

# ĐÁNH GIÁ ĐA DẠNG DI TRUYỀN VÀ CẤU TRÚC QUẦN THỂ MỘT SỐ ĐÀN CÁ CHÉP Ở VIỆT NAM

Vũ Thị Trang<sup>1\*</sup>, Lưu Thị Hà Giang<sup>1</sup>, Phạm Hồng Nhật<sup>1</sup>,  
Lê Văn Khôi<sup>1</sup>, Bùi Thị Ánh Nguyệt<sup>2</sup>, Nguyễn Hữu Đức<sup>2</sup>, Đặng Thị Lua<sup>1,2</sup>

## TÓM TẮT

Nghiên cứu này đánh giá đa dạng di truyền, cấu trúc quần thể, phát sinh loài của 7 đàn cá chép bao gồm: cá chép trắng Việt Nam (VN), cá chép chọn giống V1 (V1), cá chép vảy Hungary (HV), cá chép vàng Indonesia (IN), cá chép vảy Hungary đàn Tata (TAT), cá chép vảy Hungary đàn Szarvas (SZA) và cá chép Cộng hòa Séc (SE) dựa trên phân tích trình tự gen COI và D-loop. Kết quả chỉ ra rằng, đàn VN và V1 là đàn có đa dạng di truyền cao nhất trong khi các đàn HV, TAT, SZA, SE và IN có đa dạng di truyền thấp khi xét trên cả 2 gen COI và D-loop. Nghiên cứu đã xác định được 53 haplotype, trình tự các haplotype đã được công bố trên GenBank - NCBI, với mã số MW450991 - MW451003 (gen COI) và MW463062 - MW463101 (gen D-loop). Về sai khác di truyền, đáng chú ý nhất là sự sai khác di truyền xét trên gen COI giữa đàn SE với tất cả các đàn cá chép còn lại, đặc biệt là giữa SE và HV, giữa SE và SZA (đều có  $F_{ST} = 1$ ), cho thấy không có sự chia sẻ về mặt di truyền giữa các đàn này. Kết quả phân tích AMOVA dựa trên gen COI cho thấy các đàn có cấu trúc quần thể rõ ràng do sự đóng góp đáng kể về khác biệt di truyền giữa các đàn. Cây phát sinh loài phân tử cho thấy mối quan hệ di truyền xa giữa SE và các đàn còn lại (xét trên gen COI) và mối quan hệ di truyền gần gũi của 7 đàn (xét trên gen D-loop). Nghiên cứu đã góp phần tạo cơ sở dữ liệu nguồn gen tham chiếu và phục vụ cho các nghiên cứu cần thông tin về nguồn gen của một số đàn cá chép ở Việt Nam hiện nay, đồng thời phục vụ công tác chọn giống và bảo tồn nguồn gen.

Từ khóa: *COI, D-loop, cá chép, đa dạng di truyền, haplotype.*

## 1. ĐẶT VẤN ĐỀ

Cá chép (*Cyprinus carpio*) thuộc bộ Cypriniformes và họ Cyprinidae, được coi là họ cá nước ngọt lớn nhất và phân bố rộng rãi ở hầu hết các quốc gia trên thế giới nhưng rất phổ biến ở châu Á và một số nước châu Âu (Weber và Brown, 2011). Cá chép được coi là đối tượng tiềm năng cho nuôi trồng thủy sản thương mại vì nó có khả năng thích ứng rất cao với cả môi trường nuôi và thức ăn (Rahman, 2015). Sản lượng cá chép chiếm 3,4% (4,4 triệu tấn năm 2015) trong tổng sản lượng thủy sản toàn cầu. Đây là loài cá chiếm tầm quan trọng ở vị trí thứ 3 về sản lượng nuôi trồng thủy sản trên thế giới và 97,3% sản lượng cá chép là do nuôi trồng thủy sản (Karnai và Szucs, 2018).

Hiện nay, ở nước ta, Viện Nghiên cứu Nuôi trồng Thủy sản I đang lưu giữ 6 đàn cá chép để phục vụ công tác bảo tồn nguồn gen như sau: cá chép vảy Hungary, cá chép trắng Việt Nam, cá chép chọn giống V1, cá chép vàng Indonesia, cá chép vảy

Hungary đàn Tata và cá chép vảy Hungary đàn Szarvas. Ngoài các đàn cá trên, thực tế trong các doanh nghiệp và trang trại tư nhân còn lưu giữ nguồn gen của một số đàn cá khác có giá trị di truyền như đàn cá chép nhập từ Cộng hòa Séc. Mặc dù là loài cá nuôi quan trọng đem lại giá trị kinh tế cao, tuy nhiên trong những năm gần đây, chương trình chọn giống trên cá chép chưa được thực hiện và hầu như chưa có công trình nghiên cứu nào liên quan đến thông tin di truyền của các đàn cá chép hiện có ở Việt Nam được công bố gần đây. Cùng với đó, nghiên cứu sử dụng gen DNA ty thể trong đánh giá đa dạng di truyền và cấu trúc di truyền trên các đàn cá chép khác nhau là chưa được thực hiện. Các đoạn gen như COI (Cytochrome c Oxidase subunit I) và D-loop (Displacement loop) thuộc hệ gen ty thể là các trình tự không những được sử dụng nhiều nhất cho các phân tích phát sinh loài trong các nghiên cứu tiến hóa (Kohlmann và Kersten, 2013), phân loại như là các mã vạch di truyền (Hubert và ctv., 2008) mà còn được sử dụng thường xuyên trong các nghiên cứu về đa dạng và cấu trúc di truyền trên nhiều loài động vật thủy sản (Ward và ctv., 2005). Do vậy, nghiên cứu này nhằm tạo ra một cơ sở dữ liệu nguồn gen tham

<sup>1</sup> Viện Nghiên cứu Nuôi trồng Thủy sản I

<sup>2</sup> Học viện Nông nghiệp Việt Nam

\*Email: vttrang@rial.org

chiều và phụ trợ cho các nghiên cứu cần thông tin về di truyền của một số đàn cá chép hiện có ở Việt Nam.

**2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU**

**2.1. Vật liệu nghiên cứu**

Thông tin chi tiết về các đàn cá chép (*Cyprinus carpio*) sử dụng trong nghiên cứu được trình bày trong bảng 1.

Mẫu vây ngực của các cá thể thuộc các đàn cá nghiên cứu (như mô tả trong bảng 1) được thu, bảo quản trong Ethanol 96% và được chuyển về Phòng thí nghiệm Di truyền – Chọn giống thuộc Trung tâm Công nghệ Sinh học Thủy sản (Viện Nghiên cứu Nuôi trồng Thủy sản I), bảo quản trong tủ 4 C cho đến khi tách chiết ADN.

**Bảng 1. Thông tin mẫu các đàn cá chép sử dụng trong nghiên cứu**

TT	Tên đàn	Nguồn gốc	Thời gian bắt đầu bảo tồn, lưu giữ tại Việt Nam	Kí hiệu mẫu	Số lượng (mẫu/đàn)	Địa điểm lưu giữ
1	Cá chép trắng Việt Nam	Việt Nam	1988*	VN	17	Viện Nghiên cứu Nuôi trồng Thủy sản I
2	Cá chép chọn giống V1	Việt Nam	2005*	V1	20	
3	Cá chép vây Hungary	Hungary	1996*	HV	20	
4	Cá chép vàng Indonesia	Indonesia	1988*	IN	20	
5	Cá chép vây Hungary đàn Tata	Hungary	2019	TAT	20	
6	Cá chép vây Hungary đàn Szarvas	Hungary	2019	SZA	20	
7	Cá chép đàn Séc	CH Séc	2015	SE	20	Doanh nghiệp tư nhân Trung tâm phát triển Công nghệ thủy sản Việt Nam, địa chỉ: Đình Bảng, Từ Sơn, Bắc Ninh.
<b>Tổng số</b>					<b>137</b>	

\*Nguyễn Quang Huy, 2017.

**2.2. Phương pháp nghiên cứu**

**Tách chiết ADN:** Trong nghiên cứu này đã tách chiết ADN theo phương pháp kết tủa muối (Sambrook và Russell, 2001).

**Phản ứng PCR khuếch đại gen (PCR1):** Nghiên cứu sử dụng đoạn môi khuếch đại đoạn gen COI và D-loop thuộc mtDNA của loài cá chép *Cyprinus*

*carpio*. Hai đoạn môi này được thiết kế dựa trên phần mềm Primer3Plus (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) trên cơ sở đoạn gen COI và Dloop của loài cá này đã được công bố trên ngân hàng gen NCBI với mã số NC001606 (Chang và ctv., 1994). Thông tin chi tiết về đoạn môi được trình bày ở bảng 2.

**Bảng 2. Thông tin về trình tự đoạn môi sử dụng trong nghiên cứu**

Tên môi	Trình tự môi	Nhiệt độ gắn môi (°C)	Kích thước sản phẩm PCR (bp)
COI-F. <i>C. Carp</i>	5'-CACGCAGGAGCATCAGTAGA-3'	58	819
COI-R. <i>C. Carp</i>	3'-TGTTTCAGGTGCTGTGGAGAG-5'	58	
DL-F. <i>C. Carp</i>	5'-GGAGGTAGCACTCCCTTTATG-3'	58	715
DL-R. <i>C. Carp</i>	5'-CACACGTTCTTGGTCCTCCT-3'	58	

Phản ứng khuếch đại được thực hiện với tổng thể tích 50  $\mu$ l bao gồm: 25 $\mu$ l MyTaq™ Mix 2 $\times$  (Bioline), 2  $\mu$ l mỗi loại mỗi xuôi và mỗi ngược (nồng độ mỗi 10  $\mu$ M), 2  $\mu$ l ADN khuôn (~ 200 ng) và nước cất. Điều kiện phản ứng đối với chu kỳ nhiệt nhân đoạn gen COI, Dloop là: giai đoạn khởi đầu ở 94°C trong 2 phút; tiếp đó là 35 chu kỳ bao gồm (giai đoạn biến tính ở 94°C trong 30 giây; giai đoạn gắn mỗi ở 58°C trong 15 giây; giai đoạn kéo dài ở 72°C trong 30 giây); giai đoạn kết thúc kéo dài ở 72°C trong 5 phút và giữ ở 4°C. Sau khi phản ứng kết thúc lấy 3 $\mu$ l sản phẩm điện di trên gel agarose 1,5% để kiểm tra kết quả sản phẩm PCR.

**Tinh sạch sản phẩm PCR1:** Sử dụng kit tinh sạch sản phẩm PCR, QIAquick PCR Purification Kit, Qiagen. Quy trình tinh sạch theo hướng dẫn của nhà sản xuất.

**Phản ứng PCR giải trình tự (PCR2):** Thành phần của phản ứng PCR giải trình tự trong tổng thể tích 20  $\mu$ l bao gồm: 8  $\mu$ l DTCS Quick Start Kit, 0,5-10  $\mu$ l sản phẩm PCR1 đã tinh sạch đảm bảo nồng độ 50 ng/100 fmol, 2  $\mu$ l Primer (1,6 pmol/ $\mu$ l hoặc 1,6  $\mu$ M). Sau đó cho nước cất lên tổng thể tích 20  $\mu$ l. Chu trình nhiệt cho phản ứng PCR2 là 30 chu kỳ bao gồm: 96°C trong 20 giây, 50°C trong 20 giây và 60°C trong 4 phút. Cuối cùng giữ ở 4°C.

**Tinh sạch sản phẩm PCR2:** Tinh sạch bằng Ethanol (Merk) theo quy trình sau: Chuyển toàn bộ sản phẩm PCR2 (20  $\mu$ l) sang ống eppendorf đã chứa 5  $\mu$ l dung dịch ngừng phản ứng bao gồm: 2  $\mu$ l 3M Natri Acetat pH 5,2; 2  $\mu$ l EDTA 100 mM; pH 8,0, 1  $\mu$ l Glycogen 20 mg/ml. Thêm 75  $\mu$ l Ethanol 95% lạnh bảo quản trong -20°C, vortex, ly tâm 13.000 vòng/phút trong 20 phút tại 4°C. Bỏ dịch nổi phía trên, giữ lại ADN kết tủa ở đáy ống. Rửa kết tủa bằng 200  $\mu$ l Ethanol 70% lạnh bảo quản trong -20°C, vortex, ly tâm 13.000 vòng/phút trong 3,5 phút tại 4°C. Bỏ dịch nổi phía trên, giữ lại ADN kết tủa ở đáy ống. Bước này thực hiện 2 lần. Để khô tự nhiên ở nhiệt độ phòng. Hòa tan ADN kết tủa bằng 40  $\mu$ l SLS.

**Thực hiện giải trình tự trên máy GenomeLab GeXP:** Chuyển toàn bộ dung dịch ADN hòa tan trong SLS từ bước tinh sạch vào giếng tương ứng trên đĩa chạy mẫu. Nhỏ 1 giọt dầu (mineral oil – được cung kèm theo DTCS Quick Start Kit) lên trên để

tránh bay hơi mẫu. Đặt đĩa mẫu và đĩa đệm vào máy giải trình tự. Cài đặt và chạy chương trình.

### 2.3. Xử lý số liệu

Các trình tự gen được kiểm tra chất lượng bằng phần mềm Finch TV 1.4.0 (<http://www.geospiza.com>). Tiếp theo, chương trình ClustalW Multiple Alignment trong BioEdit 7.1.9 (Hall, 1999) được sử dụng nhằm so sánh, căn chỉnh trình tự, sau đó được kiểm tra và xác định mức độ tương đồng. Các chỉ số đa dạng di truyền kiểu đơn - Haplotype diversity (h) và đa dạng nucleotit - Nucleotide diversity ( $\pi$ ) được tính toán bằng cách sử dụng phần mềm DnaSP v6.10.01 (Rozas và ctv., 2017). Phần mềm Arlequin v.3.11 (Excoffier và ctv., 2005) được dùng để đánh giá sai khác di truyền, mức độ khác biệt bên trong và giữa các đàn qua phân tích  $F_{ST}$  (Population pairwise  $F_{ST}$ ) và phương sai phân tử AMOVA.

Cây phát sinh loài phân tử thể hiện mối quan hệ di truyền giữa các đàn cá nghiên cứu được xây dựng dựa trên phương pháp UPGMA (Unweighted Pair Group Method with Arithmetic mean) gồm 8 trình tự, trong đó có 7 trình tự phổ biến (consensus sequence) được xây dựng cho từng đàn và 1 trình tự đã được công bố trên NCBI (mã số NC001606; Chang và ctv., 1994). Các đơn vị phân loại liên kết được nhóm lại với nhau trong thử nghiệm bootstrap (1000 lần lặp lại) được hiển thị bên cạnh các nhánh. Cây được vẽ theo tỷ lệ, với độ dài các nhánh bằng đơn vị của khoảng cách tiến hóa được sử dụng để suy ra cây phát sinh loài. Các phân tích tiến hóa được thực hiện trong MEGA X (Kumar và ctv., 2018).

## 3. KẾT QUẢ VÀ THẢO LUẬN

### 3.1. Đa dạng haplotype và nucleotide

Tất cả các trình tự trước khi đưa vào phân tích bằng phần mềm đều được xác định mức độ tương đồng với trình tự gen đã được công bố trên ngân hàng gen NCBI dựa trên công cụ BLAST. Kết quả cho thấy, các trình tự gen nghiên cứu có độ bao phủ rất cao (100%) và độ tương đồng cao (97,7-100%) so với các trình tự đã được công bố. Sau khi loại bỏ các vùng trình tự nhiễu (đoạn đầu và đoạn cuối) cho kích thước đoạn gen nghiên cứu là 760 bp đối với gen COI và 695 bp đối với gen D-loop. Kết quả phân tích đa dạng di truyền của các đàn cá chép dựa trên 2 gen này được thể hiện ở bảng 3.

Bảng 3. Tổng hợp đa dạng haplotype và nucleotide của các đàn cá chép phân tích trên đoạn gen COI và D-loop

STT	Đàn cá	COI					D-loop				
		N	n	p	h	$\pi$	N	n	p	h	$\pi$
1	VN	17	6	4	0,772	0,00311	17	16	15	0,993	0,01355
2	V1	20	8	3	0,863	0,00534	20	15	11	0,958	0,01879
3	HV	20	1	0	0	0	20	4	1	0,5	0,00339
4	IN	20	3	0	0,484	0,00182	20	7	2	0,821	0,00722
5	TAT	20	2	1	0,189	0,00025	20	1	0	0	0
6	SZA	20	1	0	0	0	20	5	1	0,511	0,00120
7	SE	20	1	0	0	0	20	8	4	0,826	0,01188
Tổng số		137	13	8	0,732	0,00374	137	40	34	0,829	0,01234

Ghi chú: N: số trình tự được phân tích; n: số lượng haplotype; p: số lượng haplotype đặc trưng; h: đa dạng haplotype;  $\pi$ : đa dạng nucleotide.

Đối với gen COI, kết quả nghiên cứu đã chỉ ra có 13 haplotype khác nhau cho 7 đàn cá chép, trình tự các haplotype này đã được công bố trên GenBank - NCBI số hiệu MW450991 - MW451003. Cụ thể, đàn cá chép V1 có nhiều haplotype nhất (8), trong khi đó 3 đàn HV, SZA, SE chỉ có 1 haplotype duy nhất. Tuy V1 xuất hiện ở nhiều dạng haplotype nhất nhưng đàn VN mới là đàn có nhiều haplotype đặc trưng nhất (4), tiếp đến là đàn V1 (3), cuối cùng là đàn TAT (1). Bốn đàn cá còn lại (bao gồm IN, HV, SE và SZA) không có haplotype đặc trưng. Giá trị đa dạng haplotype h (trung bình là 0,732) và đa dạng nucleotide  $\pi$  (trung bình là 0,00374) nhìn chung là cao khi so sánh với một số loài cá nước ngọt khác đã được nghiên cứu. Đàn V1 có đa dạng haplotype và đa dạng nucleotide cao nhất ( $h = 0,863$ ,  $\pi = 0,00534$ ). Trong khi đó, 3/7 đàn được nghiên cứu có đa dạng haplotype và đa dạng nucleotide đều bằng 0 (đàn HV, SZA và SE). Điều đó phản ánh không có sự biến đổi di truyền ở đoạn gen nghiên cứu giữa các cá thể được phân tích trong các đàn này. Theo nghiên cứu của Zhao và ctv (2020) sử dụng chỉ thị khuếch đại đoạn gen COII (Cytochrome c oxidase submit 2) để nghiên cứu biến dị di truyền trên 196 mẫu cá chép thuộc 5 quần thể tự nhiên và 10 đàn nội địa (6 đàn Trung Quốc, 4 đàn châu Âu) đã cho thấy đa dạng haplotype ( $h = 0,761$ ) cao hơn nhưng đa dạng nucleotide ( $\pi = 0,0024$ ) là thấp hơn nghiên cứu của chúng tôi.

Đối với gen D-loop, kết quả phân tích 137 trình tự thuộc 7 đàn cá nghiên cứu đã chỉ ra có 40 haplotype khác nhau, trình tự các haplotype này đã được công bố trên GenBank với số hiệu là MW463062 - MW463101. Đàn VN là đàn có số lượng haplotype nhiều nhất (16 haplotype/17 mẫu nghiên

cứu) và cũng là đàn có nhiều haplotype đặc trưng nhất (15). Tiếp đó là đàn V1 với n và p lần lượt là 15 và 11. Trong khi đó đàn TAT chỉ có 1 dạng haplotype và không có haplotype đặc trưng dẫn đến h và  $\pi$  đều bằng 0. Các đàn cá có nhiều haplotype dẫn đến đa dạng haplotype và đa dạng nucleotide cao và ngược lại. Trong đó đáng chú ý là đàn VN có h cao nhất (0,993), tiếp đó là V1 (0,958). Các đàn có nguồn gốc từ Hungary (HV, TAT và SZA) nhìn chung có các giá trị n, p, h,  $\pi$  thấp, các đàn còn lại ở mức độ trung bình. So sánh với nghiên cứu của Thái Thanh Bình và ctv (2006) sử dụng phương pháp kết hợp giữa giải trình tự ADN và phân tích đa hình cấu trúc sợi đơn (SSCP) tại gen D-loop trên 968 mẫu cá chép thu tại các vùng khác nhau ở Việt Nam, Trung Quốc, Ấn Độ, Indonesia, Đài Loan cho thấy đa dạng haplotype (trung bình 0,4) thấp hơn nghiên cứu của chúng tôi.

Trong nghiên cứu này, các đàn HV, TAT, SZA, SE và IN là kết quả của các chương trình chọn giống hoặc đã được nhập nội từ nhiều năm trước đây. Sau khi đạt được các mục tiêu mà chương trình chọn giống đã thiết kế (như sinh trưởng nhanh, tỷ lệ sống cao, hệ số chuyển đổi thức ăn thấp) thì các đàn cá chọn giống hoặc đàn nhập nội chỉ được tái đàn để lưu giữ qua các năm. Đây có thể là lý do chính dẫn đến việc làm giảm biến dị di truyền trong các đàn cá này dựa trên các đoạn gen đã nghiên cứu. Ngược lại, đàn VN và V1, xuất hiện ở nhiều dạng haplotype và cũng là những đàn có nhiều haplotype đặc trưng nhất khi xét trên cả 2 gen COI và D-loop. Điều này có thể được giải thích bởi nguồn gốc của hai đàn cá này. Đối với đàn cá V1, đây là cá chép chọn giống thế hệ chọn lọc thứ 6 của những cá lai 3 máu (hay còn gọi là con lai kép) giữa cá chép Việt Nam, Hungary và

Indonesia (tiến hành từ 1984 - 1995), trong khi đó đàn VN là đàn có nguồn gốc tự nhiên.

### 3.2. Sai khác di truyền

Trong nghiên cứu này, sự sai khác di truyền giữa các đàn cá chép được ước lượng theo giá trị  $F_{ST}$ . Kết quả phân tích sự sai khác di truyền của 7 đàn cá khi so sánh từng cặp với nhau được trình bày ở bảng 4.

Kết quả ở bảng 4 cho thấy sự sai khác di truyền của 7 đàn cá khi so sánh từng cặp với nhau xét trên cả 2 gen COI và D-loop phần lớn đều là sai khác có ý nghĩa ( $P < 0,05$ ). Không có sự sai khác di truyền giữa 3 đàn HV, TAT, SZA trên cả gen COI và D-loop. Điều này có thể được giải thích bởi 3 đàn này đều có nguồn gốc từ Hungary. Trong đó đàn HV là đàn cá nhập nội và được tái đàn, lưu giữ ở nước ta từ năm 1996 cho đến nay. Hai đàn TAT và SZA được nhập về nước ta từ năm 2019. Đối với các trường hợp có sự sai khác di truyền lớn, đáng chú ý nhất là sự sai khác di truyền xét trên gen COI giữa đàn SE với tất cả các

đàn còn lại, đặc biệt là giữa SE và HV, giữa SE và SZA (đều có  $F_{ST} = 1$ ), SE và TAT ( $F_{ST} = 0,96356$ ). Theo các nhà khoa học, khi  $F_{ST} = 1$ , thì các quần thể là cố định, điều đó có nghĩa là giữa các quần thể không chia sẻ bất kỳ alen nào với nhau, tức là không sinh sản với nhau, chúng hoàn toàn bị cô lập với nhau. Việc  $F_{ST}$  trong nghiên cứu này có giá trị bằng 1 có thể được giải thích là do: (1) Các đàn này bị cô lập với nhau hoàn toàn trong thời gian dài. Thực tế cho thấy các đàn này có sự khác biệt về nguồn gốc, có cách ly địa lý, có cách quản lý, lưu giữ đàn nghiêm ngặt nên việc cách ly sinh sản hoàn toàn có thể xảy ra. (2) Gen sử dụng trong nghiên cứu là gen thuộc mtDNA chỉ di truyền theo dòng mẹ. (3) Với một số đàn số lượng cá thể giới hạn, việc sinh sản để tái đàn trong thời gian dài có thể làm gia tăng tỷ lệ giao phối cận huyết. (4) Số lượng mẫu phân tích có thể chưa đủ lớn (20 mẫu/đàn) dẫn đến xác suất bắt gặp alen giảm.

**Bảng 4. Sai khác di truyền giữa các đàn cá chép dựa trên phân tích đoạn gen COI và D-loop**

Tên đàn	VN	V1	HV	IN	TAT	SZA	SE
VN		0,10759	0,55177	0,18306	0,64502	0,61416	0,14309
V1	0,17266		0,24344	0,16203	0,32819	0,29974	0,14660
HV	0,52722	0,15132		0,68513	0,12341	0,04279	0,60900
IN	0,43078	0,32554	0,76842		0,77525	0,74356	0,16137
TAT	0,49163	0,13876	0,05263	0,72368		0,07018	0,69364
SZA	0,52722	0,15132	0,00000	0,76842	0,05263		0,66725
SE	0,73754	0,60830	1,00000	0,86640	0,96356	1,00000	

*Ghi chú:  $F_{ST}$  dựa trên gen COI (phía dưới đường chéo);  $F_{ST}$  dựa trên gen D-loop (phía trên đường chéo); Sai khác có ý nghĩa khi  $P < 0,05$ ; Giá trị  $F_{ST}$  có ý nghĩa được in đậm; Giá trị  $F_{ST}$  không có ý nghĩa được in nghiêng.*

### 3.3. Cấu trúc di truyền

Kết quả phân tích AMOVA được thể hiện ở bảng 5.

**Bảng 5. Phân tích phương sai phân tử (AMOVA) của các đàn cá chép qua phân tích trình tự gen COI và D-loop**

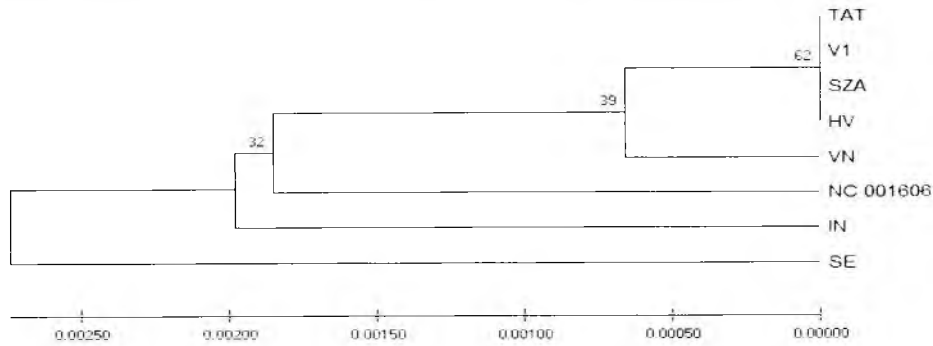
Gen	Nguồn biến động	Độ tự do (df)	Tổng bình phương (SS)	Thành phần biến động	Phần trăm biến động	$F_{ST}$
COI	Giữa các đàn	6	122,394	1,01213 <i>Va</i>	62,79	0,62791**
	Trong một đàn	130	77,971	0,59977 <i>Vb</i>	37,21	
	<i>Tổng</i>	<i>136</i>	<i>200,365</i>	<i>1,61190</i>		
D-loop	Giữa các đàn	6	278,239	2,22658 <i>Va</i>	44,15	0,44148**
	Trong một đàn	130	366,191	2,81686 <i>Vb</i>	55,85	
	<i>Tổng</i>	<i>136</i>	<i>644,431</i>	<i>5,04343</i>		

*Ghi chú: Va, Vb: số lượng thành phần phương sai, \*\*  $P < 0,0001$ .*

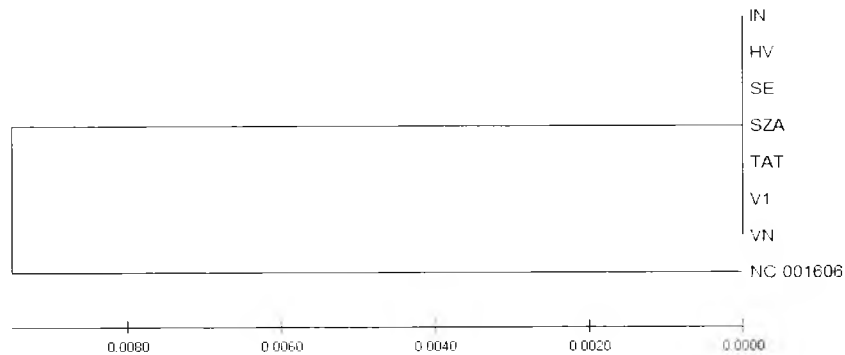
Kết quả phân tích AMOVA cho thấy phần trăm biến động giữa các đàn (62,79%) lớn hơn sự biến động trong một đàn (37,21%) trên gen COI, trong khi trên gen D-loop thì xu hướng ngược lại (44,15% và 55,85%). Kết quả đó chỉ ra rằng có sự sai khác về di truyền giữa các đàn hoặc các cá thể được nghiên cứu, đồng thời cho thấy các đàn cá có cấu trúc quần thể rõ ràng do sự đóng góp đáng kể về khác biệt di

truyền giữa các đàn khi xét trên gen COI. Giá trị  $F_{ST}$  là 0,62719 (COD), 0,44148 (D-loop) cho thấy sự biến đổi di truyền tổng thể giữa 7 đàn ở mức độ lớn. Theo Zhao và ctv (2020) khi phân tích AMOVA dựa trên gen COII cho các quần thể cá chép cho thấy phần lớn biến đổi di truyền là xảy ra trong quần thể (72,71%).

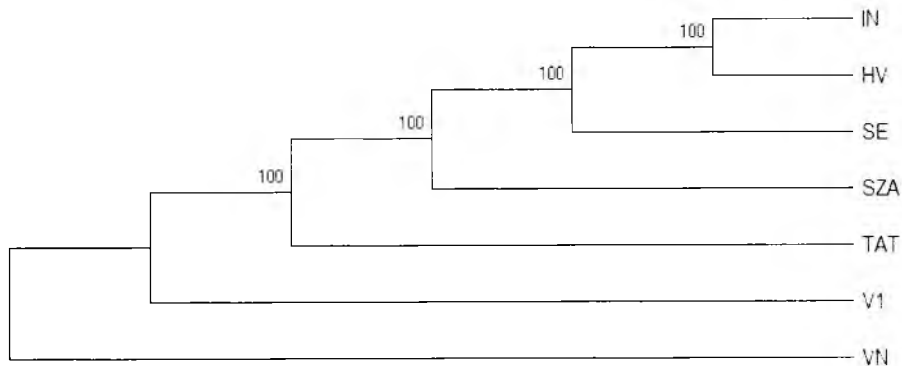
3.4. Mối quan hệ di truyền



Hình 1. Mối quan hệ di truyền (UPGMA) của 7 đàn cá chép xây dựng dựa trên gen COI



Hình 2. Mối quan hệ di truyền (UPGMA) của 7 đàn cá chép xây dựng dựa trên gen D-loop



Hình 3. Mối quan hệ di truyền (UPGMA) của 7 đàn cá chép xây dựng dựa trên gen D-loop không bao gồm trình tự NC 001606

Hình 1 cho thấy đàn SE sau đó là đàn IN có mối quan hệ di truyền xa hơn với các đàn còn lại cũng như với NC 001606. Nghiên cứu sự phân nhánh trên nhánh phụ, một lần nữa cho thấy mối quan hệ di

truyền gần gũi của 3 đàn có nguồn gốc từ Hungary là HV, TAT và SZA. Đàn chọn giống V1 cũng thuộc nhánh này do đó là con lai của cá có nguồn gốc từ Hungary. Mối quan hệ gần gũi này được thể hiện

bằng chiều dài nhánh rất nhỏ hoặc bằng 0. Xét trên gen D-loop, hình 2 cho thấy phân thành 2 nhánh trong đó 7 đàn nghiên cứu vào một nhánh, nhánh còn lại chỉ có NC 001606. Điều đó có nghĩa 7 đàn cá có mối quan hệ di truyền gần gũi khi xét trên gen D-loop. Để hiểu rõ hơn mối quan hệ di truyền của 7 đàn này dựa trên gen D-loop, tiếp tục xây dựng cây phát sinh loài phân tử với phương pháp tương tự và dữ liệu chỉ loại bỏ NC 001606 (Hình 3). Theo hình 3, đàn VN sau đó là đàn V1 là những đàn có mối quan hệ di truyền xa hơn với các đàn còn lại.

#### 4. KẾT LUẬN

Kết quả phân tích đa dạng di truyền của 7 đàn cá chép dựa trên phân tích trình tự gen COI và D-loop thuộc mtDNA cho thấy cá chép VN và V1 có đa dạng haplotype và nucleotide cao nhất. Các đàn cá nghiên cứu có sai khác di truyền lớn, đặc biệt là sai khác giữa cá chép SE với tất cả các đàn cá còn lại. Kết quả phân tích AMOVA dựa trên gen COI cho thấy các đàn cá có cấu trúc quần thể rõ ràng. Cây phát sinh loài phân tử thể hiện mối quan hệ di truyền xa giữa SE và các đàn cá còn lại (xét trên gen COI) và mối quan hệ di truyền gần gũi của 7 đàn cá nghiên cứu (xét trên gen D-loop).

#### LỜI CẢM ƠN

Nhóm tác giả xin chân thành cảm ơn Viện Nghiên cứu Nuôi trồng Thủy sản I đã cấp kinh phí để thực hiện nghiên cứu này.

#### TÀI LIỆU THAM KHẢO

1. Chang Y. S., Huang F. L., Lo T. B., 1994. The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. *J. Mol. Evol.* 38 (2), 138-155.
2. Excoffier L., Laval G., Schneider S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform.* (1) 47-50.
3. Hall T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95-98.
4. Hubert N., Hanner R., Holm E., Mandrak N. E., Taylor E., Burrige M., Watkinson D., Dumont P., Curry A., Bentzen P., Zhang J., 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS one*, 3(6), p.e2490.

5. Karnai, L. and I. Szúcs., 2018. Outlooks and perspectives of the common carp production. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu XX*: 64-71.

6. Kohlmann K and Kersten P., 2013. Deeper insight into the origin and spread of European common carp (*Cyprinus carpio carpio*) based on mitochondrial D-loop sequence polymorphisms. *Aquaculture*, 376-379:97-104.

7. Kumar S., Stecher G., Li M., Knyaz C., and Tamura K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.

8. Nguyễn Quang Huy, 2017. Báo cáo tổng kết nhiệm vụ "Bảo tồn, lưu giữ nguồn gen và giống thủy sản". Viện Nghiên cứu Nuôi trồng Thủy sản 1. Bộ Nông nghiệp và Phát triển nông thôn.

9. Rahman M. M., 2015. Effects of co-cultured common carp on nutrients and food web dynamics in rohu aquaculture ponds. *Aquacult Environ Interact* 6:223-232.

10. Rozas J., Ferrer-Mata A., Sánchez-DelBarrio J. C., Guirao-Rico S., Librado P., Ramos-Onsins S. E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular biology and evolution* 34(12): 3299-3302.

11. Sambrook J and Russell D., 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

12. Thai B. T., Pham T. A., Austin C. M., 2006. Genetic diversity of common carp in Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA control region. *Aquaculture* 258(1):228-240.

13. Ward R. D., Zemlak T. S., Innes B. H., Last P. R. and Hebert P. D., 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462):1847-1857.

14. Weber M. J. and Brown M. L., 2011. Relationships among invasive common carp, native fishes and physicochemical characteristics in upper Midwest (USA) lakes. *Ecology of Freshwater Fish*, 20(2):270-278.

15. Zhao Y., Zheng X., Zhu X., Kuang Y., Sun X., 2020. Genetic variation of common carp *Cyprinus carpio* L. in China based on mitochondrial COII gene. *Aquaculture Reports* 18: 100462.

**GENETIC DIVERSITY AND POPULATION STRUCTURE OF COMMON CARP POPULATIONS IN VIETNAM**

Vu Thi Trang<sup>1\*</sup>, Luu Thi Ha Giang<sup>1</sup>, Pham Hong Nhat<sup>1</sup>,

Le Van Khoi<sup>1</sup>, Bui Thi Anh Nguyet<sup>2</sup>, Nguyen Huu Duc<sup>2</sup>, Dang Thi Lua<sup>1,2</sup>

<sup>1</sup>*Research Institute for Aquaculture No.1 (RIA1)*

<sup>2</sup>*Vietnam National University of Agriculture (VNUA)*

*Email: vttrang@ria1.org*

**Summary**

In this study, genetic diversity, population structure, and phylogenetics of 7 common carp strains including Vietnamese white carp (VN), breeding carp V1 (V1), Hungarian scaly carp (HV), Indonesian goldfish (IN), Hungarian scaly carp Tata (TAT), Hungarian scaly carp Szarvas (SZA) and Czech Republic carp (SE) based on COI and D-loop sequences analysis. Results showed that VN and V1 were the highest genetic diversity populations while HV, TAT, SZA, SE, and IN had low genetic diversity in both COI and D-loop genes. This study identified a total of 53 haplotypes and all haplotype sequences were deposited in the GenBank database, accession numbers MW450991-MW451003 (for COI gene) and MW463062-MW463101 (for D-loop gene). In terms of genetic differentiation, the most notable was the genetic difference in the COI gene between SE and all other populations, specifically, between SE and HV, SE and SZA (both had  $F_{ST} = 1$ ), this suggests that there was no genetic sharing between these populations based on studied genes. All populations had a clear population structure due to the significant contribution of genetic differences among populations based on COI AMOVA analysis. Phylogenetic presented a long genetic relationship between SE and other remaining populations (for COI gene) and a close genetic relationship of 7 studied populations (for D-loop gene). Our research contributed to creating a database of references and supporting studies that need information on genetic resources of some common carp populations in Vietnam as well as serving for breeding and conserving genetic resources.

**Keywords:** *COI, D-loop, common carp, genetic diversity, haplotype.*

**Người phản biện:** PGS.TS. Nguyễn Hữu Ninh

**Ngày nhận bài:** 18/02/2021

**Ngày thông qua phản biện:** 18/3/2021

**Ngày duyệt đăng:** 25/3/2021