

TIỀM NĂNG CHỌN GIỐNG HỆ GEN MỘT SỐ TÍNH TRẠNG SINH TRƯỞNG TRÊN CÁ CHIM VÂY VÀNG VÂY NGẮN (*Trachinotus anak*)

Vũ Văn Sáng^{1,*}, Lưu Thị Hà Giang², Hà Khánh Linh¹, Vũ Văn In³,
Phạm Hồng Nhật², Phạm Thế Hải¹, Lê Tất Thành⁴

¹ Khoa Sinh học, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội

² Viện Khoa học Thủy sản Việt Nam

³ Khoa Công nghệ và Kỹ thuật tiên tiến, Trường Đại học Việt Nhật, Đại học Quốc gia Hà Nội

⁴ Ban Kế hoạch - Tài chính, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

* Email: vuvansangts50@gmail.com

TÓM TẮT

Cá chim vây vàng vây ngắn (*T. anak*) là đối tượng nuôi hải sản mới nổi ở châu Á, tuy nhiên các nguồn dữ liệu hệ gen phục vụ cho công tác chọn giống vẫn còn hạn chế. Nghiên cứu này đánh giá tiềm năng của phương pháp chọn giống hệ gen đối với các tính trạng liên quan đến tăng trưởng ở cá chim vây vàng vây ngắn bằng cách sử dụng các chỉ thị đa hình nucleotide đơn (SNP). Nghiên cứu được thực hiện trên 300 cá chim vây vàng vây ngắn bằng việc thu thập các số liệu kiểu hình gồm chiều dài tổng thể, chiều dài thân và chiều dài đầu. Mẫu vây cá được thu thập cho từng cá thể, liên kết với dữ liệu kiểu hình, sau đó tách chiết ADN và giải trình tự gen. 11.356 chỉ thị SNP được sử dụng để ước tính các thông số di truyền bằng phương pháp dự đoán tuyến tính hệ gen (GBLUP). Kết quả phân tích kiểu hình cho thấy, giá trị biến thiên giữa các cá thể từ 7,93 - 11,28%. Hệ số di truyền hệ gen cho các tính trạng nghiên cứu có giá trị tương đối cao (0,60 - 0,70). Tương quan di truyền tích cực và cao giữa các tính trạng sinh trưởng (0,85 - 0,92), cho thấy các đặc điểm hình thái này có mối liên hệ di truyền chặt chẽ và có thể được kiểm soát bởi các vùng hệ gen chung. Kết quả này chỉ ra rằng, việc chọn lọc một tính trạng tăng trưởng có thể dẫn đến sự cải thiện đồng thời ở các tính trạng liên quan khác. Nhìn chung, nghiên cứu này cho thấy, thông tin hệ gen thu được từ các chỉ thị SNPs cung cấp những hiểu biết quan trọng về cấu trúc di truyền của các tính trạng tăng trưởng ở cá chim vây vàng vây ngắn, đồng thời khẳng định tiềm năng ứng dụng chọn giống hệ gen nhằm thúc đẩy cải thiện di truyền trong các chương trình chọn giống cá chim vây vàng vây ngắn.

Từ khóa: Chỉ thị phân tử SNP, tính trạng tăng trưởng, cá chim vây vàng vây ngắn.

1. ĐẶT VẤN ĐỀ

Cá chim vây vàng vây ngắn (*Trachinotus anak*) là một loài cá biển có giá trị kinh tế quan trọng ở châu Á, đặc biệt tại Việt Nam, Philippines, Trung Quốc và Indonesia, nơi hoạt động nuôi trồng thủy sản đã phát triển nhanh chóng trong những năm gần đây [1]. Loài này được xem là một trong những đối tượng tiềm năng cho nuôi thương phẩm ở khu vực châu Á. Tuy nhiên, hiện nay cá chim vây vàng vây ngắn đang được nuôi tại nước ta thường được gọi bằng ba tên khoa học khác nhau (*T. ovatus*, *T. anak* và *T. falcatus*). Do đó, đã thực hiện nghiên cứu xác định tên khoa học của loài cá chim vây vàng vây ngắn đã được thực hiện bằng gene ty thể COI và khẳng định được loài được nuôi tại Việt Nam có tên khoa học là *T. anak* [2]. Hiện nay, các phương

thức nuôi chủ yếu đối với *T. anak* bao gồm nuôi lồng và nuôi ao. Sản lượng của loài này đã tăng nhanh nhờ việc áp dụng rộng rãi mô hình nuôi lồng biển quy mô lớn. Tuy nhiên, sự phát triển nhanh chóng của nghề nuôi *T. anak* trong khi thiếu các biện pháp quản lý hiệu quả đã dẫn đến một số thách thức, bao gồm thiếu nguồn cá giống chất lượng cao, bùng phát dịch bệnh và suy giảm tốc độ tăng trưởng. Do đó, việc xem xét các yếu tố di truyền trong quản lý đàn cá bố mẹ, cùng với việc áp dụng các phương pháp lựa chọn cá bố mẹ chất lượng tốt ở giai đoạn sớm, là cần thiết nhằm đảm bảo sự phát triển nhanh và bền vững của nghề nuôi. Các tính trạng liên quan đến tăng trưởng được xem là những yếu tố quan trọng ảnh hưởng trực tiếp đến hiệu quả kinh tế trong nuôi trồng thủy sản. Trong đó, các đặc điểm hình thái

như chiều dài tổng thể, chiều dài thân và chiều dài đầu được coi là những chỉ tiêu quan trọng trong các chương trình chọn giống nhằm cải thiện khả năng tăng trưởng của loài.

Chọn giống hệ gen là một chiến lược hiệu quả để dự đoán biểu hiện kiểu hình của các cá thể dựa trên các chỉ thị phân tử mật độ cao trên toàn bộ hệ gen [3]. So với phương pháp chọn giống truyền thống dựa trên phá hệ, cách tiếp cận này cho phép khai thác hiệu quả hơn biến dị di truyền trong cùng một gia đình, đồng thời đã được chứng minh là giúp nâng cao độ chính xác dự đoán và thúc đẩy tiến bộ di truyền trong các chương trình chọn giống [4]. Trong các chương trình chọn giống cá, chọn giống hệ gen ngày càng được áp dụng rộng rãi, đặc biệt trong cải thiện các tính trạng tăng trưởng và khả năng kháng bệnh. Tuy nhiên, mặc dù đã được triển khai rộng rãi trong ngành chăn nuôi, việc ứng dụng phương pháp này trong nuôi trồng thủy sản vẫn còn tương đối hạn chế. Một trong những nguyên nhân chính là chi phí cao liên quan đến giải trình tự toàn bộ hệ gen và các SNP-chip thương mại, đặc biệt đối với các loài chưa có hệ gen đầy đủ. Phương pháp giải trình tự đoạn gen (Genotyping-by-Sequencing - GBS) đã nổi lên như một giải pháp thay thế tiết kiệm chi phí hơn so với các SNP-chip. Trong số các phương pháp giải trình tự gen, công nghệ DArTseq™ kết hợp kỹ thuật giảm độ phức tạp của hệ gen truyền thống của DArT với công nghệ giải trình tự thế hệ mới [5]. Phương pháp này có thể được điều chỉnh phù hợp với từng loài thông qua việc lựa chọn enzyme cắt hạn chế và các vùng hệ gen thích hợp, cho phép tạo ra số lượng lớn các chỉ thị trên toàn bộ hệ gen với chi phí tương đối thấp. Tùy thuộc vào mật độ chỉ thị, chi phí giải trình tự gen cho mỗi mẫu có thể chỉ khoảng 20 - 35 USD. Do đó, việc đánh giá sự cân bằng giữa chi phí giải trình tự gen và độ chính xác dự đoán hệ gen là rất cần thiết nhằm thúc đẩy ứng dụng thực tiễn của chọn giống hệ gen trong các loài nuôi trồng thủy sản như cá chim vây vàng vây ngắn. Ngoài ra, tiềm năng của công nghệ DArTseq™ trong việc ước tính độ chính xác dự đoán hệ gen ở cá chim vây vàng vây ngắn hiện vẫn chưa được nghiên cứu. Đặc biệt, đây là nghiên cứu báo cáo hệ số di truyền hệ gen, chính xác dự đoán hệ gen của đối tượng cá chim vây vàng vây ngắn (*T. anak*).

Mục tiêu chính của nghiên cứu này là đánh giá tiềm năng của chọn giống hệ gen trong việc cải thiện các tính trạng kinh tế quan trọng gồm chiều dài tổng thể, chiều dài đầu và chiều dài thân ở cá chim vây vàng vây ngắn. Các chỉ thị SNP trên toàn bộ hệ gen được tạo ra bằng công nghệ DArTseq™ từ một quần thể cá chim vây vàng vây ngắn thương mại và các chỉ thị này được sử dụng để ước tính các thông số di truyền, tương quan di truyền hệ gen và mức độ chính xác của dự đoán hệ gen.

2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thiết kế thí nghiệm

Cá chim vây vàng vây ngắn (*T. anak*) được sử dụng trong nghiên cứu này được thu từ một trại nuôi lồng biển tại tỉnh Quảng Ninh, Việt Nam. Cá ban đầu được ương trong các bể xi măng tại trại giống cho đến khi đạt giai đoạn cá giống (khoảng 60 ngày tuổi), sau đó được chuyển sang nuôi trong các lồng biển ngoài khơi cho giai đoạn nuôi thương phẩm. Tất cả các lồng nuôi được đặt tại vùng nước ven biển có độ sâu từ 7 - 8 m và có kích thước lồng giống nhau (3 m × 3 m × 4 m). Điều kiện nuôi được tiêu chuẩn hóa giữa các lồng, bao gồm: Loại thức ăn, chế độ cho ăn, mật độ thả nuôi (500 cá thể/lồng) và các thông số môi trường (độ mặn 28,0 - 31,0‰; nhiệt độ 22 - 30°C). Từ quần thể ban đầu khoảng 2.000 cá thể, 500 cá giống khỏe mạnh, có kích thước tương đối đồng đều và không có dấu hiệu bệnh lý được lựa chọn để nuôi lớn phục vụ thí nghiệm. Nhóm cá này được chuyển sang một lồng riêng nhằm đảm bảo điều kiện nuôi đồng nhất. Quá trình lựa chọn nhằm loại bỏ các cá thể bất thường và tạo ra một quần thể tương đối đồng nhất cho các phân tích tiếp theo. Cá được cho ăn thức ăn thương mại chứa 35% protein thô và được cung cấp sục khí liên tục trong suốt chu kỳ nuôi. Tất cả các cá thể được duy trì trong điều kiện này cho đến khi tiến hành thu mẫu ở thời điểm 300 ngày tuổi.

Mẫu vây cá được thu thập từ 300 cá thể cá chim vây vàng vây ngắn đại diện cho hai nhóm kích thước khác nhau về khối lượng thu hoạch trung bình của quần thể và được thu từ quần đàn gồm 500 cá thể cá chim vây vàng. Cụ thể, 150 cá thể kích thước lớn được chọn ngẫu nhiên từ nhóm 30% cá thể có kích thước lớn nhất của quần thể, trong khi 150 cá thể kích thước nhỏ được lấy mẫu từ nhóm 30% cá thể nhỏ nhất. Tất cả các cá thể

đều có cùng độ tuổi, có nguồn gốc từ một quần đàn, cùng điều kiện lồng nuôi và chế độ chăm sóc. Khi thu mẫu mô, các dữ liệu kiểu hình của từng cá thể được ghi nhận tương ứng, bao gồm: Chiều dài tổng thể, chiều dài đầu, chiều dài thân. Tuy nhiên, do hạn chế về nguồn lực nên nghiên cứu không thể giải trình tự toàn bộ 500 cá thể cá chim vây vàng vây ngắn, có thể dẫn tới sai số nhỏ nhưng đã tối ưu mô hình hóa chọn lọc để giảm thiểu sai số có thể tạo ra trong quá trình thu mẫu. Mẫu vây cá được bảo quản trong dung dịch ethanol 80% và lưu trữ ở nhiệt độ -20°C cho đến khi tiến hành tách chiết ADN.

2.2. Tình trạng nghiên cứu và thu thập số liệu kiểu hình

Chiều dài tổng thể được ghi nhận với độ chính xác đến 0,01 cm, đo từ chóp miệng đến cuối vây đuôi. Ngoài ra, chiều dài đầu được đo từ chóp miệng đến mép sau của nắp mang, trong khi chiều dài thân được xác định bằng hiệu số giữa chiều dài tổng thể và chiều dài đầu.

2.3. Tách chiết ADN, xây dựng thư viện ADN và giải trình tự gen

ADN hệ gen được tách chiết từ khoảng 30 mg mô vây cá bằng bộ kit phân giải ADN mô và máu (QIAGEN, Đức) theo hướng dẫn của nhà sản xuất. Việc phát hiện và giải trình tự gen SNP được thực hiện theo phương pháp mô tả bởi Kilian và cs (2012) [5]. Tóm tắt quy trình như sau: ADN hệ gen được cắt và gắn adaptor bằng hệ enzyme cắt giới hạn PstI-SphI. Khác với phương pháp ban đầu chỉ sử dụng một adaptor tương thích, quy trình này sử dụng hai adaptor đặc hiệu cho hai đầu cắt của enzyme PstI và SphI. Adaptor tương thích với PstI bao gồm: Trình tự gắn với flow-cell của Illumina, vị trí gắn mỗi giải trình tự và vùng barcode “staggered” có độ dài biến đổi [6], trong khi adaptor còn lại chứa vị trí gắn flow-cell cùng với đầu cắt tương thích với SphI. Các đoạn ADN thu được sau đó được khuếch đại bằng PCR với các điều kiện chu kỳ như sau: Biến tính ban đầu ở 94°C trong 1 phút; tiếp theo là 30 chu kỳ gồm 94°C trong 20 giây, 58°C trong 30 giây và 72°C trong 45 giây; và bước kéo dài cuối cùng ở 72°C trong 7 phút. Các thư viện ADN sau khi khuếch đại được xử lý trên hệ thống Illumina c-Bot để thực hiện khuếch đại cầu (bridge amplification) và sau đó được giải

trình tự với chiều dài 77 bp bằng nền tảng Illumina HiSeq 2500. Các trình tự đọc có barcode không chính xác được loại bỏ dựa trên các tiêu chí lọc nghiêm ngặt. Trung bình khoảng 2,5 triệu trình tự đọc chất lượng cao cho mỗi mẫu được giữ lại để phát hiện marker. Để giảm thiểu sự dư thừa trình tự, các đoạn đọc giống hệt nhau được gộp lại thành các tệp “fastqcoll” và tiếp tục được xử lý bằng các thuật toán của DArT PL nhằm hiệu chỉnh các base có chất lượng thấp. Bộ dữ liệu sau khi xử lý được phân tích bằng pipeline DArTsoft14 để xác định các marker SNP và SilicoDArT. Việc xác định SNP được thực hiện bằng cách phân cụm tất cả các tag trên các thư viện bằng thuật toán C++ của DArT với ngưỡng khoảng cách tối đa là ba. Chất lượng marker được đánh giá thông qua các bộ lọc cân bằng alen và kiểm định phân ly Mendel trên khoảng 1.000 phép lai kiểm soát, từ đó xác định điểm nhất quán và giữ lại các locus có độ tin cậy cao và sai số thấp. Độ tin cậy của các SNP được củng cố thêm nhờ độ sâu giải trình tự vượt quá 30 lần đọc cho mỗi marker.

2.4. Độ chính xác dự đoán hệ gen

Trong nghiên cứu này, ba phương pháp dự đoán hệ gen đã được sử dụng, bao gồm BayesA [7], BayesCπ [8] và GBLUP [9]. Trong phương pháp GBLUP, phương sai của ảnh hưởng mỗi SNP được giả định là bằng nhau trong phân bố hệ gen [7, 9]. Phương pháp BayesA giả định rằng ảnh hưởng của các marker tuân theo phân bố t, trong khi BayesCπ giả định rằng chỉ một tỷ lệ nhỏ các SNP có ảnh hưởng khác không [7]. Tổng cộng 11.356 marker SNP có tần số alen nhỏ (minor allele frequency) > 0,05 được sử dụng cho phân tích 300 cá thể cá chim vây vàng vây ngắn. Mô hình 1 được sử dụng cho đánh giá GBLUP được viết như sau:

$$y = Xb + Za + e \text{ (Mô hình 1)}$$

Trong đó: y là vector các giá trị kiểu hình; b là vector các hiệu ứng cố định (giới tính); a là vector các hiệu ứng di truyền ngẫu nhiên; e là vector sai số dư; và X và Z là các ma trận thiết kế tương ứng cho các hiệu ứng cố định và hiệu ứng ngẫu nhiên. Giả định phân bố của các hiệu ứng ngẫu nhiên là phân bố chuẩn đa

biến với trung bình bằng 0 và $\text{var} \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} G\sigma_a^2 & 0 \\ 0 & I\sigma_e^2 \end{bmatrix} \begin{bmatrix} G\sigma_a^2 & 0 \\ 0 & I\sigma_e^2 \end{bmatrix}$, trong đó $\sigma_a^2 \sigma_e^2$ là phương

sai di truyền cộng gộp và $\sigma_p^2 \sigma_p^2$ là the phương sai kiểu hình và được xác định là tổng của phương sai kiểu gen và phương sai môi trường ($\sigma_e^2 \sigma_e^2$); G là ma trận được xây dựng như sau:

$$G = \frac{AA'}{\sum_i \sum_j 2b_i(1-b_i)} \quad G = \frac{AA'}{\sum_i \sum_j 2b_i(1-b_i)}$$

trong đó A là ma trận marker đã được chuẩn hóa (centered marker matrix), tổng ở mẫu số được tính trên tất cả các locus *iii*, và $b_i b_i$ là tần số alen tại locus *iii*. Mô hình 2 được áp dụng cho các phương pháp Bayesian. Mô hình thông kê được biểu diễn như sau:

$$y = Xb + \sum A g_j + e \text{ (Mô hình 2)}$$

Trong đó: y là vector kiểu hình; Xb là thành phần yếu tố cố định, với b là vector các hiệu ứng cố định bao gồm thể hệ và giới tính; $\sum A_{ij} g_j$ là giá trị giống hệ gen ước tính (gEBV) của cá thể *i*, trong đó A là ma trận marker và g là hiệu ứng của các marker. Tất cả các hiệu ứng marker, phương sai kiểu hình, phương sai sai số và các giá trị gEBV được ước tính bằng các phương pháp BayesA và BayesCπ thông qua gói phần mềm BGLR trong R [10]. Chuỗi lấy mẫu MCMC Gibbs được chạy với 15.000 vòng lặp, trong đó 3.000 vòng lặp đầu tiên được loại bỏ làm giai đoạn burn-in. Đối với phương pháp GBLUP, tất cả các thành phần phương sai và giá trị gEBV được phân tích bằng phần mềm ASReml phiên bản 4.2 [11].

2.5. Phương pháp kiểm định chéo (Cross validation)

Phương pháp kiểm định chéo 5 lần (five-fold cross-validation) được sử dụng để đánh giá độ chính xác dự đoán (tập huấn luyện 80% và tập kiểm định 20%). Kiểu hình của quần thể kiểm định được ẩn đi và được dự đoán từ quần thể huấn luyện. Độ chính xác dự đoán được tính bằng hệ số tương quan giữa giá trị giống ước tính (EBVs) của tập kiểm định và kiểu hình thực tế, sau đó chia cho căn bậc hai của hệ số di truyền ước tính của toàn bộ quần thể. Giá trị độ chính xác dự đoán trung bình thu được từ các tập khác nhau được tính toán và so sánh giữa phương pháp dựa trên phả hệ và phương pháp hệ gen.

2.6. Ước tính thông số di truyền hệ gen

Các thông số di truyền được ước tính bằng phương pháp dự đoán tuyến tính không sai lệch tốt nhất dựa trên hệ gen (GBLUP). Các mô hình tuyến tính đơn biến được sử dụng để ước tính các

thành phần phương sai và hệ số di truyền của các tính trạng. Tổng cộng 11.356 SNP đã được sử dụng để ước tính các thông số di truyền trong quần đàn cá chim vây vàng vây ngắn theo mô hình 1.

Tương quan di truyền giữa các tính trạng dựa trên thông tin marker SNP được ước tính bằng các mô hình hai biến (bivariate models). Các mô hình hai biến này là sự mở rộng của các mô hình đơn biến như được mô tả bởi Mrode (2014) [12]. Các hiệu ứng cố định và hiệu ứng ngẫu nhiên trong mô hình hai biến giống như trong các mô hình đơn biến. Tất cả các phân tích được thực hiện bằng các gói phần mềm trong chương trình R, bao gồm: Sommer, BGLR, AGHmatrix và MCMCglmm.

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Tính trạng kiểu hình

Thông kê mô tả đối với các tính trạng tăng trưởng được đo trong quần đàn lấy mẫu cá chim vây vàng vây ngắn được trình bày trong bảng 1. Sự biến thiên kiểu hình đáng kể đã được ghi nhận giữa các cá thể, cho thấy sự tồn tại của mức độ biến dị lớn có thể được khai thác trong các chương trình chọn giống. Giá trị trung bình của chiều dài tổng thể là 34,58 cm, với độ lệch chuẩn 3,19 cm, dao động từ 21,00 - 40,90 cm. Chiều dài thân có giá trị trung bình 21,35 cm (SD = 2,41 cm), trong khi chiều dài đầu có giá trị trung bình 6,79 cm với độ lệch chuẩn 0,54 cm. Hệ số biến thiên (CV) phản ánh mức độ phân tán tương đối của từng tính trạng. Trong số các tính trạng được nghiên cứu, chiều dài thân thể hiện mức biến thiên cao nhất (CV = 11,28%), tiếp theo là chiều dài tổng thể (9,21%), trong khi chiều dài đầu có mức biến thiên thấp nhất (7,93%). Những kết quả này cho thấy, chiều dài thân có thể phản ứng mạnh hơn với chọn lọc so với các tính trạng hình thái khác do có mức biến thiên kiểu hình cao hơn.

Sự biến thiên kiểu hình là một điều kiện tiên quyết quan trọng cho cải thiện di truyền trong các chương trình chọn giống thủy sản, vì nó phản ánh tác động tổng hợp của các yếu tố di truyền và môi trường lên sự biểu hiện của tính trạng. Mức độ biến thiên kiểu hình tương tự ở các tính trạng hình thái cũng đã được ghi nhận ở nhiều loài thủy sản khác, cho thấy các tính trạng liên quan đến tăng trưởng thường có mức biến thiên từ trung

bình đến cao trong các quần thể nuôi [3, 4, 13]. Sự biến thiên này tạo cơ hội thuận lợi cho việc áp dụng các chương trình chọn giống nhằm cải thiện các tính trạng có ý nghĩa kinh tế. Nhìn chung, mức độ biến thiên kiểu hình quan sát được trong

quần thể cá chim vây vàng vây ngắn cho thấy quần thể này có đủ mức độ đa dạng để hỗ trợ các phân tích hệ gen và các chiến lược chọn giống nhằm cải thiện khả năng tăng trưởng.

Bảng 1. Thống kê mô tả của số liệu kiểu hình đối với các tính trạng trong quần thể cá chim vây vàng vây ngắn

Tính trạng	Đơn vị	Số mẫu	Trung bình	Độ lệch chuẩn	Giá trị nhỏ nhất	Giá trị lớn nhất	Hệ số biến thiên (%)
Chiều dài tổng thể	cm	300	34,58	3,19	21,00	40,90	9,21
Chiều dài toàn thân	cm	300	21,35	2,41	7,10	31,90	11,28
Chiều dài đầu	cm	300	6,79	0,54	5,20	8,10	7,93

3.2. Độ tin cậy của các phương pháp chọn giống hệ gen khác nhau

Bảng 2 trình bày độ chính xác dự đoán của ba tính trạng thu được từ ba phương pháp dự đoán hệ gen. Trong số ba phương pháp này, GBLUP cho độ chính xác dự đoán cao hơn đối với cả ba tính trạng; tuy nhiên, sự khác biệt về khả năng dự đoán giữa các phương pháp là không lớn, phù hợp với các nghiên cứu trước đây [14 - 18]. Do đó, phương pháp GBLUP, vốn ít yêu cầu tính toán hơn so với các phương pháp khác, đã được

sử dụng để ước tính các thông số di truyền như hệ số di truyền và tương quan di truyền đối với các tính trạng chiều dài tổng thể, chiều dài thân và chiều dài đầu. Việc xác định mối quan hệ di truyền giữa các cá thể trong phả hệ được xem là quan trọng hơn so với việc xác định sự mất cân bằng liên kết (linkage disequilibrium). Phương pháp này cũng thuận tiện hơn cho việc triển khai trong sản xuất vì được tính toán trong một bước duy nhất, cho phép sử dụng đồng thời cả thông tin hệ gen và thông tin phả hệ [15, 16].

Bảng 2. Chính xác dự đoán (trung bình $\pm$ sai số) của các phương pháp dự đoán Bayes A, BayesC π và GBLUP đối với các tính trạng chiều dài tổng thể, chiều dài toàn thân và chiều dài đầu của cá chim vây vàng vây ngắn (đơn vị tính: cm)

Phương pháp	Tính trạng		
	Chiều dài tổng thể (cm)	Chiều dài toàn thân (cm)	Chiều dài đầu (cm)
Bayes A	0,723 $\pm$ 0,011	0,753 $\pm$ 0,009	0,753 $\pm$ 0,015
BayesC π	0,735 $\pm$ 0,012	0,764 $\pm$ 0,010	0,746 $\pm$ 0,013
GBLUP	0,802 $\pm$ 0,012	0,812 $\pm$ 0,012	0,824 $\pm$ 0,014

3.3. Hệ số di truyền và tương quan di truyền hệ gen

Các ước tính hệ số di truyền hệ gen và tương quan giữa các tính trạng tăng trưởng được trình bày trong bảng 3. Hệ số di truyền của ba tính trạng ở mức trung bình đến cao, dao động từ 0,60 - 0,70, cho thấy một tỷ lệ đáng kể của sự biến thiên kiểu hình ở các tính trạng này có thể được quy định từ yếu tố di truyền cộng gộp. Cụ thể, hệ số di truyền hệ gen được ước tính là 0,70 đối với chiều dài tổng thể, 0,65 đối với chiều dài thân và 0,60 đối với chiều dài đầu. Các giá trị hệ số di truyền tương đối cao quan sát được trong nghiên cứu này cho thấy các tính trạng tăng trưởng ở cá chim vây vàng vây ngắn chịu sự kiểm soát di truyền mạnh và do đó có tiềm năng lớn đáp ứng

với chọn lọc trong các chương trình chọn giống. Hệ số di truyền của các tính trạng trong nghiên cứu này có giá trị cao hơn đáng kể so với các nghiên cứu đánh giá tiềm năng chọn giống hệ gen kháng bệnh *Edwardsiella ictaluri* trên cá tra (*Pangasianodon hypophthalmus*) [17]. Sự khác biệt về hệ số di truyền giữa tính trạng kháng bệnh và sinh trưởng trong nghiên cứu này có thể do đặc điểm loài, điều kiện ương nuôi và đặc biệt là tính trạng kháng bệnh chịu ảnh hưởng của nhiều gen cùng các tương tác gen - gen, gen - môi trường phức tạp hơn. Vì vậy, hệ số di truyền của tính trạng kháng bệnh thường thấp hơn tính trạng sinh trưởng. Giá trị hệ số di truyền cao đối với các tính trạng hình thái cũng đã được ghi nhận ở nhiều loài thủy sản khác, bao gồm động vật thân

mềm và cá, khi các phương pháp hệ gen được sử dụng để ước tính các thông số di truyền [14, 18, 19, 20]. Những kết quả này cho thấy, chọn giống hệ gen có thể nâng cao đáng kể hiệu quả của các chương trình chọn giống thông qua việc cải thiện độ chính xác dự đoán giá trị giống.

Tương quan di truyền giữa ba tính trạng nghiên cứu đều cao và có giá trị dương, dao động từ 0,85 - 0,92, cho thấy mối quan hệ di truyền chặt chẽ giữa các tính trạng hình thái của cá chim vây vàng vây ngắn. Cụ thể, chiều dài tổng thể có tương quan di truyền mạnh với cả chiều dài thân ($r_g = 0,85$) và chiều dài đầu ($r_g = 0,92$), trong khi chiều dài thân và chiều dài đầu cũng có tương quan di truyền cao ($r_g = 0,90$). Những mối tương quan di truyền mạnh này cho thấy các tính trạng này có thể do một vùng gen quy định. Tương tự, các tương quan kiểu hình giữa các tính trạng cũng ở mức cao, dao động từ 0,84 - 0,90, cho thấy giá trị tương quan kiểu hình đồng nhất với giá trị tương quan kiểu gen. Những kết quả này phù hợp với kết quả các nghiên cứu trước đây trên các loài thủy sản, trong đó các tính trạng liên quan đến tăng trưởng thường có tương quan dương mạnh do chịu ảnh hưởng của cùng một nhóm gen [21, 22]. Từ góc độ chọn giống, các tương quan dương mạnh được quan sát trong nghiên cứu này cho thấy, việc chọn lọc một tính trạng tăng trưởng, chẳng hạn như chiều dài tổng thể, có thể dẫn đến sự cải thiện đồng thời ở các tính trạng liên quan khác như chiều dài thân và chiều dài đầu. Do đó, các kết quả này hỗ trợ tính khả thi của việc áp dụng các chiến

lược chọn giống hệ gen nhằm cải thiện khả năng tăng trưởng trong các chương trình chọn giống cá chim vây vàng vây ngắn. Nghiên cứu chọn giống hệ gen là một trong những phương pháp nghiên cứu mới đang được sử dụng phổ biến rộng rãi ở các quốc gia như Mỹ, Úc, Châu Âu để chọn lọc các tính trạng khó đo đạc hoặc rất khó thu thập tính trạng kiểu hình như khối lượng thịt hoặc tính trạng kháng bệnh. Do các loài cá, đặc biệt là cá biển thường yêu cầu sinh sản theo các tổ hợp (gồm một số lượng con đực và con cái) nhất định, do đó việc áp dụng công nghệ chọn giống hệ gen càng giúp tiết kiệm nhân công lao động và nâng cao dự đoán chọn giống. Chọn giống hệ gen còn giúp cho người nuôi có thể chọn lọc được các cá thể làm bố mẹ ở giai đoạn sớm như tại thời điểm một tuổi, trong khi cá sẽ cần 3 - 4 tuổi mới thành thực, như vậy sẽ giúp tiết kiệm nhân công và chi phí thức ăn nuôi một số lượng lớn cá. Tuy nhiên, bên cạnh những ưu điểm vượt trội của phương pháp chọn giống hệ gen so với phương pháp chọn lọc truyền thống thì chi phí giải trình tự gen cũng là một rào cản. Hiện tại, giải trình tự gen bằng công nghệ DArT có chi phí tương đối phù hợp để đánh giá được hiệu quả và chính xác chọn lọc. Trong một chương trình chọn giống, khi nhiều cá thể được giải trình tự thì một nghiên cứu độc lập cần phải tiến hành để đánh giá hiệu quả chi phí của phương pháp giải trình tự, vì cần cân bằng giữa chọn lọc và hiệu quả kinh tế để tìm ra công nghệ giải trình tự gen tiết kiệm nhất mà vẫn đạt yêu cầu chọn lọc.

Bảng 3. Ước tính hệ số di truyền ($\pm$ SE) (in đậm trên đường chéo), tương quan di truyền (phía dưới đường chéo) và tương quan kiểu hình (phía trên đường chéo) đối với các tính trạng chiều dài tổng thể, chiều dài toàn thân, chiều dài đầu của cá chim vây vàng vây ngắn

Tính trạng	Chiều dài tổng thể (cm)	Chiều dài toàn thân (cm)	Chiều dài đầu (cm)
Chiều dài tổng thể	0,70 (0,05)	0,84 (0,03)	0,85 (0,04)
Chiều dài toàn thân	0,85 (0,02)	0,65 (0,02)	0,90 (0,02)
Chiều dài đầu	0,92 (0,03)	0,90 (0,05)	0,60 (0,01)

4. KẾT LUẬN

Hệ số di truyền hệ gen ở mức trung bình đến cao (0,60 - 0,70) được ghi nhận đối với chiều dài tổng thể, chiều dài toàn thân và chiều dài đầu, cho thấy các tính trạng có tiềm năng tốt cho các chương trình chọn giống. Tương quan di truyền dương và cao (0,85 - 0,92) giữa các tính trạng cho thấy việc chọn lọc một tính trạng tăng trưởng có thể dẫn đến sự cải thiện đồng thời ở các tính trạng hình thái liên quan khác. Nhìn chung, các kết quả

này chứng minh rằng, thông tin hệ gen có thể hỗ trợ cải thiện di truyền đối với khả năng tăng trưởng ở cá chim vây vàng vây ngắn, đồng thời tạo nền tảng cho các chương trình chọn giống hệ gen trong tương lai đối với loài này.

LỜI CẢM ƠN

Nhóm tác giả gửi lời cảm ơn chân thành tới các thành viên thuộc Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, Việt Nam, đã hỗ trợ trong quá trình thực hiện các thí

nghiệm. Đồng tác giả Hà Khánh Linh thực hiện khóa luận tốt nghiệp đại học chương trình đào tạo sinh học chuẩn tại Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội niên khóa 2022–2026.

Nghiên cứu này được tài trợ bởi Quỹ Phát triển khoa học và công nghệ Quốc gia (NAFOSTED) trong đề tài mã số 106.05 - 2023.64.

TÀI LIỆU THAM KHẢO

- [1]. Shadrin, A., Semenova, A. & Thanh, N. T. (2024). Are there Atlantic species of the genus *Trachinotus* (Carangidae), *T. falcatus* and *T. ovatus*, in Asian mariculture? *Journal of Ichthyology*, 64, 854 - 863.
- [2]. Vu, S. V., Putra, A., Aini, S., Laskar, B. A., Vu, I. V., Nguyen, T. T. A., Duong, C. N. & Kundu, S. (2026). Genetic identity and length-weight relationship of *Oyster Pompano* (*Trachinotus anak*) at a mariculture site in northern Vietnam. *Fish Aquatic Sciences*, 29(5), 346 - 360.
- [3]. Meuwissen, T., Hayes, B. & Goddard, M. (2013). Improvement of livestock with genomic selection. *Annual Review of Animal Biosciences*, 1, 221 - 237.
- [4]. Georges, M., Charlier, C. & Hayes, B. J. (2019). Harnessing genomic information for livestock improvement. *Nature Reviews Genetics*, 20(3), 135 - 156.
- [5]. Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H. & Hopper, C. (2012). Diversity arrays technology: A generic genome profiling technology on open platforms. In *Data Production and Analysis in Population Genomics* (pp. 67 - 89). Springer.
- [6]. Elshire, R., Glaubitz, J., Sun, Q., Poland, J., Kawamoto, K., Buckler, E. & Mitchell, S. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6, e19379.
- [7]. Hayes, B. J. & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819 - 1829.
- [8]. Habier, D., Fernando, R. L., Kizilkaya, K. & Garrick, D. J. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics*, 12, 186.
- [9]. VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91, 4414 - 4423.
- [10]. Pérez, P. & de los Campos, G. (2014). BGLR: A statistical package for whole genome regression and prediction. *Genetics*, 198, 483 - 495.
- [11]. Gilmour, A. R. & Thompson, R. (2019). ASReml update: What's new in release 4.2. VSN International Ltd., Hemel Hempstead, UK.
- [12]. Mrode, R. A. (2014). Linear models for the prediction of animal breeding values (3rd ed.). CABI.
- [13]. Zhu, X. F. (2026). Integrating GWAS and genomic selection with SNP and SV markers for enhanced prediction of complex traits in common carp. *Aquaculture*, 743622.
- [14]. Liu, G., Dong, L., Gu, L., Han, Z., Zhang, W., Fang, M. & Wang, Z. (2019). Evaluation of genomic selection for seven economic traits in Yellow drum (*Nibea albiflora*). *Marine Biotechnology*, 21, 806 - 812.
- [15]. Ma, P., Lund, M., Nielsen, U., Aamand, G. & Su, G. (2015). Single-step genomic model improved reliability and reduced the bias of genomic predictions in Danish Jersey. *Journal of Dairy Science*, 98, 9026 - 9034.
- [16]. Fernando, R. L., Cheng, H., & Garrick, D. (2016). An efficient exact method to obtain GBLUP and single-step GBLUP when the genomic relationship matrix is singular. *Genetics Selection Evolution*, 48(80), 12.
- [17]. Vu, N. T., Phuc, T. H., Oanh, K. T. P., Sang, N. V., Trang, T. T. & Nguyen, N. H. (2022). Accuracies of genomic predictions for disease resistance of striped catfish to *Edwardsiella ictaluri* using artificial intelligence algorithms. *G3-Genes, Genomes, Genetics* 12(1), jkab361, 13.
- [18]. Gutierrez, A., Matika, O., Bean, T., & Houston, R. (2018). Genomic selection for growth traits in Pacific oyster (*Crassostrea gigas*): Potential of low-density marker panels for breeding value prediction. *Frontiers in Genetics*, 9(391), 9.

- [19]. Vu, S. V., Knibb, W., Gondro, C., Subramanian, S., Nguyen, N. T., Alam, M., & O'Connor, W. (2021). Genomic prediction for whole weight, body shape, meat yield, and color traits in the Portuguese oyster *Crassostrea angulata*. *Frontiers in Genetics*, 12 (661276), 10.
- [20]. Xu, H. J. (2026). Estimation of genetic parameters for growth traits and genomic selection study in *Siniperca scherzeri*. *Marine Biotechnology*, 28(1): 8.
- [21]. Tao, Y. F., Huang, R. S., Hua, J. X., Wang, W., Lu, S. Q., Badran, M. F., ... Qiang, J. (2026). Deciphering salinity tolerance in red tilapia (*Oreochromis* spp.): Insights from a genome-wide association study and genomic selection. *Aquaculture*, 617, 743778.
- [22]. Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). Longman, Harlow, UK.

POTENTIALS FOR GENOMIC SELECTION OF GROWTH-RELATED TRAITS IN THE OYSTER POMPANO (*Trachinotus anak*)

**Vu Van Sang^{1,*}, Luu Thi Ha Giang², Ha Khanh Linh¹, Vu Van In³,
Pham Hong Nhat², Pham The Hai¹, Le Tat Thanh⁴**

¹ Faculty of Biology, University of Science, Vietnam National University, Hanoi.

² Vietnam Academy of Fishery Sciences (VAFIS)

³ Faculty of Advanced Technology and Engineering,
Vietnam Japan University, Vietnam National University, Hanoi.

⁴ Department of Planning and Finance,
Vietnam Academy of Science and Technology, Hanoi.

Abstract

The oyster pompano (*Trachinotus anak*) is an emerging aquaculture species in Asia, yet genomic resources to support selective breeding remains limited. This study evaluated the potential of genomic selection for growth-related traits in *T. anak* using single nucleotide polymorphism (SNP) markers. A total of 300 individuals from a cultured population in Northern Vietnam were phenotyped for total length, body length and head length. Fin-clip samples were collected for DNA extraction and genotyping using DArTSeq technology (genotyping by sequencing). After quality filtering, 11,356 SNP markers were used to estimate genetic parameters using a genomic best linear unbiased prediction (GBLUP) approach. Phenotypic analyses revealed moderate variation among individuals, with coefficients of variation ranging from 7.93% to 11.28%. Genomic heritability estimates were moderate to high, with values of 0.70 for total length, 0.65 for body length, and 0.60 for head length. Strong positive genetic correlations were observed among the traits (0.85–0.92), indicating that these morphometric traits are genetically related and may be controlled by common genomic regions. The results suggest that selection for one growth trait could lead to correlated improvements in other related traits. Overall, this study demonstrates that genomic information derived from DArTSeq markers provides valuable insights into the genetic architecture of growth traits in *T. anak* and supports the potential application of genomic selection to accelerate genetic improvement in oyster pompano breeding programs.

Keywords: SNP markers, growth-related traits, *Trachinotus anak*.

Ngày nhận bài: 4/3/2026

Ngày chuyển phản biện: 02/4/2026

Ngày thông qua phản biện: 20/4/2026

Ngày duyệt đăng: 7/5/2026