**Accuracies of genomic predictions for disease resistance of striped catfish to *Edwardsiella ictaluri* using artificial intelligence algorithms**

Nguyen Thanh Vua,b,c,\*, Tran Huu Phucc, Kim Thi Phuong Oanhd, Nguyen Van Sangc, Trinh Thi Tranga,b,e, Nguyen Hong Nguyena,b,\*

aSchool of Science, Technology and Engineering, University of the Sunshine Coast, 90 Sippy Downs drive, Sippy Downs 4556, Queensland, Australia; bGenecology Research Center, University of the Sunshine Coast, 90 Sippy Downs drive, Sippy Downs 4556, Queensland, Australia; cResearch Institute for Aquaculture No.2, 116 Nguyen Dinh Chieu, District 1, Ho Chi Minh city 710000, Vietnam; dInstitute of Genome Research, Vietnam Academy of Science and Technology, Cau Giay 122000, Hanoi, Vietnam; eVietnam National University of Agriculture, Gia Lam 131000, Vietnam.

\*Correspondence to N.T. Vu vunt.ria2@mard.gov.vn or ThanhVu.Nguyen@research.usc.edu.au

and N.H. Nguyen nnguyen@usc.edu.au

**Running title: genomic prediction accuracy of disease resistance in striped catfish**

*Keywords: striped catfish Pangasianodon hypophthalmus, Edwardsiella ictaluri, BNP disease, genomic prediction, BayesR, Machine learning and deep learning.*

**Supplementary Figure S1**. Mean family survival rate in percentage post challenge test. First 20 family is low resistance group and last 20 families is high resistance group.

Supplementary Figure S2. Mean family survival time in day post challenge test. First 20 family is low resistance groups, and last 20 families is high resistance group.

Supplementary Table S1. Number of SNPs used for each testing set based on its effect with p<0.00001

|  |  |  |  |
| --- | --- | --- | --- |
| Subset | Total SNP used for ssGWAS | ssGWAS SNPs for survival status | ssGWAS SNPs for survival time |
| i1t1 | 6,470 | 342 | 1,448 |
| i1t2 | 6,470 | 381 | 1,537 |
| i1t3 | 6,470 | 356 | 1,399 |
| i1t4 | 6,470 | 351 | 1,561 |
| i1t5 | 6,470 | 318 | 1,499 |
| i2t1 | 6,470 | 335 | 1,450 |
| i2t2 | 6,470 | 366 | 1,598 |
| i2t3 | 6,470 | 400 | 1,540 |
| i2t4 | 6,470 | 318 | 1,454 |
| i2t5 | 6,470 | 370 | 1,401 |
| i3t1 | 6,470 | 341 | 1,401 |
| i3t2 | 6,470 | 365 | 1,475 |
| i3t3 | 6,470 | 373 | 1,514 |
| i3t4 | 6,470 | 369 | 1,559 |
| i3t5 | 6,470 | 361 | 1,538 |
| i4t1 | 6,470 | 372 | 1,576 |
| i4t2 | 6,470 | 326 | 1,362 |
| i4t3 | 6,470 | 357 | 1,539 |
| i4t4 | 6,470 | 349 | 1,383 |
| i4t5 | 6,470 | 327 | 1,548 |
| i5t1 | 6,470 | 329 | 1,423 |
| i5t2 | 6,470 | 367 | 1,480 |
| i5t3 | 6,470 | 341 | 1,496 |
| i5t4 | 6,470 | 362 | 1,555 |
| i5t5 | 6,470 | 360 | 1,502 |
| Mean |  | 353.4 | 1,489.5 |
| Min |  | 318 | 1,362 |
| Max |  | 400 | 1,589 |

Supplementary Table S2. Trait description by family in the main challenge test using 5.328 individual fish.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Trait | Unit | n | Mean | Min | Max |
| Survival status | % | 5,328 | 43.8 | 0 | 100 |
| Survival time | day | 5,328 | 7.2 | 2 | 23 |