

SỰ ĐA DẠNG DI TRUYỀN VÙNG HV2 HỆ GEN TY THỂ CỦA MỘT SỐ NHÓM NGƯỜI VIỆT

**Đỗ Mạnh Hưng, Nguyễn Hải Hà, Phạm Nhật Khôi, Vũ Phương Nhung,
Nguyễn Văn Phòng, Nguyễn Thùy Dương, Nông Văn Hải, Nguyễn Đăng Tôn***

Viện Nghiên cứu hệ gen, Viện Hàn lâm KH và CN Việt Nam, *dtnghuyen@igr.ac.vn

TÓM TẮT: Trong vùng điều khiển D-loop của DNA ty thể, các điểm đa hình được phát hiện nhiều nhất trong các trình tự siêu biến không mã hóa 1 (HV1) và 2 (HV2). Do đó, trình tự HV1 và HV2 đã được sử dụng nhiều trong các nghiên cứu phát sinh chủng loại và khoa học pháp y. Trong nghiên cứu này, chúng tôi tiến hành giải trình tự trực tiếp và phân tích vùng siêu biến HV2 trên hệ gen ty thể của các cá thể người dân tộc Kinh, Mường, Jarai và Ê-đê thuộc hai ngữ hệ Nam-Á và Nam-Đảo, nhằm tìm hiểu sự đa dạng di truyền của một số nhóm người dân tộc thuộc hai ngữ hệ trên. Nguyên liệu nghiên cứu là mẫu máu của 169 người khỏe mạnh thuộc bốn dân tộc Kinh, Mường, Ê-đê và Jarai. Vùng HV2 được giải trình tự bằng phương pháp Sanger và được so sánh với trình tự chuẩn rCRS. Kết quả cho thấy, các mẫu nghiên cứu thuộc 79 haplogroup khác nhau, phần lớn thuộc về 3 haplogroup R, B và F. Các cá thể người Kinh và người Mường có sự đa dạng hơn về thành phần haplogroup so với người Jarai và người Ê-đê. Cả 4 nhóm cá thể đều có sự tương đồng di truyền với các quần thể người đang sinh sống trong khu vực Đông Nam Á và Đông Á.

Từ khóa: HV2, D-loop, hệ gen ty thể, haplogroup, ngữ hệ.

MỞ ĐẦU

Năm mươi tư dân tộc ở Việt Nam được chia vào 8 nhóm ngôn ngữ thuộc 5 ngữ hệ gồm: Nam-Á, Nam Đảo, Thái-Kadai, Hán-Tạng và Hmông-Đao [9]. Nam-Á là ngữ hệ phổ biến nhất ở Việt Nam với nhiều dân tộc nhất, trong đó tiêu biểu có người Kinh và người Mường. Trong khi các ngữ hệ Nam Á, Thái-Kadai, Hán-Tạng, Hmông-Đao gồm những dân tộc bản địa đã sinh sống lâu đời, ngữ hệ Nam Đảo có lịch sử đến sống ở Việt Nam muộn hơn. Theo giả thuyết “Out of Taiwan”, ngữ hệ Nam Đảo được hình thành khoảng 5.000 năm trước đã di cư và phân bố rộng rãi tại các hải đảo ở Đông Nam Á và Thái Bình Dương, Madagascar [4]. Tên ngữ hệ Nam Đảo để chỉ những tộc người sống chủ yếu trên các đảo và quần đảo phía Nam châu Á. Tuy nhiên, có một nhóm đã di cư vào đất liền một vài thế kỷ trước công nguyên trong đó có tổ tiên của người Jarai và người Ê-đê.

Có nhiều giả thuyết cho rằng người Kinh và người Mường có cùng nguồn gốc, sau này, các nhóm người sinh sống ở vùng trung du và miền núi tách ra thành một tộc riêng là người Mường. Trong khi đó, nhóm người sinh sống ở vùng đồng bằng tiếp tục bị đồng hóa bởi người

Hán là người Kinh [24]. So với người Kinh và người Mường, người Jarai và người Ê-đê sinh sống khá tách biệt trên các vùng cao nguyên Nam Trung Bộ và không có những sự di cư lớn cũng như sự đồng hóa của các dân tộc lân cận.

DNA ty thể có tốc độ tiến hóa nhanh [6], không xảy ra hiện tượng tái tổ hợp, di truyền theo dòng mẹ và số lượng bản sao lớn [2], vì vậy DNA ty thể một công cụ hữu hiệu trong nghiên cứu di truyền và tiến hóa người [12]. Các loại ty thể khác nhau thuộc các nhóm đơn bội (haplogroup) khác nhau dựa trên trình tự đặc trưng của vùng điều khiển D-loop. Trình tự HV1 và HV2 thuộc vùng điều khiển có tần số đột biến cao và nhiều điểm đa hình nên được tập trung nghiên cứu nhiều hơn cả [26]. Nhằm tìm hiểu về đa dạng di truyền của một số nhóm cá thể người dân tộc thuộc hai ngữ hệ Nam Á và Nam Đảo ở Việt Nam, chúng tôi tiến hành thu mẫu và phân tích trình tự vùng siêu biến HV2 trên DNA ty thể của các cá thể người thuộc 4 dân tộc Kinh, Mường, Jarai và Ê-đê. Đồng thời, nghiên cứu này cũng so sánh cấu trúc di truyền của 4 dân tộc nói trên với các quần thể người khác trong khu vực lân cận.

VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Nghiên cứu được tiến hành với mẫu máu ngoại vi của 169 người khỏe mạnh thuộc bốn dân tộc dân tộc Kinh (40 cá thể), Mường (47 cá thể), Ê-đê (34 cá thể) và Jarai (48 cá thể). Nguồn gốc dân tộc được xác định dựa trên thông tin khai báo tự nguyện về 3 đời trước, các đối tượng nghiên cứu bao gồm cả nam và nữ có độ tuổi từ 18 đến 50.

Bộ kit xác định trình tự nucleotide (BigDye v3.1 terminator) (Applied Biosystem) và các hóa chất cho phản ứng PCR (dNTPs, Taq DNA polymerase, MgCl₂ ...) của Fermentas.

Cặp mồi được sử dụng để khuếch đại trình tự HV2 trong nghiên cứu này có trình tự như sau: HV2F: 5'- GGT CTA TCA CCC TAT TAA CCA C -3' và HV2R: 5'- CTG TTA AAA GTG CAT ACC GCC A -3'.

DNA tổng số được tách chiết theo phương pháp của Sambrook & Rusell (2001) [21]. Phản ứng PCR khuếch đại trình tự HV2 được tiến hành với thể tích là 25 µl gồm các thành phần: 50 ng DNA, 1X đệm PCR, 8 mM MgCl₂, 700 mM mỗi dNTP, 5 mM mồi (HV2F và HV2R) và 1 U DreamTaq (Fermentas). Quá trình khuếch đại được thực hiện trên máy Veriti™ Dx 96-Well Thermal Cycler (ABI) với chu trình nhiệt: 95°C, 3 phút; 35 chu kỳ (95°C, 45 giây; 58°C, 1 phút ; 72°C, 1 phút); 72°C, 10 phút, giữ ở 4°C. Sản phẩm PCR được tinh sạch bằng GeneJET PCR Purification Kit (Thermo Scientific) theo phương pháp của nhà sản xuất. Trình tự vùng HV2 được xác định trên máy giải trình tự tự động ABI 3100 Genetic Analyzer sử dụng bộ kit BigDye Terminator v3.1 (Applied Biosystem).

Các số liệu về trình tự vùng siêu biến HV2 của tất cả các mẫu nghiên cứu được kiểm tra và loại bỏ các trình tự lỗi bằng các phần mềm DNA Sequencing Analysis v5.3.1 và SeqScape® v2.6 (ABI). Sau đó các trình tự này được sắp xếp (aligned) với nhau và so sánh với trình tự chuẩn rCRS [1] bằng công cụ Clustal W2 [17]. Các trình tự được định dạng và sắp xếp vào từng phân nhóm với phần mềm DNAsp v5 [19]. Phép tính Median-joining network được sử dụng để dựng bản đồ mạng thể hiện mối liên hệ giữa các nhóm đơn bội bằng chương

trình NETWORK4.510 [3]. Phần mềm ARLEQUIN phiên bản 3.5.1.3 [10] được sử dụng để tính toán mối tương quan phân tử trong trình tự HV2 ở các mẫu nghiên cứu so với các quần thể người khác ở Đông Nam Á và khu vực lân cận.

KẾT QUẢ VÀ THẢO LUẬN

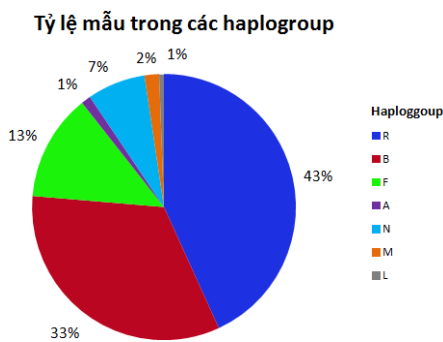
Đặc điểm trình tự HV2 của các cá thể nghiên cứu

So sánh các trình tự thu được với trình tự chuẩn của vùng HV2 (nucleotide 73-340 của DNA ty thể) trên cơ sở dữ liệu (rCRS) [1], kết quả cho thấy có 999 điểm sai khác nằm tại 52 vị trí khác nhau trên đoạn trình tự có chiều dài 267 nucleotide (tương ứng với nucleotide từ vị trí 73 - 340 trên bản đồ gen ty thể). Các trình tự này xuất hiện các điểm nóng như tại vị trí 73, 150, 152, 249 và 263, đặc biệt có 1 đoạn poly C lớn từ vị trí 303 đến 315 với 1 nucleotide T được xen vào tại vị trí 310. Số lượng đa hình ở từng mẫu không giống nhau, chủ yếu nằm trong khoảng từ 5 đến 7 đa hình trên mỗi trình tự, trong đó mẫu có nhiều đa hình nhất là 11 và mẫu có đa hình ít nhất là 3. Nhìn chung, số liệu thu được cho thấy HV2 là một trình tự có sự đa dạng lớn, phản ánh tốc độ đột biến cao của vùng này.

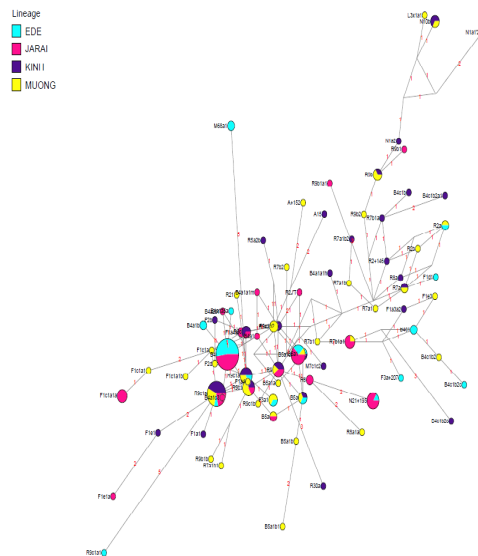
Phân nhóm đơn bội

Dựa trên đặc điểm các vị trí đa hình, 169 cá thể nghiên cứu được chia vào 79 nhóm đơn bội (haplogroup) khác nhau. Trong đó, 54 haplogroup chỉ có một cá thể duy nhất và 25 haplogroup là nhóm chung của từ 2 cá thể trở lên, lớn nhất có haplogroup là nhóm chung của 21 cá thể. Sử dụng cơ sở dữ liệu về các nhóm đơn bội trên bản đồ DNA ty thể người Haplogrep, phần lớn các haplogroup trong nghiên cứu này được định danh thuộc về các phân nhóm của 3 haplogroup là R, B và F. Cụ thể có 33 phân nhóm của haplogroup R, 20 phân nhóm của haplogroup B, 16 phân nhóm của haplogroup F, 5 phân nhóm của haplogroup N, haplogroup M và haplogroup A đều có 2 phân nhóm, và 1 phân nhóm thuộc haplogroup L. Nhóm đơn bội N (macrohaplogroup N) (bao gồm các haplogroup A, B, F, N và R) chiếm 97% các mẫu nghiên cứu bao gồm 165 mẫu thuộc 77 phân nhóm đơn bội (subhaplogroup) khác nhau.

Trong đó 43% số mẫu nằm trong haplogroup R, 33% mẫu thuộc haplogroup B, 13% mẫu thuộc haplogroup F, 7% thuộc haplogroup N và 1% thuộc haplogroup A. Trong khi đó, nhóm đơn bội M (macrohaplogroup M) (gồm các haplogroup C, D, G và M) vốn là một macrohaplogroup khá phổ biến ở khu vực Đông Nam Á chỉ phát hiện ở 3 mẫu nghiên cứu chiếm 2% và có 1 thuộc haplogroup L (hình 1). Ngoài ra, sơ đồ hình cây thể hiện phần nào mối liên quan giữa các haplogroup trong nghiên cứu cũng như quá trình hình thành và phân tách của các nhóm haplogroup khác nhau (hình 2).

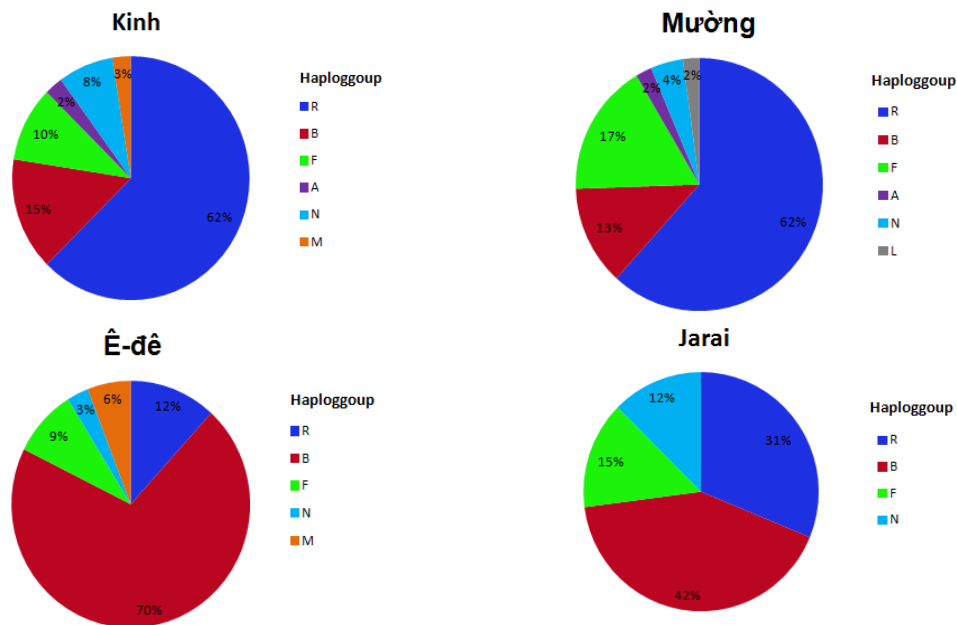


Hình 1. Biểu đồ thể hiện thành phần haplogroup của các mẫu nghiên cứu



Hình 2. Sơ đồ phân bố các haplogroup

Sơ đồ hình cây được xây dựng bằng phương pháp Median-Joining Network dựa trên trình tự HV2 (73-340) của các mẫu nghiên cứu. Các chữ số nhỏ màu đỏ thể hiện số lượng đa hình. Thành phần, số lượng các cá thể được thể hiện bằng màu và kích thước tương ứng của từng haplogroup.



Hình 3. Thành phần các haplogroup của 4 nhóm cá thể người Kinh, Mường, Jarai và Ê-đê

So sánh đặc điểm của 4 nhóm cá thể nghiên cứu

Về cơ bản, cả 4 nhóm dân tộc có thành phần chủ yếu là các nhóm đơn bội R, B, F với những đặc trưng riêng về tỷ lệ phân bố của từng haplogroup ở mỗi nhóm cá thể (hình 3). Trong các nghiên cứu về quần thể và mối quan hệ di truyền trên DNA ty thể [13, 14, 18], ngoài các haplogroup đặc trưng như R, B, F, người Kinh có một tỷ lệ khá lớn của haplogroup M [7, 27]. Tuy nhiên, trong nghiên cứu này, chúng tôi nhận thấy haplogroup M chỉ chiếm 3% số mẫu cá thể người Kinh.

Kết quả nghiên cứu cũng cho thấy, người Kinh và người Mường có sự tương đồng rất lớn về thành phần các haplogroup. Cả hai nhóm cá thể đều có 6 nhóm haplogroup khác nhau và có tỷ lệ haplogroup R là 62%, hai nhóm đơn bội B và F chiếm tỉ lệ nhỏ hơn ở hai nhóm người này (Hình 3). Kinh và Mường là hai dân tộc lớn ở Việt Nam thuộc ngữ hệ Nam Á. Kết quả nghiên cứu này cho thấy có khả năng mối quan hệ giữa hai dân tộc này khá gần gũi. Sự đa dạng về các phân nhóm đơn bội của người Kinh và người Mường cũng là đặc trưng của các quần thể người sinh sống trên phần lục địa của khu vực Đông Nam Á và Đông Á [5, 8, 22].

Ở người Jarai và người Ê-đê có sự đa dạng thấp hơn về thành phần các haplogroup. Khác với người Kinh và người Mường, hai dân tộc này có sự xuất hiện chủ yếu của haplogroup B. Ở các cá thể người Ê-đê, haplogroup B chiếm 70% trong khi ở người Jarai là 42%. Sự phổ biến của haplogroup B là đặc trưng của những nhóm người thuộc ngữ hệ Nam Đảo [23, 25]. Theo các giả thuyết cũng như những quan điểm về ngôn ngữ học, người Jarai và Ê-đê có nguồn gốc từ những người Mã Lai cổ di cư vào lục địa sinh sống một vài thế kỷ trước công nguyên [9]. Với chủ yếu các haplogroup nhóm B, hai dân tộc này đã lưu giữ đặc trưng cơ bản của các dân tộc trong ngữ hệ Nam Đảo.

So sánh di truyền với 13 quần thể người sinh sống trong khu vực lân cận

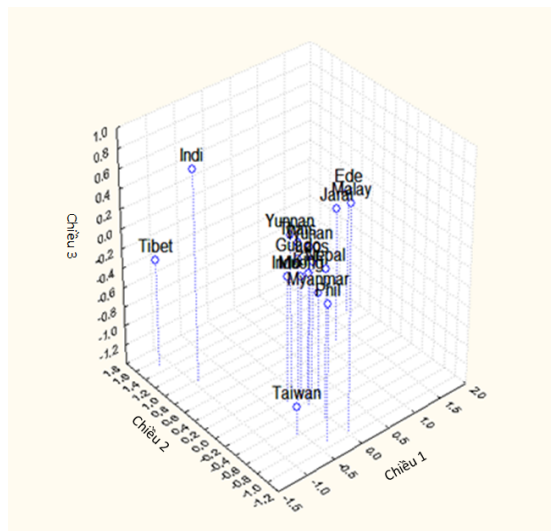
Với mục đích làm sáng tỏ về cấu trúc di truyền của các mẫu nghiên cứu, chúng tôi thực hiện các phân tích tương đồng ở 838 trình tự thuộc 13 quần thể người khác ở Đông Nam Á

và Nam Á trên đoạn trình tự HV2 (nucleotide 73 đến 340). Dữ liệu 838 trình tự HV2 của 13 quần thể người sinh sống trong khu vực Đông Nam Á và vùng lân cận dùng để so sánh trong nghiên cứu này được lấy từ các công bố trước đây gồm có người Đài Loan [25], Philipin [23], Malaysia [15], Indonesia [20], Quảng Đông, Vân Nam, Vũ Hán của Trung Quốc [27], Ấn Độ [7], Lào [5], Tây Tạng [16], Nepal [11], Myanmar [22] và Thái Lan [14]. Trình tự tham chiếu rCRS (Revised Cambridge Reference Sequence) (Genbank ID NC_012920) được lấy từ trang cơ sở dữ liệu của ty thể [1].

Mô hình thống kê phân tích phương sai phân tử (Analysis of molecular variance - AMOVA) được sử dụng để kiểm tra sự thay đổi trong cấu trúc di truyền trình tự HV2 giữa các quần thể. Giá trị của hệ số thống kê Fst (F-Statistics) thấp và tương tự có thể phản ánh nguồn gốc chung của các quần thể. Giá trị Fst giữa người Kinh và người Mường cũng như giữa hai nhóm cá thể này với các quần thể người trên lục địa Đông Nam Á và Đông Á thấp hơn hẳn so với khoảng cách với nhóm các quần thể ở khu vực Nam Á và các quần đảo phía nam. Tuy nhiên, ở người Jarai và Ê-đê lại cho thấy một giá trị khoảng cách thấp giữa hai dân tộc với nhau nhưng giá trị giữa hai dân tộc với các quần thể khác đều ở mức trung bình.

Biểu đồ không gian 3 chiều hiển thị kết quả phân tích phương sai phân tử AMOVA được mô tả ở hình 4 cho thấy, người Kinh và người Mường nằm trong một khu vực tập trung của các quần thể người trong khu vực Đông Nam Á và Đông Á, trong khi người Ê-đê và người Jarai có một khoảng cách nhất định với nhóm này (hình 4). Tương đồng với kết quả đưa ra ở trên với đặc trưng của các nhóm đơn bội trong 4 nhóm cá thể dân tộc, người Kinh và người Mường cho thấy một mối liên hệ gần gũi với các quần thể người người sinh sống trong khu vực lân cận. Điều này có thể là sự kết hợp giữa hai yếu tố đó là nguồn gốc và địa chính trị. Nguồn gốc gần gũi, địa bàn sinh sống cùng với những biến động trong lịch sử, khiến cho sự hòa huyết trở nên phổ biến tạo thành đặc trưng di truyền chung của các dân tộc trong cả một vùng rộng lớn. Trong khi đó, người Jarai và Ê-đê với lịch sử sinh sống tách biệt trên cao nguyên đã

tạo ra những đặc trưng riêng cũng như khoảng cách nhất định về di truyền với các dân tộc trong khu vực lân cận.



Hình 4. Phân tích đa chiều của khoảng cách di truyền giữa 17 quần thể người ở Đông Nam Á và khu vực lân cận

KẾT LUẬN

HV2 là một trình tự có tính đa hình rất lớn và có ý nghĩa nhất định với việc nghiên cứu di truyền của các quần thể người. Các mẫu nghiên cứu phần lớn thuộc về 3 haplogroup R, B và F, trong đó các cá thể người Kinh và người Mường có sự đa dạng hơn về thành phần haplogroup so với người Jarai và người Ê-đê. Cả 4 nhóm cá thể đều có sự tương đồng di truyền với các quần thể người đang sinh sống trong khu vực Đông Nam Á và Đông Á. Tuy nhiên, để có thể đưa ra những kết luận chính xác hơn, cần phải tiến hành phân tích với số lượng mẫu lớn hơn trên những trình tự dài hơn của DNA ty thể.

Lời cảm ơn: Công trình được tài trợ một phần từ đề tài độc lập cấp quốc gia mã số ĐTĐL.CN-05/15. Chúng tôi xin chân thành cảm ơn các đối tượng đã đồng ý tham gia và cung cấp mẫu máu để thực hiện nghiên cứu này.

TÀI LIỆU THAM KHẢO

1. Andrews R. M., Kubacka I., Chinnery P. F., Lightowlers R. N., Turnbull D. M., Howell

- N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*, 23(2): 147.
2. Avise J. C., Arnold J., Ball R. M., Bermingham E., Lamb T., Neigel J. E., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–522.
3. Bandelt H. J., Forster P., Sykes B. C., Richards M. B., 1995. Mitochondrial portraits of human populations using median networks. *Genetics*, 141(2): 743–753.
4. Bellwood P., 2006. Austronesian prehistory in southeast Asia: homeland, expansion and transformation. In: Bellwood P, Fox JJ, Tryon D, editors. *The austronesians: historical and comparative perspectives*. canberra (act): anu e press. p. 103–114.
5. Bodner M., Zimmermann B., Rock A., Kloss-Brandstatter A., Horst D., Horst B., Sengchanh S., Sanguanserm Sri T., Horst J., Kramer T., Schneider P. M., Parson W., 2011. Southeast Asian diversity: first insights into the complex mtDNA structure of Laos. *BMC Evol Biol*, 1149.
6. Brown W. M., George M., Jr., Wilson A. C., 1979. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*, 76(4): 1967–1971.
7. Chaubey G., Karmin M., Metspalu E., Metspalu M., Selvi-Rani D., Singh V. K., Parik J., Solnik A., Naidu B. P., Kumar A., Adarsh N., Mallick C. B., Trivedi B., Prakash S., Reddy R., Shukla P., Bhagat S., Verma S., Vasnik S., Khan I., Barwa A., Sahoo D., Sharma A., Rashid M., Chandra V., Reddy A. G., Torroni A., Foley R. A., Thangaraj K., Singh L., Kivisild T., Villems R., 2008. Phylogeography of mtDNA haplogroup R7 in the Indian peninsula. *BMC Evol Biol*, 8227.
8. Chen F., Wang S. Y., Zhang R. Z., Hu Y. H., Gao G. F., Liu Y. H., Kong Q. P., 2008. Analysis of mitochondrial DNA polymorphisms in Guangdong Han Chinese.

- Forensic Sci Int Genet, 2(2): 150-153.
9. Dang N. V., Chu T. S., Luu H., 2010. Ethnic Minorities in Vietnam. The Gioi Publisher
 10. Excoffier L., Lischer H. E., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*, 10(3): 564-567.
 11. Fornarino S., Pala M., Battaglia V., Maranta R., Achilli A., Modiano G., Torroni A., Semino O., Santachiara-Benerecetti S. A., 2009. Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation. *BMC Evol Biol*, 9:154.
 12. Ingman M., Kaessmann H., Paabo S., Gyllensten U., 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*, 408(6813): 708-713.
 13. Irwin J. A., Saunier J. L., Strouss K. M., Diegoli T. M., Sturk K. A., O'Callaghan J. E., Paintner C. D., Hohoff C., Brinkmann B., Parsons T. J., 2008. Mitochondrial control region sequences from a Vietnamese population sample. *Int J Legal Med*, 122(3): 257-259.
 14. Jin H. J., Tyler-Smith C., Kim W., 2009. The peopling of Korea revealed by analyses of mitochondrial DNA and Y-chromosomal markers. *PLoS One*, 4(1): e4210.
 15. Jinam T. A., Hong L. C., Phipps M. E., Stoneking M., Ameen M., Edo J., Saitou N., 2012. Evolutionary history of continental southeast Asians: "early train" hypothesis based on genetic analysis of mitochondrial and autosomal DNA data. *Mol Biol Evol*, 29(11): 3513-3527.
 16. Kang L., Zheng H. X., Chen F., Yan S., Liu K., Qin Z., Liu L., Zhao Z., Li L., Wang X., He Y., Jin L., 2013. mtDNA lineage expansions in Sherpa population suggest adaptive evolution in Tibetan highlands. *Mol Biol Evol*, 30(12): 2579-2587.
 17. Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan P. A., McWilliam H., Valentin F., Wallace I. M., Wilm A., Lopez R., Thompson J. D., Gibson T. J., Higgins D. G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21): 2947-2948.
 18. Li H., Cai X., Winograd-Cort E. R., Wen B., Cheng X., Qin Z., Liu W., Liu Y., Pan S., Qian J., Tan C. C., Jin L., 2007. Mitochondrial DNA diversity and population differentiation in southern East Asia. *Am J Phys Anthropol*, 134(4): 481-488.
 19. Librado P., Rozas J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11): 1451-1452.
 20. Redd A. J., Stoneking M., 1999. Peopling of Sahul: mtDNA variation in aboriginal Australian and Papua New Guinean populations. *Am J Hum Genet*, 65(3): 808-828.
 21. Sambrook J., Russell D. W., 2001. *Molecular Cloning*. Cold Spring Harbor Laboratory Press, New York, USA.
 22. Summerer M., Horst J., Erhart G., Weissensteiner H., Schonherr S., Pacher D., Forer L., Horst D., Manhart A., Horst B., Sanguansermsri T., Kloss-Brandstatter A., 2014. Large-scale mitochondrial DNA analysis in Southeast Asia reveals evolutionary effects of cultural isolation in the multi-ethnic population of Myanmar. *BMC Evol Biol*, 14:17.
 23. Tabbada K. A., Trejaut J., Loo J. H., Chen Y. M., Lin M., Mirazon-Lahr M., Kivisild T., De Ungria M. C., 2010. Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? *Mol Biol Evol*, 27(1): 21-31.
 24. Taylor K. W., 1983. *The Birth of Vietnam*. University of California Press, Berkeley, California.
 25. Trejaut J. A., Kivisild T., Loo J. H., Lee C. L., He C. L., Hsu C. J., Lee Z. Y., Lin M., 2005. Traces of archaic mitochondrial lineages persist in Austronesian-speaking Formosan populations. *PLoS Biol*, 3(8): e247.

26. Vigilant L., Pennington R., Harpending H., Kocher T. D., Wilson A. C., 1989. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc Natl Acad Sci U S A*, 86(23): 9350-9354.
27. Yao Y. G., Kong Q. P., Bandelt H. J., Kivisild T., Zhang Y. P., 2002. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet*, 70(3): 635-651.

GENETIC VARIATION OF MITOCHONDRIAL SEQUENCE-HV2 IN VIETNAMESE POPULATIONS

**Do Manh Hung, Nguyen Hai Ha, Pham Nhat Khoi, Vu Phuong Nhung,
Nguyen Van Phong, Nguyen Thuy Duong, Nong Van Hai, Nguyen Dang Ton**

Institute of Genome Research, VAST

SUMMARY

HV1 and HV2, which located in control D-loop region of mitochondrial, are the hyper-variable parts of mitochondrial DNA (mtDNA). Therefore, these two sequences were often used for evolutionary and forensic science. In order to determine the genetic variation of populations belong to Nam-A and Nam-Dao language systems, we sequenced and analyzed the mtDNA HV2 region of individuals from four ethnic groups: Kinh, Muong, Ede and Jarai which represent for those language systems. Study subjects are peripheral blood of 169 individuals belonging to four populations: Kinh, Muong, Ede and Jarai. HV2 regions were sequenced by Sanger sequencing. As a result, the analyzed mitochondrial control region sequences could be assigned to 79 different haplogroups, with the dominant proportion of three haplogroups: R, B and F. Donors representing Kinh or Muong population revealed higher haplogroup composition variation in comparison with those affiliated with Jarai or Ede. In addition, it was indicated that the 4 populations shared the genetic analogy with other populations inhabited in Southeast Asia and South Asia.

Keywords: D-loop, HV2, haplogroup, language system, mitochondrial DNA.

Ngày nhận bài: 21-9-2015