

## Nghiên cứu đa dạng Haplotype Y- STR của quần thể người Kinh - Việt Nam

Trịnh Tuấn Toàn<sup>1,3,\*</sup>, Hà Quốc Khanh<sup>1</sup>, Vũ Lê Lợi<sup>1</sup>,  
Nguyễn Quang Vinh<sup>2</sup>, Nguyễn Quang Huy<sup>3</sup>

<sup>1</sup>Viện Khoa học Hình sự, Bộ Công an, 99 Nguyễn Tuân, Thanh Xuân, Hà Nội, Việt Nam

<sup>2</sup>Công ty Cổ phần Phân tích Di truyền (Gentis), Tây Hồ, Hà Nội, Việt Nam

<sup>3</sup>Khoa Sinh học, Trường Đại học Khoa học Tự nhiên, ĐHQGHN,  
334 Nguyễn Trãi, Thanh Xuân, Hà Nội, Việt Nam

Nhận ngày 16 tháng 8 năm 2017

Chỉnh sửa ngày 20 tháng 9 năm 2017; Chấp nhận đăng ngày 10 tháng 10 năm 2017

**Tóm tắt:** Các mẫu ADN tách chiết từ máu, gốc tóc, móng tay hoặc móng chân của 205 người đàn ông khỏe mạnh, không có quan hệ huyết thống, thuộc dân tộc Kinh ở Việt Nam được lấy mẫu ngẫu nhiên đã được khuếch đại bằng sử dụng bộ kit AmpFISTR® Yfiler™ PCR Amplification Kit (Yfiler). Áp dụng phương pháp đếm trực tiếp và phân tích bằng phần mềm Arlequin 3.1 để xác định tần số alen, xác định khoảng cách di truyền sử dụng cơ sở dữ liệu trực tuyến Y-Chromosome haplotype. Kết quả thu được 205 haplotype, mức độ đa dạng haplotype là 1, đa dạng locus trung bình là 0,7506 và dao động từ 0,2373 ở locus DYS438 đến 0,9049 ở locus DYS385. Đây là những kết quả đầu tiên đánh giá đa dạng di truyền quần thể dựa trên 17 chỉ thị di truyền STR (short tandem repeat) trên nhiễm sắc thể giới tính Y (Y-STR) ở quần thể người Kinh (Việt Nam).

**Từ khóa:** Nhiễm sắc thể Y; Y-STR; đa dạng haplotype; quan hệ huyết thống; Yfiler kit.

### 1. Mở đầu

Ở người, nhiễm sắc thể giới tính Y nhỏ hơn nhiều so với nhiễm sắc thể giới tính X và số gen trên Y chỉ chiếm khoảng 2% tổng số gen trong hệ gen đơn bội và dài khoảng 60 Mb. Các vùng không tái tổ hợp (NRY) trên nhiễm sắc thể giới tính Y của người chiếm xấp xỉ 95% của toàn bộ nhiễm sắc thể [1]. Các đoạn lặp ngắn trên nhiễm sắc thể giới tính Y (Y-STR) có độ dài từ 2-7 bp trên vùng không tái tổ hợp (NRY), chỉ xuất hiện ở nam giới và được truyền từ nguyên vẹn từ cha sang con trai, trừ trường hợp đột

biến. Do vậy, phân tích Y-STR là rất quan trọng cho giám định ADN hình sự, như trong việc nhận dạng người chưa rõ tung tích; trường hợp bị xâm hại tình dục, khi mà phân tích ADN nhiễm sắc thể Y có thể tách riêng được kiểu gen của nam giới trong mẫu dịch âm đạo và trường hợp mẫu lẫn; xác định quan hệ huyết thống theo dòng cha trong trường hợp người cha đã chết; xác định amelogenin Y thiếu của người đàn ông; nghiên cứu sự di cư của con người, nghiên cứu tiến hóa phả hệ,... [2-3].

Tần số của haplotype trong quần thể rất quan trọng, đặc biệt với tính toán xác suất trong giám định ADN hình sự. Việc tăng cường sử dụng các marker Y-STR của người ứng dụng trong pháp y, trong nghiên cứu nhân chủng học và khảo cổ học đã tạo ra sự nỗ lực hợp tác để

\* Tác giả liên hệ. ĐT.: 84-975155747.

Email: toantt1974@gmail.com

<https://doi.org/10.25073/2588-1140/vnunst.4539>

thu thập dữ liệu haplotype từ các quần thể khác nhau và để xây dựng các cơ sở dữ liệu Y [4]. Bài báo này công bố nghiên cứu cấu trúc di truyền của người Việt (Kinh) trên chỉ thị di truyền Y-STR của nhiễm sắc thể Y và là cơ sở cho việc giám định ADN hình sự, xác định huyết thống cũng như góp phần xây dựng cơ sở dữ liệu cho nghiên cứu phả hệ.

## 2. Vật liệu và phương pháp nghiên cứu

### 2.1. Mẫu nghiên cứu

Toàn bộ 205 mẫu cá thể đàn ông người Việt Nam (dân tộc Kinh) khỏe mạnh, không có quan hệ huyết thống trực hệ với nhau được thu thập từ tháng 10/2010 đến tháng 3/2015 từ các tỉnh, thành phố lớn ở Việt Nam (Hà Nội, Đà Nẵng, TP Hồ Chí Minh) và các vùng lân cận. Đã có một số tác giả nghiên cứu về 17 chỉ thị trên nhiễm sắc thể Y (Y-STR), tuy nhiên chưa có nghiên cứu nào khảo sát với số lượng mẫu lớn (hơn 200 mẫu) đối với người Kinh - Việt Nam. Các mẫu trong nghiên cứu này được phân tích chủ yếu là mẫu tóc, một số trường hợp là mẫu móng tay, móng chân. Đối với các cá nhân đến trực tiếp phòng thí nghiệm, mẫu tiền hành thu chủ động là mẫu máu và được thu trên giấy FTA™ chuyên dụng, lượng mẫu máu lưu trên giấy FTA™ có kích thước 2x2 (mm).

### 2.2. Tách chiết và phân tích ADN

ADN của các mẫu tóc, móng được tách chiết bằng dung dịch Chelex-100™ kết hợp bổ sung proteaza K và ủ ở nhiệt độ 56°C. Với các mẫu móng chân tay: bổ sung đệm Tris - EDTA pH 8,0 để hòa tan và bảo quản ADN. Với mẫu máu thu trên giấy FTA™: rửa với dung dịch FTA™ và DDW™ và sử dụng trực tiếp mẫu FTA™ cho phản ứng PCR.

ADN được định lượng bằng phương pháp Realtime PCR sử dụng bộ kit Quantifiler™ Human DNA (Applied Biosystems, Mỹ) với hệ thống Real-Time 7500 (Applied Biosystems, Mỹ) [5].

1,0 ng ADN khuôn cho mỗi mẫu được phân tích bằng bộ kit Yfiler nhằm khuếch đại 17

locus Y-STR bao gồm: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, và GATAH4. Phản ứng PCR thực hiện trên máy khuếch đại gen GeneAmp 9700 (Applied Biosystems, Mỹ). Sản phẩm khuếch đại Y-STR được điện di và phân tích trên máy giải trình tự gen ABI 3130 Genetic Analyzer với phần mềm Genemapper-ID, sử dụng kích thước chuẩn (GeneScan-500 LIZ, Applied Biosystems, Mỹ) và thang alen chuẩn được nhà sản xuất cung cấp theo kit Y filer.

### 2.3. Phân tích thống kê

Alen được nhận diện và tính tần suất haplotype bằng cách đếm trực tiếp. Chỉ số đa dạng locus và haplotype (H) được tính bằng cách sử dụng phần mềm Arlequin 3.1 [6] theo công thức:

$$H = \frac{n}{n-1} \left( 1 - \sum_{i=1}^k P_i^2 \right)$$

Trong đó: H là chỉ số đa dạng haplotype, n là số lượng mẫu nghiên cứu và  $P_i$  là tần suất của haplotype thứ i (k là số Haplotype quan sát được). Đa dạng di truyền locus (GD) được tính toán theo cùng công thức trên, sử dụng tần số alen thay tần suất haplotype để tính toán (k là số alen nhận diện được trong locus gen khi khảo sát n mẫu nghiên cứu).

Sử dụng công cụ trực tuyến YHRD (www.yhrd.org) để thực hiện đánh giá khoảng cách di truyền của quần thể nghiên cứu với các quần thể khác đã được công bố trên cơ sở dữ liệu. Khoảng cách di truyền các quần thể ( $R_{st}$ ) là phần mở rộng của biện pháp FST được tính theo công thức:  $R_{st} = (S_b - S_w)/S_b$ . Trong đó  $S_w$  là tổng lấy theo quỹ đạo của 2 lần trung bình trọng số của những phương sai bên trong quần thể V(A) và V(B),  $S_b$  là tổng theo những quỹ đạo của 2 lần phương sai V(A+B) ở quần thể kết hợp. Giá trị P được tính toán bằng phân tích phương sai phân tử (AMOVA) với 10.000 hoán vị, sử dụng công cụ trực tuyến YHRD. Kiểm định với mức ý nghĩa  $P < 0,05$ .

Trước tiên, so sánh khoảng cách di truyền giữa quần thể nghiên cứu với các quần thể gần về vị trí địa lý để thấy được sự phù hợp giữa di truyền quần thể với vị trí địa lý. Các quần thể khác nhau được chọn gồm: Người Việt Nam có dữ liệu trên YHRD (quần thể người Hoa (số haplotype,  $n = 7$ ); người Kinh (ở TP Hồ Chí Minh,  $n = 45$ ); người Kinh (Việt Nam,  $n = 123$ ); người Mông ( $n = 59$ ); người Nùng ( $n = 7$ ); người Tày ( $n = 21$ ) và người Kinh (ở Hà Nội,  $n = 48$ )) và các quần thể thuộc các dân tộc gần về mặt địa lý và lịch sử với Việt Nam được so sánh như Trung Quốc (Jiangsu,  $n = 248$ , Jiangxi,  $n = 72$  và Đài Loan,  $n = 41$ ).

Bước tiếp theo, so sánh khoảng cách di truyền giữa quần thể nghiên cứu với các quần thể người ở các quốc gia khác xa về vị trí địa lý để thấy rõ sự khác biệt đáng kể về mặt di truyền. Các quần thể được chọn gồm: Người Brazil (Rio de Janeiro,  $n = 100$ ), người Đức (Berlin,  $n = 153$ , người Nigeria (Ibadan,  $n = 116$ , người Ethiopia (Amharic,  $n = 152$ ), người Romani (Bulgaria,  $n = 228$  và Mỹ (European American,  $n = 1744$ ).

### 3. Kết quả và thảo luận

#### 3.1. Kết quả nghiên cứu

Kết quả nghiên cứu từ tổng số 99 alen ở 17 locus Y-STR trong 205 mẫu haplotype cho thấy không có haplotype nào trùng nhau. Chỉ số đa dạng haplotype tương ứng  $H = 1$ . Locus đa hình nhất là DYS385a/b với chỉ số đa dạng di truyền locus (GD) là 0,9049 và sự đa hình ít nhất ở locus DYS438 với giá trị GD là 0,2373.

Ngoài locus DYS385, các locus DYS389II, DYS458 và DYS6335 cho các giá trị GD cao nhất. Locus khác nhau đã có số lượng các alen khác nhau. Một alen trung gian đã được nhìn thấy tại DYS385 là alen 18.2. Một null alen đã được quan sát thấy ở DYS456 (Bảng 1). Phát hiện 1 locus có 3 alen là locus DYS385a/b ở mẫu ID114: alen 13, 14, 18.2 và ở mẫu ID149: alen 13, 17, 18.

Sử dụng công cụ trực tuyến AMOVA trên cơ sở dữ liệu YHRD ([www.yhrd.org](http://www.yhrd.org)) cho thấy

quần thể người Kinh (Việt Nam) nghiên cứu khá gần với quần thể người Việt ở Việt Nam và có sự tương đồng có ý nghĩa thống kê với người Việt (Kinh) ở Hà Nội và người Việt (Kinh) ở TP Hồ Chí Minh đã được công bố trên cơ sở dữ liệu YHRD. So sánh từng cặp cho thấy quần thể người Kinh (Việt Nam) trong nghiên cứu và người Việt (Kinh) ở Hà Nội có khoảng cách di truyền  $R_{st} = 0,0223$ ,  $p = -1,0000$ ; với người Việt (Kinh) ở TP Hồ Chí Minh có ( $R_{st} = 0,0042$ ,  $p = -1,0000$ ) (Bảng 2).

Kết quả số liệu ở bảng 3 cho thấy mối liên hệ khá gần giữa quần thể nghiên cứu với quần thể người Jiangxi (Han) và Jiangsu (Han) ở Trung Quốc và quần thể Han ở Đài Loan. So sánh với các quần thể sống ở quốc gia có vị trí địa lý xa với Việt Nam và khác biệt về lịch sử và văn hóa cho thấy quần thể nghiên cứu có sự khác biệt rất đáng kể như với người Mỹ ( $R_{st} = 0,3836$ ,  $P < 0,0001$ ); người Đức ( $R_{st} = 0,3562$ ,  $p = -1,0000$ ); người Romani ( $R_{st} = 0,2918$ ,  $p = -1,0000$ ) hay với người Ethiopia ( $R_{st} = 0,3139$ ,  $p = -1,0000$ ) và đặc biệt là khác biệt đáng kể với người Nigeria ( $R_{st} = 0,5084$ ,  $p = -1,0000$ ).

#### 3.2. Thảo luận

Nghiên cứu của chúng tôi cho thấy quần thể người Kinh (Việt Nam) có tính đa dạng di truyền cao, đặc biệt locus DYS385 có tính đa hình khá cao ( $GD = 0,9049$ ). Ngoài ra chúng tôi phát hiện ra một alen lạ (null allele) ở locus DYS456 và 3 alen ở locus DYS385, một alen trung gian (18.2) ở locus DYS385, các chỉ thị này có giá trị cao trong giám định ADN hình sự và xác định quan hệ huyết thống.

Trên cơ sở dữ liệu YHRD đã có một số dữ liệu Y-STR của người Việt. Tuy nhiên, các nghiên cứu này thực hiện với số lượng mẫu nghiên cứu nhỏ, vùng thu mẫu hẹp, chưa có tính đại diện (tiêu quần thể). Nghiên cứu này là nghiên cứu đa dạng haplotype - YSTR đầu tiên ở Việt Nam có quy mô đủ lớn, khu vực thu mẫu rộng (người Kinh ở các tỉnh và thành phố lớn và vùng lân cận) và đối tượng thu mẫu được xác minh nên kết quả đảm bảo được các yếu tố khách quan trong chọn mẫu và thống kê.

Bảng 1. Tần suất alen và chỉ số đa dạng locus gen của 17 locus Y-STR

Alen	DYS456	DYS389I	DYS390	DYS389II	DYS458	DYS19	DYS393	DYS391	DYS439	DYS635	DYS392	Y_GATA_H4	DYS437	DYS438	DYS448	Alen	DYS385
6								0,0098								13,13	0,0439
7																12,19	0,0244
8														0,0049		13,19	0,1512
9								0,0049						0,0146		14,20	0,0049
10							0,0049	0,5659	0,0341		0,0049	0,1512		0,8683		12,20	0,0098
11		0,1073						0,3902	0,2146		0,0829	0,4098		0,0976		13,18,2	0,0927
12		0,2537				0,0049	0,2585	0,0293	0,5951		0,0390	0,4049		0,0146		13,18	0,2195
13	0,0293	0,4390				0,0098	0,1659		0,1415		0,7220	0,0244	0,0195			14,14	0,0244
14	0,1854	0,2000			0,0049	0,0829	0,4976		0,0146		0,1317	0,0098	0,8049			11,18	0,0439
15	0,6390				0,1171	0,3512	0,0732				0,0195		0,1610	0,0098		15,20	0,0244
16	0,1024				0,1366	0,4341							0,0146	0,0146		12,18	0,0488
17	0,0390				0,2049	0,1171								0,0488		13,14	0,0098
18					0,3512									0,5024		16,19	0,0146
18.2																11,20	0,0098
19					0,1463				0,0341						0,1561	12,13	0,0146
20					0,0390				0,0829						0,1610	13,20	0,0341
21									0,4732						0,1024	14,19	0,0146
22			0,0146						0,2293						0,0049	11,19	0,0049
23			0,2732						0,1268							14,17	0,0098
24			0,4341						0,0439							12,15	0,0049
25			0,2488						0,0098							15,19	0,0195
26			0,0293	0,0098												12,17	0,0146
27				0,0390												14,18	0,0878
28				0,2585												12,16	0,0098
29				0,2683												12,12	0,0098
30				0,3268												13,17	0,0244
31				0,0878												11,16	0,0098
32				0,0098												14,16	0,0049
null	0,0049															15,18,2	0,0146
																13,14,19	0,0049
																13,17,19	0,0049
GD	0,5471	0,6948	0,6772	0,7486	0,7832	0,6707	0,6559	0,5291	0,5812	0,7008	0,4549	0,6478	0,3273	0,2373	0,6874		0,9049

Bảng 2. Bảng giá trị  $R_{st}$  và P của quần thể người Việt (Kinh)

Quần thể người	Việt Nam [Kinh]	Việt nam [Hoa]	Tp Hồ Chí Minh, Việt Nam [Kinh]	Việt Nam [Kinh]	Việt Nam [Mông]	Việt Nam [Nùng]	Việt Nam [Tày]	Hà Nội, Việt Nam
Việt Nam [Kinh]	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Việt Nam [Hoa]	0,0568	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Tp Hồ Chí Minh, Việt Nam [Kinh]	0,0042	0,0565	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Việt Nam [Kinh]	0,0095	0,0388	0,0237	-	-1,0000	-1,0000	-1,0000	-1,0000
Việt Nam [Mông]	0,0583	0,0557	0,0489	0,0673	-	-1,0000	-1,0000	-1,0000
Việt Nam [Nùng]	0,0196	-0,0584	0,0318	-0,0116	0,0866	-	-1,0000	-1,0000
Việt Nam [Tày]	0,0377	-0,0024	0,0294	0,0125	0,0979	-0,0757	-	-1,0000
Hà Nội Việt Nam	0,0223	0,0358	0,0141	0,0276	0,0343	0,0193	0,0379	-

+  $R_{st}$  là giá trị ở dưới đường chéo, +  $P < 0,05$  có ý nghĩa thống kê

Bảng 3. Bảng giá trị  $R_{st}$  và P giữa quần thể người Việt (Kinh) và các nước

Quần thể người	Việt [Kinh]	Rio de Janeir, Braxin [African]	Bulgaria [Romani]	Jiangs, Trung Quốc [Han]	Jiangx, Trung Quốc [Han]	Ethiopia [Amharic]	Berlin, Đức [German]	Ibadan, Nigeria [Yoruba]	Đài Loan [Han]	Mỹ [European American]
Việt Nam [Kinh]	-	-1,000	-1,000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Rio de Janeiro, Braxin [African]	0,2374	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Bulgaria [Romani]	0,2918	0,1652	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Jiangsu, Trung Quốc [Han]	0,0750	0,1401	0,2199	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Jiangxi, Trung quốc [Han]	0,0249	0,1754	0,2835	0,0156	-	-1,0000	-1,0000	-1,0000	-1,0000	-1,0000
Ethiopia [Amharic]	0,3139	0,1084	0,1185	0,2339	0,2524	-	-1,0000	-1,0000	-1,0000	-1,0000
Berlin, Đức [German]	0,3562	0,1411	0,3781	0,2007	0,3011	0,3097	-	-1,0000	-1,0000	-1,0000
Ibadan, Nigeria [Yoruba]	0,5084	0,3476	0,3681	0,4567	0,5171	0,2732	0,5959	-	-1,0000	-1,0000
Đài Loan [Han]	0,0400	0,2027	0,2956	0,0313	-0,0010	0,2491	0,3505	0,5443	-	-1,0000
Mỹ [European American]	0,3836	0,1711	0,4056	0,2420	0,3209	0,3843	0,0385	0,5975	0,3728	-

Nghiên cứu tương đồng với các quần thể (có dữ liệu trên YHRD) ở vị trí địa lý gần và khác biệt di truyền với các quần thể ở những khu vực xa nên kết quả nghiên cứu là phù hợp giữa đặc điểm di truyền quần thể với đặc điểm địa lý, di cư quần thể.

Kết quả nghiên cứu bước đầu này góp phần hình thành cơ sở dữ liệu di truyền Y-STR cho mục đích giám định ADN hình sự, xác định quan hệ huyết thống, tìm tung tích nạn nhân trong tai nạn, thảm họa và xác định danh tính các Liệt sỹ hi sinh trong chiến tranh.

#### 4. Kết luận

Kết quả nghiên cứu cho thấy chỉ số đa dạng haplotype tương ứng  $H = 1$  (không có haplotype nào trùng nhau trong số 205 mẫu nghiên cứu).

Tính đa hình cao ở các locus Y-STR đối với người Kinh - Việt Nam trong đó locus đa hình nhất là DYS385a/b với chỉ số đa dạng di truyền locus (GD) là 0,9049 và sự đa hình ít nhất ở locus DYS438 với giá trị GD là 0,2373.

#### Lời cảm ơn

Các tác giả xin chân thành cảm ơn Viện Khoa học hình sự - Bộ Công an, Công ty cổ phần Phân tích di truyền (Gentis) và Khoa Sinh

học, Trường Đại học Khoa học Tự nhiên, ĐHQGHN đã giúp đỡ trong công tác phân tích mẫu. Kinh phí nghiên cứu được tài trợ từ Công ty Cổ phần phân tích di truyền (Gentis) và Viện Khoa học hình sự, Bộ Công an.

#### Tài liệu tham khảo

- [1] Charlesworth B, Charlesworth D. The degeneration of Y-chromosomes. *Philos Trans R Soc Lond B Biol Sci.* 2000 Nov 29; 355: 1563-72.
- [2] Butler JM. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis, *Forensic Science Rev*, 2003, 15: 91-111.
- [3] Gusmão L, Butler JM, Carracedo A, Gill P, Kayser M, Morling N, Prinz M, Roewer L, Tyler-Smith C, Schneider PM; DNA commission of the international Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-Strs in forensic analysis, *Forensic Sci Int.* 2006, 157:187-97.
- [4] Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes, *Forensic Sci Int.* 2001, 118:106-13.
- [5] Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques.* 1991, 10:506-13.
- [6] Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online.* 2005, 1: 47-50.

## Study on Haplotype Diversity of Y-STR Loci of the Kinh Population in Vietnam

Trinh Tuan Toan<sup>1,3</sup>, Vu Le Loi<sup>1</sup>, Ha Quoc Khanh<sup>1</sup>,  
Nguyen Quang Vinh<sup>2</sup>, Nguyen Quang Huy<sup>3</sup>

<sup>1</sup>Forensic Science Institute, Ministry of Public Security, 99 Nguyen Tuan, Thanh Xuan, Hanoi, Vietnam

<sup>2</sup>Genetic Analysis JSC (Gentis), Tay Ho District, Hanoi, Vietnam

<sup>3</sup>Faculty of Biology, VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam

**Abstract:** The study was carried out to determine allele frequencies of Y-STR loci in order to apply in the statistical validation of forensic case and paternal lineages. 17 Y-STR loci were co-amplified (including: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, and GATAH4) using DNA genome extracted from blood, hair root and finger nail samples of 205 unrelated men on Kinh population in the North and South of Vietnam.

The studied population shows that a relatively close relationship between the Jiangxi (Han) and Jiangsu (Han) populations in China and the Han population in Taiwan. Compared to other population with geographic locations far removed from Vietnam and historically and culturally diverse, there are significant differences as with Americans ( $R_{st} = 0.3836$ ,  $P < 0.0001$ ); The Germans ( $R_{st} = 0.3562$ ,  $p = -1.0000$ ); The Romans ( $R_{st} = 0.2918$ ,  $p = -1.0000$ ) or Ethiopian ( $R_{st} = 0.3139$ ,  $p = -1.0000$ ) and particularly significant differences with Nigerians ( $R_{st} = 0.5084$ ,  $p = -1.0000$ ).

The genetic diversity indexes at each locus and the haplotypes were examined; Genetic distances between the Kinh population in this study and neighboring populations and geographically remote populations are consistent with the migration history.

**Keywords:** Chromosome Y; Y-STR; haplotype; population; Yfiler kit.