VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM
VIỆN SINH THÁI VÀ TÀI NGUYÊN SINH VẬT

ĐỖ THỊ ROAN

GIẢI MÃ VÀ PHÂN TÍCH HỆ GEN VIRUS VIÊM GAN VỊT CƯỜNG ĐỘC CHỦNG NT PHÂN LẬP TẠI VIỆT NAM NĂM 2013

LUẬN VĂN THẠC SĨ SINH HỌC

Hà Nội – 12/2014

Số hoá bởi Trung tâm Học liệu – ĐHTN http://www.lrc.tnu.edu.vn

VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM
VIỆN SINH THÁI VÀ TÀI NGUYÊN SINH VẬT

ĐỖ THỊ ROAN

GIẢI MÃ VÀ PHÂN TÍCH HỆ GEN VIRUS VIÊM GAN VỊT CƯỜNG ĐỘC CHỦNG NT PHÂN LẬP TẠI VIỆT NAM NĂM 2013

Chuyên ngành: Sinh học thực nghiệm
Mã số : 60420114

LUẬN VĂN THẠC SĨ SINH HỌC

Người hướng dẫn: TS. ĐOÀN THỊ THANH HƯƠNG
Viện Công nghệ sinh học

Hà Nội – 12/2014

Số hoá bởi Trung tâm Học liệu – ĐHTN http://www.lrc.tnu.edu.vn

MỞ ĐẦU

Bệnh viêm gan vịt do virus (Duck Hepatitis Virus) được coi là một bệnh cổ điển của ngành chăn nuôi gia cầm trên thế giới cũng như tại Việt Nam. Bệnh xảy ra ở vịt con từ 1 đến 6 tuần tuổi, và đặc biệt mẫn cảm với vịt con dưới 1 tuần tuổi, tốc độ lây lan nhanh với tỷ lệ chết rất cao, có khi lên đến 95 – 100% toàn đàn. Bệnh gây thiệt hại rất lớn về kinh tế, là mối lo ngại cho các nhà chăn nuôi.

Trên thế giới, mặc dù được phát hiện từ trước những năm 1950 nhưng chỉ trong vòng 10 năm gần đây virus viêm gan vịt được nghiên cứu sâu sắc ở mức độ sinh học phân tử 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*. Theo quan điểm phân loại hiện nay, virus viêm gan vịt gồm ba genotype khác nhau: genotype I (DHAV-1), genotype II (DHAV-2) và genotype III (DHAV-3).

Tại Việt Nam, từ trước đến nay mới chỉ công bố sự phổ biến của virus viêm gan vịt thuộc genotype I, bao gồm cả các chủng cường độc và vaccine. Tuy nhiên trong những điều tra mới đây, virus viêm gan vịt thuộc genotype III đã phát hiện ở một số địa phương của Việt Nam.

Để làm sáng rõ về đặc điểm phân tử của virus thuộc genotype 3 mới phát hiện này, chúng tôi tiến hành nghiên cứu đề tài**:** ***"Giải mã và phân tích hệ gen virus viêm gan vịt cường độc chủng Ninh Thuận (NT) phân lập tại Việt Nam năm 2013"***. Đề tài được thực hiện với những mục tiêu và nội dung sau:

1. Mục tiêu:

- Thu nhận được trình tự nucleotide của toàn bộ hệ gen của một chủng virus viêm gan vịt cường độc mới phân lập tại Việt Nam năm 2013.

- Phân tích được trật tự sắp xếp của từng vùng gen và đặc điểm phân tử của các phân đoạn gen trong hệ gen.

2. Nội dung:

- Phân lập và thu thập một chủng virus cường độc viêm gan vịt mới phân lập tại tỉnh Ninh Thuận (Việt Nam) năm 2013 (Kí hiệu chủng: NT).

- Giải mã toàn bộ hệ gen của virus viêm gan vịt chủng NT (có kích thước khoảng 7.8 kb).

- Phân tích trật tự sắp xếp của các gen theo từng vùng gen.

- So sánh trình tự nucleotide của chủng NT với các chủng virus viêm gan vịt khác của Việt Nam và thế giới.

- Xây dựng cây phả hệ và vị trí phân loại.

Chương 1

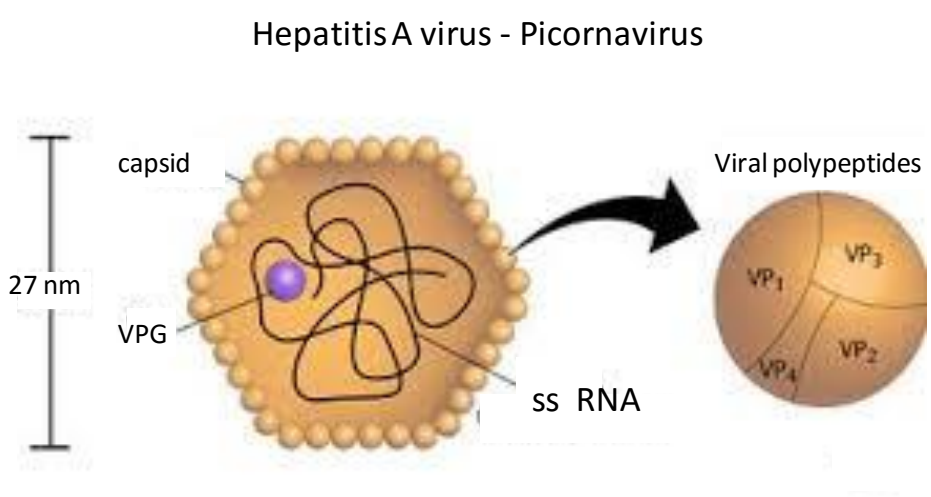
TỔNG QUAN TÀI LIỆU

1.1. ĐẶC ĐIỂM SINH HỌC CỦA VIRUS VIÊM GAN VỊT

1.1.1. Cấu tạo chung

Virus viêm gan vịt (duck hepatitis virus) thuộc họ *Picornaviridae*, có kích thước nhỏ, khoảng 20-30 nm, không có vỏ bọc ngoài cùng. Hệ gen của virus viêm gan vịt chứa acid ribonucleic (RNA) sợi đơn dương có độ dài khoảng 7600 đến 7700 nucleotide và đuôi poly A ở đầu 3', chứa duy nhất một khung đọc mở (open reading frame) mã hoá cho một protein chung (polyprotein). Protein này, sau khi tổng hợp phải trải qua nhiều lần phân cắt để tạo nên các sản phẩm protein độc lập. Hệ gen RNA của virus rất đặc biệt, có chứa một protein ở đầu 5' có tác dụng như một primer của quá trình sao chép RNA nhờ RNA polymerase (www.britannica.com/EBchecked/topic/459528/*Picornavirus*). Các virus thuộc họ *Picornaviridae* rất đa dạng (có tới trên 200 genotype) và là loại virus được biết đến từ sớm nhất. Bệnh do chúng gây ra rất đa dạng, có thể là bệnh cấp tính hay nhiễm trùng mãn tính, bao gồm cả nhiều bệnh nguy hiểm trên người và động vật.

Picornavirus rất nhỏ bé, có thể xuyên qua được màng lọc thông thường (Levine, Fabricant, 1950). Qua kính hiển vi điện tử, virus là những hạt tròn, bề mặt xù xì, không có vỏ bọc ngoài cùng (hình 1.1).

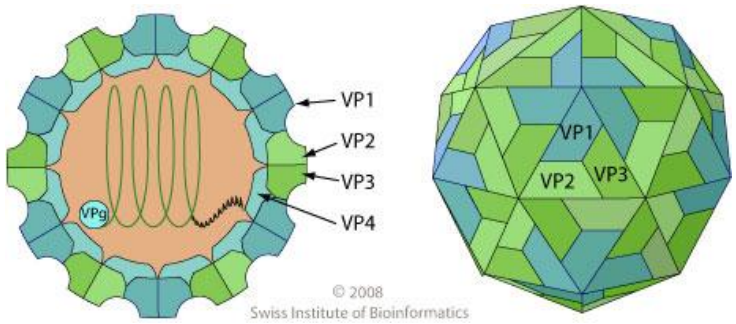


Hình 1.1. . Mô hình cấu trúc của một chủng virus thuộc họ *Picornavirus* (Hepatitis A virus). (Nguồn: www.technologyreview.com)

1.1.2. Đặc điểm cấu trúc hệ gen của virus viêm gan vịt

Trong họ *Picornaviridae*, tất cả virus đều bị thiếu cấu trúc mũ ở đầu 5' mà thay vào đó là một protein (khoảng 23 amino acid) gọi là VPg (virion protein genome-link). VPg này liên kết đồng hoá trị với đầu 5' thông qua một gốc tyrosine (Cheney *et al.*, 2003) (hình 1.2). Ở cả hai đầu của hệ gen *Picornavirus* đều có vùng không dịch mã (UTR). Vùng 5'UTR dài hơn, có độ dài khoảng từ 600 đến 1200 bp, trong khi vùng 3'UTR thường có độ dài chỉ từ 50 đến 100 bp. Vùng 5'UTR có vai trò quan trọng đối với quá trình sao mã và vùng 3'UTR có vai trò trong quá trình tổng hợp sợi âm. Tuy nhiên đầu 5' cũng có thể có vai trò trong việc làm tăng cường độc lực của virus. Trong hệ gen của *Picornavirus* có một vùng nucleotide làm điểm khởi đầu để ribosome dịch mã, gọi là IRES – Internal Ribosome Entry Site, đó là một cấu trúc di truyền hoạt động phức tạp dạng *cis* với cấu trúc bổ sung thứ cấp khởi động cho quá trình dịch mã không phụ thuộc điểm mũ của hệ gen virus (Chard *et al.*, 2006). Nằm giữa hai vùng 5' và 3' là một khung đọc mở (open frame reading) lớn, mã hóa khoảng từ 2100 đến 2400 amino acid cho cả các protein cấu trúc và không cấu trúc. Đầu cuối 3' của hệ gen virus có gắn thêm đuôi poly A, dài khoảng từ 16 đến18 nucleotide.

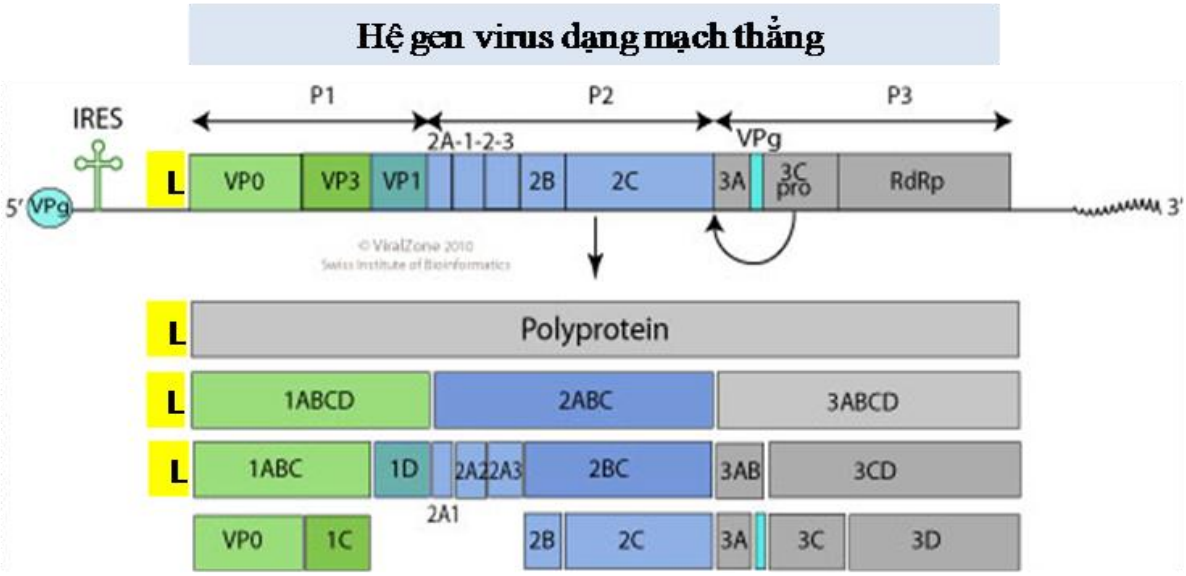
Trong chuỗi polypeptide, phần protein đầu tiên là protein dẫn, ký hiệu là L (Leader protein), tiếp theo là 4 loại protein cấu trúc bao gồm VP4 (1A), VP2 (1B), VP3 (1C) và VP1 (1D), đoạn cuối cùng gồm 7 protein không cấu trúc là protein 2A, 2B, 2C, 3A, 3B, 3C và 3D. Tuy nhiên, ở các chi virus khác nhau thì trật tự này có sự thay đổi nhỏ. Chuỗi polypeptide chung của *Picornavirus* thường bị phân giải bởi một, hai hoặc cả ba enzyme protease là $3C^{pro}$, $2A^{pro}$ và L^{pro} thành 10, 11 hoặc 12 chuỗi (Krumbholz *et al.*, 2002). Toàn bộ chuỗi các protein liên tục này gồm 2249 amino acid, được mã hóa bởi 6747 nucleotide (Ding, Zhang, 2007).



Hình 1.2. Cấu trúc không gian của *Picornavirus*. Các protein VP1, VP2 và VP3 nằm trên bề mặt vỏ kháng nguyên; protein VP4 lặn sâu vào bên trong (Nguồn: Swiss Institute of Bioinformatics, 2008).

1.1.3. Vai trò, chức năng của các gen và các protein của virus viêm gan vịt

Trong hệ gen virus, VPg – tuy là một protein nhỏ, chỉ khoảng 23 amino acid, nhưng có vai trò cực kỳ quan trọng trong quá trình sao mã. VPg là mồi cho quá trình tổng hợp sợi âm, trong khi VPgpUpUOH (một dạng chuyển hoá của VPg) là mồi cho quá trình tổng hợp sợi dương.



Hình 1.3. Cấu trúc hệ gen của một loại virus họ *Picornavirus* (Kok, McMinn, 2009).
Hệ gen chia làm ba phần chính : P1, P2, P3 mã hóa cho chuỗi polyprotein chung, sau đó tự phân cắt thành các protein thành phần.

Các protein VP1, VP2, VP3 và VP4 trong tổ hợp protein P1 có vai trò trong quá trình tạo vỏ capsid của virus. Ba loại protein đầu tiên nằm ở mặt ngoài của virus, protein VP4 ngắn hơn, nằm hoàn toàn ở bên trong vỏ capsid (Hình 1.2). Trong 4 loại protein này, VP1 có vai trò đặc biệt quan trọng, quyết định tính kháng nguyên và độc

lực của virus.

Bảy loại protein còn lại 2A, 2B, 2C, 3A, 3B, 3C, 3D là các protein không cấu trúc, tham gia vào các quá trình nhân lên của virus. Protein dẫn (L), protein 2A và 3C (L^{pro}, $2A^{pro}$ và $3C^{pro}$) là các enzyme có vai trò quan trọng trong quá trình phân cắt protein của virus (Strebel, Beck, 1986). Trong đó, L^{pro} chỉ xuất hiện trong các tế bào nhiễm *Aphtovirus* và *Cardiovirus*. Các enzyme này không chỉ phân cắt các chuỗi peptide của virus mà còn ức chế hoạt động tế bào của nhiều loại tế bào vật chủ khác nhau và gây ảnh hưởng đến chức năng nội bào. $3C^{pro}$ của *Picornavirus* có thể xâm nhập vào nhân thông qua tiền tố có mang trình tự định vị nhân (NLS) của chúng là 3CD' hay 3CD (Amineva *et al*., 2004; Sharma *et al*., 2004). $3C^{pro}$ đồng thời có thể phân cắt phần lớn các yếu tố hoặc các nhân tố điều hoà liên quan đến RNA polymerase I, II, III của tế bào. Ngoài ra, $3C^{pro}$ còn liên quan đến quá trình ức chế phiên mã của tế bào vật chủ. $2A^{pro}$ có khả năng phân cắt protein gắn có cấu trúc hộp TATA *in-vitro* nhưng lại không có khả năng ức chế quá trình phiên mã của tế bào vật chủ (Yalamachili *et al*., 1997).

Các protein khác của *Picornavirus* bao gồm 2B, 2BC, 3A và 3D liên quan đến quá trình sao chép RNA. Cụ thể :

Protein 2B/2BC

Các protein 2B hay protein 2BC được giả thiết là có khả năng biến đổi màng của các tế bào bị nhiễm . Protein 2B và tiền thân của chúng là 2BC có vai trò quan trọng trong quá trình nhân lên và gắn dính vào màng của hệ thống Golgi và phức hợp mạng lưới nội chất (endoplasmic reticulum-ER) của tế bào vật chủ, từ đó hình thành nên các cấu trúc túi hạt trên màng tế bào vật chủ bị viêm nhiễm và hình thành phức hợp viroporin (Jong *et al*., 2004). Sự tích luỹ các protein 2B và 2BC trong mạng lưới Golgi dẫn đến thay đổi tính thấm của màng sinh chất (Jong *et al*., 2008) và làm cho cấu trúc thể Golgi trở nên lỏng lẻo, từ đó có thể gây chết tế bào (Kuppeveld *et al*., 1997). Bên cạnh đó, khi protein 2B và 2BC gắn vào màng tế bào vật chủ còn gây giảm nồng độ Ca^{2+} trong phức hợp ER và Golgi, từ đó gây ức chế quá trình vận chuyển protein từ mạng lưới nội chất ER tới phức hệ Golgi (Jong *et al*., 2006).

Protein 3A

Protein 3A là một loại protein gắn màng, chúng giữ vai trò trong quá trình ức

chế tiết protein nội bào và giữ vai trò trung gian trong quá trình trình diện protein màng khi bị nhiễm virus. Khi có mặt protein A trong tế bào sẽ làm gián đoạn quá trình lưu thông từ mạng lưới nội chất tới mạng lưới Golgi, hiện tượng này cũng xuất hiện trong trường hợp tế bào bị nhiễm có protein 2B (Doedens *et al*., 1995 ; Doedens *et al*., 1997).

Protein 3AB

Theo các kết quả nghiên cứu về mặt hoá sinh, protein 3AB là một protein đa chức năng. Vùng kỵ nước trong tiểu phần 3A của protein liên kết với cấu trúc dạng túi hạt trên màng. Protein 3AB tương tác với 3D và 3CD của *Poliovirus* ở quy mô *in-vitro* (Towner *et al*., 1996; Fujita *et al*., 2007). Protein 3AB màng gắn trực tiếp với tiền thân 3CD polymerase trên cấu trúc vòng lúp của RNA *poliovirus*, từ đó kích thích hoạt tính protease của 3CD. Thêm vào đó, protein 3AB lại kích thích 3D polymerase của *Poliovirus* (Hope *et al*., 1997) và có chức năng như là một cơ chất cho 3D polymerase trong quá trình uridyl hoá VPg. Bên cạnh đó, protein 3AB của *Poliovirus* còn giữ một số vai trò khác, ví dụ như phá vỡ tính bền vững của chuỗi xoắn kép, phá vỡ tính ổn định trong cấu trúc bậc 2 của RNA và nâng cao khả năng bắt cặp của các cặp nucleotit bổ sung trong quá trình sao chép của virus (Xiang *et al*., 1995; Richard *et al*., 2006).

Protein 3B

Các protein 3B (VPg) là những peptide nhỏ có chứa từ 21 đến 23 amino acid liên kết cộng hóa trị với đầu 5’ của hệ gen *Picornavirus* thông qua liên kết 5’ tyrosyluridine trong các gốc tyrosine của VPg. VPg cũng có khả năng tương tác với 3D polymerase của *Poliovirus* và UMP để tạo thành các tổ hợp VpgpU và VpgpUpU (Paul *et al*., 1998). VPg đã uridyl hóa được sử dụng như một primer cho cả quá trình tổng hợp RNA sợi đơn âm và RNA sợi đơn dương (Pettersson *et al*., 1978).

Protein 3CD

Protein 3CD, tiền thân của protease 3C trưởng thành và polymerase 3D, có hoạt tính protease, không có hoạt tính polymerase (Harris *et al*., 1992). Protein 3CD của *Poliovirus* góp phần tích cực vào quá trình sao chép RNA của virus bằng cách tương tác với cả hai đầu 5’ và 3’ của RNA virus, làm cuộn vòng genome của virus (Parsley *et al*., 1997; Gamarnik and Andino,1997). Ngoài ra, 3CD cũng có khả năng tương tác với ribonucleoprotein C (hnRNP C) (Brunner *et al*., 2005). hnRNP C tham gia trong quá

trình tổng hợp tiền RNA thông tin ở các tế bào bình thường. hnRNP C tham gia vào quá trình sao chép RNA của virus và dạng đột biến của hnRNP C không có khả năng tương tác protein-protein sẽ ức chế quá trình tổng hợp RNA sợi đơn dương của virus (Walter *et al*., 2002 ; Brunner *et al*., 2005).

Protein 3D

RNA polymease 3D là một trong những thành phần cơ bản tham gia vào phức hợp sao chép RNA virus. 3D polymease có thể uridilyl hóa VPg và sử dụng VPg-pUpU như là một primer trong quá trình sao chép RNA.

1.1.4. Quá trình nhân lên của virus

Virus nhân lên trong tế bào chất. Đầu tiên, virus gắn vào thụ thể nằm trên màng sinh chất của tế bào chủ, sau đó theo cơ chế nhập bào và tiến hành quá trình "cởi áo" virus trong endosome hoặc lysosome.

Sau khi được giải phóng khỏi capsid, RNA hệ gen của virus được dịch mã, sản xuất các protein của virus; tiếp theo đó, các phức hợp sao chép của virus được tập hợp; RNA virus được nhân lên và hệ gen RNA mới sau tổng hợp sẽ được bao gói tạo thành các thế hệ virus con cháu (Barton, Flanegan, 1995; Molla *et al*., 1991) (Hình 1.4).

Trước tiên, RNA ngay lập tức khởi động quá trình dịch mã ở vị trí gần đầu 5' để tạo thành một polyprotein – tiền thân của các protein chức năng của virus.

Polyprotein này ban đầu được phân cắt thành 3 protein tiền chất là P1, P2 và P3. Tiếp theo đó, mỗi protein lại tiếp tục tiến hành phân cắt để tạo ra các sản phẩm protein đặc hiệu có chức năng cụ thể:

- P1 được cắt ra tạo thành VP0, VP1 và VP3. Đó là các protein capsid.
- P2 được cắt ra tạo thành các kháng nguyên 2A, 2B của vật chủ và 2C tham gia vào quá trình tổng hợp RNA.
- P3 tiến hành tự cắt thành 3A, 3B (VPg), 3C (một enzyme dùng để phân cắt hầu hết các protein) và 3D (RNA polymerase làm nhiệm vụ kéo dài chuỗi RNA mới tổng hợp trên khuôn RNA).

Để lắp ráp, trước hết virus phân cắt P1 tạo thành một đơn phân gốc (protome) 5S chưa hoàn thiện chứa VP0, VP1 và VP3. Các protome này sẽ liên kết với genome RNA để tạo thành tiền virion (provirion). Trong quá trình hoàn thiện, phần lớn hoặc tất cả các phân tử VP0 đều bị phân cắt tạo thành VP2 và VP4. Sau lắp ráp, virion trưởng