Genetic analysis of resistance in Mekong striped catfish (*Pangasianodon hypophthalmus*) to bacillary necrosis caused by *Edwardsiella ictaluri*

Running title: Genetic analysis of challenge-test data

Khoi Dinh Pham [ab], Jørgen Ødegård [ac], Sang Van Nguyen [b], Hans Magnus Gjøen [a], Gunnar Klemetsdal [a]

[a] Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1433 Ås, Norway

[b] Research Institute for Aquaculture No. 2 (RIA2), 116 Nguyen Dinh Chieu Street, District 1, Ho Chi Minh City, Vietnam

[c] AquaGen AS, P.O. Box 1240, 7462 Trondheim, Norway

Corresponding author: Khoi Dinh Pham. Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1433 Ås, Norway. Email: phamdinhkhoi@gmail.com

Acknowledgements

Funding of this research was by the Vietnamese Ministry of Agriculture and Rural Development and the Norwegian Quota scholarship scheme. Dr. Hao Van Nguyen, Dr. Phuoc Hong Le and all other RIA2 staff members are acknowledged for their contribution.

Data Availability Statement: The data that support the finding of this study can be made available on request, by the corresponding author.

Abstract

The aim of this study was to analyse four cohabitation challenge test experiments with Mekong striped catfish (*Pangasianodon hypophthalmus*) against the bacterium *Edwardsiella ictaluri*. The data was genetically analysed per experiment by three models; 1) a cross-sectional linear model, 2) a cross-sectional threshold model, and 3) a linear survival model; at both 50% mortality (for models 1 and 2) and at the end of the test (for all three models). In two of the experiments (3 and 4) that were carried out in two replicated tanks, the predicted family effects (sum of sire, dam and common environmental effects) in each tank were correlated to the family survival in the other replicated tank (cross-validation). The heritability estimates of resistance to *E. ictaluri* infection were ≤ 0.012 with the survival model, up to 0.135 - 0.220 (50% survival) and 0.085 and 0.174 (endpoint survival) for the cross-sectional linear and threshold models, respectively. The challenge test should aim for an endpoint survival that ceases naturally at 50%. Then, genetic analysis should be carried out for survival at the endpoint (reflecting susceptibility) with a simple cross-sectional linear model.

Keywords: Challenge test; Cross-sectional model; Heritability; Survival model

1. Introduction

In Vietnam, bacillary necrosis (BN) caused by the bacterium *Edwardsiella ictaluri* (*E. ictaluri*) has become a severe problem in farming of Mekong striped catfish (*Pangasianodon hypophthalmus*). Therefore, the first BN vaccine trial was conducted by Pharmaq Ltd. Vietnam in 2010 (Thanh & Berntsen 2012). The trial was successful with regard to significant lower mortality of vaccinated than non-vaccinated groups, observed in the field for 31 days. The Alpha Ject® Panga 1 vaccine was licensed in 2013 (https://www.pharmaq.no/sfiles/8/66/4/file/2013_08-cty-pharmaq-vn_thuy-

san-nam-14-so-164.pdf). However, improved resistance to BN using vaccination is costly and laborious since it must be administered to every fish produced. Therefore, the main method for treating BN at current is the use of antibiotics, but combating disease outbreaks by drug application is costly and a major concern to the environment and the consumer, and not a long-term solution to the problem (van Muiswinkel, Wiegertjes & Stet 1999). However, with fish, selective breeding for disease resistance has been proven to be an efficient strategy to prevent infectious diseases (Guy, Bishop, Woolliams & Brotherstone 2009; Lhorente, Gallardo, Villanueva, Araya, Torrealba, Toledo & Neira 2012; Taylor, Wynne, Kube & Elliott 2007). Controlled challenge testing is a widely used method for testing genetic variation in resistance (e.g., Gjedrem & Gjøen 1995). Then, individuals are typically infected by e.g. cohabitants, i.e., fish that have been injected with the causative agent (e.g., Gjøen, Refstie, Ulla & Gjerde 1997). Controlled challenge testing has been widely applied in Atlantic salmon (*Salmo salar*) breeding, demonstrating that resistance to diseases show substantial genetic variance and heritability (e.g., Gjerde, Boison, Aslam, Løvoll, Bakke, Rey & Lillehammer 2019; Yáñez, Lhorente, Bassini, Oyarzún, Neira & Newman 2014). The challenge method has also been applied in many other fish species, like common carp (*Cyprinus carpio*, Ødegård, Olesen, Dixon, Jeney, Nielsen, Way, Joiner, Jeney, Ardó, Rónyai & Gjerde 2010), Atlantic cod (*Gadus morhua L.*, Kettunen & Fjalestad 2006; Ødegård, Sommer & Præbel 2010), rainbow trout (*Oncorhynchus mykiss,* Bassini, Lhorente, Oyarzún, Bangera, Yáñez & Neira 2019), European sea bass (*Dicentrarchus labrax*, Doan Q., Vandeputte, Chatain, Haffray, Vergnet, Breuil & Allal 2017), coho salmon (*Oncorhynchus kisutch,* Barría, Doeschl-Wilson, Lhorente, Houston & Yáñez 2019), red tilapia (*Oreochromis spp.*, Sukhavachana, Poompuang, Onming & Luengnaruemitchai 2019), bighead catfish (*Clarias macrocephalus*, Srisapoome, Chatchaiphan, Bunnoy, Koonawootrittriron & Na-Nakorn 2019), Chinese tongue sole (*Cynoglossus semilaevis,*

Li, Wang, Yang, Li, Dai & Chen 2019), and Mekong striped catfish (*Pangasianodon hypophthalmus*, Vu, Sang, Trong, Duy, Dang & Nguyen 2019). The existence of genetic variation for a trait opens the possibility to improve the trait through selective breeding.

Selection of Mekong striped catfish in Vietnam has in main been carried out for increased growth rate (Vu, Sang, Phuc, Vuong & Nguyen 2019), and the authors calculated a selection response per generation of 9.3%. Moreover, Vu, Sang, Trong, Duy, Dang & Nguyen (2019) estimated genetic variance for resistance to BN across four challenge-test experiments carried out in 2010, 2011, 2012 and 2015. They assumed BN to be the same trait across experiments and found the heritability for dead/alive at maximum 29 days post-challenge to be 0.10 and 0.16, with a linear and threshold model, respectively. Here, the main objective was to analyse, experiment wise, three of the same experiments (2010 – 2012), but also an experiment carried out in 2009. In the genetic analyses, two different trait definitions were used; time until death or dead/alive, the latter measured at both 50% overall mortality and at end of the test, to assess genetic variance of disease resistance. The analysis was carried out by three genetic evaluation models; two cross-sectional models utilizing data either at the endpoint or at 50% mortality, and a survival model utilizing time until death. We evaluated these models by predicting the family survival (sum of sire, dam and common environmental effects) in one tank and correlated it to the family survival in another tank.

2. Materials and methods

By granting the research, the Vietnamese Ministry of Agriculture and Rural Development pre-approved the experiments carried out.

2.1 Data

87 The fish used in the four experiments were from different year-classes and sub-populations of the
88 breeding program for Mekong striped catfish in Vietnam, illustrated in Figure 1. This study
89 involved the year-classes produced in 2008, 2009, 2010, and 2011, being, respectively, the third
90 generation of subpopulation 2001 (G3-2001), while the others were G3-2002, G3-2003 and G4-
91 2001. In all experiments, the test-fish were the offspring from a nested mating design (one male
92 mated to two females). In 2008, 2009, 2010, and 2011 families were produced from June 16^{th} -
93 July 14^{th}, July 15^{th} - Aug 10^{th}, July 29^{th} - September 10^{th}, and from June 9^{th} - July 7^{th}, respectively.
94 Spawning was done by hormone treatment (HCG - Human Chorionic Gonadotropin). First, males
95 were stripped, milt was stored at 4^0 C, later it was split in two, mixed with samples of eggs from
96 two females, before water was added for fertilization. The fertilized eggs were washed to remove
97 sticky layers by use of tannic acid, and eggs from one female were moved to a family air-supplied
98 net-jar in one cement tank for hatching, occurring from 18 - 24 hours after fertilization. The total
99 number of families produced in 2008, 2009, 2010 and 2011 were 156, 196, 233, and 250.

100 Twenty-four hours post hatching, approximately 3,000 start-fed fry were randomly sampled from
101 each family and reared in a 1 m^3 family-fiberglass tank for about 20 days. The tanks were air
102 supplied, and about half the water was exchanged every three days. Fry were in sequence fed *ad*
103 *libitum* with newly hatched Artemia sp., Moina sp. and bloodworm (*Limnodrilus hoffmeisteri*).
104 After the 20 days, a random sample of about 300 small fingerlings from each full-sib family was
105 moved to a family hapa located in an earthen pond. Here, fish were initially fed *ad libitum* by
106 bloodworm, but within a week their diet was standard commercial pellet feed (V2-Feed, RIA2 –
107 Research Institute for Aquaculture No. 2, Ho Chi Minh City, < 2.0 mm, 22 - 28% protein). Cleaning
108 of hapas were done frequently. In 2011, 15 families were nursed in two replicated hapas.

Tagging was done when the fish were randomly netted out, hapa by hapa. Passive Integrated Transponder tags (PIT tags, Sokymat, Switzerland) were inserted from April 8th - 10th 2009 (year-class 2008 and experiment 1), January 13th - 25th, 2010 (year-class 2009 and experiment 2), December 16th, 2010, to January 9th, 2011 (year-class 2010 and experiment 3) and from December 15th - 21st, 2011 (year-class 2011 and experiment 4). Fish to be challenge tested were transferred to either of two tanks (in experiment 1, one tank was used) at the National Breeding Centre for Southern Freshwater Aquaculture (NABRECSOFA), before being transported to the Govap Experimental Center (Ho Chi Minh City), RIA2, for challenge testing. In experiment 1, the challenge test was carried out at NABRECSOFA.

The number of test-tanks in experiment 1 was one (20 m^3 each), as mentioned, whereas two were used in the last three experiments (Figure 1). Table 1 shows the number of families and test-fish in each tank of the four experiments, the latter making up a total of 2,155 (mean weight 48.8 g), 1,988 (mean weight 23.6 g), 5,689 (mean weight 20.0 g) and 6,177 (mean weight 20.0 g) fish, respectively

The cohabitation method was applied in four challenge-test experiments, started April 30th 2009, February 23rd 2010, January 14th 2011 and January 3rd 2012, respectively. Fish were transferred to the test units 22, 14, 3 and 3 days prior to challenge. Water temperature was 29.5^0 C, 29^0 C (26^0 C from day 11 until termination), 26° C and 26° C during the challenge, respectively. To get a sufficient response to the challenge, the fish in the last three experiments were stressed by halving the water level in the test tanks from one day prior to the test. The ratio of the number of cohabitants shedders to the number of test fish was ~1:7 in the first experiment and ~1:3 later. Cohabitants were infected by intraperitoneally injection (doses were: 2.5 x 10^6, 2.5 x 10^6 for half the cohabitants, and 2.5 x 10^5 for the rest, and 1 x 10^5 bacteria in the last two experiments) and released

directly into the test tanks (in the first experiment, cohabitants were located to a hapa, central in the tank). The bacteria were from a strain of *E. ictaluri* Gly09M (Southern Monitoring Center for Aquaculture Environment & Epidemic, RIA2, Ho Chi Minh City, Vietnam). In the first two experiments, dead cohabitants were removed (when sunk and floating, respectively), while in the last two experiments dead and floating cohabitants were collected into plastic baskets that were hung down into the water for another two days before removal. In experiment 3, external pathogen was added to the test tanks to reach a density of 2.5×10^6 bacteria/ml water, from day 6 post-challenge when the death of cohabitants had reached peak. This practice was continued for another 8 days. In experiment 4, addition of pathogen was started at day 3 post-challenge, and stopped at day 6, after the cohabitant mortality had reached peak. Throughout the tests, fish were daily fed with standard commercial pellets at a rate of 1% of total biomass. Random samples of dead fish were examined for presence of *E. ictaluri*, as typical colonies (Crumlish, Dung, Turnbull, Ngoc & Ferguson 2002). Kidney samples were grown on sheep blood agar plates and incubated at 30^0 C for 24 hours. In 100% of the samples, *E. ictaluri* was identified. Alive fish were biosecure-buried, following the national veterinary regulations (Department of Animal Health, Vietnam).

2.2 Statistical analyses

Initially, experiment- and tank-specific (Kaplan-Meier) survival curves were calculated. In addition, for each family survival (number of survivors to number of test fish at the start of the experiment) at the end of the tests were obtained for each tank. Because of the low mortality in experiment 2, the genetic analyses in this study had to be based on the three remaining experiments. Genetic analyses of these experiments were carried out with three different sire - dam models per experiment. In experiments 3 and 4, with replicated tanks, the models were validated by correlating

153 the predicted family effects in one tank to the corresponding observed family survival in the other
154 tank.

155 Experiment-wise analyses were conducted since only four sires and two dams from year-class 2006
156 were used in both experiments 3 and 4, while these numbers in year-class 2007, used in experiment
157 4, were five and seven, respectively (Figure 1). This led us to conclude that the genetic ties were
158 too few and that the analyses had to be carried out on a per experiment basis. In the linear model
159 (LM) and in the threshold model (TM), a binary trait (dead = 0/alive = 1) was defined at two stages:
160 At the end of the test (endpoint) and at the day the truncated mortality was closest to 50% (50%
161 mortality), which was at days 11 and 14 in tanks 1 and 2 in experiment 3 and at days 13 and 12 in
162 tanks 1 and 2 in experiment 4. Note that only endpoint mortality could be considered in experiment
163 1 since the mortality in this experiment was lower than 50%. In the linear survival model (LSM) a
164 binary variable per test day across the test period was defined as 1/0 if the fish was alive/dead on
165 test-day t, where 0 implied that there would be no further record for that fish.

166 First, for experiments 3 and 4 Kaplan-Meier trajectories of the survival curves of the two tanks
167 were compared with a log-rank test (https://www.real-statistics.com/survival-analysis/kaplan-
168 meier-procedure/log-rank-test/).

169 Then, the following cross-sectional LM was applied to the binary trait (dead/alive = 0/1, both at
170 the endpoint and at 50% mortality):

171 $$Y_{klmn} = \mu + b_1X_1 + b_2X_2 + T_k + s_l + d_m + c_{lm} + e_{klmn}$$

172 where Y_{klmn} = alive or dead (0 = dead, 1 = alive) for fish *n;* μ = the overall mean; b_1 = fixed regression
173 coefficient on number of days from spawning until tagging (X_1); b_2 = fixed regression coefficient on

number of days from tagging to start of experiment (day 0) (X_2); T_k = the fixed effect of tank k (k = 1, 2); s_l = random additive genetic effect of sire l; d_m = random additive genetic effect of dam m; c_{lm} = random common environmental effect pertaining to fullsib family lm; and e_{klmn} = random error term for fish n.

Above, the random additive genetic effects of sire and dam can be represented by a vector of sire and dam effect: $\begin{bmatrix} \boldsymbol{s} \\ \boldsymbol{d} \end{bmatrix}$. Further, E($\boldsymbol{s}$) = E($\boldsymbol{d}$) = E($\boldsymbol{c}$) = E($\boldsymbol{e}$) = **0;** Var($\boldsymbol{s}$) = Var($\boldsymbol{d}$) = $\mathbf{A}\sigma_{sd}^2$, where **A** is the additive genetic relationship matrix, and σ_{sd}^2 is the common sire-dam variance component; Var($\boldsymbol{c}$) = $\mathbf{I}\sigma_c^2$, where **I** is an identity matrix, and σ_c^2 is the common environmental variance (potentially including also maternal and dominance effects in addition to the environmental effect of hapa), and Var($\boldsymbol{e}$) = $\mathbf{I}\sigma_e^2$, with σ_e^2 being the residual variance.

Additionally, the data was analysed with a cross-sectional (probit) TM, assuming a normal underlying liability variable l that determines the categorical outcome, such that $l_{klmn} \leq 0$ gives Y_{klmn} = 0, and $l_{klmn} > 0$ gives Y_{klmn} = 1. Restricting the residual variance on the underlying liability scale to $\sigma_e^2 = 1$, the model can be written:

$$Pr\,(Y_{klmn} = 1) = Pr\,(l_{klmn} > 0) = \Phi(b_1X_1 + b_2X_2 + T_k + s_l + d_m + c_{lm})$$

where $\Phi(.)$ is the cumulative standard normal distribution function, and the other parameters are as described for the LM.

Finally, the linear survival model LSM was specified as:

$$Y_{klmnt} = \mu + b_1X_1 + b_2X_2 + T_k + \sum_{p=0}^{4} \beta_{pk} Z_p(t) + s_l + d_m + c_{lm} + e_{klmnt}$$

where Y_{klmnt} = fish *n* alive or not (dead/alive = 0/1) at test-day *t*; $Z_p(t)$= p^{th} order orthogonal polynomial of a specific day *t* (test day), with *p* = 0, 1, 2, 3 and 4; β_{pk} = p^{th} order fixed regression coefficient nested within tank *k*; e_{klmnt} = random error term for fish *n* at test-day *t* ~ $N(0, \mathbf{I}\sigma^2_e)$, and the remaining parameters as described above.

The relationships were traced back to the base, comprising a total of 2,389, 6,145 and 6,905 animals for experiments 1, 3 and 4, respectively. The data was analysed with ASReml, version 4.1 (Gilmour, Gogel, Cullis, Welham & Thompson 2015).

2.3 Heritability

Heritabilities were calculated as follows:

$$h^2 = \frac{4\sigma^2_{sd}}{2\sigma^2_{sd} + \sigma^2_c + \sigma^2_e}$$

where σ^2_{sd} is the common sire-dam variance component, σ^2_c is the common environmental variance, and σ^2_e is the residual variance.

2.4 Model comparison

To compare the predictive value of the three models, the Pearson correlation coefficient between the predicted family effects (sum of sire, dam and common environmental effects) in one tank and the mean family survival in the other tank was calculated. The validation data was always the endpoint survival. Training data was survival at the endpoint, at 50% overall survival or as time until death (endpoint for survivors). The test of the difference between the dependent correlations (the same data used by the three models) was calculated by the software of Lee & Preacher (2013).

212 Finally, Spearman rank correlation coefficients between predicted family effects across the two
213 replicate tanks were calculated at both 50% mortality and at the endpoint, with the three different
214 models.

215 **3. Results**

216 In three out of the four linear model analyses carried out across tanks in experiments 3 and 4 (both
217 at 50% mortality and at the endpoint), increased number of days from spawning till tagging
218 decreased survival ($P < 0.001$), while increased number of days from tagging to start of the
219 experiment led to enhanced survival ($P < 0.001$) (results not shown).

220 Cumulative Kaplan-Meier survival curves in the challenge-test experiments are shown in Figure 2.
221 Mortality was much lower in experiments 1 and 2 than in experiments 3 and 4. The development
222 of mortality was also different among the experiments. In the first, where the overall cumulative
223 mortality reached 25%, the earliest mortality of fish was recorded on day 8, reaching a maximum
224 of 7.6% at day 11. In experiment 2, cumulative mortality was only 3.0% and 5.7% for tanks 1 and
225 2, respectively. In experiment 3, cumulative mortality was high; 84.0% and 83.1% for tanks 1 and
226 2. Moreover, the trajectories of the survival curves were seemingly different between the two tanks,
227 with earlier initiation of mortality in tank 1 than in tank 2. In experiment 4, cumulative mortality
228 was about 87% in both tanks, with more similar survival curves. In both experiment 3 and 4, a log-
229 rank test showed that the survival curves of the two replicated tanks were significantly different
230 (with test statistics of 123.37 and 40.09, respectively, with one degree of freedom).

231 Survival at the end of the tests are shown for each family per experiment in Figure 3. In experiment
232 1, all families had surviving members, and mean fraction of survivors across families was 75%
233 (ranging from 29 to 100% among families). In experiment 2, mean survival across families in tanks

1 and 2 were 97% and 94% (range: 88 - 100% and 68 - 100%, respectively). In experiment 3, with two tanks, 140 (75% of all families) and 152 (81%) families had survivors at the end of the challenge test. Mean survival across families in tank 1 was 15% (range: 0 - 69%), while it was 16% (range: 0 - 56%) in tank 2. In experiment 4, mean survival across families were 12% in both tanks (range: 0 - 70% and 0 - 60%), and 152 and 146 families (65% and 63% of all families) had survivors at the end of the test.

Variance components and heritabilities obtained at the two stages, endpoint (all three models) and 50% overall mortality (with LM and TM), in experiments 1, 3 and 4 are presented in Table 2. Generally, as expected, the estimated heritabilities were lower for LM than for TM, because the latter estimates heritability on the underlying scale. Heritability estimates were also considerably lower with the LSM (test-day survival) than with the cross-sectional models. This is expected due to the fact that LSM models daily survival, while the LM and TM accumulates survival over the entire test period. With LSM, estimated heritability on the test-day level was ~1% in all the three experiments analysed. With the cross-sectional models, the largest heritabilities were obtained at 50% mortality, with 0.22 and 0.13 for TM and LM, respectively. At the endpoint, both genetic variance and heritability of the cross-sectional models were found to be reduced due to lower frequency of survivors. Moreover, the estimated common environmental variance for endpoint survival was somewhat enlarged in experiment 4, likely due to statistical uncertainty.

The predictive ability of the models, assessed as the Pearson correlation coefficient between the observed family survival in one tank and the predicted family survival in the other tank using survival at both 50% mortality and at the endpoint as training data, are presented in Table 3. In general, the linear models (LM and LSM) predicted family survival better than TM, while no

differences ($P < 0.05$) were obtained between LM and LSM.

Spearman rank correlations between family survival calculated across the two replicated tanks, with the three models at both 50% mortality and at the endpoint, in the same experiment are given in Table 4, showing that LM and TM correlated closely. However, with these two models, survival at 50% mortality and at the endpoint correlated moderately (< 0.72), indicating substantial reranking depending on the timing of the cut-off point. Moreover, LSM correlations with LM and TM ranged 0.81 - 0.92 at both stages.

4. Discussion

The average mortality varied much across the four challenge test experiments (5.7 – 87.7%), far away from a natural reference of 50% maximizing the phenotypic variance for a binary trait (Gjøen, Refstie, Ulla & Gjerde 1997). The four tests propose three days acclimatization of test fish prior to the challenge, with restricted water level, keeping a temperature of 26^0 C. In the challenge, cohabitant shedders should be released directly into the test tank and make up ~⅓ of the fish, and bacteria should be added directly to water. Experiments 3 and 4, with the highest mortality, suggest that any factor involving the dead cohabitants should be removed, and that additional experimentation should focus on bacteria (density) and timing for addition of bacteria to water.

The study shows that resistance to BN is heritable (Table 2). As expected, the largest estimate of heritability was obtained with the TM since these parameter estimates are on the underlying and unobserved liability scale. This heritability will only be realized given that one could observe liability directly, which is not possible in practice. One problem with the LM is that heritability estimates are frequency dependent (Gianola & Foulley 1983). Moreover, LSM estimates were even smaller than those obtained with the cross-sectional models (TM and LM) since the information

per animal is split onto several days. As expected, the heritability for survival from the cross-sectional models were higher at 50% overall mortality than at the endpoint. A main reason for this is that several families had no survivors at the endpoint (Figure 3), leaving less genetic variance to be detected. The stage of 50% mortality was chosen because it has been frequently used in challenge tests with fish (Gjøen, Refstie, Ulla & Gjerde 1997), likely because this frequency maximizes the phenotypic variance of the binary survival trait.

The two experiments 3 and 4, each with two replicated tanks, had the advantage of allowing comparison of the predictive ability of different statistical models and trait definitions, by performing a between-tank validation as also done by Gitterle, Ødegård, Gjerde, Rye & Salte (2006). Herein, the comparison was based on the use of full-sib family effects, while Gitterle, Ødegård, Gjerde, Rye & Salte (2006) based their comparison on estimated breeding values. The family effect was chosen over the additive genetic effect because of the weakness of the nested mating design in separation of the additive genetic, non-additive genetic and common environmental effects (Berg & Henryon 1998). From the results (Table 3), it can be inferred that the TM model predicts the family survival inferior to the linear models. With some families having no survivors, extreme category problems may affect the TM, which might be a reason for the inferior performance. The validation did not discriminate between the LM and the LSM (Tables 3). However, in experiment 3, with the largest difference in trajectories of survival curves between tanks and with the lower average mortality at the endpoint, the LSM was found to have the highest correlations to survival in the other tank (Table 3), which indicates that accounting for time until death may be useful as also reported by Gitterle, Ødegård, Gjerde, Rye & Salte (2006), Ødegård, Olesen, Gjerde & Klemetsdal (2006), and Ødegård, Olesen, Gjerde & Klemetsdal (2007).

The Spearman rank correlation values between family survival calculated across the two replicated tanks in the same experiment obtained at 50% mortality and at the endpoint of the challenge were moderate (≤ 0.72, Table 4), implying substantial re-ranking of family survival effects. Moreover, the corresponding correlation of family survival effects at 50% mortality to that obtained with LSM was higher (≥ 0.85), since back-truncating the test to 50% overall mortality is equivalent to a binary analysis of time-to-death, binary categorizing survival time as either long (1) or short (0). Observed survival during a challenge test may be a mixture of two underlying traits, called susceptibility and endurance (Ødegård, Madsen, Labouriau, Gjerde & Meuwissen 2011). Susceptibility is whether or not the animal is at risk of dying to the disease, while endurance is the ability of susceptible individuals to stay alive for some time (latency) during exposure (Kause & Ødegård 2012). Given that the exposure period is sufficiently long (i.e., continued until mortality ceases) the observed endpoint survival will approach the phenotypic susceptibility, while taken at earlier time-points the observed survival will be a mixture of susceptibility and endurance traits. This may explain the substantial re-ranking of families when correlating survival at 50% overall mortality with that at the endpoint. Another possibility would be to perform a more complex genetic analysis with a cure survival model (Ødegård, Gitterle, Madsen, Meuwissen, Yazdi, Gjerde, Pulgarin & Rye 2011; Ødegård, Madsen, Labouriau, Gjerde & Meuwissen 2011), attempting to separate endurance and susceptibility from survival time (mixture trait). An easier and more robust option is to only consider susceptibility at the endpoint, requiring that mortality has naturally ceased. This was not fully reached in experiments 3 and 4 where the tests ideally should have been prolonged.

In addition to susceptibility and endurance, host infectivity has received attention in genetic analysis of disease resistance data (Anacleto, Cabaleiro, Villanueva, Saura, Houston, Woolliams & Doeschl-Wilson 2019). These authors define the trait as the host's ability to infect an average

individual upon contact. However, it can be argued that if animals become non-susceptible, and this arises from the fish being resistant to the pathogen, these fish may also be less likely to spread the pathogen.

The present study has implications as to how the challenge test against BN should be carried out in Mekong striped catfish. If mortality is naturally ceasing, endpoint mortality is a measure of susceptibility and should have preference over the other measures of resistance. Preferably, this natural endpoint mortality should be attained at ~50% since this maximizes the phenotypic variance of susceptibility (Table 2) at which EBV's can be obtained by a simple cross-sectional linear model.

The models above do not allow to distinguish between individuals within a family and do not utilize the possibility to carry out within-family selection in a sib-population of untested breeding candidates. In practice, inbreeding considerations will force the breeder to select from a broader range of families, reducing the realized selection differential. Furthermore, selection accuracy will also become reduced as Mendelian sampling variation (within-family genetic variance) constitutes half the total genetic variance (likely more due to Bulmer effects in populations under selection), which is not considered through family selection (Ødegård, Baranski, Gjerde & Gjedrem 2011). In order to obtain both higher selection intensity and selection accuracy, a genomic selection program for BN resistance in Mekong striped catfish is advisable. Then, candidates can be selected based on the summed effects of markers spanning the whole genome of individual fish, allowing to utilise the whole genetic variance also when selecting among untested selection candidates. The limitation of this method is the cost of genotyping as well as the availability of a SNP array. There is work conducted to construct a high density SNP array in Mekong striped catfish, e.g. Vo, Nguyen, Nguyen & Tran (2018). Another advantage of genomic markers and genomic

relationships would be the possibility to perform a more efficient statistical correction for environmental effects common to fullsibs caused by the separate rearing of the families. Parental assignment through genetic markers could allow communal rearing of the families from a much early life stage to be used that would reduce the common environmental effect.

5. Conclusions

It is concluded that resistance to *E. ictaluri* causing BN in Mekong striped catfish is heritable. The challenge test should aim for an endpoint survival that ceases naturally at ~50%. Breeding values should preferably be calculated for endpoint survival, with a simple cross-sectional linear model. With the considerable genetic variance estimated in this study, susceptibility to BN has the potential to become considerably changed by selection over time.

Conflict of interest statement

There is no conflict of interest.

References

Anacleto, O., Cabaleiro, S., Villanueva, B., Saura, M., Houston, R.D., Woolliams, J.A. & Doeschl-Wilson, A.B. (2019) Genetic differences in host infectivity affect disease spread and survival in epidemics. *Scientific Reports,* **9,** 4924.

Barría, A., Doeschl-Wilson, A.B., Lhorente, J.P., Houston, R.D. & Yáñez, J.M. (2019) Novel insights into the genetic relationship between growth and disease resistance in an aquaculture strain of Coho salmon (*Oncorhynchus kisutch*). *Aquaculture,* **511,** 734207.

Bassini, L.N., Lhorente, J.P., Oyarzún, M., Bangera, R., Yáñez, J.M. & Neira, R. (2019) Genetic parameters for *Piscirickettsia salmonis* resistance, sea lice (*Caligus rogercresseyi*) susceptibility and harvest weight in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture,* **510,** 276-282.

Berg, P. & Henryon, M. (1998) A comparison of mating designs for inference on genetic parameters in fish. *Proc. 6th WCGALP,* **27,** 115-118.

Crumlish, M., Dung, T.T., Turnbull, J.F., Ngoc, N.T.N. & Ferguson, H.W. (2002) Identification of *Edwardsiella ictaluri* from diseased freshwater catfish, *Pangasius hypophthalmus* (Sauvage), cultured in the Mekong Delta, Vietnam. *Journal of Fish Diseases,* **25,** 733-736.

Doan Q., K., Vandeputte, M., Chatain, B., Haffray, P., Vergnet, A., Breuil, G. & Allal, F. (2017) Genetic variation of resistance to Viral Nervous Necrosis and genetic correlations with production traits in wild populations of the European sea bass (*Dicentrarchus labrax*). *Aquaculture,* **478,** 1-8.

Gianola, D. & Foulley, J.L. (1983) Sire evaluation for ordered categorical data with a threshold model. *Genetics, Selection, Evolution,* **15,** 201-224.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. & Thompson, R. (2015) ASReml User Guide Release 4.1 Structural Specification. Retrived from https://www.hpc.iastate.edu/sites/default/files/uploads/ASREML/UserGuideStructural.pdf.

Gitterle, T., Ødegård, J., Gjerde, B., Rye, M. & Salte, R. (2006) Genetic parameters and accuracy of selection for resistance to White Spot Syndrome Virus (WSSV) in *Penaeus* (*Litopenaeus*) *vannamei* using different statistical models. *Aquaculture,* **251,** 210-218.

Gjedrem, T. & Gjøen, H.M. (1995) Genetic variation in susceptibility of Atlantic salmon, *Salmo salar L.*, to furunculosis, BKD and cold water vibriosis. *Aquaculture Research,* **26,** 129-134.

Gjerde, B., Boison, S.A., Aslam, M.L., Løvoll, M., Bakke, H., Rey, S. & Lillehammer, M. (2019) Estimates of genetic correlations between susceptibility of Atlantic salmon to amoebic gill disease in a bath challenge test and a field test. *Aquaculture,* **511,** 734265.

Gjøen, H.M., Refstie, T., Ulla, O. & Gjerde, B. (1997) Genetic correlations between survival of Atlantic salmon in challenge and field tests. *Aquaculture,* **158,** 277-288.

Guy, D.R., Bishop, S.C., Woolliams, J.A. & Brotherstone, S. (2009) Genetic parameters for resistance to Infectious Pancreatic Necrosis in pedigreed Atlantic salmon (*Salmo salar*) post-smolts using a Reduced Animal Model. *Aquaculture,* **290,** 229-235.

Kause, A. & Ødegård, J. (2012) The genetic analysis of tolerance to infections: a review. *Frontiers in Genetics,* **3,** 262.

Kettunen, A. & Fjalestad, K.T. (2006) Resistance to vibriosis in Atlantic cod (*Gadus morhua* L.): First challenge test results. *Aquaculture,* **258,** 263-269.

Lee, I.A. & Preacher, K.J. (2013) Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Retrived from http://quantpsy.org/corrtest/corrtest2. In.

Lhorente, J.P., Gallardo, J.A., Villanueva, B., Araya, A.M., Torrealba, D.A., Toledo, X.E. & Neira, R. (2012) Quantitative genetic basis for resistance to *Caligus rogercresseyi* sea lice in a breeding population of Atlantic salmon (*Salmo salar*). *Aquaculture,* **324–325,** 55-59.

Li, Y., Wang, L., Yang, Y., Li, X., Dai, H. & Chen, S. (2019) Genetic analysis of disease resistance to *Vibrio harveyi* by challenge test in Chinese tongue sole (*Cynoglossus semilaevis*). *Aquaculture,* **503,** 430-435.

Srisapoome, P., Chatchaiphan, S., Bunnoy, A., Koonawootrittriron, S. & Na-Nakorn, U. (2019) Heritability of immunity traits and disease resistance of bighead catfish, *Clarias macrocephalus* Günther, 1864. *Fish & Shellfish Immunology,* **92,** 209-215.

Sukhavachana, S., Poompuang, S., Onming, S. & Luengnaruemitchai, A. (2019) Heritability estimates and selection response for resistance to *Streptococcus agalactiae* in red tilapia *Oreochromis spp. Aquaculture,* **502,** 384-390.

Taylor, R.S., Wynne, J.W., Kube, P.D. & Elliott, N.G. (2007) Genetic variation of resistance to amoebic gill disease in Atlantic salmon (*Salmo salar*) assessed in a challenge system. *Aquaculture,* **272,** S94-S99.

Thanh, P.C. & Berntsen, J.O. (2012) Towards sustainability: Vaccinated *pangasius* in Vietnam. *AQUA Culture Asia Pacific Magazine,* **May/June,** 36-37.

Van Muiswinkel, W.B., Wiegertjes, G.F. & Stet, R.J.M. (1999) The influence of environmental and genetic factors on the disease resistance of fish. *Aquaculture,* **172,** 103-110.

Vo, T.T.B., Nguyen, H.D., Nguyen, T.T.H. & Tran, H.T.T. (2018) Identification and analysis of snps in population of Vietnamese catfish (*Pangasianodon hypophthalmus*), using next generation sequencing and snp validation. *MOJ Current Research & Reviews,* **1,** 12-19.

Vu, N.T., Sang, N.V., Phuc, T.H., Vuong, N.T. & Nguyen, N.H. (2019) Genetic evaluation of a 15-year selection program for high growth in striped catfish *Pangasianodon hypophthalmus*. *Aquaculture,* **509,** 221-226.

Vu, N.T., Sang, N.V., Trong, T.Q., Duy, N.H., Dang, N.T. & Nguyen, N.H. (2019) Breeding for improved resistance to *Edwardsiella ictaluri* in striped catfish (*Pangasianodon hypophthalmus*): Quantitative genetic parameters. *Journal of Fish Diseases,* **42,** 1409-1417.

Yáñez, J.M., Lhorente, J.P., Bassini, L.N., Oyarzún, M., Neira, R. & Newman, S. (2014) Genetic co-variation between resistance against both *Caligus rogercresseyi* and *Piscirickettsia salmonis*, and body weight in Atlantic salmon (*Salmo salar*). *Aquaculture,* **433,** 295-298.

Ødegård, J., Baranski, M., Gjerde, B. & Gjedrem, T. (2011) Methodology for genetic evaluation of disease resistance in aquaculture species: challenges and future prospects. *Aquaculture Research,* **42,** 103-114.

Ødegård, J., Gitterle, T., Madsen, P., Meuwissen, T.H.E., Yazdi, M.H., Gjerde, B., Pulgarin, C. & Rye, M. (2011) Quantitative genetics of taura syndrome resistance in pacific white shrimp (*penaeus vannamei*): a cure model approach. *Genetics Selection Evolution,* **43,** 14.

Ødegård, J., Madsen, P., Labouriau, R., Gjerde, B. & Meuwissen, T.H.E. (2011) A sequential threshold cure model for genetic analysis of time-to-event data. *J. of Animal Science,* **89,** 943-950.

Ødegård, J., Olesen, I., Dixon, P., Jeney, Z., Nielsen, H.M., Way, K., Joiner, C., Jeney, G., Ardó, L., Rónyai, A. & Gjerde, B. (2010) Genetic analysis of common carp (*Cyprinus carpio*) strains. II: Resistance to koi herpesvirus and *Aeromonas hydrophila* and their relationship with pond survival. *Aquaculture,* **304,** 7-13.

Ødegård, J., Olesen, I., Gjerde, B. & Klemetsdal, G. (2006) Evaluation of statistical models for genetic analysis of challenge test data on furunculosis resistance in Atlantic salmon (*Salmo salar*): Prediction of field survival. *Aquaculture,* **259,** 116-123.

Ødegård, J., Olesen, I., Gjerde, B. & Klemetsdal, G. (2007) Positive genetic correlation between resistance to bacterial (furunculosis) and viral (infectious salmon anaemia) diseases in farmed Atlantic salmon (*Salmo salar*). *Aquaculture,* **271,** 173-177.

Ødegård, J., Sommer, A.I. & Præbel, A.K. (2010) Heritability of resistance to viral nervous necrosis in Atlantic cod (*Gadus morhua* L.). *Aquaculture,* **300,** 59-64.

453 Table 1. Number of families, sires, dams and fish in four challenge-test experiments of Mekong
454 striped catfish with *E. ictaluri*. Experiment (Exp) 1 was carried out in only one tank (t1), while
455 experiments 2, 3 and 4 had two replicated tanks.

Exp	Identity	No. of families	No. of sires	No. of dams	No. of fish
1	Exp1-t1	81	54	80	2,155
2	Exp2-t1	64	41	63	1,019
	Exp2-t2	60	40	59	969
3	Exp3-t1	187	118	183	2,944
	Exp3-t2	187	118	183	2,745
4	Exp4-t1	233	137	230	3,246
	Exp4-t2	233	137	230	2,931

456

Table 2. Estimates of additive genetic sire-dam variance (σ^2_{sd}), common environmental variance (σ^2_c) and heritability (h^2) of survival at the endpoint and at 50%-mortality (not for LSM) by use of three statistical models in three challenge-test experiments (Exp) of Mekong striped catfish with *E. ictaluri*.

		σ^2_{sd}		σ^2_c		h^2 (±SE)	
Exp	Model [†)]	Endpoint	50%	Endpoint	50%	Endpoint	50%
1 [‡)]	LM	4.6×10^{-3}	-	5.5×10^{-3}	-	0.100 (0.071)	-
	TM [§)]	5.2×10^{-2}	-	4.8×10^{-2}	-	0.180 (0.120)	-
	LSM	4.4×10^{-5}	-	7.9×10^{-5}	-	0.010 (0.008)	-
3	LM	2.9×10^{-3}	8.3×10^{-3}	1.2×10^{-3}	4.2×10^{-3}	0.085 (0.030)	0.135 (0.042)
	TM	4.9×10^{-2}	6.3×10^{-2}	2.0×10^{-2}	2.4×10^{-2}	0.174 (0.066)	0.220 (0.065)
	LSM	1.9×10^{-4}	-	1.2×10^{-4}	-	0.012 (0.004)	-
4	LM	1.1×10^{-3}	4.1×10^{-3}	4.5×10^{-3}	7.8×10^{-3}	0.044 (0.024)	0.084 (0.033)
	TM	2.4×10^{-2}	3.6×10^{-2}	1.2×10^{-1}	7.1×10^{-2}	0.083 (0.064)	0.125 (0.054)
	LSM	2.7×10^{-4}	-	4.5×10^{-4}	-	0.012 (0.004)	-

[†)] The models were: LM: Cross-sectional linear model; TM: Cross-sectional threshold model, and LSM: Linear survival model.

[‡)] Experimental survival < 50%.

[§)] In TM, the residual variance = 1.

Table 3. Pearson correlation coefficients between the observed family survival (y) in one tank (t1 or t2) to the predicted family survival (sum of sire, dam and common environmental effects = y-hat) in the other tank, in experiments (Exp) 3 and 4 both at the endpoint and at 50% mortality, calculated with either a linear model (LM), a threshold model (TM), or a linear survival model (LSM, not at 50% mortality). P-values are given for the test of difference between pairs of correlations.

Exp	r	Model	Endpoint	50%	P-value		
					Endpoint		50%
					LM	TM	LM
3	$r_{y\text{-}t1,y\text{-}hat\text{-}t2}$	LM	0.278	0.279	-	0.130	-
		TM	0.257	0.263	-	-	0.177
		LSM	0.297	-	0.534	0.219	-
	$r_{y\text{-}t2,\ y\text{-}hat\text{-}t1}$	LM	0.330	0.350	-	0.812	-
		TM	0.334	0.351	-	-	0.879
		LSM	0.352	-	0.463	0.617	-
4	$r_{y\text{-}t1,y\text{-}hat\text{-}t2}$	LM	0.407	0.336	-	< 0.001	-
		TM	0.286	0.304	-	-	0.021
		LSM	0.388	-	0.577	0.018	-
	$r_{y\text{-}t2,\ y\text{-}hat\text{-}t1}$	LM	0.381	0.283	-	< 0.001	-
		TM	0.267	0.243	-	-	0.008
		LSM	0.353	-	0.407	0.036	-

Table 4. Spearman rank correlation coefficients between predicted family survival across the two tanks (sum of sire, dam and common environmental effects) in experiments (Exp) 3 and 4 both at the endpoint and at 50% mortality, calculated with either a cross-sectional linear model (LM), a threshold model (TM), or a linear survival model (LSM, not at 50% mortality).

Exp	Model	Endpoint		50%	Endpoint - 50%		
		TM	LSM	LM	LM	TM	LSM
3	LM	0.99	0.89	-	0.71	0.72	0.92
	TM	-	0.89	0.99	0.71	0.71	0.91
4	LM	0.92	0.81	-	0.59	0.59	0.87
	TM	-	0.83	0.98	0.56	0.60	0.85

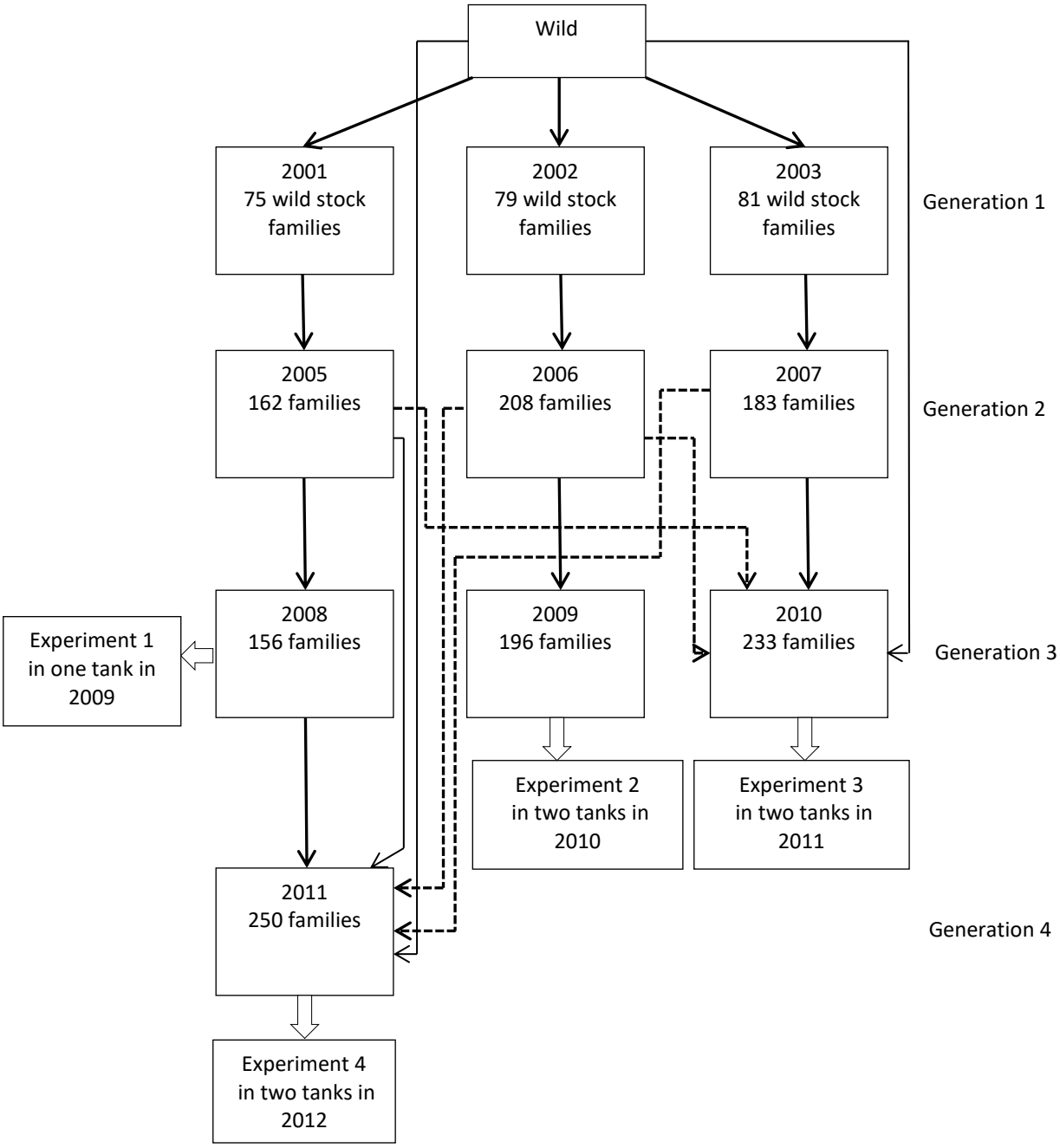
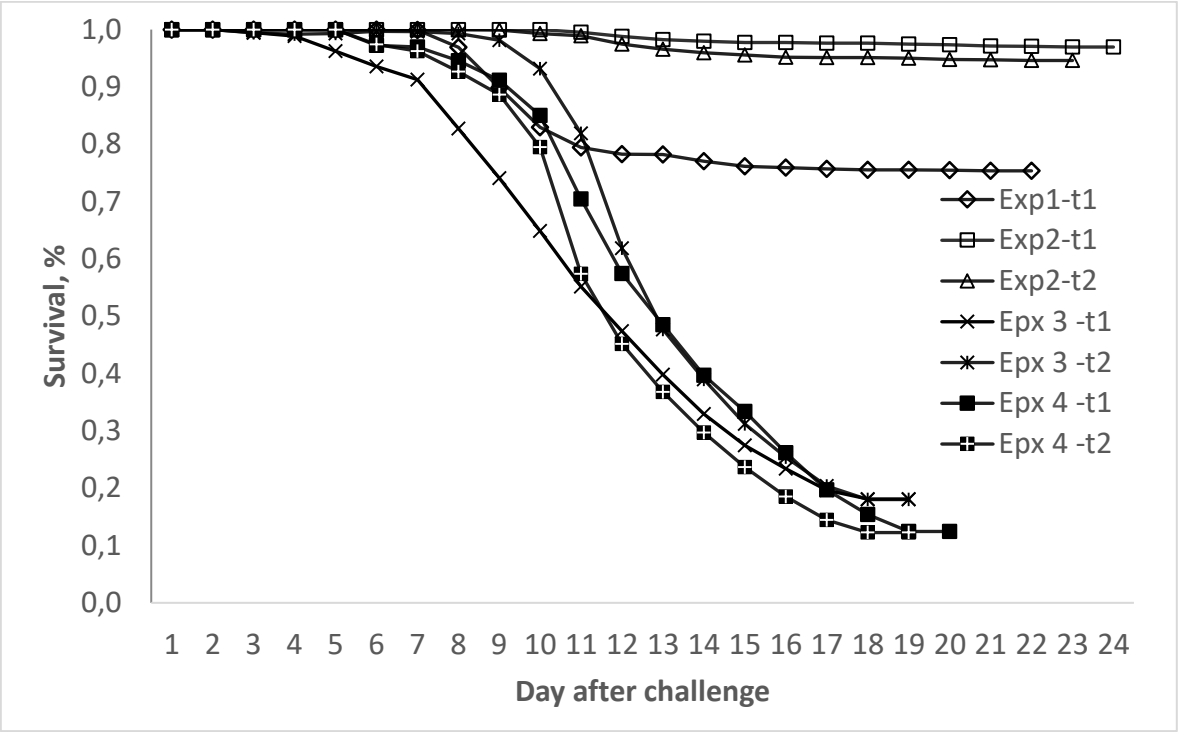


Figure 1. Schematic representation of the breeding program with Mekong striped catfish in Vietnam; year-classes, sub-populations and generations in which the challenge-test experiments were carried out is shown. Number of families per year-class is given. Dashed arrows indicate broodstock usage across year-classes.



515

516 Figure 2. Cumulative Kaplan-Meier survival curves in four challenge-test experiments of
517 Mekong striped catfish with *E. ictaluri*: Experiment (Exp) 1 was carried out in only one tank (t1),
518 while experiments 2, 3 and 4 had two replicated tanks.

519

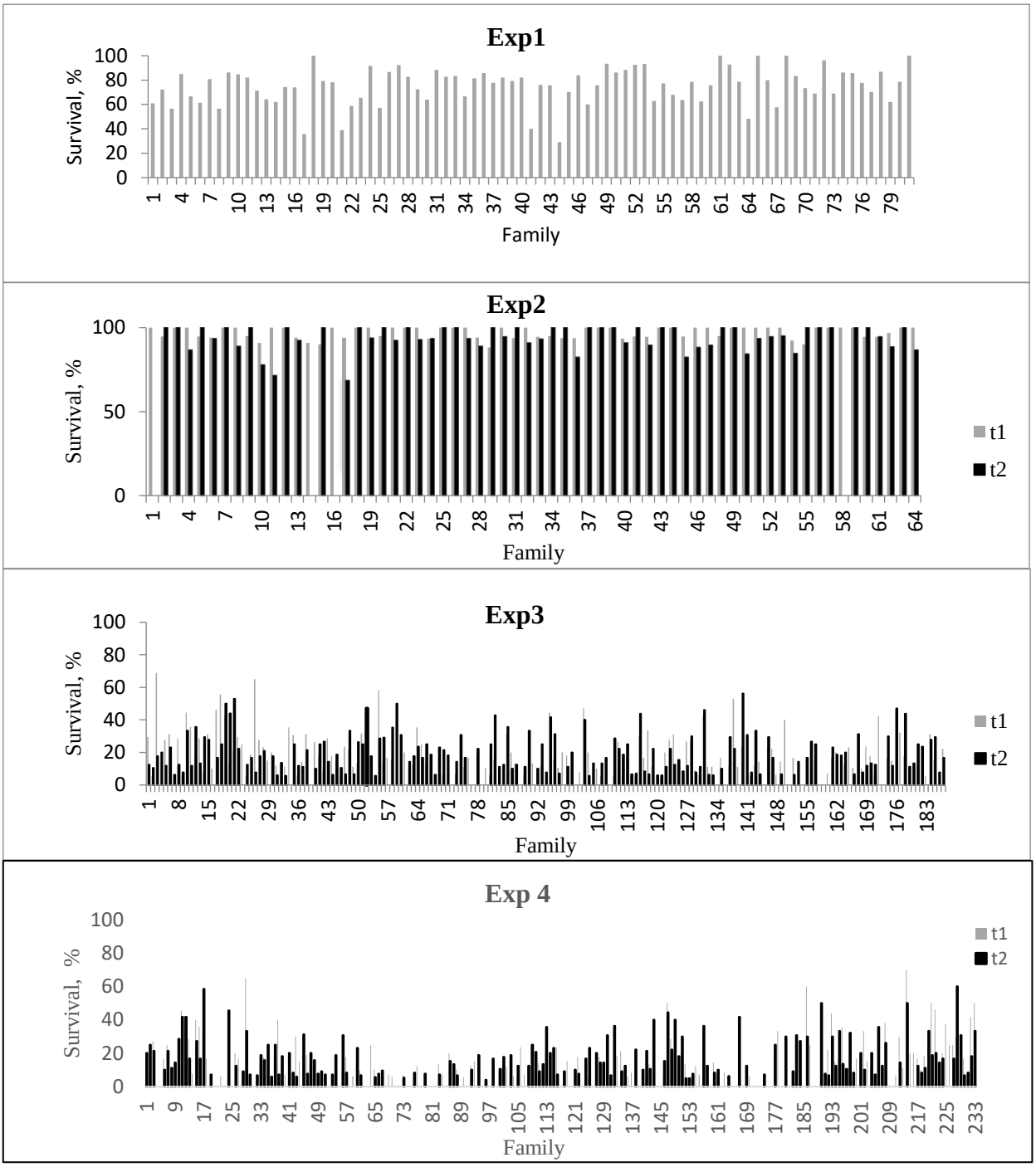


Figure 3. Terminal survival (number of survivors to number of test fish at the start of the tests) by family in four challenge-test experiments of Mekong striped catfish with *E. ictaluri*. Experiment (Exp) 1 was carried out in only one tank, while experiments 2, 3 and 4 had two replicated tanks (t1 and t2).