

Khảo sát sự đa dạng di truyền vùng D-loop hệ gen ty thể ở ba nhóm tộc người: Hà Nhì, Phù Lá, Si La

Nguyễn Văn Phòng^{1,2}, Nguyễn Thy Ngọc^{1,3}, Nguyễn Doãn Tình¹,
Nguyễn Thùy Dương^{1,2}, Nông Văn Hải^{1,2*}

¹Viện Nghiên cứu Hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

²Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

³Trường Đại học Khoa học và Công nghệ Hà Nội, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Ngày nhận bài 25/3/2019; ngày chuyển phản biện 28/3/2019; ngày nhận phản biện 24/4/2019; ngày chấp nhận đăng 29/4/2019

Tóm tắt:

Nghiên cứu này tập trung xác định các đa hình nucleotide thuộc vùng điều khiển D-loop, hệ gen ty thể của các cá thể thuộc ba tộc người cùng nhóm ngữ hệ Tạng Miến, sinh sống ở địa bàn hai tỉnh Điện Biên, Lai Châu là Hà Nhì, Phù Lá và Si La. Bằng phương pháp giải trình tự gen Sanger đã xác định được 91 điểm đa hình khác biệt so với trình tự hệ gen ty thể tham chiếu chuẩn rCRS mã số NC_012920.1. Phân tích thống kê bằng kiểm định Fisher exact (2 phía) cho thấy có 14 đa hình có sự khác biệt có ý nghĩa thống kê ($p < 0,05$) giữa các tộc người, trong đó có 4 đa hình chỉ xuất hiện ở tộc người Phù Lá gồm T16086C, C16167T, A16203G và A16318G. Các phân tích sâu hơn về khoảng cách di truyền cho thấy khoảng cách di truyền trung bình giữa các cá thể thuộc các tộc người Hà Nhì, Phù Lá, Si La lần lượt là 0,01; 0,01 và 0,008. Khoảng cách di truyền giữa các nhóm cá thể thuộc tộc người Hà Nhì - Phù Lá, Hà Nhì - Si La và Phù Lá - Si La lần lượt là 0,011; 0,0105 và 0,0106. Những số liệu này sẽ là cơ sở dữ liệu góp phần vào nghiên cứu đa dạng di truyền và nguồn gốc lịch sử của các tộc người Việt Nam.

Từ khóa: D-loop, đa dạng di truyền, Hà Nhì, hệ gen ty thể, Phù Lá, Si La.

Chỉ số phân loại: 1.6

Mở đầu

Trong 54 tộc người anh em Việt Nam, có những tộc người có nguồn gốc bản địa đã làm ăn sinh sống từ hàng ngàn năm trước, lại có những tộc người có nguồn gốc từ các nước láng giềng di cư đến giao lưu buôn bán và trở thành cộng đồng các tộc người người Việt. Hiện nay, có nhiều luồng ý kiến khác nhau về nguồn gốc của các tộc người thiểu số Việt Nam, đặc biệt là đồng bào các tộc người ở vùng núi Tây Bắc. Với ưu điểm cho kết quả nhanh, chính xác, đáng tin cậy, phương pháp nghiên cứu đa hình vùng siêu biên (D-loop), thuộc hệ gen ty thể đã được sử dụng rộng rãi trong các nghiên cứu về nguồn gốc phát sinh chủng loại ở người [1]. Nhiều công trình nghiên cứu đã khảo sát đa hình vùng D-loop thuộc hệ gen ty thể trên các tộc người khác nhau của Việt Nam như: Kinh, Mường [2], Gia rai, Ê đê [3], Tày [4], Mảng [5], tuy nhiên hiện chưa có công trình nào nghiên cứu về hệ gen ty thể của các tộc người Hà Nhì, Phù Lá và Si La ở Việt Nam cũng như trên thế giới. Theo thống kê của Ban chỉ đạo Tổng điều tra dân số và nhà ở trung ương năm 2009, ở Việt Nam tộc người Hà Nhì có dân số khoảng 21.725 người, cư trú tập trung ở các tỉnh Lai Châu, Lào Cai và rải rác trên các tỉnh/thành phố khác. Ngoài ra, người Hà Nhì còn sinh

sống ở phía nam Trung Quốc (khoảng 570.000 người) và Lào. Tộc người Phù Lá có dân số 10.944 người, sinh sống tập trung tại các tỉnh Lào Cai, Yên Bái, Hà Giang và Lai Châu của Việt Nam và rải rác khoảng 4.200 người ở các tỉnh biên giới phía nam Trung Quốc. Tộc người Si La có khoảng 709 người, sinh sống khá tách biệt tại huyện Mường Tè, tỉnh Lai Châu và một nhóm khoảng 1.800 người sinh sống tại Lào [6]. Đây là ba tộc người thiểu số cùng sinh sống trên địa bàn các tỉnh miền núi Tây Bắc, cùng thuộc nhóm ngữ hệ Tạng - Miến, và có nhiều nét tương đồng trong văn hóa, phong tục tập quán. Trong công trình này, chúng tôi sử dụng phương pháp giải trình tự Sanger nhằm xác định các đa hình thuộc vùng D-loop hệ gen ty thể của 99 người thuộc ba tộc người thiểu số Hà Nhì (33 cá thể), Phù Lá (35 cá thể) và Si La (31 cá thể) để so sánh mức độ tương đồng về vật chất di truyền của các tộc người này.

Đối tượng và phương pháp

Đối tượng nghiên cứu

Mẫu máu toàn phần thu để tách ADN được lấy từ ba quần thể người tộc người Hà Nhì, Phù Lá, Si La ở địa bàn hai tỉnh Hà Giang và Lai Châu. Các cá thể được lựa chọn

*Tác giả liên hệ. Email: vhnong@igr.ac.vn

Study of genetic polymorphisms in D-loop region of the mitochondrial genomes from three ethnic groups: Ha Nhi, Phu La, Si La

Van Phong Nguyen^{1,2}, Thy Ngoc Nguyen^{1,3},
Doan Tinh Nguyen¹, Thuy Duong Nguyen^{1,2},
Van Hai Nong^{1,2*}

¹Institute of Genome Research,
Vietnam Academy of Science and Technology
²Graduate University of Science and Technology,
Vietnam Academy of Science and Technology
³University of Science and Technology of Hanoi,
Vietnam Academy of Science and Technology

Received 25 March 2019; accepted 29 April 2019

Abstract:

Our study aimed to identify the genetic polymorphism in the D-loop region of human mitochondrial genome of three ethnic minority groups namely Ha Nhi, Phu La and Si La. These ethnic minority groups belonging to the Sino-Tibetan language family live in Northwestern Vietnam provinces: Dien Bien and Lai Chau. Using the Sanger sequencing, we identified 91 variants in the mtDNA of individuals from three ethnic groups compared to the reference mtDNA rCRS NC_012920.1. Statistical analysis was performed using the Fisher exact test (2 sides). The result showed that 14 genetic variants were significantly different among three ethnic groups ($p < 0.05$). Among them, four genetic polymorphisms were detected only in the Phu La group, including T16086C, C16167T, A16203G, and A16318G. Further analysis indicated that the mean genetic distances between individuals in each ethnic group Ha Nhi, Phu La, and Si La were 0.01, 0.01, and 0.008, respectively. The genetic distances between ethnic groups were 0.011 (Ha Nhi - Phu La), 0.0105 (Ha Nhi - Si La), and 0.0106 (Phu La - Si La). These results contribute further data to the study of genetic variations and historical origins of the Vietnamese ethnic groups.

Keywords: D-loop, genetic polymorphism, Ha Nhi, mitochondrial genome, Phu La, Si La.

Classification number: 1.6

đều là người Việt Nam, khỏe mạnh, không cùng huyết thống và có bốn ông bà nội ngoại cùng thuộc một tộc người. Mọi cá thể tham gia vào nghiên cứu đều đã được thông báo cụ thể về mục đích của nghiên cứu cũng như tiến trình lấy mẫu và đồng ý tham gia cung cấp mẫu máu cho nghiên cứu này. Đề tài đã được Hội đồng đạo đức trong nghiên cứu y sinh học của Viện Nghiên cứu Hệ gen thông qua và đồng ý cho phép tiến hành theo Quyết định số 4-2015/NCHG-HĐĐĐ.

Phương pháp nghiên cứu

Phương pháp tách chiết và khuếch đại gen: ADN tổng số được tách chiết từ mẫu máu toàn phần sử dụng bộ Kit GeneJET Whole Blood Genomic ADN Purification (Thermo Fisher Scientific Inc., Vilnius, Lithuania). Phản ứng khuếch đại các đoạn gen HV-I và HV-II thuộc vùng D-loop của hệ gen ty thể có tổng thể tích 25 μ l bao gồm: 1X đệm PCR; 2,5mM dNTP; 1 μ M từng loại mỗi tương ứng; 1 unit Taq polymerase (Fermentas); 20 ng ADN tổng số. Chu trình nhiệt phản ứng khuếch đại đoạn gen: 95°C/3 phút; 40 chu kỳ của (95°C/30 giây, 56 hoặc 58°C/20 giây, 72°C/30 giây); 72°C/5 phút; giữ 4°C.

Bảng 1. Trình tự các cặp môi sử dụng để khuếch đại các đoạn gen HV-I và HV-II.

Tên môi	Trình tự đoạn môi	Nhiệt độ gắn môi (Tm)	Kích thước sản phẩm PCR
HV-I-F	5'-CCACCATTAGCACCCAAAGC-3'	58°C	693 bp
HV-I-R	5'-CAGCGTCTCGCAATGCTATC-3'		
HV-II-F	5'-GCCTAAATAGCCACACGTTTC-3'	56°C	689 bp
HV-II-R	5'-TTTATGGGGTGATGTGAGCC-3'		

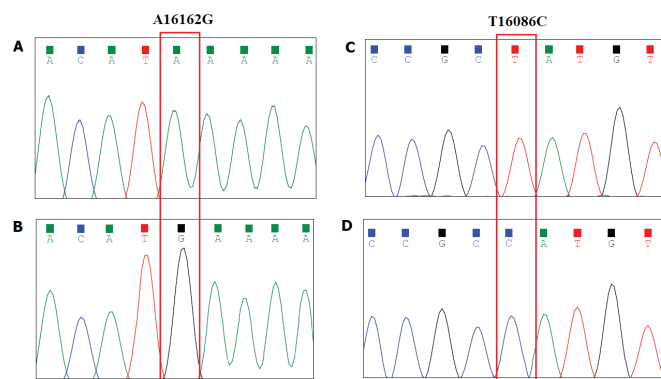
Phương pháp tinh sạch và giải trình tự: sản phẩm khuếch đại gen được tinh sạch bằng GeneJET PCR Purification Kit (Thermo Scientific) theo hướng dẫn của nhà sản xuất. Trình tự vùng D-loop hệ gen ty thể của các cá thể được xác định bằng phương pháp giải trình tự Sanger bằng Kit BigDye Terminator v3.1 (Applied Biosystems) trên máy giải trình tự ABI PRISM 3500 Genetic Analyzer. Các trình tự thu được được so sánh với trình tự ty thể tham chiếu chuẩn rCRS mã số NC_012920.1 bằng phần mềm Chromas và BioEdit.

Phương pháp phân tích thống kê: các đa hình phát hiện được trên vùng D-loop hệ gen ty thể được so sánh giữa các tộc người với nhau từng đôi một bằng kiểm định Fisher exact (2 phía) để kiểm tra xem có sự khác biệt có ý nghĩa thống kê giữa các tộc người hay không. Mức ý nghĩa được lấy với $p < 0,05$. Khoảng cách di truyền trung bình giữa các cá thể trong một tộc người và giữa các tộc người với nhau được tính bằng phần mềm MEGA X dựa trên trình tự vùng D-loop hệ gen ty thể.

Kết quả nghiên cứu và thảo luận

Xác định đa hình thuộc vùng D-loop của các cá thể thuộc ba tộc người Hà Nhì, Phù Lá, Si La

Sản phẩm phản ứng nhân gen PCR được điện di kiểm tra trên gel agarose 1% để kiểm tra kích thước. Kết quả điện di cho thấy 2 băng đặc hiệu với kích thước lần lượt 693 bp và 689 bp phù hợp với kích thước tính toán theo lý thuyết.



Hình 1. Kết quả giải trình tự xác định đa hình thuộc vùng D-loop (A. Trình tự allele 16162A của cá thể HaNhi320 thuộc tộc người Hà Nhì; B. Trình tự allele 16162G của cá thể SiLa441 thuộc tộc người Si La; C. Trình tự allele 16086T của cá thể HaNhi330 thuộc tộc người Hà Nhì; D. Trình tự allele 16086C của cá thể PhuLa604 thuộc tộc người Phù Lá).

Sản phẩm khuếch đại gen PCR sau đó được tinh sạch và giải trình tự. Kết quả đọc trình tự cho thấy tín hiệu rõ ràng, đẹp, không bị nhiễu (hình 1). Sau khi so sánh trình tự thu được với trình tự tham chiếu rCRS, chúng tôi đã phát hiện tổng cộng 91 đa hình (có trình tự khác với trình tự tham chiếu) thuộc vùng D-loop hệ gen ty thể. Số lượng điểm đa hình trên từng nhóm tộc người được liệt kê ở bảng 2.

Bảng 2. Số lượng các đa hình trung bình phát hiện được trong từng nhóm tộc người nghiên cứu.

	Tộc người		
	Hà Nhì	Phù Lá	Si La
Cỡ mẫu (người)	33	35	31
Số đa hình phát hiện	8,3±1,6	8,5±1,6	9,2±1,4
Vùng HV-I	3,9±1	4±1,4	3,8±1,3
Vùng HV-II	4,4±1,7	4,5±1,2	5,4±1,4

Sau khi có thông tin về các đa hình xuất hiện trong nghiên cứu này, chúng tôi loại bỏ các đa hình có tần số quá thấp và tiếp tục phân tích những đa hình có tần số allele hiếm (minor allele frequency - MAF) ở ít nhất một nhóm trong ba nhóm cá thể lớn hơn hoặc bằng 0,1 (10%). Kết quả còn lại 33 đa hình, trong đó có 9 đa hình thuộc vùng HV-I và 24 đa hình thuộc vùng HV-II. Thông tin và tần số xuất hiện của các đa hình được lựa chọn phân tích trên từng nhóm cá thể thuộc tộc người Hà Nhì, Phù Lá, Si La được liệt

kê trong bảng 3. Đáng chú ý, có hai điểm đa hình có tần số xuất hiện là 100% ở cả 3 tộc người trong nghiên cứu này là A73G và A263G. Ngoài ra, có 4 điểm đa hình chỉ xuất hiện ở tộc người Phù Lá mà không thấy xuất hiện trên các cá thể thuộc hai tộc người còn lại là 16086C, 16140C, 16203G và 16318G.

Bảng 3. Tần suất các đa hình trình tự vùng siêu biến D-loop thuộc hệ gen ty thể ở ba nhóm cá thể thuộc các tộc người Hà Nhì, Phù Lá và Si La.

Đa hình	Vùng ADN	Tộc người (N, %)*		
		Hà Nhì	Phù Lá	Si La
A73G	HV-I	33 (100)	35 (100)	31 (100)
C150T	HV-I	6 (18,18)	7 (20)	9 (29,03)
T152C	HV-I	7 (21,21)	8 (22,86)	1 (3,23)
T199C	HV-I	4 (12,12)	5 (14,29)	7 (22,58)
T204C	HV-I	1 (3,03)	4 (11,43)	0 (0)
G207A	HV-I	4 (12,12)	4 (11,43)	3 (9,68)
A263G	HV-I	33 (100)	35 (100)	31 (100)
T310C	HV-I	21 (63,64)	31 (88,57)	24 (77,42)
T489C	HV-I	20 (60,61)	11 (31,43)	11 (35,48)
T16086C	HV-II	0 (0)	13 (37,14)	0 (0)
T16093C	HV-II	6 (18,18)	2 (5,71)	2 (6,45)
C16108T	HV-II	4 (12,12)	1 (2,86)	17 (54,84)
G16129A	HV-II	10 (30,3)	6 (17,14)	21 (67,74)
A16162G	HV-II	4 (12,12)	1 (2,86)	17 (54,84)
C16167T	HV-II	0 (0)	10 (28,57)	0 (0)
T16172C	HV-II	8 (24,24)	2 (5,71)	20 (64,52)
A16182C	HV-II	4 (12,12)	0 (0)	7 (22,58)
A16183C	HV-II	6 (18,18)	5 (14,29)	7 (22,58)
T16189C	HV-II	9 (27,27)	9 (25,71)	10 (32,26)
A16203G	HV-II	0 (0)	10 (28,57)	0 (0)
T16217C	HV-II	3 (9,09)	2 (5,71)	3 (9,68)
C16223T	HV-II	24 (72,73)	17 (48,57)	11 (35,48)
C16261T	HV-II	2 (6,06)	0 (0)	3 (9,68)
C16290T	HV-II	6 (18,18)	6 (17,14)	0 (0)
T16297C	HV-II	0 (0)	4 (11,43)	3 (9,68)
T16298C	HV-II	9 (27,27)	4 (11,43)	0 (0)
T16304C	HV-II	5 (15,15)	13 (37,14)	17 (54,84)
T16311C	HV-II	9 (27,27)	4 (11,43)	0 (0)
A16318G	HV-II	0 (0)	10 (28,57)	0 (0)
C16327T	HV-II	6 (18,18)	2 (5,71)	0 (0)
T16362C	HV-II	9 (27,27)	13 (37,14)	7 (22,58)
T16519C	HV-II	20 (60,61)	21 (60)	23 (74,19)

* (N,%): dữ liệu đa hình được biểu diễn dưới dạng: (số cá thể mang đa hình, tỷ lệ phần trăm).

Các đa hình thuộc bảng 3 được phân tích thống kê sử dụng kiểm định Fisher exact từng đôi một giữa các nhóm tộc người để xem có sự phân bố khác nhau giữa các nhóm hay không. Kết quả cho thấy, trong 33 đa hình có 14 đa hình có sự khác biệt có ý nghĩa thống kê giữa các nhóm: gồm 4 đa hình khi so sánh giữa nhóm cá thể tộc người Hà Nhi và nhóm thuộc tộc người Phù Lá, 10 đa hình khi so sánh giữa nhóm thuộc tộc người Hà Nhi và nhóm thuộc tộc người Si La, 10 đa hình khi so sánh giữa 2 nhóm thuộc tộc người Phù Lá và Si La (bảng 4).

Bảng 4. Các đa hình có phân bố khác nhau ở các tộc người trong nghiên cứu.

Đa hình	Hà Nhi - Phù Lá ^(*)	Hà Nhi - Si La ^(*)	Phù Lá - Si La ^(*)
T152C	-	0,03	0,02
T16086C	<0,001	-	<0,001
C16108T	-	<0,001	<0,001
G16129A	-	0,003	<0,001
A16162G	-	<0,001	<0,001
C16167T	0,001	-	0,001
T16172C	0,033	0,002	<0,001
A16182C	-	-	0,003
A16203G	0,001	-	0,001
C16290T	-	0,025	0,026
T16298C	-	0,002	-
T16304C	-	0,001	-
A16318G	-	0,002	-
C16327T	-	0,025	-

^(*)Giá trị p được tính bằng kiểm định Fisher exact (2 phía) bằng phần mềm SPSS version 23.

Trình tự ADN vùng D-loop (từ vị trí nucleotide số 1 đến 576 và từ nucleotide số 16024 đến 16569) của các cá thể được sử dụng để tính toán khoảng cách di truyền giữa các cá thể trong cùng một tộc người và giữa các cá thể thuộc các tộc người khác nhau bằng phần mềm MEGA X. Kết quả cho thấy nhóm các cá thể thuộc tộc người Si La có khoảng cách di truyền trung bình giữa các cá thể thấp nhất (khoảng 0,008) so với các nhóm thuộc tộc người Hà Nhi và Si La (khoảng 0,01). Khi so sánh giữa các nhóm tộc người, chúng tôi nhận thấy khoảng cách giữa các cá thể thuộc các nhóm tộc người không khác nhau nhiều (khoảng cách trung bình $d \sim 0,01$) (bảng 5).

Bảng 5. Thông tin về khoảng cách di truyền trung bình của các tộc người tham gia trong nghiên cứu.

Khoảng cách di truyền giữa các cá thể trong một tộc người	
Hà Nhi	0,0107
Phù Lá	0,0101
Si La	0,0081

Khoảng cách di truyền giữa các tộc người			
	Hà Nhi	Phù Lá	Si La
Hà Nhi			
Phù Lá	0,011		
Si La	0,0105	0,0106	

Thảo luận

Hệ gen ty thể có đặc điểm được di truyền chủ yếu theo dòng mẹ, không hoặc rất ít xảy ra trao đổi chéo và có tỷ lệ đột biến phát sinh đa hình tương đối cao. Do đó, đa hình thuộc hệ gen ty thể là một công cụ rất hiệu quả trong các nghiên cứu về lịch sử phát sinh và di cư của các chủng tộc loài người [7-9]. Hai đa hình A73G và A263G xuất hiện với tần số 100% ở nghiên cứu này cũng đã xuất hiện với tần số 100% ở các nghiên cứu về hệ gen ty thể ở các tộc người Việt Nam khác như: tộc người Kinh và tộc người Mường [2], Thái, Tày, Mông [10], Mảng [5]. Đây cũng là các đa hình tương đối phổ biến ở các tộc người láng giềng như người Makrani ở Pakistan [11], hay người Quảng Tây ở Trung Quốc [12].

Các kết quả tính toán khoảng cách di truyền cho thấy hệ gen ty thể tương đối đồng nhất giữa các cá thể trong một tộc người và giữa ba tộc người Hà Nhi, Phù Lá, Si La. Đây là các tộc người đã sinh sống lâu đời (hơn 300 năm) cùng trên địa bàn các tỉnh miền núi Tây Bắc [13], vì vậy kết quả trên hoàn toàn có thể hiểu được. Tuy nhiên, trong nghiên cứu này chúng tôi đã phát hiện 14 đa hình có sự khác biệt có ý nghĩa thống kê giữa các nhóm cá thể thuộc ba tộc người. Điều này cho thấy rằng, mặc dù sinh sống trên cùng địa bàn, cùng thuộc một nhóm ngữ hệ Tạng - Miến và văn hóa, phong tục có nhiều điểm tương đồng, nhưng các tộc người này, vẫn có những điểm độc đáo về mặt vật chất di truyền. Đặc biệt là ở nghiên cứu này, chúng tôi phát hiện thấy 4 đa hình chỉ xuất hiện ở nhóm thuộc tộc người Phù Lá mà không xuất hiện ở hai nhóm tộc người còn lại là: T16086C, C16167T, A16203G và A16318G. Các đa hình này đã được ghi nhận xuất hiện với tần số đa dạng ở các quần thể người khỏe mạnh tại Iran [14], quần thể người Tây Tạng (Tibetan) ở chân núi Himalaya [15] và người Ma rốc [16]. Các đa hình này cũng không thấy xuất hiện hoặc xuất hiện với tần số rất thấp ở các tộc người khác sinh sống trên cùng địa bàn như tộc người Kinh, tộc người Mảng [5], hay tộc người Lô Lô, tộc người La Hù (dữ liệu chưa công bố). Vì vậy, các đa hình trên có thể coi là các đa hình đặc trưng cho hệ gen ty thể của tộc người Phù Lá. Các kết quả ban đầu này sẽ là cơ sở dữ liệu để chúng tôi tiếp tục nghiên cứu sâu hơn về lịch sử di truyền và nguồn gốc tiến hóa của các tộc người Việt Nam.

Kết luận

Nghiên cứu này đã xác định được các đa hình thuộc vùng D-loop hệ gen ty thể của nhóm các cá thể thuộc ba tộc người cùng thuộc ngữ hệ Tạng - Miến là Hà Nhì, Phù Lá và Si La. Mặc dù với cỡ mẫu còn hạn chế, nghiên cứu này bước đầu đã xác định được đa hình nucleotide ty thể giữa các cá thể thuộc ba tộc người này tương đối giống nhau và không có sự khác biệt lớn. Tuy nhiên, kết quả nghiên cứu cũng chỉ ra 14 điểm đa hình thuộc vùng D-loop có sự phân bố khác nhau có ý nghĩa thống kê giữa các nhóm cá thể thuộc các tộc người khác nhau, trong đó có 4 điểm đa hình chỉ xuất hiện ở nhóm cá thể thuộc tộc người Phù Lá mà không xuất hiện ở hai nhóm tộc người còn lại cũng như các tộc người lân cận sinh sống trên cùng địa bàn. Đây là nghiên cứu tiền đề để chúng tôi tiếp tục phân tích sâu hơn về đa dạng hệ gen ty thể cũng như toàn bộ hệ gen của các tộc người Việt Nam.

LỜI CẢM ƠN

Công trình được hoàn thành với sự tài trợ từ đề tài cấp nhà nước mã số ĐTĐL.CN-05/15, do Bộ Khoa học và Công nghệ quản lý. Chúng tôi xin chân thành cảm ơn Bộ Khoa học và Công nghệ cùng những người đã tham gia cung cấp mẫu cho nghiên cứu này.

TÀI LIỆU THAM KHẢO

[1] K. Kawabe, R. Worawut, S. Taura, T. Shimogiri, T. Nishida, S. Okamoto (2014), “Genetic Diversity of mtDNA D-loop Polymorphisms in Laotian Native Fowl Populations”, *Asian-Australasian Journal of Animal Sciences*, **27(1)**, pp.19-23.

[2] Trần Thị Thúy Hằng, Trần Văn Khánh (2017), “Đa hình thái đơn nucleotid vùng gen ty thể HV1 và HV2 trên người tộc người Kinh và tộc người Mường của Việt Nam”, *Tạp chí Nghiên cứu Y học*, **106(1)**, tr.33-40.

[3] D.M. Hung, N.H. Ha, P.N. Khoi, V.P. Nhung, V.V. Phong, N.T. Duong, N.V. Hai, N.D. Ton (2016), “Genetic variation of mitochondrial sequence-hv2 in Vietnamese populations”, *Journal of Biology*, **38(2)**, pp.243-249.

[4] Nguyễn Đăng Tôn, Nguyễn Thị Tú Linh, Vũ Hải Chi, Trần Thị Ngọc Diệp, Địch Thị Kim Hương, Bùi Thị Tuyết, Nguyễn Hải Hà, Huỳnh Thị Thu Huệ, Lê Thị Thu Hiền, Trần Thị Phương Liên, Phan Văn Hải, Nông Văn Hải (2008), “Đa hình đơn bội ADN ty thể của các cá thể người Việt Nam”, *Tạp chí Công nghệ sinh học*, **6(4)**, tr.579-590.

[5] Nguyễn Thy Ngọc, Nguyễn Bảo Trang, Nguyễn Quang Huy, Nguyễn Đăng Tôn, Nguyễn Thùy Dương (2018), “Đa hình vùng D-loop hệ gen ty thể của các cá thể tộc người Kinh và Mảng cùng trong nhóm ngữ hệ Nam Á”, *Tạp chí Công nghệ sinh học*, **16(2)**, tr.231-240.

[6] Ban chỉ đạo Tổng điều tra dân số và nhà ở trung ương (2009), *Tổng điều tra dân số và nhà ở Việt Nam - Kết quả toàn bộ*.

[7] P. Endicott, S.Y. Ho, M. Metspalu, C. Stringer (2009), “Evaluating the mitochondrial timescale of human evolution”, *Trends in Ecology and Evolution*, **24(9)**, pp.515-521.

[8] A. Achilli, U.A. Perego, C.M. Bravi, M.D. Coble, Q.P. Kong, S.R. Woodward, A. Salas, A. Torroni, H.J. Bandelt (2008), “The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies”, *PLoS One*, **3(3)**, p.e1764.

[9] N.J. Fagundes, R. Kanitz, R. Eckert, A.C. Valls, M.R. Bogo, F.M. Salzano, D.G. Smith, W.A. Silva, M.A. Zago, A.K. Ribeiro-dos-Santos (2008), “Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas”, *American Journal of Human Genetics*, **82(3)**, pp.583-592.

[10] Huỳnh Thị Thu Huệ, Hoàng Thị Thu Yến, Nguyễn Đăng Tôn, Lê Thị Thu Hiền, Nguyễn Đình Cường, Phan Văn Chi, Nông Văn Hải (2005), “Phân tích trình tự vùng điều khiển (D-loop) trên genome ty thể của 5 cá thể người Việt Nam”, *Tạp chí Công nghệ sinh học*, **3(1)**, tr.5-22.

[11] M.H. Siddiqi, T. Akhtar, A. Rakha, G. Abbas, A. Ali, N. Haider, S. Hayat, S. Masooma, J. Ahmad, M.A. Tariq (2015), “Genetic characterization of the Makrani people of Pakistan from mitochondrial DNA control-region data”, *Leg. Med. (Tokyo)*, **17(2)**, pp.134-139.

[12] X. Yang, X. Wang, H. Yao, J. Deng, Q. Jiang, Y. Guo, G. Lan, D.J. Liao, H. Jiang (2012), “Mitochondrial DNA polymorphisms are associated with the longevity in the Guangxi Bama population of China”, *Molecular Biology Reports*, **39(9)**, pp.9123-9131.

[13] Nguyễn Văn Huy, Lê Duy Đại, Nguyễn Quý Thao, Vũ Xuân Thảo (2014), *Đại gia đình các tộc người Việt Nam*, Nhà xuất bản Giáo dục Việt Nam (tái bản lần thứ 7), tr.52-114.

[14] K. Mousavizadeh, P. Rajabi, M. Alaei, S. Dadgar, M. Houshmand (2015), “Usage of mitochondrial D-loop variation to predict risk for Huntington disease”, *Mitochondrial DNA*, **26(4)**, pp.579-582.

[15] S. Bhandari, X. Zhang, C. Cui, Bianba, S. Liao, Y. Peng, H. Zhang, K. Xiang, H. Shi, Ouzhuluobu, et al. (2015), “Genetic evidence of a recent Tibetan ancestry to Sherpas in the Himalayan region”, *Scientific Reports*, **5**, p.16249.

[16] H. Charoute, R. Kefi, S. Bounaceur, H. Benrahma, A. Reguig, M. Kandil, H. Rouba, A. Bakhchane, S. Abdelhak, A. Barakat (2018), “Novel variants of mitochondrial DNA associated with Type 2 diabetes mellitus in Moroccan population”, *Mitochondrial DNA Part A: DNA Mapping, Sequencing, and Analysis 2018*, **29(1)**, pp.9-13.