors such as soil heterogeneity, fertility, gradient etc. and the screening of vast numbers of entries to compensate for the multiple number of generations of testing required to ensure line stability (Poehlman & Sleper, 1996). In many plant breeding programs, hybridisation through crossing of known varieties or varieties introduced from foreign sources such as germplasm collections, are utilised to combine in a single genotype, the cumulated desirable genes found in two or more difference parental genotypes (Lin & Poushinsky, 1983). In both cross-fertilising and self-fertilising plant species the result is numerous hybrid offspring with a wide assortment of the parental gene combinations. More so for self-fertilising species when the pedigree method is implemented, the F_2 generation can be in the order of 10^6 individuals (Poehlman & Sleper, 1996). In these trials the early generations are completely unreplicated and a proportion (often 20-30%) of the best performing test varieties are selected and their

offspring evaluated more precisely by increased replication and selection. Evaluation continues in subsequent generations with ever greater precision attributed to ever greater replication until a few varieties remain to be tested in numerous environments (Martin *et al*., 2006).

To conduct field experiments in the early generations of plant screening trials on such numerous test varieties is extremely difficult, as limitations on test capacity and acquiring enough seed material for replication is not possible (Kempton, 1984). Thus designs for varietal trials which can accommodate large numbers of test varieties, but are reliant on replication, for example incomplete block alpha designs (Patterson et al., 1978), lattice, lattice square (Yates, 1936), quasi-factorial designs (Yates, 1937) and chain block designs (Mandel, 1954) cannot be implemented. These early generation plant screening trials deviate from classical experiments as defined by Fisher (1926) where the aim is precise estimation of experimental error through the use of randomisation, replication and independence of observations, to facilitate accurate treatment comparisons. Some means of statistical control is required to adjust for the environmental variability which would otherwise be accounted for by replication.

As a means to this end reference group methodology has merit, as the reference group is a means to afford adequate statistical control over environmental variability while facilitating the evaluation of vast numbers of test entries which could not be replicated (Kempton, 1984). Presently there remain two conflicting schools of thought as to how best deploy reference entries amongst test entries (Edmondson, 2005). Systematic deployment of reference entries at regular intervals made the most intuitive sense and devotes fewer resources to the growth of reference entries (Kempton, 1984). However, many have questioned the validity of systematic designs as they fail to meet the requirements of least square methods that residual errors are independent and normally distributed, the reason for this is that systematic arrangements result in correlations between neighbouring plots (Yates, 1939). In addition they have been shown to be inefficient in accounting for spatial variability (Baker & McKenzie, 1969). The random deployment of reference entries meets the requirements of least square methods and limits the introduction of human bias. However, a vastly greater number of reference entries are required (approximately 50%) and in many instances is too cumbersome to implement efficiently (Yates, 1939).

2.1.1 Systematic Distribution of Reference Groups

The inclusion of replicated reference groups in a systematic manner enables the less precise estimation of the phenotypic performances in a large and genetically diverse population as opposed to the precise estimation of the phenotypic performances in a small genetically less diverse population (Kempton, 1984). In this manner helping to better achieve the specific

aim of the selection experiment namely, genetic gain (Yates, 1940). The optimal proportion of plots allocated to reference groups was found to be 20% (Kempton, 1984), while this decreases the selection intensity and efficiency of the design, this can be compensated for by the reduction in experimental error variance and subsequent increase in heritability (Chandra, 1994). The systematic allocation of reference varieties to the field design can be in each row or column for one-way elimination of heterogeneity or in diagonals in a chequer board design for two-way elimination of heterogeneity as seen in the figure 1 below (Kempton & Talbot, 1988). Numerous methods for adjustment of test varietal performances utilising reference varieties exist depending on the systematic layout implemented, for example in figure 1 (a) below, adjustments can be made using the mean of all reference varieties in the row or the two reference varieties on either side or the weighted means of reference varieties inversely proportional to the distance from the test plot (Kempton & Talbot, 1988).

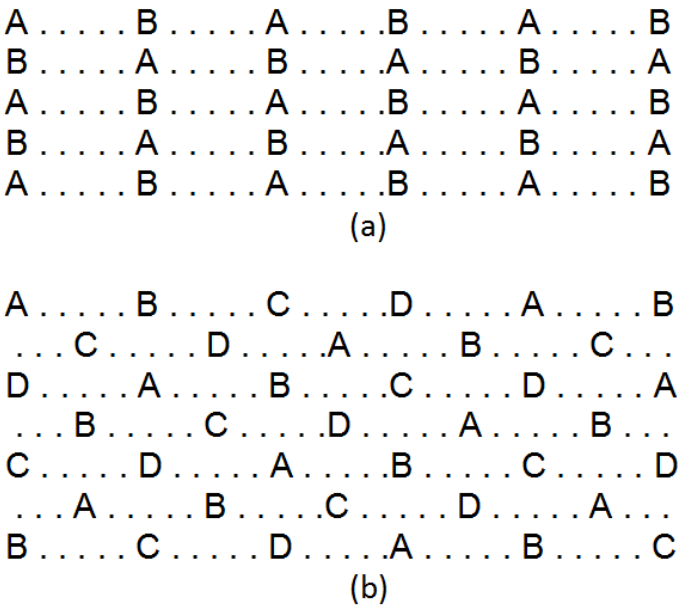

Figure 2.1 Arrangement of reference varieties among unreplicated test varieties in early generation trials: (a) systematic layout of two reference groups A and B suitable for one-dimension fertility adjustment; (b) systematic layout suitable for two-dimensional adjustment. (Adapted from Kempton & Talbot, 1988).

2.1.2 Random Distribution of Reference Groups

As opposed to the systematic arrangement of reference varieties Federer (1956) proposed a series of designs deemed ‘augmented designs’, which incorporate reference varieties randomly over the field experiment. In this method a number of reference varieties with adequate numbers for replication are selected. The reference varieties are included within blocks in any standard blocking design and then ‘augmented’ with unreplicated test varieties.

The block effects and experimental error are estimated with respect to the reference varieties and then utilised to correct the test varietal performances and for comparison of test varieties, respectively (Lin & Poushinsky, 1983). Augmented designs have the disadvantage that as much as 50% of test space is allocated to reference varieties and so their application is more limited than systematic use of reference groups in early generation plant breeding trials. An advantage however, is the increase precision with which unreplicated test variety comparisons are made due to the flexibility with which standard designs can be implemented (Sharma, 1998). Augmented designs can be used to adjust for one dimension heterogeneity by incorporating the reference groups into a randomised complete block design, linked block design or incomplete block design (Federer & Raghavarao, 1975). In addition, augmented designs allow for the adjustment of two dimension heterogeneity by incorporating reference groups into Latin square designs, Youden square designs, row column designs or modified augmented designs (Federer & Raghavarao, 1975; Federer *et al*,. 1975; Lin & Poushinsky, 1983).

An alternative to the use of control varieties as replicated reference groups is to include partial replicates of some of the test varieties as reference group to estimate experimental error. This type of design called partially replicated designs (p-rep), are particularly efficient when trials are conducted at multiple locations (Smith *et al*, 2006; Cullis *et al*, 2006). In a partially replicated design over *l* locations a proportion of test varieties (1/*l*) are replicated such that each test variety is replicated twice at each location. The advantage of using test varieties as the reference group is the increase in genetic gain (Cullis et al, 2006). Building on this design Williams *et al*., (2011) have consolidated augmented designs and p-rep designs into a single design. These augmented p-rep designs, make use of replicating a portion of test varieties twice and then determine a resolvable incomplete block design with the two replicates. The blocks in the design are then augmented with the remaining test varieties. This method combines the increase precision of augmented block designs with the increased efficiency and genetic gain of p-rep designs (Williams *et al*., 2011).

2.1.3 Temporal Distribution of Reference Groups

As opposed to the less precise estimation of spatial variation of test plots during the trials, a method to precisely estimate the spatial variation of test plots before the trial was developed by Cochran (1937) utilising information from uniformity trial data. Within this line of thinking a field is divided into plots of equal dimensions and a reference group in the form of a single variety is grown on all plots. In this instance phenotypic difference between plots can be attributed directly to environmental effects such as soil heterogeneity, gradient, differences in fertility etc. and a wealth of information can be derived from such a trial (Cochran, 1937).

The environmental effects attributed to plots can be utilised to amalgamate similar plots and be used to determine the optimum size and shape of plots (Smith, 1938; Zhang *et al*, 1994). In addition, similar adjacent plots can be grouped into blocks to further improve accuracy of future experiments conducted on the same field (Cochran, 1937). Yates (1936) also made use of uniformity trials to test efficiency of experimental designs and uniformity trial data is still used in this capacity in present times (Idrees *et al*., 2009). If the variety used in the preliminary trial is genetically uniform for instance a cross of two inbred lines, it can be seen in the equation below that any phenotypic variance between plots can be attributed to environmental effect alone since:

$$V_{Gu} = 0$$
$$V_{Pu} = V_{Eu}$$
$$V_{Gs} = V_{ps} - V_{Eu}$$

Where V is the variance and P, G and E denote the phenotypic, genotypic and environmental variance and subscripts u and s represent the uniformity trial and selection trial respectively (Kearsey & Pooni. 1996).

It was through the development of analysis of covariance (ANCOVA) (Fisher, 1925) that the correlations between uniformity trial data and selection trial data, that the phenotypic performances can be adjusted to remove environmental variance (Cochran, 1957). It was observed that precision of experiments analysed in this manner showed a dramatic increase often in the order of double, however, the generation interval for experiments and the accompanying labour also doubled (Cochran, 1937). In the instances of perennial plants where the experimental plot is a single tree or when a crop rotation method is implemented with annual crops this method is remarkably efficient (Cochran, 1937). A limitation of this method is that the temporal variation between years is not accounted for and can still have a sizeable effect. The typical statistical model for an analysis of covariance is:

$$y_{ij} = \mu + \tau i + \rho i + \beta(x_{ij} - x..) + \varepsilon ij$$

The γ_{ij} is the phenotypic performance of test individuals of the i-th treatment in the j-th replicate, while x_{ij} is the phenotypic performance of the concomitant covariate or in this instance the uniformity trial performance, on which γ_{ij} has a linear regression coefficient β. The constants μ, τ_i, and ρ_i are the true mean response and the effects of the i-th treatment and the j-th replication, respectively. The residual ε_{ij} is the residual error assumed to be normally distributed with homogeneity of variance and independence of observations (Cochran, 1957).

2.2 Reference Groups in Animal Breeding

The use of reference groups in animal breeding developed from that of plant breeding, primarily to address the temporal environmental variability between generations and the vastly different spatial variation where large numbers of breeders in vastly different environments are members of the same breeding programs. The utilisation of reference groups in multi-generation selection experiments, facilitate accurate estimation of response to selection per generation, which is made difficult by the confounding effects of improvements in animal husbandry through time (Dickerson, 1969; Hill, 1972a; Fredeen, 1986; Gall *et al.,* 1993). Typically there are two ways a breeder can direct populations towards desired improvement of phenotypic performance. Either the manipulation of genotypic frequencies using inbreeding or outbreeding to influence allelic distribution within individuals or to manipulate the allele frequencies towards increasing desirable alleles at the expense of less favourable alleles, through selection and migration (Gall *et al.,* 1993). The results of the latter method of improvement are phenotypic performances that consequently reflect in part, a genetic component and an environmental component (Falconer, 1952). Breeders are interested primarily in the estimation of the change in the genetic component to make inferences about the efficiency of the response to selection.

Comparing the expected and the realised response often proves informative in evaluating the environmental and genetic trends during selection. The expected or predicted response to selection per year is calculated as a function of the product of the heritability of the phenotype, the variability of the phenotype and the selection intensity applied to each generation, divided by the number of years required to complete the generation interval (Falconer & MacKay, 1996). The observed or realised response to selection per year will only correspond to the predicted response per year when the environmental component is identical in parent and offspring generations (Gall *et al.,* 1993).

In practice it is not possible to maintain uniform environmental conditions over multiple generations. In addition to temporal variation in environment over generations, improvements in animal husbandry practices, management procedures, nutrition etc. can result in an increasing positive environmental trend which masks genetic improvements (Hill, 1972a). From this there follows two main methods of estimating the genetic trend free of environmental effects. The most common is the use of reference groups in the form of replicated genetic material in successive generations to account for environmental trend and estimate the genetic trend as the deviation of the selected population from the replicated genetic material (Hill, 1972a; Fredeen, 1986). The second method is to estimate the genetic trend using mixed model procedures (Henderson, 1975) to estimate breeding values of all individuals based on pedigree and performance data and the deviation of mean breeding

values within each generation. The environmental trend is then simply the deviation of observed phenotypic changes from the calculated genetic trend (Gall *et al.,* 1993).

There are four approaches of utilising a reference group as replicated genetic material in estimating genetic change. Traditionally, the most common approach was to utilise an unselected control population to estimate environmental changes and to a limited extent the effects of inbreeding (Hill, 1972a). Another approach is to make use of divergent selection and utilise the divergent population in the opposing direction to that of the breeding goal as a reference group to adjust for asymmetry in response (Hill, 1972b). A third approach is to utilise a contemporary population as a reference group, through repeat mattings, genotype storage or gamete storage of parental populations before selection and directly compare performances (Goodwin *et al*., 1960; Bray *et al*., 1962; Dickerson, 1969). The most common approach in recent years is the use of mixed models to estimate genetic change, strictly speaking this is not a reference method or a least squares method, however in light of recent studies utilising mixed models in conjunction with reference groups this method warrants explanation (Sorenson *et al*., 2003).

2.2.1 Unselected control population

An unselected control population or segregating control population is a group of individuals randomly sampled from the same base population as that of the population undergoing selection (Hill, 1972a). The emphasis in the selected population is to increase the allele frequencies of beneficial alleles to the detriment of poor alleles, whereas the emphasis in the unselected control population is in maintaining the original allelic and genotypic frequencies in the base population (Gall *et al.,* 1993). In effect the genetic material in the base population is replicated in successive generations to keep genetic effects constant in order to estimate environmental changes (Fredeen, 1986).

Constraints in this method are that the effect population size must be large in comparison to the number of generations (Hill, 1972a) in order to prevent random genetic drift and inbreeding from confounding future performances of the control population. In addition the control population must not undergo any form of selection, to minimise this the replacement of breeding individuals must be done at random, the individuals selected must also not contribute to inbreeding and must deviate from the population mean the least (Fredeen, 1986). An additional source of error to be avoided is the sampling variance when recording the performance of the control population (Hill, 1972a). Lastly, the use of this method over a high number of generations can result in sufficient divergence between control and selected populations for variable responses to the environment i.e. genotype by environment interactions in the selected population (Hill, 1972a). Granted the control population and the selected population are reared in the same environment and the above

constraints are adhered to and no genotype by environment interactions, the only sources of variation in the performance of the control population in future generations are due to environmental trend (Gall *et al.,* 1993). The genetic trend for the selected population can be estimated as follows:

$$\Delta G = P_s - P_c = G_s - G_c + E_s - E_c$$

Where P, G and E are the respective phenotypic, genotypic and environmental values. The subscripts S and C denote the selected population and the control population, respectively (Gall *et al.,* 1993).

This form of reference group has successfully been implemented in selection experiments in laboratory animals such as the common fruit fly *Drosophila melanogaster* (Clayton *et al*., 1957; Frankham *et al*., 1957), the red flour beetle *Tribolium castaneum* (Bray *et al*., 1962) and laboratory mice *Mus musculus* (Esien *et al*., 1970). As well as commercial livestock species such as beef cattle (Newman *et al*., 1973), swine (Mikami *et al*., 1977) and poultry (Gowe *et al*., 1959). The use of unselected control populations are by no means limited to the above mentioned examples and species, rather for the sake of brevity and relevance, the examples above are included to provide scope of diversity and age of implementation of the method, for a more comprehensive review see Hill (1972c) and Fredeen (1986).

The use of an unselected control population in commercial aquaculture species is more prevalent, as can be seen by its application to numerous fish species such as tilapia (*Oreochromis spp*), salmonids (*Oncorhynchus spp*), common carp (*Cyprinus carpio*) and channel catfish (*Ictalurus punctatus*) (Bolivar & Newkirk, 2002; Brink *et al*., 2002; Maluwa & Gjerde, 2007; Kincaid, 1979; Hershberger *et al*, 1990; Brink, 2004; Wohlfarth *et al*., 1975; Wohlfarth & Moav, 1991; Ninh, 2009; Dunham, & Brummett, 1999). Similarly, numerous shellfish species selection programmes utilise an unselected control population in this manner, such as in oysters (*Ostrea* and *Saccoostrea spp*), freshwater prawn (*Macrobrachium rosenbergii)* and redclaw crayfish (*Cherax quadricarinatus*) (Newkirk & Haley, 1983; Nell *et al*., 1999; Malecha, 1980; McPhee & Jones, 1997).

2.2.2 Divergent Selection

In estimating the response to selection in a divergent selection experiment, selected populations act as concomitant reference groups for each other. Similarly, in divergent selection as in unselected control populations, both populations must ideally be sampled from the same original base population (Rye & Gjedrem, 2005). Individuals in each population are selected simultaneously for increased and decreased phenotypic merit, respectively.

Limitations associated with this method are in the assumption that the magnitude of genetic change will be equal in both directions, for many traits in animal breeding this has been observed to not be the case (Falconer & Mackay, 1996). Another limitation is the efficiency and cost, it is unlikely that selection in both directions will yield profitable phenotypes and the additional resources required for testing both selection groups decreases the relative efficiency (Rye & Gjedrem, 2005). However, this form of selection can facilitate the identification of markers associated with increased and decreased phenotypic performance and thus has application in quantitative trait loci mapping (Ollivier *et al*., 2007). Furthermore, after numerous generations of selection, substantial genetic changes increase the probability of genotype by environment interactions which will bias the method of estimation (Fredeen, 1986).

If both populations are subjected to identical rearing environment within each generation, the response to selection can be estimated as half the deviation of respective phenotypic performances as follows:

$$\Delta P_i = \Delta G_i + E$$

$$\Delta P_d = \Delta G_d + E$$

$$\Delta P_i - \Delta P_d = \Delta G_i + E - (-\Delta G_d + E) = 2\Delta G$$

Where, P, G and E are the respective phenotypic, genotypic and environmental values. The subscripts i and d denote the selected increase and decrease populations, respectively (Gall *et al.,* 1993).

Divergent selection has been implemented in poultry (Liu *et a*l., 1994; Terčič & Holcman, 2008), in beef cattle (Davis, 1987), in pigs (Cameron & Curran, 1995) and in sheep (Cloete *et al*., 2005). Interestingly, the use of a divergent selection line as a reference group is also widely used in the commercial selection programmes of numerous fish species such as tilapia (Oreochromis spp), channel catfish (*Ictalurus punctatus)* and common carp *(Cyprinus carpio)* (Bondari *et al*.,1983; Abucayo & Mair, 2000; Bondari, K. 1983; Vandeputte *et al*., 2002). As well as commercial shell fish species such as oysters (*Ostrea spp*) and kuruma prawn (Toro & Newkirk, 1990; Toro & Newkirk, 1991; Hetzel *et al.,* 2000).

2.2.3 Contemporary Populations

The methods of repeat matting, genotype storage and gamete storage utilise the central theme of temporal comparison of genetically equal individuals as reference groups in the same environment to eliminate environmental differences between generations (Gall *et al.,* 1993). As opposed to the above mentioned approaches, this method does not require resources invested in maintaining genetically controlled lines throughout the selection experiment. Rather the genetic material can be repeated over specific generations through

the use of repeat mattings or artificial insemination for example using frozen semen (Goodwin *et al*, 1955).

In the instance of repeat mattings, contemporary offspring are produced in successive parities of the same parental individuals (Fredeen, 1986). The performances of the offspring represent the environmental effects and the least squares regression of performances on generation number provides the environmental trend. Alternatively, testing the offspring from parents of the initial and final generations in the same environment will also provide a measure of genetic trend (Gall *et al.,* 1993). A similar approach adopted by the beef and dairy cattle industries is the use of "reference sires" in estimating the breeding values of sires (Cundiff *et al*., 1975). The method is based on the limitations on testing progeny in different herds, as differences between herds are due to genetic and environmental effects so comparison of progeny performances in separate herds cannot be compared unless a standard is present in all herds. Utilising the progeny of reference sires in all herds provides the means for a standard to partition phenotypic response into environmental and genetic components.

The use of repeat mattings has the advantages of a high degree of genetic relationship in herds or sub-populations which minimises the probability of confounding genetic and inbreeding affects and also the contributions of genotype by environment interactions (Fredeen, 1986). Limitations to the use of repeat mattings are short term, usually limited to the life span of the parents used for repeat mattings. In addition, the mortality of parents and offspring used in this method increase the sample bias (Fredeen, 1986). The traits under question need to be devoid of maternal effects which bias performances of different parities and unfortunately most commercial production traits are not entirely devoid of maternal influences (Goodwin *et al*., 1960).

The advent of improvements of cryopreservation of gametes or whole genotypes has increased the feasibility of contemporary populations in estimating genetic trends. For the most part the ability to store a replicate whole genotypes in successive generations was exemplified in plant breeding where storage of seed or vegetative reproduction where utilised to extend comparisons over a longer number of generations than repeat matings. Initially the only organisms in the kingdom *Animalia* to be cold stored as whole genotypes where insects such the model laboratory organisms for example *D. melanogaster* and *T. castaneum* (Hill, 1972a), later this was extended to commercial insects such as the Western Honey bee *Apis mellifera* (Leopold, 2007). In commercial livestock species the cryopreservation of gametes such as spermatozoa is becoming more common place especially in the beef and dairy industries (Gall *et al.,* 1993).

Storage of spermatozoa taken from sires at frequent generation intervals throughout the selection experiment can more effectively be utilised in later generations with newly

selected females (Fredeen, 1986). The current sires and spermatozoa from past sires can either be combined through mating and artificial insemination with random samples of females to produce unrelated progeny groups or with the same sample of females to produce half-sib groups. The phenotypic performances of the offspring of current sires and past sires can be used to estimate the response to selection as follows:

$$\Delta P_c - \Delta P_p = G_c + G_p + E_c - E_p = \Delta G$$

Where, P, G and E are the respective phenotypic, genotypic and environmental values. The subscripts c and p denote the progeny of current and past sires, respectively (Gall *et al.,* 1993).

Since the environmental values in the equation above will be equal the response to selection per generation can be calculated by adjusting for generation number between current and past sires. A limitation associated with the use of gamete storage for contemporary control populations is the bias from decreased viability of cryopreserved spermatozoa. In addition, the likelihood that sires will not be chosen at random for cryopreservation of spermatozoa is great, thus the probability that offspring of past sires will likely not be as inbred as current sires can confound inbreeding effects (Fredeen, 1986). Lastly if unequal numbers of progeny from sires are tested a bias in estimation will occur complex statistical analyses will be required to adjust for this likely occurrence. In the event that current sires are descendants of past sires mixed model analyses as discussed below are required to account for the genetic relationships (Gall *et al.,* 1993).

The use of contemporary populations in conventional livestock species is largely limited to application in dairy cattle (Philipson *et al*., 1994; Fredeen, 1986), this method has also been implemented in fish species such as Nile tilapia (*Oreochromis niloticus*) (Khaw *et al*., 2008) and Rainbow trout (*Oncorhynchus mykiss*)(Hörstgen-Schwark,1993) albeit to a lesser extent than the previous approaches mentioned.

2.2.4 Mixed Models with Reference Groups

Mixed models are an extremely powerful and versatile tool developed in conjunction with the relatively recent advances in statistical theory accompanied by advances in computation power (Cameron, 1997). Mixed models provide a means of predicting breeding values and estimation environmental trends in the absence of an unselected control population (Gall *et al.,* 1993). The optimal use of mixed model analyses requires a form of genetic relationship between individuals, individual identification and pedigree information, knowledge of phenotypic and genetic parameters as well as relative completeness of data for traits in question (Henderson, 1953, 1975, 1988). Many fixed effects such as year of measurement, management system, age of dam etc. can be removed directly, greatly reducing the

influence of environmental effects which would otherwise not be accounted for in standard least square regression methods (Gall *et al*., 1993; Cameron, 1997).

In selection programmes where there is a high degree of confounding between environmental and genetic effects and a full pedigree record is available, it is possible to simultaneously utilise the performance and degree of relatedness between individuals to predict breeding values of individuals and estimate environmental effects. The procedure for simultaneously predicting these effects is called Best Linear Unbiased Prediction (BLUP) (Henderson, 1953). The pedigree information enables the breeding values of all individuals in all generations to be estimated by use of Wright's numerator relationship matrix, which makes us of Wright's relationship coefficient for all pairs of individuals (Gall *et al.,* 1993). There are numerous methods to utilise BLUP depending on the specific type of records and degree or relatedness between individuals. For instance sire models are to predict sire breeding values from progeny performances (Henderson, 1973). The repeatability model for predicting breeding values of individuals with repeated performances and the animal model to predict the breeding values of all individual in the pedigree (Henderson, 1975). The animal model will can be seen as follows:

$$y = Xb + Zu + e$$

Where y is the vector of observations, X is the incidence matrix of the fixed effects; Z is the incidence matrix of the additive genetic effects. The symbol u is the vector for additive genetic effects, b is the vector for fixed effects and e is the vector of residuals (Cameron, 1993).

As multivariate vectors u and e are assumed prior to selection to have null means and the following variances:

$$Var\begin{bmatrix} u \\ e \\ y \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & AZ'\sigma_a^2 \\ 0 & R & R \\ ZA\sigma_a^2 & R & V\sigma_e^2 \end{bmatrix}$$

Where A is the relationship matrix, σ^2_a is the additive genetic variance, σ^2_e is the residual variance, $R=I\sigma^2_e$ and V= (I + ZAZ' r) for $r = \sigma^2_a/ \sigma^2_e$, $h^2 = r/(1+r)$ (Cameron, 1933; Gall *et al.,* 1993). The variance ratios in this model are either estimated for each dataset or adapted from previous studies. The mixed model equations solved to predict individual breeding values are as follows:

$$\begin{bmatrix} X'X & X'Z \\ ZiX & Z'Z + A^{-1}\gamma \end{bmatrix} \begin{bmatrix} u \\ b \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Where γ = 1/r, the response to selection is estimated as the within generation change in of the average breeding values for all individuals. The regression of the average breeding values of all individuals on generation number provides the rate of genetic change per

generation (Gall *et al.,* 1993). The above explanation is simplified for the purpose of explanation within the present study, for more in-depth reviews on the topic see Cameron (1993), Meyer and Hill (1991) and Mrode (1996).

Limitations in the use of mixed model analyses are the need for pedigree records, a degree of genetic relatedness between individuals and completeness of data (Gall *et al.,* 1993). For the most part this information is readily available in conventional livestock species but in the instances of selection programmes where the base population is sampled from wild populations this information is not available until a few generations of breeding a record keeping has taken place.

The relevance of mixed model analyses in the context of reference groups is outline by a more recent study by Sorenson *et al*, (2003), where situational studies were conducted to elucidate the value of control populations to estimate environmental trend in selection programs. It was found that the inclusion of performances of control populations in the dataset where an environmental trend was model as a fixed effect the sampling variances were around four times larger for the dataset with only the selected population than the dataset when a control population was included. However, when no environmental trend is observed and the authors specify this maybe the case if previously determined by another study, that there was no improvement in accuracy attributable to a control population.

The use of a reference group in conjunction with mixed model analyses at estimating genetic change has been implemented in sheep (Thompson & Atkins, 1994) as well as commercial aquaculture fish species such as tilapia (*Oreochromis spp*) (Ponzoni *et al*., 2005; Gall & Bakar, 2002) and rainbow trout (*Oncorhynchus mykiss*) (Su *et al*., 1997).

2.3 Reference Groups in Fish and Shellfish Breeding

The use of reference groups in the selection and breeding of aquatic species for the most part is the same as those mentioned above for animal breeding (Rye & Gjedrem, 2005). There is however specific systematic and biological constraints in aquaculture selection programmes which require novel methods of estimating the genetic components of phenotypic performances.

The numbers of rearing structures available for genetic testing are often limited, especially in developing countries, these limitations on testing space can be found through-out the production cycle from spawning, nursing, weaning to the final grow-out/testing of future brood stocks (Moav *et al*., 1976; Doyle *et al*., 1990; Rye & Gjedrem, 2005). A rearing structure in the context of aquaculture can be tanks, race-ways, grow-out baskets, ponds, net pens and aquaria. In the absence of a physical identification method such as tagging or the capacity for economically feasible genetic markers, individual identification is not possible. In the context of selection programs this means individual families or strains must

be reared separately throughout the entire life cycle which further makes inefficient use of already limited resources (Moav *et al*., 1976).

In aquaculture experiments, a large number of random environmental variables are present, which are either too costly (water temperature, water quality, oxygen content etc.) or too impractical (stocking densities, biomass of plankton, algae and benthic organisms, number of predators, photoperiod etc.) to control (Moav *et al*., 1976; Basiao & Doyle, 1990; Kocour *et al*., 2005). Resultantly, the inter-rearing structure variability is high (Buck *et al*., 1970; Uraiwan & Doyle, 1986). In the cases where individual families or strains must be reared separately throughout the selection experiment, the inter-rearing structure differences are confounded with the genetic differences of the individual families or strains in question (Moav *et al*., 1976). The major limitation from a statistical perspective is obtaining adequate replication with limited resources for the removal of random environmental variation from the measured phenotypic performances of each test family or strain (Doyle *et al*., 1990).

A further limiting factor for replication even when testing space is not limited, is obtaining enough offspring from each controlled spawning to facilitate replication and still maintain commercial stocking densities. As it is important to test the performances of families or strains under the commercial conditions in which they are expected to perform, it is necessary to maintain commercial stocking densities (Rye & Gjedrem, 2005). Even though most aquaculture species are highly fecund, most brood stock sampled from the wild will not readily spawn in captivity, in addition the technologies for new aquaculture species have to be developed (Zohar, 1989; Mylonas *et al*., 2010).

2.3.1 Internal Reference Technique

A potential solution to the statistical problems created by the biological and systematic constraints above is the inclusion of an “internal reference” group into each rearing structure or experimental unit (Kirpichnikov, 1966; Moav & Wohlfarth, 1974; Doyle *et al*., 1990). The reference group which is a specific strain, line or family is distinguishable by some form of group identification such as unique phenotypes or mechanical markers such as colour tags or fin-clipping (Moav *et al*., 1976; Vandeputte *et al*., 2002). The reference group is included within all rearing structures such that the same genetic material is replicated in each experimental unit (Doyle *et al*., 1990). The reference group is an internal statistical control of the random environmental variance specific to each experimental unit, so that differences in performance between internal control groups reflects the inter-rearing structure differences attributable to variable environmental effects (Linhart *et al*., 2002; Kocour *et al*., 2005). This enables correction for environmental variance and expression of standardised performances of test families or lines and greatly improves the statistical power of analyses and reduces

the need a high number of test replicates (Moav *et al*., 1976; Basiao *et al*., 1996; Vandeputte *et al*., 2002).

Differences in the fertility of rearing structures for example where rearing conditions are particularly favourable for growth in a particular rearing structure due to a random environmental variable such as water temperature or mortality has reduced the population density, corrections can be made to the test group as the reference group and test group will mutually exhibit enhanced performance (Doyle *et al*., 1990). Corrections for environmental variation between rearing structures can be made by deviation of the reference group performance from the test group performance for each rearing unit (Kirpichnikov, 1966; Doyle *et al*., 1990; Vandeputte *et al*., 2002). The relevant equating for replicate variance after deviating reference and test group performances is as follows:

$$V_{(tg-rg)} = V_{tg} + V_{rg} - 2Cov_{(tg,rg)}$$

Where V(tg-rg) is the variance of the corrected replicate means for a single test group and Vtg and Vrg are the uncorrected variance of test group means and reference group means respectively. Cov(tg,rg) is the covariance between test and reference groups (Basiao & Doyle, 1990).

Alternatively, correction for environmental variation may be facilitated by an ANCOVA, where the performance of the reference group is utilised as a concomitant covariate (Doyle *et al*., 1990; Basiao *et al*., 1996). A typical model for this analysis is as follows:

$$y_{ij} = \mu + Si + Ei + R_{ij}(SR_{ij}) + \varepsilon ij$$

Where, γ_{ij} is the mean specific performance of the i-th test group in the j-th environment. While R_{ij} is the mean specific performance of reference group grown the with the i-th test group in the j-th environment. The grand mean of the test group is represented by μ, S_i is the i-th test group effect, E_j is the j-th environmental effect and ε_{ij} is the random error (Basiao *et al*., 1996).

According to Kirpichnikov (1966) the most important considerations in internal control methodology are that control groups must be distinguishable from the test group, the conditions of both groups prior to testing must be uniform and initial weights must be similar at the commencement of the experiment. In addition, some statistical limitations are that the test and reference group must not interact in such a manner that one group benefits to the detriment of the other (Doyle *et al*., 1990). This can have implications for management of the experiment as animals sharing a rearing structure will interact and antagonistic behaviour leading to growth dispensations are a major source of growth variation in aquatic species, the risk of this interaction can be reduced by ensuring initial stocking sizes of groups are similar, feeding is done to satiation and grading is carried out at regular intervals (Basiao &

Doyle, 1990; Basiao *et al*., 1996). Another important consideration especially when testing is carried out in multiple locations is that no genotype by environment interaction occurs as this leads to unpredictable performances of the reference group, test group or both (Basiao & Doyle, 1990; Romana-Eguia & Doyle, 2002; Romana-Eguia *et al*., 2010). To reduce the chances of genotype by environmental interaction the reference group must preferably be very similar genetically speaking to the test groups (Gall *et al.,* 1993). Additional sources of error to be avoided are random sampling error, when sampling the population of reference group animals to allocated animals to experimental units, sampling must be done randomly to prevent sample bias (Gall *et al.,* 1993).

The internal reference method has been successfully implemented to control random environmental variance in growth performance comparisons in numerous aquaculture fish species such as common carp (*Cyprinus caprio*) (Kirpichnikov, 1966; Linhart *et al*., 2002; Vandeputte *et al*., 2002; Kocour *et al*., 2003; Kocour *et al*., 2005; Buchtová *et al*., 2006), rainbow trout (*Oncohynchus mykiss*) and brown trout (*Salmo trutta*) (Blanc *et al*., 1983; Blanc et al., 2001), tilapia (*Oreochromis spp*) (Doyle *et al*., 1990; Basiao & Doyle, 1990; Basiao *et al*., 1996; Romana-Eguia & Doyle, 2002; Romana-Eguia *et al*., 2010) and tench (*Tinca tinca*) (Kocour *et al*., 2010). Interestingly the common carp experiments made use of a phenotypic variant “mirror” which is highly distinguishable from the “scally” wild type, the rainbow trout and brown trout experiments made use of the “golden” phenotypic variant which was distinguishable from the “rainbow” and “brown” respective wild type phenotypes. The tench experiment also made use of a “golden” phenotypic variance, distinguishable from the “green” wild type and only the tilapia experiments made use of different tilapia species or hybrids thereof or tagging to distinguish the reference group from the test group.

2.3.2 Communal Rearing with Multiple Nursing Technique

Another potential solution to the statistical problems created by the unique biological and systematic constrains in aquaculture selection experiments is communal rearing with multiple nursing technique (Wohlfarth & Moav, 1985). Communal testing entails stocking distinguishable different genetic groups within the same rearing structures, in a manner analogous to offspring of reference sires within different herds utilised in conventional livestock breeding (Cundiff *et al*., 1975). Effectively this method eliminates the confounding of inter-rearing structure variability from that of genetic differences between groups under evaluation and drastically decreases the number of rearing structures required (Wohlfarth & Moav, 1985). However, the implementation of this method diverges from conventional livestock selection experiments, in that aquatic species are far more sensitive to hierarchical growth structures and growth dispensations due to antagonistic interactions. Growth

dispensations are strongly influenced by the variation in initial size of test groups (Wohlfarth & Moav, 1972, 1985, 1993).

The multiple nursing technique is a method to correct for the bias caused by variation in the initial sizes of genetic groups co-stocked communally (Wohlfarth & Moav, 1985, 1993). It is pertinent to note that communal stocking does not always require the multiple nursing technique if test groups are identical in size at co-stocking, of course in most occasions this is rarely possible. Multiple nursing entails randomly splitting each genetic group during the nursing phases prior to the experimental and stocking each separately at differential stocking densities (Wohlfarth & Moav, 1985, 1993). The objective is to induce extreme differences in the initial sizes of each sample of each genetic group prior to the experiment by manipulating the density-dependence of growth rate (Wohlfarth & Moav, 1985, 1993). An environmental correction factor is calculated as the linear regression coefficient of the difference between mean growth rates of differentially stocked genetic subgroups against the difference in the initial sizes at co-stocking if genetic subgroups. As demonstrated below:

$$\beta = \frac{(Y_L + Y_S)}{(X_L + Xs)}$$

Where β is the environmentally generated coefficient of regression of growth rate differences and initial size difference, Y and X are the respective growth rates and initial sizes and subscripts L and S denoted large subgroup and small subgroup (Wohlfarth & Moav, 1985; Wohlfarth & Milstein, 1987). The correction utilising β is a simple transformation as follows:

$$Y' = Y - \beta(X - X.)$$

Where Y' is the corrected growth rate, Y is observed growth rate, β is the environmentally generated coefficient of regression of growth rate differences and initial size differences, X is the initial weight of specific genetic group and X. is the grand mean of initial weight (Wohlfarth & Moav, 1985; Wohlfarth & Milstein, 1987).

Limitations to the method are that all groups must be distinguishable and as aquatic species are highly fecund this often involves tagging or branding vast numbers of offspring (Moav & Wohlfarth, 1976). Although this method greatly reduces the number of rearing structures required for the grow-out phases, it increases the resources requirements in the pre-experiment nursing phases, for successful implementation of the multiple nursing technique (Wohlfarth & Moav, 1985). There is evidence to suggest that communal rearing with the multiple nursing technique may lead to over estimation of selection response, as was reported in a recent study with sea bass (*Dicentrarchus labrax*), the authors believed this was due to the effects of stocking density (Vandeputte *et al.*, 2009). While even more recently a study reported that communal rearing increased growth rate and decreased generation interval and decreased the common environmental component of variance while

estimating genetic parameters of common carp (Cyprinus carpio) reared communally and separately (Ninh *et al*., 2011).

Communal rearing has been extensively utilised testing genetic groups of freshwater and marine fish such as common carp (*Cyprinus carpio*) (Wohlfarth & Moav, 1985, 1991, 1993; Wohlfarth & Milstein, 1987; Ninh *et al*., 2011), salmonids (*Oncorhynchus & Salmo spp*) (Gjerde et al., 1983; Gjerde & Gjedrem, 1984; Busak & Riddel, 1985; Refstie,1990; Herschberger *et al*., 1990), Tilapia (*Oreochromis spp*) (McGinty, 1984, 1987), channel catfish (*Ictalurus spp*) (Dunham *et al*., 1982) and sea bass (*Dicentrarchus labrax)* (Vandeputte *et al*., 2009). The use of the communal rearing method has been reported on somewhat less in shellfish breeding and shellfish culture, not surprisingly limited to the motile species with increased interaction, such as freshwater prawns (*Macrobrachium rosenbergii*) (Karplus *et al*., 1989; Karaket *et al*., 2005) and Atlantic lobsters (*Homarus spp)* (Jørstad *et al.,* 2005). It is pertinent to note that for the most part, the investigations above opted to synchronise spawning and ensure initial sizes where the same at co-stocking and thus were not always accompanied by the multiple nursing technique. In a study designed to evaluate the effect of not implementing the multiple nursing technique when sizes at co-stocking are not equal, the linear regression factor proved to inflate correction estimates without multiple nursing (Wohlfarth *et al*., 1991).

2.4 References

Abucayo, J.S., & Mair, G.C. 2000. Divergent selection for growth in the development of a female line for the production of improved genetically male tilapia (*Oeochromis niloticus* L.) In: Fitzsimmons, K., & Filho, J.C. (Eds), Tilapia Aquaculture, *Proceedings of the Fifth International Symposium on Tilapia in Aquaculture*. Rio De Janeiro, Brazil. pp. 71-72.

Baker, R.J., & R.I.H. McKenzie. 1967. Use of control plots in yield trials. *Crop Science*. **7**: 335–337.

Basiao, Z.U. & Doyle, R.W. 1990. Use of internal reference population for growth rate comparison of tilapia strains. I. In a crowded environment. In: Hirano, R. & Hanyu, I., (Eds.), *The Second Asian Fisheries Forum. Manila: Asian Fisheries Society*. pp. 503-506.

Basiao, Z.U., Doyle, R.W., & Arago, A.L. 1996. A statistical power analysis of the 'internal reference' technique for comparing growth and growth dispensation of tilapia strains. *Journal of Fish Biology*. **49** : 277-286.

Blanc, J.M., Chevassus, B., & Poisson, H., 1983. Utilisation de mutants ''golden'' comme temoins intra-lots pour le testage des performances de croissance chez la truite arc-en-ciel. *Cybium*. **7**(2): 93–103.

Blanc, J.M., Poisson, H., Vallée, F. 2001. Covariation between diploid and triploid progenies from common breeders in rainbow trout, *Oncorhynchus mykiss* (Walbaum). *Aquaculture Research*. **32**: 507-516.

Buchtová, H., Svobodová., Z. Kocour, M., & Velišek, J. 2006. Evaluation of growth and dressing out parameters of experimental scaly crossbreds in 3-year-old common carp (*Cyprinus carpio*, Linnaeus 1758). *Aquaculture Research*. **37**: 466-471.

Bolivar, R.B. & Newkirk, G.F. 2000. Response to selection for body weight of Nile tilapia (*Oreochromis niloticus*) in different culture environments. In: Fitzsimmons, K., Filho, J.C. (Eds.), *Proceedings from the Fifth International Symposium on Tilapia Aquaculture. American Tilapia Association and ICLARM*. pp. 12-23.

Bondari, K., Dunham, R.A., Smitherman, R.O., Joyce, J.A. & Castillo, S. 1983. Response to bidirectional selection for body weight in blue tilapia. *In:* Fishelson, L. & Yaron, Z. (eds.), *Proceedings of the International Symposium on Tilapia in Aquaculture*, Tel Aviv University, Tel Aviv, Israel. pp. 302-312.

Bondari, K. 1983. Response to bidirectional selection for body weight in channel catfish. *Aquaculture*. **33**: 73-81.

Bray, D.F., Bell, A.E., & King, S.C., 1962. The importance of genotype-by-environment interaction with reference to control populations. *Genetic Research*. **3**: 282-302.

Brink, D., Mair, G.C., Hoffman, L., & Beardmore, J.A. DFID Aquaculture Final Technical Report of project R7284: Genetic improvement and Utilisation of Indigenous Tilapia in Southern Africa. DFID Aquaculture and Fish Genetics Programme. pp 1-183.

Brink, D. 2004. Genetic improvement of growth rate in rainbow trout (*Oncorhynchus mykiss).* Phd Thesis, Stellenbosch University, South Africa.

Buck, D.H., Thoits, C.F., & Rose, C.R. 1970. Variation in carp production in replicate ponds. *Transactions of the American Fisheries Society*.**99**: 74-79.

Busack, C.A., & Riddell, B.E. 1985. Effect of rearing-group composition on the growth of juvenile Coho salmon in hatchery environments. *Proceeding of the 2nd International Symposium on Genetics in Aquaculture*, University of Davis, CA, 23-28 June 1985.

Cameron, N.D. 1997. Selection Indices and Prediction of Genetic Merit in Animal Breeding. CAB international. Oxon, United Kingdom.

Cameron, N.D., & Curran, M.K. 1995. Responses in carcass composition to divergent selection for components of efficient lean growth rate in pigs. *Animal Science*. **61**(2): 347-359.

Chandra, S. 1994. Efficiency of check-plot designs in unreplicated field trials. *Theoretical and Applied Genetics*. **88**: 618-620.

Clayton, G.A., Morris, J.A. & Robertson, A. 1957. An experimental check on quantitative genetic theory. I. Short-term selection experiments. *Journal of Genetics*. **55**: 131-151.

Cloete, S. W. P., Olivier, J. J., Van Wyk, J. B., Schoeman, S. J., & Erasmus, G. J. 2005. Genetic parameters and trends for hogget traits in Merino lines divergently selected for multiple rearing ability. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics.* **16**: 24-27.

Cochran, W.G. 1937. A Catalogue of Uniformity Trial Data. *Supplement to the Journal of The Royal Statistical Society*. **4** (2): 233-253.

Cochran, W.G. 1957. Analysis of Covariance: Its Nature and Uses. *Biometrics*. **13**(3): 261-281.

Cundiff, L.V., Gregory, I.C.E., & Long, C.R., 1975. Genetic variation among and within herds of Angus and Hereford cattle. *Journal of Animal Science*. **41**: 1270-1280.

Cullis, B.R., Smith, A.B., & Coombes, N. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural Biological and Environmental Statistics.* **11**: 381-393.

Davis, M.E. 1987. Divergent selection for postweaning feed conversion in beef cattle: predicted response base on an index of feed intake and gain vs. fee: gain ratio. *Journal of Animal Science*. **65**(4): 885-895.

Dickerson, G.E. 1969. Efficiency of alternative designs in selection experiments. In: Proceedings of World Conference on Animal Production. *American Dairy Science Association*, Urbana, Illinois. pp 310.

Doyle, R.W., Field, C.A. & Basiao, Z.U. 1990. A Statistical procedure for using 'reference fish' to compare the growth of genetic strains in aquaculture. In: Hirano, R. & Hanyu, I., (Eds.), *The Second Asian Fisheries Forum. Manila: Asian Fisheries Society*. pp 499-502.

Dunham, R.A., & Brummett, R.E. 1999. Response of Two Generations of Selection to Increased Body Weight in Channel Catfish, *I. punctatus*, Compared to Hybridization with Blue Catfish, *I. furcatus*, Males. *Journal of Applied Aquaculture*. **9**(3): 37-45.

Dunham, R.A., Smitherman, R.O., Chappel, J.A., Youngblood, P.N., & Bite, T.O. 1982. Communal stocking and multiple rearing technique for catfish genetics research. *Journal of the World Mariculture Society*. **13**: 261-267.

Edmondson, R.N. 2005. Past developments and future opportunities in the design and analysis of crop experiments. *The Journal of Agricultural Science*. **143**(1): 27-33.

Eisen, E.J., Legates, J.E., & Robison, O.W. 1970. Selection for 12-Day litter weight in mice. *Genetics*. **64** (3-4): 511-532.

Falconer, D.S. 1952. The problem of environment and selection. *The American Naturalist*. **86:** 293–298.

Falconer, D.S., Mackay T.F.C. 1996. Introduction to Quantitative Genetics Edit. Longman group Ltd.

Federer, W.T. 1956. Augmented (or hoonuiaku) designs. *Hawaiian Planters' Record*. **55**: 191-208.

Federer, W.T., & Raghavarao, D. 1975. On augmented designs. *Biometrics*. **31**: 29-35.

Federer, W.T., Nair, R.C., & Raghavarao,D. 1975. Some Augmented Row-Column Designs. *Biometrics*. **31**: 361–374.

Fisher, R.A. 1925. Statistical methods for research workers. Oliver and Boyd Ltd., Edinbergh.

Fisher, R.A. 1926. The arrangements of Field Experiments. *Journal of the Ministry if Great Britain*. **33**: 503-513.

Frankham, R., Jones, L.P., & Barker, J.S.F. 1968. The effects of populations size and selection intensity in selection for a quantitative character in Drosophila. I. Short-term response to selection. *Genetic Resources*. **12**: 237-248.

Fredeen, H. 1986. Monitoring Genetic Change. *Aquaculture*. **57**: 1-26.

Gall, G.A.E., Bakar, Y., & Famula, T. 1993. Estimating genetic change from selection. *Aquaculture*. **111**: 75-88.

Gall, G.A.E., & Bakar, Y. 2002. Application of mixed-model techniques to fish breed improvement: analysis of breeding-value selection to increase 98-day body weight in tilapia. *Aquaculture*. **212**: 93-113.

Gjerde, B., Gunnes, K., & Gjedrem, T. 1983. Effect of Inbreeding on Survival and Growth in Rainbow Trout. *Aquaculture*. **34**: 327-332.

Gjerde, B., & Gjedrem, T. 1984. Estimated of Phenotypic and Genetic Parameters for Carcass Traits in Atlantic Salmon and Rainbow Trout. *Aquaculture*. **36**: 97-110.

Goodwin, K., Dickerson, G.E., & Lamoreux, W.F. 1955. A technique for measuring genetic progress in poultry breeding experiments. *Poultry Science*. **34**: 1197. Abstract.

Goodwin, K., Dickerson, G.E., & Lamoureux, W.F. 1960. An experimental design for separating genetic and environmental changes in animal populations under selection. Biometrical Genetics. (Ed O. Kempthorne), Pergamon Press, New York. pp 117-138.

Gowe, R. S., Robertson, A., & Latter, B. D. H. 1959. Environment and Poultry Breeding Problems 5. The Design of Poultry Control Strains. *Poultry Science*. **38**(2): 462-471.

Henderson, C.R. 1953. Estimation of variance and covariance components. *Biometrics*. **9**: 226-252.

Henderson, C.R, 1973. Sire evaluation and genetic trends. In: *Proceedings of the Animal Breeding and Genetics Symposium in Honour of Dr. J.L Lush*, ASAS & ASDA. pp 10-41.

Henderson, CR. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics*. **31**: 423-447.

Henderson, C.R. 1988. Theoretical basis and computational methods for a number of different animal breeding models. In: G.H. Schmidt (Eds), *Proceedings Animal Model Workshop*, Edmonton, Canada, pp. 1-16.

Hershberger, W.K., Myers, J.M., Iwamoto, R.N., McAuley, W.C., & Saxton, A.M. 1990. Genetic Changes in the Growth of Coho Salmon (Oncorhynchus kisutch) in Marine Net-Pens, Produced by Ten Years of Selection. *Aquaculture*. **85**: 187-197.

Hetzel, D.J.S., Crocos, P.J., Davis, G.P., Moore, S.S., & Preston, N.C. 2000. Response to selection and heritability for growth in the Kuruma prawn, *Penaeus japonicus*. *Aquaculture*. **181**. 215-223.

Hill, W.G. 1972a. Estimation of genetic change. I. General theory and design of control populations. *Animal Breeding Abstracts*. **40**: 1-16.

Hill, W.G. 1972b. Estimation of realized heritabilities from selection experiments. I. Divergent selection. *Biometrics*. **28**: 747-766.

Hill, W.G. 1972c. Estimation of genetic change. II. Experimental evaluation of control populations. *Animal Breeding Abstracts*. **40**: 193-213.

Hörstgen-Schwark, G. 1993. Selection experiments for improving “pansize” body weight in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*. **112**: 13-24.

Idrees, N. & Khan, M.I. 2009. Design improvements using uniformity trials experimental data. *Pakistan Journal of Agricultural Sciences*. **46** (4): 315-320.

Jørstad, K. E., Prodöhl, P. A., Kristiansen, T. S., Hughes, M., Farestveit, E., Taggart, J. B., & Ferguson, A. 2005. Communal larval rearing of European lobster (*Homarus gammarus*): Family identification by microsatellite DNA profiling and offspring fitness comparisons. *Aquaculture*. **247**(1): 275-285.

Karaket, T., Poompuang, S., Na-Nakorn, U., Kamonrat, W., & Hallerman, E. M. 2011. DNA microsatellite-based evaluation of early growth performance among strains of freshwater prawn *Macrobrachium rosenbergii* de Man. *Aquaculture*. **311**(1), 115-122.

Karplus, I., Samsonov, E., Hulata, G., & Milstein, A. 1989. Social control of growth in *Macrobrachium rosenbergii*. The effect of claw ablation on survival and growth of communally raised prawns. *Aquaculture.* **80**(3): 325-335.

Kearsey, M. J., and H. S. Pooni. 1996. Genetical analysis of quantitative traits. Chapman & Hall, London.

Kempton, R.A. & Talbot, M. 1988. The Development of New Crop Varieties. *Journal of the Royal Statistical Society. A*. **5**: 227-341.

Kempton, R.A. 1984. The design and analysis of unreplicated field trials. *Vortr. Pflanzenzüchtg*. **7**: 219–242.

Khaw, H.L., Ponzoni, R.W., & Dantinf, M.J.C. 2008. Estimation of genetic change in the Gift strain of Nile tilapia (*Oreochromis niloticus*) by comparing contemporary progeny produced by males born in 1991 or in 2003. *Aquaculture*. **275**: 64-69.

Kincaid, H.L. 1979. Development of Standard Reference Lines of Rainbow Trout. *Transactions of the American Fisheries Society*. **108**: 457-461.

Kirpichnikov, V.S., 1966. Methods of progeny testing spawners in carp fish farms. *Selective Breeding of Carp. Israel Program for Scientific Translations*. **5737**: 36– 55.

Kocour, M., Linhart, O., & Gela, D. 2003. Results of comparative growing test of all-female and bisexual population in two-year-old common carp (*Cyprinus carpio* L.). *Aquaculture International*. **11**: 369-378.

Kocour, M., Linhart, O., Gela, D., & Rodina, M. 2005. Growth Performance of All-female and Mixed-sex Common Carp Cyprinus Carpio, L. Populations in the Central Europe Climatic Conditions. *Journal of the World Aquaculture Society*. **36**(1): 103-133.

Kocour, M., Gela, D., Rodina, M., & Flajšhans, M. 2010. Performance of different tench, *Tinca tinca* (L.), groups under semi-intensive pond conditions: is it worth establishing a coordinated breeding program. *Reviews in Fish Biology and Fisheries*. **20**: 345-355.

Leopold, R.A. 2007. Colony Maintenance and Mass-Rearing: Using Cold storage for Extending the Shelf-Life of Insects. In: Area-Wide Control of Insect Pests. M.J.B. Vreysen, A.S. Robinson, J. Hendrichs (Eds.). Springer, Netherlands. pp. 149-162.

Lin, C.S., & Poushinsky, G. 1983. A Modified Augmented Design for an Early Stage of Plant Selection Involving a Large Number of Test Lines without replication. *Biometrics*. **39**: 553-561.

Liu, G., Dunnington, E.A., & Siegel, P.B. 1994. Responses to Long-Term Divergent Selection for Eight-Week Body Weight in Chickens. *Poultry Science* **73**: 1642-1650.

Linhart, O., Gela, D., Rodina, M., Sˇlechtov´a, V., & Sˇlechta, V. 2002. Top-crossing with paternal inheritance of common carp (*Cyprinus carpio* L.) progeny under two altitude conditions. *Aquaculture*. **204**: 481-491.

Malecha, S. 1980. Research and Development in Freshwater Prawn, Macrobrachium rosenbergii, Culture in the united States: Current Status, and Biological Constraints With Emphasis on Breeding and Domestication. In: Sindemann, C.J. (Eds) *Proceeding of the Ninth and Tenth U.S.-Japan Meetings on Aquaculture.* pp 35-55.

Maluwa, A.O., & Gjerde, B. 2007. Response to selection for harvest body weight of *Oreochromis shiranus*. *Aquaculture*. **273**: 33-41.

Mandel, J. 1954. Chain block designs with two-way eliminations of heterogeneity. *Biometrics*. **10**: 251-272.

Martin, R.J., Chaunhan, N., Eccleston, J.A., & Chan, B.S.P. 2006. Efficient experimental designs when most treatments are unreplicated. Linear Algerbra and its Applications. **417**: 163-182.

McGinty, A.S. 1984. Suitability of communal rearing for performance testing of tilapias. *Proceedings of the Caribbean Food Crops Society*. **19**: 259-266.

McGinty, A.S. 1987. Efficacy of mixed-species communal rearing as a method of performance testing of tilapias. *Progressive Fish Culturist*. **49**: 17-20.

McPhee, C.P., & Jones, C.M.1997. Selection for weight gain in redclaw crayfish. *Association for the Advancement of Animal Breeding and Genetics*. **12**: 81-84.

Meyer, K., & Hill, W.G. 1991. Mixed model analysis of a selection experiment for food intake in mice. *Genetics Research*. **57**: 71-81.

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

Moav, R., Brody, T., Wohlfarth, G., & Hulata, G. 1976. Applications of Electrophoretic genetic markers to fish breeding. I. Advantage and Methods. *Aquaculture*. **9**: 217-228.

Mrode, R.A. 1996. Linear Models for the Prediction of Animal Breeding Values. CAB international. Oxon, United Kingdom.

Mylonas, C., Fostier, A., & Zanuy, S. 2010. Broodstock management and hormonal manipulations of fish reproduction. *General and Comparative Endocrinology*. **165**: 516-534.

Newman, J.A., Rahnefeld, G., & Fredeen, H.T., 1973. Selection intensity and response to selection for yearling weight in beef cattle. *Canadian Journal of Animal Science*. **53**: 1-12.

Nell, J.A., Smith, I.R., & Sheridan, A.K. 1999. Third generation evaluation of Sydney rock oyster *Saccostrea commercialis* (Iredale & Roughley) breeding lines. *Aquaculture*. **170**: 195-203.

Newkirk, G.F., & Haley, L.E.1983. Selection for growth rate in European oyster, *Ostrea edulis*: Response of Second generation groups. *Aquaculture*. **33**: 149-155.

Ninh, N.H. 2009. Communal or Separate Rearing of Families in Selective Breeding of Common Carp (*Cyprinus carpio* L.). Phd Thesis. University of Stirling, Scotland.

Ninh, N.N., Ponzoni, R.W., Nguyen, N.H., Wooliams, J.A., Taggart, J.B., McAndrew, B.J., & Penman, D.J. 2011. A comparison of communal and separate rearing of families in selective breeding of common carp (*Cyprinus carpio*): Estimation of genetic parameters. *Aquaculture*. **323**: 39-46.

Ollivier, L., Messer, L.A., Rothschild, M.F., & Legault, C. 1997. The use of selection experiments for detecting quantitative trait loci. *Genetical Research*. **69**: 227-232.

Patterson, H.D., Williams, E.R., & Hunter, E.A. 1978. Block designs for variety trials. *Journal of Agricultural Sciences*. **90**: 395-400.

Philipson, J., Banos, G., & Arnason, T. 1994. Present and future uses of selection index methodology in dairy cattle. *Journal of Dairy Science*, **77**(10): 3252-3261.

Poehlman, J.M., Sleper, D.A. 1996. Breeding Field Crops. Iowa State University Press, Ames, Iowa. 5th edition.

Ponzoni, R.W., Hamzah, A., Tan, S., & 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.

Refstie, T. 1990. Application of breeding schemes. *Aquaculture*. **85**: 163-l69.

Romana-Eguia, M.R.R., & 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., Ikelda, M., Basiaou, Z.U., & Taniguchi, N. 2010. Growth comparison of Asian Nile and red tilapia strain in controlled and uncontrolled farm conditions. *Aquaculture international*. **18**: 1205-1221.

Rye, M., & Gjedrem, T. 2005. Measuring genetic change. In : T. Gjedrem (Eds), *Selection and Breeding Programs in Aquaculture*, Springer, Netherlands. pp. 243–250.

Sharma, J.R., 1998. Statistical and biometrical techniques in plant breeding. New Delhi: New Age International (Pty) Limited. pp 20-27.

Smith, H.F., 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Sciences*. **28**: 1-23.

Smith, A.B., Lim, P., & Cullis, B.R. 2006. The design and analysis of multi-phase plant breeding experiments. *Journal of Agricultural Science.* **144**. 393–409.

Sorenson, D., Guldbrandtsen, B., & Jensen, J. 2003. On the need for a control line in selection experiments: A likelihood analysis. *Gentics, Selection, Evolution*. **35**: 3-20.

Su, G.S., Liljedahl, L.E., & Gall, G.A.E. 1997. Genetic and environmental variation of female reproductive traits in rainbow trout (Oncorhychus mykiss). *Aquaculture*. **154**: 115-124.

Terčič, D., & Holcman, A. 2008. Long-term Diverngent Selection for 8-Week Body weight in Chickens – A Review of Experiments. *Acta agriculturae Slovenica.* **92**(2): 131-138.

Thompson R., Atkins K.D. 1994. Sources of information for estimating heritability from selection experiments, *Genetic Research*. **63**: 49-55.

Toro, J.E., & Newkirk, G.F. 1990. Divergent selection for growth rate in the European oyster *Ostrea edulis*: response to selection and estimation of genetic parameters. *Marine Ecology Progress Series*. **62**: 219-227.

Toro, J.E., & Newkirk, G.F. 1991. Response to artificial selection and realized heritability estimate for shell height in the Chilean oyster *Ostrea chilensis*. *Aquatic Living Resources*. **4**: 101-108.

Uraiwan, S., & Doyle, R.W. 1986. Replicate variance and the choice of selection procedure for tilapia (*Oreochromis niloticus*) stock improvement in Thailand. *Aquaculture*. **57**: 93-98.

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

Vandeputte, M., Dupont-Nivet, M., Haffray, P., Chavanne, H., Cenadelli, S., Parati, K., Vidal, M.O., Vergnet, A., & Chatain, B. 2009. Response to domestication and selection for growth in the European sea bass (*Dicentrarchus labrax*) in separate and mixed tanks. *Aquaculture*. **286**: 20–27.

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

Wohlfarth, G.W., & Moav, R. 1985. Communal Testing, A Method of Testing the Growth of Different Genetic Groups of Common Carp in Earthen Ponds. *Aquaculture*. **48**: 143-157.

Wohlfarth, G.W., & Moav, R.1991. Genetic testing of common carp in cages 1. Communal versus separate testing. *Aquaculture*. **95**: 215-223.

Wohlfarth, G.W., & Moav, R. 1993. Genetic testing of common carp in cages. 2. Influence of Variation in Initial Weight on Weight Gain. *Aquaculture*. **109**: 245-256.

Wohlfarth, G.W., & Milstein, A. 1987. Predicting Correction Factors for Differences in Initial Weight among Genetic Groups of Common Carp in Communal Testing. *Aquaculture*. **60**: 13-25.

Wohlfarth, G.W., Nagy, A., & McAndrew, B.J. 1991. A mistaken method for correcting potential bias in genetic testing of common carp, *Cyprinus carpio* L., and tilapias, *Oreochromis* ssp. *Aquaculture and Fisheries Management*. **22**: 309-316.

Williams, E., Piepho, H., & Whitaker, D. 2011. Augmented p-rep designs. *Biometric Journal*. **53**(1): 19-27.

Yates, F. 1936. A new method of arranging variety trials involving a large number of varieties. *Journal of Agricultural Sciences*. **26**: 424-455

Yates, F. 1937. A further note on the arrangement of variety trials: quasi-latin squares. *Annuls of Eugenics.* **8:** 319-332.

Yates, F. 1939. The Comparative Advantages of Systematic and Randomised Arrangements in the Design of Agricultural and Biological Experiments. *Biometrika*. **30**(3): 440-466.

Yates, F. 1940. Modern experimental design and its function in plant selection. *Empire Journal of Experimental Agriculture*. **8**: 222-230.

Zhang, R., Warrick, A.W., & Myers, D.E. 1994. Heterogeneity, plot shape effect and optimum plot size. *Geoderma*. **62**: 183-197.

Zohar, Y., 1989. Fish reproduction: its physiology and artificial manipulation. In: Shilo, M., Sarig, S. (Eds.). *Fish Culture in Warm Water Systems: Problems and Trends*. CRC Press, Boca Raton. pp 65-119.

Chapter 3

3.1 MATERIALS AND METHODS

3.1.1 Formation of Base Population

With the commencement of the Innovation Fund Abalone Breeding Project in 2005, a base population was formed through the pooling of wild *Haliotis midae* individuals, randomly sampled from the Walker Bay region of the Overberg in the Western Cape of South Africa. The resulting base population consisted of 800 sexually mature individuals, distributed randomly amongst the five participating commercial abalone farm hatcheries.

The five participating commercial abalone farms are representative of over 75% of the total abalone production in South Africa, even though they are all located within the Walker Bay region (Brink *et al*., 2011). Two of the commercial farms: Irwin & Johnson Ltd, Abalone Division (I&J) and Roman Bay Sea Farm (Pty.) Ltd (RB) are located at Danger Point and at Roman Bay near Gansbaai respectively. The three remaining commercial farms: Abagold (Pty.) Ltd (Aba), HIK Abalone (Pty.) Ltd. (HIK) and Aquafarm Development Company (Pty.) Ltd. (AF) are located near the New Harbour in Hermanus.

The procedure of randomly sampling from wild populations and the subsequent pooling thereof, aimed at achieving maximal genetic diversity within the base population. The examination of 12 species-specific microsatellite markers by Swart (2012) indicated the success of this exercise with heterozygosity values ranging from 0.778 to 0.824, average allele number 18,01 and an inbreeding coefficient of 0.121.

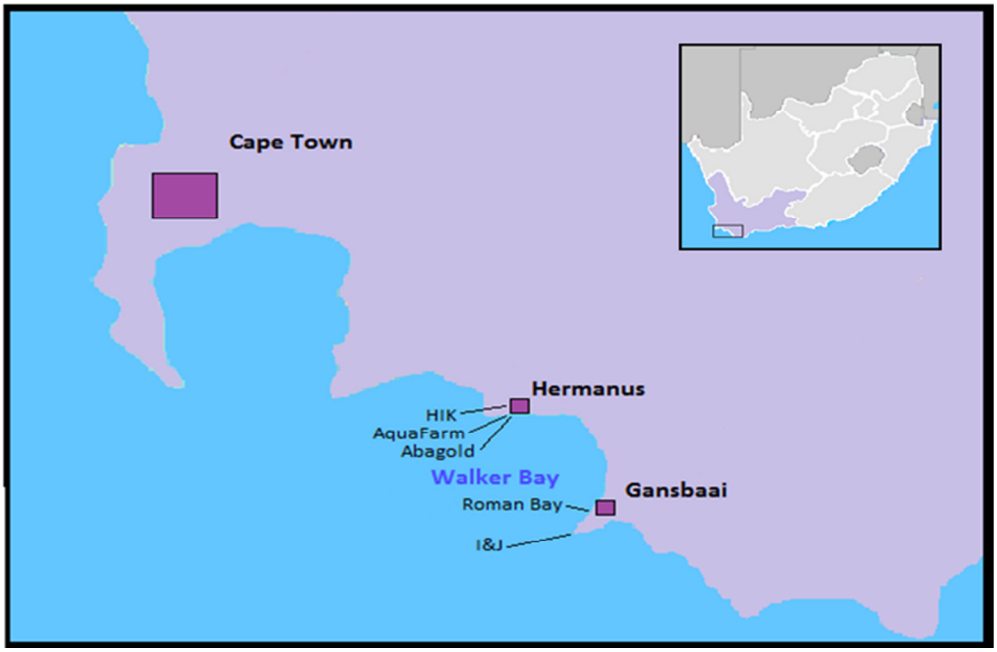

Figure 3.2 Locations of participating abalone farms within South Africa.

3.1.2 Synchronised Individual Spawning and Mating Design

In large, *H. midae* sampled from wild populations do not spawn readily in captivity (Sales & Brits, 2001). For this reason, a random sample of individuals was drawn from the base population at each of the five locations and conditioned for spawning in the 2008-2009 summers spawning season. The conditioning of brood stock was carried out according to the specific protocols of each of the five hatcheries. Protocols for the conditioning of *Haliotis midae* have been summarised by Fleming (1999). The process predominantly entails an acclimatisation period to allow brood stock to adjust to the new hatchery environment, and strictly controlled temperature and photoperiod regimes.

Spawning of brood stock was synchronised through chemical induction between and within each location, in an effort to establish maximal number of test families given limited hatchery capacities, while reducing the possible effect of age differences between test families. The risks of introducing hatchery bias was rejected due to findings of a previous study conducted at each of the five hatcheries utilising the same base population, which revealed no significant differences in growth rates between test families reared until completion of the weaning phase which is approximately 24 months after spawning (Vlok, 2012).

Individuals conditioned for spawning were kept individually in separate containers throughout the conditioning and spawning process. Sperm and eggs were collected from the tanks of individuals successfully induced to spawn and mixed for the establishment of families. The mating design was aimed at the production of unrelated full sibling families to enable the evaluation of maximal diverse families. However, limitations in the success of induced synchronised spawning and the time constraints on the fertilising capability of sperm resulted in the establishment of a small proportion of half sibling families. Successful spawning of 118 sires and 125 dams culminated in the establishment of 191 families over a 100 day period completed in mid-December 2008. The fertilisation dates of all families were recorded.

As there remains no recognised or reliable method to tag abalone individuals until they reach the size of 10mm at roughly 6 months after spawning, all families were maintained and cultured separately throughout fertilisation and the developmental and pre-experimental phases according to the specific protocols of each participating hatchery. During which time, no evaluation on growth performances was conducted. A detailed overview of the spawning, development and weaning of *Haliotis midae* is supplied in the work of Genade et al., (1988).

Upon completion of the weaning phase the families were moved to grow out baskets under commercial stocking densities of 1000 animals per basket. Families with adequate

numbers for replicates while maintaining commercial stocking densities were retained for this purpose. Of 191 families to successfully complete the weaning stage, 6 had adequate numbers for the inclusion of a replicate at each of the five locations; the remaining 185 families had enough numbers to facilitate a single unreplicated basket. The 6 families were randomly assigned to 5 replicate baskets each and relocated to each of the five locations under standard biosecurity and quarantine procedures.

3.1.3 Reference Group and Tagging

A single full sibling family was selected for use as the internal reference group on the basis of success during settlement and subsequent development stages, culminating in approximately 17 250 animals upon completion of the weaning phase.

The internal reference group was tagged to allow for individual identification of members for the duration of the trial. The specific method of tagging was described by Brink et al, (2009) and specially developed for use in the Innovation Fund Abalone Breeding Project. The reference individuals were anaesthetised in CO_2 saturated baths for 5 minutes and arranged over hydrated sponge mats with the dorsal shell surface exposed. Paper towel was used to remove excess water and sessile organisms from the shell surface and allowed to air dry for a few minutes. A cyanoacrylate adhesive agent, namely Super Glue TM was applied near the apex spire in the posterior region of the shell, upon which a coloured bee tag was carefully placed. Upon polymerisation of the cyanoacrylate and confirmation of secure adhesion of the coloured tag, the animals were returned to their respective grow out baskets to await allocation into test family baskets.

3.1.4 Experimental Design and Growth Trial

The growth trial design utilised was an unreplicated design at two locations, with the exception of the six full sibling families which are included at both locations as repeats in a standard design format. The test families and a random sample of 50 tagged internal reference group individuals were included within a basket holding system used in commercial production of abalone.

A typical basket holding system is of dimension 55cm (W) x 94cm (L) x 56cm (H) and contains vertical plates to act as substrate for the abalone. The baskets were labelled and randomly allocated to land-based flow through tanks of dimension 2 m (W) x 5 m (L) x 1 m (H) at each location. Throughout the period of evaluation the test families and corresponding internal reference groups were subjected to the standard rearing conditions of each location such as stocking densities, types of feed, feeding practices and methods, flow rates and handling. In the context of evaluation and selection it is necessary to subject individuals to the commercial rearing environment in which their progeny are expected to perform. As it

was too costly and impractical to standardise the rearing conditions over both locations and only within locations, the possibility of location effects cannot be overlooked (Gjedrem, 2005).

From the approximate age of 24 months after spawning, at regular six month intervals, families were anaesthetised in CO_2 saturated baths to facilitate removal from the basket without unduly stressing individuals. Families are removed for the purpose of re-adjusting stocking densities or “splitting” and the cleaning of tanks and baskets to remove waste, benthic organisms and sessile organisms. A proportion of the test family individuals were randomly sampled and placed in a new basket according to commercial stocking densities, the remainder of individuals were absorbed into the commercial production of each location. In addition, individuals were screened for shell damage and removed, as shell damage permanently retards growth rate and introduces a bias into family performance estimates (Schoonbee. 2008). Special care was taken to ensure families were kept separate and all tagged reference animals were identified and included within the new basket. The baskets were correctly labelled and re-allocated randomly to tanks over the location.

Note, the original experimental design was 6 full sibling test families repeated over all five locations, which were then ‘augmented’ with the remainder of the unreplicated test families at each location, in a manner analogous to an augmented partially replicated (p-rep) design (Smith *et al.*, 2006; Cullis *et al.*, 2006). However, a biosecurity risk in the form of a viral outbreak in the region of the two Gansbaai commercial farms RB and I&J, prior to the onset of the growth trial prompted the relocation of the RB and I&J test families to HIK and ABA respectively. During the relocation and quarantine process two replicates of each of the 6 full sibling repeat test families were lost. Resultantly, a single set of the repeat families was allocated to HIK and a duplicate of repeat families to ABA. Furthermore, records pertaining to spawning information of the AF test families were lost, while these families remain within the selection program they are excluded from the present investigation into the internal reference technique.

3.1.5 Sampling and Measurement

Measurements were obtained through non-destructive sampling of the test families and internal reference groups during the re-adjustment of stocking densities at intervals of every six months as described above. A random sample of the test family and internal reference group is drawn by means of a diagonal transect over each group and the first sixteen individuals beneath the transect line are taken as the samples respectively. The samples are transferred into separate mesh containers to facilitate the removal of excess water prior to measurement.

The weight of each individual in the sample is measured in grams to the nearest 0.1 g by means of an AND EK-300i electronic balance (digital scale). The length of each individual in the sample is measured in mm to the nearest 0.1mm, as the greatest distance between the posterior and anterior extremities of the shell by means of a Mitutoyo IP67 digital calliper. All length and weight data was recorded by a digital data logger. The last measurements were taken after 48 months after spawning which corresponds with the end of the commercial production cycle.

3.1.6 Definition of Traits and Statistical Analysis

The measurement of shell length and body weight was aimed at estimating the latent underlying factor of growth rate of test families and internal reference groups. Growth rate was estimated as Average Daily Weight Gain at 5 years (ADWG) in grams/day and Average Daily Length Gain at 5 years (ADLG) in millimetres/day. ADWG and ADLG can be defines as follows:

ADWG = (Average final sample body weight – average initial sample body weight) /
(Age of sample at final weight – age of sample at initial weight)

Where:

Average sample body weight = (Sum of individual body weights in sample) /
Total number in sample

ADLG = (Average final sample shell length – Average initial sample shell length) /
(Age of sample at final length – age of sample at initial length)

Where

Average sample length = (Sum of individual shell lengths in sample) /
Total number in sample

The data from the five age intervals was analysed using Statistical Analysis System Enterprise Guide 9.1 (SAS Institute Inc. 2012) and R version 2.15.2 (R Core Team, 2012) software. The ADLG and ADWG of test families and reference groups were estimated as the coefficient of linear regression, by means of pros GLM (General Linear Models) in SAS 9.1 as the least square regression of the respective length and weight of samples against time.

Summary statistics and Box-plot analyses were generated for ADLG and ADWG of test families and respective reference groups to identify possible outliers, observations outside the 95% confidence intervals or 1.5 times the interquartile range were evaluated for error and where possible corrected or removed. The data sets pertaining to unreplicated

performances generated in this study were analysed in a step-wise manner to assess the validity of the internal reference group in terms of the following:

- The effect of time on tag loss and possible bias in tag loss in the reference group samples.
- The possible effects of variation in size ratio of reference group to test group at initial co-stocking.
- Evaluate the effect of location on test family and reference group performances.
- Evaluate the correlation between test family and internal reference family performances.

The data sets pertaining to the six repeat families and respective internal reference groups generated in this study were analysed in a step-wise manner to assess the validity of the internal reference group in terms of the following:

- Testing for possible genotype by environment interactions.
- Testing for possible antagonistic interactions between test families and reference groups.
- Compare coefficients of determination between statistical models with various methods of implementing reference group information.
- Evaluate the change in replicate variance due to reference group correction.
- Evaluate the correlation between test family and internal reference family performances.

3.2 Summary

Wild populations of *H. midae* were sampled and pooled to form a genetically diverse base population. A sample of wild individuals from the base population was conditioned and induced for synchronised individuals spawning. Families produced were reared separately throughout the fertilisation to weaning phases. At age 18 months families were stocked in baskets at an initial density of 1000 individuals per basket. Families with adequate numbers for replication were retained. An extremely successful full sibling family was tagged and included in each basket at an initial density of 50 tagged individuals per basket. Test families were cultured at two locations for a period of 30 months according to each locations specific culture methods. At intervals of six months test families were anaesthetised and stocking densities re-adjusted, concurrently random samples of 16 were drawn from the test families and reference groups and individuals body weight and shell length measured and recorded.

The objective of the growth trial was to evaluate the maximal number of test families and select individuals from families with superior growth rate, for consolidation into the brood stock population. The relocation of families due to a biosecurity concern and the subsequent reduction in number of test locations provided an opportunity to evaluate the workings of the

internal reference group method. The objective of the present study is to evaluate the use of the internal reference group method within the growth trial and implement the optimal use of reference group information towards identifying test families with superior growth rate.

Data was analysed *ad posteriori* with the specific objectives of testing the validity of the reference group method and assessing the efficiency of various methods of implementing the reference group information.

3.3 References

Swart, L. 2012. Microsatellite markers as a tool in genetic enhancement and husbandry of *Haliotus midae*: a South African case study.Msc Thesis, Stellenbosch University, South Africa.

Brink, D., Van Der Merwe, A.E, Roodt-Wilding, R., Hatting, A., Yearsley, R., Landman, R., Van Wyk, S., & Janse, L. 2011. Management of Innovation Fund Abalone Breeding Project. Hermanus, South Africa. Personal communication.

Fleming A.E. 1999. Conditioning of Australian abalone broodstock. Marine and Freshwater Research Institute, Queenscliff, Australia.

Genade, A.B., Hirst, A.L. & Smit, C.J. 1988. Observations on the spawning, development and rearing of the South African abalone *Haliotis midae* Linn. *South African Journal of Marine Science*. **6**: 3-12.

Brink, D., Van Der Merwe, M., Vlok, A.C., Hepple, J., Slabbert, R., 2009. Tagging methodology for juvenile and adult abalone (*Haliotis midae*). Poster presentation during AASA International Aquaculture Conference, Swakopmund, Namibia.

Smith, A.B., Lim, P. & Cullis, B.R. 2006. The design and analysis of multi-phase plant breeding experiments. *Journal of Agricultural Science* **144**: 393‐409.

Cullis, B. R., Smith, A. B. and Coombes, N. E. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological and Environmental Statistics*. **11**: 381‐393.

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

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

SAS Institute Inc. 2012. *What's New in SAS® 9.3*. Cary, NC: SAS Institute Inc.

Schoonbee, L., 2008. The effect of triploidy on the growth and survival of the indigenous abalone, Haliotis midae, over a 24-month period under commercial rearing conditions. M.Sc. thesis. Stellenbosch University. Stellenbosch, South Africa

Vlok, A.C. 2012. Growth assessment. Unpublished M.Sc. thesis. Stellenbosch University, South Africa

Sales, J. & Britz, P.J. 2001. Research on abalone (Haliotis midae L.) cultivation in South Africa. *Aquaculture Research*. **32**: 863-874

3.4 Photographic Outlay

Figure 1: Brood stock conditioning tanks

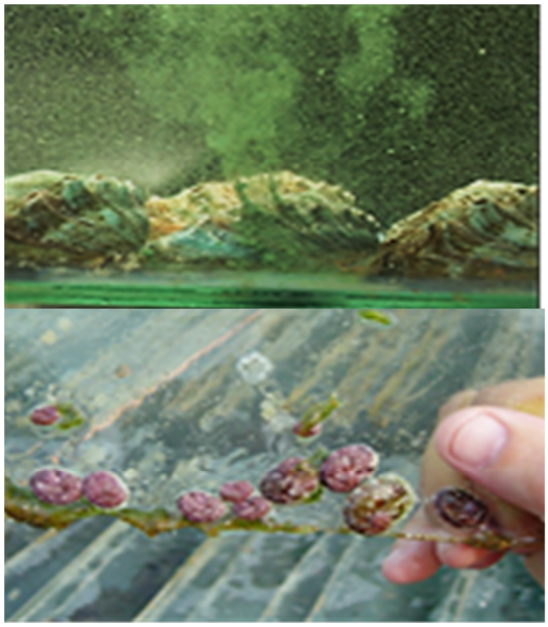

Figure 2: Female abalone spawning & settlement of juveniles

Figure 3: Larval rearing tanks

Figure 4: Abalone weaning tanks

Figure 5: Juvenile abalone in weaning cone

Figure 6: Tagged juvenile reference group

Template 1: Facilities and activities associated with spawning and developmental phases in the Innovation Fund Abalone Breeding Project.

Figure 7: Grow-out basket & substrate

Figure 8: Data logger, digital balance & digital calliper

Figure 9: Grow-out tanks at HIK (Pty.) Ltd

Figure 10: Diagonal transect sampling

Figure 11: Grow-out tanks at Abagold (Pty.) Ltd

Template 2: Facilities and activities associated with growth-out phase and measurement in the Innovation Fund Abalone Breeding Project.

Chapter 4

RESULTS

4.1 Evaluation of Unreplicated Test Groups and Internal Reference Groups

4.1.1 Tag Loss

The loss of tagged reference animals due to failed tag adhesion, encrusting organisms masking the tags, damaged of tagged individuals during handling, mortality, escape and human counting error, was reflected in decreased samples sizes at each measurement interval. Descriptive statistics of tagged animals relative to the sample sizes at each measurement interval are summarised in Table 4.1 below.

Table 4.1 Descriptive statistics of sample sizes due to tag loss through time.

Time since tagging (Months)	Mean number of tagged animals	Standard Deviation	Median	Count	Min.	Max.
0	15.93	0.277	16	173	14	16
6	15.59	1.42	16	173	2	16
12	15.05	2.88	16	173	2	16
18	12.65	4.26	15	173	0	16
24	8.83	4.65	9	173	0	16

The data pertaining to tag loss was tested for deviations from normality and homoscedasticity through the Shapiro-Wilk test and Levene's test for equality of variances (Shapiro & Wilk, 1965; Levene, 1960). Both tests proved significant for deviations from normality ($p<0.001$) and homoscedasticity ($p<0.001$) and subsequent transformation thereof failed to remedy the deviations. It was therefore decided to proceed with non-parametric testing by means of a Kruskal Wallis non-parametric analysis of variance, the results of which (Table 4.2) show at least one of the measurement intervals differs significantly from the rest ($p<0.001$) (Kruskal & Wallis, 1952).

In order to determine the measurement intervals where tag loss first becomes significant, pairwise Wilcoxon signed-rank tests (Wilcoxon, F. 1945) were used to compare tag loss at all measurement intervals with the first measurement at 0 months after tagging where tag loss is close to zero (Table 4.3). The Wilcoxon signed-rank test comparisons found significant differences between all measurement intervals ($p<0.001$) with the exception of 6 months and 12 months after tagging ($p=0.0741$).

Table 4.2 Kruskal Wallis non-parametric analysis of variance for tag loss at measurement intervals.

Time since tagging (Months)	N	Sum of Scores	Expected under Ho	Std Dev under Ho	Mean Score	Kruskal Wallis test
0	173	98 665.0	74 563	2 541.61	579.32	
6	173	91 422.5	74 563	2 541.61	528.45	
12	173	85 716.0	74 563	2 541.61	495.47	
18	173	62 997.0	74 563	2 541.61	364.14	
24	173	32 290.5	74 563	2 541.61	186.65	
Chi-square						361.13
DF						4
Pr<Chi-square						<0.001

Table 4.3 Pairwise Wilcoxon-signed rank test for significant differences in tag loss between measurement intervals

Time since tagging (Months)	0	6	12	18	24
0					
6	0.0011*				
12	<0.001**	0.0741			
18	<0.001**	<0.001**	<0.001**		
24	<0.001**	<0.001**	<0.001**	<0.001**	

* significant at the 0.05 significance level
** significant at the 0.01 significance level

A graphical representation of how the number of samples of internal reference groups which fall below 16, 10 and 0, increases due to tag loss through time (Figure 4.1) is presented below.

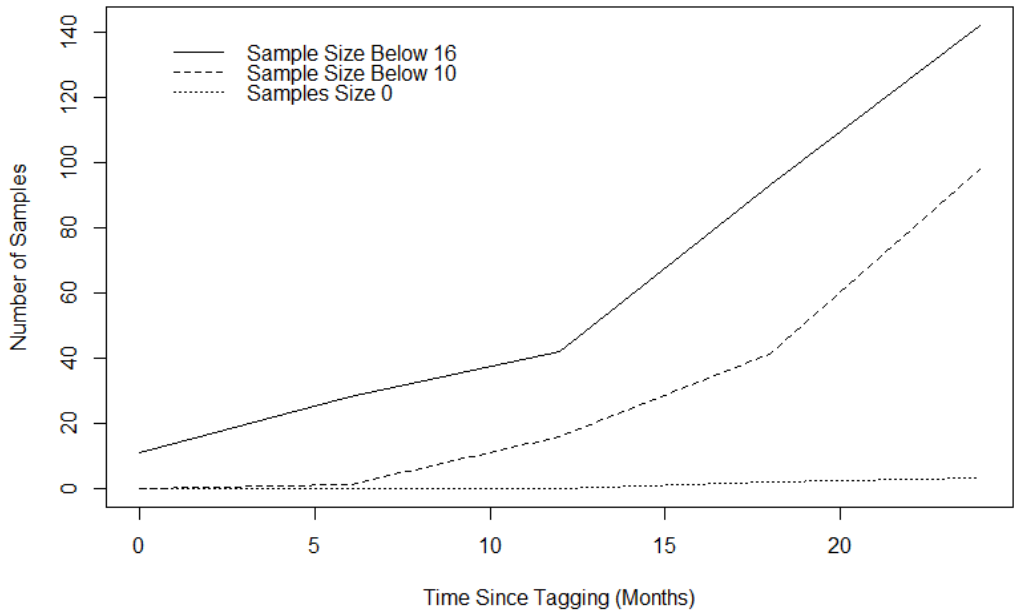

Figure 4.1 Graph of number of internal reference samples which fall below 16,10 and 0 due to tag losses over time.

4.1.2 Location Effect

The average daily weight gain (ADWG) and average daily length gain (ADLG) of internal reference groups was calculated as the regression coefficient of the respective linear regression analyses. Even though samples sizes greatly decreased through time due to tag loss, the effects acting on tag loss are seen to act randomly and do not bias the estimates. However, the decreases in sample size does lower the accuracy of the linear regression analyses and increase the confidence limits near the extremities of the regression and this has been identified as a weakness. The experimental units with complete loss of internal groups were excluded from further analysis as this would bias linear regression analyses. Furthermore, a small number of experimental units (baskets) were temporarily lost at each location preventing measurement and “splitting” which effectively meant these units were not under the same stocking densities which would bias performances and were thus also excluded from further analyses and notes as a managerial weakness. Descriptive statistics for ADWG and ADLG of internal reference groups and test families are summarised by location (Table 4.4) below.

Table 4.4 Descriptive statistics for test families and internal reference groups by location.

Location		Abagold			HIK		
		Mean[A]	Std Dev[B]	CV[C]	Mean[A]	Std Dev[B]	CV[C]
ADLG	Test	$5.86x10^{-2}$	$5.47x10^{-3}$	9.3	$5.83x10^{-2}$	$2.99x10^{-3}$	5.1
	Reference	$5.91x10^{-2}$	$4.62x10^{-3}$	7.8	$5.47x10^{-2}$	$3.41x10^{-3}$	6.2
ADWG	Test	$1.45x10^{-1}$	$2.30x10^{-2}$	15.8	$1.50x10^{-1}$	$1.80x10^{-3}$	11.7
	Reference	$1.37x10^{-1}$	$1.84x10^{-2}$	13.4	$1.45x10^{-1}$	$9.40x10^{-3}$	6.5

[A] Means (g/day) for ADLG and (mm/day) for ADWG
[B] Standard deviation of means
[C] Coefficient of variation (%)

The ADLG and ADWG of internal reference groups were tested for deviations from normality and homoscedasticity by means of Shapiro-Wilk test and Levene's test of equality of variances (Shapiro & Wilk, 1965; Levene, 1960). Both ADLG and ADWG of internal reference groups were found to deviate significantly from normality and homoscedasticity at the 5% significance level. Transformation of ADLG by a $\log_{10}$ transformation resulted in non-significant deviations from normality and homoscedasticity with p=0.5150 and p=0.0558, respectively. However, all subsequent transformations of ADWG failed to resolve the deviations from normality and homoscedasticity.

The test for significant differences between performances of internal reference groups at locations for ADLG was carried out by means of a parametric t test, the results of which were highly significant (p<0.001) as presented in Table 4.5 below. Test for significant differences between performances of internal reference groups at locations for ADWG was carried out by means of a Wilcoxon signed-rank test, the results of which were also highly significant (p<0.001) as presented in Table 4.5 below (Wilcoxon, 1945).

Table 4.5 Tests for significant differences between traits of interest (ADLG & ADWG) at each of the two locations (Abagold and HIK).

ADLG			ADWG	
Parametric t-test			Wilcoxon signed-rank test	
Test Statistic	Df	p	Test Statistic (W)	p
-7.127	165.85	$2.33x10^{-1}$	2393.5	0.0003

To better visually highlight the differences in performance of internal reference groups at each location, a scatterplot of ADLG versus ADWG is included in Figure 4.3 below. In addition the phenotypic correlation between ADLG and ADWG and respective 95% confidence intervals were calculated by means of least squares regression utilising R version 2.15.2 and overlaid (R Core Team, 2012).

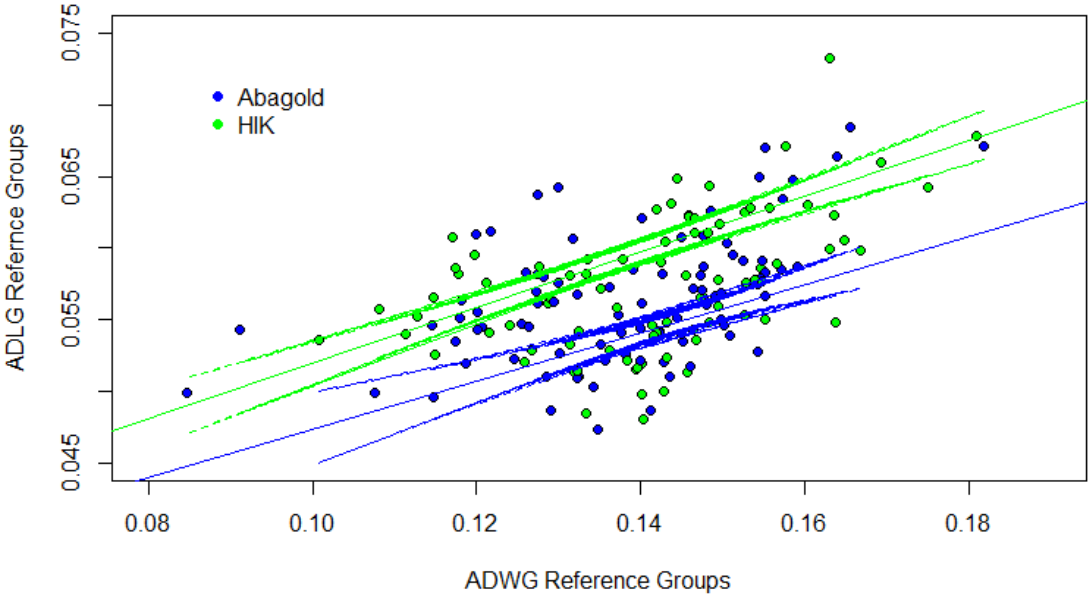

Figure 4.2 Scatterplot of average daily weight gain (ADWG) (g/day) and average daily length gain (ADLG)(mm/day) of reference groups at each location (solid points), correlation between ADWG and ADLG of reference groups at each location (solid line) and accompanying 95% confidence intervals (dotted line).

4.1.3 Antagonistic Interaction Effects and Initial Size Variability

Due to the inherent lack of replication of test families, a two-way analysis of variance with an interaction term, to determine if test and reference groups are interacting, is not possible. However, the evaluation of the effect of co-stocking reference groups with test groups of differing sizes due to age differences may prove informative for future implementation. To this end, the size ratio of internal reference group to respective test group was calculated utilising sample means drawn at the commencement of co-stocking. The evaluations were conducted separately for each location to remove confounding of certain size ratio classes with that of location effects. Subsequent testing for deviation of the residuals from normality and homoscedasticity of size ratio classes were found to not be significant, and no transformation of data required (Shapiro & Wilk, 1965; Levene, 1960).

A one-way analysis of variance (ANOVA) was performed using the General Linear Models (GLM) procedure of SAS 9.1, to evaluate the effect of size differences at co-stocking on the ADLG and ADWG of internal reference groups at each location (SAS Institute, 2011). The following model was fitted for the main effect (size differences at co-stocking):

$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$

Where Y_{ij} is the j^{th} observation of the i^{th} treatment (size ratio), μ is the common mean, α_i is the effect of size ratio at initial co-stocking, and ε_{ij} is the random error.

The results of which can be found in Table 4.6 for ADLG and Table 4.7 for ADWG. The effect of size ratio at initial co-stocking was found to be highly significant at the location HIK for ADLG ($p<0.001$) and ADWG ($p<0.05$). Due to the extreme differences in the number of observations in each class for the main effect, multiple comparisons to determine the upper and lower thresholds by means of least square means (LS Means) is not possible. However, boxplots of size ratio classes are still visually informative to this end and are included for ADWG (Figure 4.3) and ADLG (Figure 4.4) at location HIK (R Core Team, 2012).

Table 4.6 One-way ANOVA for main effect of size ratio of reference group to test group at each of the locations (Abagold and HIK).

Average Daily Length Gain

Abagold						HIK					
Source	d.f.	Mean square	*F*-ratio	*P*	R^2	Source	d.f.	Mean square	*F-ratio*	*P*	R^2
Size Ratio	8	$1.74x10^{-5}$	0.791	$6.12x10^{-1}$	0.08	Size Ratio	5	$8.56x10^{-5}$	12.83	$5.93x10^{-9***}$	0.46
Error	78	$2.20x10^{-5}$				Error	73	$6.68x10^{-6}$			

* Significant at 0.05
** Significant at 0.01
***Significant at 0.001

Table 4.7 One-way ANOVA for main effect of size ratio of reference group to test group at each of the locations (Abagold and HIK).

Average Weight Length Gain

Abagold						HIK					
Source	d.f.	Mean square	*F*-ratio	*P*	R^2	Source	d.f.	Mean square	*F-ratio*	*P*	R^2
Size Ratio	8	$3.90x10^{-3}$	1.502	0.17	0.13	Size Ratio	5	$2.18x10^{-4}$	2.73	$2.60x10^{-2*}$	0.15
Error	78	$2.53x10^{-2}$				Error	73	$7.99x10^{-5}$			

* Significant at 0.05
** Significant at 0.01
*** Significant at 0.001

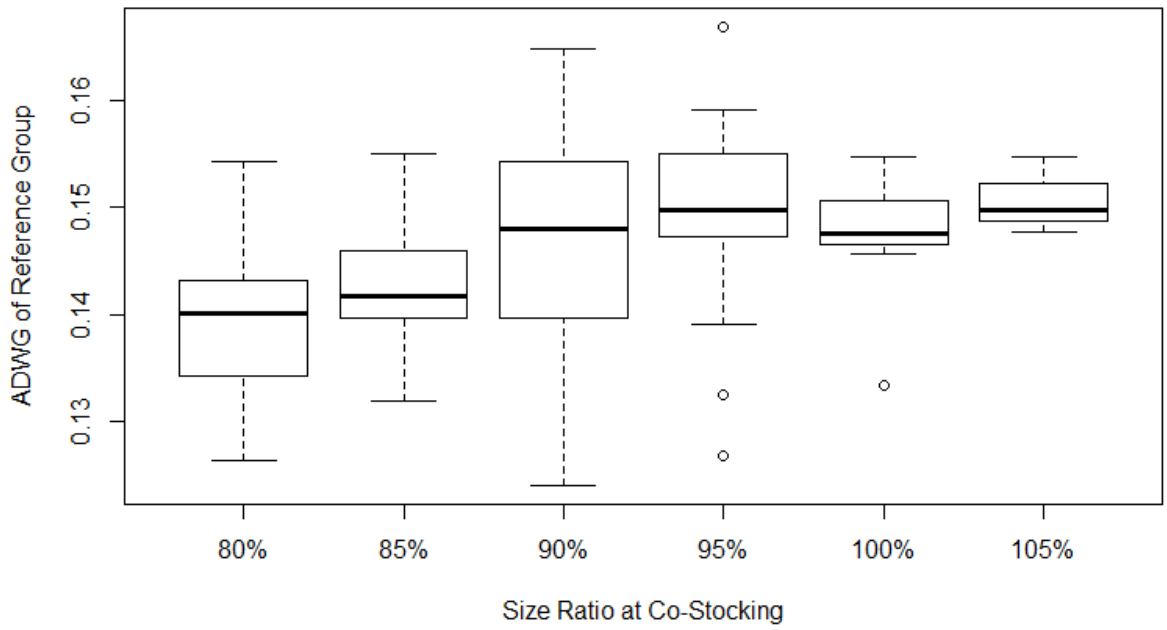

Figure 4.3 Boxplots of the ADWG of size ratio classes (internal reference group to test group) at initial co-stocking for location HIK.

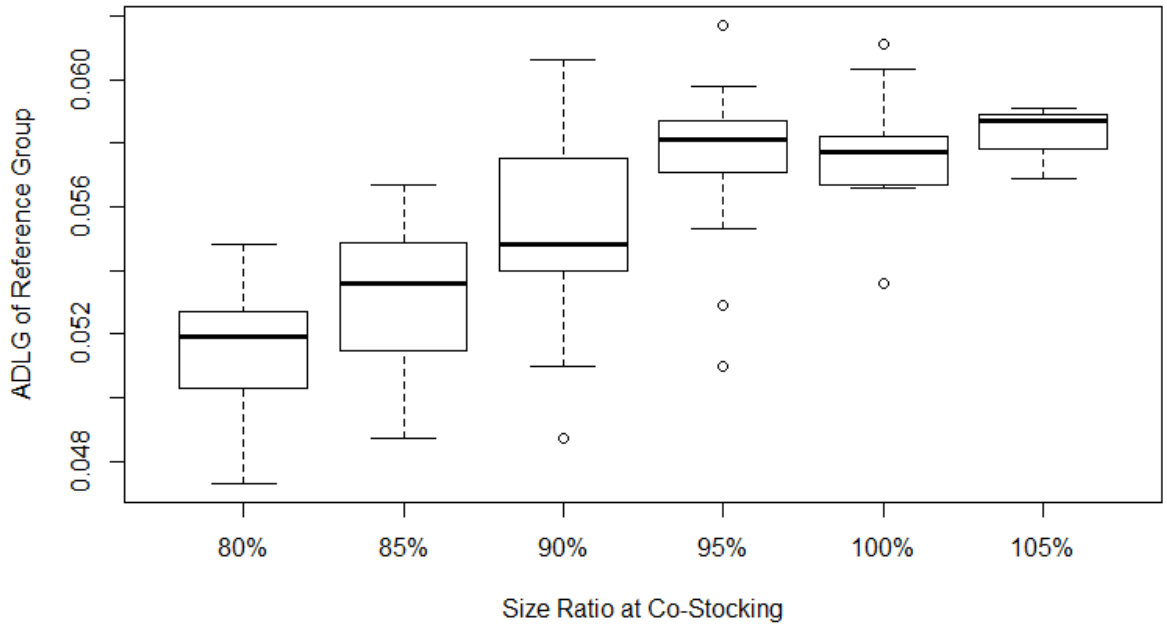

Figure 4.4 Boxplots of the ADLG of size ratio classes (internal reference group to test group) at initial co-stocking for location HIK.

4.1.4 Correlation between Unreplicated Test Family Performance and Reference Performance

The internal reference group is utilised as a measure of the within experimental unit environment, so that if an experimental unit (basket) experiences favourable environmental conditions due to uncontrollable environmental factors such as lower stocking density due to mortalities or slightly higher water temperatures, this will be reflected as an increase in the performance of the internal reference group as well as the respective test family (Kirpichnikov, 1966). Provided no antagonistic interactions occur between test family and internal reference group and no genotype by environment interactions occur, the correlation between internal reference group and test family is expected to be strong and positive (Basiao & Doyle, 1990).

A scatter plot of ADWG and ADLG of reference groups against ADWG and ADLG of test families for each location are included below in Figure 4.5 and Figure 4.6 respectively. The correlations between the performances of test families and reference group for each of the locations were calculated by means of pearson's correlation coefficient and fitted as a solid line, estimated by least squares regression in R version 2.15.2 (R Core Team, 2012). Pearson's correlation coefficient for ADWG at location Abagold is 0.10 and location HIK is 0.33, for ADLG at location Abagold is 0.21 and location HIK is 0.24 (Pearson, 1986).

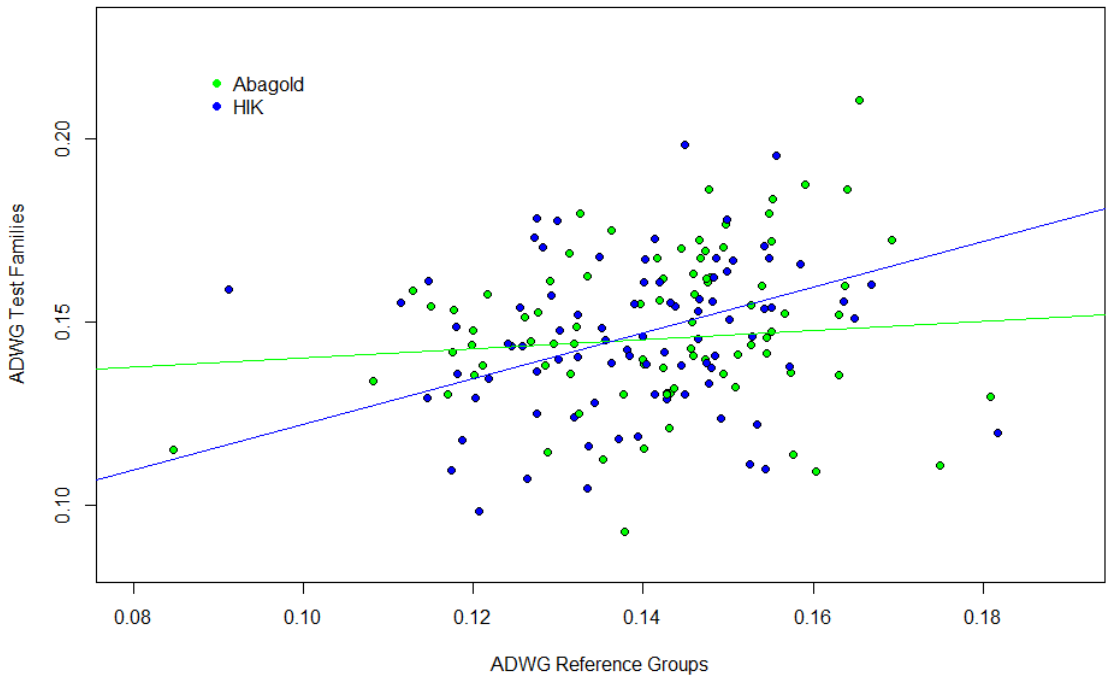

Figure 4.5 Scatterplot of average daily weight gain (ADWG) (g/day) of reference groups vs test families at each location grown in the same basket (solid points), correlation between ADWG of reference groups and test families at each location (solid line).

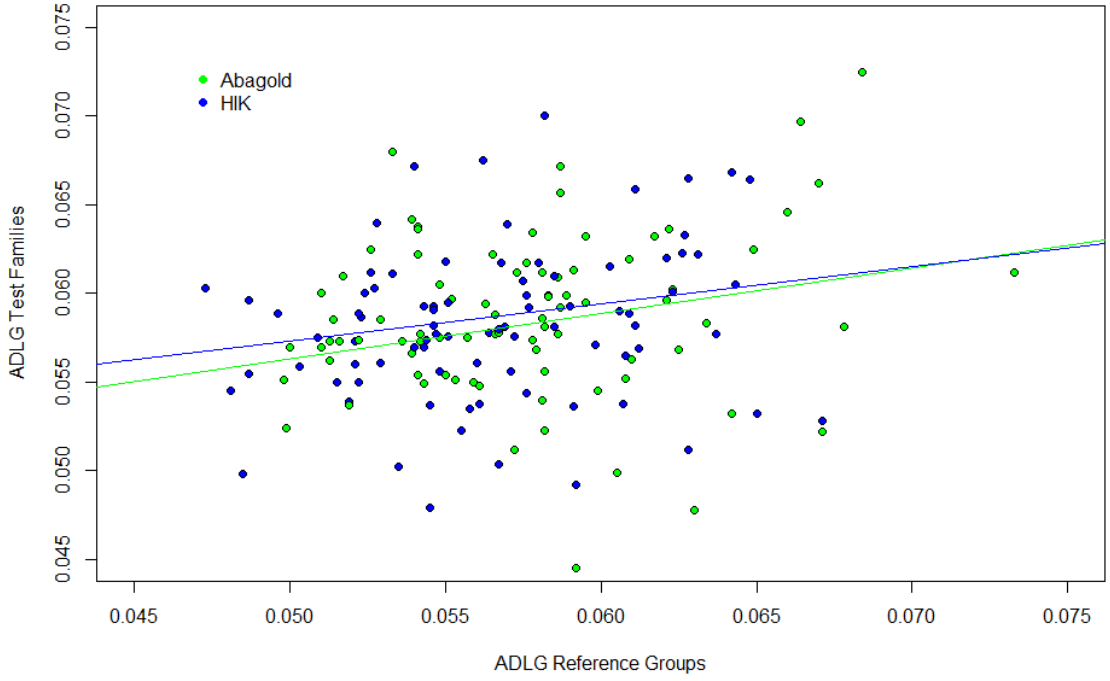

Figure 4.6 Scatterplot of average daily length gain (ADLG) (mm/day) of reference groups vs test families at each location grown in the same basket (solid points), correlation between ADLG of reference groups and test families at each location (solid line).

4.2 Evaluation of Replicated Test Groups and Internal Reference Groups

4.2.1 Genotype by Environment Interaction

As six repeat test families are in duplicate at location Abagold but are not replicated at location HIK the inclusion of an interaction term for genotype by environment interaction in a two-way ANOVA table is not possible. None the less, a graphical representation of test family performance over location for ADWG and ADLG by means of an interaction plot (R version 2.15.2) is possible. However, rank order changes in performances of families over location, usually indicative of genotype by environment interaction, cannot be interpreted as such because the possible effects of random uncontrollable environmental factors cannot be ruled out. The interaction plots of ADWG and ADLG of repeat test families can be found below in figure 4.7 and figure 4.8, respectively.

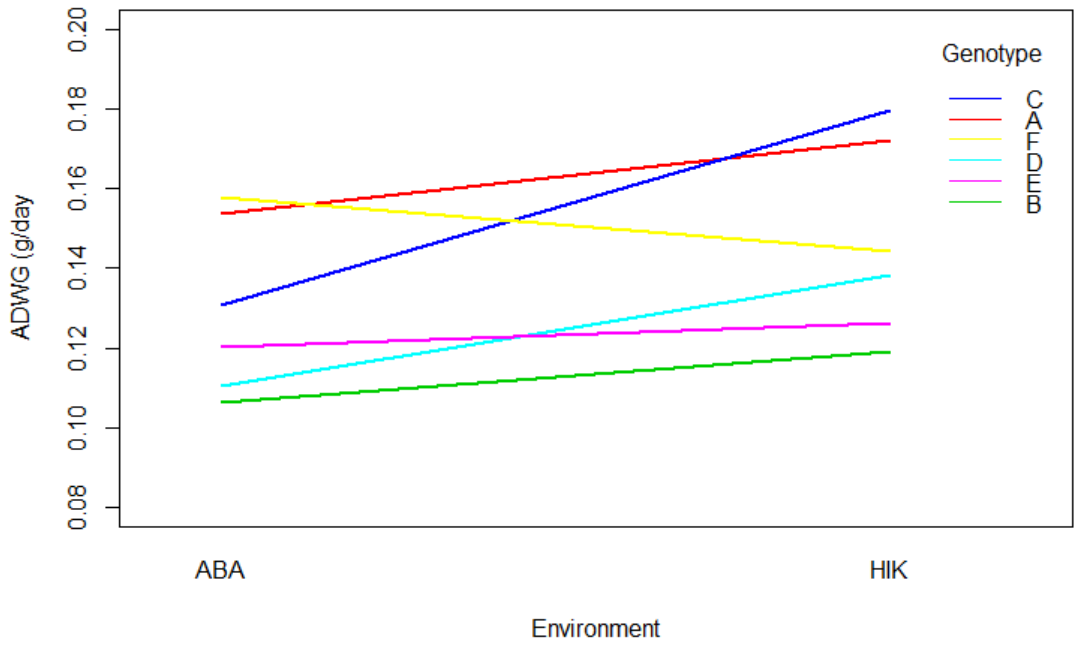

Figure 4.7 Interaction plot of test family ADWG by environment.

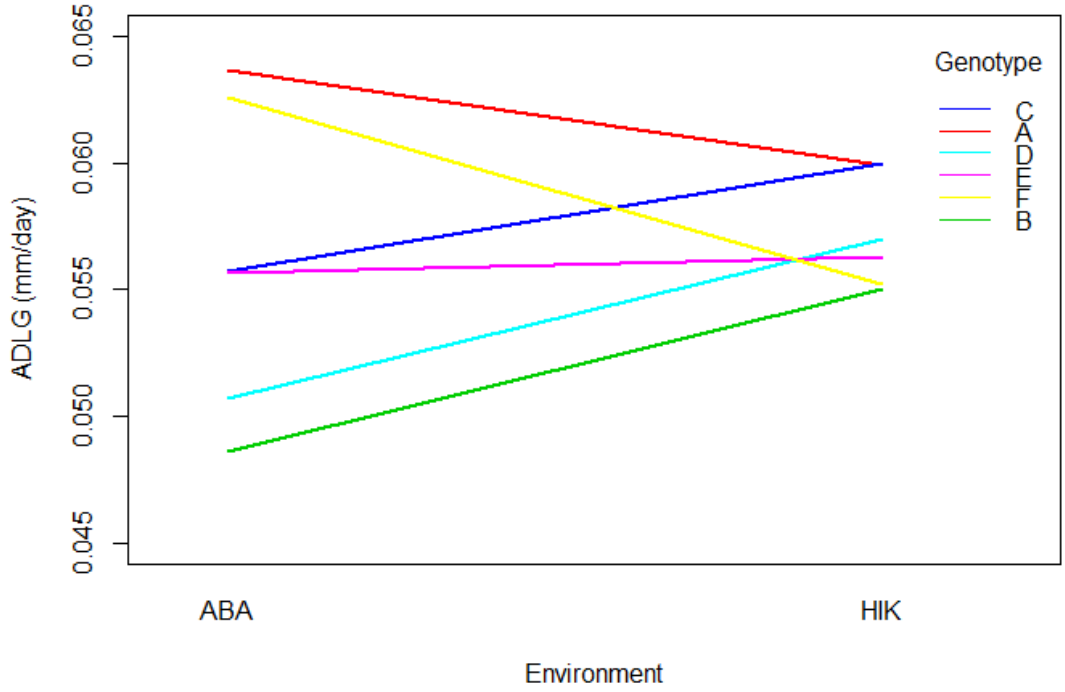

Figure 4.8 Interaction plot of test family ADLG by environment.

4.2.2 Optimal use of Reference Group

In order to determine if either the use of the reference group for manual correction of test family performances or as a concomitant covariate has any statistical merit in reducing environmental variance, it would be pertinent to compare these methods to the conventional two-way analysis of variance of family performances without the use of reference group

information. To this end, the repeat test family and respective internal reference group ADLG and ADWG was tested for deviations of the residuals of the fitted models to follow, from normality by means of a Shapiro-Wilk test, the results of which were all not significant (Shapiro & Wilk, 1965). In addition, the variances of factor groups of the fitted models to follow were tested for homoscedasticity by means of Levene's test for equality of variances, which were also not significant (Levene, 1960).

The assumptions being met, a two-way analysis of variance (ANOVA) was performed using the General Linear Models (GLM) procedure of SAS 9.1, to determine family and location effects on the ADLG and ADWG of test families (SAS Institute, 2011). The following model was fitted for the main effects (Family and Location):

$$y_{ijk} = \mu + \alpha_i + \beta_k + \varepsilon_{ijk}$$

Where Y_{ijk} is the j^{th} observation of the i^{th} treatment (Family) at the k^{th} location, μ is the common mean, α_i is the family effects,β_k is the location effects and ε_{ij} is the random error. The results of which can be found below in Table 4.8(a) and Table 4.9(a) for ADLG and ADWG, respectively.

In addition, the test family performances where corrected by means of directly subtracting their respective internal reference group performances to obtain corrected ADWG and ADLG of test families. When the reference group information is utilised in this manner, the assumptions that no genotype by environment interaction or antagonistic interactions have occurred, must be met. Due to the lack of adequate replication, neither of these assumptions can be tested. The two-way analysis of variance was conducted utilising the General Linear Models (GLM) procedure of SAS 9.1, to determine family and location effects on the corrected ADLG and ADWG of test families (SAS institute, 2011). The same model was fitted as the model mentioned above for the uncorrected test family performances. The results of which, are summarised in Table 4.8(b) and Table 4.9(b) below.

Table 4.8 Two-way ANOVA for average daily length gain (ADLG) of replicated test families: uncorrected (a) and corrected (b).

Average Daily Length Gain (ADLG)					
Source	d.f	Mean square	*F*-ratio	*P*	R^2
(a) *Uncorrected Family Performance*					
Family	5	$2.85x10^{-4}$	2.664	0.082	0.31
Location	1	$4.76x10^{-6}$	0.222	0.647	
Error	11	$2.36x10^{-4}$			
(b) *Corrected Family Performance*					
Family	5	$3.76x10^{-4}$	1.993	0.158	0.41
Location	1	$2.87x10^{-4}$	7.618	0.019*	
Error	11	$4.15x10^{-4}$			

* Significant at 0.05
** Significant at 0.01
***Significant at 0.001

Table 4.9 Two-way ANOVA for average daily weight gain (ADWG) of replicated test families: uncorrected (a) and corrected (b).

Average Weight Length Gain (ADWG)					
Source	d.f	Mean square	*F*-ratio	*P*	R^2
(a) *Uncorrected Family Performance*					
Family	5	$1.26x10^{-3}$	5.695	0.001^{***}	0.62
Location	1	$1.10x10^{-3}$	4.993	0.047^{*}	
Error	11	$2.21x10^{-4}$			
(b) *Corrected Family Performance*					
Family	5	$3.12x10^{-3}$	15.59	$1.14x10^{-4***}$	0.84
Location	1	$3.93x10^{-3}$	20.32	$8.90x10^{-4***}$	
Error	11	$1.93x10^{-4}$			

* Significant at 0.05
** Significant at 0.01
***Significant at 0.001

In order for the reference group performances to be utilised as a concomitant covariate in an analysis of covariance, the additional assumption of common slope between test family performances and reference group performances is required. The General Linear Models procedure (GLM) in SAS 9.1 was utilised to determine communality of slopes between test family and reference group performances for ADLG and ADWG (SAS institute, 2011). The following model was fitted for interaction between main effects (Family) and covariate (Reference):

$$y_{ijk} = \mu + \alpha_i + \beta_k + \gamma_{ik} + \alpha\gamma_{ik} + \varepsilon_{ijk}$$

Where Y_{ijk} is the j^{th} observation of the i^{th} treatment (Family) at the k^{th} location. While γ_{ik} is the mean performance of the reference group grown with the i^{th} treatment in the j^{th} environment and $\alpha\gamma$ is the interaction term. μ is the common mean, α_i is the family effects, β_k is the location effects and ε_{ij} is the random error. The results of which can be found below in Table 4.10(a) and Table 4.11(a) for ADLG and ADWG, respectively.

The following covariance model was fitted for the interpretation of main effects of (Family) and the covariate (Reference):

$$y_{ijk} = \mu + \alpha_i + \gamma_i + \varepsilon_{ijk}$$

where Y_{ij} is the j^{th} observation of the i^{th} treatment (Family). While γ_{ik} is the mean performance of the reference group grown with the i^{th} treatment. μ is the common mean, α_i is the family effects and ε_{ij} is the random error. The results of which can be found below in Table 4.10(b) and Table 4.11(b).

Table 4.10 ANCOVA of average daily length gain (ADLG) of test families with covariate reference group: Interaction model (a) and Covariance model (b).

Average Daily Length Gain (ADLG)					
Source	d.f	Mean square	*F*-ratio	*P*	R^2
(a) *Interaction model*					
Family	5	5.72×10^{-5}	2.50	0.152	0.25
Reference	1	2.16×10^{-6}	0.09	0.771	
Family x Reference	5	1.98×10^{-5}	0.85	0.853	
Error	6	2.33×10^{-4}			
(b) *Covariance model*					
Family	5	5.72×10^{-5}	2.635	0.084	0.30
Reference	1	2.16×10^{-6}	0.1	0.748	
Error	11	2.17×10^{-5}			

* Significant at 0.05
** Significant at 0.01
***Significant at 0.001

Table 4.11 ANCOVA of average daily weight gain (ADWG) of test families with covariate reference group: Interaction model (a) and Covariance model (b).

Average Daily Weight Gain (ADWG)					
Source	d.f	Mean square	*F*-ratio	*P*	R^2
(a) *Interaction model*					
Family	5	$1.26x10^{-3}$	11.66	$4.80x10^{-3**}$	0.81
Reference	1	$2.04x10^{-4}$	1.90	0.218	
Family x Reference	5	$5.36x10^{-4}$	4.97	0.038*	
Error	6	$1.08x10^{-4}$			
(b) *Covariance model*					
Family	5	$1.26x10^{-3}$	4.16	0.023*	0.48
Reference	1	$2.15x10^{-4}$	0.68	0.428	
Error	11	$3.01x10^{-4}$			

* Significant at 0.05
** Significant at 0.01
***Significant at 0.001

Statistical power analyses where conducted on the non-significant family factors for ADLG in Table 4.8 above, by means of the power anova test function in R version 2.15.2 (R Core Team, 2012). The results of which concluded the power of the anova F-ratio test for family effects on uncorrected family performances was approximately 0.76. The power of the same test on the corrected performance was approximately 0.61. By means of substituting error terms, it is estimated that approximately four replications for each family utilising corrected performances is required to attain the same power as the uncorrected performances.

The efficiency of the internal reference group as a covariate in the analysis of covariance was estimated as the ratio of the error variance without adjustment and the adjusted error variance. For the relevant equations for calculating the error variance after adjustment see Steel *et al*, (1997). The efficiency of the reference group adjustment was calculated as 0.98 and 11.72 latter being that of ADLG and the former being ADWG.

4.2.3 Correlation between Repeat Family Performances and Reference Group Performance

The correlation between repeat test family performances and internal reference group performances were calculated by means of pearson's correlation coefficient for ADWG and ADLG, -0.24 and -0.06 respectively (Pearson, 1986). Scatterplots of test family (A through F) performances vs internal reference group performances for ADWG and ADLG can be found in Figure 4.9 and Figure 4.10 below. In addition, the correlation was fitted to scatterplots by means of least squares regression in R version 2.15.2 (R Core Team, 2012).

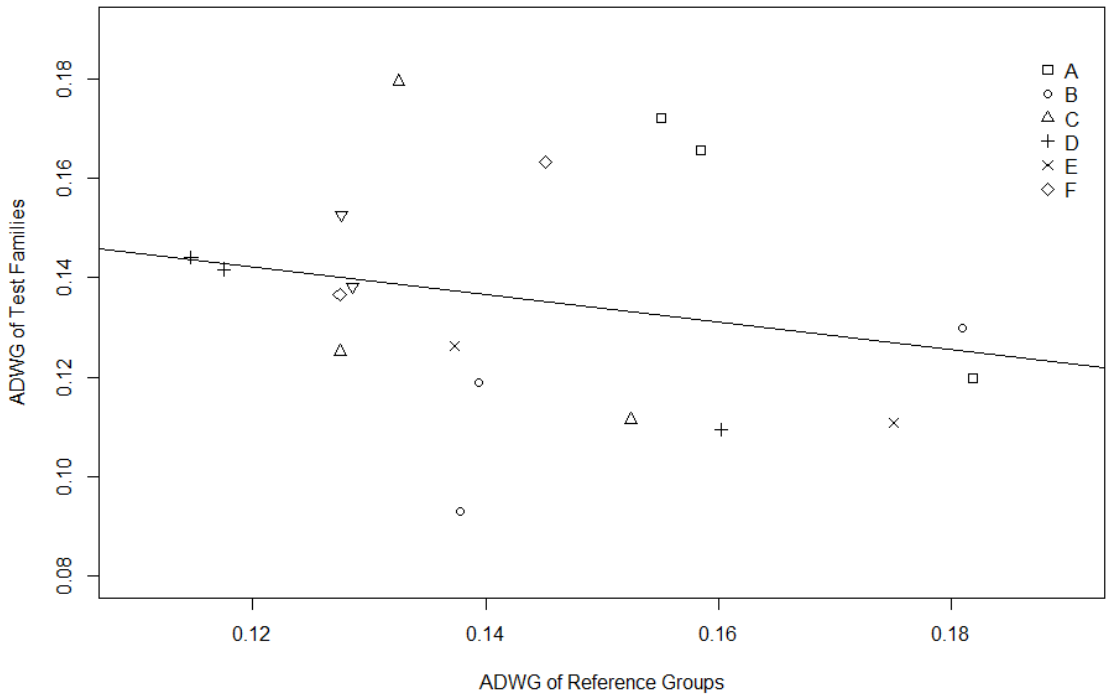

Figure 4.9 Scatterplot of average daily weight gain (ADWG) (g/day) of reference groups vs test families (A through F) at each location, grown in the replicate environment (solid points), correlation between ADWG of reference groups and test families at each location (solid line).

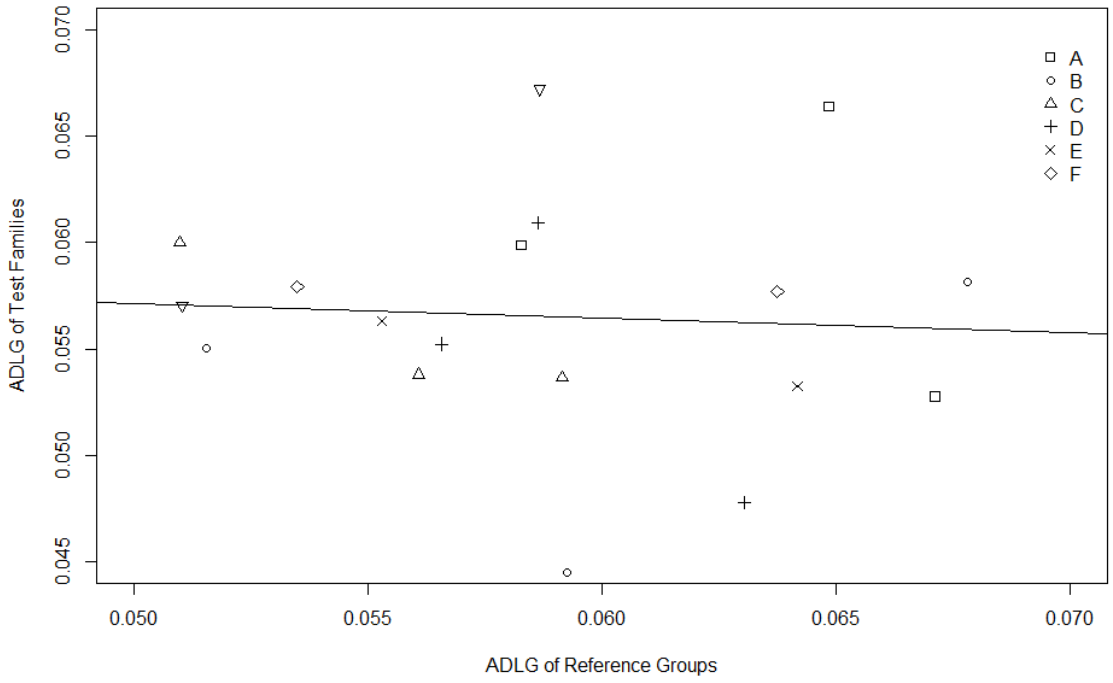

Figure 4.10 Scatterplot of average daily length gain (ADLG) (g/day) of reference groups vs test families (A through F) at each location, grown in the replicate environment (solid points), correlation between ADLG of reference groups and test families at each location (solid line).

4.3 Summary

The unreplicated test family and internal reference group data was utilised to evaluate the validity of implementing the internal reference technique within the Innovation Fund Abalone Breeding Project. It was found that tag loss is a limiting factor towards accurate estimation of family performances but does not bias family performances. It is recommended that reference animals are retagged at 6 months intervals. The comparison of internal reference performances over locations indicates location effects are significant but evidence suggests an external factor may be inflating the estimation of location effects. Without replication no estimation of antagonistic interactions or genotype by environment interactions was possible. However, there were significant differences in the performances of size ratio classes of internal reference groups to test groups at initial co-stocking. Visual evaluation of performances of size ratio classes reveals internal reference groups which were smaller than their respective test families had reduced performances. The phenotypic correlations between test families and internal reference groups were moderate in magnitude and positive, suggesting there are random environmental factors within experimental units affected both groups in a similar manner although, the moderate to weak strength of correlations further indicates the effects of an external factor which affects internal reference groups and test families differently.

The replicated test families and respective test families were utilised to test for genotype by environment interactions and antagonistic interactions as well as evaluate the relative efficiency attained by various methods of utilising reference group information to reduce replicate variance. The replication proved to be insufficient to test for genotype by environment interactions as test family performances were confounded with inter-rearing structure effects at one of the locations. For ADWG significant antagonistic interactions between test families and reference groups was detected, for ADLG measured on the same individuals no significant antagonistic interactions were detected. The manual correction of test family performances utilising internal reference performances resulted in marginal increases in the adjusted coefficient of determination. An examination of the change in replicate variance due to the correction revealed that for most families yielded a reduction in replicate variance, while for some of the families the replicate variance increased, further suggesting antagonistic interactions are introducing bias into internal reference group performances. It was estimated by statistical power analysis that the manual correction utilising the reference group information requires an extra replicate to achieve the same sensitivity as the analysis of variance without the correction. The use of the internal reference group as a covariate yielded a negligible change in the adjusted coefficient of determination for ADLG and a large decrease for ADWG. The efficiency of the internal reference adjustment was estimated as 0.98 for ADWG and 11.72 for ADLG. The correlation

between repeat test family performances and internal reference groups was very weak in magnitude and negative, a clear indication that one of the groups benefits from supressing the other group's growth.

4.4. References

Basiao, Z.U. & Doyle, R.W. 1990. Use of internal reference population for growth rate comparison of tilapia strains. I. In a crowded environment. In: Hirano, R. & Hanyu, I., (Eds.), *The Second Asian Fisheries Forum. Manila: Asian Fisheries Society*. pp. 503-506.

Kirpichnikov, V.S. 1966. Methods of progeny testing spawners in carp fish farms*. Selective Breeding of Carp. Israel Program for Scientific Translations*. **5737**: 26-55.

Kruskal, W.H., & Wallis, W.A. 1952. Use of ranks in one-criterion variance analysis. *Journal of American Statistical Association*. 47: 583-621.

Levene, Howard (1960). "Robust tests for equality of variances". In Ingram Olkin, Harold Hotelling*, et alia*. Contributions to Probability and Statistics: Essays in Honour of Harold Hotelling. Stanford University Press. pp. 278–292.

R Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

SAS Institute Inc. 2011. Base SAS® 9.3 Procedures Guide. Cary, NC: SAS Institute Inc.

Shapiro, S.S., & Wilk, M.B. 1965**.** An analysis of variance test for normality (complete samples). *Biometrika*, **52**: 591-611.

Steel, R. G. D., Torrie, J. H., & Dickey, D. A. 1997. *Principles and procedures of statistics: A Biometric Approach*. 3rd edition. McGraw-Hill, New York, USA. ch 17.6.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin*. **1**(6): 80-83.

Pearson, K. 1896. Mathematical Contributions to the Theory of Evolution. On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. *Proceedings of the Royal Society of London*, **60***:* 359-367.

Chapter 5

DISCUSSION AND CONCLUSION

5.1 Evaluation of Unreplicated Test Groups and Internal Reference Groups

5.1.1 Tag Loss

A vital aspect of implementing the reference group method is that reference groups can be distinguished from test groups within the same basket (Kirpichnikov, 1966). Within this growth trial a total of 50 tagged reference group individuals from a single full sibling family were included within each experimental unit (basket) and random samples were drawn at 6 month time intervals for measurement. It was observed that near the end of the 3 year growth trial the number of tagged individuals within each experimental unit had decreased to below 16 for a large proportion of the experimental units.

To estimate the tag losses through time, the samples sizes where analysed as the dependent variable, as no information regarding the exact number of tagged individuals over and above the samples sizes drawn from within each basket at each measurement interval was recorded. The decrease in samples sizes reflected the combined effects of the following factors: failed adhesion of tags, encrusting organisms masking tags, damaged tagged individuals due to handling, mortalities, escape and human error. The Kruskal-Wallis non-parametric ANOVA proved highly significant ($P<0.001$) for differences in samples sizes between measurement intervals (Table 4.3) and pairwise comparisons were significant for all comparisons ($p<0.001$) except that of 6-12 months after tagging ($p=0.07$) (Table 4.3). As can be seen in Table 4.2 and Figure 4.1, even at 0 months after tagging some samples sizes where below 16, this is attributed to human sampling error. As can be seen in Figure 4.1, the number of samples below 16 and 10, which correspond to a respective loss of 68% and 80% of tagged individuals with each basket, greatly increases after 12 months.

Even though the factors influencing tag loss can be seen to act randomly and not bias estimates, they do increase the standard error of mean estimations. To prevent this scale of tag loss in future trials implementing the bee tag method of tagging, it would be strongly advised to re-tag at intervals of 6 months. Regardless of the fact that there was no significant differences between sample sizes of 6 to 12 months after tagging, as sample size numbers fail to reflect the true tag loss of the remaining tagged individuals not drawn into samples within each basket.

5.1.2 Location Effect

Within the context of the abalone breeding program as a whole, past practice has been to conduct growth trials over all five locations and make selections from within each location

(Brink *et al*., 2011). From the perspective a combine family selection program, selection over locations would further improve between family selection intensity and therefore genetic response (Gjedrem, 2005). Prior to the present study, no single, full sibling family has resulted in the settlement numbers large enough for the replication required to estimate location effects over all five locations. Within the present study, a single full sibling family was included within all experimental units, from all five locations. Unfortunately, due to biosecurity concerns, the families spawned over the 2007-2008 spawning season of all five locations were tested at two locations. None the less, evaluation of location effects between the two locations provides insight into the statistical control required for future trials over multiple locations.

Significant differences were observed between ADLG and ADWG of internal reference groups for each location. By means of a log transformation the ADLG data met the requirements of a parametric t-test ($p<0.001$). However, no transformations of ADWG met the requirements of parametric testing and was thus tested by means of a non-parametric Wilcoxon signed-rank test ($p<0.001$) (Table 4.5). An examination of the coefficients of variation for ADLG and ADWG of internal reference groups at each location sheds some light onto the difficulties in meeting the requirement of parametric testing. For location Abagold the CV for ADLG is 7.8% and for ADWG the CV is 13.4%, which is not what one would expect considering Pearson's 2nd dimension rule (Pearson, 1896; Zar, 1984). Furthermore, the CV for ADLG and ADWG at location HIK are 5.1% and 6.5%, respectively (Table 4.4). When we consider that the internal reference group is a single full sibling family, we can see that these discrepancies cannot be entirely be due to within family genetic variation and must be attributed to environmental factors. The reason for these differences is that while stringent systems were implemented to maintain the same animal husbandry practices within each location, husbandry practices were not standardised over locations. While geographically and oceanographically these two locations do not differ, a host of factors such as stocking densities, management practices, systematic conditions and nutrition are confounded with location.

When the above mentioned coefficients of variation are compared with that of the test families at each location, it can be seen that the coefficients of variation of test families are consistently higher, albeit marginally, the only exception being that of ADLG at location HIK. Moreover, as the differences between test families and internal reference groups within each location can be attributed almost entirely to genetic factors, a rough indication emerges that the environmental contributions to variation in these traits is proportionally large. Although, interpreting the differences in coefficients of variation in this manner must be approached with caution. Most notably for ADLG at location HIK, where the CV of the internal reference group (6.1%) is larger than that of the test families (5.1%). As the internal reference group is

the genetic control, for the CV to be greater than that of the test families grown in the same experimental units, is indicative of some external factor, for instances antagonistic interactions between test and reference groups or genotype by environment interaction. This discrepancy must be closely monitored throughout the trial and noted for future trials implementing the internal reference method.

Interestingly, for ADLG the mean performances of internal reference groups was greater at location Abagold ($5.91x10^{-2}$ mm/day) compared with that of location HIK ($5.47x10^{-2}$ mm/day), juxtaposed to ADWG where the mean performances where greater at location HIK ($1.45x10^{-1}$ g/day) compared to that of location Abagold ($1.37x10^{-1}$ g/day) (Table 4.4). This is further illustrated by the differences in phenotypic correlation between ADLG and ADWG at each of the locations, indicated by the solid lines fitted to the spread of data in Figure 4.2. Again these differences in phenotypic correlation can be explained by differences in management procedures such as stocking density, which effect growth but do not alter the profitability of the business model (Brink *et al*., 2011).

5.1.3 Antagonistic Interaction Effects and Initial Size Variability

The inherent lack of replication of internal reference groups with the same test families, with the exception of the six repeated tests families to be discussed later, makes the testing for test family to internal reference group interactions impossible. To date no studies have investigated the effects of antagonistic interactions between abalone due to size differences. Although, the management at the five locations have observed growth dispensations within baskets due to larger abalone out competing smaller abalone for resources (Brink *et al*., 2011). The differences in spawning dates of test families and the internal reference family has led to size differences at initial co-stocking. The size ratio of internal reference group to respective test families at co-stocking was calculated and analysis of variance conducted to determine if there are any significant differences in ADLG and ADWG of internal reference groups co-stocked with varying sizes of test families. To eliminate the confounding effects of location with that of size ratio classes, the effects of co-stocking internal reference groups with test families of different sizes was conducted separately for each location.

For ADLG and ADWG of internal reference groups at location Abagold the eight size ratio classes at initial co-stocking, ranged from the internal reference group being 80% the size of test families to 120% the size of test families. For both ADLG and ADWG the effect of co-stocking the internal reference groups with test families of difference size ratio classes proved non-significant and only explained 8% and 13% of the variance within the respective models (Table 4.6 and Table 4.7). However, for ADLG and ADWG of internal reference groups at location HIK the 5 size ratio classes at initial co-stocking, ranged from internal reference group being 80% the size of test families to 105% the size of test families. For

both ADLG and ADWG the effects of co-stocking the internal reference groups with test families of different size ratio classes proved significant with ADWG ($p<0.05$) and ADLG ($p<0.001$). In addition the size ratio classes for ADWG and ADLG explained 15% and 46% of the variance in the model. The highly significant p value for ADLG at HIK and the proportionally larger amount of explained variance is consistent with the inflated CV discussed above and supports the theory that antagonistic interactions maybe present.

It would prove informative for future trials implementing the internal reference method if, pairwise comparisons of size ratio classes could be utilised to set upper and lower boundaries for size differences at initial co-stocking and furthermore could be utilised in future trials to prevent any antagonistic interactions. However, the highly variable number of observations in each size ratio class statistically detracts from the calculation and interpretation of least square means comparisons (LS means). Nevertheless, the visual evaluation of boxplots for size ratio classes at location HIK for ADLG (Figure 4.3) and ADWG (Figure 4.4) proves informative towards this end. A trend emerges that the smaller the internal reference group to test family ratio at co-stocking the more reduced the internal reference performance. Upon visual evaluation as a rough guideline, it would appear a deviation of 5% either side of the 1:1 size ratio will not yield antagonistic effects. This must be noted for future trials implementing the internal reference method in *Haliotis midae*

5.1.4 Correlation between Unreplicated Test Family Performance and Reference Performance

The measure of correlation between test families and their respective internal reference groups is an indication of how successful the internal reference group is, as an internal measure of random environmental factors in each experimental unit (Doyle *et al*., 1990). As random uncontrollable environmental factors (for instance food availability due to deceased numbers as a result of mortalities or higher water temperatures), will affect the test family and reference group in similar ways leading to positively correlated growth (Basiao & Doyle, 1990). The correlations of test family to reference family growth for ADWG and ADLG were positive at both locations, albeit weak. Interestingly, the correlations for ADWG were more variable (0.1 at Abagold and 0.33 at HIK) than those of ADLG (0.21 at Abagold and 0.24 at HIK), as is highlighted by the fitted lines in Figure 4.5 and Figure 4.6. These estimates are far lower than those found in strain experiments of common carp (*C.carpio*), Nile tilapia (*O.niloticus*) and rainbow trout (*O.mykiss*) in multiple environments utilising the internal reference method, where the correlation estimates ranged from (0.83 - 0.97), (0.56 – 0.90) and (0.80) in each respective experiment (Basiao & Doyle, 1990; Vandeputte *et al*., 2002;Blanc *et al*., 1983). These weak positively correlated performances could further

indicate the effects of factors such as antagonistic interactions, although without replication it is not possible to statistically verify this.

5.2 Evaluation of Replicated Test Groups and Internal Reference Groups

5.2.1 Genotype by Environment Interaction

Six of the test families achieved high enough settlement numbers to facilitate replication in triplicate over locations. Unfortunately, during the logistics of relocating families to the two locations, three of the replicates of each family were lost. Resultantly, the test families were replicated in duplicate at location Abagold and but not replicated at location HIK. The lack of adequate replication at location HIK precluded the inclusion of an interaction term in a two-way analysis of variance for family effects and location effects, utilised to test for genotype by environment interactions. The results of the interaction plots (Figure 4.7 & Figure 4.8) would normally indicate genotype by environment interaction, however, the lack of replication at location HIK confounds basket effects with location effects and random basket effects, therefore causality cannot be inferred for genotype by environment interaction.

5.2.3 Optimal use of Reference Group

Conventionally, when assessing the optimal implementation of the internal reference group method in significantly different environments five statistical models are utilised to validate key assumptions and efficiency of implementation (Basiao & Doyle, 1990; Basiao *et al*., 1996; Vandeputte *et al*., 2002; Romana-Eguia *et al*., 2010). The first of which is a two-way analysis of variance with an interaction term for genotype by environment interaction, as explicated above the inherent lack of adequate replication precluded this. While previous studies conducted within the abalone breeding program, to assess ADLG and ADWG (Vlok, 2012) and yield traits (Van Schalkwyk, H. J. 2012) have failed to find significant genotype by environment interactions, these studies implemented an experimental design where few test families were repeated over location and then in a cascade type design, where families that were replicated over locations where only done so over two of the five locations. According to Doyle *et al* (1990), even if there is a suspicion that genotype by environment interactions have occurred, it is necessary to analyse environments separately when utilising the internal reference method. However, within the present study there is not enough replication to assess the efficiency of the internal reference group within a single location, so even if genotype by environment interactions occurred, it would not be possible to continue statistically analysing repeat family information. While the lack of replication detracts from the statistical merit of models, interpretation of models may still prove informative for future implementation of the internal reference method.

The second statistical model utilised to evaluate key assumptions, is an analysis of covariance with an interaction term to assess whether any interaction has occurred between the test family and the internal reference group. As can be seen in Table 4.10a the interaction term for ADLG of test families and internal reference group is non-significant at the 0.05 significance level and therefore the internal reference group can be utilised and the main effects of models interpreted. However, for ADWG of test families and internal reference group is significant (p=0.038) (Table 4.11a), therefore antagonistic interactions have occurred between test families and internal reference group and the internal reference group cannot be implemented and interpreted with any certainty. It is pertinent to note that while no significant interactions between test families and internal reference groups were detected for ADLG, but were detected for ADWG, these traits were recorded on the same individuals and so any interpretation of statistical models implementing internal reference information is likely to be unreliable. For the purpose of the present study it is pertinent to continue with interpretation of statistical models to identify further flaws to be considered in future growth trails implementing the internal reference method and to determine how interaction between test families and internal reference groups affects the statistical analyses.

Without the utilisation of any internal reference group information a two-way analysis of variance to test the main effects of family and location on ADLG and ADWG of test families was conducted. For ADLG of test families, the effects of family and location where not significant at the 0.05 significance level and only explained 31% of the variance in the model (Table 4.8a). For ADWG of test families, the effects of family and location both proved significant at the 0.05 significance level and explained 62% of the variation within the model (Table 4.9a). Comparatively, once manual correction by subtraction of internal reference group performance directly from respective test family performance, the same model for main effects location and family was carried out. For ADLG the variance explained by the main effects in the model as determined by the adjusted coefficient of determination increased to 41% and the effect of location proved significant at the 0.05 significance level (Table 4.8b). A similar improvement was also observed with ADWG, where the total variance explained in the model increased to 84% and main effects for location and family proved significant at the 0.001 significance level. According to Baisiao & Doyle (1990) this increase in explained variance is only possible if the covariance between test families and internal reference families is positive and large, resulting in a decrease in the variance of corrected performances. However, in their study utilising internal reference groups toward investigating the performances of different strains of Nile tilapia (*Oreochromis niloticus*) in crowded environments, the variances of uncorrected performances were consistently lower than the variances of corrected performances, ranging from 76%-26% (Basiao & Doyle 1990). Within

the presently study, the variances of uncorrected performances of four of the test families were higher than the variances of corrected families for ADLG and ADWG (54%-13%). Interestingly, the correction by reference group performances for two of the test families actually resulted in the variances of corrected performances being greater than the variances of uncorrected performances (257%-308%). An examination of the size ratio of internal reference group to test family at co-stocking for these two families, revealed both were between 80%-85%, this could indicate the source of inflated variances is due to antagonistic interactions between these two test families and their internal reference groups.

An analysis of covariance utilising the internal reference group is usually more efficient than manual correction in reducing replicate variance due to environmental variation than manual correction utilising the internal reference group performances (Neter *et al*., 1985; Basiao & Doyle, 1990). The analysis of covariance testing for main effects family did not contain test for location effects as the growth of the internal reference group is used to describe environmental difference (Basiao *et al*., 1996). For ADLG the effects of family and the use of the internal reference group proved non-significant at the 0.05 significance level and for ADWG only family effects proved significant at the 0.05 significance level. Interestingly, for both the ANCOVA models of ADLG and ADWG the adjusted coefficient of variation values showed the model explained less variance than the two-way analysis of variance of uncorrected performances. For ADLG the ANCOVA explained 30% of the variance and the ANOVA of uncorrected performances explained 31%. Similarly, for ADWG the ANCOVA explained 48% of the variance and the ANOVA of uncorrected performances explained 62%. This is not surprising since, the key assumption of common slopes has been violated by interaction between test families and reference groups and the model has poor replication over locations. These results are contradictory to those found by Basiao *et al*, (1996) in a study to evaluate the performances of different strains of Nile tilapia (*Oreochromis niloticus*) to crowding stress, where the use of the reference fish increased the fit of the model from 6.2% to 90.8%. Similarly, in a study to evaluate strains of Nile tilapia fry to salinity stress the use of the internal reference fish increased the fir of the model from 10.6%-63.8% (Basiao *et al*., 2005). However, in both studies there was 8-10 replicates of each strain in each of the locations or environments and the assumptions of no genotype by environment interaction and test strain to internal reference group interaction where satisfied (Basiao *et al*, 1996; Basiao *et al*., 2005).

Within the present experimental design, the incorrect implementation of the internal reference group information can be seen to have negatively affected the validity of comparisons. The statistical power analysis of the non-significant family term in Table 4.10 for the uncorrected family performances was 0.76; effectively this means the probability of incorrectly stating there is no significant difference is 24%. The statistical power analysis of

the corrected test family ADLG performances decreased to 0.61. By substitution of error terms, it would appear instead of decreasing replicate number, increasing replicates to 4 will attain the same statistical power as that of the ANOVA with no correction by reference group information. This decrease in statistical power can be attributed to the increases in within family variances attributed to the few families where antagonistic interactions are believed to have occurred. The efficiency was decreased for ADWG by 2% when the internal reference was used as a covariate, but increased 11.7 fold for ADLG. The conflicting information regarding the decrease in statistical power of the manual correction and the increase in efficiency of the covariate adjustment for ADLG using internal reference information is a discrepancy which must be carefully monitored in future trials as it has drastic implications for which method of correction should be implemented.

5.3.3 Correlation between Repeat Family Performances and Reference Group Performance

The correlation between test families and respective internal reference groups are negative for both ADLG (-0.24) and ADWG (-0.06). However, a visual evaluation of the spread of data shows an even spread of the data around correlation 0 for most of the families (Figure 4.9 and 4.10). However, a small number of the families show a strong negative trend, which greatly detracts from the credibility of the reference group information and further indicates antagonistic interactions. As has been noted previously, one of each family basket is confounded with the HIK location effects, this should not affect the correlation of between internal reference and test family unless genotype by environment interactions have occurred. The correlations between test family and internal reference group are not in agreement with the studies discussed previously for the correlation between unreplicated test families and reference groups. However, these correlations are in agreement with the findings of Doyle et al, (1990) in a study designed to induce antagonistic interactions between test strains and internal reference strains through increased stocking densities, where high stocking densities resulted in moderate negative correlations.

5.3 Conclusion

According to Kirpichnikov (1966) and Basiao et al, (1990) the most important factors when implementing the internal reference method are:

- Reference group individuals must be absolutely distinguishable from test individuals
- Conditions prior to stocking must be the same for all individuals
- When testing in multiple environments, no genotype by environment interaction must occur
- The weights (size) of reference group individuals must be the same as test individuals at co-stocking

- The test or reference group must not benefit from suppressing the other (antagonistic interactions)

As revealed by the investigation into tag loss, the implementation within the present study does not always meet this standard. However, until alternative methods of tagging are developed, the current method will suffice provided reference group animals are re-tagged every six months. Additionally, there is no reliable method of tagging individuals prior to the commencement of the trial at two years of age. Fortunately, a previous study the Performance Recording Scheme has determined that there are no significant differences in the growth of abalone at each of the locations prior to the commencement of the trials, which satisfies the second factor listed above (Vlok, 2012).

It has been established that there are significant differences in the growth performances of abalone at the two locations in this study. Undoubtedly, statistical control is required to correct for location effects. As has been demonstrated in the present study, without adequate replication of experimental units over locations, it is not possible to ascertain if genotype by environment interactions have occurred and therefore verify the creditability of the internal reference group performances as a means of correcting for location effects.

The present experimental design failed to ensure the sizes of reference and test animals are the same at co-stocking, due to limitations on hatchery resources and technical difficulties in synchronised spawning. In addition, there is evidence to suggest reference groups that were sufficiently smaller than their test family counterparts had reduced performances. Furthermore, analysis of the limited replicated test families also suggests significant antagonistic interactions occurred between test families and references groups.

A holistic view of the comparisons of different methods of implementing internal reference group information suggests there is merit in implementing the reference group method toward reducing replicate variance. Provided the factors listed above are met, in particular that of antagonistic interactions.

5.4 Recommendations

Within the context of the present study, it is recommended that the selection of families be made within each location. As location effects are significantly different and it is not possible determine if genotype by environment interactions have occurred. In addition, it is recommended that the reference group information not be utilised to correct for basket differences, as there is evidence that antagonistic interactions between test family and reference groups has introduced a bias into the reference group performances. Even though basket differences are confounded with family differences, it would prove more accurate to fit

a regression line to test family performances of ADLG and ADWG at each location and select the highest ranking families along the regression line.

5.4.1 Recommendations for Future Growth Trials

As long as the unique set constraints within the abalone breeding program, such as limitations in individual identification and limitations in settlement numbers preventing replication at commercial stocking densities, hold true, basket differences will be confounded with family differences and as has been demonstrated by the large variability in growth performances of internal reference groups (albeit inflated due to antagonistic interactions), basket differences are large. Thus the replicate numbers required to reduce environmental variation will be high, if commercial stocking densities are to be maintained, the number of individuals per family required may very well exceed the settlement numbers achieved to date. In addition, the testing at multiple locations where management procedures are not all standardised, means that families must be replicated over locations to correct for location effects and ascertain if genotype by environment interactions occur. Furthermore, the lack of parental performance information due to brood stock acquired from the wild populations precludes effective implementations of mixed model equations.

Until these constraints are remedied there are two methods which can be implemented to reduce environmental variability and reduce replicate number. The first of which is communal testing, where families are reared separately until tagging is possible and families can be mixed within multiple baskets (Wohlfarth & Moav, 1985). However, this method corrects for growth based on initial sizes, utilising the multiple nursing method (Wohlfarth & Moav, 1972). The drawback to this method is that it requires at least 100 tagged individuals per family (Moav & Wohlfarth, 1976), with the constraint of tag loss; this method would require re-tagging every 6 months. In a trial such as the present with 173 families, the communal testing method would require the tagging of 173 000 individuals every 6 months. Additionally, the communal testing method forgoes the added within family selection intensity, as it cannot facilitate selection over multiple locations. Furthermore, genotyping of individuals is required when individuals lose tags, which is a further strain on resources such as time, personal, and funds (Moav & Wohlfarth, 1976).

The second method which can be implemented to reduce environmental variability and reduce replicate number is of course the internal reference method (Kirpichnikov, 1966, Moav & Wohlfarth, 1976). As has been determined in the current investigation, the spawning period must be reduced and steps taken to ensure the reference group is very nearly the same size as the test group it is co-stocked with and reference group individuals must be re-tagged every 6 months. Additionally, attention must be paid to location effects and satisfying the key assumption of no genotype by environment interactions.

At this point there are two scenarios which may be employed to better realise the potential of the internal reference group method. Firstly, test families which attain sufficient settlement numbers can be replicated over locations a manner similar to augmented p-rep designs (Williams *et al*., 2011). The benefits of this type of design are two-fold, they facilitate testing for genotype by environment interaction and they act as a failsafe for adjusting for location effects in the advent that internal reference groups experience antagonistic interactions. Additionally, augmented p-rep designs which are balanced will enable a more conclusive evaluation of the relative changes in statistical power associated with the internal reference manual correction and efficiency of internal reference group as a covariate. The second scenario which can be employed in the event test families do not realised adequate settlement numbers is to forego testing at multiple locations and conduct growth trials at a centralised location. Numerous breeding programs implement centralised testing stations, where breeders submit genetic material for testing. This has the benefit of reducing environmental variability, particularly when index selection and mixed model equations are utilised (Gjedrem, 2005; Jansens, 2012). Additionally, testing at a single location provides the flexibility to implement the internal reference methods or communal testing method.

5.4.2 Recommendations for Future Studies

Further development of synchronised spawning and reproductive technology is required. The success of synchronised spawning has successively improved since the commencement of the breeding program (Brink *et al*., 2011). However, for the most part most families only realise settlement numbers adequate for a single basket under commercial stocking densities (1000 individuals). Increases in settlement numbers will benefit the breeding program primarily in two ways, firstly it will aid replication under commercial stocking densities, which provides greater flexibility in implementing experimental designs and reducing the confounding of basket effects with family effects. Secondly it increases selection intensity and therefore genetic response (Gjedrem, 2005). Furthermore, improvements in the synchronisation of spawning events will facilitate the reduction in the spawning interval and thus size differences between test families as well as between test families and the internal reference group. Furthermore, improvements in the synchronisation of spawning events will provide greater flexibility of mating structures and therefore genetic parameter estimation (Falconer, 1981).

Further research and development of tagging methods are required. At present there is no method of tagging abalone prior to the commencement of the commercial grow-out phase (Brink *et al*., 2011), limited group identification is available for the commercial grow-out phase (Brink *et al*., 2009) and individual identification is available for the post-harvest traits (Van Schalkwyk, 2012). As can be seen in the breeding programs of multiple

aquaculture species, the introduction of high thorough-put, reliable tagging methods which facilitate individual identification, marks the transition to mixed model equations, genetic parameter estimation and breeding value estimation (Gjedrem, 2012).

For the continued use of the internal reference method there is a necessity to determine the upper and lower boundaries for size ratio differences at co-stocking which do not result in antagonistic interactions. The present study revealed that a possible 5% deviation in size may not result in antagonistic interactions, but this was not confirmed statistically by means of least square means comparisons. An experimental design which tests the size ratio differences at which antagonistic interactions occur would be particularly useful, especially if no further developments in synchronised spawning are achieved.

Further studies regarding genotype by environment interactions are required for the dispersal of improved genetic material from the abalone breeding program. The five locations in the abalone breeding program are all located within the Walker Bay region; there are abalone production units located in the warmer east coast of South Africa and colder west coast of South Africa and Namibia, which are not members of the Innovation Fund Abalone Breeding Program (AFASA, 2010). As was found in the Norwegian Salmon Breeding program, no significant genotype by environment interactions occurred over the entire country, this facilitated the use and the sale of the improved strains over the entire region (Gunnes & Gjedrem, 1978; Gjedrem, 2012). An analysis of the population genetic structure of *H.midae* in South Africa revealed differentiation between east and west coast populations (Van der Merwe, 2009), while a cross-breeding trial failed to find any significant heterosis (Voster, 2003), genotype by environment interactions may still occur. Furthermore with the increased interest of abalone ranching, any genotype by environment interactions may have economic implications for the sale of improved genetic strains on a royalty basis.

5.5 References

AFASA. 2010. Abalone Sector Report. Available from: http://www.aasa-aqua.co.za/site/wp-content/uploads/2010/12/Abalone-Sector-Report-August-2010-_1_1.pdf. [Accessed 24/10/12].

Basiao, Z.U. & Doyle, R.W. 1990. Use of internal reference population for growth rate comparison of tilapia strains. I. In a crowded environment. In: Hirano, R. & Hanyu, I., (Eds.), *The Second Asian Fisheries Forum. Manila: Asian Fisheries Society*. pp. 503-506.

Basiao, Z.U., Doyle, R.W., & Arago, A.L. 1996. A statistical power analysis of the 'internal reference' technique for comparing growth and growth dispensation of tilapia strains. *Journal of Fish Biology*. **49** : 277-286.

Basiao, Z.U., Eguia, R.V., & Doyle, R.W. 2005. Growth response of Nile tilapia fry to salinity stress in the presence of an 'internal reference' fish. *Aquaculture Research*. **36**:712-720.

Blanc, J.M., Chevassus, B., & Poisson, H., 1983. Utilisation de mutants ''golden'' comme temoins intra-lots pour le testage des performances de croissance chez la truite arc-en-ciel. *Cybium*. **7**(2): 93–103.

Blanc, J.M., Poisson, H., Vallée, F. 2001. Covariation between diploid and triploid progenies from common breeders in rainbow trout, Oncorhynchus mykiss (Walbaum). *Aquaculture Research*. **32**: 507-516.

Brink, D., Van Der Merwe, M., Vlok, A.C., Hepple, J., Slabbert, R., 2009. Tagging methodology for juvenile and adult abalone (*Haliotis midae*). Poster presentation during AASA International Aquaculture Conference, Swakopmund.

Brink, D., Van Der Merwe, A., Roodt-Wilding, R., Hatting, A., Yearsley, R., Landman, R., van Wyk, S., & Janse, L. 2011. Management of Innovation Fund abalone breeding program. Hermanus, South Africa. Personal communication.

Falconer, D.S. 1981. Introduction to Quantitative Genetics. New York: Longman.

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

Gjedrem, T. 2012. Genetic improvement for the development of efficient global aquaculture: A personal opinion review. *Aquaculture*. **344-349**:12-22.

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

Janssens, S. 2012. Research Unit Livestock Genetics, Division of Gene Technology, K.U. Leuven, Leuven Belgium. Personal communication.

Neter, J., Wasserman, W. & Kutner, M.H. 1985. Applied linear statistical models. Richard Irwin, Inc. Homewood, Illinois, USA. pp. 127.

Pearson, K. 1896. Regression, heridity, and panmixia. P*hilosophical Transactions of the Royal Society of London*. (A), **187**:253–318

Romana-Eguia, M.R.R., Ikelda, M., Basiaou, Z.U., & Taniguchi, N. 2010. Growth comparison of Asian Nile and red tilapia strain in controlled and uncontrolled farm conditions. *Aquaculture international*. **18**: 1205-1221.

Van der Merwe, A.E. 2009. Population genetic structure and demographical history of South African abalone, *Haliotis midae*, in a conservation context. Ph.D Thesis. Stellenbosch University, South Africa.

Van Schalkwyk, H. J. 2012. Assessment of yield traits between family groups of cultured abalone (*Haliotis midae*) in South Africa. M.S.c thesis. Stellenbosch University, South Africa.

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

Vlok, A.C. 2012. Growth assessment. Unpublished M.Sc. thesis. Stellenbosch University, South Africa

Voster, G. 2003. A Quantitative Genetic Analysis of the Effect of Crossbreeding on the Growth Rate of the South African Abalone, *Haliotis midae*. MS.c Thesis. Stellenbosch University, South Africa.

Williams, E., Piepho, H., & Whitaker, D. 2011. Augmented p-rep designs. *Biometric Journal*. **53** (1): 19-27.

Wohlfarth, G.W.,& Moav, R. 1972. The regression of weight gain on initial weight in carp: I. Methods and results. *Aquaculture*. **1**:7-28.

Wohlfarth, G.W.,& Moav, R. 1985. Communal testing, a method of testing growth rate of different genetic groups of common carp in earthen ponds. *Aquaculture*. **48**:143-157.

Zar, J.H. 1984. Biostatistical Analysis. Englewood Cliffs, NJ; Prentice-Hall.