A Cure Survival Model for Genetic Analysis of Taura Syndrome Virus Resistance in Pacific White Shrimp (*Penaeus vannamei*)

J. Ødegård[*], M. H. Yazdi[†], T. Gitterle[‡†], P. Madsen[§], B. Gjerde[*] and M. Rye[†]

Introduction

In standard survival models for time-to-event data (i.e., time until death), all individuals are typically assumed to be at risk. As a consequence, censoring is explained by a limited follow-up period. However, when analyzing survival time from challenge tests with specific pathogens (typical for aquaculture breeding programs) this assumption will be invalid given that a fraction of the population is non-susceptible (cure fraction), i.e., tolerant or completely resistant individuals. If so, mortality is expected to reach a plateau at the point when most susceptible individuals have died. Ideally, selection for non-susceptibility would often be preferred over endurance (time until death), as the latter is more likely to postpone mortality rather than avoiding it in the long run. Still, follow-up period is often limited due to practical considerations, and survivors are thus expected to be a mixture of non-susceptible and susceptible individuals with censored lifespans. A mixture cure model (Farewell 1982) is a survival model that attempts to distinguish susceptible and non-susceptible survivors, which may be of great value in analysis of time-to-event data containing a cure fraction. The aim of the study was to employ a cure model to survival data from challenge testing of Pacific white shrimp (*Penaeus vannamei*) with the taura syndrome virus (TSV).

Material and methods

Data. Shrimp originated from a Colombian Pacific white shrimp population, with parents partly selected for TSV resistance through a combined individual and family based selection program (Cock *et al.* 2009). The dataset contained individuals from seven different batches, including three consecutive generations. Each batch was tested separately in three different tanks. Parents were used across several batches, and there were thus good genetic ties between the different groups. Shrimp from the first batch were orally infected with TSV-infected minced muscle tissue for seven consecutive days at a feeding intensity of 10% of the tank biomass per day. Due to low mortality animals were infected through intramuscular injections with 20 µm of a purified inoculum of the pathogen from the second batch

[*] Nofima Marin, P.O. Box 5010, NO-1432 Ås, Norway
[†] Akvaforsk Genetics Center AS, Sjølseng, NO-6600 Sunndalsøra, Norway
[‡] CENIACUA, Bogotá: Cra 9 B No. 113-60, Colombia
[§] Department of Genetics and Biotechnology, Faculty of Agricultural Sciences (DJF), Aarhus University, DK-8830 Tjele, Denmark

onwards. For each test, mortalities were recorded on an hourly basis until no dead animals were recorded for 24 hours.

Statistical analysis. Survival times in hours were transformed to test-day (24h) survival scores with the number of records per individual equal to the number of days until death from the time of first observed mortality. For each period, the individual was scored as dead (= 1) if it was recorded as dead during this period and as alive (= 0) otherwise. The model was based on a survival score cure model developed by Ødegård *et al.* (2009). In principle, the model is a bivariate threshold model, where the first trait (endurance) includes binary test-day survival scores, given susceptibility, and the second trait (susceptibility) includes the putative binary susceptibility statuses (0 = non-susceptible, 1 = susceptible). The corresponding underlying liabilities of the two traits were analyzed with the following model:

$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda_1} \\ \boldsymbol{\lambda_2} \end{bmatrix} = \begin{bmatrix} \mathbf{X_1}\boldsymbol{\mu_1} + \mathbf{Z_t}\mathbf{t} + \mathbf{Z_{a1}}\mathbf{a_1} + \mathbf{e_1} \\ \mathbf{X_2}\boldsymbol{\mu_2} + \mathbf{Z_{a2}}\mathbf{a_2} + \mathbf{e_2} \end{bmatrix}$, where $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$ are vectors of liabilities associated with survival scores (endurance) and susceptibility statuses, respectively, $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu_1}' & \boldsymbol{\mu_2}' \end{bmatrix}'$ is a vector of "fixed" effects of batch-tank on the two traits, $\mathbf{t} \sim N\left(\mathbf{0}, \mathbf{I}\sigma_t^2\right)$ is a vector of random (batch-tank) test-day effects on $\boldsymbol{\lambda_1}$, $\mathbf{a} = \begin{bmatrix} \mathbf{a_1}' & \mathbf{a_2}' \end{bmatrix}' \sim N\left(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}\right)$ is a vector of random additive genetic effects of all individuals, $\mathbf{e} = \begin{bmatrix} \mathbf{e_1}' & \mathbf{e_2}' \end{bmatrix}' \sim N\left(\mathbf{0}, \mathbf{I}\right)$ is a vector of random residuals associated with both traits, **G** is the genetic co-variance matrix, σ_t^2 is the variance of test-day effects, **A** is the additive genetic relationship matrix and **I** denote an identity matrix of appropriate size. For a survivor *i*, surviving *k* days in a given test, individual susceptibility status was sampled from a Bernoulli distribution with a fully conditional probability for susceptibility (Ødegård et al., 2010) calculated as:

$$\tau_i = \frac{\Pr(Z_i = 1)\prod_{j=1}^{k}\left(1 - \Pr(S_{ij} = 0)\right)}{\left(1 - \Pr(Z_i = 1)\right) + \Pr(Z_i = 1)\prod_{j=1}^{k}\left(1 - \Pr(S_{ij} = 0)\right)},$$

where $\Pr(Z_i = 1)$ is the prior probability of susceptibility for animal *i* (given the associated location parameters) and $\prod_{j=1}^{k}\left(1 - \Pr(S_{ij} = 0)\right)$ is the likelihood of the survival scores for the entire test period (given the associated location parameters). Putative survival scores were set equal to the observed ones for putative susceptible individuals and missing for putative non-susceptible individuals. Genetic (co)variance components were estimated based on parental breeding values only (Ødegård *et al.* 2010), while all other dispersion and location parameters were estimated as in a standard bivariate animal threshold model. The Gibbs sampler was run for 340,000 rounds, discarding the first 40,000 rounds as burn-in. Of the remaining 300,000 rounds, samples from every 100 rounds were kept. Analysis was

performed using the Gibbs sampling module in the DMU software package (Madsen and Jensen 2007).

Results and discussion

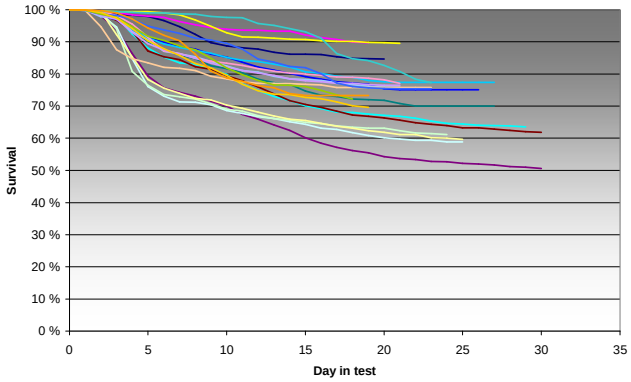

Figure 1: Kaplan-Meier survival curves for the different TSV challenge tests

Table 1: Descriptive statistics of the data set

Item	
Shrimp with data	15,261
Full-sib families	513
Sires	266
Dams	484
Generations with data	3
Batches with data	7
Challenge test tanks per batch	3
Average mortality (across tests)	28%
Average time until death[α] (across tests)	193h

[α]Excluding individuals with censored lifespans.

Descriptive statistics of the data set are given in Table 1, and Kaplan-Meier survival curves for the different tanks and batches are given in Figure 1. Earlier studies have shown considerable heritability for TSV resistance, typically in the range 0.2 – 0.3 (Cock *et al.* 2009). Preliminary analyses of survival at end of test using a standard sire-dam model (results not shown), indicated that heritability for TSV resistance was high (0.41±0.03 / 0.27±0.02 for threshold/linear models), and that common environmental family effects (e.g., effect of common rearing) were not significant. Results for the cure model are shown in Table 2. The posterior mean of the fraction susceptible shrimp were ~50%, although only 28% of the shrimp die (Table 1). Furthermore, the estimated underlying heritability of susceptibility was higher (0.54±0.04) than for crude survival at end of test, while the estimated underlying (test-day) heritability of survival scores was moderate (0.21±0.03). The genetic correlation between the two traits was not significantly different from zero (-0.06±0.07), indicating that selection for increased time until death would be suboptimal if the

aim is to improve the long-term survival. This result is also relevant for the current disease challenge practice in many aquaculture breeding programs, where challenge tests are often terminated at intermediate frequencies as this is considered optimal for analysis of binary traits (i.e., dead/alive). However, early termination at still increasing mortality is likely to shift the focus of selection towards endurance rather than susceptibility. If possible, challenge tests should ideally run until mortality naturally ceases, as this maximizes the potential importance of susceptibility status on the recorded end-survival.

Table 2: Posterior means of parameters

Parameters	Posterior means	Posterior SD
Proportion susceptible shrimp	0.47	0.02
Test-day variance for survival scores[α]	0.14	0.02
Genetic variance - survival scores[α]	0.28	0.05
Genetic variance - susceptibility[α]	1.21	0.19
Genetic correlation[α]	-0.06	0.07
Heritability - survival scores[αβ]	0.21	0.03
Heritability - susceptibility[αβ]	0.54	0.04

[α]On the underlying liability scale, [β]h^2 = genetic variance/(genetic variance +1)

Conclusion

Cure model analysis of TSV resistance indicates that endurance and susceptibility to the disease are clearly distinct genetic traits. If the aim is to improve long-term survival under exposure to TSV, the breeding goal should aim at reducing susceptibility rather than increasing time until death, which can be done through a cure survival model. However, genetic evaluation of susceptibility would benefit if challenge tests are run until mortality naturally ceases. This has implications for the current testing practice in many aquaculture breeding programs, where challenge tests are commonly terminated at intermediate and still increasing mortality levels.

Acknowledgments

The research was co-funded by Akvaforsk Genetics Center AS (AFGC) and The Research Council of Norway in project no. 192331/S40.

References

Cock, J. *et al.* (2009). *Aquaculture*, 286:1-11.
Farewell, V. T. (1982). *Biometrics*, 38:1041-1046.
Madsen, P. and Jensen, J. (2007). DMU: a user's guide. A package for analysing multivariate mixed models. Version 6, release 4.7:
Ødegård, J. *et al.* (2009). *J. Anim. Sci.* (submitted).
Ødegård, J. *et al.* (2010). *Gen. Sel. Evol.* (submitted).