DOI: https://doi.org/10.59294/HIUJS2026003

Nghiên cứu xây dựng cơ sở dữ liệu xác định genotype HCV và bước đầu định lượng HCV bằng kỹ thuật digital PCR

Huỳnh Thị Thu Thảo[1], Hà Thị Anh[1], Nguyễn Thị Nga[1], Vũ Thị Hải Yến[1], Nguyễn Tuấn Anh[2,3,*]

[1]Trường Đại học Quốc tế Hồng Bàng

[2]Trung tâm Đào tạo và Chẩn đoán Y Sinh học Phân tử, Bệnh viện Đại học Y Dược Thành phố Hồ Chí Minh

[3]Bệnh viện Đại học Y Dược Thành phố Hồ Chí Minh - Cơ sở 2

TÓM TẮT

Đặt vấn đề: Virus viêm gan C (Hepatitis C virus, HCV) có tính đa dạng di truyền cao, gây nhiều thách thức trong xác định và định lượng giữa các genotype/subtype. Mục tiêu: Xây dựng bộ dữ liệu hệ gen HCV nhằm xác định genotype/ subtype; đồng thời bước đầu phát triển xét nghiệm digital PCR (dPCR) nhằm phát hiện/định lượng HCV RNA. Đối tượng và phương pháp: Trình tự hệ gen HCV được thu thập và chọn lọc nhằm xây dựng bộ dữ liệu tham chiếu xác định genotype/subtype. Phân tích in silico được thực hiện để lựa chọn bộ mồi và mẫu dò phù hợp cho phản ứng PCR. Trên cơ sở này, xét nghiệm dPCR được thử nghiệm trên mẫu huyết thanh lâm sàng và được đối chiếu với real-time PCR. Kết quả: Bộ dữ liệu tham chiếu chuẩn đại diện cho 86 subtype thuộc 8 genotype HCV đã được thu thập, hỗ trợ phân loại genotype/subtype dựa trên bằng chứng tiến hóa. Xét nghiệm dPCR định lượng thành công tải lượng virus ở các mẫu dương tính và đồng thuận với real-time PCR. Kết luận: Nghiên cứu cung cấp dữ liệu tham chiếu và quy trình kỹ thuật ban đầu cho các ứng dụng lâm sàng trong tương lai, đồng thời làm cơ sở cho các nghiên cứu sâu về genotype và chẩn đoán phân tử HCV bằng dPCR.

Từ khóa: *Virus viêm gan C, khảo sát in silico, định lượng, digital PCR*

1. ĐẶT VẤN ĐỀ

Virus viêm gan C hiện đang là một thách thức y tế công cộng toàn cầu, với khoảng 50 triệu người nhiễm mạn tính [1, 2]. HCV là căn nguyên quan trọng gây bệnh gan tiến triển, bao gồm xơ gan và ung thư biểu mô tế bào gan, ước tính khoảng 242,000 trường hợp tử vong trong năm 2022 (WHO, 2025). Đến nay, vẫn chưa có vắc-xin hiệu quả để phòng ngừa HCV; do đó, các chiến lược phòng ngừa, chẩn đoán sớm và giám sát dịch tễ - phân tử giữ vai trò then chốt trong kiểm soát bệnh. Các thuốc kháng virus tác động trực tiếp (DAAs) với hiệu quả rất cao (tỷ lệ đạt đáp ứng virus bền vững trên 95%), do đó, việc xác định nhiễm HCV là hết sức quan trọng, nhằm điều trị kịp thời và giảm gánh nặng bệnh tật [1, 2]. Việt Nam là quốc gia này nằm trong nhóm 20 nước có tỷ lệ hiện mắc HCV cao nhất thế giới, với khoảng 1 - 1.5 triệu người nhiễm mạn tính, tương ứng khoảng 1% dân số trưởng thành [2]. Tuy nhiên, phân bố nhiễm HCV rất không đồng nhất giữa các nhóm quần thể. Ở cộng đồng chung, các nghiên cứu gần đây ghi nhận mức nhiễm tương đối thấp (khoảng 0.3%), trong khi một số nhóm nguy cơ cao có tỷ lệ nhiễm tăng mạnh, đặc biệt ở người tiêm chích ma túy, với ước tính dao động từ 58% đến 74%, cũng như ở những người từng truyền máu hoặc đang chạy thận nhân tạo [1, 2]. Trước các hậu quả lâu dài và nghiêm trọng của nhiễm HCV mạn tính, Việt Nam cần có các kế hoạch tăng cường giám sát phân tử và triển khai các chiến lược can thiệp có mục tiêu [3].

Dữ liệu về genotype HCV tại Việt Nam đặc biệt quan trọng, do tỷ lệ genotype lưu hành ở nước ta khác biệt so với xu hướng toàn cầu. Trong khi genotype 1 là kiểu gen phổ biến nhất trên thế giới [4], Việt Nam là một trong số ít quốc gia mà genotype 6 chiếm ưu thế [2]. Các khảo sát trên quần thể nhiễm HCV tại Việt Nam, cho thấy rằng genotype 1 và 6 chiếm khoảng 90 - 95% số trường hợp nhiễm [5]. Đáng chú ý, genotype 6, đặc biệt subtype 6a, lưu hành phổ biến và được ghi nhận chiếm khoảng một nửa tổng số ca nhiễm HCV [6]. Genotype 6 cũng là kiểu gen có mức đa dạng di truyền cao nhất, bao gồm nhiều subtype, từ đó làm gia tăng thách thức cho chẩn đoán và điều trị (khó khăn trong việc phân biệt

Tác giả liên hệ: Nguyễn Tuấn Anh
Email: anh.nt@umc.edu.vn

các biến thể thuộc genotype 6 với genotype 1) [2]. Việc xác định genotype không chỉ hỗ trợ lựa chọn chiến lược điều trị phù hợp, mà còn phát hiện sự xâm nhập của các genotype mới trong quần thể [2].

Hiện nay, phần lớn hoạt động giám sát và nghiên cứu HCV tại Việt Nam vẫn dựa trên các kỹ thuật PCR truyền thống và real-time PCR để phát hiện/ định lượng virus [7]. Tuy nhiên, bên cạnh những đóng góp quan trọng của các phương pháp PCR hiện hành, một hướng phát triển mới trong chẩn đoán HCV là ứng dụng kỹ thuật dPCR nhằm tăng độ nhạy và nâng cao độ chính xác trong định lượng virus. dPCR là công nghệ PCR thế hệ thứ ba, cho phép định lượng tuyệt đối acid nucleic thông qua việc phân chia mẫu thành nhiều vi phân đoạn, khuếch đại ở điểm cuối và suy luận nồng độ theo mô hình thống kê Poisson [8]. Ưu điểm nổi bật của dPCR là khả năng phát hiện tín hiệu ở mức rất thấp, đặc biệt hữu ích trong các tình huống như giai đoạn nhiễm sớm khi tải lượng virus còn thấp hoặc khi cần đánh giá các điểm cuối điều trị.

Nghiên cứu này được thực hiện với hai nội dung chính. Thứ nhất, lựa chọn và tổng hợp các trình tự hệ gen HCV nhằm xây dựng cơ sở dữ liệu tham chiếu hỗ trợ việc xác định genotype và subtype HCV. Thứ hai, chúng tôi tiến hành thử nghiệm kỹ thuật dPCR phát hiện/ định lượng HCV, qua đó đặt nền tảng thực nghiệm cho việc ứng dụng công nghệ này trong nghiên cứu và chẩn đoán HCV. Quy trình bao gồm lựa chọn bộ mồi và mẫu dò bằng các phân tích *in silico*, sau đó là thử nghiệm nhằm chứng minh khả năng phát hiện và định lượng HCV bằng kỹ thuật dPCR.

2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Thu thập trình tự

Các trình tự nucleotide HCV đại diện được thu thập từ các cơ sở dữ liệu công khai (GenBank và các cơ sở dữ liệu chuyên biệt về HCV) dựa trên thông tin chú giải genotype và subtype. Trình tự được lựa chọn theo các tiêu chí như sau: (i) genotype và subtype được báo cáo rõ ràng; (ii) hệ gen virus hoàn chỉnh (≥ 9 kb); và (iii) các thông tin về hệ gen đồng thuận giữa bản ghi trong cơ sở dữ liệu và bài báo liên quan. Các trình tự bị loại nếu có bằng chứng tái tổ hợp, chứa tỷ lệ nucleotide không xác định cao (N > 5% tổng chiều dài), có dữ liệu về hệ gen không đầy đủ, không nhất quán, hoặc thiếu cơ sở thực nghiệm.

2.2. Phân tích phát sinh loài

Tất cả các trình tự sau khi sàng lọc theo tiêu chí được tải về ở định dạng FASTA và căn chỉnh bằng công cụ MAFFT với thiết lập mặc định cho hệ gen virus. Bộ dữ liệu đã căn chỉnh sau đó được kiểm tra và hiệu chỉnh thủ công nhằm loại bỏ các sai lệch trong căn chỉnh, các đoạn chèn/xóa lớn (indel) và các vùng căn chỉnh kém tin cậy. Cây phát sinh loài được xây dựng theo phương pháp maximum likelihood (ML). Mô hình tiến hóa tối ưu được lựa chọn bằng phần mềm ModelFinder. Phân tích ML được thực hiện bằng phần mềm IQ-TREE với 1,000 lần lặp bootstrap để đánh giá độ tin cậy của các nút phân nhánh. Cây phát sinh loài được trực quan hóa bằng phần mềm iTOL v7.

2.3. Khảo sát *in silico* bộ mồi và mẫu dò nhằm phát hiện/ định lượng HCV

Các thông số của mồi và mẫu dò, bao gồm nhiệt độ nóng chảy (T_m), hàm lượng GC và khả năng hình thành cấu trúc thứ cấp (hairpin, self-dimer và hetero-dimer), được đánh giá bằng công cụ IDT OligoAnalyzer (https://sg.idtdna.com/pages/tools/oligoanalyzer). Để khảo sát độ đặc hiệu và mức độ bao phủ các genotype, các trình tự hệ gen HCV đại diện cho các genotype được thu thập từ cơ sở dữ liệu GenBank và căn chỉnh bằng phần mềm MAFFT v7. Vị trí bắt cặp của mồi và mẫu dò được căn chỉnh cùng với các trình tự genotype HCV nhằm đánh giá mức độ tương đồng và xác định các sai khác nucleotide. Độ đặc hiệu tiếp tục được khảo sát bằng công cụ NCBI Primer-BLAST đối chiếu với cơ sở dữ liệu nucleotide non-redundant (nr) (loại trừ các trình tự HCV) nhằm đánh giá nguy cơ khuếch đại ngoài mục tiêu, bao gồm hệ gen người và các hệ gene virus khác.

2.4. Y đức, thu thập mẫu, tách chiết acid nucleic tổng và tổng hợp cDNA

Nghiên cứu được Trường Đại học Quốc tế Hồng Bàng phê duyệt về mặt đạo đức dựa trên kết luận của Hội đồng Đạo đức trong nghiên cứu Y sinh, theo Quyết định số 46/PCT-HĐĐĐ-ĐT. Mẫu máu toàn phần được thu từ các đối tượng đến sử dụng dịch vụ chẩn đoán tại Trung tâm Sinh học Phân tử, Bệnh viện Đại học Y Dược Thành phố Hồ Chí Minh - Cơ sở 2, trong giai đoạn tháng 4 đến tháng 8 năm 2023. Các mẫu được phân thành hai nhóm: HCV-dương tính và HCV-âm tính, dựa trên kết quả xét nghiệm real-time PCR tại Trung tâm Sinh học Phân tử, Bệnh viện Đại học Y Dược Thành phố Hồ Chí Minh - Cơ sở 2. Mẫu huyết thanh được tách từ mẫu máu toàn

phần, sau đó, được tách chiết acid nucleic tổng bằng bộ sinh phẩm AccuRive Viral sDNA/RNA Prep Kit (EX-DNA02.1F, Khoa Thuong BioTech Company, Việt Nam), theo hướng dẫn của nhà sản xuất. Sản phẩm tách chiết được bảo quản ở -80°C kèm chất ức chế RNase để duy trì tính toàn vẹn RNA cho đến khi sử dụng. Khoảng 1 µg acid nucleic tổng (bao gồm DNA/RNA) được sử dụng làm khuôn mẫu cho phản ứng RT-PCR bằng bộ sinh phẩm SensiFAST cDNA Synthesis Kit (#BIO-65054, Bioline). Mẫu cDNA được bảo quản ở -20°C để phục vụ các thí nghiệm tiếp theo.

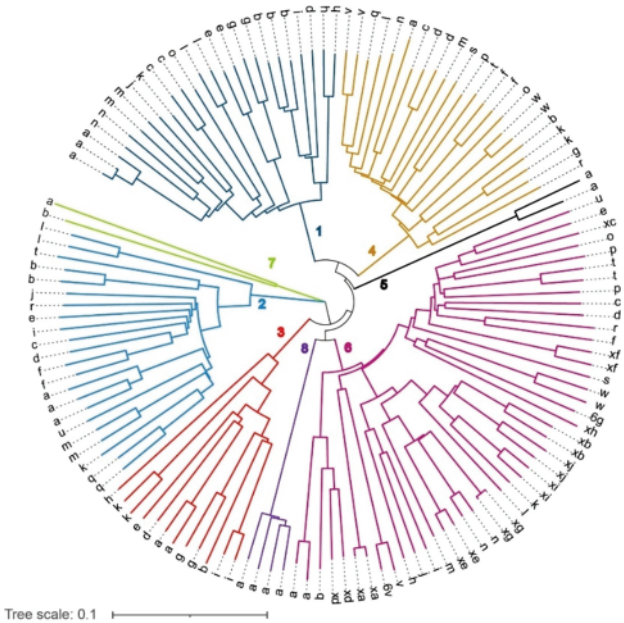
2.5. Thử nghiệm kỹ thuật dPCR nhằm phát hiện/ định lượng HCV trên mẫu huyết thanh lâm sàng

Kỹ thuật dPCR được thử nghiệm với bộ mồi và mẫu dò (được khảo sát trong nghiên cứu) trên 10 mẫu huyết thanh lâm sàng nhằm định lượng tải lượng HCV. Phản ứng dPCR được thực hiện bằng bộ sinh phẩm Absolute Q DNA Digital PCR Master Mix (#A52490, Thermo Fisher Scientific) trên máy QuantStudio Absolute Q Digital PCR System (Thermo Fisher Scientific, Hoa Kỳ). Thành phần phản ứng dPCR bao gồm: Master mix 1×, mồi xuôi 500 nM, mồi ngược 500 nM, mẫu dò 150 nM và 100 ng mẫu cDNA. Chương trình nhiệt được thiết lập như sau: 95°C trong 10 phút, sau đó 40 chu kỳ gồm 96°C trong 5 giây, 60°C trong 5 giây và 72°C trong 20 giây. Mẫu chứng âm (không chứa DNA/cDNA) được đưa vào mỗi lần chạy để kiểm soát nguy cơ nhiễm. Kết quả định lượng được phân tích bằng phần mềm QuantStudio Absolute Q Analysis Software (Thermo Fisher Scientific, Hoa Kỳ), với khoảng tin cậy 95% dựa trên mô hình phân phối Poisson.

3. KẾT QUẢ

3.1. Bộ dữ liệu tham chiếu phân loại genotype và subtype HCV

Bộ dữ liệu tham chiếu được lựa chọn, tổng hợp và nhằm bảo đảm bao phủ tối đa các genotype và subtype của HCV. Bộ dữ liệu bao gồm các trình tự thuộc genotype 1 - 8 và 86 subtype đã được công bố (Hình 1, Bảng 1). Đối với mỗi subtype, tối đa ba trình tự được lựa chọn, ưu tiên các trình tự hệ gen hoàn chỉnh (kích thước hệ gen lớn hơn 9 kb). Genotype 1 được đại diện bởi các subtype 1a-1o; genotype 2 bao gồm các subtype 2a-2u; và genotype 3 bao gồm các subtype 3a-3k. Mức độ bao phủ tương tự cũng đạt được đối với genotype 4 (4a-4w), genotype 5 (5a) và genotype 6 có tính đa dạng cao (6a-6xf). Ngoài ra, các genotype được mô tả gần đây như genotype 7 (7a, 7b) và genotype 8 (8a) cũng được đưa vào bộ dữ liệu.



Hình 1. Cây phát sinh loài của các genotype và subtype HCV dựa trên hệ gen hoàn chỉnh (> 9.0 kb). Bộ dữ liệu bao gồm các trình tự đại diện của tám genotype HCV (1 - 8), với các subtype được ký hiệu theo chữ cái (a, b, c, ...). Trên cây phát sinh loài, từng nhánh mang các màu sắc khác nhau tương ứng với từng genotype, bao gồm: 1 (xanh đậm), 2 (xanh nhạt), 3 (đỏ), 4 (cam), 5 (đen), 6 (hồng), 7 (xanh lá) và 8 (tím)

Cây phát sinh loài cho thấy các trình tự hệ gen HCV được phân loại rõ ràng thành tám nhóm genotype riêng biệt (1 - 8). Cấu trúc phân nhánh thu được phù hợp với khung phân loại genotype đã được thiết lập trong các nghiên cứu trước đây. Mỗi genotype hình thành một nhóm đơn ngành (monophyletic), phản ánh rằng các subtype trong cùng genotype có chung nguồn gốc và tách biệt với các genotype khác, qua đó cho thấy quá trình tiến hóa của HCV đã dẫn đến sự hình thành các genotype và subtype khác nhau. Đáng chú ý, genotype 1, 2 và 6 thể hiện mức đa dạng subtype cao, phù hợp với các báo cáo trước đây về tính đa dạng của các genotype này. Ngược lại, các genotype 5, 7 và 8 tạo thành các cụm nhánh tương đối nhỏ, phản ánh số lượng subtype hạn chế và có thể liên quan đến quá trình tiến hóa xảy ra muộn hơn. Nhìn chung, bộ dữ liệu đã bao quát phần lớn các genotype và subtype HCV; đồng thời, phân tích

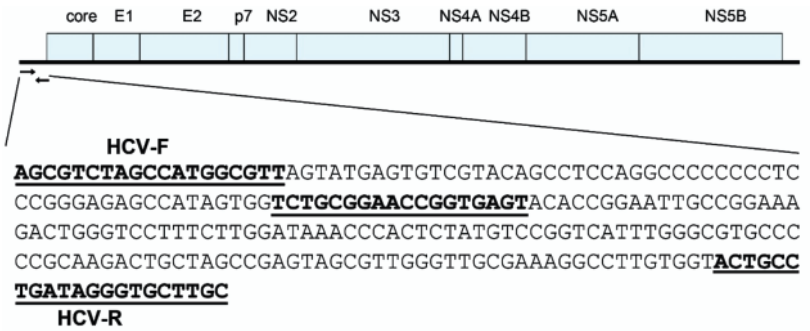
phát sinh loài đã xác nhận tính nhất quán về mặt tiến hóa của các genotype và subtype, qua đó hỗ trợ việc sử dụng bộ dữ liệu này trong các nghiên cứu xác định genotype và subtype HCV.

Bảng 1. Bộ dữ liệu hệ gen phân loại genotype (1 - 8) và subtype HCV

Mã số	Phân loại	Mã số	Phân loại	Mã số	Phân loại	Mã số	Phân loại
M62321	1a	KC844042	2f	FJ462438	4k	EF424625	6p
M67463	1a	DQ155561	2i	FJ839870	4l	EU408328	6r
AF009606	1a	JF735113	2j	FJ462433	4m	EU408329	6s
D90208	1b	AB031663	2k	FJ462441	4n	EU246939	6t
M58335	1b	KC197235	2l	FJ462440	4o	EF632071	6t
AJ238799	1b	KC197240	2l	FJ462431	4p	EU246940	6u
AY051292	1c	JX227967	2m	FJ462434	4q	EU158186	6v
D14853	1c	JF735111	2m	FJ462439	4r	EU798760	6v
KJ439768	1d	FN666429	2q	JF735136	4s	EU643834	6w
KC248194	1e	FN666428.2	2q	FJ839869	4t	DQ278892	6w
KJ439769	1e	JF735115	2r	HQ537009	4v	EU408332	6xa
KJ439770	1g	KC197238	2t	JX227959	4v	EU408330	6xa
AM910652.2	1g	JF735112	2u	FJ025855	4w	KJ567645	6xb
KC248199	1h	D28917	3a	FJ025856	4w	JX183552	6xb
KC248198	1h	D17763	3a	Y13184	5a	KJ567651	6xc
KJ439772	1i	D49374	3b	AF064490	5a	KM252789	6xd
KJ439773	1j	KJ470619	3d	Y12083	6a	KM252790	6xd
KJ439774	1k	KJ470618	3e	AY859526	6a	KM252792	6xe
KC248193	1l	JX227954	3g	D84262.2	6b	JX183557	6xe
KC248197	1l	JF735123	3g	EF424629	6c	KJ567647	6xf
KJ439778	1m	JF735121	3h	D84263.2	6d	KJ567646	6xf
KJ439782	1m	JX227955	3i	DQ314805	6e	MH492360	6xg
KJ439775	1n	FJ407092	3i	DQ835760	6f	MH492361	6xg
KJ439781	1n	D63821	3k	D63822	6g	MG879000	6xh
KJ439779	1o	JF735122	3k	D84265.2	6h	MZ504973	6xi
D00944	2a	Y11604	4a	DQ835770	6i	JX183549	6xi
AB047639	2a	FJ462435	4b	DQ835769	6j	MZ171127	6xj
AY746460	2a	FJ462436	4c	D84264.2	6k	DQ278891	6xj
D10988	2b	DQ418786	4d	EF424628	6l	EF108306	7a
AB030907	2b	FJ462437	4d	DQ835767	6m	KX092342	7b
D50409	2c	EU392175	4f	DQ278894	6n	MH590699	8a
JF735114	2d	EF589161	4f	DQ835768	6n	MH590698	8a
JF735120	2e	FJ462432	4g	EF424627	6o	MH590701	8a
KC844050	2f	EU392173	4k	EF424626	6p	MH590700	8a

3.2. Bộ mồi và mẫu dò phát hiện/ định lượng HCV

Trình tự mồi và mẫu dò được lựa chọn từ nghiên cứu của Yang và cộng sự (2015) [9]. Trong nghiên cứu này, việc đánh giá bộ mồi và mẫu dò được thực hiện bằng các phân tích *in silico*. Bộ oligonucleotide được sử dụng gồm mồi xuôi HCV-F (AGCGTCTAGCCATGGCGTT), mồi ngược HCV-R (GCAAGCACCCTATCAGGCAGT) và mẫu dò Probe-HCV (HEX-TCTGCGGAACCGGTGAGT-MGB), được thiết kế nhằm khuếch đại một vùng trình tự bảo tồn trên vùng 5'UTR của hệ gen HCV. Vùng 5'UTR là một trong những thành phần bảo tồn nhất của hệ gen HCV và thường được sử dụng làm trình tự đích trong chẩn đoán phân tử [9]. Việc thiết kế mồi và mẫu dò tại vùng trình tự này cho phép khuếch đại hiệu quả tất cả genotype và subtype khác nhau với sai khác trình tự tối thiểu, qua đó góp phần nâng cao độ nhạy của xét nghiệm (Hình 2).



Hình 2. Sơ đồ cấu trúc hệ gen HCV và vị trí bắt cặp của mồi và mẫu dò. Sơ đồ hệ gen HCV với các vùng mã hóa được chú thích, bao gồm Core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A và NS5B. Phần phóng đại vùng 5'-UTR minh họa vị trí bắt cặp của bộ mồi và mẫu dò; vị trí của mồi xuôi, mồi ngược và mẫu dò được gạch chân

Mồi xuôi HCV-F dài 19 nucleotide, có nhiệt độ nóng chảy (T_m) 66.2°C và hàm lượng GC 57.9%. Mồi ngược HCV-R dài 21 nucleotide, với T_m = 66.5°C và hàm lượng GC 61.9%. Mẫu dò Probe-HCV dài 18 nucleotide và T_m 61.1°C. Kích thước sản phẩm khuếch đại dự kiến của cặp mồi là 235 bp. Bộ mồi được thiết kế với ba nucleotide cuối chứa các nucleotide G/C (GTT, mồi xuôi) và (AGT, mồi ngược), góp phần ổn định liên kết bắt cặp nhờ tương tác hydro mạnh hơn. Nhìn chung, các thông số vật lý của bộ mồi và mẫu dò, bao gồm chiều dài, T_m, hàm lượng GC và thành phần trình tự ở đầu 3', đều nằm trong khoảng khuyến nghị cho hiệu năng tối ưu của phản ứng PCR (Bảng 2).

Bảng 2. Trình tự mồi và mẫu dò được sử dụng trong nghiên cứu

Tên mồi và mẫu dò	Trình tự (5'-3')	Chiều dài (bp)	T_m (°C)	Kích thước sản phẩm PCR (bp)
HCV-F	AGCGTCTAGCCATGGCGTT	19	66.2	235
HCV-R	GCAAGCACCCTATCAGGCAGT	21	66.5	
Probe-HCV	**HEX**-TCTGCGGAACCGGTGAGT-**MGB**	18	61.1	-

Chú thích: HEX - reporter; MGB - quencher

Bộ mồi và mẫu dò được đánh giá khả năng hình thành cấu trúc thứ cấp thông qua giá trị năng lượng tự do (ΔG) tối thiểu bằng công cụ OligoAnalyzer, bao gồm cấu trúc kẹp tóc (hairpin), tự tạo dimer (self-dimer) và tự dimer dị hợp (hetero-dimer) (Bảng 3). Kết quả dự đoán cho thấy giá trị ΔG hình thành cấu trúc kẹp tóc của mồi và mẫu dò dao động từ -0.78 đến -1.51 kcal/mol, phản ánh xu hướng gấp cuộn nội phân tử thấp. Khả năng hình thành cấu trúc self-dimer cũng hạn chế, với ΔG trong khoảng -4.16 đến -6.68 kcal/mol, cho thấy nguy cơ tương tác không đặc hiệu giữa mồi-mồi hoặc mẫu dò - mẫu dò ở mức thấp. Bên cạnh đó, giá trị ΔG hình thành cấu trúc hetero-dimer giữa mồi xuôi và mồi ngược là -6.21 kcal/mol, nằm trong ngưỡng chấp nhận. Nhìn chung, các phân tích *in silico* cho thấy bộ mồi và mẫu dò được lựa chọn có đặc tính nhiệt động học thuận lợi, ít bị ảnh hưởng bởi cấu trúc thứ cấp, phù hợp cho ứng dụng trong phản ứng PCR/dPCR.

Bảng 3. Năng lượng tự do (ΔG) của cấu trúc hairpin, self-dimer và hetero-dimer của của mồi và mẫu dò dự đoán bằng công cụ OligoAnalyzer

Tên mồi và mẫu dò	Hairpin ΔG (kcal.mole^{-1})	Self-dimer ΔG (kcal.mole^{-1})	Hetero-dimer ΔG (kcal.mole^{-1})
HCV-F	-1.3	-4.16	-6.21
HCV-R	-0.78	-4.67	
Probe-HCV	-1.51	-6.68	

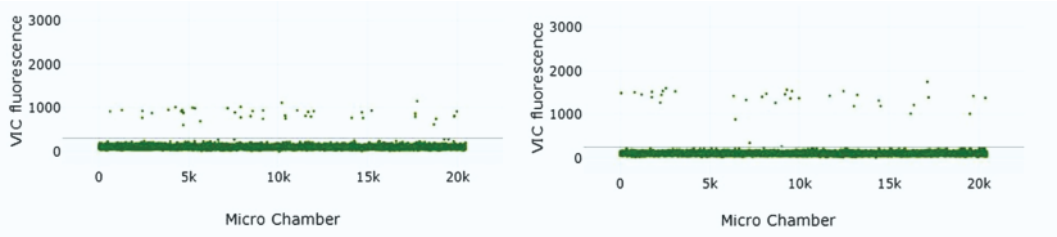
Dựa trên kết quả căn chỉnh trình tự, mồi xuôi và mồi ngược đều đạt mức tương đồng 100% với tất cả các hệ gen HCV được lựa chọn. Sản phẩm khuếch đại dự kiến có kích thước 235 bp, từ vị trí nucleotide 35 đến 269 trên trình tự hệ gen HCV. Đồng thời, trình tự mẫu dò cũng cho thấy mức tương đồng 100% trên toàn bộ các hệ gen HCV. Các kết quả này khẳng định tính đặc hiệu của bộ mồi và mẫu dò bằng các xét nghiệm dựa trên phương pháp PCR. Bên cạnh đó, công cụ NCBI Primer-

BLAST được sử dụng nhằm đánh giá khả năng bắt cặp/khuếch đại ngoài mục tiêu, đối với hệ gen người (taxid: 9606), HBV (taxid: 10404), HDV (taxid: 39759) và HIV (taxid: 11676). Kết quả cho thấy bộ mồi và mẫu dò không bắt cặp với bất kỳ trình tự nào trong các cơ sở dữ liệu nêu trên.

Hình 3. Vị trí bắt cặp của mồi và mẫu dò trên vùng 5'UTR ở hệ gen HCV. Vùng bắt cặp của mồi xuôi HCV-F và mồi ngược HCV-R được ký hiệu bằng các thanh màu xanh; vị trí bắt cặp của mẫu dò Probe-HCV được ký hiệu bằng thanh màu xám. Mã truy cập của các chủng HCV được hiển thị ở bên trái của trình tự

3.3. Bước đầu thử nghiệm kỹ thuật dPCR nhằm phát hiện/ định lượng HCV

Kỹ thuật dPCR được thử nghiệm trên 10 mẫu huyết thanh lâm sàng, bao gồm 5 mẫu dương tính với HCV và 5 mẫu âm tính, được chẩn đoán bằng kỹ thuật real-time PCR. Kết quả định lượng tải lượng HCV của từng mẫu được trình bày trong Bảng 4. Kỹ thuật dPCR cho thấy sự đồng thuận với kỹ thuật real-time PCR, đạt mức 100% trong phân loại dương/âm tính. Ở các mẫu dương tính, tải lượng HCV dao động trong khoảng từ 6.54×10^5 đến 5.29×10^6 bản sao/mL huyết thanh. Trên các biểu đồ kết quả, ghi nhận sự phân tách tốt giữa hai quần thể dương tính và âm tính (Hình 4). Tuy nhiên, một số vi phản ứng xuất hiện tín hiệu huỳnh quang trung gian ("rain"), các vi phản ứng có tín hiệu huỳnh quang không rõ ràng nằm giữa cụm dương tính và cụm âm tính. Quan sát này cho thấy cần tiếp tục tối ưu các điều kiện phản ứng dPCR nhằm cải thiện độ phân tách tín hiệu và tăng độ tin cậy của định lượng, đặc biệt trong các mẫu có tải lượng virus thấp.



Hình 4. Biểu đồ kết quả thử nghiệm phát hiện/ định lượng HCV (mẫu Trial-HCV-12 và Trial-HCV-13) bằng kỹ thuật dPCR. Mỗi điểm trên biểu đồ biểu thị cho từng vi phản ứng. Đường nền chia các vi phản ứng thành hai nhóm: Dương tính (điểm nằm trên đường nền, khuếch đại mục tiêu HCV) và âm tính (điểm nằm dưới đường nền, không khuếch đại)

Bảng 4. Kết quả định lượng tải lượng HCV bằng phương pháp dPCR trên mẫu lâm sàng

STT	Mẫu	Real-time PCR	dPCR (bản sao/ mL)
1	Trial-HCV-08	Dương tính	8.03×10^5
2	Trial-HCV-10	Dương tính	5.29×10^6
3	Trial-HCV-11	Dương tính	6.54×10^5
4	Trial-HCV-12	Dương tính	2.08×10^6
5	Trial-HCV-13	Dương tính	2.16×10^5
6	Control-01	Âm tính	Âm tính
7	Control-02	Âm tính	Âm tính

STT	Mẫu	Real-time PCR	dPCR (bản sao/ mL)
8	Control-03	Âm tính	Âm tính
9	Control-04	Âm tính	Âm tính
10	Control-05	Âm tính	Âm tính

4. BÀN LUẬN

Trong nghiên cứu này, chúng tôi trình bày hai nội dung chính: (1) xây dựng bộ dữ liệu tham chiếu phục vụ xác định genotype và subtype HCV một cách tin cậy; (2) thử nghiệm bộ mồi và mẫu dò nhằm đặt nền tảng xây dựng quy trình phát hiện/ định lượng HCV bằng kỹ thuật dPCR. Nghiên cứu chỉ dừng ở mức thiết lập các nền tảng ban đầu và đánh giá sơ bộ, chưa triển khai các phân tích chuyên sâu hoặc ứng dụng trên quy mô lâm sàng; đây là hạn chế của nghiên cứu. Cùng với đó, một hạn chế của nghiên cứu là cỡ mẫu tối thiểu (n = 10), do đó kết quả chưa có giá trị đại diện cho quần thể. Tuy nhiên, nghiên cứu được thực hiện theo hướng thử nghiệm bước đầu nhằm đánh giá khả năng hoạt động của bộ mồi và mẫu dò trong quy trình dPCR, nên không đặt mục tiêu suy rộng dịch tễ hay khái quát hóa. Tuy vậy, các phân tích và thử nghiệm bước đầu này có ý nghĩa thực tiễn, tạo cơ sở cho nhóm nghiên cứu phát triển các hướng nghiên cứu tiếp theo ở mức độ chuyên sâu hơn về dịch tễ học phân tử và chẩn đoán HCV. Trong thời gian tới, chúng tôi dự kiến: (1) ứng dụng kỹ thuật giải trình tự thế hệ mới (next-generation sequencing, NGS) để đánh giá đa dạng di truyền của các genotype và subtype HCV đang lưu hành tại Việt Nam; (2) triển khai nghiên cứu song song tập trung vào phát hiện và định lượng HCV bằng kỹ thuật dPCR. Nhìn chung, nghiên cứu này cung cấp nguồn dữ liệu và các vật liệu ban đầu quan trọng, làm tiền đề cho các nghiên cứu sâu hơn về hệ gen HCV và phát triển các xét nghiệm chẩn đoán liên quan.

Bộ dữ liệu phân tử được tổng hợp trong nghiên cứu này mang lại giá trị quan trọng trong việc xác định genotype và subtype của HCV [10]. Bằng cách sàng lọc, loại bỏ các trình tự bị gán nhãn sai, chất lượng thấp và/hoặc có dấu hiệu tái tổ hợp, chúng tôi xây dựng một bộ dữ liệu tham chiếu chuẩn với các tiêu chí được xác định rõ ràng. Khác với nhiều công cụ phân loại genotype trực tuyến, thường ít công bố chi tiết về bộ trình tự tham chiếu được sử dụng, bộ dữ liệu trong nghiên cứu này cung cấp genotype và subtype kèm theo bằng chứng phát sinh loài. Do đó, nguồn dữ liệu này góp phần chuẩn hóa phân loại HCV, đặc biệt đối với các subtype hiếm hoặc còn ít được nghiên cứu tại Việt Nam, thay vì phụ thuộc vào các công cụ phân loại trực tuyến.

Việc ứng dụng kỹ thuật dPCR để định lượng HCV mang lại các ý nghĩa quan trọng về phương pháp và lâm sàng. dPCR cho phép định lượng tuyệt đối mà không cần đường chuẩn, đồng thời cải thiện độ chính xác ở mức tải lượng virus thấp so với real-time PCR truyền thống [11]. Điều này đặc biệt phù hợp trong theo dõi bệnh nhân điều trị bằng thuốc kháng virus tác động trực tiếp (DAAs), khi việc phát hiện tồn dư virus tối thiểu một cách chính xác và khả năng lặp lại cao, điều này đóng vai trò then chốt trong đánh giá đáp ứng điều trị và nguy cơ tái phát [11]. Bên cạnh đó, độ nhạy cao của kỹ thuật dPCR giúp cải thiện khả năng phát hiện HCV trong các tình huống thách thức như nhiễm ẩn, giai đoạn nhiễm cấp sớm, hoặc ở người suy giảm miễn dịch với tải lượng virus trong máu rất thấp [12]. Cuối cùng, việc phát triển quy trình dPCR cho HCV tạo nền tảng kỹ thuật cho các xét nghiệm đa mồi trong tương lai, nhắm tới nhiều virus viêm gan (HAV, HBV, HDV, …) trong cùng một phản ứng, qua đó tinh gọn chẩn đoán viêm gan do virus trong lâm sàng.

5. KẾT LUẬN

Dựa trên các trình tự hệ gen HCV đã công bố, nghiên cứu này đã xây dựng một bộ dữ liệu tham chiếu được lựa chọn có hệ thống và được củng cố bằng phân tích phát sinh loài, nhằm xác định genotype và subtype HCV một cách chính xác. Đồng thời, nghiên cứu bước đầu triển khai ứng dụng kỹ thuật dPCR nhằm phát hiện/ định lượng HCV. Bộ mồi và mẫu dò được thiết kế trên vùng gen bảo tồn của HCV; các phân tích *in silico* cho thấy bộ mồi và mẫu dò có tính đặc hiệu đối với HCV và phù hợp với phương pháp real-time PCR. Thông qua việc tích hợp xác định genotype và subtype HCV và đổi mới công nghệ chẩn đoán (phương pháp dPCR), nghiên cứu này cung cấp cơ sở phương pháp luận quan trọng, góp phần hỗ trợ các nỗ lực quốc gia hướng tới mục tiêu loại trừ HCV tại Việt Nam.

LỜI CẢM ƠN

Nghiên cứu này được Trường Đại học Quốc tế Hồng Bàng cấp kinh phí thực hiện dưới mã số đề tài GVTC18.14.

TÀI LIỆU THAM KHẢO

[1] B. Flower *et al.*, "Seroprevalence of Hepatitis B, C and D in Vietnam: A systematic review and meta-analysis," *Lancet Reg. Heal. Pacific*, vol. 24, 2022.

[2] C. Le Ngoc *et al.*, "Differential prevalence and geographic distribution of hepatitis C virus genotypes in acute and chronic hepatitis C patients in Vietnam," *PLoS One*, vol. 14, no. 3, p. e0212734, 2019.

[3] A. Berto *et al.*, "Current challenges and possible solutions to improve access to care and treatment for hepatitis C infection in Vietnam: a systematic review," *BMC Infect. Dis.*, vol. 17, no. 1, p. 260, 2017.

[4] S. Zeuzem *et al.*, "Faldaprevir and deleobuvir for HCV genotype 1 infection," *N. Engl. J. Med.*, vol. 369, no. 7, pp. 630-639, 2013.

[5] D. Rapoud *et al.*, "Towards HCV elimination among people who inject drugs in Hai Phong, Vietnam: study protocol for an effectiveness-implementation trial evaluating an integrated model of HCV care (DRIVE-C: DRug use & Infections in ViEtnam-hepatitis C)," *BMJ Open*, vol. 10, no. 11, p. e039234, 2020.

[6] B. Flower *et al.*, "High cure rates for hepatitis C virus genotype 6 in advanced liver fibrosis with 12 weeks sofosbuvir and daclatasvir: The Vietnam SEARCH study," in *Open forum infectious diseases*, Oxford University Press US, 2021, p. ofab267.

[7] J. D. Scott and D. R. Gretch, "Molecular diagnostics of hepatitis C virus infection: a systematic review," *Jama*, vol. 297, no. 7, pp. 724-732, 2007.

[8] A. A. Morley, "Digital PCR: A brief history," *Biomol. Detect. Quantif.*, vol. 1, no. 1, pp. 1-2, 2014.

[9] Q. Yang *et al.*, "Evaluation of suitable control genes for quantitative polymerase chain reaction analysis of maternal plasma cell-free DNA," *Mol. Med. Rep.*, vol. 12, no. 5, pp. 7728-7734, 2015.

[10] D. B. Smith *et al.*, "Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource," *Hepatology*, vol. 59, no. 1, pp. 318-327, 2014.

[11] R. Kanagal-Shamanna, "Digital PCR: principles and applications," *Clin. Appl. PCR*, pp. 43-50, 2016.

[12] M. Frías *et al.*, "Evaluation of hepatitis C viral RNA persistence in HIV-infected patients with long-term sustained virological response by droplet digital PCR," *Sci. Rep.*, vol. 9, no. 1, p. 12507, 2019.

A dataset for HCV genotype identification and preliminary HCV quantification using digital PCR

Huynh Thi Thu Thao, Ha Thi Anh, Nguyen Thi Nga, Vu Thi Hai Yen, Nguyen Tuan Anh

ABSTRACT

Introduction: Hepatitis C virus (HCV) exhibits significant genetic diversity, posing challenges in detection and quantification across different genotypes. Research objectives: This study aims to collect representative HCV genome sequences to construct a robust genotype- and subtype-specific reference dataset reflecting both global and Vietnam-specific diversity. In parallel, it seeks to initiate the development of a dPCR-based assay for sensitive and accurate detection and quantification of HCV RNA. Subjects and methods: We constructed a curated reference dataset comprising representative HCV genotypes and subtypes and conducted in silico analyses to select suitable primers and probe for dPCR assay. Results: A panel of reference sequences representing 86 validated HCV subtypes across all eight recognized genotypes was compiled to facilitate accurate genotype and subtype identification. Additionally, preliminary dPCR experiments using clinical serum samples showed high concordance with real-time PCR and successfully quantified viral loads in HCV-positive samples. Conclusion: These foundational results provide key resources and protocols for future clinical applications and large-scale studies on HCV genotype identification and molecular diagnostics using dPCR.

***Key words**: Hepatitis C virus, in silico analysis, quantification, digital PCR*

Received: 12/12/2025
Revised: 13/01/2026
Accepted for publication: 15/01/2026