Aquaculture

ASSESSMENT OF THE LEVEL OF HETEROZYGOSITY IN THE TASMANIAN ATLANTIC SALMON (*SALMO SALAR*) POPULATION USING SINGLE NUCLEOTIDE POLYMORPHISM MARKERS

S.Dominik[1, 2], J.M. Henshall[1,2], P.D. Kube[1,3], H. King[4], S. Lien[5], M. Kent[5] and N.G. Elliott[1,3]

[1]CSIRO Food Futures Flagship
[2]CSIRO Livestock Industries, Armidale, NSW 2350
[3]CSIRO Marine and Atmospheric Research, Hobart, TAS 7000
[4]Salmon Enterprises of Tasmania Pty. Ltd., Wayatinah, TAS 7140
[5]CIGENE, Norwegian University of Life Sciences, Aas, Norway

SUMMARY

The level of heterozygosity in the Tasmanian Atlantic salmon population was investigated using SNP data from 93 fish from unselected broodstock and 2991 single nucleotide polymorphisms (SNPs). Three approaches were used as measures of the likelihood of two alleles being identical by descent: two homozygosity indices, which have not been used on SNP data before and an inbreeding coefficient for the SNP markers based on observed and expected heterozygosity. As expected, the two homozygosity indices yielded highly correlated results, because biallelic SNP markers are all equally informative. The observed and expected heterozygosity at the SNP loci were very similar and that was reflected in intermediate inbreeding coefficients. All three analysis methods indicated a moderate likelihood that alleles are identical by descent, and offspring tested came from unrelated parents. Most markers were in Hardy-Weinberg-equilibrium. The SNP data support previous microsatellite studies concluding that the Tasmanian Atlantic salmon population has a moderate level of genetic diversity, as indicated by homozygosity indices and the inbreeding coefficient, and that hatchery protocols over the 40 years since the population was established in Australia have maintained a genetically healthy population.

INTRODUCTION

Salmon Enterprises of Tasmania Pty. Ltd. (SALTAS) embarked on a selective breeding program (SBP) in 2004 (Elliott and Kube 2009). One of the main factors determining potential genetic gain is the level of relatedness present in the population. Genetic diversity in the Tasmanian Atlantic salmon population was previously investigated based on sparse microsatellite markers (Innes and Elliott 2006). That study concluded that despite loss of alleles, heterozygosity had been maintained in the Tasmanian population when compared to its progenitor population. It was further concluded that sufficient genetic diversity was present to support a selective breeding program, but regular monitoring was recommended to detect any decline genetic diversity. Stochastic simulation studies by Henshall and Dominik (unpublished) modelled the history of the population and predicted the average rate of inbreeding to be below 1% per generation.

The aim of this study was to assess the level of heterozygosity in offspring of unselected broodstock of the founder generations of the SALTAS selective breeding program based on single nucleotide polymorphisms (SNP). Three approaches including two measures of homozygosity, which have not previously been applied to SNP data, are used.

MATERIALS AND METHODS

The Tasmanian Atlantic salmon population. The history of the Australian Atlantic salmon population is described in detail in Reilly *et al.* (1999). Two key events led to the establishment of the Tasmanian Atlantic salmon population. The first event was the importation of around 400,000

ova from Nova Scotia, Canada into Gaden, New South Wales (NSW), Australia between 1963 and 1968. The second event occurred between 1984 and 1986 when 570,000 ova were brought from NSW into Tasmania leading to the establishment of the SALTAS hatchery. These imports formed the foundation of Australia's leading aquaculture industry based in Tasmania. Following those events the population has been closed to any further importations and SALTAS has been producing broodstock and smolt for the industry over the last 20 years.

Sample preparation and genotyping. DNA was extracted from fin-clippings of 95 offspring across three year classes of founder families (2004 – 2006) from the SALTAS selective breeding program. Fish were chosen to be unrelated with the exception of the inclusion of the parents of one of the individuals. DNA was diluted in TE (10mM Tris pH 7.5, 1mM EDTA). A minimum of 20ul per sample at 50ng/ul (total = 1ug DNA) was prepared to facilitate pipetting.
The 95 samples plus one control sample were genotyped using a 15,225 SNP iSelect Atlantic salmon chip developed by Illumina (www.illumina.com) for the Centre for Integrative Genetics (CIGENE; www.cigene.no) as described by Kent *et al.* (2009). BeadStudio software (Illumina, 2007) was used to generate the allelic data, based on identifiable clusters of the normalised intensity data. SNPs were classed into polymorphic and monomorphic SNPs in unique and duplicated genomic regions. Only SNPs from unique regions were used for further analysis.

Measures of identity by descent. Two approaches that provide a measure of the likelihood of alleles to be identical by descent at the individual level were used.

Internal relatedness Index. The measure of internal relatedness (IR) was developed by Amos *et al.* (2001). It accounts for differences in allelic frequencies. IR compares pairs of loci based on their number of homozygotes and frequencies of alleles. IR ranges from IR = -1 for complete heterozygosity to IR = 1 for complete homozygosity of an individual across all loci.

$$IR = \frac{(2H - \sum f_i)}{(2N - \sum f_i)}$$

with H = number of homozygote loci
N = the total number of loci
f_i = frequency of the ith allele contained in the genotype

Homozygosity by loci Index. Aparicio *et al.* (2006) used the homozygosity by loci (HL) index to improve on the Amos *et al* (2001) approach and to reflect differences in the level of information of loci; the IR was suggested to underestimate heterozygous individuals carrying rare alleles (a possible issue with SNP data). HL = 0 indicates heterozygosity across all loci and HL = 1 reflects complete homozygosity.

$$HL = \frac{\sum E_h}{\sum E_h + \sum E_j}$$

with E_h = expected heterozygosity of loci that an individual carries in homozgyosity
E_j = expected heterozygosity of loci that an individual carries in heterozygosity

Coefficient of inbreeding. The simplest parameter of Wright's F-statistics (Wright 1951) was used. The inbreeding coefficient (F) describes the probability that two alleles at a locus taken at

random in a population are identical by descent. F can range between -1, when there is no inbreeding, to 1 if all alleles are identical by descent.

$$F = \frac{(H_e - H_o)}{H_e}$$

with H_e = expected heterozygosity based on allele frequencies
H_o = observed heterozygosity

RESULTS AND DISCUSSION

Of the 15,225 SNPs contained on the Atlantic salmon SNP chip, 2991 markers were found to be both polymorphic and present in unique genomic regions in our samples. Since the SNP chip was developed using data from the Norwegian breeding population, it was expected that only a proportion of the markers would be polymorphic in the Tasmanian Atlantic salmon population. The genotyping results showed very few missing values with a mean of 92.7 samples per SNP and on average 2982.5 SNPs per sample genotyped.

Table 1. Summary statistics of the homozygosity measures (IR, HL), for the observed and expected heterozygosity (H_o and H_e), the allele frequency for one of the alleles of each SNP (Allelefreq) and the inbreeding coefficient of the loci (F) for 93 samples of Atlantic salmon.

	Mean	Maximum	Minimum	Standard deviation
IR	-0.01	0.15	-0.20	0.05
HL	0.64	0.69	0.54	0.02
H_o	0.25	1.00	0.00	0.18
H_e	0.25	0.50	0.01	0.17
Allelefreq	0.46	0.99	0.01	0.35
F	-0.01	1.00	-1.00	0.15

Alleles cover the possible range of frequencies with some SNPs having rare alleles, which are reflected in the maximum and minimum values of the Allelefreq (Table 1). 87.5% of the polymorphic SNPs were in Hardy-Weinberg-Equilibrium (HWE) as assessed by a Chi-square test ($P< 0.05$), which is confirmed in Figure 1. It shows that the relationships between genotype and gene frequencies are close to what is expected when markers are in HWE.

The mean inbreeding coefficient is close to zero, which is a result of no major discrepancies in the observed (H_o) and expected (H_e) heterozygosities. Four loci had F = -1, which reflects no inbreeding and allele frequencies of both alleles of 0.5. Thirteen SNPs resulted in F = 1. In these loci, no heterozygote genotypes were observed.

IR and HL are distributed in the intermediate range of possible values, indicating a moderate likelihood of alleles being identical by descent, which indicates moderate genetic diversity in the original broodstock. As expected both measures were highly correlated with $r^2 = 0.97$, due to the biallelic nature of the markers.

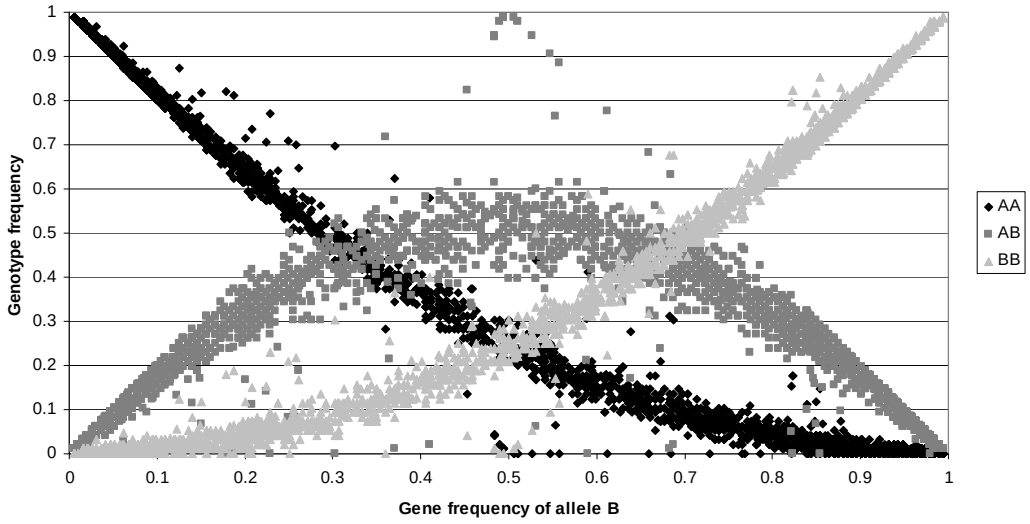

Figure 1. Relationship between genotype frequencies and gene frequencies for the two SNP alleles (A and B) in 93 Tasmanian Atlantic salmon samples.

CONCLUSION

The results of this study using high density biallelic SNPs support those obtained with sparser highly polymorphic microsatellites. Despite anecdotal evidence of small effective population size in the early years of the domestication of Atlantic salmon in Australia, both demonstrate maintenance of heterozygosity and indicate a healthy level of genetic diversity in Tasmanian Atlantic salmon population at the start of its selective breeding program. The genetic diversity found across 2991 SNP markers suggests sufficient diversity for genetic progress in the SALTAS selective breeding program. SNP technology can provide a useful tool to monitor the level of heterozygosity, as suggested by Reilly *et al.* (1999), and should be used alongside other diversity measures to avoid a decline in genetic diversity in the population.

ACKNOWLEDGEMENT

This research was supported by the Fisheries Research and Development Cooperation.

REFERENCES

Amos, W., Worthington Wilmer, J., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D. & Coulson, T. (2001). *Proc. Roy. Soc. Lond. B* **268**, 2021-2027.

Aparicio, J. M., Ortego, J. & Cordero, P. J. (2006) *Mol. Ecol.* **15**, 4659-4665.

Elliott, N.G. and Kube, P.D (2009) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **18:**362.

Innes, B.H. and Elliott N.G. (2006) *Aquaculture Research* **37**, 563 – 569.

Illumina (2007) Technical Note: Illumina DNA analysis. Publication Number 970-2007-005 06 August 2007.

Kent, M.P., Hayes, B., Xiang, Q., Berg, P.R., Gibbs, R.A., Lien, S. (2009) *Proc. PAG XVII*, San Diego, CA.

Reilly, A., Elliott, N.G., Grewe, P.M, Clabby, C., Powell, R., Ward, R.D. (1999) *Aquaculture* **173**: 459.

Wright, S. (1951). *Ann. Eugen. Lond.* **15**: 323.