

GIẢI MÃ TOÀN BỘ HỆ GEN VÀ PHÂN TÍCH XÁC ĐỊNH VỊ TRÍ PHÂN LOẠI CỦA CHŨNG VIRUS VIÊM GAN VỊT CƯỜNG ĐỘC DN2 THUỘC GENOTYPE III (DHAV-3) LẦN ĐẦU TIÊN PHÁT HIỆN TẠI VIỆT NAM

Đoàn Thị Thanh Hương¹, Nguyễn Bá Hiên², Trần Xuân Hạnh³, Lê Thanh Hòa¹

¹Viện Công nghệ sinh học, Viện Khoa học và Công nghệ Việt Nam

²Trường Đại học Nông nghiệp Hà Nội

³Công ty Thuốc Thú y trung ương (NAVETCO), thành phố Hồ Chí Minh

TÓM TẮT

Toàn bộ chuỗi nucleotide của hệ gen virus viêm gan vịt cường độc lần đầu tiên phân lập tại Việt Nam, ký hiệu DN2, gồm 7777 bp đã được giải mã và xác định thuộc genotype III (DHAV-3). Chuỗi nucleotide của toàn bộ hệ gen DN2 đã được thu nhận bằng phương pháp PCR phân đoạn, sử dụng 5 cặp mồi móc gối vào nhau, với khuôn là cDNA được chuyển đổi bằng hexamer xác suất và giải trình tự trực tiếp. Trật tự hệ gen DHAV-3 chủng DN2, sắp xếp gồm 3 phần chính, lần lượt gồm một vùng không mã hóa ở đầu 5' gọi là 5'UTR có độ dài là 652 bp, tiếp đến là một khung đọc mở (ORF) kích thước 6756 bp, mã hoá cho một polypeptide chung, cuối cùng là vùng không mã hóa ở đầu 3' (3'UTR) độ dài 369 bp, cuối cùng nối tiếp bằng một dãy polyA. Khung đọc mở gồm 6756 bp, mã hoá 2251 amino acid, gồm 12 phân đoạn protein bao gồm L; VP0, VP3, VP1 (thuộc tổ hợp P1); 2A1, 2A2, 2B, 2C (thuộc tổ hợp P2); 3A, 3B, 3C, 3D (thuộc tổ hợp P3). Sự phân cắt thành 12 protein thành phần được xảy ra tại các điểm cắt có sự tham gia của các amino acid L/G, Q/G, Q/S, E/L và NPG/P. So sánh với chủng AP-04114 (GenBank: DQ812093, Hàn Quốc) cùng thuộc DHAV-3 cho thấy, sắp xếp gen và số lượng nucleotide mỗi gen là giống nhau, chỉ vùng 3'UTR là sai khác 2 nucleotide. So sánh với các chủng thuộc DHAV-1 (chủng C80, GenBank: DQ864514) và DHAV-2 (chủng 04G, GenBank: EF067923) cho thấy, kích thước các phân đoạn gen là tương tự, ngoại trừ phần gen của VP1 có độ dài 720 bp ở DHAV-3 và 714 bp ở DHAV-1. Vùng không mã hóa của DHAV-1 (chủng C80) có kích thước ngắn nhất (626 bp của 5'UTR và 315 bp của 3'UTR).

Từ khóa: DHAV-1, DHAV-2, DHAV-3, hệ gen, phân đoạn gen, viêm gan vịt, vùng mã hoá, vùng không mã hóa

ĐẶT VẤN ĐỀ

Bệnh viêm gan vịt là bệnh truyền nhiễm cấp tính xảy ra ở vịt con 2 - 6 tuần tuổi, tỷ lệ chết cao, gây thiệt hại nặng nề cho chăn nuôi vịt ở các nước trên thế giới, trong đó có Việt Nam (OIE, 2008; Woolcock, 2003). Virus viêm gan vịt (Duck Hepatitis Virus A - DHAV) cho đến gần đây được phân loại gồm 3 genotype: genotype I (DHAV-1), genotype II (DHAV-2) và genotype III (DHAV-3) (Wang *et al.*, 2008) khác với trước đây phân loại chỉ căn cứ vào kết quả huyết thanh học (Woolcock, 2003; OIE, 2008). DHAV-1 phát hiện ở hầu hết các nước trên thế giới; DHAV-2 ở Đài Loan; và DHAV-3 mới phát hiện gần đây ở Hàn Quốc và Trung Quốc (Kim *et al.*, 2007; 2009). DHAV thuộc họ *Picornaviridae*, có cấu trúc hệ gen là RNA sợi đơn dương, gồm 7.600 đến 7.700 nucleotide (tuỳ từng chủng), ngoài vùng không mã hóa (UTR) ở đầu 5' và 3', chỉ chứa duy nhất một khung đọc mở (open reading frame, ORF) gồm 3 tổ hợp gen là P1, P2 và

P3 (Kim *et al.*, 2006; 2007; Tseng, Tsai, 2007a,b; Ding, Zhang, 2007; Tseng *et al.*, 2007). Polypeptide được phân cắt tạo nên 12 phân đoạn protein thành phần gồm L, VP0, VP3, VP1, 2A1, 2A2, 2B, 2C, 3A, 3B, 3C, 3D; trong nhiều trường hợp chúng tiếp tục phân cắt thành các polypeptide bé hơn (Ding, Zhang, 2007).

Trên thế giới, mặc dù được phát hiện từ trước những năm 1950 (Levine, Hofstad, 1945), nhưng nghiên cứu virus viêm gan vịt chỉ dừng ở mức độ khảo sát lâm sàng bệnh tích, chẩn đoán virus học và phát triển vaccine truyền thống để phòng chống (OIE, 2008). Tuy nhiên, trong vòng 5 năm nay, virus viêm gan vịt được nghiên cứu sâu sắc về gen và hệ gen, đặc biệt về ứng dụng sinh học phân tử trong giám định, nghiên cứu hệ gen, phân loại và tìm hiểu quan hệ sinh học trong họ *Picornaviridae* (Tseng, Tsai, 2007a,b). Có nhiều quan điểm khác nhau về phân loại đã được đưa ra để phân biệt các chủng của các type virus viêm gan vịt khác nhau mới được phát hiện trên thế giới (Kim *et al.*, 2006; 2007; Wang *et*

al., 2008). Cho đến nay, gần 50 hệ gen hoàn chỉnh của virus viêm gan vịt đã được giải mã và phân tích, trong đó có trên 30 chủng thuộc DHAV-1, hai chủng thuộc DHAV-2 và 15 chủng thuộc DHAV-3, cũng như một số phân đoạn gen/vùng gen của một số chủng khác nhau đã được thu nhận (<http://www.picornaviridae.com/avihepatovirus/dhav/dhav.htm>).

Từ trước tới nay, tại Việt Nam, virus gây bệnh viêm gan vịt chỉ mới xác định thuộc genotype I (DHAV-1), bao gồm cả một số chủng virus vaccine hiện đang được sử dụng (Lê Thanh Hòa *et al.*, 1986; Nguyễn Đức Lư, Vũ Như Quán, 2002; Đoàn Thị Thanh Hương *et al.*, 2009). Trong một nghiên cứu gần đây, bằng cặp mồi đặc hiệu của DHAV-3, chúng tôi đã phát hiện và xác định một số chủng virus viêm gan vịt (DN1, DN2, NC, BM) thuộc genotype III (DHAV-3), lần đầu tiên được phân lập tại Việt Nam (Đoàn Thị Thanh Hương *et al.*, 2010). Mức độ tương đồng về amino acid của protein kháng nguyên VP1 giữa virus cường độc thuộc DHAV-3 với một chủng vaccine thuộc DHAV-1 đang sử dụng tại Việt Nam là rất thấp, chỉ có 74 - 77% (Đoàn Thị Thanh Hương *et al.*, 2010). Với mục đích có được dữ liệu gen và hệ gen của một chủng thuộc DHAV-3 của Việt Nam, chúng tôi tiến hành giải mã và phân tích toàn bộ hệ gen của chủng DN2 phân lập tại Đồng Nai năm 2009.

Trong bài báo này, chúng tôi giới thiệu quá trình thu nhận và giải mã hệ gen chủng DN2 thuộc DHAV-3 và những đặc điểm tổng thể về cấu trúc của hệ gen phân tích từ chủng này.

NGUYÊN LIỆU VÀ PHƯƠNG PHÁP

Virus cường độc chủng DN2 và tách chiết RNA chứa hệ gen virus

Mẫu bệnh phẩm viêm gan vịt phân lập ở vịt bệnh tại tỉnh Đồng Nai năm 2009, được tiếp truyền qua phôi trứng vịt 11 ngày tuổi tại Bộ môn Vi sinh vật - Truyền nhiễm, Khoa Thú y, Trường Đại học Nông nghiệp Hà Nội. Virus viêm gan vịt cường độc ký hiệu DN2, được giám định là thuộc type III (DHAV-3) bằng chỉ thị phân tử gen kháng nguyên VP1 (Đoàn Thị Thanh Hương *et al.*, 2010).

RNA hệ gen của virus được tách chiết bằng bộ sinh phẩm QIA-amp Viral Mini kit (QIAGEN Inc.) theo hướng dẫn của nhà sản xuất và bảo

quản ở -20°C cho đến khi sử dụng.

Chuyển đổi cDNA từ RNA bằng mồi xác suất hexamer

DNA bổ sung (cDNA) được tổng hợp theo phương pháp chuyển đổi từ RNA hệ gen của virus bằng mồi xác suất (hexamer) sử dụng bộ kit và enzyme Maxima Reverse Transcriptase của hãng Fermentas. Phản ứng cDNA được thực hiện dùng tích 20 μl gồm có các thành phần sau: Cho 2 μl (100 ng/ μl) RNA tổng số; 1 μl (100 picomol/ μl) mồi hexamer, bổ sung nước khử ion DEPC tới thể tích 13 μl . Ủ ở 65°C trong 5 phút, lấy ra cho vào đá. Chuẩn bị tiếp các thành phần gồm: 4 μl đệm 5X AMV buffer; 0.5 μl RiboLock Ribonuclease inhibitor (40 u/ μl); 2 μl dNTP (10 mM), 0.5 μl AMV Maxima Reverse Transcriptase (20 u/ μl). Hỗn hợp lại với nhau để có 20 μl ; trộn nhẹ và ly tâm nhẹ; ủ 25°C trong 5 phút, rồi ủ tiếp ở 50°C trong 60 phút. Sản phẩm cDNA được bảo quản ở -20°C cho đến khi sử dụng, để thực hiện một loạt các phản ứng PCR thu nhận từng vùng gen tiếp nối vào nhau.

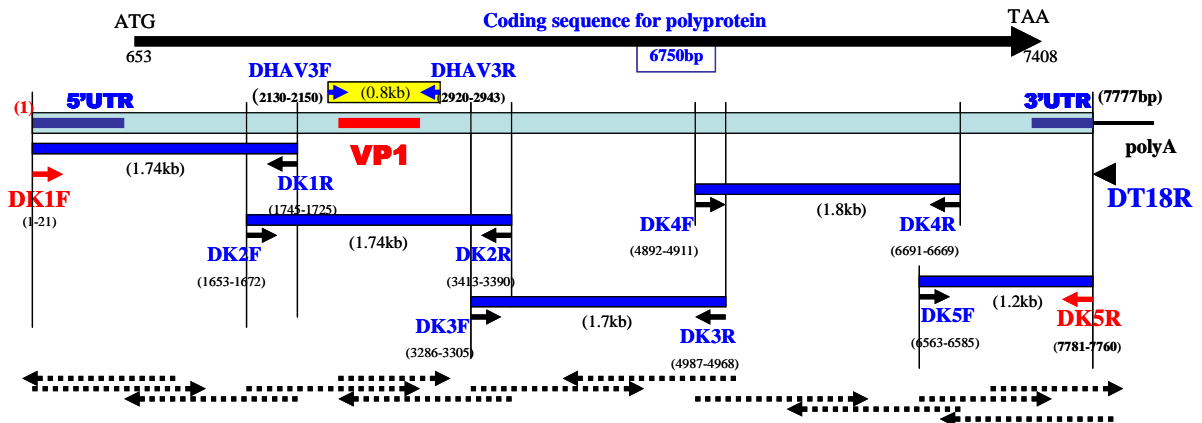
Chiến lược thiết kế mồi, thực hiện PCR và giải trình tự

Chiến lược thiết kế mồi: So sánh trình tự chuỗi nucleotide của tất cả 15 chủng DHAV-3 hiện có trong Ngân hàng gen, từ đó, tìm và chọn những vùng đồng nhất để thiết kế chính thức 12 mồi (6 mồi xuôi và 6 mồi ngược) và một số mồi phụ trợ, để thực hiện PCR và giải trình tự trực tiếp (Bảng 1). Mỗi cặp mồi được thiết kế sao cho khi thu được sản phẩm PCR của hệ gen, các đoạn DNA có trình tự lồng vào nhau ít nhất khoảng từ 80 đến 100 nucleotide ở các đầu 5' và 3'. Với các chuỗi nucleotide sẽ được sắp xếp gói vào nhau như vậy, sẽ thu nhận được chuỗi cuối cùng của toàn bộ hệ gen (Hình 1).

Chiến lược thực hiện PCR và xử lý sản phẩm PCR: Các phản ứng PCR được thực hiện với chu trình nhiệt: $94^{\circ}\text{C}/5$ phút; 35 chu kỳ [$94^{\circ}\text{C}/1$ phút, $50^{\circ}\text{C}/1$ phút, $72^{\circ}\text{C}/2$ phút] và chu kỳ cuối cùng ở $72^{\circ}\text{C}/10$ phút. Bảo quản sản phẩm ở -20°C cho đến khi sử dụng. Sản phẩm PCR được kiểm tra trên thạch agarose 1%, tinh sạch bằng bộ sinh phẩm AccuPrep™ PCR Purification kit của hãng BIONEER (Hàn Quốc), theo hướng dẫn của nhà sản xuất. Các sản phẩm PCR thu được từ các vùng gen được giải trình tự trực tiếp, trên máy Applied Biosystems (ABI) Model 3730XL.

Bảng 1. Danh sách liệt kê mỗi sử dụng cho việc thu nhận và giải trình tự toàn bộ hệ gen của virus viêm gan vịt cường độc chủng DN2 phân lập tại Việt Nam.

Tên mỗi	Trình tự nucleotide	Ký hiệu sản phẩm PCR	Kích thước (kb)
DHAV3F	5' ATGCGAGTTGGTAAGGATTTTCAG 3'	VP1	
DHAV3R	5' GATCCTGATTACCAACAACCAT 3'	Để giám định đoạn gen VP1	~0,8
DK1F	5' TTTGAAAGCGGCTGTGGTGTAG 3'		
DK1R	5'- GAGAGCAAAAGTTGGCATTTCG -3'	Đ1	~1,74
DK2F	5'- GTGGGTGATTTTCAGTGGGC -3'		
DK2R	5'- AACCAACCTCGTAAGTGAGCACG -3'	Đ2	~1,74
DK3F	5'- TGGAATCACTTGTTCTGTGCC -3'		
DK3R	5'- AGACTGCCATCCCTCATTGC -3'	Đ3	~1,7
DK4F	5'- GCGTAGGTTTCCAATCCGAC -3'		
DK4R	5'- GCTAACAGATTGTCCCCTCAGC -3'	Đ4	~1,8
DK5F	5'- TTTGACTACACAGTTGCGTTCC -3'		
DK5R	5' AGGGTGGGGAGGAATAGTAAAG 3'	Đ5 và	và
DT18R	5' TTTTTTTTTTTTTTTTTT 3'	Đ5'	~1,22



Hình 1. Sơ đồ bố trí mỗi thực hiện PCR và quá trình giải mã hệ gen virus viêm gan vịt cường độc chủng DN2 thuộc DHAV-3. Mỗi tên dài trên cùng là ORF chung kích thước 6750 bp (ATG: mã khởi đầu; TAA: mã kết thúc; chiếm vị trí 653-7408 trong chuỗi nucleotide của hệ gen); các mỗi tên ngắn biểu thị vị trí và chiều 5'→3' của các mỗi với tên mỗi (Bảng 1) cùng với vị trí trong hệ gen; các vạch giới hạn giữa 2 mỗi là sản phẩm PCR ước tính thu được; VP1 đã được xác định vị trí trong hệ gen bằng cặp mỗi DHAV3F-DHAV3R; polyA là chuỗi adenine thêm vào ở cuối hệ gen khi sao chép; mỗi tên đứt đoạn biểu thị chuỗi nucleotide sau giải trình tự cho độ dài toàn bộ hệ gen là 7777 bp. Vùng không mã hóa ở đầu 5' hệ gen là 5'UTR ở đầu 3' là 3'UTR.

Xử lý chuỗi nucleotide sản phẩm gen và toàn bộ hệ gen

Các chuỗi nucleotide sản phẩm được chỉnh sửa bằng chương trình SeqEd1.03, sắp xếp và xử lý chuỗi bằng chương trình AssemblyLIGNv1.9 và hệ chương trình MacVector 8.2 (Accelrys Inc.) trên

máy Macintosh. Xác định ORF chung (mã hóa cho polyprotein), 5'UTR (untranslation region) và 3'UTR bằng hệ chương trình MacVector 8.2 (Accelrys Inc.). Xác định các đoạn gen thành phần và điểm cắt của polyprotein bằng phương pháp so sánh, sắp xếp, đối chiếu sử dụng chương trình GENEDOC2.5 và MacVector 8.2. Thành phần

amino acid được thu nhận bằng bộ mã của vi sinh vật bậc thấp (vi khuẩn) có trong Ngân hàng gen, sử dụng chương trình GENEDOC2.5. Phân tích quan hệ phả hệ bằng chương trình MEGA4.1 (Tamura *et al.*, 2007).

Phân tích đặc điểm sắp xếp gen, thành phần nucleotide trong hệ gen và vị trí phân cắt của polypeptide

Chuỗi nucleotide hệ gen của chủng virus viêm gan vịt DN2, được truy cập vào Ngân hàng gen

(<http://www.blast.ncbi.nlm.nih.gov/blast.cgi>) bằng chương trình BLAST, đối chiếu với trình tự tương ứng của các chủng DHAV đã được công bố. Một chủng thuộc DHAV-3 (chủng AP-04114 (Hàn Quốc, DQ812093); một chủng 04G (Đài Loan, EF067923) thuộc DHAV-2 và một chủng C80 (Trung Quốc, DQ864514) thuộc DHAV-1, được chọn để phân tích đặc điểm hệ gen cùng với chủng DN2 của Việt Nam (Bảng 2). Các điểm phân cắt polypeptide của protein chung thành các protein thành phần, được xác định dựa vào phân tích của Ding, Zhang (2007).

Bảng 2. Danh sách các chủng DHAV-3 của thế giới sử dụng trong phân tích so sánh thành phần gen và mối quan hệ nguồn gốc phả hệ với chủng DN2 của Việt Nam.

TT	Ký hiệu chủng	Số đăng ký Ngân hàng gen	Nước phân lập	DNA hệ gen sử dụng trong phân tích
1	AP-03337	DQ256132	Hàn Quốc	ORF (6756 bp)
2	AP-04009	DQ256133	Hàn Quốc	ORF (6756 bp)
3	AP-04203	DQ256134	Hàn Quốc	ORF (6756 bp)
4	AP-04114	DQ812093	Hàn Quốc	ORF (6756 bp)
5	B63	EU747874	Trung Quốc	ORF (6759 bp)*
6	G	EU755009	Trung Quốc	ORF (6756 bp)
7	C-GY	EU352805	Trung Quốc	ORF (6756 bp)
8	C-FS	EU877916	Trung Quốc	ORF (6756 bp)
9	GD	GQ122332	Trung Quốc	ORF (6756 bp)
10	SD01	GQ485310	Trung Quốc	ORF (6756 bp)
11	SD02	GQ485311	Trung Quốc	ORF (6756 bp)
12	C-YDF	GU066821	Trung Quốc	ORF (6756 bp)
13	C-BLZ	GU066822	Trung Quốc	ORF (6756 bp)
14	C-YCZ	GU66823	Trung Quốc	ORF (6756 bp)
15	C-YCW	GU066824	Trung Quốc	ORF (6756 bp)

*Ghi chú: Chủng B63 (Trung Quốc, EU747874) có ORF kích thước 6759 bp, nhiều hơn các chủng khác trong DHAV-3 một bộ mã (ATA).

KẾT QUẢ VÀ THẢO LUẬN

Thu nhận sản phẩm PCR và xác định trình tự nucleotide của hệ gen chủng DN2

Để giám định trước một cách chính xác chủng DN2 thuộc type III (DHAV-3) của virus cường độc viêm gan vịt hay không, chúng tôi thiết kế cặp mồi DHAV3F – DHAV3R thu nhận đoạn gen kích thước 0,8 kb bằng RT-PCR chứa toàn bộ chuỗi gen VP1, tách dòng và giải trình tự trực tiếp. Quá trình thực hiện, phân tích đoạn DNA chứa phần gen VP1 đã

được chúng tôi công bố trong một bài báo trước đây và kết quả phân tích cho thấy, chủng DN2 chính xác là virus viêm gan vịt thuộc DHAV-3 (Đoàn Thị Thanh Hương *et al.*, 2010). Các trình tự gen VP1 của các chủng và toàn bộ hệ gen virus viêm gan vịt của Việt Nam thu được trong nghiên cứu này đang được đăng ký trong Ngân hàng gen.

Sau khi xác định là DHAV-3, chúng tôi tiến hành chuyển đổi RNA của chủng DN2 thành DNA bổ sung (cDNA) bằng phương pháp chuyển cDNA sử dụng hexamer xác suất (Fermentas Inc). Lấy 2 µl

cDNA của DN2 đã chuyển đổi làm khuôn cho mỗi phản ứng, chúng tôi tiến hành thực hiện cùng lúc 5 phản ứng PCR, với các cặp mồi lần lượt là DK1F-DK1R, DK2F-DK2R, DK3F-DK3R, DK4F-DK4R, DK5F-DK5R/và DT18R, cho các sản phẩm PCR là: Đ1 = ~1,74 kb, Đ2 = ~1,74 kb, Đ3 = ~1,7 kb, Đ4 = ~1,8 kb, Đ5 = ~1,2/và Đ5' = ~1,22 kb (Bảng 1, Hình 1). Sản phẩm PCR thu được có chất lượng rất tốt, do vậy, được tinh sạch bằng bộ sinh phẩm AccuPrep™ PCR Purification Kit (Bioneer, Hàn Quốc) và giải trình tự trực tiếp. Sau khi phân tích và ghép lồng các chuỗi nucleotide của các phân đoạn Đ1, Đ2, Đ3, Đ4, Đ5/Đ5' với nhau, chúng tôi thu được 7777 bp là tổng độ dài của toàn bộ hệ gen của virus viêm gan vịt chủng DN2 và một dãy polyA ở đầu 3' (đang đăng ký Ngân hàng gen).

Đặc điểm hệ gen của virus viêm gan vịt cường độc chủng DN2

Trật tự sắp xếp hệ gen

Chuỗi nucleotide (7777 bp) của toàn bộ hệ gen của virus viêm gan vịt chủng DN2 của Việt Nam được truy cập vào Ngân hàng gen sử dụng chương trình BLAST. Kết quả cho thấy, chủng DN2 thuộc về genotype 3 (DHAV-3), với hệ số tương đồng đạt 98 - 100% với 15 chủng hiện có đã đăng ký trong Ngân hàng gen (Bảng 2). Chúng tôi chọn chủng cường độc AP-04114 (Hàn Quốc) cùng genotype III với DN2, chủng 04G (genotype II) và chủng vaccine của Trung Quốc C80 (genotype III) để so sánh. Kết quả được thống kê ở bảng 3.

Bảng 3. Thống kê đặc điểm hệ gen của chủng virus cường độc DN2 của Việt Nam, so sánh với chủng AP-04114 (DHAV-3), chủng 04G (DHAV-2) và chủng C80 (DHAV-3) của thế giới.

Tổ hợp gen	Gen	DHAV-3						DHAV-2			DHAV-1			
		DN2			AP-04114			04G			C80			
		Vị trí trong hệ gen	Kích thước		Vị trí cắt	Kích thước		Vị trí cắt	Kích thước		Vị trí cắt	Kích thước		Vị trí cắt
			bp	aa		bp	aa		bp	Aa		bp	aa	
	5'UTR	1 – 652	652	-	652	-	652	-	656	-	-	626	-	-
	L	653 – 742	90	30	L/G	90	30	L/G	90	30	L/G	90	30	L/G
P1	VPO	743 – 1420	678	226	Q/G	678	226	Q/G	678	226	Q/G	678	226	Q/G
	VP3	1421 - 2131	711	237	Q/G	711	237	Q/G	711	237	Q/G	711	237	Q/G
	VP1	2132 - 2851	720	240	E/L	720	240	E/S	714	238	E/S	714	238	E/S
P2	2A1	2852 - 2911	60	20	NPG/P	60	20	NPG/P	60	20	NPG/P	60	20	NPG/P
	2A2	2912 - 3766	855	285	Q/S	855	285	Q/S	855	285	Q/S	855	285	Q/S
	2B	3767 - 4123	357	119	Q/S	357	119	Q/S	357	119	Q/S	357	119	Q/S
	2C	4124 - 5122	999	333	Q/G	999	333	Q/G	999	333	Q/G	999	333	Q/S
P3	3A	5122 - 5401	279	93	Q/S	279	93	Q/S	279	93	Q/S	279	93	Q/S
	3B	5402 - 5503	84	34	Q/S	84	34	Q/S	84	28	Q/S	84	28	E/T
	3C	5504 – 6046	561	181	Q/G	561	181	Q/G	561	187	Q/G	561	187	Q/G
	3D	6047 - 7408	1362	453	-	1362	453	-	1362	453	-	1362	453	-
	3'UTR	7409 - 7777	369	-	-	367	-	-	366	-	-	315	-	-
Tổng cộng			7777	2251		7775	2251		7772	2249		7691	2249	

Ghi chú: Vị trí cắt giữa các phân đoạn polypeptide thành phần của protein chung được xác định dựa trên phương pháp phân tích chủng C80 của Ding, Zhang (2007). Khung gen chứa VP1 của các chủng được bôi đậm để nhận biết.

Hệ gen của virus viêm gan vịt chủng DN2 cũng giống như tất cả các virus viêm gan vịt khác đã được công bố, có cấu trúc đặc trưng của họ *Picomaviridae*,

gồm một khung đọc mở (ORF) duy nhất mã hóa cho một protein chung (polyprotein) gồm 12 protein thành phần và thêm ở hai đầu tận cùng 5' và 3' của hệ gen là hai

vùng không mã hóa (untranslated region = UTR).

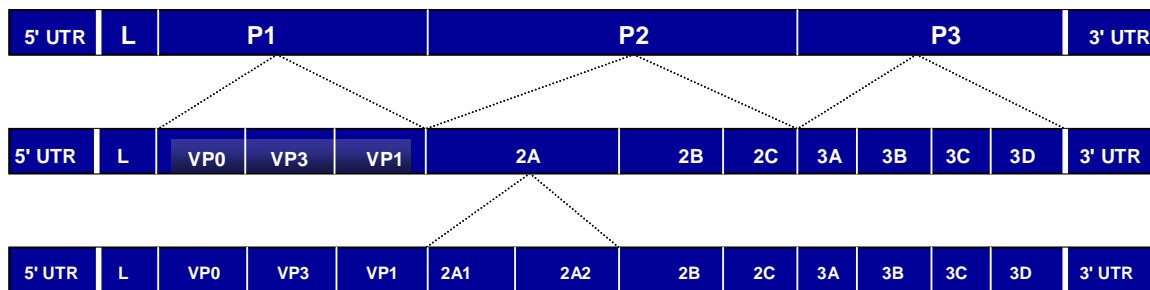
Khung đọc mở có kích thước 6756 bp, được xác định từ vị trí nucleotide thứ 653 đến 7408 trong chuỗi nucleotide của hệ gen, mã hóa cho một protein chung chứa 2251 amino acid. Vùng không mã hóa ở đầu cuối 5' của hệ gen (5'UTR) có kích thước 652 bp; và vùng không mã hóa ở đầu cuối 3' của hệ gen (3'UTR) có kích thước 369 bp (Bảng 3). Sau vùng 3'UTR, ở đầu tận cùng 3', hệ gen được adenyl hóa (polyA).

Dựa vào trật tự của các gen trong hệ gen của các chủng virus viêm gan vẹt đã được công bố, trật tự sắp xếp

của các gen trong hệ gen của chủng DN2 đã xác định và cơ bản cũng giống như ở các virus khác trong họ *Picornaviridae* (Bảng 3; Hình 2).

Trật tự hệ gen được khái quát như sau: Theo sự phân bố, hệ gen DHAV-3 chủng DN2 được sắp xếp gồm 5'UTR; L; tổ hợp P1 (VP0; VP3; VP1); tổ hợp P2 (2A: 2A1 và 2A2; 2B; 2C); tổ hợp P3 (3A; 3B; 3C; 3D); và 3'UTR. Protein chung do chuỗi nucleotide gồm 6756 bp mã hóa, sau vài lần phân cắt, tạo nên 12 protein thành phần tham gia hoạt động sống của virus (Hình 3).

SẮP XẾP HỆ GEN DHAV



Hình 3. Trật tự sắp xếp các gen trong hệ gen virus viêm gan vẹt nhóm A (DHAV) nói chung, bao gồm cả DHAV-3 chủng DN2 phân lập tại Việt Nam.

Bảng 4. Thành phần nucleotide của hệ gen và tỷ lệ (%) của mỗi loại nucleotide sử dụng ở virus viêm gan vẹt chủng DN2.

Base	Toàn bộ hệ gen		Khung đọc mở (ORF)	
	Số lượng	Tỷ lệ (%)	Số lượng	Tỷ lệ (%)
A	2144	27,6	1911	28,3
T	2292	29,5	1973	29,2
G	1724	22,2	1499	22,2
C	1617	20,7	1373	20,3
Tổng số	7777	100	6756	100

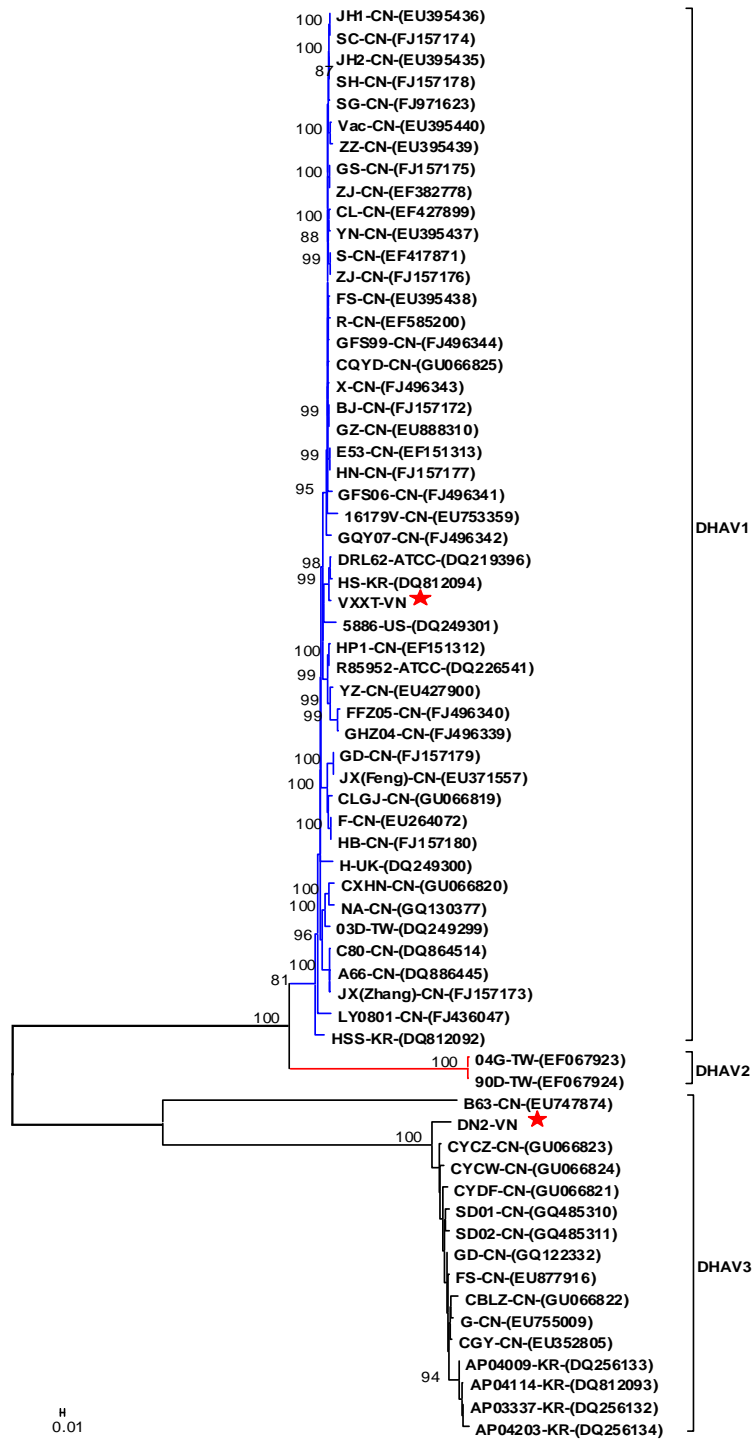
Thành phần nucleotide cấu tạo hệ gen

Thành phần nucleotide (A, T, G, C) cấu tạo nên toàn bộ chuỗi nucleotide của hệ gen (7777 bp) và khung đọc mở (6756 bp) của chủng DN2 được xác định và thống kê tại Bảng 4. Kết quả cho thấy, số lượng các base sử dụng cho toàn bộ hệ gen và cho khung đọc mở là tương đối

bằng nhau; trong đó A (27,6% và 28,3%); T (29,5% và 29,2%); G (22,2% và 22,2%); và C 20,7 và 20,3%). Như vậy, nhìn chung A+T chiếm khoảng 58% và G+C khoảng 42%, bảo đảm cân đối tỷ lệ thành phần nucleotide cấu tạo của một hệ gen (Bảng 4).

Xác định vị trí phân loại của chủng virus viêm gan vẹt cường độc DN2 của Việt Nam

Hiện nay, trên thế giới đã có 66 hệ gen của virus viêm gan vẹt (48 thuộc genotype I; 2 thuộc genotype II; 16 thuộc genotype III) được giải mã toàn bộ và phân tích đầy đủ (<http://www.picornaviridae.com/avihepatovirus/avihepatovirus.htm>), trong đó kể cả 1 chủng DHAV-1 (chủng vaccine VXXT) và 1 chủng DHAV-3 (chủng DN2) của Việt Nam (Đoàn Thị Thanh Hương *et al.*, 2010), trình bày trong bài báo này. Toàn bộ chuỗi nucleotide của khung đọc mở (6750 bp) của 48 chủng thuộc DHAV-1; 6756 bp của 2 chủng thuộc DHAV-2 và 6756 bp của 16 chủng thuộc DHAV-3 (chủng B63 là 6759 bp), được phân tích bằng chương trình GENEDOC2.5 và xác lập cây phả hệ bằng chương trình MEGA4.1.



Hình 4. Mối quan hệ phả hệ và vị trí phân loại của các chủng virus viêm gan vệt Việt Nam và thế giới dựa trên cơ sở phân tích thành phần nucleotide của toàn bộ khung đọc mở (6750 bp và 6756 bp), bằng chương trình MEGA4.1 bằng phương pháp kết nối liền kề NJ (Neighbour-Joining), sử dụng độ tin cậy 1000 bootstrap. Dấu ngôi sao chỉ dẫn vị trí của virus chủng độc viêm gan vệt chủng DN2 chúng tôi phân lập và chủng vaccine VXXT đang sử dụng tại Việt Nam.

Kết quả phân tích phả hệ trình bày ở Hình 4 cho thấy, các chủng thuộc về 3 genotype tập hợp riêng biệt thành 3 nhóm chính. Phần lớn các chủng viêm gan vịt thuộc về genotype I (48/66 chủng so sánh), chủ yếu là các chủng của Trung Quốc, Mỹ, Anh, Hàn Quốc, Đài Loan và chủng virus vaccine VXXT đang sử dụng tại Việt Nam cũng nằm trong nhóm genotype I (DHAV-1) (Hình 4). Virus viêm gan vịt genotype II (DHAV-2) chỉ bao gồm hai chủng virus mới phân lập ở Đài Loan (chủng 04G và 90D), tạo nên một nhánh hoàn toàn riêng biệt. Virus cường độc viêm gan vịt chủng DN2 của Việt Nam thuộc về genotype III (DHAV-3) cùng với các chủng chỉ mới được phân lập ở Hàn Quốc và Trung Quốc.

Như vậy, phân tích trình tự xác lập cây phả hệ xây dựng từ trình tự nucleotide của toàn bộ khung đọc mở của hệ gen cho thấy sự tồn tại của ba genotype riêng biệt trong quần thể virus viêm gan vịt trên thế giới. Vị trí phân loại của chủng virus viêm gan vịt cường độc DN2 của Việt Nam được xác định chính xác là: Group IV: ssRNA positive-strand viruses; Order: Nidovirales; Family: Picornaviridae; Genus: Avihepatovirus; (Genotype: DHAV-3); Strain: DN2.

KẾT LUẬN

Virus viêm gan vịt cường độc chủng DN2 phân lập tại Việt Nam thuộc genotype III (DHAV-3), với toàn bộ hệ gen có kích thước là 7777 bp, chứa một khung đọc mở gồm 6756 bp, mã hóa cho một protein chung gồm 12 protein thành phần, có cấu trúc đặc trưng của họ *Picornaviridae*, có sự sắp xếp gen liên tục là: 5'UTR/L/VP0/VP3/VP1/2A1/2A2/2B/2C/3A/3B/3C/3D/3'UTR. DHAV-3 là genotype virus viêm gan vịt lần đầu tiên phát hiện tại Việt Nam; lần đầu tiên được giải mã và phân tích toàn bộ hệ gen để xác định vị trí phân loại, thuộc: Họ (family): **Picornaviridae**, Chi (genus): **Avihepatovirus**; Genotype: **DHAV-3**; Chủng (strain): **DN2**.

Lời cảm ơn: Chúng tôi cảm ơn sự hỗ trợ kinh phí của đề tài nhánh do PGS. TS. Lê Thanh Hòa chủ nhiệm, thuộc chương trình Công nghệ sinh học Nông nghiệp “Ứng dụng kỹ thuật sinh học phân tử để xây dựng danh mục giống virus gia cầm quốc gia” do Trung tâm Kiểm nghiệm Thuốc thú y Trung ương chủ trì (2010 - 2011). Cảm ơn Phòng Thí nghiệm trọng điểm Công nghệ gen (Viện Công nghệ sinh học) bảo đảm trang thiết bị để thực hiện công trình này.

TÀI LIỆU THAM KHẢO

- Ding C, Zhang D (2007) Molecular analysis of duck hepatitis virus type 1. *Virology* 361: 9-17.
- Đoàn Thị Thanh Hương, Nguyễn Bá Hiên, Lê Thanh Hòa (2009) Nghiên cứu đặc tính phân tử virus nhược độc vaccine viêm gan vịt tại Việt Nam qua khảo sát chuỗi gen kháng nguyên VP1. *Tạp chí Công nghệ Sinh học* 7(1): 19-26.
- Đoàn Thị Thanh Hương, Nguyễn Bá Hiên, Lê Thanh Hòa (2010) Phát hiện lần đầu tiên genotype III (DHAV-3) virus gây bệnh viêm gan vịt tại Việt Nam bằng phương pháp giám định phân tử. *Tạp chí Công nghệ Sinh học* 8(2): 145-152.
- Kim MC, Kim MJ, Kwon YK, Lindberg AM, Joh SJ, Kwon HM, Lee YJ, Kwon JH (2009) Development of duck hepatitis A virus type 3 vaccine and its use to protect ducklings against infections. *Vaccine* 27: 6688-6694.
- Kim MC, Kwon YK, Joh SJ, Kim SJ, Tolf C, Kim JH, Sung HW, Lindberg AM and Kwon JH (2007) Recent Korean isolates of duck hepatitis virus reveal the presence of a new genome and serotype when compared to duck hepatitis virus type 1 type strains. *Arch Virol* 152(11): 2059-2072.
- Kim MC, Kwon YK, Joh SJ, Lindberg AM, Kwon JH, Kim JH, Kim SJ (2006) Molecular analysis of duck hepatitis virus type 1 reveals a novel lineage close to the genus *Parechovirus* in the family *Picornaviridae*. *J Gen Virol* 87: 3307-3316.
- Lê Thanh Hòa, Nguyễn Như Thanh, Nguyễn Bá Hiên (1986) Tính thích ứng ổn định và bệnh tích trên phôi gà của virus vacxin viêm gan vịt chủng TN. *Tạp chí Khoa học và Kỹ thuật Nông nghiệp* 284: 87-89.
- Levine PP, Hofstad MS (1945) Duck disease investigation. Annual Report of the New York State Veterinary College, Ithaca: 55-56.
- Nguyễn Đức Lưu, Vũ Như Quán (2002) Bệnh viêm gan virus vịt. *Khoa học và kỹ thuật Thú y* 9(1): 87-90.
- OIE (2008) Chapter 2.3.8. Duck virus hepatitis. *Terrestrial Manual 2008* (http://www.oie.int/eng/normes/mmanual/2008/pdf/2.03.08_DVH.pdf).
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.
- Tseng CH, Knowles NJ, Tsai HJ (2007) Molecular analysis of duck hepatitis virus type 1 indicates that it should be assigned to a new genus. *Virus Res* 123: 190-203.
- Tseng CH, Tsai HJ (2007a) Sequence analysis of a duck picornavirus isolate indicates that it together with porcine enterovirus type 8 and simian picornavirus type 2 should be assigned to a new picornavirus genus. *Virus Res* 129(2): 104-114.
- Tseng CH, Tsai HJ (2007b) Molecular characterization of a

new serotype of duck hepatitis virus. *Virus Res* 126: 19-31.
Wang L, Pan M, Fu Y, Zhang D (2008) Classification of duck hepatitis virus into three genotypes based on molecular evolutionary analysis. *Virus Genes* 37(1): 52-59.

Woolcock PR (2003) Duck hepatitis. In: Saif YM, Barnes HJ, Glisson JR, Fadly AM, McDougald LR, Swayne DE (Eds.). *Disease of Poultry*, 11th ed. *Iowa State University Press, Ames, IA*: 343-354.

COMPLETE GENOME SEQUENCE AND TAXONOMIC CLARIFICATION OF THE DN2 STRAIN OF DUCK HEPATITIS A VIRUS GENOTYPE III (DHAV-3) FIRST TIME ISOLATED IN VIETNAM

Doan Thi Thanh Huong¹, Nguyen Ba Hien², Tran Xuan Hanh³, Le Thanh Hoa^{1,*}

¹*Institute of Biotechnology, Vietnam Academy of Science and Technology*

²*Hanoi Agriculture University*

³*National Veterinary Company (NAVETCO), Ho Chi Minh City*

SUMMARY

The entire nucleotide sequence of 7777 bp for the complete genome of a virulent strain of duck hepatitis A virus (DHAV) first time isolated in Vietnam, designated as DN2, was obtained and taxonomically identified to genotype III (DHAV-3). The complete genomic sequence was obtained by segmented PCR amplification, using 5 overlapping anchor-primer pairs, with cDNA template converted from the RNA genome by random hexamer and by direct sequencing. The deduced gene order consists of 3 major parts, subsequently, a non-coding region of 652 bp at 5' (5'UTR) as first region, followed up by an open reading frame (ORF) encoding for a polyprotein; and last, a non-coding region of 369 bp at 3' (3'UTR), ending with a poly-A sequence. ORF of the genome comprises of 6756 bp coding 2251 amino acids, including 12 compositional polypeptides such as L; VP0, VP3, VP1 (P1 complex); 2A1, 2A2, 2B, 2C (P2 complex); and 3A, 3B, 3C, 3D (P3 complex). Cleavage of 12 compositional proteins occurred at sites of amino acids L/G, Q/G, Q/S, E/L, NPG/P as determined according to Ding, Zhang (2007). Compared with genome of the AP-04114 strain (Korea, GenBank: DQ812093) of DHAV-3 genotype, it revealed that gene order and size of each compositional genes in the open reading frame in the DN2 genome are the similar, except 2 nucleotides missing in the 3'UTR. Compared to a strain of DHAV-1 (C80, GenBank: DQ864514) and DHAV-2 (strain 04G, GenBank: EF067923), it revealed that the size of all DNA region coding for 11 proteins is similar, with an exception of VP1 that size of this gene is of 720 bp in the strains DHAV-3 and 714 bp in the strains of DHAV-1. Non-coding region in the strains of DHAV-1 (ie. strain C80) is the shortest (626 bp at 5'UTR and 315 bp at 3'UTR).

Từ khóa: *coding-region, DHAV-1, DHAV-2, DHAV-3, duck hepatitis, genome, genomic region, non-coding region*

* Author for correspondence: Tel: 84-4-37567297; E-mail: imibtvn@gmail.com