SUPERIORITY OF A BIOCHEMICALLY-BASED APPROACH FOR THE DETECTION OF A QUANTITATIVE TRAIT LOCUS AFFECTING ULTIMATE pH

E. Navajas[1], D.J. Garrick[1], A.B. Pleasants[2] and C.A. Morris[2]

[1]Inst. of Vet., Animal & Biomed. Sciences, Massey University, New Zealand
[2]AgResearch, Ruakura Agricultural Research Centre, Hamilton, New Zealand

INTRODUCTION

The traditional approach to identify quantitative trait loci (QTL) in most traits has been based on the combined analysis of molecular and phenotypic information by searching for significant associations. Different methodologies are available to find the location of a QTL on a chromosome and to estimate the additive and/or dominant effects. However, the knowledge of biological processes underlying the expression of a specific trait has not been expressly included in the standard techniques.

The ultimate pH (pH_U) in beef is an economically relevant characteristic, where values higher than 5.5 negatively affect meat attributes such as tenderness and colour (Smith *et al.*, 1996; Purchas and Aungsupakorn, 1993). The frequency distribution of muscle pH_U is non-normal, presenting a peaked primary mode at 5.5 and a longer upper tail with a suggestion of a secondary mode at high values. This skewed distribution has made the studies of pH_U more difficult. In particular, logarithmic transformations fail to derive a normal probability density, thus affecting the inference that can be made.

The pH_U is determined by the post-mortem accumulation of lactic acid in the muscle. This is produced by the conversion of glycogen stored in the muscle before slaughter. A mathematical model that describes the conversion of glycogen to lactic acid in a manner suitable for statistical analysis has been proposed by Pleasants *et al.* (1999). This paper evaluates the utility of this approach in the detection of a QTL for pH_U in a simulated back-cross design.

A DYNAMICAL MODEL FROM GLYCOGEN TO ULTIMATE pH

After slaughter, muscle pH declines from neutral values to around 5.5, as the result of the breakdown of muscle glycogen into lactic acid. The following pair of differential equations states the relationship between pH of beef and muscle glycogen (G):

$$dG/dt = (5.5 - pH)y(G) \qquad \textbf{[1a]}$$
$$k(dpH/dt) = (5.5 - pH)(7.2 - pH)y(G) \qquad \textbf{[1b]}$$

where k is a constant related to the rates of each equation and y(G) is the function that describes how the loss of G varies.

The rate of change of muscle glycogen depends on the availability of muscle glycogen and the muscle pH [1a]. If the muscle glycogen levels at slaughter are low, the supply of substrate for glycolysis is limited resulting in a lower concentration of lactic acid and a higher pH_U. On the other hand, the breakdown of muscle glycogen ceases when pH is approximately 5.5 even if

residual glycogen remains undegraded probably due to enzymatic inactivation by low pH. Similarly, the rate of change of pH is affected by the concentration of muscle glycogen and by the buffering capacity of the muscle [1b]. Muscle pH falls slowly at values close to neutrality, due to the higher buffering capacity, and at pH levels near to 5.5 when the glycogen breakdown is inhibited. These facts are incorporated into the equations [1] above.

Thus, by dividing [1b] by [1a] we have $dpH/dG = 7.2 - pH$ and the solution is:

$$pH_U = 7.2 - \varepsilon\, e^{Go} \qquad\qquad \textbf{[2]}$$

where Go is the muscle glycogen at slaughter and ε recognises that muscle pH in live animals must be marginally less than 7.2 or else the evolution will not proceed. Transforming [2] and manipulating the result:

$$Go = \ln(7.2 - pH_U /\ \varepsilon) \qquad\qquad \textbf{[3]}$$

Using these results, and the fact that pH_U 5.5 is an upper bound, Pleasants *et al*. (1999) showed that the probability function of pH_U can be expressed as a mixed probability density in terms of the frequency of Go:

$$P[pH_U] = \alpha N(\ln(1.7/\varepsilon), \sigma^2_{em}) + (1-\alpha)N(\mu_{GoT}, \sigma^2_{GoT}) \qquad\qquad \textbf{[4]}$$

The peak in the actual pH_U distribution corresponds to the animals with Go equal to or greater than the amount required to achieve a pH_U of 5.5. This parameter is called Q and is calculated as $\ln(1.7/\varepsilon)$. Go values equal or higher than Q lead to a pH_U of 5.5, while those below Q result in pH_U between 5.5 and 7.2, depending on the Go values. The variation around the peak is due to the instrumental error associated with the measurement of pH and other random factors. This is assumed to be a random normal variable with mean zero and variance σ_{em}^2. On the other hand, the tail in the pH distribution is explained by the animals with Go lower than Q and with pH_U described directly by equation [2]. The total variance in the second normal distribution in the mixture includes measurement error variance σ_{em}^2 but also reflects the uncertainty in pH_U from Go, transformed from pH_U based on [3] and named as GoT in [4]. The proportion of each distribution is indicated by parameters α and $(1-\alpha)$, with α equal to $\int_Q^\infty f(Go)\, d(Go)$.

APPLICATION BY SIMULATING A BACK-CROSS DESIGN
Phenotypic data for Go was simulated using a back-cross design with a model that included additive polygenic effects and one QTL effect. pH_U records were derived in the back-cross animals from their Go records and then backtransformed to the underlying glycogen based on the novel approach. All these variables were then analysed to search for the simulated QTL.

Go values were generated for two inbred lines that included four genetic markers, equally spaced at 20 cM intervals, and one QTL located 5 cM from the second marker. Two alleles were simulated for each genetic marker and the QTL. In the base population, additive polygenic values for each animal were sampled from a distribution $N(0, \sigma_a^2)$, with σ_a^2 the additive polygenic standard deviation, and added to the additive effect of the QTL that was assumed equivalent to one σ_a. For the F1 and back-cross animals the additive polygenic values were calculated as $\frac{1}{2}gs + \frac{1}{2} gd + \phi$, where gs and gd are the additive genetic values of sires and dams, respectively, and ϕ is Mendelian sampling with $\phi \sim N(0, \sigma_{MS}^2)$ and where the Mendelian

sampling standard deviation (σ_{MS}) is equivalent to $\sqrt{\frac{1}{2}\sigma_a^2}$. The transmitted haplotypes were simulated based on the Haldane mapping function. In order to obtain the phenotypic records, residual components were generated from a distribution $N(0, \sigma_e^2)$, where σ_e is the residual standard deviation, and added to the genetic components.

The pH_U in the back-cross were calculated using the glycogen data by applying [2] for the individuals with Go lower than the truncation point Q and assigning a pH_U of 5.5 for those presenting Go equal or higher than Q. The measurement error sampled from a distribution $N(0, \sigma_{em}^2)$ was added.

Then, assuming that actual muscle glycogen was unknown, the pH_U data was backtransformed to predicted glycogen (P_{Go}) for each animal. Based on the derivation by Pleasants *et al.* (1999) that pH_U is a mixture of two normal distributions, the means, variances and proportions (α) of both sub-populations were estimated by maximum likelihood using the iterative algorithm proposed by Hosmer (1973). For each animal the P_{Go} was calculated as:

$$P_{Go} = w_{pH1}.\ EG_H + w_{pH2}.\ EG_L$$

where w_{pH1} and w_{pH2} are the weights for each record according to the probability of belonging to each sub-population. They are functions of the means and variances of the two normal distributions that were previously estimated. EG_H and EG_L are the expected values of glycogen. EG_L corresponds to those muscle glycogen levels that are below Q and is estimated based on the actual pH_U according to [3]. EG_H is related to the glycogen distribution that is above the truncation point, being equivalent to the mean of that proportion of the distribution.

$$EG_H = \mu_H = \mu_{glycogen} + (\sigma_{glycogen}.\ h_Q/p_Q)$$

where μ_H is the mean of a truncated normal distribution with lower limit equal to Q, $\mu_{glycogen}$ is the mean of muscle glycogen in the live animals, $\sigma_{glycogen}$ is the standard deviation of muscle glycogen, h_Q is the high of the standard normal curve at the truncation point (Q) and p_Q is the total area to the right of Q, which was previously estimated as the proportion of the corresponding sub-population in the mixture. Assuming $\sigma_{glycogen}$, the unknown $\mu_{glycogen}$ was estimated by Newton's method (Press *et al.*, 1986).

Estimation of the QTL position was obtained by flanking-marker regression analysis along the mapped chromosome. The following one-QTL regression model was fitted at 1-cM intervals:
$y = \mu + \beta x + e$, where y is the observed phenotypes for Go, P_{Go}, pH_U and log_{pH}, which is a logarithmic transformation of pH_U; μ is the overall mean; β is the regression coefficient; x is the probability of having inherited a paternal Q1, given the observed marker genotypes and marker/QTL positions; and e is a random residual. The *t*-test for an overall significance level of 1% was calculated by the permutation test.

Figure 1 shows the absolute value of the *t*-test along the mapped chromosome for the four analysed traits. The *t*-test curves for all traits exceed the significant threshold and reveal the presence of a possible QTL positioned between 27 cM (Go and P_{Go}) and 29 cM (pH_U and log_{pH}). However, pH_U and log_{pH} presented the lowest peak in the test statistic. Both curves were very similar indicating that the logarithmic transformation did not modify the power

compared to the actual pH_U. On the other hand, P_{Go} showed clearer evidence of a QTL by having a higher profile that is closer to Go.
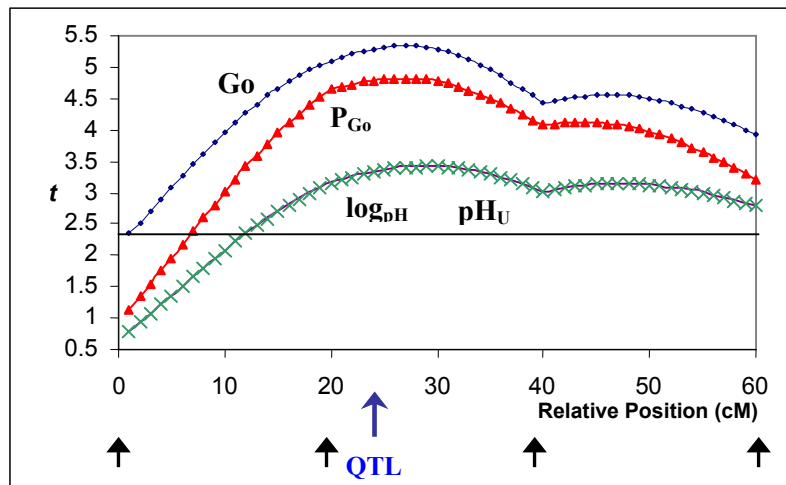

Figure 1. Curves of *t*-test absolute value along the chromosome for muscle glycogen (Go), ultimate pH (pH_U), logarithmic transformation of pH_U (log_{pH}) and predicted muscle glycogen (P_{Go}). (Arrows indicate positions of markers)

CONCLUSIONS

The power in the detection of the QTL affecting Go and therefore pH_U depends on the trait analysed. The logarithmic transformation of pH_U (log_{pH}) did not increase the power compared to pH_U, since they both presented the lowest test statistics. By applying the biochemical knowledge of the pathways that connect Go and pH_U it is possible to transform the actual pH_U data to the underlying Go. This new trait (P_{Go}) presented a higher peak, close to the values observed in Go, indicating that it is possible to improve the significance in QTL searching based on the new approach.

REFERENCES

Hosmer, D.W. (1973) *Biometrics* **29** : 761-770.

Pleasants, A.B., Wake, G.C. and Daly, C.C. (1999) *Proc MODSIM* **3** : 813-816.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986) «Numerical recipes the art of scientific computing». Cambridge University Press, New York, USA.

Purchas, R.W. and Aungsupakorn, R. (1993) *Meat Sci.* **34** : 163-178.

Smith, D.R., Wright, D.R. and Muir, P.D. (1996) *Proc NZ Soc.Anim.Prod.* **56** : 187-192.