Tạp chí Công nghệ Sinh học **17**(3): 411-418, 2019

PHÁT HIỆN BIẾN THỂ MỚI TRÊN GEN *MYOPALLADIN* Ở BỆNH NHÂN CƠ TIM BẰNG KỸ THUẬT GIẢI TRÌNH TỰ TOÀN BỘ VÙNG MÃ HOÁ (EXOME)

Bùi Chí Bảo[1,2,3#], Nguyễn Minh Hiệp[4#], Nguyễn Mạnh Công[5], Phạm Thị Thu Trang[6], Lương Thị Thắm[6], Hà Thị Thanh Ngà[6], Vũ Bảo Quốc[7], Phạm Hồ Thuật Khoa[7], Nguyễn Thị Huỳnh Nga[7,✉]

[1]*Khoa Y, Đại học Quốc gia Thành phố Hồ Chí Minh*
[2]*Trung tâm Y Sinh học Phân tử, Trường Đại học Y Dược, Thành phố Hồ Chí Minh*
[3]*Đơn vị Sinh học Phân tử Di truyền, Bệnh viện Nhi đồng 2, Thành phố Hồ Chí Minh*
[4]*Trung tâm Công nghệ bức xạ, Viện nghiên cứu hạt nhân, Thành phố Đà Lạt*
[5]*Đơn vị nghiên cứu chức năng gen, Công ty Cổ phần Công nghệ Y khoa DNA, Thành phố Hồ Chí Minh*
[6]*Khoa Sau Đại học, Trường Đại học Đà Lạt, Thành phố Đà Lạt*
[7]*Khoa Sinh học, Trường Đại học Đà Lạt, Thành phố Đà Lạt*

[#]Các tác giả có mức độ đóng góp ngang nhau
[✉]Người chịu trách nhiệm liên lạc. E-mail: nganth@dlu.edu.vn

Ngày nhận bài: 17.12.2018
Ngày nhận đăng: 26.7.2019

TÓM TẮT

Ở nhóm bệnh cơ tim vô căn, sự trùng lặp trong biểu hiện lâm sàng gây hạn chế cho việc chẩn đoán và điều trị, do đó, việc nghiên cứu về các gen gây bệnh là rất cần thiết. Trong dự án này, chúng tôi ứng dụng kỹ thuật giải trình tự thế hệ mới (NGS) nhằm khảo sát 142 gen có liên quan đến các bệnh cơ tim ở 9 bệnh nhân người Việt Nam tại Bệnh viện Nhi Đồng 2 và Bệnh viện Đại học Y Dược, Thành phố Hồ Chí Minh. Bằng chiến lược giải trình tự toàn bộ vùng mã hoá (WES), qua quá trình phân tích và sàng lọc phân tử, chúng tôi đã xác định được 65 biến thể hiếm thuộc 18 gen trên nhóm bệnh nhân người Việt Nam nói trên, bao gồm 28 biến thể đồng hợp tử và 37 biến thể dị hợp tử. Tần số biến thể của gen *TTN* chiếm nhiều nhất với 13 biến thể. Tiếp theo là các gen *SYNE1*và *MYPN* lần lượt có 9 và 8 biến thể. Gen *SYNE2* có 6 biến thể. Mỗi gen *NDUFV2* và *SCN5A* có 5 biến thể, gen *COX15* có 4 biến thể. Chúng tôi xác định được 2 biến thể của mỗi gen *DMD*, *KCNE1, NEBL*, *RBM20* và 1 biến thể của các gen *AKAP9*, *CAV3*, *DSC2*, *DSG2*, *DSP, MYBPC3* và *MYH6*. Trong đó, có 1 biến thể mới là đột biến dị hợp tử c.1527C>G của gen *MYPN*. Các kết quả nghiên cứu di truyền trên sẽ góp phần vào việc chẩn đoán phân tử và tầm soát bệnh tim mạch được triệt để hơn.

Từ khoá: *Bệnh cơ tim vô căn, Biến thể gen, Đột biến, Giải trình tự thế hệ mới (NGS), Giải trình tự toàn bộ vùng mã hoá (WES)*

MỞ ĐẦU

Bệnh cơ tim vô căn (cardiomyopathy) là nhóm bệnh đa gen gây rối loạn ở nguyên bào cơ tim, dẫn đến suy tim, loạn nhịp tim hoặc đột tử (Sisakian, 2014). Các nhóm bệnh đã được phân loại bao gồm: bệnh cơ tim giãn nở (DCM), bệnh cơ tim phì đại (HCM), bệnh cơ tim hạn chế (RCM) và bệnh cơ tim thất phải gây loạn nhịp (ARVC) (Simpson *et al.,* 2017). Những bệnh này có tính đa dạng di truyền và liên quan đến các đột biến hiếm ở một số lượng lớn các gen, nhiều loại gen này trùng lặp với các bệnh lý cơ tim khác nhau (Simpson *et al.,* 2017). Sự trùng lặp trong biểu hiện lâm sàng giữa các bệnh tim mạch đã gây hạn chế cho việc chẩn đoán và điều trị.

Giải trình tự Sanger là một quy trình xác định chính xác trình tự các nucleotide của phân tử DNA, được ứng dụng rộng rãi đối với các rối loạn chủ yếu liên quan đến một gen gây bệnh đơn lẻ (Chen *et al.,* 2014). Tuy nhiên, phương pháp sàng lọc này rất mất thời gian và cho mức độ không đồng nhất di truyền

cao. Cho đến nay, với hơn 100 gen liên quan đến bệnh cơ tim được xác định, có thể được giải quyết hiệu quả hơn bằng cách sắp xếp trình tự thông tin cao, được gọi là giải trình tự thế hệ mới (NGS). Trong đó, kỹ thuật giải trình tự gen toàn bộ vùng mã hoá (WES) là một ứng dụng của công nghệ NGS nhằm xác định các biến thể của tất cả các vùng mã hoá (exome) của các gen mục tiêu theo từng loại bệnh đã biết, giúp sàng lọc, chẩn đoán và đánh giá biến thể (Faita *et al.,* 2012; Morini *et al.,* 2015).

Trong nghiên cứu này, chúng tôi sẽ thiết kế một bảng mẫu dò tùy chỉnh có chứa 142 gen liên quan đến các bệnh cơ tim. Chúng tôi sẽ tiến hành sàng lọc phân tử ở các đối tượng nghiên cứu nhằm tìm ra các biến thể mới liên quan đến bệnh cơ tim tương ứng, góp phần vào việc chẩn đoán phân tử đối với bệnh cơ tim được chính xác hơn, nhằm xác định sớm quá trình phát triển loạn nhịp tim và quản lý lâm sàng tốt hơn đối với bệnh nhân cơ tim.

ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Đối tượng

Bệnh nhân người Việt Nam được lập hồ sơ hội chứng, di truyền và quản lý lâm sàng tại Bệnh viện Nhi đồng 2, Thành phố Hồ Chí Minh và Bệnh viện Đại học Y Dược, Thành phố Hồ Chí Minh. Nhóm đối chứng gồm 4 đối tượng nghiên cứu khoẻ mạnh, không biểu hiện bệnh (NH, normal human). Nhóm tiếp theo gồm 3 bệnh nhân cơ tim thất phải gây loạn nhịp (ARVC, arrhythmogenic right ventricular cardiomyopathy). Nhóm còn lại gồm 2 bệnh nhân có bệnh lý về da (SD, skin disorder).

Phương pháp thu mẫu, tách chiết và tinh sạch DNA tổng số

Một lượng 3 – 5 mL máu tĩnh mạch được lấy trực tiếp từ bệnh nhân, cho vào ống chống đông chứa ethylenediaminetetraacetic acid (EDTA), trên ống có ghi đầy đủ thông tin của bệnh nhân. Ống mẫu được bảo quản trong thùng đựng mẫu ở nhiệt độ 2 – 8^{o}C để đưa đi tách DNA.

DNA tổng số (gDNA) từ 200 µL mẫu máu bệnh nhân được tách chiết và tinh sạch bằng bộ kit Qiagen theo hướng dẫn của nhà sản xuất. Sau khi tách chiết, các mẫu DNA được đo bằng máy NanoDrop có nồng độ trong khoảng 100 – 200ng/µL và giá trị OD 260/280 trong khoảng 1,8 – 1,9. Mẫu DNA sau đó được chuyển đi giải trình tự toàn bộ vùng mã hóa trên hệ thống HiSeq 4000 (Illumina) tại công ty Macrogen.

Phương pháp chuẩn bị và thiết kế mẫu hệ gen

Phân mảnh sử dụng nền tảng Illumina như sau: 1 µg gDNA được cắt thành 100 – 900 bp mảnh với Covaris E210 (Covaris, Inc., Woburn, MA) để chạy trên thư viện NGS. Thư viện cho giải trình tự được dựa trên nguyên tắc số mẫu dò (probe) được thiết kế cho hệ gen hoặc cho nhóm gen mục tiêu AgilentSureSelectRT Reagent. Chúng tôi cũng tạo số lượng mẫu dò cho nhóm gen mục tiêu 142 gen của bệnh cơ tim. Đây là mẫu dò thư viện được ký hiệu "Pan 146 Cardiomyo". Trong bước này, chúng tôi thu thập trình tự mã hóa (exon) của toàn bộ 142 gen, sau đó sử dụng phần mềm Afident Probe Library Select để chạy tìm mẫu dò. Tổng cộng có 2.072.000 mẫu dò bao phủ cho 142 gen được đặt hàng riêng cho nhóm nghiên cứu. Thư viện có thể được định lượng bằng PCR định lượng (qPCR) bằng cách sử dụng Bộ định lượng thư viện Kapa (Kapa Biosystems, Inc., Wilmington, MA) trên 7900HT (Applied Biosystems, Foster City, CA). Giải trình tự được thực hiện trên máy HiSeq 4000 với bộ kit TruSeq 3000/4000 SBS Kit v3 (protocol HiSeq 3000/4000 System User Guide Part # 15066496, Rev. A HCS 3.3.20). Dữ liệu sau cùng được xuất ra và gửi về ở dạng file *.Fastq.

Để đối chiếu các đoạn đọc ngắn lên ngân hàng gen người (UCSC/hg19), chúng tôi dùng phần mềm Burrows Wheeler Aligner (BWA). Việc xử lý tiếp theo là sắp xếp, hợp nhất và loại bỏ trùng lặp cho các tệp BAM được thực hiện bằng cách sử dụng SAMtools và Picard (http://picard.sourceforge.net/index.shtml). Để xuất các biến thể (các SNP và các INDEL ngắn), chúng tôi sử dụng các phần mềm Platypus and Genome Analysis Toolkit (GATK). Các biến thể được chú thích bởi phần mềm ANNOVAR. Các bước lọc biến thể tiếp theo, việc chú giải biến thể và kiểm tra sau chức năng của các đột biến được thực hiện trên phần mềm Geneticist Assitant. Để tìm ra các biến thể tiềm năng, nghiên cứu sẽ giữ lại các biến thể không đồng nghĩa (nonsynonymous), có MAF ≤ 0.005 và được dự đoán gây bệnh bằng tất cả các công cụ dự đoán chức năng base và xuất biến thể. Các biến thể có điểm chất lượng tối thiểu được loại bỏ không xem xét. Các tần số allele nhỏ của các biến thể được xác định từ ba database: ExAC database, 1000 Genome Project database và Exome Variant Server. Các biến thể sai nghĩa (missense variant) được xem là "potentially pathogenic" nếu được phân loại đồng thời là "damaging" ở SIFT, "deleterious" ở PROVEAN, "possibly" hoặc "probably damaging"

ở PolyPhen-2 và "disease causing" ở MutationTaster.

KẾT QUẢ VÀ THẢO LUẬN

Thông tin của nhóm đối tượng nghiên cứu

Thông tin của nhóm đối tượng nghiên cứu được trình bày ở Bảng 1. Tất cả các mẫu trong nghiên cứu này được thu thập từ bệnh nhân người Việt Nam tại Bệnh viện Nhi đồng 2, Thành phố Hồ Chí Minh và Bệnh viện Đại học Y Dược, Thành phố Hồ Chí Minh khi đã có sự đồng ý của bệnh nhân.

Nhiều nghiên cứu cho thấy có một mối liên hệ mật thiết giữa các bệnh cơ tim và bệnh lý về da vì độ vững chắc của các cầu nối bám được quyết định bởi phần lớn những protein liên kết ở cả tế bào tim và da (Alcalai *et al.,* 2003; McKoy*et al.,* 2000; Norgett *et al.,* 2000). Do đó, chúng tôi tiến hành khảo sát cả trên hai đối tượng bệnh nhân có bệnh lý về da.

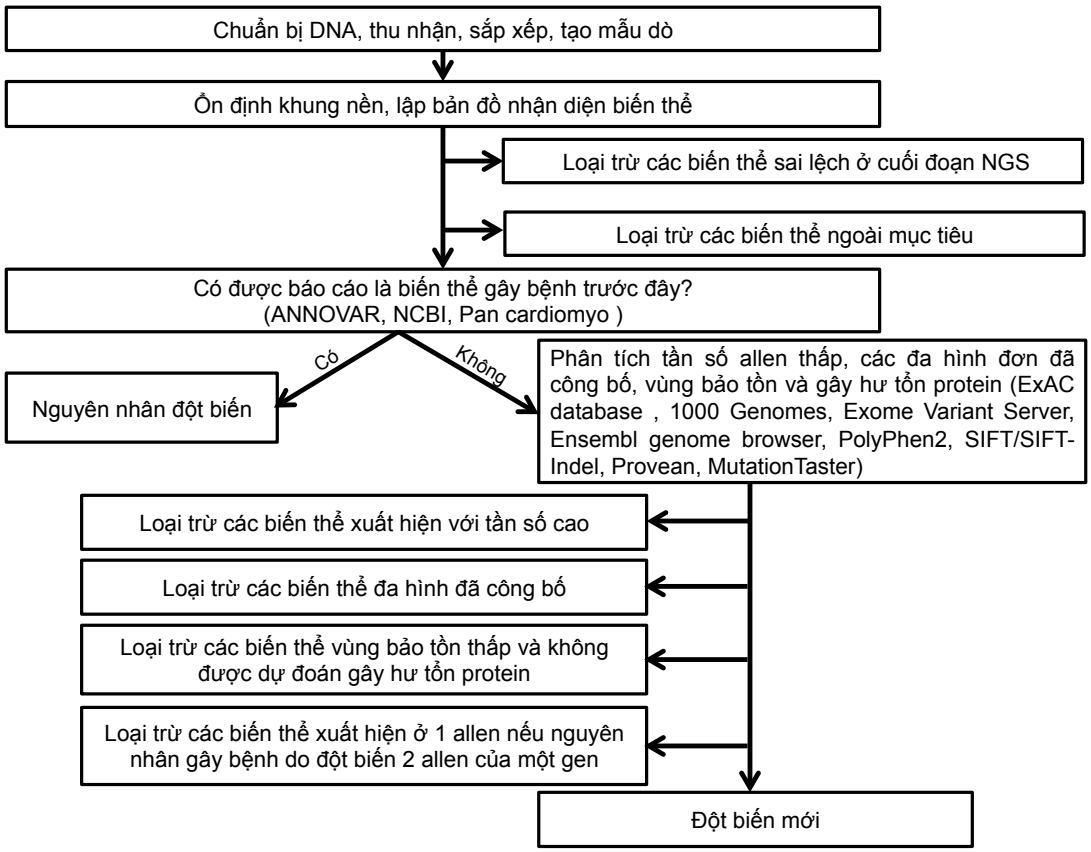
Bảng 1. Thông tin của nhóm đối tượng nghiên cứu.

STT	Mã số	Giới tính	Tuổi	*Gen* (số lượng biến thể) ĐỒNG HỢP/DỊ HỢP	Biến thể tiềm năng	Đặc điểm lâm sàng
1	ARVC-1	Nữ	4 tháng tuổi	*CAV3* (1) HET; *DMD* (2) HET; *DSP* (1) HET; *KCNE1* (1) HOM; *MYH6* (1) HET; *NEBL* (1) HET; *RBM20* (1) HOM; *SCN5A* (1) HET; *TTN* (6) HET	-	Khó thở, hẹp eo động mạch chủ, nhiễm sắc thể bất thường
2	ARVC-2	Nam	3 tháng tuổi	*KCNE1* (1) HOM; *NEBL* (1) HET; *RBM20* (1) HOM; *SCN5A* (1) HOM	-	Khó thở, bất thường về gen *SFTPB*, *SFTPC*
3	ARVC-3	Nữ	2 tuổi	*COX15* (1) HET; *MYPN* (1) HET; *NDUFV2* (1) HOM; *SCN5A* (1) HET; *SYNE2* (1) HET	*MYPN* (1) HET; *SYNE2* (1) HET	Khó thở, chỉ số tim ngực >55%
4	SD-1	Nam	11 tháng tuổi	*COX15* (1) HOM; *MYPN* (5) HOM; *NDUFV2* (1) HET; *SYNE1* (1) HET; *SYNE2* (2) HET; *TTN* (2) 1 HET, 1 HOM	*MYPN* (2) HOM; *SYNE1* (1) HET; *SYNE2* (2) HET	Dày sừng ở lòng bàn tay (bẩm sinh)
5	SD-2	Nam	19 tuổi	*COX15* (1) HOM; *MYBPC3* (1) HET; *MYPN* (1) HET; *SYNE2* (2)HET; *TTN* (1) HOM	*MYPN* (1) HET	Dày sừng ở lòng bàn tay
6	NH-1	Nam	38 tuổi	*COX15* (1) HET; *MYPN* (1) HOM; *SCN5A* (1) HET; *SYNE1* (7) 6 HOM, 1 HET; *SYNE2* (1) HET; *TTN* (1) HOM	*MYPN* (1) HOM; *SYNE2* (1) HET	Đối chứng (người khỏe mạnh)
7	NH-2	Nữ	31 tuổi	*AKAP9* (1) HET; *NDUFV2* (1) HOM; *SYNE1* (1) HET; *TTN* (1) HOM	*AKAP9* (1) HET; *SYNE1* (1) HET	Đối chứng (người khỏe mạnh)
8	NH-3	Nam	12 tuổi	*DSC2* (1) HET; *NDUFV2* (1) HET; *TTN* (1) HET	*DSC2* (1) HET	Đối chứng (người khỏe mạnh)
9	NH-4	Nam	38 tuổi	*DSG2* (1) HET; *NDUFV2* (1) HOM; *SCN5A* (1) HOM; *TTN* (1) HOM	*DSG2* (1) HET	Đối chứng (người khỏe mạnh)

Ghi chú: ARVC (arrhythmogenic right ventricular cardiomyopathy): bệnh nhân cơ tim thất phải gây loạn nhịp; HET (heterozygous): dị hợp tử; HOM (homozygous): đồng hợp tử; NH (normal human): đối tượng nghiên cứu không biểu hiện bệnh; SD (skin disorder): bệnh nhân có bệnh lý về da.

Quy trình giải trình tự gen toàn bộ vùng mã hoá (exome) và lọc biến thể

Chúng tôi đã thiết kế một pan 146 cardiomyo tùy chỉnh có chứa 142 gen. Bảng điều khiển gen tùy chỉnh bao gồm tất cả các exon của mỗi gen và các vị trí nối. Chúng tôi ứng dụng phương pháp giải trình tự gen thế hệ mới Illumina. Bằng cách sử dụng bảng điều khiển tùy chỉnh, chúng tôi xác định được bộ số liệu các đột biến gen và biến dị di truyền ở vùng mã hoá liên quan đến bệnh cơ tim ở 9 đối tượng nghiên cứu. Dung lượng giải trình tự của mỗi mẫu vào khoảng 5,9 Gbp, độ bao phủ (depth coverage) từ 65x đến 105x. Phần lớn các lần đọc có chất lượng base cao, với Q30 (Phred-score) là 96%.

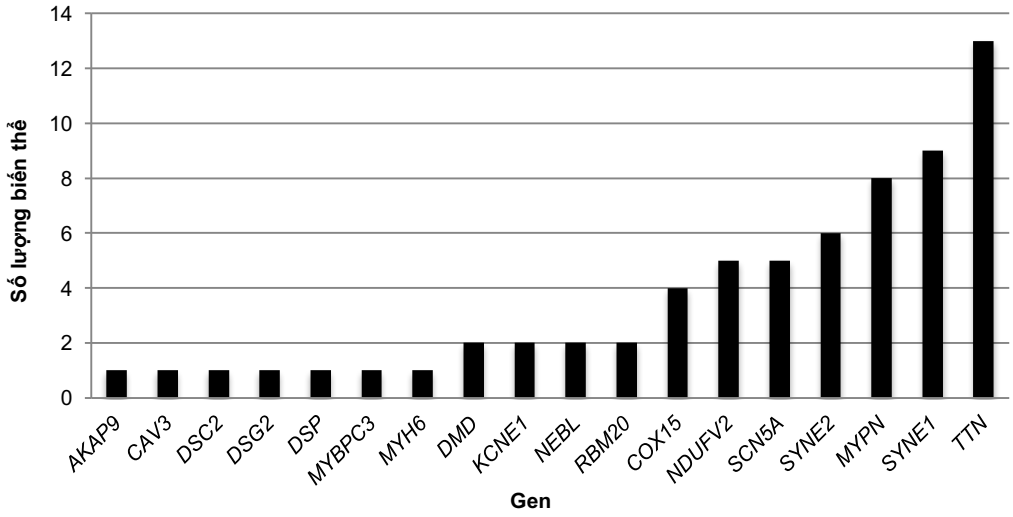


Hình 1. Quy trình phân tích và lọc biến thể. NCBI (National Center for Biotechnology Information): Trung tâm Thông tin Công nghệ sinh học Quốc gia, Pan cardiomyo: panel bệnh cơ tim, PolyPhen (Polymorphism Phenotyping): phân tích tính đa hình về kiểu hình, SIFT (Sorting Intolerant From Tolerant): công cụ dự đoán chức năng.

Kết quả đánh giá biến thể

Trong số 9 đối tượng nghiên cứu, chúng tôi phát hiện 65 biến thể thuộc 18 gen. Trong đó có 28 biến thể ở dạng đồng hợp tử và 37 biến thể dị hợp tử. Có 2 biến thể thuộc nhiễm sắc thể X ở dạng dị hợp tử. Trong 65 biến thể có 45 biến thể thuộc loại "vô nghĩa" (nonsynonymous), chiếm 69,2%, 14 biến thể "sai nghĩa" (missense), chiếm 21,5%, 3 biến thể "thuộc vị trí nối" (splice site), (chiếm 4,6%), 1 biến thể "dừng" (stop-gain), chiếm 1,5%, 1 biến thể "ngược hướng" (upstream), chiếm 1,5% và 1 biến thể "lệch khung đọc" (frameshift), chiếm 1,5%.

Tần số biến thể của gen *TTN* chiếm nhiều nhất với 13 biến thể. Tiếp theo là các gen *SYNE1* và *MYPN* lần lượt có 9 và 8 biến thể. Gen *SYNE2* có 6 biến thể. Mỗi gen *NDUFV2* và *SCN5A* có 5 biến thể, gen *COX15* có 4 biến thể. Ngoài ra, chúng tôi xác định được 2 biến thể của mỗi gen *DMD*, *KCNE1, NEBL*, *RBM20* và 1 biến thể của các gen *AKAP9*, *CAV3*, *DSC2*, *DSG2*, *DSP, MYBPC3* và *MYH6* **(**Hình 2**)**. Trong đó, có 1 biến thể mới là đột biến dị hợp tử c.1527C>G của gen *MYPN*.



Hình 2. Biểu đồ thể hiện tần số xuất hiện của các biến thể ở 9 đối tượng nghiên cứu. Ở 9 đối tượng nghiên cứu, tổng cộng 65 biến thể được xác định xuất hiện ở 18 gen. Số lượng biến thể ở gen *TTN* là cao nhất với 13 biến thể, gen *SYNE1* với 9 biến thể, gen *MYPN* với 8 biến thể. Gen *SYNE2* có 6 biến thể. Mỗi gen *NDUFV2* và *SCN5A* đều có 5 biến thể. Gen *COX15* có 4 biến thể. Mỗi gen *DMD, KCNE1, NEBL* và *RBM20* có 2 biến thể. Các gen *AKAP9, CAV3,DSC2, DSG2, DSP, MYBPC3* và *MYH6* có 1 biến thể.

Hiện nay, công nghệ giải trình tự NGS là cách tiếp cận mạnh mẽ để khám phá toàn diện các đột biến di truyền đối với hàng loạt các bệnh lý ở người (Di Resta *et al.,* 2018). Trong đó, kỹ thuật WES là một ứng dụng của công nghệ NGS nhằm xác định các biến thể trên tất cả các vùng mã hóa trong hệ gen. WES ngày càng được sử dụng rộng rãi trong nghiên cứu các bệnh hiểm nghèo, khó chẩn đoán, không phát hiện rõ nguyên nhân như ung thư, tim mạch, thần kinh, v.v. (Boemer *et al.,* 2017; Qiao D *et al.,* 2018; Goh, Choi, 2012; Haskell *et al.,* 2018; Golbus *et al.,* 2014). Trong nghiên cứu này, chúng tôi đã sử dụng các phần mềm và thuật toán tin sinh học để xác định các biến đổi di truyền mới có liên quan đến các bệnh về cơ tim và những bệnh lý khác ảnh hưởng đến hệ tim mạch.

Ở nhóm 3 bệnh nhân ARVC, có 13 gen được xác định bao gồm *CAV3, COX15, DMD, DSP, KCNE1, MYH6, MYPN, NDUFV2, NEBL, RBM20, SCN5A, SYEN2* và *TTN* với 24 biến thể (Bảng 1). Trong đó gen *TTN* chiếm số lượng biến thể nhiều nhất với 6 biến thể và gen *SCN5A* với 3 biến thể, xuất hiện ở cả 3 bệnh nhân. Mỗi gen *KCNE1, NEBL, RBM20* được xác định có 2 biến thể và các gen còn lại có 1 biến thể. Đáng lưu ý là các gen *CAV3, COX15, KCNE1, MYH6, MYPN, NDUFV2, NEBL* và *RBM20* thường không xuất hiện ở bệnh lý ARVC, điều này cho thấy sự trùng lặp của nhiều kiểu gen và kiểu hình ở bệnh cơ tim.

Đối với nhóm 4 đối tượng nghiên cứu khỏe mạnh không có biểu hiện bệnh lý cơ tim (nhóm NH), công nghệ giải trình tự NGS đã xác định được 10 gen là *AKAP9, COX15, DSC2, DSG2, MYPN, NDUFV2, SCN5A, SYNE1, SYNE2, TTN* với 23 biến thể gây bệnh cơ tim **(**Bảng 1). Gen *SYNE1* được phát hiện có nhiều biến thể nhất với 8 biến thể và gen *TTN* có 4 biến thể xuất hiện ở tất cả các đối tượng được khảo sát. Kết quả cho thấy phổ xuất hiện các đột biến liên quan đến bệnh cơ tim là khá cao, tương ứng với nguy cơ tiền ẩm mắc bệnh ở nhóm đối tượng trên, thể hiện cả ở 3 loại bệnh lý cơ tim gồm bệnh ARVC, DCM và HCM. Tuy nhiên, thông qua các phần mềm dự đoán sinh học bao gồm PolyPhen2, MutationTaster và SIFT, chúng tôi nhận thấy phần lớn những biến thể phát hiện được ở người khỏe mạnh là những đột biến vô nghĩa, hoặc vị trí đột biến không làm ảnh hưởng đến cấu trúc của protein. Vì vậy, đối với những bệnh đa gen, mức độ biểu hiện bệnh ở từng cá thể còn liên quan đến nhân tố ngoại cảnh cùng với số lượng và loại đột biến, cũng như phương thức di truyền phức tạp khi mỗi gen đều có tác động tích lũy, việc kết hợp kết quả chẩn đoán lâm sàng của bệnh nhân với việc nghiên cứu di truyền phân tử mới có thể khẳng định được

liệu đột biến có phải là nguyên nhân gây bệnh.

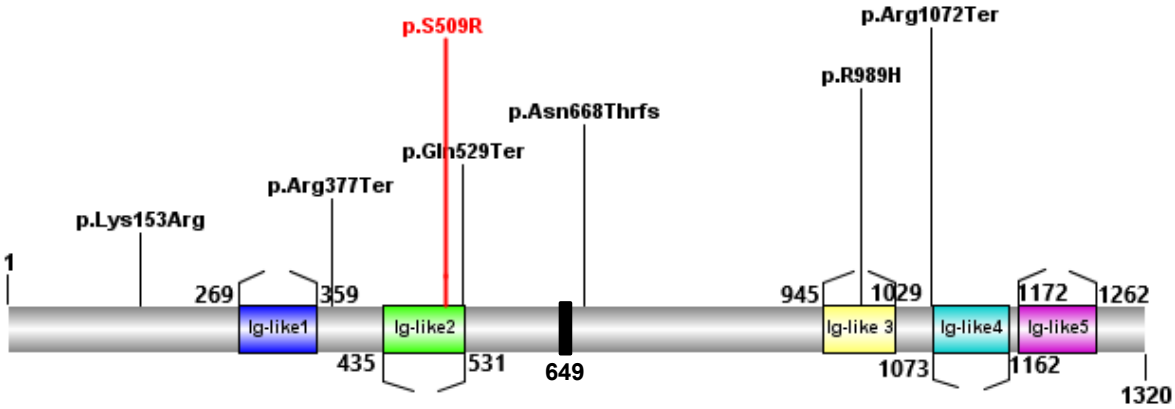
Hai bệnh nhân SD mang bệnh lý về da mà chúng tôi nghi ngờ có liên quan đến bệnh tim mạch được phát hiện có 7 gen gồm *COX15, MYBPC3, MYPN, NDUFV2, SYNE1, SYNE2, TTN* với 18 biến thể (Bảng 1). Gen mang nhiều biến thể nhất là *MYPN* với 6 biến thể, tiếp theo là các gen *SYNE2* với 4 biến thể và gen *TTN* với 3 biến thể. Hai bệnh nhân này mang số lượng biến thể gây các bệnh cơ tim cao nhất trong số 3 nhóm đối tượng nghiên cứu được khảo sát. Kết quả cho thấycác bệnh nhân trên có nguy cơ cao mắc bệnh tim mạch.

Đáng lưu ý là chúng tôi đã tìm thấy một đột biến mới nonsynonymous ở dạng dị hợp tử là c.1527C>G ở gen *MYPN*. Gen *MYPN* mã hoá cho protein myopalladin nằm trên nhiễm sắc thể 10q21.3 (Bang *et al.;* 2001, Florescu *et al.,* 2016). Đột biến trên xảy ra ở exon thứ 15, trong đó C được thay thế bằng G ở vị trí 1527 trong cDNA, dẫn đến việc thay thế amino acid Ser (S) bằng Arg (R) ở vị trí 509 trên protein MYPN (Hình 3). Đột biến này có tần số allele là 0,5205 theo ExAC. Đột biến có khả năng ảnh hưởng đến cơ tim ở bệnh nhân có bệnh lý về da.

Bảng 2. Các gen gây bệnh cơ tim xếp theo từng nhóm đối tượng nghiên cứu.

STT	Đối tượng nghiên cứu	Gen	Loại bệnh cơ tim lâm sàng
1	ARVC	*CAV3*	HCM, LQTS
		COX15	HCM
		DMD	ARVC, DCM, HCM
		DSP	ARVC, DCM, HCM, RCM
		KCNE1	AF, LQTS
		MYH6	DCM, HCM
		MYPN	DCM, HCM
		NDUFV2	HCM
		NEBL	DCM, HCM
		RBM20	DCM, HCM
		SCN5A	AF, ARVC, DCM, HCM, LQTS, RCM
		SYNE2	ARVC, DCM, HCM
		TTN	ARVC, DCM, HCM
2	SD	*COX15*	HCM
		MYBPC3	DCM, HCM
		MYPN	DCM, HCM
		NDUFV2	HCM
		SYNE1	ARVC, DCM, HCM
		SYNE2	ARVC, DCM, HCM
		TTN	ARVC, DCM, HCM
3	NH	*AKAP9*	DCM, HCM, LQTS, RCM
		COX15	HCM
		DSC2	ARVC, DCM, HCM
		DSG2	ARVC, DCM, HCM, RCM
		MYPN	DCM, HCM
		NDUFV2	HCM
		SCN5A	AF, ARVC, DCM, HCM, LQTS, RCM
		SYNE1	ARVC, DCM, HCM
		SYNE2	ARVC, DCM, HCM
		TTN	ARVC, DCM, HCM

Ghi chú: AF (atrial fibrillation): rung nhĩ; ARVC (arrhythmogenic right ventricular cardiomyopathy): bệnh cơ tim thất phải gây loạn nhịp; CM (cardiomyopathy): bệnh cơ tim nói chung; DCM (dilated cardiomyopathy): bệnh cơ tim dãn; HCM (hypertrophic cardiomyopathy): bệnh cơ tim phì đại; LQTS (long QT syndrome): hội chứng QT kéo dài; RCM (restrictive cardiomyopathy) bệnh cơ tim hạn chế. Trong bảng có thể xác định sự trùng lặp về kiểu hình rất rõ, đồng thời thể hiện các cặp kiểu gen - kiểu hình mới trong nhóm bệnh nhân ARVC.



Hình 3. Sự phân bố của các đột biến trên domain của protein myopalladin. Protein myopalladin được mã hóa bởi gen *myopalladin (MYPN)*. Protein myopalladin là thành phần cấu tạo nên cơ tim và cơ xương, có phân tử lượng 145 kDa, bao gồm 1320 amino acid. Myopalladin có 5 trình tự lặp immunoglobulin (Ig) và một vùng giàu proline (màu đen, vị trí amino acid thứ 649). Hình vẽ mô tả sự phân bố của 6 đột biến đã được công bố (cập nhật đến năm 2019) và vị trí biến thể mới được phát hiện (màu đỏ) ở bệnh nhân cơ tim người Việt Nam.

KẾT LUẬN

Kết quả nghiên cứu di truyền ở công trình này cho thấy rõ vai trò của các biến thể liên quan đến các bệnh lý cơ tim, góp phần vào việc chẩn đoán phân tử và tầm soát bệnh tim mạch được triệt để hơn. Việc giải trình tự và sàng lọc phân tử ở người bình thường hoặc bệnh nhân tim mạch là rất quan trọng, giúp cho việc xác định sớm quá trình phát triển loạn nhịp tim và quản lý lâm sàng tốt hơn đối với bệnh nhân cơ tim.

Lời cảm ơn: *Nghiên cứu này được tài trợ bởi Quỹ phát triển khoa học và công nghệ Quốc gia (NAFOSTED) trong đề tài mã số 106-YS.01-2016.39.*

TÀI LIỆU THAM KHẢO

Alcalai R, Metzger S, Rosenheck S, Meiner V, Chajek-Shaul T (2003) A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder and woolly hair. *J Am Coll Cardiol* 42: 319-327.

Bang ML, Mudry RE, McElhinny AS, Trombitás K, Geach AJ, Yamasaki R, Sorimachi H, Granzier H, Gregorio CC, Labeit S (2001) Myopalladin, a novel 145-kilodalton sarcomeric protein with multiple roles in Z-disc and I-band protein assemblies. *J Cell Biol* 153 (2): 413-27.

Boemer F, Fasquelle C, d'Otreppe S, Josse C, Dideberg V, Segers K, Guissard V, Capraro V, Debray FG, Bours V (2017) A next-generation newborn screening pilot study: NGS on dried blood spots detects causal mutations in patients with inherited metabolic diseases. *SciRep*7: 17641.

Chen L, Cai Y, Zhou G, Shi X, Su J, Chen G, Lin K (2014) Rapid Sanger sequencing of the 16S rRNA gene for identification of some common pathogens. *PLoS One* 9: e88886.

Di Resta C, Galbiati S, Carrera P, Ferrari M (2018) Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. *EJIFCC* 29: 4-14.

Faita F, Vecoli C, Foffa I, Andreassi MG (2012) Next generation sequencing in cardiovascular diseases. *World J Cardiol* 4: 288-295.

Florescu C, Rogoveanu I, Vere CC, Târtea GC, Târtea EA, Mogoantă L (2016)From molecular mechanism to morphological changes in cardiomyopathy.*RomJMorpholEmbryol* 57: 1207-1214.

Goh G, Choi M(2012) Application of whole exome sequencing to identify disease-causing variants in inherited human diseases. *Genomics Inform* 10: 214-219.

Golbus JR, Puckelwartz MJ, Dellefave-Castillo L, Fahrenbach JP, Nelakuditi V, Pesce LL, Pytel P, McNally EM (2014) Targeted analysis of whole genome sequence data to diagnose genetic cardiomyopathy. *Circ Cardiovasc Genet* 7: 751-759.

Haskell GT, Adams MC, Fan Z, Amin K, Guzman Badillo RJ, Zhou L, Bizon C, Chahin N, Greenwood RS, Milko LV, Shiloh-Malawsky Y, Crooks KR, Strande N, Tennison M, Tilley CR, Brandt A, Wilhelmsen KC, Weck K, Evans JP, Berg JS (2018) Diagnostic utility of exome sequencing in the evaluation of neuromuscular disorders. *Neurol Genet* 4: e212.

McKoy G, Protonotarios N, Crosby A, Tsatsopoulou A, Anastasakis A, Coonar A, Norman M, Baboonian C,

Jeffery S, McKenna WJ (2000) Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease). *Lancet* 355: 2119-2124.

Morini E, Sangiuolo F, Caporossi D, Novelli G, Amati F (2015) Application of next generation sequencing for personalized medicine for sudden cardiac death. *Front Genet* 6: 55.

Norgett EE, Hatsell SJ, Carvajal-Huerta L, Cabezas JC, Common J, Purkis PE, Whittock N, Leigh IM, Stevens HP, Kelsell DP (2000) Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. *Hum Mol Genet* 9: 2761-2766.

Qiao D, Ameli A, Prokopenko D, Chen H, Kho AT, Parker MM, Morrow J, Hobbs BD, Liu Y, Beaty TH, Crapo JD, Barnes KC, Nickerson DA, Bamshad M, Hersh CP, Lomas DA, Agusti A, Make BJ, Calverley PMA, Donner CF, Wouters EF, Vestbo J, Paré PD, Levy RD, Rennard SI, Tal-Singer R, Spitz MR, Sharma A, Ruczinski I, Lange C, Silverman EK, Cho MH (2018) Whole exome sequencing analysis in severe chronic obstructive pulmonary disease. *Hum Mol Genet* 27: 3801-3812.

Simpson S, Rutland P, Rutland CS (2017) Genomic insights into cardiomyopathies: a comparative cross-species review. *Vet Sci* 4: 19.

Sisakian H (2014) Cardiomyopathies: Evolution of pathogenesis concepts and potential for new therapies. *World J Cardiol* 6: 478-494.

WHOLE EXOME SEQUENCING IDENTIFIED A NOVEL *MYOPALLADIN* GENE MUTATION IN A CARDIOMYOPATHY PATIENT

Bui Chi Bao[1,2,3], Nguyen Minh Hiep[4], Nguyen Manh Cong[5], Pham Thi Thu Trang[6], Luong Thi Tham[6], Ha Thi Thanh Nga[6], Vu Bao Quoc[7], Pham Ho Thuat Khoa[7], Nguyen Thi Huynh Nga[7]

[1]*School of Medicine, Vietnam National University, Hochiminh City*
[2]*The Center for Molecular Biomedicine, University of Medicine and Pharmacy, Hochiminh City*
[3]*Department of Molecular Genetics, Children's Hospital, Hochiminh City*
[4]*Radiation Technology Center, Nuclear Research Institute, Dalat City*
[5]*Department of Post-graduate, Dalat University, Dalat City*
[6]*Functional Genomics Center, DNA Medical Technology Company, Hochiminh City*
[7]*Department of Biology, Dalat University, Dalat City*

SUMMARY

Cardiomyopathies (CMs) are a heterogenous group of disorders that affects the heart muscle. In cardiomyopathies, phenotypic overlapping among the inherited cardiovascular diseases (CVDs) limits the ability to establish a diagnosis based solely on clinical features. Here, we developed a next generation sequencing (NGS) assay to analyze a panel of 142 known cardiomyopathy genes in 9 Vietnamese patients from Children Hospital 2, Hochiminh City and Medical University Hospital, Hochiminh City, Vietnam. Whole exome sequencing (WES) - a technique which determines the variations of all coding regions (exons) of the known genes - validated a total of 65 rare variants in 18 cardiomyopathy genes among the studied Vietnamese unrelated patients. Of 65 variants identified, 28 variants were homozygous and the other 37 ones were heterozygous. Among the 65 variants, *TTN* gene variants accounted the most for 13 mutations, which are known to be benign. Other groups of 9 and 8 mutations belong to *SYNE1* and *MYPN* genes, respectively. Ten out of 65 mutations distributed equally to *NDUFV2* and *SCN5A* gene variants. We detected 6 and 4 variants for *SYNE2* and *COX15* genes, respectively. Each gene of *DMD*, *KCNE1, NEBL* and *RBM20* has 2 variants. A single variant was detected for *AKAP9*, *CAV3*, *DSC2*, *DSG2*, *DSP*, *MYBPC3* and *MYH6* genes. Especially, among them, we found a novel heterozygous nonsynonymous mutation c.1527C>G on the *MYPN* gene. These genetic results support the "pan-cardiomyopathy panel" approach, by which the molecular diagnosis of cardiomyopathies, early identification of arrhythmia development and better clinical management of cardiomyopathic patients are applied.

Keywords: *Cardiomyopathy, Gene variant, Next generation sequencing (NGS), Mutation, Whole exome sequencing (WES)*