

SỰ PHÂN BỐ CÁC ĐA HÌNH NUCLEOTIDE ĐƠN CỦA NHÓM ĐƠN BỘI O TRÊN NIỄM SẮC THỂ Y Ở NGƯỜI VIỆT NAM

Nguyễn Đăng Tôn, Nguyễn Thùy Dương, Nông Văn Hải

Phòng thí nghiệm trọng điểm Công nghệ gen, Viện Công nghệ sinh học

TÓM TẮT

Các chi thị trên vùng không trao đổi chéo của nhiễm sắc thể Y vẫn được sử dụng trong nhiều lĩnh vực nghiên cứu khác nhau, bao gồm nghiên cứu về tiến hóa, ứng dụng trong khoa học hình sự, di truyền y học và trong việc tái xây dựng lại cây phát sinh chủng loại người. Việt Nam có 54 dân tộc anh em, trong đó dân tộc Kinh là dân tộc có số dân đông nhất, chiếm 87% dân số. Cho đến nay, vẫn chưa có công trình nào công bố về các nhóm đơn bội trên nhiễm sắc thể Y ở người Việt Nam. Trong công trình này, chúng tôi công bố số liệu đầu tiên về sự phân bố và tần số của các nhóm đơn bội của nhiễm sắc thể Y trên đối tượng là người Kinh của Việt Nam. Để tìm hiểu về sự phân bố của các nhóm đơn bội trên nhiễm sắc thể Y ở người Kinh của Việt Nam, chúng tôi đã sử dụng 10 chỉ thị hai allele thuộc nhóm đơn bội O trên nhiễm sắc thể Y là M175, P186, P191, P196, M119, M50, M103, M110, SRY465 và JST022454. Kết quả phân tích cho thấy, 6 trên 10 chi thị có sự đa hình trong tổng số 248 mẫu nghiên cứu. Nhóm đơn bội O1a*-M119 có tần số phân bố cao nhất là 62,9% trong tổng số 248 mẫu nghiên cứu. Nhóm đơn bội O*-M175/P186/P191/P196 có tần số phân bố là 10,1%. Nhóm đơn bội O1a2 (M50, M103 và M110) và O2b-SRY465 không phát hiện thấy đa hình trong các mẫu nghiên cứu.

Từ khóa: Đa hình nucleotide đơn (SNP), nhóm đơn bội O, nhiễm sắc thể Y, người Việt Nam

ĐẶT VẤN ĐỀ

Với đặc điểm là di truyền theo phụ hệ (dòng bố) và không xảy ra hiện tượng trao đổi chéo trên phần lớn chiều dài của nhiễm sắc thể, nhiễm sắc thể Y là công cụ hữu hiệu để nghiên cứu về di truyền quần thể theo phụ hệ, tương tự như mtDNA trong nghiên cứu di truyền quần thể theo mẫu hệ (dòng mẹ). Sự đa hình của nhiễm sắc thể Y không chỉ có ý nghĩa khoa học trong phân tích đa dạng di truyền quần thể, nghiên cứu nguồn gốc tiến hóa loài người, mà nó còn có ý nghĩa đối với việc điều tra tội phạm, giám định hình sự và pháp y. Nghiên cứu sự đa hình của nhiễm sắc thể Y ở người còn cho phép chúng ta tìm hiểu về các bệnh liên quan đến nhiễm sắc thể Y như chứng thiếu hoặc không có khả năng sinh tinh trùng, ung thư và một số bệnh thường gặp khác...

Các nghiên cứu về tiến hóa người cũng như di truyền quần thể người trên cơ sở vùng đặc hiệu giới tính nam trên nhiễm sắc thể Y (Male Specific region on Y chromosome - MSY) của nhiễm sắc thể Y trên thế giới được bắt đầu từ giữa thập kỷ 80 của thế kỷ trước, khi những đa hình hai allele và đa allele trên nhiễm sắc thể này được khám phá. Đặc biệt, từ sau năm 2000, khi các nhà khoa học công bố bản dự thảo về bộ gen người và với sự cải thiện của các kỹ thuật sử dụng trong các nghiên cứu về genome học thì các

nghiên cứu về đa hình nhiễm sắc thể Y cũng được tăng lên một cách không ngừng. Vào giữa năm 2008, tổng số chi thị đa hình 2 allele, chủ yếu là các chi thị đa hình nucleotide đơn (SNP), đã được phát hiện là 599 chi thị, được chia thành 311 nhóm đơn bội (haplogroup) thuộc 20 nhóm phân bố lớn (clade) được đánh số từ A - T (Karafet *et al.*, 2008).

Châu Á là châu lục có số lượng các nhóm đơn bội lớn nhất, với 15 trong tổng số 20 nhóm phân bố lớn có mặt ở châu lục này. Tại khu vực Đông Á, các nhóm đơn bội có tần số phân bố lớn nhất là C, D, O và N (Ding *et al.*, 2000; Hammer *et al.*, 2001; Jin *et al.*, 2003; Jin, Su, 2000; Su *et al.*, 2000; Tajima *et al.*, 2002).

Nhóm đơn bội O có mức độ đa dạng rất cao, được xác định bằng 4 SNP (M175, P186, P191 và P196), có 44 SNP chia thành 30 phân nhóm, và một cận nhóm O*. Nhóm O là nhóm chính ở Đông Á. Nhóm O cũng tìm thấy ở Trung Á và châu Đại Dương với tần số trung bình hoặc thấp (Deng *et al.*, 2004; Karafet *et al.*, 2001; Su *et al.*, 1999; Underhill *et al.*, 2001). Nhóm đơn bội này có thể phát sinh tại Đông Á, sau đó di cư đến vùng Nam Thái Bình Dương. Một nhánh của nhóm đơn bội này mở rộng đến đảo Đài Loan, Indonesia, Melanesia, Micronesia và Polynesia. Nhánh O3 là một phân nhóm chính ở Trung Quốc,

trong khi đó, O1 và O2a được tìm thấy ở các quần thể người ở Đông Nam Á: Malaysia, Việt Nam, Indonesia và Nam Trung Quốc. Nhánh O2b xuất hiện với tần số cao ở các nhóm người Nhật Bản và Triều Tiên (Nonaka *et al.*, 2007; Xue *et al.*, 2006).

Cho đến nay, các nghiên cứu trong nước về đặc điểm bộ gen của các dân tộc Việt Nam còn rất ít ỏi và các thông tin có được còn rất hạn chế. Liên quan đến nhiễm sắc thể Y, cho đến nay mới có 3 công bố trên đối tượng người Việt Nam của các tác giả nước ngoài (Karafet *et al.*, 2001; Li *et al.*, 2008a; Scheinfeldt *et al.*, 2006). Tuy nhiên, các công trình này đều được thực hiện ở nước ngoài, với số mẫu cá thể nhỏ. Ở trong nước, cho đến thời điểm này, nghiên cứu đặc điểm cấu trúc phân tử của nhiễm sắc thể Y ở người Việt Nam vẫn là vấn đề hoàn toàn mới.

Để tìm hiểu về tần số phân bố của nhóm đơn bội O ở người Việt Nam, nghiên cứu này tiến hành phân tích 10 chỉ thị SNP là đại diện của nhóm đơn bội O* (M175, P186, P191 và P196), O1 (M119, M50, M103 và M110) và O2b (SRY465 và JST022454).

NGUYÊN LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Mẫu nghiên cứu

Các nghiên cứu được tiến hành với 248 mẫu máu ngoại vi của các cá thể nam khỏe mạnh, tình nguyện cho máu để nghiên cứu khoa học, có độ tuổi từ 18 đến 50, thuộc dân tộc Kinh thu thập ở miền Bắc Việt Nam. Nguồn gốc dân tộc dựa trên cơ sở người tình nguyện cho máu tự khai báo 3 đời. Các mẫu máu được bảo quản trên giấy FTA (Flinders Technology Associates) ở nhiệt độ 4°C.

Nguyên liệu, hóa chất

Bộ kit xác định trình tự nucleotide (Bigdye 3.1 terminator) và kit thực hiện phản ứng SBE (SNaPshot) đặt mua của hãng Applied Biosystems, Mỹ.

Các cặp môi dùng để nhân các đoạn chứa các chỉ thị SNP và môi dùng để thực hiện phản ứng SBE được thiết kế dựa trên cơ sở công bố trước đây (Bảng 1) (Sanchez *et al.*, 2003) và trên cơ sở trình tự nucleotide của nhiễm sắc thể Y đã công bố trên ngân hàng Genbank, được đặt tổng hợp tại hãng IDT (Canada). Các SBE primer được tinh sạch qua hệ thống HPLC.

Các hóa chất dùng cho PCR thông thường (dNTP, Taq DNA polymerase, MgCl₂...) của hãng Fermentas.

Phương pháp nghiên cứu

DNA tổng số được tách chiết theo phương pháp Chelex từ mẫu máu ngoại vi được lưu giữ trên giấy FTA (Walsh *et al.*, 1991).

PCR được tiến hành với tổng thể tích 25 µl với 10 ng DNA tổng số, 1X PCR buffer, 8 mM MgCl₂, 700 mM mỗi dNTP, 5 mM primer và 1 U of AmpliTaq Gold[®] DNA polymerase (Applied Biosystems), với chu trình nhiệt: biến tính 94°C trong 5 phút, 35 chu kỳ (95°C - 30 giây, 60°C - 30 giây và 65°C - 30 giây) và 7 phút ở 65°C. Phản ứng được thực hiện trên máy GenAmp PCR system 9700.

Năm µl sản phẩm PCR của mỗi một SNP được trộn với nhau và được tinh sạch để loại bỏ primer và dNTP dư thừa qua cột MiniElute (Qiagen) theo hướng dẫn của nhà sản xuất. Hỗn hợp sản phẩm PCR tinh sạch được hòa 12 µl dung dịch TE 10 mM pH 8,0. Phản ứng SBE được tiến hành theo phương pháp đã được công bố trước đây (Nguyễn Đăng Tôn *et al.*, 2009).

Tần số xuất hiện của từng chỉ thị SNP và tần số của từng nhóm đơn bội được tính trên cơ sở số cá thể xuất hiện đa hình trong tổng số mẫu nghiên cứu, trên cơ sở các công bố trước đây. Thông số AMOVA (Analysis of Molecular variance - AMOVA) (Excoffier *et al.*, 1992) được tính bằng phần mềm Arlequin 3.0 (Excoffier *et al.*, 2005), trên cơ sở kết hợp với các số liệu về đa hình SNP của các dân tộc sống ở khu vực Đông Á. Thông số khoảng cách di truyền (*Fst*) được sử dụng để thiết lập bản đồ phân bố và tương quan (MDS - Multidimensional Scaling) giữa các nhóm người bằng phần mềm STATISTICA 7.1 (StatSoft Inc).

KẾT QUẢ VÀ THẢO LUẬN

Lựa chọn các chỉ thị SNP

Nhóm đơn bội O chiếm khoảng 80 đến 90% nam giới ở khu vực Đông và Đông Nam châu Á, hầu như không tìm thấy ở các nơi khác trên thế giới. Chỉ thị M175 không xuất hiện ở Tây Siberia, Tây Á và châu Âu và đặc biệt là hoàn toàn không tìm thấy ở châu Phi và châu Mỹ, một số nhánh của nhóm đơn bội O xuất hiện với tần số tương đối ở các bộ lạc ở Nam Á, các nhóm người thuộc nhóm ngôn ngữ Altaic ở

Trung Á và các nhóm người thuộc nhóm ngôn ngữ Austronesia ở châu Đại Dương (Deng *et al.*, 2004; Karafet *et al.*, 2001; Su *et al.*, 1999; Underhill *et al.*, 2001).

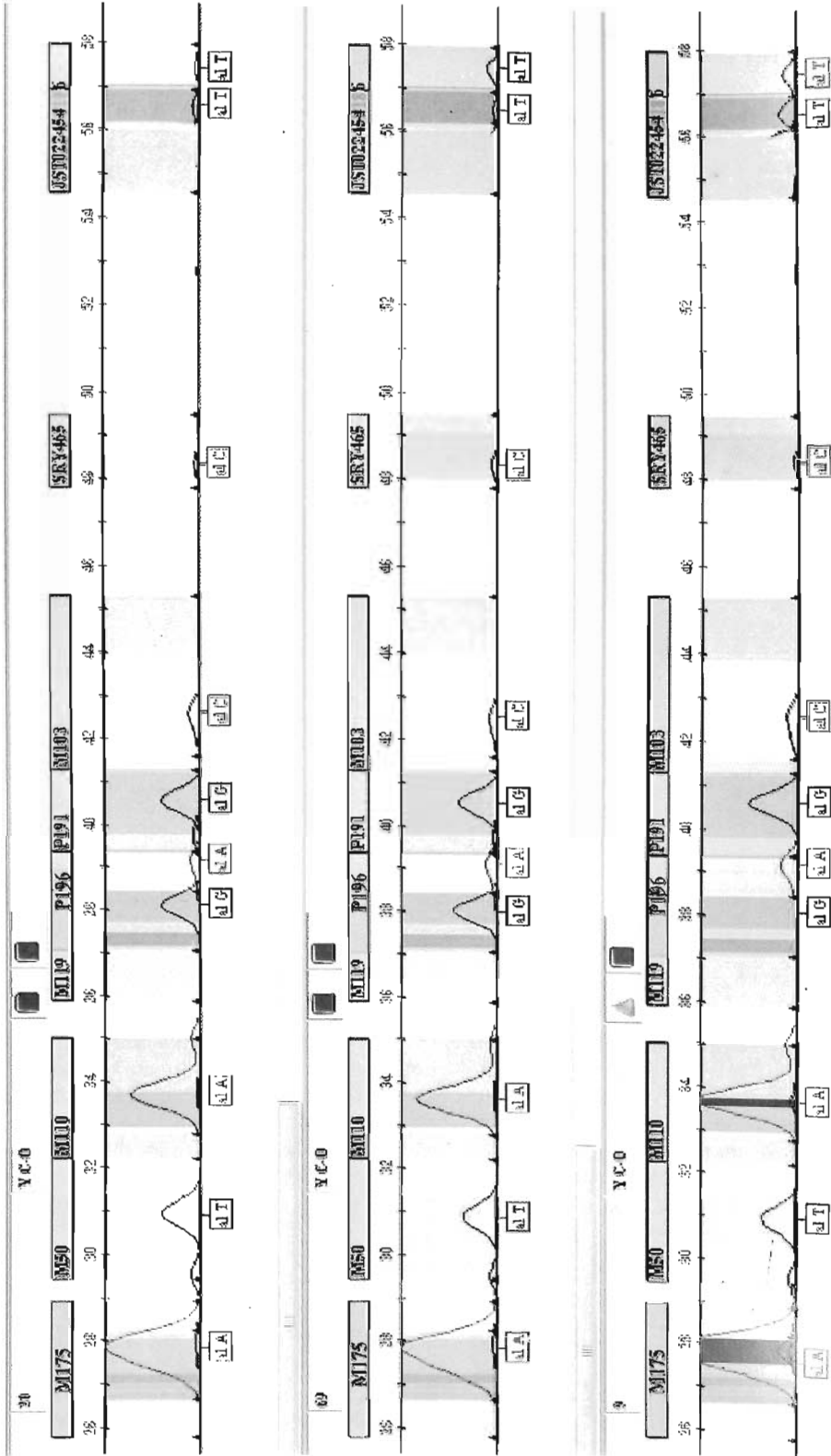
Bảng 1. Các chỉ thị SNP, trình tự primer dùng để nhân đoạn mang các chỉ thị SNP thuộc nhóm đơn bội O.

Chỉ thị SNP	Nhóm đơn bội	Chiều	Primer (5' - 3')	Kích thước (bp)
M175	O	F	ACCCAATCAACTCAACTCC	178
		R	CCATGTACTTTGTCCAATGC	
		S	GCACATGCCTTCTCACTTCTC	
P196	O	F	ACTCAGGGCTGTCTGTTAG	152
		R	GGCCTAGGCATCAAACCTTTA	
		S	GAAAGTCTGACAACAAATCTTAGGGGAACCCATTG	
P191	O	F	TGAGGACTGAGCGAAAATA	177
		R	GGTAGGAGCCAAGAGTGGTA	
		S	AGAAACCTTATTAAGAAAGGTATAGTGTTCAAATGTA	
P186	O	F	TGACATGTCCAATGCTGACA	168
		R	CCTGGTCTCAAATGATCCTC	
		S	CGTGAAAGTCTGACAAAATTGAAATGTGGAAATAATAGA ATTTCCCTGATTTT	
M119	O1a	F	ATGGGTTATTCCAATTCAGC	180
		R	TAAATAGGGAAATGCCAAGG	
		S	GACAGATAATTCTGCTCTTTAAAATTGCTATTTTA	
M50	O1a2	F	TAGAGGCCACTCTTGGTCTT	192
		R	CAGTGTGACCAGTCCTGAGA	
		S	AAGTCTGACAAGCCTACCCAAACCACCC	
M110	O1a2	F	AGACTCACTGGGTTCTCAGC	180
		R	CTTTTCCCCGCTCTAAAGTC	
		S	CACTGTTAATGCAACAGTTTACAAGAACATA	
M103	O1a2	F	TACTGCTCCAAGGACACAGA	209
		R	AATGGCATATTGGGTTCAAG	
		S	CACGTCGTGAAAGTCTGACAACCTCTTGCCCTAGGTTCTCC	
SRY465	O2b	F	GGAGGCACAGAAATTACAGG	237
		R	CTAGCTGGTGTCTCCATTCTT	
		S	GGTGCCACGTCGTGAAAGTCTGACAACGCTTCGGTACTCTGC AG	
JST022454	O2b	F	ACTTATCCTTTGGCCTTCCT	191
		R	GAAACTTGGTGGCATTTCCT	
		S	GGTGCCACGTCGTGAAAGTCTGACAATCCTTTGTA ATTGCAGATAAAGAGGGAT	

Nhóm đơn bội O được xác định bằng 4 chỉ thị SNP là M175, P186, P191 và P196 (Karafet *et al.*, 2008). Để tìm hiểu về sự phân bố của nhóm đơn bội O ở người Việt Nam, chúng tôi tiến hành phân tích sự phân bố của cả 4 chỉ thị trên. Chỉ thị M175 có đa hình là - 5bp (\pm /TTCTC) từ vị trí 14018100 đến 14018104 trên nhiễm sắc thể Y. Chỉ thị P186 và P196 có đa hình C \rightarrow A tương ứng với vị trí 7628568 và 14263707, còn chỉ thị P191 có đa hình là A \rightarrow G ở vị trí 13924509 trên nhiễm sắc thể Y.

Các nghiên cứu trước đây chỉ ra rằng, các nhánh O1 và O2 có tần số phân bố phổ biến ở khu vực Đông Nam Á, trong khi đó nhánh O3 phân bố ở Trung Quốc và một số nơi của châu Đại Dương.

Cùng với 4 chỉ thị xác định nên nhóm đơn bội O, trong nghiên cứu này cũng tiến hành phân tích một số chỉ thị thuộc các nhánh phát sinh từ O, gồm các nhánh O1a, O2b. Các chỉ thị thuộc các nhóm đơn bội này bao gồm: chỉ thị M119 xác định nhánh O1a; các chỉ thị M50, M103 và M110 thuộc nhánh O1a2; 2 chỉ thị SRY465 và JST022454 thuộc nhóm đơn bội O2b. M119 với điểm đa hình A \rightarrow C tại vị trí 20222073 trên nhiễm sắc thể Y. Hai chỉ thị M50, M110 có điểm đa hình là T \rightarrow C tại vị trí 20328060 và 20366485 còn M103 có đa hình là C \rightarrow T tại vị trí 20395526 trên nhiễm sắc thể Y. Hai chỉ thị thuộc nhóm đơn bội O2b là SRY465 và JST022454 có điểm đa hình tương ứng là C \rightarrow T và T \rightarrow C tại vị trí 2715180 và 22919969 trên nhiễm sắc thể Y.



Hình 1. Xác định 10 chỉ thị SNP thuộc nhóm đơn bội O ở một số mẫu nghiên cứu. M175, M50, M119, P196, P191, M103, SRY465, JS1022454 và P186 là các chỉ thị SNP thuộc nhóm đơn bội O ở đã được phân tích. Trên hình là kết quả xác định các chỉ thị SNP trên hệ thống máy giải trình tự tự động ABI 3100 Genetic Analyzer. A, G, al G, al C, al T chỉ các đỉnh ứng với các SNP tại vị trí đa hình.

Xác định chỉ thị SNP

Trong 10 chỉ thị SNP thuộc nhóm đơn bội O và các nhánh thuộc nhóm này được nghiên cứu (Hình 1), đã phát hiện thấy 6 chỉ thị SNP có đa hình. Cả 4 chỉ thị xác định nhóm đơn bội O là M175, P186, P191 và P196 đều xuất hiện đa hình trong các mẫu nghiên cứu. Trong đó, chỉ thị có mức độ đa hình lớn nhất là P191, với 157 trong tổng số 248 mẫu nghiên cứu mang đa hình G, chiếm 63,31%. Tiếp theo là chỉ thị M175, với 25 cá thể mang đa hình mất 5 bp (-TTCTC), chiếm 10,08%. Hai chỉ thị còn lại là P186 và P196 có mức độ đa hình thấp với tần số tương ứng là 1,61 và 4,03% trong tổng số các mẫu đã được phân tích. Chỉ thị M119 xác định nhánh O1a và JST022454 thuộc nhánh O2b cũng xuất hiện đa hình trong các mẫu phân tích. Trong tổng số 248 mẫu nghiên cứu, 156 mẫu có đa hình C ở chỉ thị M119, chiếm 62,9%, trong khi đó chỉ thị JST022454 có 21 cá thể mang đa hình, chiếm 8,47%. Các chỉ thị thuộc nhóm O1a1 là M50, M103 và M110 và chỉ thị SRY465 thuộc nhóm O2b không xuất hiện đa hình trong nghiên cứu này.

Nhóm đơn bội O*-M175/P186/P191/P196

Theo phương pháp phân loại nhóm đơn bội trên nhiễm sắc thể Y đã được đưa ra năm 2002 (Y Chromosome Consortium 2002) và được cập nhật lại năm 2008 (Karafet *et al.*, 2008), nhóm đơn bội O được xác định bằng 4 chỉ thị (M175, P186, P191 và P196). Cận nhóm O* chỉ mang 1 trong 4 chỉ thị đã nêu ở trên, không mang các đa hình thuộc các nhánh phát sinh từ O. Từ những số liệu thu được cho thấy, có 3 cá thể mang các đa hình thuộc cận nhóm O*-M175, 1 cá thể mang đa hình thuộc cận nhóm O*-P186 và 21 thuộc cận nhóm O*-P191 với tần số phân bố tương ứng là 1,2, 0,4 và 8,46%. Trong số 248

mẫu nghiên cứu không có cá thể nào thuộc cận nhóm O*-P196. Tuy nhiên, trong nghiên cứu này, chúng tôi chưa có điều kiện để phân tích toàn bộ các chỉ thị của các nhánh được sinh ra từ nhóm đơn bội O, vì vậy những số liệu thống kê trên đây có thể bị thay đổi khi có những phân tích đầy đủ về đa hình các chỉ thị của các nhánh thuộc nhóm đơn bội này.

Do chưa có công bố nào liên quan đến 3 chỉ thị P191, P186 và P196 nên tần số phân bố của nhóm O* được tính bằng tổng tần số của O-M175, O-P186, O-P191 và O-P196. Cách tính này cũng đã được một số tác giả trên thế giới đã sử dụng trong các công trình của mình. Với cách tính như vậy thì tần số của nhóm đơn bội O* ở số mẫu nghiên cứu người Kinh đã nghiên cứu là 10,10%. Các công bố trước đây cũng chỉ ra rằng, nhóm đơn bội O* phân bố khá rộng rãi ở các dân tộc sống tại khu vực Đông Á với tần số phân bố trung bình từ 0,68% ở người Mông Cổ đến 5,77% ở người Manchu (Bảng 2). Tuy nhiên, do trước đây, các tác giả mới chỉ phân tích có một chỉ thị M175 nên tần số phân bố của nhóm đơn bội này thấp hơn so với số liệu được công bố trong công trình này.

Trong một điều tra tổng thể về đa hình nhiễm sắc thể Y ở các nhóm người sống trong khu vực trung đại lục Á - Âu, người ta đã tìm thấy nhóm đơn bội O* - M175 xuất hiện với tần số: 2,5% (1/40 cá thể) ở người Tajik ở Samarkand, 4,5% (1/22) ở người Crimean Tatar sống ở Uzbekistan, 1,5% (1/68) ở người Uzbek sống ở Surkhandarya, 1,4% (1/70) ở người Uzbek sống ở Khorezm, 6,3% (1/16) của người Tajik sống ở Dushanbe, 1,9% (1/54) của người Kazakh sống ở Kazakhstan, 4,5% (2/41) của người Uyghur sống ở Kazakhstan và 31,1% (14/45) của người Triều Tiên (Karafet *et al.*, 2005).

Bảng 2. Tần số phân bố của các nhóm đơn bội trên nhiễm sắc thể Y ở 44 quần thể người sống ở Đông Á.

Các nhóm dân tộc	Tần số xuất hiện các nhóm đơn bội (%)					
	n	O*	O1*	O1b-M50	O2b*	Khác
Đông Bắc Á						
Nhật Bản	47		3,40		4,26	62,56
Triều Tiên	43	1,35	4,00		32,43	41,29
Tạng	75	1,33				62,67
Ewenki	26		7,69			61,54
Hani	34					82,35
Hui	20		8,57			22,46
Xibe	41		7,32		2,44	53,66

Hán ở phía Bắc	44		9,80		85,65
Daur	39	2,56	5,13		61,54
Oroqen	31	3,23			77,42
Buryat	81				37,37
Manchu	35	5,77	2,86		5,38
Mông Cổ	147	0,68	0,68		42,86
Buyi	35		2,86		91,43
Qiang	33		15,15		66,67
Uygurs	68		2,94		61,76
Đông Nam Á					
Balinese	551		18,00		79,80
Java	53	2,00	23,00		73,00
Malaysia	50		8,00		86,00
Moluccas	34		6,00		7,00
Nusa Tenggara	31		23,00		54,00
Philippines	115		28,00		67,00
S. Borneo	40	5,00	15,00		74,00
Ami	53	4,00	43,00		53,00
Atayal	50		92,00	8,00	
Batak	18		22,20		77,80
Bunun	50		74,00		26,00
Campuchia	26		3,80	3,80	84,80
Dong	10		2,00	1,00	7,00
Đông Indonesia	55		13,00		54,00
Kinh	248	10,1	62,93		8,47
Li	34	5,88	27,30		66,82
Mông	57		7,18		87,72
Thái Lan	20		5,00	5,00	85,00
Paiwan	53	4,00	96,00		
She	34		14,76		82,35
Southern Han	280		16,80		83,20
Đài Loan bản địa	43	5,00	79,00		16,00
Hán sống ở Đài Loan	82		1,98		81,77
Tujian	49		8,16		67,35
Yami	40		7,00		28,00
Yao	60		1,67		78,33
Yizu	43				97,67
Zhuang	28	5,26	17,90		76,84

Ghi chú: n: số cá thể được tiến hành nghiên cứu; Các số liệu này được tổng hợp từ các công bố khác nhau (Karalet *et al.*, 2001; 2005; Kayser *et al.*, 2001; 2003; 2006; Li *et al.*, 2008a; 2008b; Su *et al.*, 1999; Xue *et al.*, 2006).

Nhóm đơn bội O1a*-M119

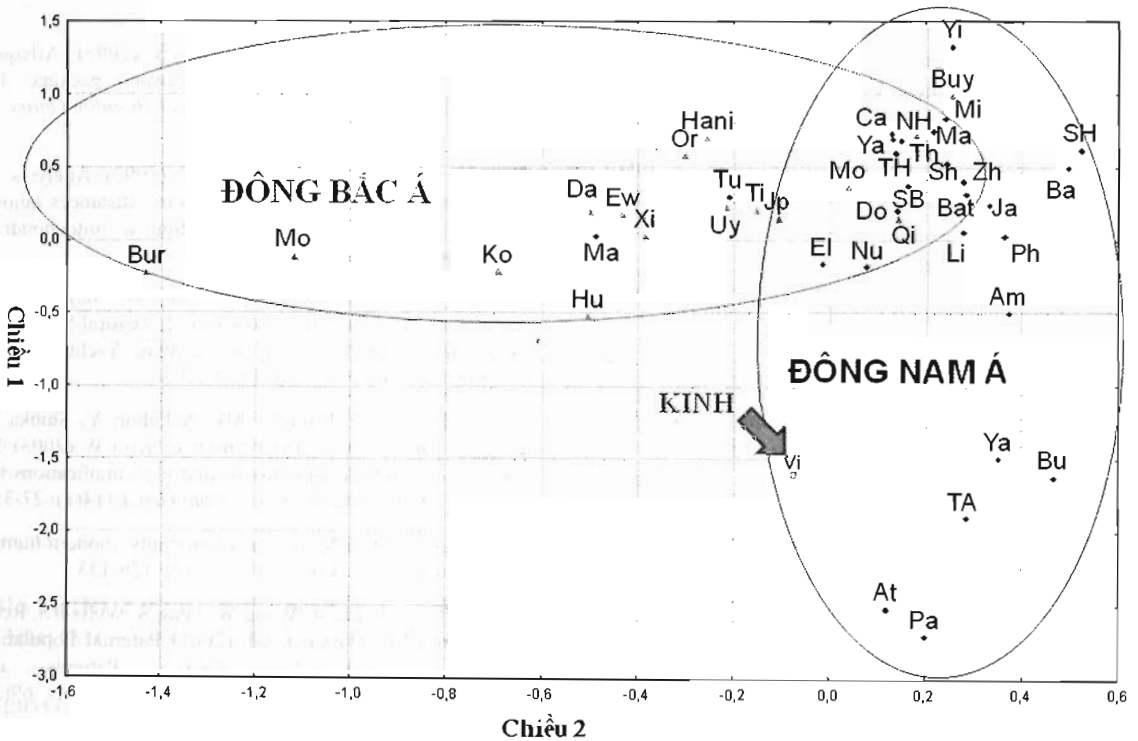
Kết quả phân tích ở cũng cho thấy, có 156 cá thể có đa hình thuộc nhánh O1a*-M119, với tần số phân bố là 62,90%. Nhánh O1a được chia ra thành 2 nhánh nhỏ là O1a1 và O1a2. O1a2 được xác định bằng 3 chỉ thị SNP là M50, M103 và M110. Cả 3 chỉ thị này đều không xuất hiện đa hình trong các mẫu nghiên cứu.

Có thể nói, nhóm đơn bội O1a*-M119 có tần số phân bố tương đối lớn và phân bố địa lý rộng ở khu vực Đông Á. Với 44 dân tộc sống ở khu vực này được các tác giả khác nhau nghiên cứu có 39 dân tộc có mặt nhóm đơn bội này với tần số từ 0,68% ở người Mông Cổ sống ở Đông Bắc Á cho đến 96% ở người Paiwan (dân bản địa sống ở Đài Loan) (Li *et al.*, 2008a; b).

Nhóm đơn bội O2b*-SRY465/JST022454

Nhóm đơn bội O2b được xác định bằng hai chỉ thị SRY46 JST022454, được tìm thấy chủ yếu ở Đông Bắc Á, từ người Daur của Mông Cổ đến người Nhật của Nhật Bản. O2b cũng được tìm thấy ở các nhóm người ở Đông Nam Á, trong đó có Indonesia, Thái Lan và Việt Nam và một số ở Bangladesh. Nhóm đơn bội này có tần số phân bố cao nhất và có mức đa dạng nhất ở người Nhật Bản và Triều Tiên, tuy nhiên, nhóm đơn bội này hiếm hoặc rất ít xuất hiện ở các dân tộc người sống ở Trung Quốc.

Trong 2 chỉ thị xác định nhánh O2b là SRY465 và JST022454 chỉ có chỉ thị JST022454 là có xuất hiện đa hình trong các mẫu nghiên cứu, với tần số xuất hiện là 8,47%.



Hình 2. Biểu đồ hai chiều tương quan giữa các nhóm dân tộc sống ở các khu vực địa lý khác nhau (MDS - Multidimensional Scaling). Mo: Mông Cổ; Ko: người Hàn; Ba: Balinese, Ja: Java, Ma: Malaysia, Mo: Moluccas, Nu: Nusa Tenggara, Ph: Philippines, SB: S. Borneo, At: Atayal, Bat: Batak, Bunun, Bur: Buryat, Buy: Buyí, Ca: Campuchia, Da: Daur, Do: Dong, In: Indonesia, Ew: Ewenki, Hu: Hui, Jp: Nhật Bản, Vi: Kinh, Ma: Manchu, Mi: Mông, TH: người Thái ở Đông Bắc Thái Lan, NH: Hán phía Bắc, Or: Oroqen, Pa: Paiwan, Qi: Qiang, Sh: She, SH: Hán phía Nam, TA: Đài Loan bản địa, HT: Hán Đài Loan, Ti: người Tạng, Tu: Tujian, Uy: Uyгур, Xi: Xibe, Ya: Yami, Yi: Yizu, Zh: Zhuang.

Sự phân bố của các nhóm đơn bội nhiễm sắc thể Y ở người Việt Nam

Trong số 10 chi thị SNP thuộc nhóm đơn bội O được xác định trên tổng số 248 mẫu nghiên cứu là các cá thể thuộc dân tộc Kinh/Việt sống ở miền Bắc Việt Nam, 6 chi thị SNP xuất hiện đa hình trong các mẫu nghiên cứu. Các chi thị có đa hình được xếp vào 3 nhóm đơn bội với tần số xuất hiện từ 8,47% ở nhóm đơn bội O2b cho đến 62,93% (156/248) ở nhóm đơn bội O1a*-M119.

Mức độ đa dạng di truyền chung trên cơ sở phân tích bằng phần mềm Arlequin là 0.5801 ± 0.0343 cho thấy, người Kinh ở Việt Nam có mức độ đa dạng di truyền lớn. Để tìm hiểu mối tương quan giữa dân tộc Kinh/Việt sống ở miền Bắc Việt Nam so với các dân tộc khác sống ở khu vực Đông Nam Á và Đông Bắc Á, chúng tôi đã tiến hành so sánh các số liệu thu được trong nghiên cứu này với các số liệu đã được công bố trước đây (Karafet *et al.*, 2001; Kayser *et al.*, 2006, Su *et al.*, 1999, Xue *et al.*, 2006) (Bảng 2). Toàn bộ số liệu về tần số phân bố các nhóm đơn bội của các dân tộc chúng tôi tiến hành phân tích trong nghiên cứu này được chia thành 2 nhóm theo phân bố địa lý là nhóm Đông Bắc Á và Đông Nam Á.

Giá trị khoảng cách di truyền giữa các quần thể dựa trên AMOVA (Analysis of Molecular Variance), được tính trên cơ sở tần số phân bố các nhóm đơn bội của nhiễm sắc thể Y của các dân tộc, được sử dụng để xây dựng biểu đồ tương quan giữa các quần thể sống ở các khu vực địa lý khác nhau.

Kết quả phân tích trên biểu đồ MDS (Multidimensional Scaling) (Hình 2), trên cơ sở số liệu về khoảng cách di truyền giữa các dân tộc khác nhau cho thấy, các dân tộc thuộc hai khu vực địa lý, Đông Bắc Á và Đông Nam Á, được phân chia thành 2 khu vực tương đối rõ ràng trên biểu đồ. Dân tộc Kinh/Việt được phân bố cùng với các dân tộc sống ở phía Nam Á.

KẾT LUẬN

Nhóm đơn bội O1a*-M119 có tần số phân bố trong tổng số 248 mẫu nghiên cứu là 62,9%, nhóm đơn bội O*-M175/P186/P191/P196 có tần số phân bố là 10,1%, và O2b*-JST022454 có tần số phân bố là 8,47%. Trong khi đó, nhóm đơn bội O1a2 không có đa hình trong các mẫu nghiên cứu.

Đây là những số liệu đầu tiên về tần số phân bố của nhóm đơn bội O trên nhiễm sắc thể Y ở người

Việt Nam được tiến hành nghiên cứu ở trong nước với số lượng mẫu đủ lớn.

Lời cảm ơn: Công trình này được hỗ trợ kinh phí của Bộ Khoa học và Công nghệ cho nhiệm vụ nghiên cứu thường xuyên của Phòng thí nghiệm trọng điểm Công nghệ gen 2009 - 2010 "Nghiên cứu đặc điểm đa hình nucleotide đơn (SNPs) của hệ gen đơn bội (DNA ty thể và nhiễm sắc thể Y) ở người Việt Nam".

TÀI LIỆU THAM KHẢO

Deng W, Shi B, He X, Zhang Z, Xu J, Li B, Yang J, Ling L, Dai C, Qiang B, Shen Y, Chen R (2004) Evolution and migration history of the Chinese population inferred from Chinese Y-chromosome evidence. *J Hum Genet* 49(7): 339-348.

Ding YC, Wooding S, Harpending HC, Chi HC, Li HP, Fu YX, Pang JF, Yao YG, Yu JG, Moyzis R, Zhang Y (2000) Population structure and history in East Asia. *Proc Natl Acad Sci USA* 97(25): 14003-14006.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol Bioinform Online* 1: 47-50.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479.

Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegural SL (2001) Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol* 18: 1189-1203.

Jin HJ, Kwak KD, Hammer MF, Nakahori Y, Shinka T, Lee JW, Jin F, Jia X, Tyler-Smith C, Kim W (2003) Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans. *Hum Genet* 114(1): 27-35.

Jin L, Su B (2000) Natives or immigrants: modern human origin in east Asia. *Nat Rev Genet* 1(2): 126-133.

Karafet T, Xu L, Du R, Wang W, Feng S, Wells RS, Redd AJ, Zegural SL, Hammer MF (2001) Paternal Population History of East Asia: Sources, Patterns, and Microevolutionary Processes. *Am J Hum Genet* 69(3): 615-628.

Karafet TM, Lansing JS, Redd AJ, Reznikova S, Watkins JC, Surata SP, Arthawiguna WA, Mayer L, Bamshad M, Jorde LB, Hammer MF (2005) Balinese Y-chromosome perspective on the peopling of Indonesia: genetic contributions from pre-neolithic hunter-gatherers, Austronesian farmers, and Indian traders. *Hum Biol* 77(1): 93-114.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, Hammer MF (2008) New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res* 18(5): 830-838.

Kayser M, Brauer S, Cordaux R, Casto A, Lao O, Zhivotovsky LA, Moysé-Faurie C, Rutledge RB, Schiefenhoel W, Gil D, Lin AA, Underhill PA, Oefner PJ, Trent RJ, Stoneking M (2006) Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. *Mol Biol Evol* 23(11): 2234-2244.

Kayser M, Brauer S, Weiss G, Schiefenhoel W, Underhill P, Shen P, Oefner P, Tommaseo-Ponzetta M, Stoneking M (2003) Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *Am J Hum Genet* 72(2): 281-302.

Kayser M, Brauer S, Weiss G, Schiefenhoel W, Underhill PA, Stoneking M (2001) Independent histories of human Y chromosomes from Melanesia and Australia. *Am J Hum Genet* 68(1): 173-190.

Li D, Li H, Ou C, Lu Y, Sun Y, Yang B, Qin Z, Zhou Z, Li S, Jin L (2008a) Paternal Genetic Structure of Hainan Aborigines Isolated at the Entrance to East Asia. *PLoS ONE*, 3(5): 10.1371/journal.pone.0002168.

Li H, Wen B, Chen SJ, Su B, Pramoongjago P, Liu Y, Pan S, Qin Z, Liu W, Cheng X, Yang N, Li X, Tran D, Lu D, Hsu MT, Deka R, Marzuki S, Tan CC, Jin L (2008b) Paternal genetic affinity between western Austronesians and Daic populations. *BMC Evol Biol* 8(146): doi:10.1186/1471-2148-1188-1146.

Nguyễn Đăng Tôn, Nguyễn Thùy Dương, Nông Văn Hải (2009) Sự phân bố các đa hình nucleotide đơn của nhóm đơn bội C trên nhiễm sắc thể Y ở người Việt Nam. *Tap chí Công nghệ Sinh học* 7(1): 11-18.

Nonaka I, Minaguchi K, Takezaki N (2007) Y-chromosomal binary haplogroups in the Japanese population and their relationship to 16 Y-STR polymorphisms. *Ann Hum Genet* 71: 480-495.

Sanchez JJ, Børsting C, Hallenberg C, Buchard A,

Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs—a model with 35 Y chromosome SNPs. *Forensic Sci Int* 137(1): 74-84.

Scheinfeldt L, Friedlaender F, Friedlaender J, Latham K, Koki G, Karafet T, Hammer M, Lorenz J (2006) Unexpected NRY chromosome variation in Northern Island Melanesia. *Mol Biol Evol* 23(8): 1628-1641.

Su B, Jin L, Underhill P, Martinson J, Saha N, McGarvey ST, Shriver MD, Chu J, Oefner P, Chakraborty R, Deka R (2000) Polynesian origins: insights from the Y chromosome. *Proc Natl Acad Sci USA* 97(15): 8225-8228.

Su B, Xiao J, Underhill P, Deka R, Zhang W, Akey J, Huang W, Shen D, Lu D, Luo J, Chu J, Tan J, Shen P, Davis R, Cavalli-Sforza L, Chakraborty R, Xiong M, Du R, Oefner P, Chen Z, Jin L (1999) Y-Chromosome evidence for a northward migration of modern humans into Eastern Asia during the last Ice Age. *Am J Hum Genet* 65(6): 1718-1724.

Tajima A, Pan IH, Fucharoen G, Fucharoen S, Matsuo M, Tokunaga K, Juji T, Hayami M, Omoto K, Horai S (2002) Three major lineages of Asian Y chromosomes: implications for the peopling of east and southeast Asia. *Hum Genet* 110(1): 80-88.

Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65(Pt 1): 43-62.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10: 506-513.

Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, Du R, Fu S, Li P, Hurles ME, Yang H, Tyler-Smith C (2006) Male Demography in East Asia: A North-South Contrast in Human Population Expansion Times. *Genetics* 172(4): 2431-2439.

Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12(2): 339-348.

DISTRIBUTION OF SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs) OF Y CHROMOSOME HAPLOGROUP O IN VIETNAMESE

Nguyen Dang Ton, Nguyen Thuy Duong, Nong Van Hai*

Key Laboratory of Gene Technology, Institute of Biotechnology

SUMMARY

Markers on the non-recombining portion of the human Y chromosome continue to have applications in

* Author for correspondence: Tel: 84-4-37562934; Fax: 84-4-38363144; E-mail: vhnong@ibt.ac.vn

many fields including evolutionary biology, forensics, medical genetics, and genealogical reconstruction. Vietnam has 54 ethnic nationalities, of which Kinh is the majority with 87% of population. Until now, Y chromosome and mtDNA haplogroups determining paternal and maternal lineages, respectively, in Vietnamese have not been investigated. In this report, we present data on distribution and frequency of Y haplogroups in Kinh ethnic nationality of Vietnam. To investigate a distribution of Y chromosome variation in Kinh ethnic nationality, ten binary markers on the Y chromosome M175, P186, P191, P196, M119, M50, M103, M110, SRY465, JST022454 belonging to haplogroup O, were genotyped by primer extension in 248 normal male samples, collected from North Vietnam provinces. The results showed that six of the 10 binary markers gave polymorphisms in the samples. Haplogroup O1a*-M119 was observed with highest frequency (62.9%), while haplogroups O*-M175/P186/P191/P196 was with 10.1% distribution. Haplogroups O1a2 (M50, M103, and M110) and O2b-SRY465 were not detected. Taking together, Y chromosome of Vietnamese Kinh ethnic nationality was most frequently distributed in the haplogroup O1a*, followed by haplogroups O*, and O2b. The haplogroup O1a2 was not found in the samples investigated.

Keywords: Haplogroup O, Y chromosome, single nucleotide polymorphisms (SNPs), Vietnamese