Aquaculture Science and Management

RESEARCH

Open Access

Impact of genetic and genomic factors on host traits in striped catfish infected with *Edwardsiella ictaluri*

Nguyen Hong Nguyen[1,2*], Nguyen Thanh Vu[3,4], Tran Thi Mai Huong[2] and Tran Huu Phuc[3]

Abstract

Infectious diseases pose a significant threat to sustainable agricultural production. While conventional quantitative genetic theory has successfully enhanced the resistance of animal individuals or populations, it fails to consider epidemiological factors. As a result, it may not adequately capture the maximum genetic gains in selected populations. This study aims to address this limitation by employing a genetic-epidemiological model that enables the estimation of genetic parameters for three host traits: susceptibility, infectivity, and recovery (or removal/mortality). We conducted our analysis on a population of striped catfish (*Pangasianodon hypophthalmus*) exposed to the *Edwardsiella ictaluri* pathogen, which causes bacillary necrosis of pangasius (BNP) disease, through challenge test experiments using injection and cohabitation methods. A total of 560 individuals (490 offspring and 70 parents) were evaluated for disease resistance, measured as the time (in days) from the challenge test to death. Our analysis using the genetic-epidemiological model revealed significant heritability in the epidemiological host traits. The genetic variances for infectivity were found to be greater than those for susceptibility and mortality. Additionally, genetic correlations of susceptibility with infectivity and mortality were moderate and negative, while those between infectivity and mortality were positive. Significant SNPs obtained from our genome-wide scan exhibited small additive genetic and non-significant (or incomplete) dominant effects, suggesting polygenic nature of epidemiological host traits. Genomic prediction accuracies for the transition time between susceptibility and infectivity, as well between infectivity and mortality were moderate to high (0.16 – 0.73). These findings suggest promising prospects for improving epidemiological host traits in genetic programs to enhance the overall resilience of the striped catfish population. The selection index approach yielded a predicted genetic gain ranging from 5.5 – 10.3% per generation for the epidemiological host traits. The accuracy of the selection index was moderate (0.585). Our study provides fundamental genetic parameters for modelling alternative selection strategies aimed at increasing disease resilience to infectious diseases in striped catfish and other aquaculture species.

Keywords Genetic-epidemiological model, Genetic variances, Heritability, Disease resistance, Genetic improvement

*Correspondence:
Nguyen Hong Nguyen
nnguyen@usc.edu.au
Full list of author information is available at the end of the article

Introduction

Infectious diseases caused by microorganisms, including parasites, bacteria, and viruses, have significant impacts on individual health, economics, food security, and animal welfare. In the agricultural sector, estimated economic losses attributed to infectious diseases reach up to 50% of turnover in developing countries and approximately 20% in developed countries [1]. An illustrative example of the economic losses for striped catfish alone due to *Edwardsiella ictaluri*, especially during the fingerling stage, ranged from 48 to 94 million US dollars per annum [2]. Theoretical predictions and modelling approaches indicate that the severity of infectious diseases could escalate due to changing environments and intensified production systems, leading to increased stress, immunosuppression, and frequent disease outbreaks. As a result, these diseases compromise the fitness and performance of hosts by disrupting vital biological processes, such as reallocating nutritional resources for fighting infections rather than for production, and by reducing host immunity, consequently impacting animal welfare [3].

Conventional methods for controlling infectious diseases comprise chemical interventions (e.g., antibiotics), husbandry and management practices (e.g., sanitation and disinfection), culling or isolating sick individuals, and implementing control measures like movement restrictions and vaccination. However, none of these methods prove to be cost-effective, permanent, or sustainable in the long run [4]. Genetics has demonstrated as a powerful tool for understanding and mitigating pathogen risks while enhancing host resilience to infectious diseases through genetic selection [5]. In farmed animals, including aquaculture species, genetic selection approach used to improve disease resistance (tolerance) relies on survival data collected from pathogen challenge tests [6–8] or field environments [9]. Typically, survival data is recorded as presence or absence of a disease of interest. Other measures including pathogen load [10], immune response [11] or indicators of animal health [12] are less frequently recorded. These traits are analysed using various quantitative genetic models, such as linear mixed model for continuous expressions or the liability threshold (generalised) linear mixed model for categorical characters [13]. These mixed models account for systematic effects, random factors, and pedigree relationships, resulting in substantial improvements in disease resistance, with average genetic gains ranging from 5 to 15% per generation. For instance, in fish species like Atlantic salmon, resistance to yerisiosis has increased by approximately 15% [14], while white leg shrimp (crustaceans) and Pacific oysters (molluscs) have gained approximately 7–8% for resistance to white spot disease and 10% for OsHV-1 virus, respectively [15, 16].

Despite these successes, quantitative genetic models have inherent limitations as they do not encompass epidemiological traits or parameters, making them suboptimal for genetic programs targeting infectious diseases. Recently, there has been a growing interest in incorporating epidemiological models to aid in risk identification, determination of epidemic duration, evaluation of infection rates, recoveries, and mortalities, as well as the assessment of intervention impacts on these parameters [17]. The epidemiological model focuses on three primary parameters: susceptibility (S), infectivity (I), and recovery (R), forming the SIR model. Susceptibility refers to individuals who are prone to be infected or are more likely to be infected. Infectivity represents the ability of an infected individual to transmit the infection, with higher infectivity indicating an increased risk of transmission and epidemic outbreaks. Recovery signifies the duration it takes for an individual to recuperate after infection and can also refer to the time until removal or death occurs. In our study, we refer it as removal or mortality. In the SIR model, two parameters need to be estimated: transmission (β) and recovery (γ) rates [18]. They are important inputs to model epidemiological genetics of animals. Thanks to these features, the SIR model has been integrated with quantitative genetics theory, leading to the development of the genetic-epidemiological model.

The genetic-epidemiological model (GEM) made its initial debut in the study of gastrointestinal infections in sheep [19], highlighting the underestimation of response to selection for increased pathogen load in this parasite when the epidemiological model was disregarded. Subsequently, the GEM was further refined to predict the impact of selective breeding on reducing the prevalence of infectious diseases in terrestrial and aquatic animal species [20–22]. Moreover, the GEM model enables the estimation of genetic effects for distinct traits, including susceptibility, infectivity, and mortality. A review of published literature (unpublished results) reveals substantial genetic variations in these traits, confirming their genetically distinct nature. Genetic correlations between susceptibility and infectivity exhibit a negative relationship, while those between infectivity and recovery demonstrate a positive association [23]. With reliable genetic parameter estimates for these traits when the data are available, they can be amalgamated into a selection index, facilitating a multi-trait selection program to enhance the resilience of animals/populations against infectious diseases. Furthermore, this approach offers a pathway to improving overall population resilience rather than solely focusing on the individual level [24].

To date, only a few studies have reported genetic variances for susceptibility, infectivity, and recovery traits in aquaculture species, specifically anaemia virus in Atlantic salmon [25], *Philasterides dicentrarchi* in turbot [23] and cyprinid herpes virus type 3 (CyHV-3) in common carp [26]. Hence, the present study aims to utilise the GEM model to estimate epidemiological traits of striped catfish (*Pangasianodon hypophthalmus*) exposed to the *E. ictaluri* pathogen in a challenge test experiment. The ultimate objective of our study is to refine breeding strategies that enhance the overall resilience of the striped catfish population against *E. ictaluri*, a causative agent of bacillary necrosis of pangasius (BNP) disease in this species.

Materials and methods

Animals

The experimental animals in this study were obtained from a selective breeding program aimed at enhancing disease resistance to the Bacillary Necrosis of Pangasius (BNP) caused by *E. ictaluri* in striped catfish at Research Institute for Aquaculture No.2 (RIA2), Vietnam [7]. In 2020, the first generation of catfish involving 266 broodstock (166 females and 100 males selected from a separate breeding program for high growth) was produced employing a nested mating design with a male-to-female ratio of 1:2. Following the breeding protocol described in our previous studies [27], a total of 166 families (comprising 32 full-sib families and 134 half-sib families) were successfully produced in 2021.

The fry from each family were raised in separate fiberglass tanks (1.5m^3) for a duration of three weeks before being transferred to a net hapa system installed in an earthen pond. During this period, they were fed a high protein diet (40%) three times a day. Upon reaching an average body weight of 15–20 g (about 2–3-month-old), 100 fingerlings per family were randomly collected and individually identified using PIT (Passive Integrated Transponder) tags. Subsequently, half of each family was allocated for grow-out in ponds, while the other half that was pathogen-free based on our PCR test (averaging three fish per family) was utilised for pathogen challenge tests to assess *E. ictaluri* resistance.

Challenge test and data collection

The challenge test comprised a total of 5,328 individuals derived from 166 families, with an average of 32 individuals per family. Firstly, the experimental fish were acclimatised in cement tanks for approximately two weeks. Following this acclimation period, an equal number of fish from each family were randomly assigned to different cement tanks, each with a capacity of 10 m3, to conduct the challenge test using the cohabitation method [7].

The cohabitant fish (originated from the same families as the experimental animals with an average weight of 16.7 ± 6.1 g) were initially injected with the *E. ictaluri* pathogen at a dosage of 10^6 CFU/0.2 ml per fish [7]. Two days after the injection, they were introduced into the cement tanks (10 m^3) to rear together with the experimental fish at a ratio of 1 cohabitant fish to 3 experimental fish (or approximately 30% cohabitant fish or averaging 24 indviduals in each tank). The bacterial density necessary for disease infection was maintained by adding the bacteria to the experimental tanks on day 4, at a density of 10^5 CFU per one ml of rearing water, that is 1 L of medium containing 10^9 CFU of the pathogen in each 10 m^3-tank. The duration of the experiment spanned 23 days, during which no fish mortality was observed. The bacterial strain tested from our previous experiments at RIA2 was used in this study.

During the experimental period, the feeding rate was reduced from 3.0% to 1.5% of the total biomass in each tank. The highest mortality rate occurred on day 5, and deceased fish were sampled for laboratory PCR testing to confirm that the symptoms leading to their demise (such as white spots in the spleen, liver, and kidney) were indeed caused by the *E. ictaluri* pathogen. Upon completion of the experiment, all surviving fish were properly disposed, following the biosecurity burial procedures in accordance with the regulations set by the national veterinary authority (the Department of Animal Health, Vietnam).

Throughout the challenge test, fish that were dead due to the infection were collected every three hours until 10 PM and subsequently in the next morning (5 AM), and their clinical symptoms were recorded. These collected data were used to compute two indicators of E. *ictaluri* resistance, namely survival status and survival time. Survival time was defined as a continuous trait, measured from the beginning of the test until the point of the fish's demise, expressed in days. In this study, we analysed only the mortality time. When phenotypic measurements were made, fin tissue samples were also collected from individual fish for genotyping. Only 490 samples (12–15 fish per family × 40 families from a single generation), along with 70 parental specimens, were randomly chosen for the genotyping and subsequent genetic and genomic analyses.

Genotype data

DNA samples were extracted from the dorsal fin tissues of 560 individual fish and subsequently analysed using Diversity Arrays Technology sequencing (DArTseq™). DArTseq™ is an innovative approach that combines complexity reduction techniques from DArT (Diversity Arrays Technology) with cutting-edge next-generation

sequencing platforms [28]. Detailed information of the library constructions is given in Vu et al. [29].

The resulting sequences underwent thorough analysis through proprietary DArT analytical pipelines. SNP calling involved clustering all tags from all libraries included in the DArTsoft14 analysis, utilising DArT PL's C++ algorithm with a threshold distance of 3. The pipeline employed a reference-free approach. The average call-rate for variants was 99%, and for individual samples, it was 92%. This led to the identification of 14,154 single nucleotide polymorphisms (SNPs) using DArT analytical pipelines (Supplementary Figure S1).

To ensure data quality, these SNPs were subjected to additional quality control measures using the dartR package [30], where loci with call rates below 95%, minor allele frequencies below 5%, and significant departures from Hardy–Weinberg equilibrium ($P<0.05$) were excluded. After this stringent filtering, 6,470 high-quality SNPs remained (averaging 209 SNPs per chromosome, ranging from 74 to 314), forming the basis for subsequent genetic and genomic analyses.

Estimation of genetic (co)-variances

We utilised a genetic-epidemiological model to analyse our data. This model assumes that the spread of disease within a contact group (i.e., in the same challenge test tank) follows the SIR model, where individuals' transition rates are influenced by both systematic and random effects [31]. The SIR model, as described in Pooley et al. [18] and [32], classifies individuals as susceptible to infection (S), infected (I), and mortality, that is dead fish were removed from the experimental facilities (R).

The time-dependent force of infection for a susceptible individual, denoted as j, is represented by $\lambda j(t)$, indicating the probability per unit time of becoming infected. For those individuals who do become infected, their infectious duration is assumed to follow a gamma distribution with a mean of wj and a shape parameter of k. We can express these quantities as follows:

$$\lambda j(t) = \beta e^{Gz} e^{g_j} \sum_i e^{f_i},\ w_j = (\gamma e^{r_j})^{-1} \tag{1}$$

In Eq. 1, β and γ represent the population average transmission and recovery rates, respectively. Gz, known as the "group effect" (where z indexes the contact group), accounts for group-specific factors that influence the speed of an epidemic within one contact group relative to another. These factors may include variations in animal management conditions, environmental differences, or pathogen strains with varying virulence. Gz is treated as a random effect with a standard deviation of σ_G, following a normal distribution. The expression for w_j does not include a group effect because we assume that the environment primarily affects the speed of infection spread within groups, rather than individual susceptibility, infectivity, or mortality.

Additionally, Eq. 1 involves g_j, f_i, and r_j, which represent the fractional deviation in individual j's susceptibility, individual i's infectivity, and mortality, respectively, compared to the population. These deviations can be decomposed into various main factors denoted as g, f, and r, respectively (Eq. 2 below).

$$\begin{aligned} \boldsymbol{g} &= \mu + \boldsymbol{X}\boldsymbol{b}_{\mathbf{g}} + \boldsymbol{a}_{\mathbf{g}} + \boldsymbol{e}_{\mathbf{g}} \\ \boldsymbol{f} &= \mu + \boldsymbol{X}\boldsymbol{b}_{\mathbf{f}} + \boldsymbol{a}_{\mathbf{f}} + \boldsymbol{e}_{\mathbf{f}} \\ \boldsymbol{r} &= \mu + \boldsymbol{X}\boldsymbol{b}_{\mathbf{r}} + \boldsymbol{a}_{\mathbf{r}} + \boldsymbol{e}_{\mathbf{r}} \end{aligned} \tag{2}$$

where μ is a SNP effect; $\boldsymbol{b}_g$, $\boldsymbol{b}_f$ and $\boldsymbol{b}_r$ corresponds to fixed effects, i.e., spawning batch and age from birth to commencement of the challenge test. $\mathbf{X}$ represents a design matrix that can account for differences in traits. The additive genetic contributions $\mathbf{a}=(\boldsymbol{a}_g, \boldsymbol{a}_f, \boldsymbol{a}_r)$ capture the relationships in trait values between different individuals and follow a multivariate normal distribution with a mean of zero and a covariance matrix $\mathbf{A}\otimes\boldsymbol{\Omega}$. Here, $\mathbf{A}$ represents the pedigree relationship matrix estimated from all 560 individuals, and $\boldsymbol{\Omega}$ is a 3×3 covariance matrix that describes potential correlations between traits (susceptibility, infectivity, and mortality). Finally, the residual contributions $\boldsymbol{\varepsilon}=(\boldsymbol{\varepsilon}_g, \boldsymbol{\varepsilon}_f, \boldsymbol{\varepsilon}_r)$ in Eq. (2) account for all other variations. These residuals also follow a multivariate normal distribution with a mean of zero and a covariance matrix $\mathbf{I}\otimes\boldsymbol{\Psi}$, where $\mathbf{I}$ is the identity matrix reflecting the assumption of uncorrelated residuals between individuals, and $\boldsymbol{\Psi}$ is a 3×3 covariance matrix that characterizes environmental correlations between traits.

Our analysis of the SIR model utilised the SIRE 2.0 package, which leverages the Monte Carlo Markov Chain (MCMC) approach for Bayesian inference of genetic parameters [18]. By employing default settings, with 10,000 samples and burn-in of 2000 samples in three independent runs, we ensured convergence (estimated effective sample size of over 200 for each model parameter).

The heritability for host traits was calculated as $h^2 = \frac{\widehat{\sigma}_a^2}{\widehat{\sigma}_a^2+\widehat{\sigma}_e^2}$ where $\widehat{\sigma}_a^2$ is the additive genetic variance, and the residual variance ($\widehat{\sigma}_e^2$). The genetic and phenotypic correlations between the traits were estimated as: $r = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}}$, where the numerator represents covariance between the two traits and the denominator specifies the genetic or phenotypic variance of individual traits (x and y).

Accuracy of genomic prediction

Since the disease status (susceptibility and infectivity) of individual animals couldn't be determined reliably under group testing in the tanks, we conducted an evaluation of the genomic prediction accuracy for the transition time between susceptibility and infectivity (i.e., the time taken for individual *i* to transition from S → I), as well as between infectivity and mortality (i.e., the time taken for individual *i* to transition from I → R). These data were extracted from the SIR model analysis (Genotype data section) and analysed using the genomic best linear unbiased prediction (GBLUP) method in ASReml [33]. In each trait model, we used two sets of SNPs: the full set of 6370 SNPs and a subset of highly significant SNPs ($P < 1e\text{-}05$). The univariate GBLUP along with PBLUP models employed the same fixed and random effects as Eq. 2. To obtain the accuracy of genomic prediction for these traits, we performed a five-fold cross-validation over 20 iterations and computed the correlation between the predicted and actual phenotypes within the validation set, which was then divided by the trait heritability. Additionally, when subsets of SNPs were evaluated, the five-fold cross validation used only the significant SNPs identified from GWAS that was conducted separately for each training data subsets to avoid any possible bias in the prediction accuracy for the trait studied.

Estimation of SNPs effects

In the SIR model, the SNP contribution to the host traits for individual *j* depends on its genotype as the followings:

$$g_j^{SNP} = \begin{matrix} S_g \\ S_g\Delta_g \\ -S_g \end{matrix}, \begin{matrix} S_f \\ f_i^{SNP} = S_f\Delta_f \\ -S_f \end{matrix}, \begin{matrix} S_r \\ f_i^{SNP} = S_r\Delta_r \\ -S_r \end{matrix} \Bigg\} \begin{matrix} \text{if } j \text{ is } AA \\ \text{if } j \text{ is } AB \\ \text{if } j \text{ is } BB \end{matrix}$$

The parameters s_g, s_f and s_r denote half the difference in trait values between the homozygote genotypes (*AA* and *BB*) and Δ_g, Δ_f and Δ_r represent the degree of dominance (a value of 1 (−1) indicates complete dominance of the *A* (*B*) allele over the *B* (*A*) allele), with a value of zero indicating no dominant effect.

We estimated the additive genetic and dominant effects for top significant SNPs obtained from our genome-wide association study (GWAS) for two traits: transition time between susceptibility and infectivity as well between infectivity and mortality. Data for these two traits were extracted from the output generated by Sire 2.0, as described above (heading 2.3). Our GWAS analysis used a multi-locus mixed model. This approach incorporated fixed effects of spawning batch and age, and a kinship matrix was included as a random factor to account for family structure (Supplementary Figure S2). The GWAS analysis was conducted for each training subsets using the blupf90 family program [34], involving three main steps: 1) renumf90 to renumber the data into a standard format used by the program, 2) blupf90 + to perform quality control on the genotype data, compute the breeding values of each individual, and generate the weighted genomic relationship matrix (default for single-step analyses), and 3) the subprogram postgsf90, which reads breeding values along with mapping information for each marker to compute SNP effects and *p*-values. We utilised the CMplot R package to generate Manhattan plots from BLUPf90 outputs, with the significance threshold set at 0.05 divided by the number of markers used (6470 SNPs), approximately 5.111. The GWAS models were the same as those used to compute heritability, with identical fixed and random effects as described above. Regarding genome annotation, we initially established a local SwissProt database, pivotal for subsequent analysis with the BLAST2GO software. Employing blastx with default software parameters, we translated nucleotide fasta sequences into protein annotations. Subsequently, through meticulous mapping and annotation, we unveiled the corresponding GO terms.

Prediction of genetic gain

Finally, we predicted genetic gain using selection index approach. The breeding goal of our genetic program aimed at improving resistance (or resilience) of striped catfish to *E. ictaluri* disease. The traits included in the breeding objective (goal) are susceptibility, infectivity and mortality. However, the selection criterion is only mortality because this trait is easily measured and available in most of genetic improvement programs for aquaculture species, including striped catfish. Economic values for the traits in the breeding goal were arbitrarily set as one. The genetic parameters used are given in Tables 1 and 2. The selection index in a matrix notation is written as:

$$I = bx$$

where I = index value, b = a weighting factor, and x = phenotypic information (susceptibility, infectivity and mortality). A detailed description of the selection index approach is given in Cameron [35].

In the selection index, we made the following assumptions based on a typical structure of a selective breeding program for aquaculture species [36]: i) the pedigree consisted of 100 families (50 sires and 100 dams), ii) there were 20 female and 20 male offspring tested per family that were potential selection candidates, iii) the proportions of selected animals were 15% in females and 7.5% in males, and iv) selection was based on individual performance.

The annual genetic gain (Δ_G) was calculated as:

Table 1 Additive genetic variances and heritability estimates for epidemiological host traits

Traits	Additive genetic variance		Total variance		Heritability	
	Mean	95% CI	Mean	95% CI	Mean	95% CI
Susceptibility	0.48	0.41 — 0.52	1.01	0.45 – 2.10	0.59	0.25 — 0.93
Infectivity	1.76	0.79 — 2.45	4.70	3.59 – 5.45	0.36	0.21 — 0.46
Mortality	0.78	0.37 — 1.01	3.74	3.23 – 4.01	0.20	0.11 — 0.25

$$\Delta_G = [(i_F)(\sigma_I) + (i_M)(\sigma_I)]/(L_F + L_M)$$

where σ_I is the standard deviation of the index, i is the selection intensity ($i_F = 1.554$ and $i_M = 1.887$), and L is the generation interval (three years in both sexes). We assume that in each generation, a total of 4000 fish are recorded (100 families times 40 individuals per family), out of which different proportions of males and females were selected. The proportion of selected females and males was 0.15 and 0.075, corresponding selection intensities of 1.554 (i_F) and 1.887 (i_M), respectively. The number of selected females and males was three times (i.e., 300 females and 150 males) greater than that actually needed, to allow for losses and unsuccessful matings.

With the assumptions as described above, together with the genetic parameters estimated in this study, we predicted genetic gain for each trait in the breeding objective, using SelAction [37]. The genetic gains for host traits were also compared with that of a conventional selection program for a single trait (i.e., only survival/mortality in the breeding goal).

Results

Genotype and phenotype

This study consisted of 560 individuals, including 490 offspring and 70 parents from a single generation, which were genotyped. These individuals were sampled from a population of 5,328 fish and were subjected to pathogen challenge tests involving *E. ictaluri*. Mortality events occurred relatively early, approximately 1.5 days after the start of the experiments, with a significant increase between 5 and 10 days (Fig. 1). The overall mortality rate for the population was 32.5%.

In this experiment, there were 10 distinct contact groups corresponding to 10 spawning batches, each was tested in a different tank with varying size (ranging from 22 to 140 individuals). There were significant differences in mortality times observed between these contact groups ($P < 0.05$), the difference between the highest and lowest groups was 5.4%. Among the contact groups, group #5 exhibited the highest transmission rate, followed by group #7, while group #6 had the lowest transmission rate (Supplementary Table S1).

Genetic variances and heritability estimates for epidemiological host traits

Genetic variances and heritability (h^2) estimates for epidemiological host traits in striped catfish are statistically significant, as indicated by the 95% confidence intervals (CI) presented in Table 1. A 95% CI implies that the interval includes the true value in 95 out of 100 studies. Clearly, all three epidemiological host traits had h^2 estimates that were statistically significant, as their 95% CI did not encompass a zero effect (Table 1).

Among the traits, Infectivity exhibited a larger genetic variance compared to susceptibility and mortality. However, the heritability estimate for susceptibility was higher than that for infectivity, likely due to the larger residual variance observed in the latter trait (2.94 *vs.* 0.53). It is worth mentioning that the heritability estimates for susceptibility may be less reliable due to the wide range of its 95% CI. This suggests a larger variation in the susceptibility than other host traits in the population. In contrast, the heritability estimate for mortality aligns with our independent estimate ($h^2 = 0.16$) derived from a

Table 2 Phenotypic (above the diagonal) and genetic (below the diagonal) correlations among epidemiological host traits (95% confidence interval in bracket)

Trait	Susceptibility	Infectivity	Mortality
Susceptibility		-0.26 (-0.98 — 0.02)	-0.12 (-0.57 — 0.06)
Infectivity	-0.40 (-0.98 — -0.07)		0.81 (0.66 — 0.91)
Mortality	-0.37 (-1.03 — -0.08)	0.45 (-0.58 — 0.80)	

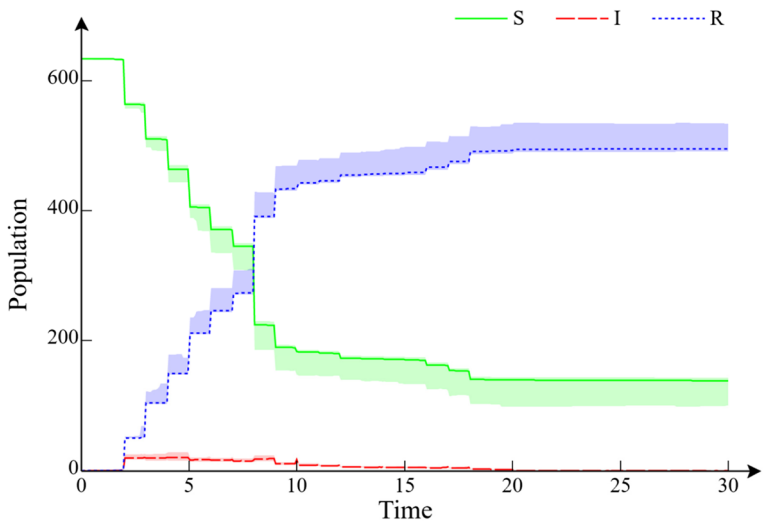

Fig. 1 Population plot for susceptibility (S), infectivity (I) and mortality or removal (R). The Y-axis represents the population size (including parents of sires and dams), and X-axis shows experimental time beyond 23 days to demonstrate the stabilised S, I and R trends

Table 3 Genetic (V_A) and environmental (V_E) variance, heritability (h^2) and accuracy (± SD) of genomic prediction for transition time between susceptibility and infectivity (Time1) and between infectivity and mortality/death removal (Time2)

Trait	V_A	V_E	h^2	Accuracy (only significant SNPs)	Accuracy (Full set 6470 SNPs)	PBLUP
Time1	3.51	6.95	0.33 ± 0.08	0.74 ± 0.23	0.57 ± 0.18	0.58 ± 0.18
Time2	0.92	5.38	0.14 ± 0.08	0.63 ± 0.61	0.36 ± 0.40	0.50 ± 0.37

Number of significant SNPs (P < 1e-05): 519 – 1410

PBLUP Best Linear Unbiased Prediction using pedigree information

conventional quantitative linear mixed model using the restricted maximum likelihood method.

Genetic correlations (r_g) among epidemiological host traits

Epidemiological host traits in our study exhibit genetic correlations, as shown in Table 2. The genetic correlations between susceptibility and infectivity, as well as susceptibility and mortality, were negative (−0.40 and −0.37, respectively). Although these estimates were statistically significant (their 95% CI did not include zero), the confidence intervals were relatively large due to the limited sample size of our dataset.

Infectivity showed a positive genetic correlation with mortality (r_g = 0.45). However, this estimate was not statistically significant, as indicated by its 95% CI.

At the phenotypic level, the correlations among traits matched the sign and direction of the genetic correlation estimates. However, only the phenotypic correlation between infectivity and mortality reached statistical significance (0.81, 95% CI = 0.66 – 0.91).

Accuracy of genomic prediction

In Table 3, we present the estimated accuracies of genomic prediction for two crucial traits: the transition time between susceptibility and infectivity as well between infectivity and mortality. The prediction models were constructed using two sets of SNPs, comprising the full set of 6370 SNPs and a subset of only significant SNPs (P < 1e-05) (Supplementary Figures S2). Interestingly, we observed slightly higher accuracies when employing only the significant SNPs in our prediction models. This was likely by the trait characteristic, or the high quality of the significant SNPs as compared to that of the full set. The prediction accuracies for both traits when the significant SNPs were employed were also higher than those obtained from PBLUP model (Table 3). Overall, the accuracies for both traits were found to be moderate, indicating promising potential for the application of genomic selection to improve epidemiological host traits within this population of stripped catfish. The SNP heritability for these traits were low to moderate (0.14 – 0.33).

Table 4 Additive and dominant effects of highly significant SNPs ($P < 0.5 \times 10^{-6}$). Confidence interval in parenthesis

Effect	Trait	SNP1	SNP2	SNP3	SNP4	SNP5
Additive genetic effect	Susceptibility	−0.166 (−0.323 — −0.003)	−0.165 (−0.394 — 0.145)	0.10023 (−0.17902 — 0.36046)	0.099263 (−0.18184 — 0.36457)	0.085369 (−0.17638 — 0.30048)
	Infectivity	2.838 (2.507 — 2.996)	1.018 (−2.847 — 2.996)	−0.41215 (−2.8554 — 2.9973)	−0.41381 (−2.8610 — 2.9971)	−0.46826 (−2.7479 — 2.9971)
	Mortality	2.007 (1.825 — 2.172)	0.109 (−2.229 — 1.484)	−0.39373 (−1.8884 — 1.7189)	−0.39331 (−1.8873 — 1.7174)	−0.40703 (−1.7500 — 1.7058)
		0.361 (−0.705 — 0.966)	−0.1011 (−0.951 — 0.906)	−0.27582 (−0.98661 — 0.95406)	−0.27845 (−0.98630 — 0.95614)	−0.27753 (−0.98620 — 0.94617)
Dominant effect	Susceptibility	0.476 (0.036 — 0.916)	0.416 (−0.961 — 0.987)	0.33717 (−0.96260 — 0.99563)	0.33513 (−0.96260 — 0.99609)	0.18858 (−0.97108 — 0.99534)
	Infectivity	0.913 (0.722 — 0.999)	0.319 (−0.998 — 0.998)	0.33961 (−0.87130 — 0.99548)	0.34061 (−0.87225 — 0.99552)	0.31178 (−0.88221 — 0.99430)
	Mortality	−0.166 (−0.323 — −0.0029)	−0.165 (−0.394 — 0.145)	0.10023 (−0.17902 — 0.36046)	0.099263 (−0.18184 — 0.36457)	0.085369 (−0.17638 — 0.30048)

Detailed information of SNPs provided in Table S2

Table 5 Predicted genetic gain using selection index

Estimate	Multi-trait selection			Single trait
	Susceptibility	Infectivity	Mortality	Mortality
Genetic gain in trait unit (days)	0.860	0.903	0.488	0.597
Genetic gain expressed as percent of the population mean	9.8	10.3	5.5	6.8
Variance of the index	0.458			0.059
Variance of the breeding goal	1.341			0.333
Accuracy of index	0.585			0.422

Additive genetic and dominant effects of SNPs

Table 4 shows the additive genetic and dominant effects of the top five significant SNPs obtained from our GWAS analysis for removal (death) time (Supplementary Table S2). Regarding the first SNP, we observed a negative additive genetic effect on susceptibility, while its effects on infectivity and mortality were positive. For the remaining SNPs, both the additive genetic and dominant effects were not significant, as their confidence intervals included the value zero. However, the dominant effects on epidemiological host traits were different from zero or one (−1), suggesting they follow an incomplete dominant mode of inheritance. These results highlight the complex mechanisms underlying genetic variation in disease susceptibility and transmission dynamics.

Predicted genetic gain

Table 5 presents genetic gains per generation for individual traits in the breeding objectives, standard deviation of the index and of the breeding goal, accuracy of selection, and overall gain in economic units. Genetic gains in susceptibility, infectivity and mortality are as expected from their genetic variances and genetic correlations (0.860, 0.903 and 0.488 days or 9.8, 10.3 and 5.5% per generation, respectively). The variance of the index and of the breeding goal was 0.458 and 1.341, respectively. Accuracy of the selection index was moderate (0.585). When comparing these values with those of a single trait selection (i.e., for mortality only), the multi-trait selection index showed somewhat greater accuracy while yielding slightly lower genetic gain (Table 5).

Discussion

Genetic variation in host traits

In this study, we have demonstrated the presence of heritable genetic variations in three important epidemiological host traits: susceptibility, infectivity, and mortality. As expected, the genetic variance observed for infectivity was found to be greater than that of susceptibility and mortality within this specific population. These results support previous findings that there exists significant genetic variation underlying disease prevalence, particularly in terms of host infectivity. Until now, only two studies focusing on aquaculture species have reported significant genetic variances for epidemiological host traits. In a study on Atlantic salmon, Chose-Topping et al.

[25] discovered that fish with varying resistance capacities to the infectious salmon anaemia virus displayed different levels of susceptibility and endurance against the disease. Similarly, an analysis of both simulated and empirical data in turbot confirmed significant genetic variances in host infectivity, as reported by Prentice et al. [23]. Additionally, investigations conducted on farmed animals utilised an extended version of the standard threshold linear mixed model to jointly estimate genetic parameters for susceptibility and recovery in dairy cattle. These studies, conducted by Kulkarni et al. [38] and Barden et al. [39], revealed that susceptibility and recovery are distinct genetic traits. These findings suggest that genetic selection can be utilised not only to enhance disease resistance (tolerance/resilience) and/or decrease infectivity but also to identify individuals who are less susceptible and less likely to transmit diseases within a population [22]. Simulation studies have further provided evidence supporting our conclusions that genetic selection can effectively reduce the prevalence of infectious diseases in farmed animals and ecological populations [40, 41].

Moreover, simultaneous improvements in multiple epidemiological traits through genetic selection can be realised, given the positive genetic correlations observed between susceptibility, infectivity, and mortality in this population of striped catfish. Our estimates align with those reported in the study conducted on turbot by Prentice et al. [23]. Collectively, the published information, combined with our study results, highlights the potential for genetic selection to enhance disease resistance, decrease infectivity, minimise susceptibility, and reduce disease transmission within our population. This knowledge can contribute significantly to the development of strategies aimed at mitigating the impact of infectious diseases on aquaculture species and farmed animals.

Accuracy of genomic prediction and SNPs effects

To date, genomic prediction has primarily focused on disease status or disease resistance indicators in animals and plants. Although the prediction accuracy for infectious diseases caused by bacteria, viruses, and parasites in these species has generally been moderate to low (0.25—0.65) [42–45], our evaluation of prediction accuracy for epidemiological host traits falls within this range. Surprisingly, when the full set of SNPs was analysed, the prediction accuracy using GBLUP model was not higher than those obtained from PBLUP; however, the GBLUP model performed better than PBLUP when significant SNPs were employed, such as 0.73 vs. 0.58 for the transition time from susceptibility to infectivity. Collectively, these results suggest the potential for implementing genome-based selection programs to expedite genetic advancements in these traits, although a comparison should be made with future predictions for disease status (susceptibility and infectivity) once these data are reliably determined in this population. In addition to determining genetic parameters that offer fundamental insights into the quantitative genetic basis of epidemiological host traits, we also investigated SNP effects to unravel the genomic architecture underlying the susceptibility, infectivity, and mortality of striped catfish. Our findings support the evidence that multiple genes, each exerting a minor effect, contribute to the control of the three epidemiological host traits. Furthermore, no significant or incomplete dominant effects of the top SNPs ($P < 2.25927E\text{-}37$) on these traits were observed, indicating a complex genetic regulation. Moreover, we attempted to map the sequences of these significant SNPs to the available published genome assemblies for striped catfish [46]. Nevertheless, none of the highly significant SNPs ($P < 1e\text{-}05$) were associated with immune response. Majority of the SNPs linked with known biological functions are not significant. Examples of these genetic variants include those involved in the positive regulation of interferon production (IFN-α and IFN-β), positive regulation of interleukin-6 production (IL6), regulation of inflammatory responses (NF-κB genes), leukocyte migration during immune responses (chemokine receptor and ligand genes), mediation of cytokine signaling (JAK1, JAK2, and STAT1), and defence processes (defensin beta, DEFB gene) (Supplementary Figure S3). However, potential candidate genes were not validated in this population, which is due to the short sequence obtained by genotyping by sequencing platform (only 68bp), and the incomplete genome assembly for the studied species. Therefore, future studies should aim to increase the depth of sequence coverage or re-sequencing a larger number of indviduals and families to obtain more comprehensive insights into the biological functions and pathways of genetic variants and functional mutations involved in epidemiological host traits of this striped fish population. Moreover, due to the potential overestimation of the prediction accuracy via the five-fold cross-validation approach [44], there is a need to verify these results using progeny performance data in future generations of the selection program aimed at improving disease resistance to *E. ictaluri* in this population.

Prospects to improve individual and population resilience

Additionally, our findings provide genetic parameters to model selection strategies to improve the overall resilience of the population [24]. Population resilience encompasses the capacity of a group of animals to maintain high production performance even in the face of pathogen challenges, or their ability to remain

minimally affected by exposure to infectious diseases [47, 48]. By emphasising the improvement of population resilience, we can realise two key advantages: (i) capturing the direct effects of host genetics on the production performance and fitness of individuals, and (ii) reducing the environmental pathogen load to which the population is exposed. This approach is particularly applicable to addressing BNP disease, which primarily stems from the shedding of the *E. ictaluri* pathogen by members of the striped catfish population. *E. ictaluri* is mostly spread from dead infected catfish to healthy individuals in the same population (horizontal transmission). So far, vertical transmission (from parent to offspring or from *H. pangasius* to other species) has not been reported, but the presence of the bacteria in the gonads suggests it could be possible [49].

To demonstrate potential benefits of this approach, we used genetic parameters estimated in the present study as the primary inputs within a selection index framework to model the genetic gain achieved through a selective breeding program targeting three epidemiological traits, using the SIR model. The results demonstrated that the multi-trait selection index prolonged the times to susceptibility, infectivity, and mortality from 0.488 to 0.903 days or 5.5 to 10.4% per generation. Our approach can be expanded to include new traits, such as performance production and fitness-related traits when their genetic correlations with susceptibility, infectivity and mortality are available. Broadening the breeding objectives of the genetic program in striped catfish by incorporating production traits into the selection index would maximise genetic gains and economic return for aquaculture enterprises. The multi-trait selection index approach not only improve growth performance and disease resistance but also effectively curtailed undesired increase in other traits, e.g., feed intake. Despite these promising results, genetic selection aimed at enhancing population resilience to *E. ictaluri* should consider other factors, such as disease dynamics. Understanding disease dynamics and identification of key epidemiological factors causing BNP disease by *E. ictaluri* [50] are essential for effective genetic selection strategies. While incorporating disease prevalence, transmission routes, and pathogen evolution into the models can refine the selection process and optimise resilience outcomes, these types of information or data are currently not available in our studied species, which merit further studies. Additionally, disease surveillance and continuous monitoring of population resilience and disease prevalence is crucial for assessing the effectiveness of genetic selection programs. Regular evaluations of genetic improvement programs allow for adjustments and refinements to breeding objectives and selection strategies as new information becomes available [51].

Advantages and limitations

As demonstrated in our study where only the time to death was available, the SIR model enabled the estimation of genetic parameters for three different host traits (susceptibility, infectivity, and mortality). It also offers flexibility to analyse a range of data types, namely disease status, diagnostic test results, time censoring start or end of epidemics [18, 32]. More importantly, this approach assists the improvement in the overall resilience of the population rather than in individual animals. However, also note that the genetic-epidemiological SIR model employed in our study is built upon a set of assumptions, such as homogeneous mixing, equal chances of transmission, and the pivotal role of host transmission. However, it is important to acknowledge that numerous infectious diseases involve the transmission of pathogens through vectors other than direct individual-to-individual contact. For instance, wild hosts like predatory or migratory animals can serve as carriers. Additionally, some diseases exhibit low rates of host transmission, while certain pathogens remain dormant in the environment. The transmission and severity of infectious diseases are also influenced by environmental factors and management practices. Furthermore, certain diseases are caused by multiple pathogens, making it challenging to quantify individual (or population) resistance/resilience in these situations [52]. Although our current experimental design, including family structure and contact groups, enabled the detection of genetic variances and correlations for epidemiological traits, future studies should incorporate comprehensive data spanning multiple generations and in-depth pedigrees. In our study, we only observed the time of death, as it was practically unfeasible to record the times of susceptibility and infection due to the lack of observable clinical signs during group testing in tanks. To address this limitation, challenge test experiments should be combined with innovative phenotyping tools, such as water cameras, in combination with laboratory diagnostic tests to capture individual-level susceptibility and infection information. Throughout our experiments, we employed PCR to frequently confirm the demise of fish, ensuring that they were indeed deceased due to *E. ictaluri*. However, secondary infections caused by other pathogens (e.g., *Aeromonas hydrophila*) could not be entirely eliminated. Minor fluctuations in ambient temperature within the experimental tanks were beyond our control. Also, a combination of cohabitation and direct addition of the pathogen to tanks was not accounted for in the SIR model, which might have impacted genetic parameter estimates in our study.

Furthermore, epidemiological host traits may respond differently to developmental stages in striped catfish. We didn't have resources to obtain the pathogen load data of individual fish at different stages to identify exactly disease status (susceptibility and infectivity) of the population. Collectively, these factors may have influenced the statistical power of our study in distinguishing the various genetic components associated with epidemiological host traits. Addressing these factors can effectively enhance the resilience of populations in genetic selection program, ultimately improving production performance and minimizing the impact of infectious diseases. Therefore, future studies should take these factors into consideration to enhance the reliability of genetic parameters for epidemiological traits in striped catfish and other aquaculture species.

Concluding remarks and suggestions

Our study highlights the significant presence of heritable genetic variations in epidemiological host traits, including susceptibility, infectivity, and mortality. The heritability estimates for these traits are generally moderate. Genetic correlations between susceptibility and infectivity, as well as between susceptibility and mortality were negative, while a positive correlation was observed between infectivity and mortality. Genomic prediction for transition time from susceptibility to infectivity and from infectivity to mortality achieved moderate level of accuracy, suggesting possibilities for performing genome-based selection programs to improve epidemiological host traits in striped catfish. However, these results should be compared with those obtained using actual data on disease status (susceptibility and infectivity) when these traits are reliably determined in this population. Using a selection index approach, we made predictions regarding genetic gains, which ranged from 5.5 to 10.3% per generation. While these findings are promising, we strongly recommend further analysis with an increased accumulation of data especially from progeny testing schemes to obtain more reliable estimates of the genetic parameters and genomic prediction accuracies for epidemiological host traits. In addition to genetic selection aimed at reducing susceptibility, infectivity and mortality, indirect selection for enhanced immune response can be practised through recording immunological traits such as IgM, IgG, and lymphocyte counts. Breeding animals with stronger immune systems, enhanced innate defence mechanisms, and reduced pathogen shedding can effectively mitigate infectivity rates. Selectively bred animals with superior immune responses and faster recovery times, based on the identification and targeting of immune-related traits or markers associated with expedited recovery rates also can enhance the overall resilience of the population. In addition, future studies should prioritise the identification of genes responsible for these traits to deepen our understanding of the complex genomic architecture underlying these characteristics. Such knowledge will significantly contribute to future genetic improvement endeavours, ultimately enhancing the resistance to infectious diseases in the population of striped catfish under investigation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s44365-025-00006-6.

Supplementary Material 1.

Acknowledgements
We would like to thank the team at National Breeding Centre for Southern Aquaculture (NABRECSOFA), Research Institute for Aquaculture (RIA2) for their hard work in handling the animals and data collection.

Authors' contributions
NHN: Conceptualisation, Resources, Methodology, Formal Analysis, Writing -original draft, reviewing and editing, supervision NTV: Conduct of the experiment, Data curation and analysis, draft preperation and approval of the submission. TTMH: Data management, analysis, editing. THP: Funding aqusition, supervision, experiments, data collection and analysis. All authors approved to submit the manuscript.

Funding
This study was jointly funded by Ministry of Agriculture and Rural Development (MARD), Vietnam (Project tittle: Breeding for disease resistance to Bacillary Necrosis of Pangasius for striped catfish, 2018 – 2020) and higher degree by research fundings from University of the Sunshine Coast, Australia.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Author details
[1]School of Science, Engineering and Technology, University of the Sunshine Coast, Locked Bag 4, Maroochydore DC, QLD 4558, Australia. [2]Center for Bioinnovation, University of the Sunshine Coast, Locked Bag 4, Maroochydore DC, QLD 4558, Australia. [3]Research Institute for Aquaculture No.2,, 121 Nguyen Dinh Chieu Street, District 1, Ho Chi Minh City, Vietnam. [4]Tropical Futures Institute, James Cook University, 149 Sims Drive, Singapore 387380, Singapore.

Received: 4 December 2024 Accepted: 29 January 2025
Published online: 08 February 2025

References

1. FAOSTAT. Statistical Database. Rome: Food and Agriculture Organization of the United Nations; 2020.
2. Vu N. Genetics of the bacterial (Edwardsiella ictaluri) disease resistance in striped catfish (Pangasianodon hypophthalmus). *PhD thesis*. University of the Sunshine Coast, Queensland; 2021.
3. Segner H, Sundh H, Buchmann K, Douxfils J, Sundell KS, Mathieu C, Ruane N, Jutfelt F, Toften H, Vaughan L. Health of farmed fish: its relation to fish welfare and its utility as welfare indicator. Fish Physiol Biochem. 2012;38(1):85–105.

4. Robinson NA, Robledo D, Sveen L, Daniels RR, Krasnov A, Coates A, Jin YH, Barrett LT, Lillehammer M, Kettunen AH, et al. Applying genetic technologies to combat infectious diseases in aquaculture. Rev Aquac. 2023;15(2):491–535.
5. Bishop SC, Axford RF, Nicholas FW, Owen JB. Breeding for disease resistance in farm animals: CABI; 2010.
6. Van Khang P, Van Nha V, Nguyen NH. Resistance to Streptococcus iniae and its genetic associations with traits of economic importance in Asian seabass (Lates calcarifer). J Fish Dis. 2019;42(12):1657–66.
7. Vu NT, Sang NV, Trong TQ, Duy NH, Dang NT, Nguyen NH. Breeding for improved resistance to Edwardsiella ictaluri in striped catfish (Pangasianodon hypophthalmus): Quantitative genetic parameters. J Fish Dis. 2019;42(10):1409–17.
8. Trang TT, Hung NH, Ninh NH, Knibb W, Nguyen NH. Genetic variation in disease resistance against White Spot Syndrome Virus (WSSV) in Liptopenaeus vannamei. Front Genet. 2019;10:264.
9. Nguyen NH, Vu NT. Threshold models using Gibbs sampling and machine learning genomic predictions for skin fluke disease recorded under field environment in yellowtail kingfish Seriola lalandi. Aquaculture. 2022;547: 737513.
10. Phuthaworn C, Nguyen NH, Quinn J, Knibb W. Moderate heritability of hepatopancreatic parvovirus titre suggests a new option for selection against viral diseases in banana shrimp (Fenneropenaeus merguiensis) and other aquaculture species. Genet Sel Evol. 2016;48(1):64.
11. Van Sang N, Dung TTP, Phuong VH, Nguyen NH, Thinh NH. Immune response of selective breeding striped catfish families (Pangasiandon hypophthalmus) to Edwardsiella ictaluri after challenge. Aquaculture. 2023;572: 739515.
12. Cavallino L, Rincón L, Scaia MF. Social behaviors as welfare indicators in teleost fish. 2023;10.
13. Gilmour A, Anderson RD, Rae AL. The analysis of binomial data by a generalized linear mixed model. Biometrika. 1985;72(3):593–9.
14. Taylor RS, Carvalheiro R, Patchett AL, Verbyla KL, Carson J, Wynne JW, Evans BS, Lind CE, King H. Genetic and genomic analyses of resistance to yersiniosis in Atlantic salmon (Salmo salar) assessed by tank challenge. Aquaculture. 2023;564: 739088.
15. Trang TT, Hung NH, Ninh NH, Nguyen NH. Selection for improved white spot syndrome virus resistance increased larval survival and growth rate of Pacific Whiteleg shrimp, Liptopenaeus vannamei. J Invertebr Pathol. 2019;166: 107219.
16. Dégremont L, Garcia C, Allen SK Jr. Genetic improvement for disease resistance in oysters: a review. J Invertebr Pathol. 2015;131:226–41.
17. Milgroom MG. Epidemiology and SIR Models. In: Biology of Infectious Disease: From Molecules to Ecosystems. Cham: Springer International Publishing. 2023 p. 253–268.
18. Pooley CM, Marion G, Bishop SC, Bailey RI, Doeschl-Wilson AB. Estimating individuals' genetic and non-genetic effects underlying infectious disease transmission from temporal epidemic data. PLoS Comput Biol. 2020;16(12): e1008447.
19. Bishop S, Stear M. Genetic and epidemiological relationships between productivity and disease resistance: gastro-intestinal parasite infection in growing lambs. J Animal Science. 1999;69(3):515–24.
20. Doeschl-Wilson AB, Davidson R, Conington J, Roughsedge T, Hutchings MR, Villanueva B. Implications of host genetic variation on the risk and prevalence of infectious diseases transmitted through the environment. Genetics. 2011;188(3):683–93.
21. Anacleto O, Cabaleiro S, Villanueva B, Saura M, Houston RD, Woolliams JA. Doeschl-Wilson ABJSr: Genetic differences in host infectivity affect disease spread and survival in epidemics. 2019;9(1):4924.
22. Tsairidou S, Anacleto O, Woolliams J, Doeschl-Wilson A. Enhancing genetic disease control by selecting for lower host infectivity and susceptibility. Heredity. 2019;122(6):742–58.
23. Prentice J, Pooley C, Tsairidou S, Wong RP, Anacleto O, Saura M, Bailey R, Marion G, Villanueva B, Doeschl-Wilson A: Transmission experiment in turbot shows high genetic variation in host infectivity affecting disease spread and survival. In: Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges: 2022: Wageningen Academic Publishers; 2022. p. 728–731.
24. Doeschl-Wilson A, Knap P, Opriessnig T, More SJ. Livestock disease resilience: from individual to herd level. Animal. 2021;15: 100286.
25. Chase-Topping ME, Pooley C, Moghadam HK, Hillestad B, Lillehammer M, Sveen L, Doeschl-Wilson A. Impact of vaccination and selective breeding on the transmission of infectious salmon anemia virus. Aquaculture. 2021;535: 736365.
26. Dorfman B, Marcos-Hadad E, Tadmor-Levi R, David L. Disease resistance and infectivity of virus susceptible and resistant common carp strains. Sci Rep. 2024;14(1):4677.
27. Vu NT, Van Sang N, Phuc TH, Vuong NT, Nguyen NH. Genetic evaluation of a 15-year selection program for high growth in striped catfish Pangasianodon hypophthalmus. Aquaculture. 2019;509:221–6.
28. Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C. Diversity arrays technology: a generic genome profiling technology on open platforms. In: Data production and analysis in population genomics. Springer; 2012: 67–89.
29. Vu NT, Ha TTT, Thuy VTB, Trang VT, Nguyen NH. Population Genomic Analyses of Wild and Farmed Striped Catfish Pangasianodon Hypophthalmus in the Lower Mekong River. Journal of Marine Science and Engineering. 2020;8(6):471.
30. Gruber B, Unmack PJ, Berry OF, Georges A. dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol Ecol Resour. 2018;18(3):691–9.
31. Doeschl-Wilson A, Marion G, Pooley C. New tools and insights to enable breeding for reduced disease transmission. In: Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges: 2022: Wageningen Academic Publishers; 2022. p. 724–727.
32. Pooley C, Marion G, Bishop S, Doeschl-Wilson A. Optimal experimental designs for estimating genetic and non-genetic effects underlying infectious disease transmission. Genet Sel Evol. 2022;54(1):1–22.
33. Gilmour AR, Gogel BJ, Cullis BR, Welham S, Thompson R. ASReml user guide release 4.0. 2002.
34. Misztal I, Tsuruta S, Strabel T, Auvray B, Druet T, Lee D. BLUPF90 and related programs (BGF90). In: Proceedings of the 7th world congress on genetics applied to livestock production: 2002: Montpellier. 2002. p. 743.
35. Cameron ND. Selection Indices and Prediction of Genetic Merit in Animal Breeding: CAB International. 1997.
36. Ponzoni RW, Nguyen NH, Khaw HL, Ninh NH. Accounting for genotype by environment interaction in economic appraisal of genetic improvement programs in common carp Cyprinus carpio. Aquaculture. 2008;285(1):47–55.
37. Rutten MJM, Bijma P, Woolliams JA, van Arendonk JAM. SelAction: Software to Predict Selection Response and Rate of Inbreeding in Livestock Breeding Programs. J Hered. 2002;93(6):456–8.
38. Kulkarni PS, Biemans F, de Jong MCM, Bijma P. On the origin of the genetic variation in infectious disease prevalence: Genetic analysis of disease status versus infections for digital dermatitis in Dutch dairy cattle. J Anim Breed Genet. 2021;138(6):629–42. https://doi.org/10.1111/jbg.12635.
39. Barden M, Anagnostopoulos A, Griffiths BE, Li B, Bedford C, Watson C, et al. Genetic parameters of sole lesion recovery in Holstein cows. J Dairy Sci. 2023;106(3):1874–88. https://doi.org/10.3168/jds.2022-22064.
40. Hulst AD, de Jong MC, Bijma P. Why genetic selection to reduce the prevalence of infectious diseases is way more promising than currently believed. Genetics. 2021;217(4):iyab024.
41. Bijma P, Hulst AD, de Jong MC. The quantitative genetics of the prevalence of infectious diseases: hidden genetic variation due to Indirect Genetic Effects dominates heritable variation and response to selection. Genetics. 2022;220(1):141.
42. Vallejo RL, Fragomeni BO, Cheng H, Gao G, Long RL, Shewbridge KL, MacMillan JR, Towner R, Palti Y. Assessing accuracy of genomic predictions for resistance to infectious hematopoietic necrosis virus with progeny testing of selection candidates in a commercial rainbow trout breeding population. Front Vet Sci. 7:590048. https://doi.org/10.3389/fvets.2020.590048.
43. Allal F, Nguyen NH. Genomic Selection in AquacultureAquaculture Species. In: Genomic Prediction of Complex Traits: Methods and Protocols. Edited by Ahmadi N, Bartholomé J. New York, NY: Springer US; 2022: 469–491.

44. Vallejo RL, Pietrak MR, Milligan MM, Gao G, Tsuruta S, Fragomeni BO, Long RL, Peterson BC, Palti Y. Genetic architecture and accuracy of predicted genomic breeding values for sea lice resistance in the St John River aquaculture strain of North American Atlantic salmon. Aquaculture. 2024;586:740819.
45. Nguyen NH. Genetics and Genomics of Infectious Diseases in Key Aquaculture Species. Biology. 2024;13(1):29.
46. Kim OTP, Nguyen PT, Shoguchi E, Hisata K, Vo TTB, Inoue J, Shinzato C, Le BTN, Nishitsuji K, Kanda M, et al. A draft genome of the striped catfish, Pangasianodon hypophthalmus, for comparative analysis of genes relevant to development and a resource for aquaculture improvement. BMC Genomics. 2018;19(1):733.
47. Berghof TV, Poppe M, Mulder HA. Opportunities to improve resilience in animal breeding programs. Front Genet. 2019;9:692.
48. Knap PW, Doeschl-Wilson A. Why breed disease-resilient livestock, and how? Genet Sel Evol. 2020;52:1–18.
49. Machimbirike VI, Crumlish M, Dong HT, Santander J, Khunrae P, Rattanarojpong T. Edwardsiella ictaluri: A systemic review and future perspectives on disease management. Rev Aquac. 2022;14(3):1613–36.
50. Kirkeby C, Brookes VJ, Ward MP, Dürr S, Halasa T. A practical introduction to mechanistic modeling of disease transmission in veterinary science. Front Vet Sci. 7:546651. https://doi.org/10.3389/fvets.2020.546651.
51. Nguyen NH. Genetic improvement for important farmed aquaculture species with a reference to carp, tilapia and prawns in Asia: achievements, lessons and challenges. Fish Fish. 2016;17(2):483–506.
52. Putz AM, Harding JC, Dyck MK, Fortin F, Plastow GS, Dekkers JC, Canada P. Novel resilience phenotypes using feed intake data from a natural disease challenge model in wean-to-finish pigs. Front Genet. 2019;9:660.

Publisher's Note