

BỘ GIÁO DỤC VÀ ĐÀO TẠO

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



HÀ HỮU HẢO

**NGHIÊN CỨU MỘT SỐ CHỈ THỊ TRÊN
NHIỄM SẮC THỂ Y ĐỂ ỨNG DỤNG TRONG GIÁM
ĐỊNH PHÁP Y**

LUẬN ÁN TIẾN SĨ SINH HỌC

HÀ NỘI – 2023

**VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM
HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ**

.....***.....

**NGHIÊN CỨU MỘT SỐ CHỈ THỊ TRÊN
NHIỄM SẮC THỂ Y ĐỂ ỨNG DỤNG TRONG GIÁM
ĐỊNH PHÁP Y**

Chuyên ngành: Công nghệ sinh học

Mã số: 9 42 02 01

LUẬN ÁN TIẾN SỸ SINH HỌC

Người hướng dẫn khoa học:

1. GS.TS. Chu Hoàng Hà
2. PGS.TS. Lê Văn Sơn

Hà Nội – 2023

LỜI CAM ĐOAN

Tôi xin cam đoan:

Luận án là công trình nghiên cứu của tôi và một số kết quả cùng nghiên cứu, cộng tác với các nhà khoa học khác;

Các số liệu và kết quả trình bày trong luận án là trung thực, một phần đã được công bố trên các tạp chí khoa học chuyên ngành cũng như các hội nghị trong nước và quốc tế với sự đồng ý và cho phép của đồng tác giả; những kết quả còn lại trong luận án chưa được tác giả nào công bố trong bất kỳ công trình nào khác.

Hà Nội, ngày tháng năm 2023

Nghiên cứu sinh

Hà Hữu Hào

LỜI CẢM ƠN

Lời đầu tiên, tôi xin bày tỏ lòng biết ơn sâu sắc tới GS.TS. Chu Hoàng Hà - Viện Công nghệ sinh học - Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã hướng dẫn, chỉ bảo, giúp đỡ tận tình trong suốt quá trình tôi thực hiện luận án. Thầy không chỉ truyền thụ cho tôi nhiều kiến thức chuyên môn mà còn giúp tôi bồi đắp lòng say mê, sự nghiêm túc, tính cẩn thận trong nghiên cứu khoa học. Đó là nền tảng cho quá trình thực hiện luận án là hành trang giúp tôi tự tin vững bước trên con đường khoa học sau này.

Tôi cũng xin được bày tỏ lòng biết ơn sâu sắc đến PGS.TS. Lê Văn Sơn - Viện Công nghệ sinh học - Viện Hàn lâm Khoa học và Công nghệ Việt Nam - Người thầy đã sát sao điều dẫu, truyền lại cho tôi phương pháp mới về nghiên cứu phân tích trình tự hệ gen ty thể cũng như niềm say mê nghiên cứu về Hệ gen học.

Tôi xin chân thành cảm ơn sự giúp đỡ của Ban phụ trách đào tạo Học viện Khoa học và Công nghệ và Viện Công nghệ sinh đã tận tình hướng dẫn tôi hoàn thành mọi thủ tục trong suốt quá trình học tập, làm nghiên cứu sinh tại học viện.

Tôi cũng xin chân thành cảm ơn TS.BS. Nguyễn Đức Như, Viện trưởng Viện Pháp y quốc gia đã nhiệt tình giúp đỡ, động viên và tạo điều kiện thuận lợi cho tôi trong suốt thời gian thực hiện luận án này.

Trong suốt quá trình thực hiện Đề tài nghiên cứu, tôi đã nhận được sự giúp đỡ tận tình về chuyên môn của các nhà khoa học, các cán bộ nghiên cứu công tác tại Khoa Y sinh học – Viện Pháp y quốc gia. Họ đã luôn bên cạnh giúp đỡ và cổ vũ tôi trong suốt thời gian qua.

Với tất cả lòng biết ơn, tôi xin dành cho bố mẹ, gia đình và bạn bè đã luôn tin tưởng, thông cảm, động viên, tạo điều kiện và chia sẻ khó khăn trong thời gian qua, giúp tôi hoàn thành tốt luận án này.

Nghiên cứu sinh

Hà Hữu Hào

MỤC LỤC

MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN	4
1.1. Phân tích ADN trong giám định pháp y	4
1.2. Các loại chỉ thị phân tử (marker).....	5
1.3. Tổng quan chỉ thị STR.....	8
1.3.1. Đặc điểm trong bộ gen	8
1.3.2. Phân loại STR	9
1.4. STR trên nhiễm sắc thể giới tính Y	11
1.5. Các hướng ứng dụng của chỉ thị STR trên nhiễm sắc thể giới tính Y	12
1.5.1. Y-STR trong nghiên cứu cấu trúc di truyền quần thể (genetic structure).....	12
1.5.2. Ứng dụng các Y-STR trong xác định huyết thống	14
1.5.3. Ứng dụng của các Y-STR trong lĩnh vực khoa học hình sự	15
1.6. Phương pháp phân tích các chỉ thị Y-STR.....	19
1.6.1. Phân tích dựa trên phương pháp điện di mao quản.....	19
1.6.2. Phân tích sử dụng các bộ kit Y-STR thương mại hóa	25
1.6.3. Phân tích sử dụng chiến lược mini STR	28
1.7. Tầm quan trọng của việc tính toán tần suất alen các chỉ thị STR	30
1.8. Tình hình nghiên cứu các chỉ thị Y-STR.....	31
1.8.1. Tình hình nghiên cứu trên thế giới.....	31
1.8.2. Tình hình nghiên cứu tại Việt Nam	35
CHƯƠNG 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU	37
2.1. Vật liệu nghiên cứu	37
2.1.1. Mẫu nghiên cứu.....	37
2.1.2. Hoá chất nghiên cứu.....	37
2.2. Thiết bị chính được sử dụng	39
Quantus™ Fluorometer.....	39

2.3. Phương pháp nghiên cứu.....	39
2.3.1. Phương pháp tách chiết ADN	39
2.3.2. Phương pháp định lượng ADN	42
2.3.3. Phương pháp điện di gel.....	42
Phương pháp điện di được sử dụng để xác định sự có mặt của ADN đồng thời phân biệt sự khác nhau về kích thước giữa những đoạn ADN quan tâm.	42
2.3.4. Phương pháp PCR.....	43
2.3.5. Phương pháp điện di mao quản.....	45
2.3.6. Phương pháp giải trình tự gen.....	46
2.3.7. Phân tích và xử lý số liệu	46
2.4. Đạo đức trong nghiên cứu.....	47
CHƯƠNG III. KẾT QUẢ VÀ BÀN LUẬN.....	48
3.1. Kết quả tách chiết ADN tổng số từ các mẫu nghiên cứu.....	49
3.1.1. Mẫu tách chiết theo phương pháp Chelex.....	49
3.1.2. Mẫu tách chiết bằng kit thương mại	50
3.2. Kết quả khảo sát 29 chỉ thị Y-STR từ 2 bộ kit PPY23 và Yfiler Plus	51
3.2.1. Xây dựng hồ sơ Y-STR.....	51
3.2.2. Bảng phân bố tần suất alen thuộc 29 chỉ thị Y-STR.....	53
3.2.3. Đánh giá đặc điểm các alen thuộc 29 chỉ thị Y-STR.....	70
3.2.4. Độ đa hình của 29 chỉ thị Y-STR.....	73
3.2.5. Độ đa dạng haplotype (HD) và khả năng phân biệt.....	76
3.3. Kết quả nghiên cứu một số chỉ thị mini Y-STR mới.....	77
3.3.1. Lựa chọn chỉ thị mini Y-STR	77
3.3.2. Tối ưu hoá phản ứng PCR.....	78
3.3.3. Giải trình tự xác định cấu trúc lặp các mini Y-STR	81
3.3.4. Tỷ lệ khuếch đại thành công các mini Y-STR.....	83
3.4. Hiệu quả của các chỉ thị Y-STR trong xét nghiệm ADN	86
3.4.1. Hiệu quả trong phân tích mẫu đã bị phân huỷ	86

3.4.3. Hiệu quả sử dụng Y-STR trong phân tích mẫu lẫn (mixture sample)	95
3.4.1. Hiệu quả trong việc tăng khả năng phân biệt giữa các cá thể.....	98
3.4.4. Ứng dụng chỉ thị Y-STR để so sánh khoảng cách di truyền giữa các quần thể	102
KẾT LUẬN VÀ KIẾN NGHỊ	107
KẾT LUẬN	107
KIẾN NGHỊ	108
DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN	
LUẬN ÁN TIẾN SĨ	109
TÀI LIỆU THAM KHẢO	110
PHẦN PHỤ LỤC	118

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ CÁI VIẾT TẮT

Tên viết tắt	Tên đầy đủ
ADN	Acid deoxyribonucleic
AMOVA	Analysis of molecular variance (phân tích phương sai phân tử)
AS-STR	Autosomal STR (STR trên nhiễm sắc thể thường)
Bp	Base pair (Cặp base)
CE	Capillary electrophoresis (Điện di mao quản)
CRI	Combined Paternity Index (chỉ số có quan hệ huyết thống kết hợp)
DC	Discriminating capacity (chỉ số về khả năng phân biệt)
DI	Degradation index (Chỉ số phân huỷ)
ADN	Deoxyribonucleic acid
HD	Haplotype diversity (Độ đa dạng Haplotype)
HUGO	Human Genome Organisation (Tổ chức về bộ gen người)
GD	Gene diversity (độ đa dạng gen)
ISFN	International Society for Forensic Genetic (Hiệp hội Di truyền Pháp y)
LR	Likelihood ratio (tỷ số tương đồng)
MDS plot	Multidimensional scaling plot (Đồ thị phân bố không gian đa chiều)
MHL	Minimal haplotype loci (Bộ haplotype tối thiểu)
Multiplex	Multiplex Polymerase chain reaction (Phản ứng PCR đa môi)
NIST	National Institute of Standards and Technology (Viện tiêu chuẩn và
NGS	Next generation sequencing (giải trình tự gen thế hệ mới)
NST	Nhiễm sắc thể
OL	Off ladder (Lệch thang)
PAGE	Polyacrylamide gel electrophoresis (Điện di trên gel polyacrylamide)
PCR	Polymerase chain reaction (Phản ứng chuỗi trùng hợp)
PD	Power of discrimination (Khả năng phân biệt)
PP Y23	PowerPlex [®] Y23 System
RFLP	Restriction Fragment Length Polymorphism (đa hình chiều dài đoạn

RM Y-STR	Rapidly mutating Y – STR (Y-STR có tốc độ đột biến nhanh)
RFU	Relative fluorescence units (đơn vị đo tín hiệu huỳnh quang tương đối)
SNP	Single nucleotide polymorphism (đơn hình nucleotit)
SSR	Simple sequence repeats (Trình tự lặp lại đơn giản)
STR	Short tandem repeat (Trình tự lặp lại ngắn)
SWGAM	Scientific Working Group on ADN Analysis Methods (Hiệp hội các
VNTRs	Variable Number of Tandem Repeats (tiểu vệ tinh)
X-STR	STR trên nhiễm sắc thể giới tính X
YHRD	Y-STR Haplotype Reference Database (Cơ sở dữ liệu tham khảo
Y-STR	Y Chromosome - Short tandem repeat (STR trên nhiễm sắc thể giới
YPlus	Yfiler™ Plus PCR Amplification Kit (bộ kit khuếch đại Yfiler™ Plus
Yfiler	AmpFLSTR™ Yfiler™ PCR Amplification Kit

DANH MỤC HÌNH VẼ, ĐỒ THỊ

Hình 1. 1. Cấu trúc lặp điển hình của một STR	8
Hình 1. 2. Mô hình di truyền các marker theo kiểu tái tổ hợp (trên NST thường), theo dòng cha (trên NST Y) và theo dòng mẹ (trên ty thể)	12
Hình 1. 3. Mô hình biểu thị sự phân hoá các nhóm Y haplotype khác nhau bắt nguồn từ 1 tổ tiên chung [17]	13
Hình 1. 4. Cấu trúc MDS plot và cây phân loại mô tả mối quan hệ giữa các quần thể dựa trên dữ liệu Y – halotype [18]	14
Hình 1. 5. So sánh việc lập hồ sơ ADN dựa trên STR trên NST thường và trên NST Y.....	16
Hình 1. 6. Ví dụ về 1 hồ sơ Y-STR từ 1 cá thể với mỗi locus chỉ gồm 1 alen (ngoại trừ locus DYS385 gồm tổ hợp 2 locus DYS385a và b)	18
Hình 1. 7. Minh họa quá trình khuếch đại cùng lúc 3 locus sử dụng 3 cặp mồi khác nhau trong phản ứng multiplex PCR	21
Hình 1. 8. Các bước hoạt động chính của một hệ thống điện di mao quản [35].....	23
Hình 1. 9. Kết quả phân tích Y-STR từ 1 mẫu nam giới với các locus DYS389I, DYS439, DYS437, DYS389II bị đột biến lặp đoạn với 2 alen trong 1 locus Y-STR chỉ hơn kém nhau 1 đơn vị lặp [35]	25
Hình 1. 10. Minh họa về vị trí khuếch đại cùng 1 locus STR giữa cặp mồi PCR kích thước lớn và cặp mồi mini STR [35]	30
Hình 1.11. Sự gia tăng số lượng hồ sơ Y-STR trên YHRD.org từ năm 1999 đến 2018 (số năm biểu thị từ 1 đến 59 [56]	33
Hình 2.1. Vị trí 29 locus Y-STR trên bản đồ NST Y	44
Hình 2.2. Sơ đồ thí nghiệm trong nghiên cứu	48
Hình 3. 1. Ảnh điện di ADN tổng số tách chiết bằng Chelex® 100 trên gel agarose 0.8%	49
Hình 3. 2. Dạng đỉnh lặp xuất hiện ở locus DYS481 khi phân tích với bộ PPY23 (mẫu phân tích ký hiệu ID101) và Dạng đỉnh gai xuất hiện ở locus DYS533 khi phân tích với bộ Yfiler Plus (mẫu phân tích ký hiệu ID56).....	52

Hình 3. 3. Số liệu thống kê số haplotype đóng góp từ quần thể người Việt Nam trên YHRD	70
Hình 3. 4. Số alen của 29 locus Y-STR dựa trên 2 bộ kit PPY23 và YPlus	71
Hình 3. 5. Các trường hợp mất locus trong hồ sơ Y-STR (A) vị trí mất 18/29 locus Y-STR trên bản đồ NST Y (B) hồ sơ Y-STR của trường hợp mất 4/23 locus Y-STR.....	72
Hình 3. 6. Xếp hạng độ đa hình các chỉ thị Y-STR theo giá trị từ cao đến thấp.....	75
Hình 3. 7. Ảnh điện di sản phẩm sau PCR với 10 locus Y-STR trên gel Polyacrylamide 6% (Sử dụng thang chuẩn ILS 500 (Promega - Mỹ)).....	79
Hình 3. 8. Ảnh điện di trên gel polyacrylamide các phản ứng PCR đa mồi	80
Hình 3. 9. Kết quả giải trình tự locus DYS505	81
Hình 3. 10. Kết quả giải trình tự locus DYS508	82
Hình 3. 11. Kết quả giải trình tự locus DYS522	82
Hình 3. 12. Kết quả giải trình tự locus DYS388	82
Hình 3. 13. Kết quả tạo thang chuẩn mini Y-STR với chỉ thị DYS643.....	83
Hình 3. 14. Hồ sơ Y-STR của mẫu hài cốt sử dụng bộ kit PPY23	84
Hình 3. 15. So sánh hồ sơ Y-STR từ mẫu hiện trường (trái) và mẫu nghi phạm (phải) vụ án kí hiệu V234	87
Hình 3. 16. So sánh mini STR từ mẫu hiện trường và mẫu nghi phạm vụ án kiệu V213 ..	88
Hình 3. 17. So sánh hồ sơ Y-STR giữa mẫu nạn nhân (trái) và mẫu nghi phạm (phải) vụ án ký hiệu V319	89
Hình 3. 18. So sánh mini STR từ mẫu mẫu nạn nhân và mẫu nghi phạm vụ án.....	89
Hình 3. 19. So sánh mini STR từ mẫu mẫu hài cốt và mẫu thân nhân	93
Hình 3. 20. Hồ sơ STR phân tích từ mẫu dịch âm đạo có dạng mẫu lẫn nhiều nguồn ADN	96
Hình 3. 21. Đồ thị so sánh khả năng tạo hồ sơ Y-STR từ mẫu lẫn từ bộ PPY23 và YPlus	97
Hình 3. 22. Sơ đồ biểu thị khoảng cách di truyền giữa các quần thể trong không gian 2 chiều	105

DANH MỤC CÁC BẢNG BIỂU

Bảng 1.1. Danh sách các locus Y-STR có mặt trong các bộ kit thương mại phổ biến hiện nay [20]	27
Bảng 2.1. Kit và hoá chất chính sử dụng trong nghiên cứu	37
Bảng 2.2. Trình tự mỗi khuếch đại 10 mini Y-STR trong nghiên cứu.....	38
Bảng 2.3. Thiết bị chính được sử dụng trong nghiên cứu	39
Bảng 3. 1. Kết quả realtime PCR của mẫu xương, răng ký hiệu R1-R5.	51
Bảng 3. 2. Tần suất phân bố chung của 25 locus Y-STR thuộc hai bộ kit PPY23 và YPlus (4 chỉ thị DYS385a/b và DYF387S1 được thống kê riêng).....	53
Bảng 3. 3. Tần suất phân bố các alen thuộc locus DYS576.....	54
Bảng 3. 4. Tần suất phân bố các alen thuộc locus DYS389I.....	55
Bảng 3. 5. Tần suất phân bố các alen thuộc locus DYS448.....	55
Bảng 3. 6. Tần suất phân bố các alen thuộc locus DYS389II.....	56
Bảng 3. 7. Tần suất phân bố các alen thuộc locus DYS19.....	56
Bảng 3. 8. Tần suất phân bố các alen thuộc locus DYS391.....	56
Bảng 3. 9. Tần suất phân bố các alen thuộc locus DYS481.....	57
Bảng 3. 10. Tần suất phân bố các alen thuộc locus DYS549.....	57
Bảng 3. 11. Tần suất phân bố các alen thuộc locus DYS533.....	58
Bảng 3. 12. Tần suất phân bố các alen thuộc locus DYS438.....	58
Bảng 3. 13. Tần suất phân bố các alen thuộc locus DYS437.....	59
Bảng 3. 14. Tần suất phân bố các alen thuộc locus DYS570.....	59
Bảng 3. 15. Tần suất phân bố các alen thuộc locus DYS635.....	60
Bảng 3. 16. Tần suất phân bố các alen thuộc locus DYS390.....	60
Bảng 3. 17. Tần suất phân bố các alen thuộc locus DYS439.....	60
Bảng 3. 18. Tần suất phân bố các alen thuộc locus DYS392.....	61
Bảng 3. 19. Tần suất phân bố các alen thuộc locus DYS643.....	61
Bảng 3. 20. Tần suất phân bố các alen thuộc locus DYS393.....	61
Bảng 3. 21. Tần suất phân bố các alen thuộc locus DYS458.....	62

Bảng 3. 22. Tần suất phân bố các alen thuộc locus DYS456.....	62
Bảng 3. 23. Tần suất phân bố các alen thuộc locus YGATA-H4.....	63
Bảng 3. 24. Tần suất phân bố các alen thuộc locus DYS460.....	63
Bảng 3. 25. Tần suất phân bố các alen thuộc locus DYS627.....	63
Bảng 3. 26. Tần suất phân bố các alen thuộc locus DYS518.....	64
Bảng 3. 27. Tần suất phân bố các alen thuộc locus DYS449.....	64
Bảng 3. 29. Tần suất phân bố các tổ hợp alen thuộc locus DYS385a/b.....	65
Bảng 3. 30. Tần suất phân bố các alen thuộc locus DYF387S1.....	66
Bảng 3. 31. Tần suất phân bố các tổ hợp alen thuộc locus DYF387S1.....	66
Bảng 3.32: Hồ sơ Y-STR của 2 mẫu so sánh từ nghi phạm và dấu vết hiện trường..	68
Bảng 3.33: Hồ sơ Y-STR của 2 mẫu so sánh từ chú và cháu trai	68
Bảng 3. 34. Thống kê số lượng Haplotype Y-STR ở một số nước trên thế giới được công bố trên YHRD	69
Bảng 3. 35. Độ đa dạng gen của 29 locus Y-STR trong nghiên cứu và so sánh với giá trị trung bình của quần thể người châu Á	73
Bảng 3. 36. Độ đa dạng của các mini Y-STR mới trong một số quần thể	78
Bảng 3. 37. Kết quả tối ưu hóa nồng độ mỗi khuếch đại 10 locus Y-STR	79
Bảng 3. 38. Bảng thống kê tỷ lệ khuếch đại thành công 10 locus mini Y-STR trên 30 mẫu đã bị phân hủy	85
Bảng 3.39: So sánh trình tự vùng HV2 giữa mẫu hài cốt và mẫu thân nhân	91
Bảng 3.40: So sánh trình tự vùng HV1 giữa mẫu hài cốt và mẫu thân nhân	92
Bảng 3. 41. So sánh các chỉ số di truyền thu được từ tập hợp các chỉ thị Y-STR có kích thước < 220 bp trong 2 bộ PPY23 và YPlus.....	94
Bảng 3. 42. Danh sách 12 locus mới có trong bộ kit PPY23 và YPlus.....	99
Bảng 3. 43. So sánh số lượng haplotype quan sát được và các chỉ số thống kê thu được từ tập hợp các marker Y-STR trong các bộ haplotype từ dữ liệu 200 mẫu với bộ YPlus	100

Bảng 3. 44. So sánh số lượng haplotype quan sát được và các chỉ số thống kê thu được từ tập hợp các marker Y-STR trong các bộ haplotype từ dữ liệu 200 mẫu với bộ PPY23	101
Bảng 3. 45. Kết quả phân tích khoảng cách di truyền theo cặp dựa trên các giá trị Rst giữa quần thể Việt Nam trong nghiên cứu với các quần thể được chọn. Giá trị P được hiển thị phía trên đường chéo, giá trị rst phía dưới đường chéo	104
Bảng 3.46. Khoảng cách di truyền giữa các quần thể người Việt Nam	105

MỞ ĐẦU

Tính cấp thiết của luận án

Ngày nay cùng với sự phát triển vượt bậc của công nghệ sinh học, phân tích ADN đã trở thành công cụ đắc lực không thể thiếu trong xét nghiệm mối quan hệ huyết thống, giám định pháp y, điều tra hình sự... Việc phân tích ADN giúp giải quyết từ những vụ việc mang tính dân sự như xác định cha cho con, tranh chấp quyền thừa kế... đến các vụ án hình sự giết người, hiếp dâm... ADN là bằng chứng có tính pháp lý cao và có giá trị quyết định trong các phiên tòa. Hơn thế nữa, trong những vụ thiên tai, thảm họa với số lượng người tử vong lớn hay khi việc nhận dạng thông thường không thể thực hiện thì giám định ADN để xác định danh tính nạn nhân trở thành biện pháp khả thi nhất.

Sự lựa chọn các chỉ thị phân tử là các đoạn gen (locus gen) có tính bền vững, đa hình cao, mang đặc trưng cho từng cá thể trở thành đối tượng cho việc nghiên cứu ứng dụng ADN trong khoa học điều tra, pháp y. Các đoạn lặp lại ngắn - STR (short tandem repeat) - là những trình tự ADN lặp lại liên tiếp với mỗi đơn vị lặp lại gồm 1 - 6 bp được nghiên cứu, lựa chọn và trở thành chỉ thị phân tử phổ biến nhất hiện nay. Các locus STR đã được khảo sát, hệ thống lại tạo thành các bộ kit thương phẩm với khả năng khuếch đại đồng thời nhiều locus STR một cách nhanh chóng, chính xác, có độ nhạy cao.

Song song với các locus STR trên cặp nhiễm sắc thể thường, việc nghiên cứu ứng dụng các locus STR trên nhiễm sắc thể giới tính X và Y cũng đang thu hút sự quan tâm của các nhà khoa học. Các STR trên NST giới tính Y (gọi tắt là Y-STR) có đặc điểm chính là chỉ di truyền theo dòng cha mà không có sự tái tổ hợp. Phân tích Y-STR giúp giải quyết nhiều trường hợp khi việc sử dụng các STR trên nhiễm sắc thể thường không thể đáp ứng được yêu cầu giám định, đặc biệt là giám định ADN đối với các mẫu từ nam giới như trong việc xác định mối quan hệ cha - con, anh - em trai, các vụ án hiếp dâm...

Hiện nay việc phân tích Y-STR chủ yếu thực hiện qua các bộ kit thương mại hoá mà phổ biến nhất là bộ PowerPlex® Y23 system (hãng Promega) phân tích 23 Y-STR và bộ Yfiler™ Plus PCR Amplification Kit (hãng Thermo Fisher Scientific) phân tích 27 Y-STR. Hai bộ kit tỏ rõ hữu hiệu với nhiều ứng dụng khác nhau: phân biệt cá thể nam trong quần thể, xác định quan hệ huyết thống theo dòng cha, giám định hình sự trên mẫu có nồng độ

ADN thấp, mẫu lẫn từ nhiều nguồn, mẫu vi vết. Do cấu trúc di truyền của từng quần thể là khác nhau nên việc đánh giá độ đa hình, tính phù hợp của các STR nói chung và Y-STR nói riêng là điều cần thiết trước khi áp dụng trong quần thể. Mỗi phòng thí nghiệm, viện nghiên cứu được khuyến cáo nên xây dựng bảng tần suất phân bố các alen với từng quần thể người khác nhau nhằm phục vụ cho việc tính toán độ tin cậy, xác suất có quan hệ huyết thống. Việc sử dụng các bộ kit trên để nghiên cứu và ứng dụng các chỉ thị Y-STR trong khoa học kỹ thuật, hình sự, giám định huyết thống đã được áp dụng phổ biến trên thế giới. Tuy vậy tại Việt Nam hiện nay vẫn chưa có nhiều nghiên cứu chi tiết để khảo sát tần suất phân bố, độ đa hình, tính phù hợp, khả năng ứng dụng của các chỉ thị Y-STR trên quần thể đại diện cho nam giới người Việt. Các nghiên cứu mới chỉ dừng lại ở số ít các chỉ thị trên quy mô mẫu nhỏ. Việc khảo sát thêm các locus Y-STR nhằm phục vụ cho việc tính toán độ tin cậy hồ sơ Y-STR, xác suất có quan hệ theo dòng cha, ứng dụng trong khoa học hình sự là thực sự cần thiết và có ý nghĩa thực tiễn cao. Vì vậy chúng tôi tiến hành thực hiện đề tài “Nghiên cứu một số chỉ thị trên nhiễm sắc thể Y để ứng dụng trong giám định pháp y” với mục tiêu:

1. Lựa chọn và đánh giá các đặc điểm liên quan (tần suất phân bố alen, độ đa hình, khả năng phân biệt cá thể ...) của 29 chỉ thị Y-STR trong quần thể nam giới Việt Nam dân tộc Kinh.
2. Khảo sát một số chỉ thị có kích thước nhỏ trên NST Y (mini Y-STR) nhằm ứng dụng trong phân tích các mẫu có ADN đứt gãy nhiều, mẫu đã bị phân hủy.
3. Đánh giá tính ứng dụng của các chỉ thị Y-STR trên nhiều loại mẫu khác nhau (mẫu lẫn, mẫu có chất lượng kém, mẫu đã phân hủy).

Nội dung nghiên cứu

1. Lựa chọn đối tượng nghiên cứu, tiến hành thu thập mẫu và tách chiết ADN.
2. Thực hiện quy trình phân tích chỉ thị Y-STR từ các mẫu nghiên cứu. Lập bảng phân bố tần suất alen của các chỉ thị locus Y-STR và tính toán các chỉ số liên quan: số alen, độ đa hình, độ đa dạng haplotype, khả năng phân biệt cá thể.
3. Lựa chọn và tối ưu điều kiện khuếch đại một số mini STR trên NST Y và bước đầu ứng dụng trong công tác giám định ADN.

4. Khảo sát khả năng tạo chỉ thị Y-STR trên nhiều loại mẫu giám định khác nhau như mẫu lần từ nhiều nguồn, mẫu hài cốt lâu năm, mẫu vi vết.

Đóng góp mới của luận án

1. Đây là nghiên cứu đầu tiên ở Việt Nam khảo sát toàn diện 29 chỉ thị Y-STR trong quần thể nam giới dân tộc Kinh, Việt Nam với tổng cộng 400 mẫu nghiên cứu và số chỉ thị Y-STR lớn hơn so với nghiên cứu trước đây trong quần thể người Kinh, Việt Nam. Nghiên cứu đã xây dựng bảng phân bố tần suất alen của 29 chỉ thị Y-STR là cơ sở để tính toán độ hiếm của một hồ sơ ADN trong quần thể cũng như chỉ số có quan hệ họ hàng (kinship index) giữa các hồ sơ ADN. Kết quả nghiên cứu đã tìm ra những chỉ thị Y-STR có số alen lớn, độ đa hình cao, tiềm năng ứng dụng lớn trong phân biệt giữa các cá thể nam như DYS385, DYS387S1, DYS518, DYS 627, DYS458 và các chỉ thị có alen hiếm gặp trong quần thể.

2. Đây là nghiên cứu đầu tiên ở Việt Nam thực hiện việc khảo sát các mini Y-STR - là các chỉ thị có kích thước ngắn dưới 200 bp, có hiệu quả cao trong việc phân tích các mẫu đã bị phân hủy so với các chỉ thị Y-STR thông thường trong bộ kit thương mại.

3. Nghiên cứu đã khảo sát tiềm năng ứng dụng của các chỉ thị Y-STR trong bộ kit thương mại và các mini Y-STR theo nhiều hướng mới trong lĩnh vực giám định pháp y tại Việt Nam bao gồm: giám định các mẫu có quan hệ huyết thống theo dòng cha, giám định mẫu trong khoa học hình sự đặc biệt là với trường hợp mẫu lần giữa nam và nữ, giám định mẫu lâu năm, mẫu đã bị phân hủy. Bên cạnh đó kết quả từ 400 hồ sơ Y-STR còn được ứng dụng để xây dựng khoảng cách di truyền giữa quần thể người Việt Nam trong nghiên cứu với các nhóm quần thể khác trên thế giới.

CHƯƠNG 1. TỔNG QUAN

1.1. Phân tích ADN trong giám định pháp y

Giám định pháp y là một ngành khoa học sử dụng những thành tựu trong lĩnh vực y học, sinh học, hoá học, vật lý học, tin học... để đáp ứng những yêu cầu của pháp luật trong hoạt động tố tụng hình sự và dân sự thông qua hoạt động giám định khi được các cơ quan trung cầu. Các thành tựu từ lĩnh vực sinh học và công nghệ sinh học được đặc biệt ứng dụng trong lĩnh vực giám định pháp y, phản ánh những đặc trưng riêng biệt của từng cá thể để từ đó có thể truy nguyên cá thể [1].

Từ những năm 1930, dấu vân tay được đưa vào ứng dụng trong pháp y và trở thành công cụ cần thiết trong phòng xét nghiệm để nhận dạng cá thể. Tuy nhiên dấu vân tay cũng bộc lộ nhiều khiếm khuyết khi áp dụng trong thực tế như chỉ có thể thu mẫu ở đầu ngón tay, có thể bị thay đổi qua phẫu thuật, hung thủ không để lại dấu vân tay tại hiện trường....

Chỉ trong vòng hơn 60 năm kể từ khi cấu trúc ADN được Watson và Crick công bố ngành sinh học đã phát triển mạnh mẽ, được ứng dụng rộng rãi. Một trong những ứng dụng quan trọng là sử dụng đặc tính ADN đặc trưng của từng cá thể vào giám định pháp y để nhận dạng, phân biệt giữa các cá thể. Được bắt đầu vào giữa những năm 1980 kỹ thuật phân tích ADN được coi là một cuộc cách mạng trong khoa học hình sự của nhân loại. Giống như dấu vân tay, mỗi người đều có đặc trưng riêng về cấu trúc di truyền của mình được xác định bằng trình tự nucleotide trên ADN. Việc phân tích ADN nhằm xác định các đặc trưng riêng của từng cá thể hình thành nên thuật ngữ kỹ thuật là dấu vân ADN (DNA fingerprint), dấu vân di truyền (genetic fingerprinting) hay lập hồ sơ ADN (ADN profiling) và mở ra hướng nghiên cứu mới là giám định ADN [2]. Kỹ thuật này lần đầu tiên được mô tả vào năm 1984 bởi nhà khoa học người Anh Alec Jeffreys. Tiến sĩ Jeffreys nhận thấy tại các khu vực nhất định trong chuỗi ADN có các đoạn ADN được lặp đi lặp lại nhiều lần liên tiếp. Ông cũng phát hiện ra rằng số lượng các trình tự lặp lại có thể khác nhau ở các cá thể nhằm ứng dụng vào việc phân biệt giữa các cá thể. Bằng kỹ thuật xác định sự thay đổi chiều dài của các chuỗi ADN có trình tự lặp lại, Tiến sĩ Jeffreys đã phát hiện ra các kỹ thuật phân tích nhận dạng cá thể từ ADN. Do đó, Alec Jeffreys đã công bố rằng: “ADN là duy nhất ở

mỗi cá thể, trong đó có những cặp base được di truyền từ cha và mẹ sang con. Cấu trúc của ADN không thay đổi từ lúc còn là phôi thai cho đến suốt cuộc đời”.

Cho đến nay giám định ADN trong khoa học hình sự và pháp y đã phổ biến và phát triển mạnh mẽ. Hầu hết các nước đều có phòng xét nghiệm ADN cho mục đích hình sự và dân sự. Ở Việt Nam kỹ thuật phân tích ADN được áp dụng vào những năm cuối của thập niên 90 và thực sự phát triển trong khoảng 10 năm trở lại đây.

Hiện nay, giám định ADN có 2 loại chính theo mục đích ứng dụng đó là: giám định để xác định mối quan hệ huyết thống và giám định pháp y.

+ Giám định ADN để xác định mối quan hệ huyết thống

Mỗi người đều thừa hưởng các trình tự ADN từ bố và mẹ, các trình tự đặc biệt này có thể được sử dụng để xem xét mối quan hệ huyết thống. Tùy theo trình tự được lựa chọn việc giám định ADN có thể xác định nhiều mối quan hệ huyết thống khác nhau:

- Dựa vào các chỉ thị trên NST thường: thường là các đoạn lặp lại trung bình (VNTR) và đoạn lặp lại ngắn (STR) có thể xác định mối quan hệ trực hệ cha/mẹ - con.

- Dựa vào chỉ thị trên NST giới tính X hoặc Y có thể xác định mối quan hệ huyết thống phụ hệ như anh – em trai, ông nội – cháu trai, chị – em gái, bà nội – cháu gái.

- Dựa vào chỉ thị trên ADN ty thể có thể xác định mối quan hệ huyết thống phụ hệ theo dòng mẹ như bà ngoại – cháu ngoại, anh/chị - em cùng mẹ, dì – cháu....

+ Giám định pháp y

Phân tích ADN được ứng dụng trong giám định pháp y, khoa học hình sự với mục đích xác định huyết thống trong những vụ việc dân sự, hình sự, xác định danh tính của hài cốt liệt sĩ, mồ mả bị thất lạc, những nạn nhân bị chết trong các thiên tai, thảm họa hoặc với mục đích xin thị thực di dân. Ngoài ra bằng cách so sánh ADN thu từ dấu vết, bằng chứng thu được ở hiện trường vụ án với ADN của người bị tình nghi có thể xác định người liên quan đến vụ án. Với độ chính xác cao việc giám định ADN trong điều tra tội phạm đang ngày càng phổ biến và là bằng chứng không thể chối cãi trước Tòa án.

1.2. Các loại chỉ thị phân tử (marker)

Về cơ bản chỉ thị (marker) là một dấu hiệu, một đặc trưng có thể nhận biết được giúp chúng ta phân biệt thứ này với thứ khác. Ví dụ, khi muốn phân biệt giữa lúa tẻ và lúa nếp chúng ta có thể quan sát hình thái của hạt; hạt lúa nếp thường tròn hơn hạt lúa tẻ. Khi đó,

hình dạng hạt là một loại chỉ thị, gọi là chỉ thị hình thái (morphological marker). Tương tự, chỉ thị phân tử (molecular marker) hay chỉ thị di truyền (genetic marker) cũng là các dấu hiệu, hoặc các đặc trưng có tính phân biệt giữa các cá thể. Điểm khác biệt là những chỉ thị này không dựa trên hình thái bên ngoài mà là dựa trên sự khác biệt về trình tự ADN của mỗi sinh vật [3].

Chỉ thị phân tử thường là các đoạn ADN ngắn đã biết vị trí trên nhiễm sắc thể, được di truyền cho thế hệ sau. Chỉ thị ADN được hình thành từ các loại đột biến ADN khác nhau như thay thế (đột biến điểm), sắp xếp lại (thêm vào hoặc bớt đi nucleotide) hoặc các sai sót trong sao chép các đoạn ADN lặp lại liền kề. Các chỉ thị ADN thường nằm ở các vùng không phiên mã. Khác với các chỉ thị hình thái và sinh hóa, chỉ thị ADN thường không giới hạn về số lượng, không ảnh hưởng bởi yếu tố môi trường và giai đoạn phát triển của cá thể. Chỉ thị ADN được sử dụng nhiều trong nghiên cứu quan hệ di truyền, phát sinh chủng loại và phân loại phân tử; trong lập bản đồ liên kết di truyền, nhận biết gen; và trong chọn giống, định danh cá thể và xác định mối quan hệ huyết thống giữa các cá thể...

Mỗi loại chỉ thị ADN được phát triển bằng một kỹ thuật tương ứng. Một chỉ thị ADN lý tưởng cần phải có các tiêu chí sau: có độ đa hình cao (có nhiều alen) và phân bố đều trong genome; cho sự phân biệt rõ sự khác nhau về di truyền, tạo nhiều chỉ thị độc lập và chính xác; đơn giản, nhanh và ít tốn kém; cần ít mẫu và ADN; có tính ổn định cao; kết quả phân tích lặp lại thống nhất trong các nghiên cứu, mức độ sai sót thấp nhất, phân tích số liệu dễ và chính xác.

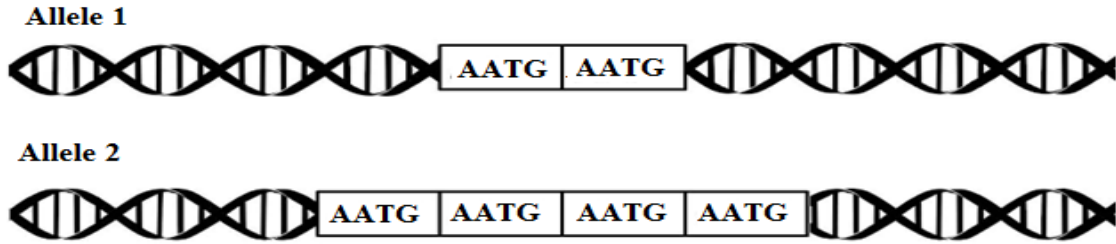
Kể từ khi chỉ thị ADN đầu tiên được phát triển và ứng dụng cho đến cuối những năm 90 của thế kỷ XX, hàng loạt chỉ thị ADN được ra đời [4]. Trong lĩnh vực giám định pháp y, xác định huyết thống có thể kể đến các loại phổ biến bao gồm:

- Các tiểu vệ tinh (minisatellite): Đó là các trình tự base lõi lặp lại gồm 6 – 100bp, gọi là các tiểu vệ tinh VNTRs (Variable Number of Tandem Repeats - Số lượng thay đổi các trình tự base lặp lại), số lần lặp lại có thể đến hàng ngàn lần. Sự sai khác trong số lần lặp lại có thể tạo ra các alen có kích thước từ 500 bp đến hơn 30 kb. Các VNTRs này hiện diện rải rác tại nhiều vị trí trên bộ gen, là tiểu vệ tinh đa vị trí (multilocus minisatellite); hay chỉ có tại một vị trí trên bộ gen, là tiểu vệ tinh đơn vị trí (single-locus

minisatelite) [5]. VNTRs là loại đa hình đầu tiên được sử dụng trong hồ sơ ADN và được sử dụng thành công trong nhiều vụ án giám định trước đây. Xét nghiệm dấu vân tay ADN mà Jeffereys thực hiện năm 1985 là xét nghiệm phát hiện các tiểu vệ tinh đa vị trí bằng kỹ thuật phát hiện sự đa hình về chiều dài các đoạn ADN của bộ gen bị cắt bởi enzyme cắt hạn chế (Restriction Fragments Length Polymorphism = RFLP). Tuy nhiên việc sử dụng VNTRs vẫn có nhiều hạn chế như đòi hỏi lượng ADN đầu vào phải lớn, không áp dụng được với các mẫu đã bị phân hủy, việc diễn giải dữ liệu gặp nhiều khó khăn vì kích thước lớn [6].

- Đa hình đơn nucleotide (SNP-single nucleotide polymorphism): đây là loại đa hình đơn giản nhất, được hiểu là sự khác biệt ở 1 nucleotide trong chuỗi ADN. SNPs được hình thành khi xảy ra đột biến trong quá trình nhân đôi ADN. SNP thường chỉ có 2 alen do đó không có tính đa hình cao và không phù hợp với các đặc tính lý tưởng của 1 chỉ thị phân tử dùng trong pháp y. Tuy nhiên số lượng SNP trong bộ gen là rất lớn nên khi kết hợp phân tích đồng thời nhiều SNP sẽ làm tăng hiệu quả phân biệt giữa các cá thể. Người ta ước tính rằng để đạt được khả năng phân biệt tương đương khi phân tích 10 locus STR thì cần phải phân tích đồng thời từ 50 đến 80 SNP. Với công nghệ hiện nay việc phân tích nhiều SNP vẫn đang bị giới hạn và khá phức tạp [7].

- ADN vi vệ tinh (microsatellites) hay còn lại trình tự các đoạn lặp lại ngắn (short tandem repeat- STR) hoặc đoạn lặp lại đơn giản (simple sequence repeats - SSR). Các đoạn lặp lại ngắn được bắt đầu nghiên cứu từ những năm cuối thập niên 1980, được định nghĩa là những trình tự ADN đặc hiệu có chiều dài khoảng dưới 400 bp (nên gọi là ngắn - short), được cấu thành từ khoảng 5 đến 50 đơn vị lặp lại, mỗi đơn vị lặp lại có chiều dài khoảng 1 đến 6 bp (gọi là đoạn lặp lại - short tandem) (hình 1.1). Với nhiều ưu điểm vượt trội như có kích thước ngắn, độ đa hình cao, kỹ thuật phân tích đơn giản STRs được coi là tiêu chuẩn vàng trong di truyền học pháp y đương đại và sẽ được đề cập sâu trong các phần tiếp theo của luận án [8].



Hình 1. 1. Cấu trúc lặp điển hình của một STR

1.3. Tổng quan chỉ thị STR

1.3.1. Đặc điểm trong bộ gen

Danh pháp: Tên locus STR được đặt theo tên của gen nếu locus này nằm một phần hoặc nằm toàn bộ trong gen. Ví dụ marker STR TH01 từ gen tyrosine nằm trên NST số 11. Chữ “TH” xuất phát từ chữ cái đầu tyrosine hydroxylase. Phần “01” của ký hiệu “TH01” xuất phát từ vùng intron 1 (vùng không mã hóa protein) của gen tyrosine hydroxylase. Đôi khi tiếp đầu ngữ HUM được thêm vào đầu danh pháp của chỉ thị này để xác định đó là từ bộ gen người (Human). Vì vậy, locus STR này sẽ được gọi chính là HUM TH01 hay TH01. Các marker ADN nằm ngoài vùng gen thì được xác định bởi vị trí của chúng trên NST. Ví dụ, các locus STR D5S818 và DYS19 đó là những marker không nằm trong vùng gen. Trong trường hợp này chữ D có nghĩa là ADN. Con số tiếp theo là số thứ tự của NST hoặc tên của NST giới tính X, Y. Chữ “S” thực chất là bản sao duy nhất của ADN marker. Con số cuối cùng là vị trí nucleotide nằm trên NST. Con số này là duy nhất đối với marker ADN nhận dạng cá thể. Ví dụ, STR D16S539 có nghĩa là D: ADN, 16: nhiễm sắc thể số 16, S: trình tự bản sao đơn lẻ (single copy sequence), 539: vị trí thứ 539 xác định trên NST số 16 [9].

Sự phân bố STR: STR được tìm thấy trong hầu hết các loài sinh vật nhân sơ và nhân chuẩn. Ở người, STR xuất hiện trên tất cả 22 cặp nhiễm sắc thể thường và trên cả 2 nhiễm sắc thể giới tính X và Y. STR xuất hiện rải rác trên các NST với ước tính cứ khoảng 2.000 cặp base lại xuất hiện 1 STR trong bộ gen người và chiếm khoảng 3% tổng chiều dài bộ gen. Ước tính có khoảng hơn 1 triệu STR trong bộ gen người.

Phần lớn STR nằm trong vùng không mã hóa (non-coding region) trong khi chỉ khoảng 8% nằm trong vùng mã hóa. Mật độ phân bố các STR trên các NST cũng có sự

khác nhau, ví dụ NST số 19 có mật độ phân bố STR cao hơn so với các NST khác [10]. Ở người, dạng STR phổ biến nhất thường có nhiều nucleotide A, ví dụ như A, AC, AAAN, AAN hoặc AG.

Sự phân bố STR giữa các loài khác nhau cũng khác nhau. Dữ liệu nghiên cứu đã chỉ ra rằng ở chuột và các loài động vật gặm nhấm có tỷ lệ các đoạn lặp lại nhiều hơn hẳn ở người. Mật độ các STR thường có xu hướng tỷ lệ thuận với kích thước của bộ gen. Trong số các loài sinh vật nhân chuẩn được giải trình tự hoàn chỉnh, STR xuất hiện nhiều nhất ở động vật có vú. Tuy nhiên ở thực vật mật độ STR lại tỷ lệ nghịch với kích thước bộ gen với đặc thù là dạng (TA)_n là dạng xuất hiện nhiều nhất.

Tốc độ đột biến của STR: Các chuỗi ADN duy nhất trong hệ gen có tỷ lệ đột biến rất thấp (khoảng 10^{-9} qua mỗi thế hệ), trong khi tỷ lệ đột biến trong chuỗi STR thường cao hơn vài bậc, dao động từ 10^{-6} đến 10^{-2} qua mỗi thế hệ. Tốc độ đột biến STR là đặc trưng đối với từng sinh vật. Ví dụ, tỷ lệ đột biến STR trong nấm men và con người lần lượt là 10^{-5} nt và $10^{-3} - 10^{-5}$ qua mỗi lần phân chia tế bào. Các ước tính trực tiếp về tỷ lệ đột biến microsatellite đã được thực hiện ở nhiều sinh vật, từ côn trùng đến người. Ở loài *Schistocerca gregaria*, tỷ lệ đột biến microsatellite ước tính là $2,1 \times 10^{-4}$ mỗi thế hệ cho mỗi locus [11]. Tỷ lệ đột biến STR trong dòng tế bào sinh tinh của con người cao hơn từ 5 đến 6 lần so với dòng tế bào sinh trứng tế bào sinh trứng và dao động trong khoảng từ 0 đến 7×10^{-3} trên mỗi locus qua mỗi thế hệ.

Sự di truyền của STR: STR nằm trên NST thường và nhiễm sắc thể giới tính X đều được di truyền qua các thế hệ theo quy luật Mendel, bởi vậy tần số các alen của từng locus STR trong quần thể cũng tuân theo đúng định luật cân bằng Hardy – Weinberg [12].

Một cá thể có một locus đồng hợp tử sẽ có cùng số lần lặp lại trên cả hai nhiễm sắc thể, trong khi một cá thể dị hợp tử sẽ có số lần lặp lại khác nhau trên hai nhiễm sắc thể. Những vùng xung quanh locus STR, được gọi là vùng hai bên (flanking region) có thể có cùng trình tự. Điều này rất quan trọng bởi vì những vùng hai bên có thể được dùng như môi của phản ứng PCR khi nó sẽ khuếch đại STR và vùng hai bên này sẽ bảo tồn giữa các giống hay đôi khi giữa các họ trong cùng một loài.

1.3.2. Phân loại STR

Trên cơ sở các đơn vị lặp lại khác nhau, STR có thể được phân loại thành các loại khác nhau. Một mặt dựa theo chiều dài của đơn vị lặp lại chính, các STR được phân loại

thành các dạng lặp lại mono- (1), di- (2), tri- (3), tetra- (4), penta- (5) và hexa- (6) nucleotide. STR phổ biến nhất trong hệ gen của người là dạng lặp lại dinucleotide. Trong số các dinucleotide, (CA)_n là dạng lặp lại thường xuyên nhất, tiếp theo là (AT)_n, (GA)_n và (GC)_n là loại lặp lại hiếm gặp nhất.

Mặt khác dựa theo cấu trúc của đơn vị lặp lại, STR được phân loại thành ba loại chính [13]:

- Dạng lặp lại hoàn hảo (lặp lại đơn giản) chỉ chứa một đơn vị lặp đi lặp lại, ví dụ: (GTG)₁₅.

- Dạng lặp lại không hoàn hảo (lặp lại gián đoạn với sự thay thế bazơ), ví dụ: (GTG)₇CTCTG(GTG)₈.

- Dạng lặp lại phức tạp (gồm nhiều đơn vị lặp lại khác nhau), ví dụ: (GTG)₈(AT)₁₆.

Dựa theo vị trí phân bố trên Nhiễm sắc thể STR có thể được phân thành 3 nhóm chính: STR trên NST thường (autosomal STR – AS STR), nhiễm sắc thể trên NST giới tính X (X-STR) và STR trên NST giới tính Y (Y-STR).

Ngoài ra dựa theo vị trí phân bố có thể phân 3 loại chính:

- STR trên NST thường (autosomal STR – AS STR): đặc điểm chính là thường gồm 2 alen trong 1 vị trí locus. STR trên NST thường được dùng phổ biến trong xét nghiệm mối quan hệ trực hệ và định danh cá thể. Hiện nay có rất nhiều bộ kit cho phép phân tích 24 – 30 STR trên NST thường.

- STR trên NST giới tính X: NST giới tính X và Y là duy nhất và khác với NST thường ở một vài khía cạnh. Ở nữ giới, NST giới tính tạo thành cặp tương đồng XX giống như NST thường. Tuy nhiên, mặc dù cơ thể có nhiều hơn 1 NST X nhưng chỉ 1 NST X trong tế bào được biểu hiện. Những bản sao khác của NST X bị bất hoạt theo giả thuyết Lyon giải thích tại sao cơ thể chỉ có 1 NST X hoặc 3 NST X hoặc nhiều NST X vẫn có khả năng sống sót. Hiện nay, hãng Qiagen đã phát triển bộ kit Investigator Argus X-12 Kit gồm 12 chỉ thị STR trên NST X. Việc phân tích các locus trên NST X làm tăng cơ hội cho các trường hợp yêu cầu phân tích quan hệ huyết thống không trực hệ khi chỉ có mẫu so sánh là một người họ hàng xa, đặc biệt là trong tìm người thân sau chiến tranh hoặc sau các cuộc di dân như bà nội – cháu gái, chị - em gái cùng cha...

- STR trên NST giới tính Y: so với STR trên NST thường và STR trên NST giới tính X, STR trên NST Y có điểm khác biệt nổi bật là chỉ gồm 1 alen trong mỗi vị trí locus. Điều này trở thành ưu điểm lớn trong việc diễn giải dữ liệu, phân tích quan hệ cá thể theo dòng cha.

1.4. STR trên nhiễm sắc thể giới tính Y

STR nằm trên NST Y thường được gọi tắt là Y-STR (Y - short tandem repeat). Locus STR đầu tiên được xác định trên NST Y là DYS19 [14]. Vào khoảng giữa thập niên 90 của thế kỷ XX, chỉ có một số ít các locus Y-STR được phát hiện với khả năng ứng dụng cao trong các nghiên cứu khoa học hình sự. Chúng bao gồm các locus Y-STR có kiểu lặp lại gồm 2 nucleotide (YCA I, YCA II, YCA III, DYS288), kiểu lặp lại gồm 3 nucleotide (DYS461, DYS462) và kiểu lặp lại gồm 4 nucleotide (DYS19, DYS288, DYS385, DYS390, DYS391, DXYS156Y, DYS393). Hầu hết các cặp mồi Y-STR khuếch đại một sản phẩm PCR, trong khi một số cặp mồi Y-STR khuếch đại hai hoặc nhiều sản phẩm PCR (YCA I, YCA II, YCA III, DYF371, DYS385, G10123). Nhận thấy tiềm năng của việc ứng dụng Y-STR trong xác định huyết thống và giám định trong điều tra hình sự các nhà khoa học không ngừng nghiên cứu và tìm ra nhiều locus Y-STR mới. Năm 2006, các nhà khoa học đã tiến hành giải trình tự toàn bộ NST Y với diễn giải đầy đủ và xác định được khoảng hơn 400 marker Y-STR riêng rẽ có tiềm năng ứng dụng trong giám định ADN phục vụ điều tra các vụ án [15].

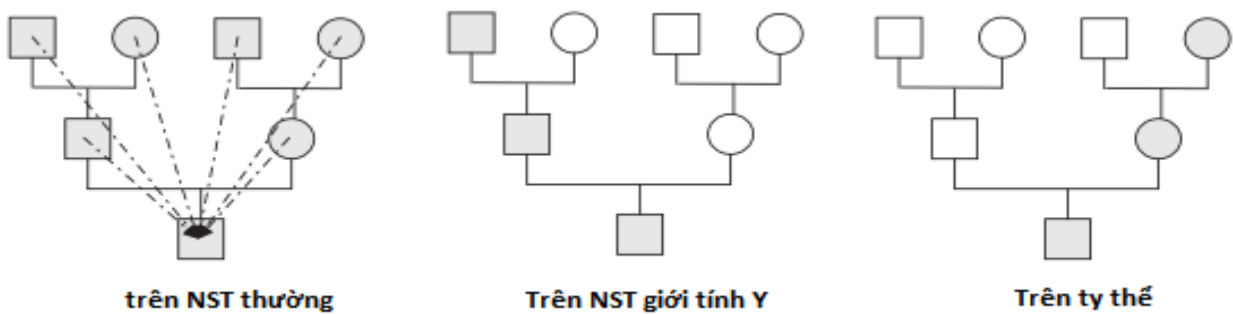
Danh pháp của Y-STR cũng tuân theo quy tắc đặt tên chung cho STR và được thống nhất tên bởi Tổ chức về bộ gen người (Human Genome Organisation – HUGO). Một số công ty thử nghiệm có các định dạng khác nhau cho cách viết các marker Y-STR. Ví dụ: marker DYS455 có thể được viết là DYS455, DYS 455, DYS # 455 hoặc DYS # 455. Tuy nhiên, tiêu chuẩn khoa học được HUGO và NIST (Viện tiêu chuẩn và công nghệ quốc gia Hoa Kỳ) chấp nhận là DYS455.

Ngoài các đặc điểm chung của các STR như đã trình bày ở trên STR trên nhiễm sắc thể Y có một số điểm khác biệt như sau:

Do hầu hết Y-STR nằm trong vùng không tái tổ hợp của NST Y nên với các locus Y-STR sẽ chỉ tồn tại ở dạng đồng hợp tử (mỗi locus sẽ chỉ có 1 alen tương ứng). Đặc điểm

này khiến cho việc phân tích dữ liệu, so sánh đối chiếu giữa các mẫu trở nên đơn giản hơn nhiều so với các STR trên NST thường.

Không giống như các marker STR trên NST thường, các marker STR trên NST Y được liên kết trên cùng một NST và không có sự tái tổ hợp giữa các marker qua các thế hệ [16]. Do đó định luật cân bằng Hardy - Weinberg (thường được áp dụng để tính toán tần suất các alen trên NST thường) không phù hợp để tính toán tần số các alen ở mỗi locus Y-STR. Để xác định được tần số các STR đặc trưng trên NST Y dữ liệu phải được so sánh giữa nhiều tập mẫu khác nhau và trên số lượng đủ lớn để đại diện chính xác cho tần số của các haplotype trong quần thể đang nghiên cứu. Bởi vậy nhiều mẫu Y-STR được tập hợp và bổ sung vào các cơ sở dữ liệu phục vụ hữu ích cho việc tra cứu. Tuy nhiên, hiện nay cơ sở dữ liệu về Y-STR còn quá ít so với các dữ liệu về STR trên NST thường đã thu được.



Hình 1.2. Mô hình di truyền các marker theo kiểu tái tổ hợp (trên NST thường), theo dòng cha (trên NST Y) và theo dòng mẹ (trên ty thể)

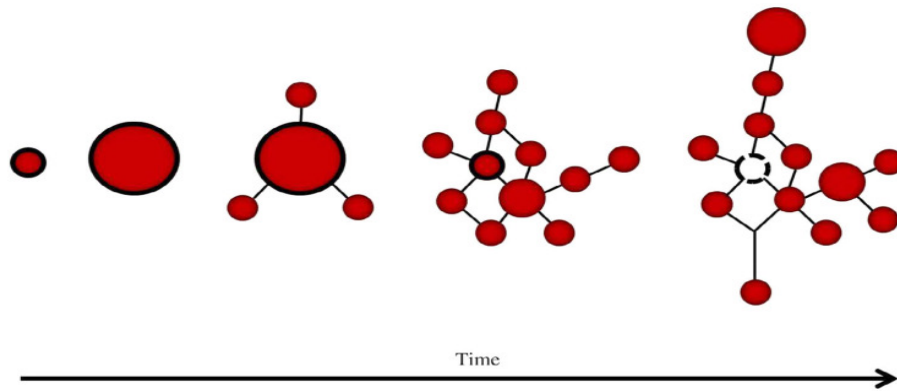
Do NST Y chỉ có ở nam giới nên Y-STR cũng được di truyền theo dòng cha từ ông nội đến bố rồi đến các con trai. Trên lý thuyết nếu không có đột biến xảy ra trên Y-STR tất cả nam giới trong cùng một dòng họ sẽ có kiểu Haplotype giống nhau trong một tập hợp các Y-STR được phân tích (Hình 1.2).

1.5. Các hướng ứng dụng của chỉ thị STR trên nhiễm sắc thể giới tính Y

1.5.1. Y-STR trong nghiên cứu cấu trúc di truyền quần thể (genetic structure)

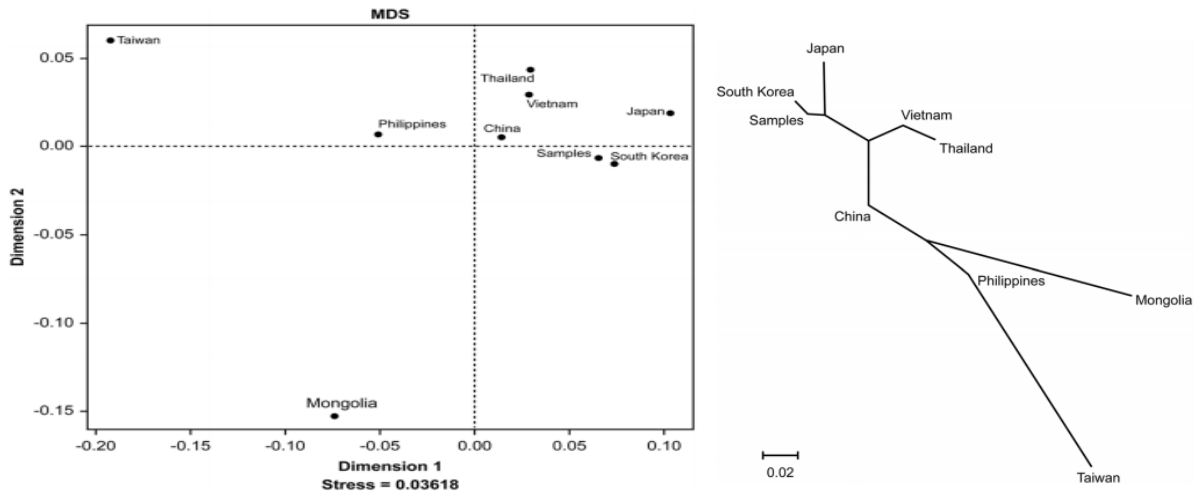
Cấu trúc di truyền của một quần thể được đặc trưng bởi số lượng quần thể trong đó, tần số của các biến thể di truyền khác nhau (alen) trong mỗi quần thể, tần số kiểu gen (genotype) và mức độ phân lập di truyền của các quần thể. Tần số của haplotype trong quần thể cũng rất quan trọng, đặc biệt với tính toán xác suất trong giám định ADN. Việc tăng

cường sử dụng các chỉ thị Y-STR của người ứng dụng trong giám định ADN, trong nghiên cứu nhân chủng học và khảo cổ học đã tạo ra sự nỗ lực hợp tác để thu thập dữ liệu haplotype từ các quần thể khác nhau và để xây dựng các cơ sở dữ liệu ADN trên NST Y. Việc phân tích cơ sở dữ liệu và mô hình hóa phân phối xác suất của các haplotypes Y-STR giúp khám phá ra cấu trúc quần thể có liên quan ẩn trong một quần thể tham chiếu bao gồm các haplotypes được tổ chức trong các cụm gốc có nguồn gốc từ người sáng lập được kết nối bởi các sự kiện đột biến (Hình 1.3).



Hình 1. 3. Mô hình biểu thị sự phân hoá các nhóm Y haplotype khác nhau bắt nguồn từ 1 tổ tiên chung [17]

Dựa vào bảng số liệu tần số alen của từng locus Y-STR trong quần thể nghiên cứu, sử dụng công cụ online AMOVA (phương pháp phân tích phương sai phân tử) trên trang yhrd.org có thể tính toán khoảng cách di truyền theo cặp (giá trị R_{st}) và xác suất kết hợp (P value) giữa quần thể trong nghiên cứu với một số quần thể khác để từ đó thấy được mối quan hệ giữa các quần thể, đặt ra các giả thuyết về nguồn gốc liên quan theo dòng cha giữa các quần thể. Một ví dụ minh họa như trong nghiên cứu quần thể của Kyu Sik Jeong đã sử dụng cơ sở dữ liệu Y-haplotype và công cụ AMOVA để xây dựng cấu trúc MDS plot mô phỏng mối quan hệ giữa các quần thể. Kết quả cho thấy người Việt Nam có quan hệ gần gũi với các quần thể người Thái Lan, Trung Quốc, Hàn Quốc nhưng có khoảng cách xa so với quần thể người Mông Cổ và Đài Loan [18]. Phương pháp neighbor-joining dựa trên giá trị R_{st} cũng được sử dụng để xây dựng cây phân loại của 8 quần thể (Hình 1.4).



Hình 1. 4. Cấu trúc MDS plot và cây phân loại mô tả mối quan hệ giữa các quần thể dựa trên dữ liệu Y – halotype [18]

Bên cạnh các STR trên NST Y trong nghiên cứu cấu trúc di truyền quần thể cũng sử dụng các chỉ thị SNP trên vùng không trao đổi chéo của NST để đánh giá cấu trúc di truyền theo dòng cha giữa các nhóm quần thể khác nhau, nguồn gốc và sự phân tách các nhóm theo nhiều hướng.

1.5.2. Ứng dụng các Y-STR trong xác định huyết thống

“Hồ sơ ADN” (DNA profile) là một phần nhỏ trích từ cấu trúc toàn bộ của chuỗi ADN (của một cá thể) đủ để phản ánh đặc thù riêng nhất của mỗi cá thể không lẫn lộn với những cá thể khác. Và cũng từ đó mà thuật ngữ “Vân tay ADN” (DNA fingerprint) - cũng chính là “hồ sơ ADN”, do một nhà Di truyền học người Anh Alec Jeffreys đề xuất, nói lên một đặc trưng duy nhất cho mỗi cá thể giống như là dấu vân tay (không có hai người bất kỳ có cùng một vân tay) [2]. “Vân tay ADN” của 1 cá thể là hoàn toàn giống nhau ở tất cả các tế bào, tổ chức mô, tạng phủ trong một cơ thể; và vân tay ADN là không thể thay đổi dưới bằng bất cứ tác động nào từ bên ngoài. Việc lập hồ sơ ADN dựa trên các marker STR đã bắt đầu phát triển từ những năm 1990 và trở thành công cụ không thể thiếu trong lĩnh vực giám định pháp lý, định danh cá thể. Nhiều nước trên thế giới đã đề xuất việc sử dụng hồ sơ ADN thay thế cho chứng minh thư/thẻ căn cước thông thường. Tương tự như vậy, một hồ sơ Y-STR cũng có thể được tạo lập. Tuy nhiên, do đặc điểm di truyền theo dòng cha mà không có sự tái tổ hợp nên hồ sơ Y-STR không có giá trị phân biệt các cá thể (PD: power of discrimination) cao như các STR trên NST thường.

Bởi vì các kiểu haplotype Y-STR được chia sẻ giữa những người đàn ông có cùng quan hệ theo dòng cha, nên việc phân tích Y-STR là phù hợp để giải quyết tranh chấp về quan hệ cha con, các câu hỏi liên quan đến quan hệ huyết thống theo dòng cha và tìm kiếm nguồn gốc gia đình [19]. Phân tích Y-STR cũng đặc biệt phù hợp trong các trường hợp thiếu hụt thông tin, một trong những trường hợp đó là người cha giả định của một đứa trẻ nam đã chết và không thể thu mẫu để xét nghiệm. Về lý thuyết, bằng cách sử dụng các Y-STR tiêu chuẩn với tỷ lệ đột biến trung bình, những người thân theo dòng cha của người cha giả định đã chết (như bố, anh/em trai của người cha giả định) sẽ chia sẻ cùng một Y-STR haplotype với người cha giả định và cũng giống với con trai của người cha giả định đã chết. Qua đó thân thế của đứa trẻ có thể được xác định.

Phân tích Y-STR cũng phù hợp với các trường hợp nhận dạng nam giới liên quan đến tử thi, hài cốt như nạn nhân trong thảm họa và nhận dạng người mất tích.

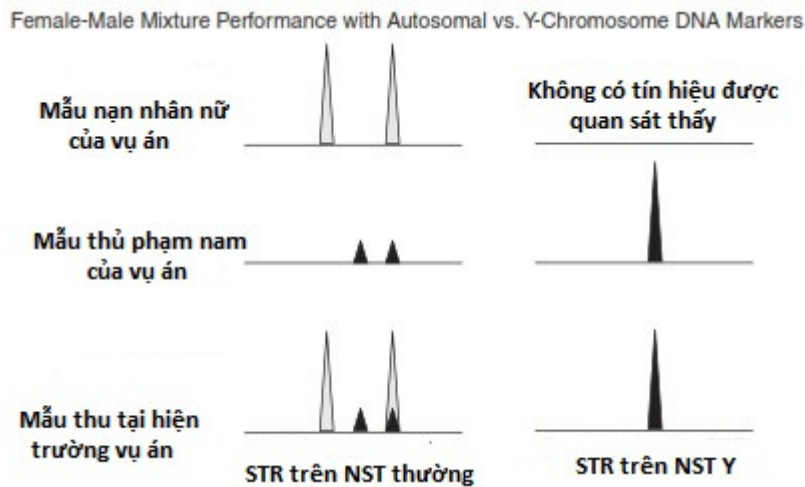
1.5.3. Ứng dụng của các Y-STR trong lĩnh vực khoa học hình sự

Các hồ sơ ADN tiêu chuẩn dựa trên một tập hợp các STR trên NST thường có độ đa hình cao, được lựa chọn kỹ lưỡng là điều kiện lý tưởng trong việc xác định thủ phạm duy nhất để lại các dấu vết tại hiện trường. Ngày nay, các điều tra viên có thể dễ dàng khớp nối hồ sơ STR của những người đã từng phạm tội bị kết án, đã từng có tiền án được lưu trữ trong ngân hàng dữ liệu của cơ quan điều tra với hồ sơ STR thu được từ dấu vết hiện trường vụ án được so sánh và tìm kiếm sự trùng khớp. Tuy nhiên, việc khớp nối hồ sơ tự động như vậy có thể không thành công trong trường hợp thủ phạm thực sự của vụ án chưa được lưu hồ sơ STR trong ngân hàng dữ liệu. Đặc biệt trong trường hợp hồ sơ STR thu từ dấu vết hiện trường là hỗn hợp của nhiều hơn 1 người (vụ án có nhiều hơn một thủ phạm hoặc mẫu lẫn giữa thủ phạm và nạn nhân) thì việc khớp nối hồ sơ STR sẽ gặp nhiều khó khăn. Chỉ có trong điều kiện rất lý tưởng, chẳng hạn như nồng độ ADN của 1 trong 2 thủ phạm có trong dấu vết hiện trường nhiều hơn hẳn người còn lại mới có thể lập nên hồ sơ STR hoàn chỉnh đủ để so sánh khớp nối, trong khi phần lớn các trường hợp khác không thể thực hiện việc so sánh trùng khớp [20]. Lúc này việc mở rộng phân tích Y-STR có thể giúp giải quyết vấn đề.

Bắt đầu từ năm 1992, các STR có độ đa hình cao nằm trên phần không tái tổ hợp của NST giới tính Y đã được công bố và ứng dụng ngay vào việc giải quyết các vụ án hình sự [14]. Ngày càng có nhiều marker Y-STR được nghiên cứu để tạo ra nhiều haplotype Y

khác nhau [21-25]. Hiện nay có tới 27 marker Y-STR phổ biến có trong các bộ kit thương mại (bộ kit Yfiler™ Plus PCR Amplification Kit –hãng Thermo Fisher Scientific) [26].

Do tính đa dạng haplotype cao, các công cụ này cho phép phân tích đặc tính mẫu thu được theo dòng cha. Hơn nữa, các bộ kit Y-STR thương mại này cho phép phát hiện và phân tích ADN từ mẫu nam giới có trong các mẫu hỗn hợp có nồng độ ADN từ người nữ chiếm ưu thế. Cụ thể, trong các vụ án hiếp dâm mẫu thu tại hiện trường thường là hỗn hợp của cả người nam và người nữ (chủ yếu là mẫu dịch trong âm đạo người bị hiếp dâm), trong đó lượng ADN thu từ người nữ thường vượt trội hơn so với nồng độ ADN thu từ người nam (thủ phạm hiếp dâm). Khi phân tích hồ sơ STR trong NST thường do nồng độ ADN vượt trội từ tế bào biểu mô của người nữ và do tiềm năng chia sẻ một số locus STR giống nhau giữa người nam với người nữ nên rất khó có thể lập được hồ sơ STR trên NST thường từ thủ phạm nam của vụ án [27]. Do chỉ xuất hiện ở nam giới nên việc phân tích hồ sơ Y-STR là lựa chọn tối ưu để so sánh giữa mẫu người nghi phạm và mẫu để lại hiện trường.



Hình 1. 5. So sánh việc lập hồ sơ ADN dựa trên STR trên NST thường và trên NST Y

Với mẫu hỗn hợp thu tại hiện trường nếu sử dụng STR trên NST thường rất khó có thể phân tách, lập được hồ sơ ADN của cả nạn nhân và thủ phạm, tuy nhiên nếu dùng Y-STR có thể lập riêng được hồ sơ Y-STR của thủ phạm là nam giới mà không xuất hiện các locus của nạn nhân nữ.

Các khuyến nghị về phân tích pháp y với Y-STR đã được thành lập bởi Ủy ban di truyền học pháp y quốc tế [28,29] và các bộ kit Y-STR đã được kiểm định một cách hợp pháp [30,31] cho phép các điều tra viên không chỉ loại trừ các nghi phạm nam khỏi liên

quan đến việc phạm tội thông qua hồ sơ Y-STR không khớp với mẫu thu tại hiện trường mà còn xác định nguồn gốc gia đình của người nam có trong dấu vết hiện trường. Ví dụ, một nghiên cứu gần đây về hàng trăm trường hợp tấn công tình dục, trong đó việc phân tích hồ sơ Y-STR được áp dụng cùng với hồ sơ STR NST thường cho thấy rằng 1/10 trong số các trường hợp này sẽ vẫn không đưa ra được kết luận nếu không sử dụng các chỉ thị Y-STR và hơn nữa, phân tích Y-STR haplotype có khả năng nhận dạng nhiều người phù hợp gấp ba lần so với hồ sơ STR trên NST thường [32].

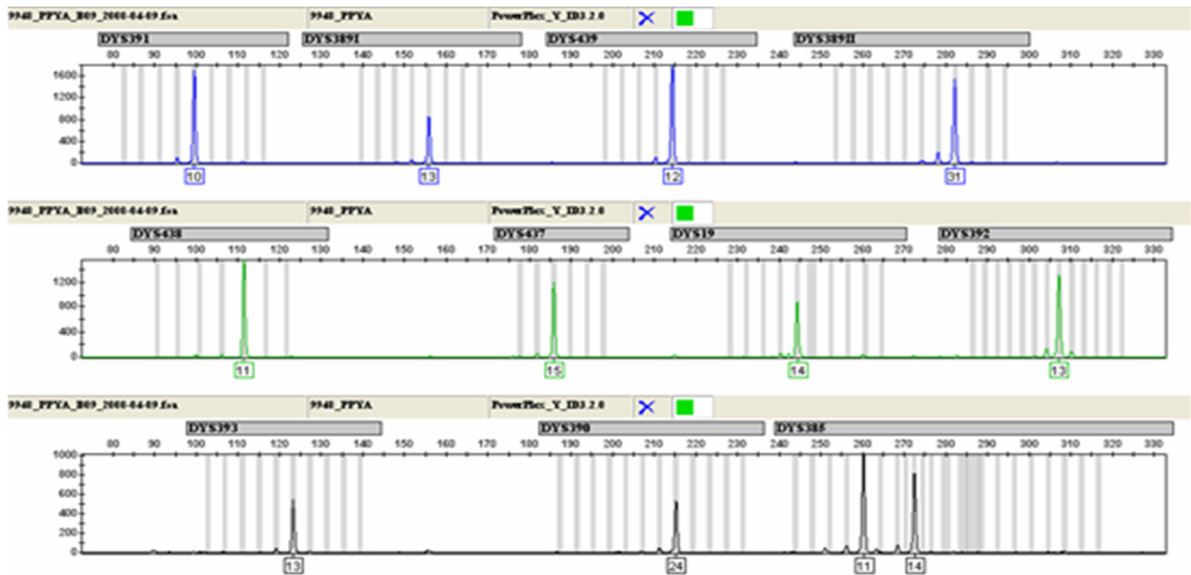
Phân tích các STR trên NST Y có ưu điểm hơn so với trên các NST thường bởi một số đặc điểm chính sau:

Thứ nhất, không cần tách chiết riêng các thành phần từ mẫu hỗn hợp của người nam và nữ trong các vụ án hình sự đặc biệt là trong các vụ xâm hại tình dục. Thông thường, để có thể phân tích STR trên NST thường từ mẫu tinh dịch có trong mẫu dịch thu ở âm đạo hoặc mẫu quần/áo của một nữ giới bị hiếp dâm cần một số phương pháp để tách riêng thành phần tinh trùng ra khỏi mẫu lẫn trong quá trình tách chiết ADN. Quá trình này đòi hỏi nhiều thời gian, hóa chất, công sức và không phải lúc nào cũng đạt hiệu quả nếu số lượng tinh trùng có trong mẫu hỗn hợp quá ít hoặc đã bị thoái hóa qua thời gian, ngoài ra còn phụ thuộc rất nhiều vào kinh nghiệm của cán bộ thực hiện các thí nghiệm. Tuy nhiên với việc phân tích các Y-STR sẽ không đòi hỏi quá trình này do đó giảm đáng kể các công đoạn trong tách chiết ADN.

Thứ hai, một dữ liệu Y-STR đầy đủ vẫn có thể được tạo ra từ lượng ADN nam cực nhỏ có trong mẫu vật hoặc ở dạng mẫu lẫn nơi mà lượng ADN của nữ giới chiếm ưu thế. Đã có rất nhiều nghiên cứu chứng minh khả năng phân tích thành công hồ sơ Y-STR trong mẫu có tỷ lệ ADN nữ/ADN nam giới chênh lệch nhau rất lớn. Prinz và cộng sự cho thấy trong các thí nghiệm mô hình mà thành phần mẫu của nam giới nhỏ trong hỗn hợp nam/nữ có thể được phát hiện lên đến tỷ lệ 1 :4000 với 0.4 ng ADN nam giới ban đầu [32]. Sibille báo cáo kết quả phân tích Y-STR là dương tính trong khoảng 1/3 các trường hợp xâm hại tình dục với kết quả định tính các mẫu thu dịch âm đạo ban đầu là “âm tính” đối với tinh dịch [33]. Hiệu suất, độ nhạy của phân tích Y-STR trong phản ứng multiplex PCR phụ thuộc vào việc thiết kế môi, việc tối ưu hóa các điều kiện phản ứng. Các bộ kit Y-STR thương mại phổ biến hiện nay là PowerPlex[®]Y23 (hãng Promega) và Yfiler[®] Plus (Thermo

Fisher scientific) đã phát triển rất tốt, chứng minh độ nhạy cao trong việc phát hiện lượng ADN nam ban đầu rất nhỏ (dưới 0.5 ng) [34].

Thứ ba, mỗi locus trên NST Y chỉ gồm 1 alen, do các locus Y-STR chủ yếu nằm trong phần không bắt cặp tương đồng trên NST Y nên ngoại trừ một số trường hợp đặc biệt, mỗi cặp mồi chỉ khuếch đại cho 1 alen trên 1 locus Y-STR. Do đó việc phân tích sau điện di mao quản cũng được thực hiện nhanh chóng, dễ dàng hơn khi so với STR trên NST thường. Ngay cả trong trường hợp phân tích hỗn hợp từ nhiều nam giới khác nhau, hồ sơ Y-STR riêng rẽ của từng người cũng có thể được xác định.



Hình 1. 6. Ví dụ về 1 hồ sơ Y-STR từ 1 cá thể với mỗi locus chỉ gồm 1 alen (ngoại trừ locus DYS385 gồm tổ hợp 2 locus DYS385a và b)

Thứ tư, khả năng khuếch đại nhiều Y-STR giúp tiết kiệm đáng kể thời gian, hóa chất. Cũng giống như các STR trên NST thường, phân tích đồng thời nhiều locus Y-STR có thể được thông qua phản ứng multiplex PCR và việc sử dụng nhiều loại màu (dye) huỳnh quang khác nhau. Sự phát triển của các bộ kit thương mại đã được kiểm định, nghiên cứu càng hữu ích cho ứng dụng Y-STR trong khoa học hình sự (Hình 1.6).

Với những ưu điểm nêu trên, việc phân tích hồ sơ Y-STR có thể áp dụng trong điều tra hiện trường tội phạm trong các trường hợp cụ thể sau đây:

Trường hợp thứ nhất, mẫu lần từ người nam và nữ trong đó nồng độ ADN mẫu nữ chiếm ưu thế hơn hẳn mẫu nam (thường là mẫu thu dịch trong âm đạo của một nữ giới bị hiếp dâm).

Trường hợp thứ hai, nghi phạm bị cáo buộc xâm hại tình dục trong đó kết quả thử định tính phát hiện dấu vết tinh dịch là âm tính hoặc không tìm thấy xác tinh trùng khi quan sát dưới kính hiển vi.

Trường hợp thứ ba, mẫu vật thu từ hiện trường các vụ xâm hại tình dục có kết quả thử định tính phát hiện vết tinh dịch là dương tính nhưng không phát hiện được ADN của thủ phạm do các alen tại các locus STR trên NST thường của thủ phạm nằm dưới ngưỡng phát hiện.

Trường hợp thứ tư, trong các vụ xâm hại tình dục có mẫu thu tại hiện trường là mẫu lẫn của nam và nữ hoặc của 2 người nam trở lên khi phân tích với hồ sơ STR trong nhân.

Trường hợp thứ năm, vết tinh dịch đã để lâu, phần lớn các tế bào tinh trùng được dự đoán là đã bị thoái hóa và phân tách các thành phần trong tách chiết ADN không thành công hoặc rui ro.

Trường hợp thứ sáu, mẫu hung khí trong các vụ cố ý gây thương tích, giết người là mẫu hỗn hợp từ nhiều người nam hoặc mẫu lẫn của nam giới và nữ giới.

1.6. Phương pháp phân tích các chỉ thị Y-STR

1.6.1. Phân tích dựa trên phương pháp điện di mao quản

Có một số phương pháp dùng trong phân tích STR như PCR với cặp mồi đặc hiệu và phát hiện trên ảnh điện di gel, phân tích đa hình chiều dài các phân đoạn cắt giới hạn (RFLP - Restriction Fragment Length Polymorphism)... Hiện nay với sự phát triển của sinh học phân tử quy trình phân tích STR tại các phòng xét nghiệm đã được chuẩn hóa dựa trên phương pháp PCR đa mồi có gắn huỳnh quang kết hợp với điện di mao quản cho phép phân tích đồng thời nhiều STR với độ nhạy, độ chính xác cao, tiết kiệm thời gian và công sức. Các bước cơ bản của một quy trình phân tích STR bao gồm: tách chiết ADN, định lượng ADN sau tách chiết, PCR khuếch đại các locus STR, điện di phát hiện sản phẩm sau PCR, phân tích kết quả [35].

Tách chiết ADN

Tùy thuộc vào nồng độ ADN có trong mẫu sinh phẩm để lựa chọn phương pháp tách chiết ADN phù hợp. Phương pháp được sử dụng phổ biến với mẫu có lượng ADN nhiều (mẫu tóc, niêm mạc miệng, mô tươi, máu...) là phương pháp Chelex® 100. Với những mẫu

có lượng ADN ít hoặc mẫu đã bị phân hủy việc sử dụng các bộ kit tách chiết ADN chuyên biệt được áp dụng nhằm tăng hiệu quả tách chiết.

Định lượng ADN

Sau tách chiết lượng ADN thu được cần phải định lượng chính xác và kiểm tra chất lượng ADN. Xác định đủ lượng ADN đầu vào cho phản ứng PCR sẽ cho kết quả tốt nhất. Điều này cực kỳ quan trọng đối với những mẫu ADN trong pháp y. Phương pháp định lượng phổ biến là điện di trên gel agarose, ngoài ra còn có thể định lượng bằng Real-time PCR.

Thực hiện phản ứng PCR khuếch đại các locus Y-STR

Phân tích STR nói chung và Y-STR nói riêng được thực hiện bằng kỹ thuật PCR (Polymerase Chain Reaction). Đây là kỹ thuật nhân bản ADN trong ống nghiệm dựa vào việc tăng giảm nhiệt theo chu kỳ, nhờ đó mà các mẫu nghiên cứu dù chứa ADN với số lượng rất ít vẫn có thể nhân bản thành nhiều bản sao đặc hiệu giống nhau về kích thước và trình tự [36].

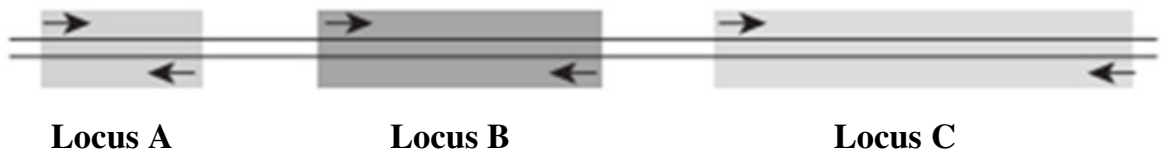
Trong hệ gen của người có nhiều STR được nghiên cứu và có khả năng ứng dụng trong phân biệt cá thể. Để đạt được khả năng phân biệt giữa các cá thể số lượng STR phân tích phải đủ lớn. Số lượng STR được lựa chọn phụ thuộc vào độ đa hình, số các alen trong 1 locus và tần suất phân bố các alen trong từng quần thể người nhất định. Thông thường số lượng STR được lựa chọn là từ 17 locus STR trở lên. Tuy nhiên nếu thực hiện việc khuếch đại riêng rẽ từng STR trong phản ứng PCR sẽ đòi hỏi tiêu tốn hóa chất, công sức rất nhiều. Sự ra đời của kỹ thuật PCR đa môi (multiplex PCR) đã giúp giải quyết vấn đề này. Phản ứng PCR đa môi (Multiplex PCR) được định nghĩa là sự khuếch đại đồng thời của nhiều vùng ADN trong một phản ứng PCR duy nhất. Trong phản ứng này nhiều vùng gen đích có thể được khuếch đại nhờ bổ sung đồng thời nhiều cặp môi. Như một phương pháp cải tiến của phản ứng PCR thông thường, phương pháp này giúp rút ngắn thời gian thực hiện mà lại không ảnh hưởng tới kết quả thí nghiệm. Kể từ lần đầu tiên được mô tả trong 1988, multiplex PCR đã được áp dụng trong nhiều lĩnh vực nghiên cứu ADN như phát hiện đột biến mất trong gen, định lượng mẫu, xác định tác nhân gây bệnh [37].

Ưu điểm lớn nhất của kỹ thuật multiplex PCR là có thể khuếch đại đồng thời nhiều locus STR trong một phản ứng. Điều này giúp tiết kiệm thời gian, chi phí, công sức rất nhiều so với việc khuếch đại từng locus STR riêng rẽ. Phản ứng multiplex PCR giúp tiết kiệm đáng kể chi phí hóa chất đặc biệt là ADN polymerase, môi trong điều kiện thiếu hụt.

Kết hợp với phương pháp đánh dấu môi huỳnh quang, điện đi mao quản và sử dụng phần mềm phân tích dữ liệu, hiện nay toàn bộ thời gian phân tích một bộ gồm 17 - 24 locus trong xét nghiệm huyết thống chỉ mất khoảng 2 giờ.

Thêm vào đó trong phản ứng multiplex PCR có thể đánh giá được hiệu quả khuôn ADN (độ tinh sạch, nồng độ) được đưa vào tốt hơn so với trong phản ứng PCR đơn lẻ.

Một ưu điểm khác của phản ứng multiplex PCR là có thể đánh giá được trường hợp âm tính giả (do phản ứng PCR thất bại) hoặc dương tính giả (do mẫu nhiễm) tốt hơn.



Hình 1. 7. Minh họa quá trình khuếch đại cùng lúc 3 locus sử dụng 3 cặp mồi khác nhau trong phản ứng multiplex PCR

Cũng giống như phản ứng PCR thông thường, việc tối ưu hóa điều kiện phản ứng trong multiplex PCR cũng là bước quan trọng và thậm chí còn phức tạp hơn nhiều lần so với trong phản ứng PCR thông thường [38]. Một số điều kiện chính trong phản ứng multiplex PCR cần được tối ưu:

Độ dài mồi: Phản ứng multiplex PCR liên quan đến việc thiết kế một số lượng lớn các mồi do đó yêu cầu mồi được thiết kế phải có độ dài thích hợp. Thông thường, mồi phải có độ dài ngắn khoảng 18-22 bazơ.

Nhiệt độ nóng chảy (T_m): các mồi có T_m tương tự nhau, tốt nhất là từ 55°C - 60°C nên được sử dụng. Đối với các trình tự có hàm lượng GC cao nên sử dụng mồi có T_m cao hơn (tốt nhất là 75°C - 80°C). Có thể chấp nhận sự thay đổi T_m trong khoảng từ 3° - 5°C đối với các mồi sử dụng trong cùng một phản ứng

Tính đặc hiệu: trong multiplex điều quan trọng là phải đảm bảo tính đặc hiệu của các đoạn mồi được thiết kế nhằm khuếch đại các trình tự mục tiêu khác nhau trong một phản ứng duy nhất

Tránh hình thành Primer Dimer: các mồi được thiết kế trong cùng 1 hỗn hợp phản ứng phải được kiểm tra xem có hình thành các dimer hay không. Sự hình thành dimer sẽ dẫn đến sự khuếch đại không đặc hiệu.

Ngoài ra để thu được lượng các sản phẩm khuếch đại sau PCR có sự cân bằng với năng suất tương tự nhau là một nhiệm vụ đầy thách thức. Ngày nay, với sự sẵn có của nhiều

bộ kit thương mại, các phòng thí nghiệm đơn lẻ thường hiếm khi thực hiện các thí nghiệm tối ưu hóa phản ứng multiplex PCR nữa. Thay vào đó, các nghiên cứu kiểm định nội bộ thường tập trung vào việc đánh giá hiệu suất của phản ứng multiplex PCR với các điều kiện khác nhau xung quanh các tham số tối ưu được cung cấp trong protocol của bộ kit. Ví dụ, hiệu suất PCR thể hiện độ cao của các đỉnh tín hiệu (peak) STR thu được theo protocol từ các kit thương mại có thể được đánh giá ở nhiệt độ gắn mồi tối ưu (ví dụ 59°C) và nhiệt độ cao hoặc thấp hơn nhiệt độ gắn mồi tối ưu 2-4°C (ví dụ 55°C, 57°C, 61°C và 63°C).

Điện di phát hiện sản phẩm PCR

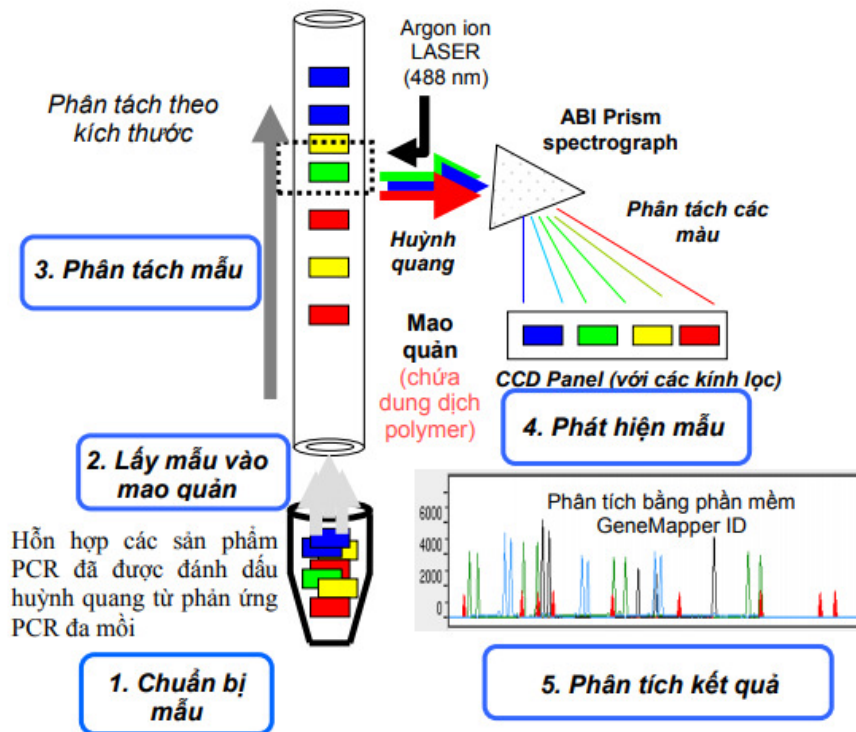
Bước tiếp theo sau phản ứng PCR là quá trình phân tách và phát hiện các sản phẩm khuếch đại. Phản ứng multiplex PCR khuếch đại các STR có thể tạo ra một hỗn hợp các phân tử ADN có kích thước khác nhau (trong khoảng từ 100 đến 400 bp) và cần được phân tách khỏi nhau một cách rõ ràng. Đặc biệt là cần phân tách các alen có khoảng cách gần nhau (ví dụ như 2 alen 9.3 và 10 của locus TH01 chỉ hơn kém nhau 1 bp). Việc này đòi hỏi phải có phương pháp phân tách ADN hiệu quả. Ngoài ra, điều quan trọng là phương pháp phân tách có thể tiến hành lặp lại và mang lại kết quả có thể so sánh kết quả phân tích giữa các phòng thí nghiệm với nhau.

Để phân tách các phân tử có kích thước khác nhau, cần phải có quá trình kéo các mảnh có kích thước khác nhau tách rời nhau qua quá trình gọi là điện di. Một số phương pháp điện di có thể được sử dụng bao gồm (1) Điện di trên gel polyacrylamide (PAGE) kết hợp với phát hiện bằng phương pháp nhuộm bạc hoặc máy quét gel huỳnh quang (nếu mồi được gắn huỳnh quang) và (2) Điện di mao quản (CE) với cảm ứng huỳnh quang laser. Hiện nay điện di mao quản với nhiều ưu điểm như: cho độ phân giải cực cao, tốc độ và độ nhạy cao đối với việc phân tích các mẫu rất nhỏ, tự động hóa, không cần sử dụng gel, mẫu có thể được bổ sung thêm... đang là phương pháp phổ biến trong phân tích STR [39].

Để có thể được nhận biết bởi cảm biến laser và phân tách được các locus có kích thước gần nhau trong hệ thống máy điện di mao quản đòi hỏi phân tử ADN phải được đánh dấu huỳnh quang (fluorescence dye). Đánh dấu huỳnh quang vào các đoạn ADN có thể được thực hiện bằng nhiều cách khác nhau. Tuy nhiên, phương pháp phổ biến nhất được sử dụng đánh dấu các alen STR là sử dụng các dye huỳnh quang để đánh dấu ở đầu 5' của một mồi PCR (mồi xuôi hoặc mồi ngược) [40]. Mỗi loại màu này sẽ phát huỳnh quang ở các bước sóng khác nhau. Quá trình đánh dấu được thực hiện ngay từ khi tổng hợp mồi. Sau

đó, sử dụng các môi này để PCR nhân bội các locus STR tương ứng sẽ tạo các sản phẩm PCR được đánh dấu. Các sản phẩm PCR này được điện di phân tách trên hệ thống điện di mao quản. Hai sợi ADN được phân tách trong quá trình điện di và chỉ sợi ADN được đánh dấu sẽ được phát hiện bằng thiết bị phát hiện nhờ cảm biến laser.

Việc gắn môi huỳnh quang có nhiều ưu điểm như giúp kết quả phiên giải được đọc dễ dàng vì chỉ có một trong hai mạch của phân tử ADN được phát hiện. Ngoài ra, trong phản ứng multiplex PCR khuếch đại đồng thời nhiều locus STR việc đánh dấu môi bằng huỳnh quang giúp dễ dàng phân biệt các sản phẩm khuếch đại có kích thước tương tự [41].



Hình 1. 8. Các bước hoạt động chính của một hệ thống điện di mao quản [35]

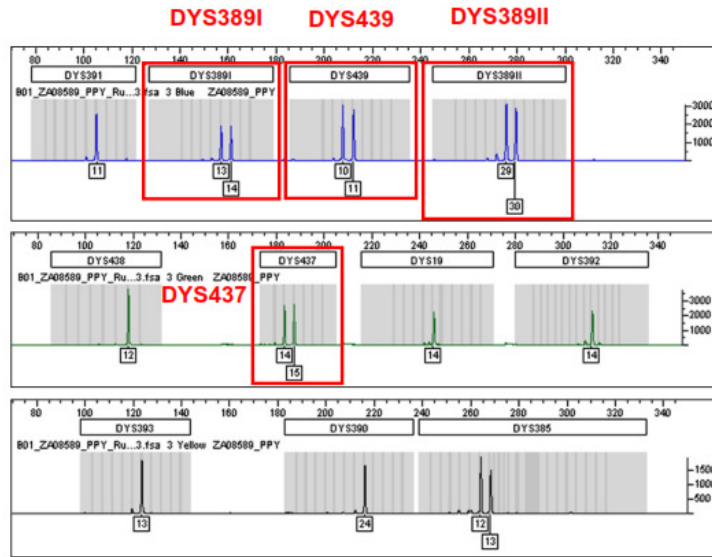
Phân tích xử lý số liệu

Phần mềm thu nhận dữ liệu sẽ giúp tính toán, xử lý dữ liệu thô và hiển thị phổ dữ liệu dưới dạng các nhiều đỉnh tín hiệu (peak). Bước cuối cùng trong việc tạo ra hồ sơ STR là việc gắn tên các locus, tên alen của từng locus tương ứng với các đỉnh tín hiệu. Khi gắn tên alen, các đỉnh tín hiệu chưa biết sẽ được so sánh với thang alen và kích thước của tín hiệu đó phải rơi vào một khe với độ rộng trong khoảng ± 0.5 bp so với kích thước tương ứng ở thang alen chuẩn. Nếu kích thước của đỉnh tín hiệu không nằm khoảng này sẽ bị phân

loại thành nhóm lệch thang (OL - off ladder) và đôi khi phải tiến hành phân tích lại [42]. Sự so sánh các đỉnh tín hiệu chưa biết với thang alen có thể được thực hiện thủ công hoặc tự động bằng các phần mềm chuyên dụng. Kết quả cuối cùng sẽ là một hồ sơ (profile) trong đó các alen của từng locus STR được biểu thị dưới dạng các chữ số hay chính là số lần lặp lại trong lõi STR.

Với STR trong nhân tế bào tùy thuộc vào dạng đồng hợp tử hay dị hợp tử tín hiệu phân tích sau điện di mao quản sẽ hiển thị ở dạng 1 hoặc 2 peak. Tuy nhiên, với Y-STR do mỗi locus thường chỉ có 1 bản sao trên NST Y nên khi phân tích kết quả thông thường tại 1 locus STR trên NST Y sẽ chỉ xuất hiện 1 peak tương ứng. Việc xuất hiện 2 hoặc nhiều peak hơn tại mỗi locus thường là do 2 nguyên nhân chính sau:

Mẫu hỗn hợp (thường là mẫu thu từ hiện trường vụ án) có thể được nhận biết bằng hồ sơ ADN xuất hiện nhiều hơn 1 alen tại mỗi locus Y-STR. Thông thường trong trường hợp này sẽ có 2 hoặc 3 alen cùng xuất hiện tại mỗi locus (phụ thuộc vào số lượng nguồn mẫu khác nhau có trong mẫu hỗn hợp) và sự cân bằng giữa các đỉnh tín hiệu (peak) cũng không còn. Ngoài ra trên thực tế có nhiều vùng trên NST Y có thể bị mất hoặc được nhân đôi, nhân ba ở một số tế bào hoặc cá thể qua quá trình phân bào. Do đó, thực tế này có thể làm phức tạp kết quả phiên giải mẫu. Một số các mẫu đã được quan sát với nhiều Y-STR được nhân đôi như minh họa trong Hình 1.9. Bản sao thường sở hữu các alen chỉ hơn kém nhau 1 đơn vị lặp lại và sở hữu chiều cao peak tương đương nhau. Nghiên cứu với cặp mẫu cha - con đã được chứng minh sự xuất hiện của đột biến nhân đôi, nhân ba và xóa 1 đơn vị lặp lại trong các locus Y-STR [43].



Hình 1. 9. Kết quả phân tích Y-STR từ 1 mẫu nam giới với các locus DYS389I, DYS439, DYS437, DYS389II bị đột biến lặp đoạn với 2 alen trong 1 locus Y-STR chỉ hơn kém nhau 1 đơn vị lặp [35]

Trong giám định hình sự, cần phải phân biệt rõ 2 hiện tượng khi xuất hiện nhiều hơn 1 alen trong mỗi locus Y-STR: (1) do đột biến lặp 1 đoạn (vùng) trên NST Y với các locus đột biến xuất hiện liền kề nhau trên NST Y; (2) là do nhiễm từ nhiều nguồn khác nhau, thường xuất hiện nhiều peak tại mỗi locus ở các locus Y-STR không liền kề nhau.

Một nguyên nhân khác có thể dẫn đến xuất hiện không đầy đủ các locus Y-STR trong hồ sơ phân tích là do chất lượng của mẫu ADN. Với những mẫu thu từ hiện trường vụ án đã tiếp xúc lâu với môi trường trong điều kiện: vi sinh vật, độ ẩm, nhiệt độ cao sẽ khiến ADN bị phân hủy, đứt gãy, không còn nguyên vẹn. Điều này dẫn đến việc khi phân tích Y-STR sẽ chỉ phát hiện được các alen có kích thước nhỏ, tỷ lệ thành công giảm theo kích thước của các alen. Các alen có kích thước lớn có thể bị đột biến trong hồ sơ ADN, dễ gây nhầm lẫn khi phân tích kết quả.

1.6.2. Phân tích sử dụng các bộ kit Y-STR thương mại hóa

Độ hiếm gặp của một profile các STR trên NST thường sẽ tỷ thuận với số lượng STR được phân tích. Hay nói cách khác, càng nhiều STR được phân tích sẽ tạo càng nhiều hồ sơ dữ liệu STR khác nhau giữa các mẫu trong quần thể. Điều này cũng đúng với các Y-STR. Tuy nhiên trong số hơn 400 locus Y-STR đã được biết đến việc lựa chọn Y-STR cần thiết có thể ứng dụng được trong giám định pháp y phải dựa trên một số yêu cầu sau:

Thứ nhất, các locus Y-STR phải có tính đa hình, có nhiều alen khác nhau. Điều này giúp các chuyên gia phân tích tạo ra số lượng Y haplotype nhiều nhất nhằm phân biệt các cá thể một cách hiệu quả nhất.

Thứ hai, các Y-STR phải có kích thước ngắn từ 100 - 300 bp. Do các đoạn ADN ngắn có tính bền vững cao hơn, ít bị đứt gãy hơn khi có tác động của điều kiện tự nhiên và quá trình nhân gen PCR có thể được thực hiện dễ dàng và hiệu quả hơn. Đối với những đoạn ADN có tính đa hình cao nhưng kích thước lớn, trong thực tế chỉ có thể thực hiện phản ứng PCR với những mẫu sinh phẩm còn tươi, mới hoặc được bảo quản trong những điều kiện tốt. Ngoài ra, điện di mao quản có độ phân giải cao thì khó phân tách chính xác đối với các băng ADN có kích thước lớn. Vì vậy, không nên thiết kế các STR có kích thước lớn hơn 400 bp.

Thứ 3, các locus phải có trình tự các vùng phân bố hai bên STR đặc hiệu (đơn bản, không có SNP) để việc thiết kế mồi có thể dễ dàng được thực hiện.

Dựa trên các tiêu chí đó, ban đầu một bộ gồm 9 locus Y-STR tối thiểu ban đầu được chọn khi cơ sở dữ liệu dân số về Y-STR bắt đầu được tạo ra vào cuối những năm 1990 ở châu Âu. Những locus đó bao gồm marker chỉ có một bản sao: DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392 và DYS393 và một marker đa bản sao DYS385 a/b (thường tạo ra hai sản phẩm khuếch đại với một cặp mồi). Tập hợp này được gọi là bộ lõi hay bộ haplotype Y tối thiểu (minimal haplotype loci - MHL) [47]. Vào năm 2003, khi Mỹ tiến hành lựa chọn bộ lõi Y-STR, Hiệp hội khoa học làm việc trong lĩnh vực phân tích ADN (SWGAM) đã khuyến cáo bổ sung thêm 2 locus mới là DYS438 và DYS439 so với bộ lõi Y-STR trước đây, thành tổng số 11 locus.

Ngay sau đó, một loạt các bộ kit thương mại Y-STR được phát triển dựa trên bộ lõi này. PowerPlex®Y, bộ Y-STR thương mại đầu tiên từ tập đoàn Promega năm 2003 đã thêm DYS437 vào bộ lõi của SWGDAM. Bộ kit AmpFlSTR® Yfiler (viết tắt: Yfiler) của hãng Applied Biosystem năm 2004 bổ sung vào 5 locus mới so với bộ PowerPlex® Y: DYS448, DYS456, DYS458, DYS635 và Y-GATA-H4, thành tổng số 17 locus Y-STR [48].

Vào năm 2012, PowerPlex® Y23 system (hãng Promega) được phát hành bao gồm 17 locus của bộ Yfiler và thêm 6 locus mới: DYS481, DYS533, DYS549, DYS570, DYS576, và DYS643 [49]. Bộ kit PowerPlex® Y23 System (viết tắt: PPY23) cho phép

khuếch đại và phát hiện đồng thời 23 locus với hệ thống 5 dye huỳnh quang. Bộ kit được cấu thành với phần lớn các locus có độ đa hình cao, khả năng phân biệt lớn với kích thước khuếch đại nhỏ hơn. Do kích thước khuếch đại nhỏ hơn, đặc biệt có thể khuếch đại được lượng ADN trong mẫu vi vết, tiềm năng phân biệt các mẫu khi sử dụng bộ PP Y23 có thể được đạt được rất lớn. Bộ kit cũng cung cấp tất cả các thành phần cần thiết cho việc khuếch đại các vùng Y-STR trong ADN người như ADN polymerase, môi, đệm....

Bộ kit cũng được thiết kế để tương thích với nhiều hệ thống máy điện di mao quản như 3130, 3130xl, 3500 và 3500xl và đã được chứng minh hiệu quả cao trong nghiên cứu cấu trúc di truyền quần thể và trong lĩnh vực hình sự [49,50].

Bảng 1.1. Danh sách các locus Y-STR có mặt trong các bộ kit thương mại phổ biến hiện nay [20]

Y-STR marker	Các bộ kit thương mại Y-STR chủ yếu hiện nay				
	Bộ tối thiểu	PowerPlex® Y ^a	AmpF/STR® Yfiler® ^b	PowerPlex® Y23 ^a	Yfiler™ Plus ^b
DYS19	✓	✓	✓	✓	✓
DYS385a/b	✓	✓	✓	✓	✓
DYS389I	✓	✓	✓	✓	✓
DYS389II	✓	✓	✓	✓	✓
DYS390	✓	✓	✓	✓	✓
DYS391	✓	✓	✓	✓	✓
DYS392	✓	✓	✓	✓	✓
DYS393	✓	✓	✓	✓	✓
DYS437		✓	✓	✓	✓
DYS438		✓	✓	✓	✓
DYS439		✓	✓	✓	✓
DYS448			✓	✓	✓
DYS456			✓	✓	✓
DYS458			✓	✓	✓
DYS635			✓	✓	✓
Y-GATA-H4			✓	✓	✓
DYS481				✓	✓

Y-STR marker	Các bộ kit thương mại Y-STR chủ yếu hiện nay				
	Bộ tối thiểu	PowerPlex® Y ^a	AmpF/STR® Yfiler® ^b	PowerPlex® Y23 ^a	Yfiler™ Plus ^b
DYS533				✓	✓
DYS549				✓	
DYS570				✓	✓
DYS576				✓	✓
DYS643				✓	
DYS449					✓
DYS460					✓
DYS518					✓
DYS627					✓
DYF387S1a/b					✓

Năm 2014, hãng Applied Biosystem (Thermo Fisher Scientific) cho ra đời bộ kit Y-STR mới có tên gọi là Yfiler™ Plus PCRAmplification Kit (gọi tắt là YPlus) với 6 dye huỳnh quang, bổ sung thêm 10 locus Y-STR mới so với bộ Yfiler trước đây, trong đó bao gồm 7 Y-STR có tốc độ đột biến nhanh bao gồm: DYF387S1a/b, DYS449, DYS518, DYS570, DYS576, DYS627 nâng tổng số locus Y-STR được phân tích lên thành 27 locus. Tính đến thời điểm này, đây cũng là bộ kit thương mại có số lượng các Y-STR nhiều nhất [51]. Hiện nay PPY23 và YPlus là 2 bộ kit thương mại về Y-STR đang được sử dụng rộng rãi trong cả lĩnh vực nghiên cứu và ứng dụng trên thế giới.

1.6.3. Phân tích sử dụng chiến lược mini STR

Trong thực tế việc nhận dạng cá thể không phải lúc nào cũng được thực hiện với các mẫu có chất lượng tốt (mẫu thu trực tiếp từ cơ thể sống, các mẫu còn giữ nguyên vẹn về cấu trúc và đủ nồng độ ADN phân tích). Rất nhiều trường hợp phân tích đã bị biến tính hoặc ở tình trạng phân hủy mạnh như mẫu dấu vết hình sự, mẫu hài cốt lâu năm. Sau khi cơ thể chết, máu ngừng cung cấp oxy và dinh dưỡng cho các mô, dẫn đến giảm sản xuất adenosine triphosphate (ATP) của tế bào. Mức ATP nội bào xác định quá trình nào trong số hai quá trình chết tế bào mà một tế bào cụ thể sẽ trải qua: apoptosis hoặc hoại tử. Màng tế bào bị phá vỡ và sự phân hủy bào quan theo sau và các enzym lysosome được giải phóng

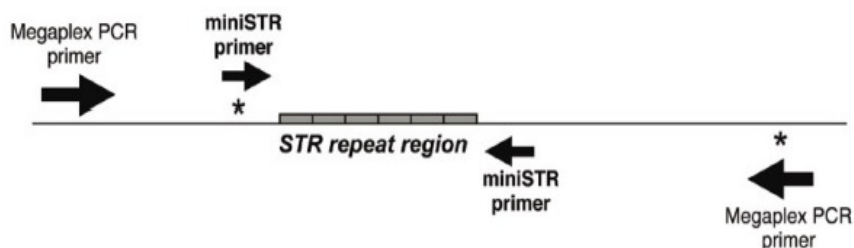
tạo ra một kiểu phân hủy ADN ngẫu nhiên, ADN bị bẻ gãy thành các đoạn dài nhỏ bởi các enzym nội sinh. Điều này làm cho việc khuếch đại ADN trở nên khó khăn. Một khi chiều dài đoạn ADN trung bình giảm xuống dưới 300bp, thông tin di truyền sẽ bị mất đáng kể do thiếu ADN khuôn mẫu thích hợp để khuếch đại

Các quá trình phân hủy tích lũy theo thời gian, trong khi các điều kiện môi trường như nhiệt độ, độ ẩm và độ pH làm thay đổi tốc độ và mức độ tích cực của quá trình phân hủy. Lửa, độ ẩm cao (ví dụ mẫu ngâm trong nước, chôn trong đất) và sự xâm nhập của vi sinh vật là tất cả các yếu tố ảnh hưởng đến sự phân hủy của ADN còn sót lại. Sự phát triển của vi sinh vật đã được chứng minh là một nguyên nhân đáng kể gây ra tổn thương ADN. Các vi sinh vật tiết ra các enzym tiêu hóa làm tổn thương ADN thông qua các loại oxy phản ứng và đứt gãy sợi kép.

Ngoài ra, sự hiện diện của chất ức chế trong mẫu cũng có thể là một thách thức đối với phản ứng PCR. Các chất ức chế PCR là nguyên nhân phổ biến nhất của sự thất bại trong quá trình khuếch đại mặc dù có đủ lượng ADN đầu vào [42]. Phản ứng PCR có chứa chất ức chế dẫn đến việc mất alen từ các locus STR có kích thước lớn hoặc thất bại hoàn toàn trong việc khuếch đại tất cả các locus. Các kết quả này có thể bị nhầm lẫn với các mẫu bị phân hủy nghiêm trọng (mẫu có nồng độ ADN thấp).

Để khắc phục tình trạng đó một số nghiên cứu đã chỉ ra rằng về mặt lý thuyết cấu trúc ADN có thể được phục hồi hiệu quả từ các mẫu ADN đã bị phân hủy bằng cách khuếch đại từng đoạn ADN có kích thước nhỏ hay còn gọi là chiến lược mini STR [44].

Bằng cách di chuyển vị trí bắt cặp mỗi lại gần nhất có thể đến vùng lõi lặp của STR có thể giảm kích thước của sản phẩm PCR đồng thời vẫn giữ được thông tin vùng lõi cần thiết như các STR bình thường. Chiến lược này được gọi là khuếch đại các mini STR. Việc phân tích một số locus mini-STR này sẽ góp phần bổ sung thông tin về hồ sơ ADN (kiểu gen) cá thể trong những trường hợp phân tích STR thông thường không thu được một hồ sơ kiểu gen đầy đủ. Tỷ lệ thành công của các hệ mini-STR so với các hệ STR “truyền thống” trên các mẫu bị phân hủy mạnh đã được chứng minh: Mini-STR có thể cung cấp một hồ sơ ADN với thông tin di truyền gấp 3 lần và đầy đủ hơn hệ STR tiêu chuẩn. Các nghiên cứu trên mẫu bị phân hủy mạnh cho thấy rằng các locus mini-STR có kích thước dưới 120 bp đều cho kết quả tốt, các trình tự trên 150 bp thường khó thu được kết quả hơn.



Hình 1. 10. Minh họa về vị trí khuếch đại cùng 1 locus STR giữa cặp mỗi PCR kích thước lớn và cặp mỗi mini STR [35]

Tương tự như vậy các locus Y-STR có kích thước nhỏ có thể được phân tích trong một số ứng dụng như tìm người thân mất tích theo dòng cha trong chiến tranh, thiên tai, hỏa hoạn..., phân tích ADN từ mẫu vi vết thu tại hiện trường các vụ án với số lượng bản sao ADN thấp và ADN đã bị phân cắt.... Phương pháp này cũng khắc phục nhược điểm của việc phân tích Y-STR so với việc sử dụng các bộ kit thương mại có sẵn. Năm 2007, Myung Jin Park và cộng sự đã sử dụng phương pháp trên để giảm kích thước của 17 locus Y-STR (trong số 27 locus Y-STR của bộ kit AmpFI STR[®] Yfiler[™] - Applied Biosystem) và đã chứng minh được ưu thế vượt trội trong việc phân tích các mẫu ADN có chất lượng kém từ mẫu hài cốt hơn 50 năm [45]. Năm 2017, Y.Gu và cộng sự cũng thiết kế lại các cặp mỗi nhằm khuếch đại đồng thời 3 locus DYS385a/b, DYS389II và DYS392 có kích thước nhỏ hơn khoảng 140 bp so với trong bộ kit PowerPlex[®] Y (Promega) nhằm áp dụng trong việc phân tích ADN đã bị phân hủy trong các vụ án hình sự [46].

1.7. Tầm quan trọng của việc tính toán tần suất alen các chỉ thị STR

Tần suất alen chính là tỉ lệ xuất hiện của một alen nào đó trong quần thể ở một locus gen trên NST. Thực tế, trong xét nghiệm ADN có hàng chục locus gen được nghiên cứu và sử dụng ở dạng bộ Kit thương phẩm, các locus gen này nằm trên các NST khác nhau từ cặp số 1 đến cặp số 23.

Trong giám định huyết thống cha/mẹ - con, các nhà nghiên cứu vẫn phải sử dụng tần suất alen trên cơ sở dựa vào quy luật di truyền về khả năng cho nhận các alen của bố mẹ với con cái. Một người đàn ông bất kỳ thỏa mãn điều kiện cho người con một nửa số alen của mình thì người đó có khả năng là bố đẻ. Tuy nhiên, chúng ta cần biết rằng trong một số trường hợp phải tính cả xác suất để xem có người đàn ông nào khác cũng có khả năng như vậy không? Nếu phân tích ít locus gen (từ 10 -12) thậm chí đến 16 locus gen thì vẫn có

trường hợp một vài người đàn ông khác nhau cùng cho con những alen tương tự, nhưng người bố đẻ chỉ có một. Điều này vẫn có thể xảy ra là do cha đẻ chỉ truyền cho con một alen ở mỗi locus gen chứ không phải toàn bộ kiểu gen của mình.

Bởi vậy để tính toán được xác suất trùng hợp giữa hai hồ sơ STR cũng như độ tin cậy của việc có quan hệ huyết thống thì việc xác định tần suất alen của các locus STR là rất quan trọng. Theo khuyến cáo của Hiệp hội pháp y quốc tế (ISFN) mỗi một quần thể, dân tộc cần tự xây dựng bảng tần suất phân bố các alen của riêng mình và phải được kiểm định chính xác nhằm phục vụ cho việc phân tích quan hệ huyết thống, điều tra hình sự, tìm người mất tích trong thiên tai, thảm hoạ... [52].

Thông thường, các chuyên gia phân tích ADN sẽ sử dụng những locus gen khác nhau trên NST khác nhau để đảm bảo tính phân ly độc lập cũng như tính đa hình cao. Trong mỗi locus gen lại có nhiều alen khác nhau vì vậy sẽ có rất nhiều đặc điểm khác nhau được phân tích. Để xác định được tần suất alen, các nhà khoa học cần nghiên cứu, khảo sát về nhiều STR thuộc số quần thể, dân tộc người nhất định. Điều này sẽ giúp xác định tính đa hình của locus STR đó, sự khác biệt giữa các tộc người, tỷ lệ đột biến cao hay thấp để từ đó quyết định có nên sử dụng STR đó hay không [53].

Bên cạnh đó, việc thu thập mẫu và lựa chọn kích thước mẫu là rất quan trọng. Các nhà nghiên cứu cần phải thu thập được mẫu đã biết chính xác về nguồn gốc: tộc người, không có quan hệ huyết thống, ở các vùng địa lý khác nhau trên các vùng lãnh thổ. Để đảm bảo độ tin cậy, số lượng mẫu phân tích cần lớn. Khi đó, kết quả phân tích thống kê tần số alen và kiểu gen phải phù hợp giữa mẫu nghiên cứu và quần thể lý thuyết có thể sử dụng tiêu chuẩn khi bình phương (χ^2). Chỉ khi nào đạt được tiêu chuẩn này thì tần suất alen này mới có giá trị và đủ tin cậy để sử dụng xác định trong truy nguyên cá thể và xác định quan hệ huyết thống.

Số lượng locus gen: Bên cạnh tần suất alen, độ tin cậy của xét nghiệm còn phụ thuộc vào cả số lượng locus tham gia phân tích. Số lượng locus STR phân tích càng nhiều, tần suất của 1 số alen nào đó càng thấp thì mẫu phân tích đó càng có độ chính xác cao, hạn chế tối đa khả năng trùng lặp ngẫu nhiên mẫu trong quần thể.

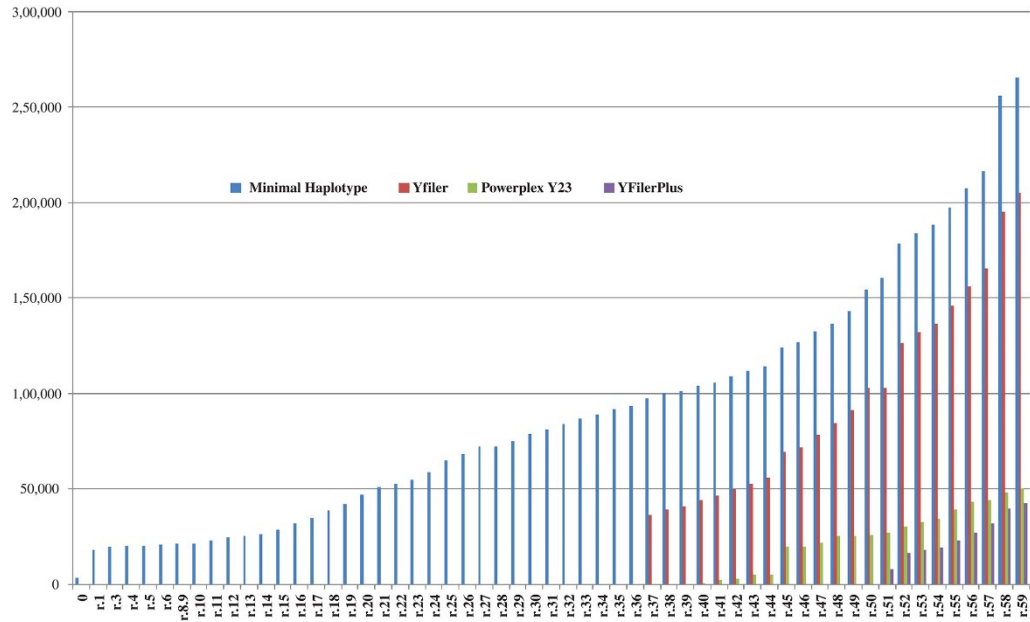
1.8. Tình hình nghiên cứu các chỉ thị Y-STR

1.8.1. Tình hình nghiên cứu trên thế giới

Những năm gần đây, hướng nghiên cứu về Y-STR đặc biệt là tần suất các locus alen Y-STR xuất hiện ở những nam giới cư trú tại các quần thể, quốc gia nhất định đã được các phòng thí nghiệm ở nhiều nước trên thế giới quan tâm và nghiên cứu. Với việc sử dụng các bộ kit thương mại sẵn có đặc biệt là 2 bộ kit PPY23 và YPlus việc tính toán tần suất alen, độ đa hình của các Y-STR, tần suất phân bố các haplotype đang được nghiên cứu rộng rãi tại nhiều quần thể người trên thế giới và liên tục cập nhật trên hệ thống cơ sở dữ liệu chung.

Việc xây dựng bảng tần suất alen các locus Y-STR của một bộ kit mới trở thành điều kiện cần thiết của mỗi cơ sở nghiên cứu, phòng thí nghiệm khi tiến hành áp dụng cho việc xét nghiệm huyết thống, giám định ADN liên quan đến hình sự theo khuyến cáo của Hiệp hội các nhà khoa học làm việc trong lĩnh vực phân tích ADN (SWGDM) [54].

Hiện nay cơ sở dữ liệu Y-STR về di truyền và quần thể lớn nhất đang được sử dụng rộng rãi nhất là Cơ sở dữ liệu tham khảo Haplotype Y-STR (YHRD - Y Chromosome Haplotype Reference Database, trang web <https://yhrd.org/>), được Lutz Roewer và các đồng nghiệp tại Đại học Humbolt ở Berlin, Đức tạo ra và đã có sẵn trên Internet từ năm 2000 [55]. Tính đến năm 2018, YHRD chứa kết quả từ hơn 265,324 hồ sơ Y-STR từ 1251 quần thể, nhóm nghiên cứu thuộc hơn 100 nước trên thế giới với số lượng Y-STR là 29 locus [56] (hình 1.11). YHRD cũng bao gồm các công cụ online cho phép kiểm định số liệu Y-STR, công cụ AMOVA để tính toán khoảng cách di truyền giữa các quần thể dựa trên số liệu Y-STR được cập nhật. Số lượng hồ sơ Y-STR không ngừng gia tăng trên YHRD trong những năm gần đây cho thấy hướng nghiên cứu Y-STR đang ngày thu hút sự quan tâm của các nhà khoa học.



Hình 1.11. Sự gia tăng số lượng hồ sơ Y-STR trên YHRD.org từ năm 1999 đến 2018 (số năm biểu thị từ 1 đến 59 [56])

Trên thế giới, có thể thấy các bộ kit Y-STR thương mại hóa đang ngày càng được phát triển với xu hướng tăng số lượng các Y-STR, với mục tiêu làm tăng hiệu quả của Y-STR trong việc phân biệt giữa cá thể trong quần thể, khả năng ứng dụng Y-STR trong nghiên cứu khoa học hình sự, phát hiện ADN từ mẫu vi vết. Tuy nhiên, theo các nghiên cứu trong các quần thể người khác nhau trên thế giới, các bộ kit này chưa phân biệt được toàn bộ các cá thể nam trong một quần thể người nhất định, đặc biệt là ở các quần thể dân tộc thiểu số. Ví dụ nghiên cứu trên quần thể gồm 268 nam giới người Ghana sử dụng bộ PPY23 chỉ cho khả năng phân biệt là 0.9627 [57]. Khảo sát trên quần thể người Mãn Chu tại tỉnh Liêu Ninh, Đông Bắc Trung Quốc cho khả năng phân biệt của bộ kit PPY23 là 0.9847 [58]. Với nghiên cứu trên 1225 nam giới người Hán, tại tỉnh Hà Nam, miền Trung Trung Quốc sử dụng bộ YPlus khả năng phân biệt cũng chỉ đạt 98.94 % [59].

Giữa những người không có quan hệ huyết thống theo dòng cha vẫn có thể có cùng một kiểu haplotype Y-STR giống nhau. Điều này dẫn đến nhu cầu phân tích thêm các locus Y-STR khác để tăng khả năng phân biệt giữa các cá thể và cũng để phù hợp với từng quần thể dân cư với cấu trúc di truyền khác nhau. Do đó, bên cạnh các bộ kit thương mại sẵn có nhiều phòng thí nghiệm cũng phát triển các bộ mỗi riêng nhằm khuếch đại các locus Y-STR khác nhau phục vụ cho nhiều mục đích: nghiên cứu di truyền quần thể, tăng khả năng

ứng dụng Y-STR trong khoa học hình sự... Điển hình như năm 2007, nghiên cứu của Lim và cộng sự đã phát triển phản ứng multiplex PCR khuếch đại 52 locus Y-STR nhằm đánh giá độ đa hình trong quần thể người Trung Quốc [60]. Amato và cộng sự cũng phát triển phản ứng multiplex PCR khuếch đại 10 locus Y-STR có độ đa hình cao hơn so với các locus đã có trong các bộ kit thương mại trước đó nhằm phù hợp hơn với các nghiên cứu trong quần thể người Nam Phi [61].

Một số tác giả cũng đề xuất thêm hướng ứng dụng mới của Y-STR trong việc lập phả hệ, xác định nguồn gốc dòng họ cha của nam giới. Việc sử dụng nhiễm sắc thể Y trong lập phả hệ di truyền đã dẫn đến các ứng dụng hữu ích trong một số ngành nghiên cứu cụ thể như: nghiên cứu lịch sử gia đình, nhân khẩu học, nhân chủng học, khoa học pháp y, di truyền dân số và tiến hóa nhiễm sắc thể giới tính [62].

Trên thế giới nhu cầu áp dụng Y-STR trong các lĩnh vực đang ngày càng gia tăng. Tuy nhiên cũng còn một số vấn đề chính cần giải quyết. Ví dụ như khi tăng số lượng locus Y-STR thì cũng sẽ làm tăng số kiểu haplotype có thể tìm thấy trong quần thể và đòi hỏi phải phân tích, thống kê trên một số lượng mẫu cực lớn để đạt đủ độ tin cậy. Hơn nữa, với hai ứng dụng chính của các bộ kit Y-STR hiện nay là: phân tích nguồn gốc tiến hóa quần thể theo dòng cha thích hợp với các Y-STR có tốc độ đột biến cao và phân tích nhận dạng cá thể trong gia đình phù hợp với các Y-STR có tốc độ đột biến thấp đòi hỏi các việc thiết kế bộ kit Y-STR phải cân nhắc và lựa chọn các marker hết sức cẩn thận.

Hơn nữa việc phân tích quá nhiều locus Y-STR (khoảng hơn 27 locus) sẽ tiến tới giới hạn phân tách của phương pháp điện di mao quản đang được sử dụng phổ biến hiện nay [20].

Các vấn đề nêu trên có thể được khắc phục bằng cách sử dụng nhiều loại màu (dye) đánh dấu huỳnh quang hơn, đòi hỏi sự phát triển của nhiều hóa chất, thiết bị mới. Ngoài ra, các bộ kit Y-STR khác nhau để có thể được phát triển nhằm phục vụ cho 2 hướng ứng dụng riêng biệt của Y-STR. Ngoài ra, công nghệ giải trình gen thế hệ mới (NGS) hiện nay có thể đóng vai trò thay thế phương pháp điện di mao quản với khả năng phân tích số lượng lớn Y-STR, như đã được áp dụng cho STR trên NST thường, Y-STR và SNP với bộ kit ForenSeq (Illumina). Tuy nhiên, các nền tảng NGS hiện tại phù hợp với khả năng phân tích ADN có kích thước giới hạn trong phạm vi nhất định. Các lần đọc trình tự ngắn thu được

phù hợp với nhiều Y-STR có tốc độ đột biến trung bình thấp nhưng chúng không đủ cho các Y-STR có tốc độ đột biến nhanh với trình tự dài của đoạn lặp lại. Do đó, các công nghệ NGS cho phép giải trình tự các đoạn ADN có kích thước dài, hay phân tích mẫu chất lượng thấp cũng cần được phát triển trong tương lai [63].

Một vấn đề khác đó là việc lựa chọn các locus Y-STR nhằm mục đích phân biệt giữa các cá thể trong một quần thể nhất định cũng cần được cân nhắc chính xác bởi lý do NST Y là NST có tốc độ tiến hóa cao, giữa các quần thể khác nhau có sự khác nhau trong tần suất phân bố các locus Y-STR ngược lại giữa các quần thể có cùng nguồn gốc có thể có sự tương đồng lớn trong phân bố tần suất các locus Y-STR.

1.8.2. Tình hình nghiên cứu tại Việt Nam

Tại Việt Nam các locus Y-STR đang được sử dụng tại một số ở Trung tâm xét nghiệm ADN, các Viện nghiên cứu nhằm xác định quan hệ huyết thống theo dòng cha, điều tra các vụ án tội phạm khi dấu vết để lại tại hiện trường rất ít. Tuy nhiên các nghiên cứu tần suất locus Y-STR trên nam giới của một số dân tộc ở Việt Nam còn hạn chế. Một số nghiên cứu công bố mới chỉ khảo sát được tần suất từ 13 đến 17 locus Y-STR.

Cụ thể, năm 2003 Nguyễn Quang Tuyên và các cộng sự Nhật Bản đã tính toán tần suất các alen của 13 STR trên NST Y từ 119 người đàn ông Việt không có quan hệ họ hàng [64].

Năm 2013 Vũ Lê Lợi đã tính toán tần suất Haplotype, tính chỉ số Haplotype và chỉ số các locus gen dựa trên số liệu phân tích các alen thuộc hệ 17 Y-STR trên NST Y của 80 người đàn ông Việt (Kinh) ở các vùng miền của Việt Nam [65]. Mới đây các nhà khoa học Argentina đã phối hợp với các nhà khoa học Việt Nam tiến hành khảo sát tần suất 17 locus Y-STR từ 218 mẫu nam giới thuộc một số nhóm dân tộc chính tại Việt Nam bao gồm dân tộc Kinh, Tày, Mông, Nùng, Hoa, Thái [66].

Có thể nói hiện nay vẫn chưa có nghiên cứu toàn diện trên một số lượng mẫu đủ lớn để đặc trưng cho quần thể người Việt Nam. Hiện tại cũng chưa có nghiên cứu nào mở rộng số lượng locus gen nghiên cứu lên 23-30 locus. Việc mở rộng số lượng locus Y-STR nghiên cứu là điều hết sức cần thiết để đánh giá được tần suất phân bố, đặc điểm của từng locus Y-STR trong quần thể người Việt Nam để từ đó lựa chọn và tạo nên một tập hợp Y-STR phù

hợp, đạt được độ tin cậy cũng như khả năng phân biệt hai cá thể nam giới không có quan hệ huyết thống với nhau trong quần thể người Việt Nam một cách chính xác nhất.

Đặc biệt trong lĩnh vực xét nghiệm huyết thống, điều tra khoa học hình sự hiện nay tại Việt Nam hiện nay vẫn chủ yếu dựa vào các chỉ thị STR trên NST thường. Các chỉ thị này thường chỉ có ý nghĩa trong xét nghiệm huyết thống trực hệ và nhiều khi không áp dụng được trong nhiều trường hợp: các mẫu phân tích thiếu hụt thông tin, mẫu nạn nhân trong thảm họa hàng loạt, mẫu có chứa nhiều nguồn ADN....Việc nghiên cứu khảo sát các thêm các chỉ thị Y-STR – là các chỉ thị đơn alen, di truyền ít biến đổi theo dòng cha là vô cùng cần thiết, đáp ứng nhu cầu giám định trong nhiều trường hợp. Do đó đề tài “***Nghiên cứu một số chỉ thị trên nhiễm sắc thể Y để ứng dụng trong giám định pháp y***” được thực hiện với mục tiêu chính là khảo sát đặc điểm các chỉ thị Y-STR mới này trên mẫu nam giới Việt Nam nhằm đánh giá tính phù hợp trong phân tích Y-STR ở quần thể người Việt Nam. Ngoài ra trong nước cũng chưa có một nghiên cứu nào về các mini Y-STR mới nằm ngoài các bộ kit thương mại do đó đề tài này cũng đặt mục tiêu khảo sát thêm một số locus Y-STR mới có độ đa hình cao nhằm làm tăng khả năng phân biệt giữa các cá thể nam giới trong quần thể người Việt Nam.

Bên cạnh đó, đề tài cũng nghiên cứu việc ứng dụng các chỉ thị Y-STR theo nhiều hướng khác nhau: xét nghiệm huyết thống theo dòng cha, giám định mẫu nam giới trong các vụ án hình sự, giám định mẫu đã bị phân huỷ.

CHƯƠNG 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Vật liệu nghiên cứu

2.1.1. Mẫu nghiên cứu

2.1.1.1. Đối với nghiên cứu khảo sát đa hình locus Y-STR thuộc bộ kit PPY23 và YPlus

Chúng tôi tiến hành nghiên cứu trên tổng số 400 mẫu nam giới dân tộc Kinh ở miền bắc Việt Nam trong đó mỗi bộ kit Y-STR sử dụng 200 mẫu khác nhau. Kích cỡ mẫu nghiên cứu được lựa chọn theo nguyên lý ước lượng cho một tỷ lệ (không biết tổng thể).

Tiêu chuẩn lựa chọn mẫu:

+ Mẫu được cung cấp bởi những người đã đến xét nghiệm ADN tại Viện Pháp y quốc gia.

+ Mẫu được lựa chọn ngẫu nhiên, giữa các mẫu được xác định không có quan hệ huyết thống theo dòng cha, thuộc dân tộc Kinh dựa trên hộ khẩu, chứng minh nhân dân.

+ Mẫu được thu là mẫu máu (mẫu máu khô thấm trên thẻ FTA), tóc (nhỏ trực tiếp 5-10 sợi tóc có chân) hoặc niêm mạc miệng (mẫu thu trên que tăm bông).

Mẫu sau khi thu được bảo quản nhiệt độ 4°C.

2.1.1.2. Đối với nghiên cứu trên mẫu đã bị phân huỷ

Nghiên cứu sử dụng 30 mẫu xương hài cốt được thu thập từ: Nghĩa trang liệt sĩ Hữu Nghị Việt Lào, huyện Anh Sơn, tỉnh Nghệ An (15 mẫu) và Nghĩa trang liệt sĩ huyện Định Hóa, tỉnh Thái Nguyên (15 mẫu).

2.1.1.3. Đối với nghiên cứu trên mẫu lẫn (mixture sample)

50 mẫu dịch âm đạo thu từ nạn nhân nữ trong những vụ án hiếp dâm, tấn công tình dục được cung cấp bởi cơ quan công an các tỉnh, thành phố trung cầu giám định ADN.

Mẫu đã được tách chiết ADN và bảo quản ở điều kiện -20°C.

2.1.2. Hoá chất nghiên cứu

Bảng 2.1. Kit và hoá chất chính sử dụng trong nghiên cứu

TT	Tên kit, hóa chất	Hãng sản xuất	Nước sản xuất
1	Đệm PBS	Thermo scientific	Mỹ
2	Đệm TE, proteinase K	Promega	Mỹ
3	Chelex [®] 100	Biorad	Mỹ
4	QIAamp [®] DNA micro kit.	Qiagen	Đức
5	QIAamp [®] DNA investigator kit	Qiagen	Đức
6	GoTaq [®] G2 Master Mixes	Promega	Mỹ
7	Primer	IDT	Singapore

8	EDTA 0,5M pH 8	Thermo scientific	Mỹ
9	Tris-base	Merck	Đức
10	Ethanol 100%	Merck	Đức
11	NaOH	Merck	Đức
12	AgNO ₃	Fisher	Mỹ
13	Acetic acid	Merck	Đức
14	Formaldehyde	Merck	Đức
15	Polycarylamid	Merck	Đức
16	GelRed TM Nucleic Acid Gel Stain	Fisher	Mỹ
17	Nước DEPC	Promega	Mỹ
18	Wizard SV Gel and PCR Clean - Up System	Promega	Mỹ
19	BigDye TM Terminator 3.1	Thermo scientific	Mỹ
20	BigDye TM Xterminator Purification kit	Thermo scientific	Mỹ
21	Pop-7 TM polymer	Thermo scientific	Mỹ

Bảng 2.2. Trình tự mỗi khuếch đại 10 mini Y-STR trong nghiên cứu

STT	Primer ID	Mũi xuôi (chiều 5'-3') Mũi ngược (chiều 3'-5')	Kích thước băng (bp)
1	DYS533	TCAATCTTCTACCTATCATCTTTCT TCAGTTCTTAACCAACA	107 – 127
2	DYS19	CTACTGAGTTTCTGTTATAGT ATGGCCATGTAGTGAGGACA	177 – 213
3	DYS481	AGGAATGTGGCTAACGCTG ACAGCTCACCAGAAGGTTG	119 – 149
4	DYS643	GCCATGCCTGGTTAACTACT TAACCAAACACCACCCATTCC	121 – 166
5	YGATA-H4	ATGCTGAGGAGAATTTCCAA CTATTCATCCATCTAATCTATCCATT	121 – 141
6	DYS505	TCTGGCGAAGTAACCCAAAC TCGAGTCAGTTCACCAGAAGG	164 – 176
7	DYS508	CCCAAATTCCAGCTTATTATTCC AGGTGATAGATGATATGCAGAGAGT	105 – 129
8	DYS522	TGGATAGACATAGGTGACAGATGAT GGAACCTCATGGTGGACATC	112 – 132
9	DYS388	AGCGAGAGTCCGTCTCAAAT TCATGTGAGTTAGCCGTTTAGC	91 – 106
10	DYS460	GAGGAATCTGACACCTCTGACA GTCCATATCATCTATCCTCTGCCTA	101 – 121

2.2. Thiết bị chính được sử dụng

Bảng 2.3. Thiết bị chính được sử dụng trong nghiên cứu

STT	Tên thiết bị	Model	Hãng sản xuất (nước)
1	Hệ thống điện di mao quản và giải trình tự gen	ABI 3500 Genetic Analyzers	Thermo Scientific, Mỹ
2	Bộ điện di đứng	Emperor Penguin P10 DS	Thermo scientific – Mỹ
3	Máy ly tâm nhanh	Mini centrifuge	Labnet - Mỹ
4	Hệ thống máy Realtime PCR	7500	Thermo scientific - Mỹ
5	Máy đo nồng độ ADN huỳnh quang	Quantus™ Fluorometer	Promega - Mỹ
6	Máy lắc ổn nhiệt	Vortemp 56TM	Labnet - Mỹ
7	Máy PCR	ProFlex	Thermo scientific - Mỹ
8	Tủ an toàn sinh học cấp II	AC2-4E1	Esco - Singapore

Các thiết bị phụ trợ khác:

- + Găng tay tiệt trùng, khẩu trang, giấy khử trùng.
- + Pipette các loại: 10 µl, 20 µl, 100 µl, 200 µl, 1000 µl và các loại đầu côn tương ứng.
- Ống eppendorf các loại
- + Ống ly tâm các loại.
- + Đĩa 96 giếng
- + Capillary array 36,50 cm (Applied Biosystems).

2.3. Phương pháp nghiên cứu

2.3.1. Phương pháp tách chiết ADN

Tách chiết ADN với 2 tiêu chí chính:

- (1) Đảm bảo mẫu giám định đủ lượng ADN để phân tích.
- (2) Đảm bảo lượng ADN tinh sạch để phân tích.

Nguyên tắc gồm 3 giai đoạn:

- (1) Phá vỡ màng tế bào và giải phóng axit nucleic (ADN và ARN)
- (2) Biến tính protein.
- (3) Tách ADN ra khỏi protein đã bị biến tính cùng các tạp chất khác.

Tùy theo từng loại mẫu: máu, lông/tóc, răng hay xương sẽ tiến hành tách chiết ADN theo phương pháp tách chiết theo các phương pháp khác nhau để đảm bảo thu được lượng ADN tối đa nhất.

2.3.1.1. Phương pháp tách chiết ADN bằng Chelex® 100

Đây là phương pháp tách chiết được chấp nhận sử dụng trong các phòng thí nghiệm ADN, phù hợp để sử dụng đối với các mẫu máu, lông, tóc và cả mẫu mô chưa bị phân hủy [67]. Trong nghiên cứu phương pháp dựa trên các bước thực hiện như sau:

Bước 1: Xử lý mẫu.

- Đối với mẫu máu:

+ Mẫu máu khô trên thẻ lưu mẫu, gạt tiết trùng: Lấy dao tách khoảng ¼ diện tích vòng tròn trên thẻ hoặc 0,5 cm² vết máu khô trên gạt thành những mảnh nhỏ cho vào ống eppendorf 1,5 ml vô trùng.

+ Mẫu máu tươi: Lấy khoảng 5 - 50 µl máu vào ống eppendorf 1,5 ml vô trùng.

- Đối với mẫu lông, tóc có chân: Lấy dao cắt 2 - 4 chân lông, tóc cho vào ống eppendorf 1,5ml.

Bước 2: Bổ sung 1ml đệm PBS vào trong ống eppendorf, vortex đều 5 - 10 giây.

Bước 3: Ly tâm ở nhiệt độ phòng 12000 vòng/phút trong vòng 5 phút.

Bước 4: Hút bỏ dịch nổi, giữ lại cặn (Lặp lại bước 2 - 4 từ 1 đến 2 lần để làm sạch mẫu).

Bước 5: Bổ sung 150 - 200 µl dung dịch Chelex 10% và 10 - 15 µl dung dịch Proteinase K (10mg/ml).

Bước 6: Ủ ở nhiệt độ 56°C trong khoảng 30 - 60 phút.

Bước 7: Đun mẫu ở nhiệt độ 100°C trong 8 phút để đảm bảo tất cả màng tế bào bị phá vỡ và protein bị biến tính.

Bước 8: Ly tâm 12000 vòng/phút trong 2 - 3 phút. Sau đó hút dịch nổi sang một ống eppendorf 1,5ml mới và bảo quản ở nhiệt độ -20°C.

2.3.1.2. Phương pháp tách chiết ADN bằng Qiamp® DNA micro kit

Phương pháp này được sử dụng đối với các mẫu có nồng độ ADN thấp, mẫu thu thập tại hiện trường các vụ án, mẫu dịch âm đạo và khó tách chiết ADN. Bộ kit QIAamp DNA Micro Kit sử dụng công nghệ đã được thiết lập tốt để tách chiết ADN trong nhân và ty thể từ lượng hoặc kích thước mẫu nhỏ. Bộ kit kết hợp các đặc tính liên kết chọn lọc của

màng silica với thể tích rửa giải linh hoạt từ 20 đến 100 μ l. Quy trình này phù hợp với nhiều loại vật liệu mẫu như lượng máu nhỏ, thể máu và mẫu mô nhỏ. Quy trình được thiết kế để đảm bảo rằng không có sự lây nhiễm chéo từ mẫu này sang mẫu khác và cho phép xử lý an toàn các mẫu có khả năng lây nhiễm. ADN được rửa giải trong Buffer AE hoặc nước và sẵn sàng để sử dụng trong các phản ứng khuếch đại hoặc để bảo quản ở -20°C . ADN tinh khiết không chứa protein, nuclease và các tạp chất khác.

Cơ chế của bộ kit gồm 4 bước chính: 1. Ly giải ADN khỏi tế bào: Mẫu được ly giải trong điều kiện biến tính cao ở nhiệt độ cao với sự hiện diện của proteinase K và Buffer ATL. 2. Gắn ADN lên màng silica gel của cột QIAamp MinElute. Các điều kiện về muối và pH đảm bảo rằng protein và các chất gây nhiễm khác có thể ức chế phản ứng PCR và các phản ứng enzym khác không bị giữ lại trên màng của cột QIAamp MinElute. 3. Rửa cột: Trong khi axit nucleic vẫn liên kết với màng của cột QIAamp MinElute, các chất gây nhiễm được rửa trôi một cách hiệu quả bằng cách sử dụng Buffer AW1 và sau đó là Buffer AW2. Để làm sạch ADN trong nhân, một bước rửa duy nhất với Buffer AW2 là đủ để loại bỏ các chất gây ô nhiễm. 4. Thôi ADN: ADN được rửa giải từ cột QIAamp MinElute bằng cách sử dụng một thể tích nhỏ của Buffer AE hoặc nước cất.

Các bước tách chiết cụ thể tuân theo hướng dẫn của nhà sản xuất.

2.3.1.3. Phương pháp tách chiết ADN bằng QIAamp® DNA investigator kit

Phương pháp này được sử dụng đối với các mẫu xương, răng hài cốt đã bị phân hủy, có nồng độ ADN thấp, ADN bị đứt gãy nhiều. Cơ chế của bộ kit và các bước tương tự như QIAamp® DNA micro kit.

Trước khi tách chiết mẫu cần được xử lý theo các bước:

- * Làm sạch bằng cơ học
 - Làm sạch bên ngoài mẫu xương/ răng bằng cơ học.
 - Đối với mẫu xương, cưa nhỏ thành các miếng khoảng 1 - 2 cm, mài sạch phần đã phân hủy bên ngoài và bên trong ống xương bằng máy mài.
- * Làm sạch bằng hoá chất
 - Rửa xương/ răng bằng Natri hypochlorit 1% - 5%, trong 20 - 40 giây (tùy theo chất lượng mẫu).
 - Rửa lại bằng nước khử ion vô trùng, không chứa ADN (lặp lại 3 lần).

- Ngâm mẫu xương/ răng trong Ethanol 100% trong 10 phút.
- Để khô mẫu xương/ răng ở nhiệt độ phòng, trong 3 - 5 giờ (hoặc cho đến khi mẫu khô hoàn toàn).

* Nghiền mẫu xương/ răng

- Nghiền xương/ răng bằng máy nghiền chuyên dụng sử dụng trong phòng thí nghiệm.

- Lấy từ 300 - 500 mg bột xương/ răng đã nghiền vào ống ly tâm 15 ml.

- Quy trình tách chiết: Các bước tách chiết tuân theo hướng dẫn của nhà sản xuất.

2.3.2. Phương pháp định lượng ADN

2.3.2.1. Đo nồng độ ADN

Đo nồng độ ADN sau tách chiết bằng máy QuantusTM Fluorometer (Promega – Mỹ)

- Hiệu chỉnh máy với QuantiFlour Dye theo hướng dẫn của nhà sản xuất.

- Trộn đều hoặc vortex 1 μ l mẫu với 200 μ l QuantiFlour[®] ONE dsADN Dye trong ống PCR 0.5 ml.

- Ủ hỗn hợp phản ứng trong 5 phút ở nhiệt độ phòng (cần tránh ánh sáng).

- Đặt ống PCR vào máy, đóng nắp và đọc kết quả hiện trên màn hình.

2.3.2.2. Định lượng bằng Realtime PCR

Sau khi tách chiết ADN ta tiến hành định lượng ADN bằng phương pháp Realtime PCR theo Trio DNA Quantification Kits (Applied Biosystem, Mỹ) trên máy 7500 Real-Time PCR (Applied Biosystem, Mỹ) theo hướng dẫn của nhà sản xuất với các bước chính bao gồm:

- Xây dựng đường chuẩn

- Chuẩn bị phản ứng

- Tiến hành phản ứng: chạy trên hệ thống 7500 Real-Time PCR

- Phân tích dữ liệu

2.3.3. Phương pháp điện di gel

Phương pháp điện di được sử dụng để xác định sự có mặt của ADN đồng thời phân biệt sự khác nhau về kích thước giữa những đoạn ADN quan tâm.

Nghiên cứu sử dụng phương pháp điện di Polyacrylamide 6% để phát hiện sản phẩm sau PCR. Sản phẩm sau điện di được phát hiện bằng phương pháp nhuộm bạc.

Phương pháp điện di Polyacrylamide được sử dụng để phân tách các đoạn ADN có kích thước nhỏ (5 - 500 bp) phù hợp với nghiên cứu các đoạn mini Y-STR có kích thước nhỏ dưới 200 bp.

2.3.4. Phương pháp PCR

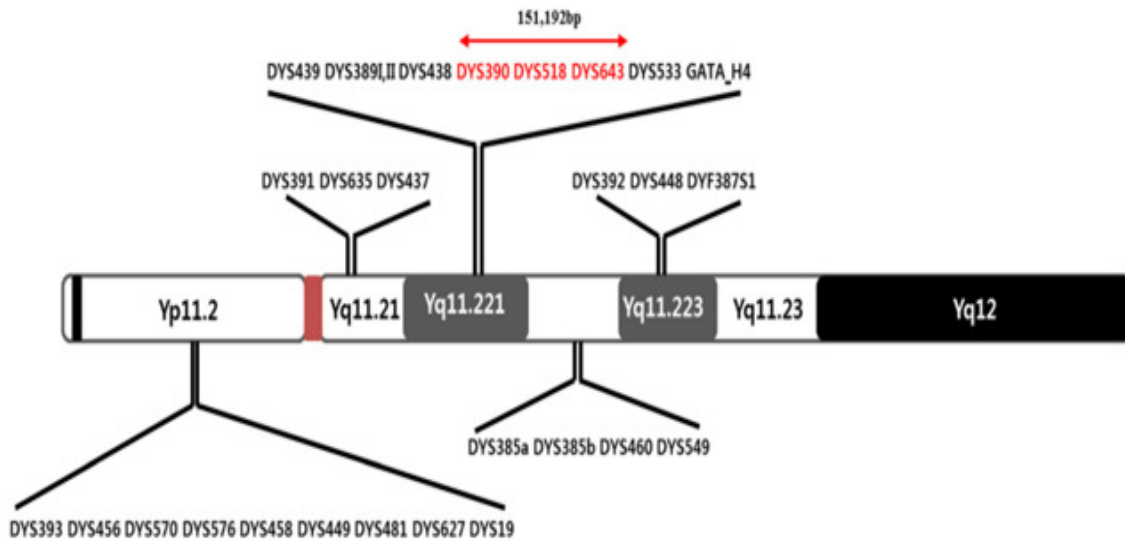
2.3.4.1. Phương pháp PCR phân tích Y-STR theo bộ kit PPY23 và YPlus

Tính đến nay có khoảng 400 marker Y-STR riêng rẽ được nghiên cứu rõ về cấu trúc và vị trí trên NST. Tuy nhiên để được lựa chọn thành locus có tiềm năng ứng dụng trong giám định ADN cần đáp ứng một số yêu cầu như: có độ đa hình cao, số alen tương đối lớn, cấu trúc lặp không quá phức tạp, đã được nghiên cứu rõ về trình tự gen, phương pháp phân tích đơn giản, có độ chính xác khi đọc kết quả. Dựa trên các tiêu chí đó trong đề tài này chúng tôi lựa chọn 29 locus Y-STR để khảo sát trong quần thể nam giới người Kinh ở Việt Nam, sử dụng hai bộ kit Y-STR thương mại phổ biến hiện nay là bộ PowerPlex®Y23 System (hãng Promega, gọi tắt là PPY23) và bộ Yfiler™ Plus PCR Amplification Kit (hãng Thermo Fisher Scientific, gọi tắt là YPlus). Hai bộ kit sử dụng cũng đã được chứng minh có độ tin cậy cao và được sử dụng rất nhiều trong việc lập bảng tần suất phân bố alen các Y-STR của các quần thể trên thế giới [48,49,50].

29 chỉ thị Y-STR được lựa chọn cũng đã được nghiên cứu rõ trên nhiều quần thể khác nhau cho thấy có độ đa hình cao, tiềm năng ứng dụng trong xét nghiệm huyết thống, phân tích hình sự. Ngoài ra các locus này cũng có mặt trong nhiều bộ kit thương mại theo khuyến cáo của các tổ chức phân tích ADN (SWGDM, NIST...) với phương pháp phân tích đơn giản là phản ứng PCR đa môi có gắn huỳnh quang và điện di mao quản. Do đó cho phép phân tích và đọc dữ liệu đồng thời của nhiều locus trong thời gian ngắn.

Thông tin chi tiết về cấu trúc lặp, số alen phổ biến thường gặp, tần suất đột biến, kích thước của từng locus được trình bày chi tiết trong phụ lục I. Kích thước các locus, khoảng alen dao động trong 2 bộ kit có thể khác nhau tùy theo việc thiết kế môi. Trường hợp DYS385 và DYSF387S1 cùng 1 cặp môi nhưng khuếch đại đồng thời 2 đoạn lặp khác nhau có vị trí nằm gần nhau, nên vẫn được tính là 4 locus DYS385a, DYS385b, DYF387S1a và DYF387Sb [73].

Physical map of STR used in Y23 and YPlus



Hình 2.1. Vị trí 29 locus Y-STR trên bản đồ NST Y

Phản ứng PCR nhằm mục đích khuếch đại các trình tự ADN quan tâm. Trong bộ PowerPlex® Y23 System phản ứng PCR là phản ứng đa môi (multiplex PCR) có gắn huỳnh quang cho phép khuếch đại đồng thời 23 chỉ thị Y-STR. Trong bộ Yfiler™ Plus PCR Amplification Kit khuếch đại đồng thời 27 chỉ thị Y-STR.

Thành phần phản ứng và chu trình PCR với hai bộ kit PowerPlex® Y23 System (hãng Promega - Mỹ) và Yfiler™ Plus PCR Amplification Kit (Thermo Fisher Scientific – Mỹ) tuân theo đúng hướng dẫn của nhà sản xuất [68,69].

2.3.4.2. Phương pháp PCR với 10 chỉ thị mini Y-STR

Trình tự các cặp môi nhân các locus Y-STR có kích thước ngắn: Dựa vào các nghiên cứu đã công bố trước đó chúng tôi đã lựa chọn các cặp môi phù hợp, tiến hành đặt môi để thực hiện phản ứng khuếch đại 10 chỉ thị mini Y-STR nêu trên [45, 46, 71].

- Thành phần phản ứng PCR:

2x PCR Master mix	: 12,5 µl
Môi xuôi 10pmol/ µl	: 0.5 µl
Môi ngược 10pmol/ µl	: 0.5 µl
ADN khuôn	: 1 µl
H ₂ O	: 10,5 µl

- Chu trình nhiệt được thực hiện trên máy PCR (Eppendorf-Đức)

Nhiệt độ	Thời gian	Số chu kỳ
96°C	5 phút	
94°C	30 giây	30 – 38 chu kỳ**
56-58°C*	30 giây	
72°C	60 giây	
60°C	10 phút	
10°C	∞	

*Nhiệt độ gắn môi tùy thuộc vào từng locus Y-STR

**Số chu kỳ thay đổi tùy thuộc nồng độ ADN (5 – 50 ng/μl) trong mẫu nghiên cứu

2.3.5. Phương pháp điện di mao quản

Đây là một phương pháp phổ biến cho sự phân tích điện di và có ứng dụng ngày càng nhiều trong nghiên cứu sinh hóa. Điện di mao quản là một biến đổi của điện di trên gel sử dụng cùng nguyên tắc phân tách dựa trên điện tích, kích thước của phân tử, nhưng được thực hiện trong ống mao dẫn với chất gel hoặc polyme lỏng. Các ống mao dẫn được chuẩn bị bằng silica nung chảy, và mỗi ống mao dẫn có đường kính trong từ 50-100μm và dài 25-100 cm. Các mẫu được tiêm vào ống mao dẫn chứa vật liệu polyme và được tách nhanh hơn nhiều so với phương pháp điện di trên gel thông thường.

Các bước thực hiện

- Trước khi điện di mao quản, sản phẩm PCR được pha loãng để lấy 1 μl.
- Chuẩn bị mẫu điện di:
 - Bộ kit PPY23: Thành phần phản ứng với tổng thể tích 11 μl: ILS-500 0,5 μl + Hi-Di Formamide: 9,5 μl + Sản phẩm PCR đã pha loãng: 1.0 μl
 - Bộ kit YPlus: Thành phần phản ứng với tổng thể tích 11 μl: GeneScan™ – 600 LIZ™ Size Standard v2.0: 0.4 μl + Hi-Di Formamide: 9,5 μl + Sản phẩm PCR đã pha loãng: 1.0 μl.

Với thang chuẩn Allelic ladder: 1,0 μl allelic ladder + 10,0 μl hỗn hợp mix kể trên.

Đối với mẫu trắng (blank): 11,0 μl Hi-Di Formamide

Các mẫu sau khi mix được tra vào giếng trên đĩa 96 giếng, sốc ở nhiệt độ 95°C, trong 3 phút. Sau đó chuyển ngay lập tức vào đá lạnh để biến tính 3 phút và cho vào hệ thống máy điện di mao quản 3500 Genetic Analyzer (Applied Biosystems - Mỹ) để phân tích với chương trình được cài đặt theo hướng dẫn của nhà sản xuất.

2.3.6. Phương pháp giải trình tự gen

Sản phẩm khuếch đại 10 locus STR được giải trình tự sử dụng phương pháp giải trình tự gen Sanger. Sản phẩm sau PCR được tinh sạch bằng kit Wizard SV Gel and PCR Clean - Up System (Promega - Mỹ). Trước khi tinh sạch cần điện di trên gel agarose để thu được băng đặc hiệu, cắt băng và làm theo hướng dẫn của nhà sản xuất theo các bước chính:

- Thực hiện phản ứng PCR với BigDye Terminator v3.1 Cycle Sequencing Kit.
- Tinh sạch sản phẩm sau chạy bigdye: sử dụng BigDye Xterminator purification kit.
- Tiến hành giải trình tự: Quá trình giải trình tự gen được thực hiện trên máy hệ thống máy ABI 3500 Genetic Analyzer (Thermo Scientific, Mỹ) với cấu hình cài đặt theo đúng hướng dẫn của nhà sản xuất. Dữ liệu được thu nhận trên phần mềm 3500 series Data collection software 3.

2.3.7. Phân tích và xử lý số liệu

Dữ liệu phân tích sau điện di mao quản được phân tích bằng phần mềm GeneMapper® ID-X 1.3 (Applied Biosystems - Mỹ).

Dữ liệu sau giải trình tự được đọc và phân tích trên phần mềm BioEdit, hiệu chỉnh sử dụng phần mềm SeqA, so sánh trình tự các mẫu sử dụng phần mềm ApE để đưa ra trình tự đúng của mẫu.

Tần số của từng alen trong các locus, tần suất của từng haplotype, tổng số alen và tổng số haplotype trong các nghiên cứu được xác định bằng Excel 2013 và phần mềm Arlequin 1.3.

Độ đa dạng về gen (gene diversity – GD) của từng locus Y-STR được tính theo công thức: $GD = n*(1 - \sum pi^2)/(n-1)$.

Haplotype diversity (HD) theo công thức của Nei [70]

$$HD = n*(1 - \sum pi_{th}^2)/(n-1)$$

Khả năng phân biệt (Discrimination capacity - DC) được tính theo công thức:

$$DC = h/n$$

Trong đó n là số mẫu, p_i là tần số của các alen tương ứng, p_{th} là tần số các haplotype tương ứng, h là số haplotype khác nhau được quan sát thấy trong quần thể. Ngoài ra chúng tôi cũng sử dụng công cụ online AMOVA trên website <http://yhrd.org> để tính toán khoảng cách di truyền quần thể (R_{st}) và giá trị xác suất kết hợp (P value) giữa quần thể nghiên cứu và các quần thể lân cận khác

Tính tỷ số tương đồng (LR- likelihood ratio) giữa hai hồ sơ Y-STR tuân theo Định lý Bayes. dựa theo công thức [111]:

$$LR = \frac{P(E|H_0)}{P(E|H_1)}$$

Tỷ số LR là tỷ lệ của 2 xác suất với cùng một chứng cứ với 2 giả thuyết khác nhau. Trong đó $P(E|H_0)$ là giả thiết giữa hai hồ sơ Y-STR có quan hệ theo dòng cha hay mẫu thu được tại hiện trường đúng là của hung thủ, $P(E|H_1)$ là giả thiết giữa hai hồ sơ Y-STR không có quan hệ theo dòng cha hay dấu vết ở hiện trường không phải của thủ phạm mà là của một người ngẫu nhiên khác không liên quan.

Từ đó tính được xác suất có quan hệ huyết thống hay còn gọi là chỉ số Combined Paternity Index (CPI) theo công thức:

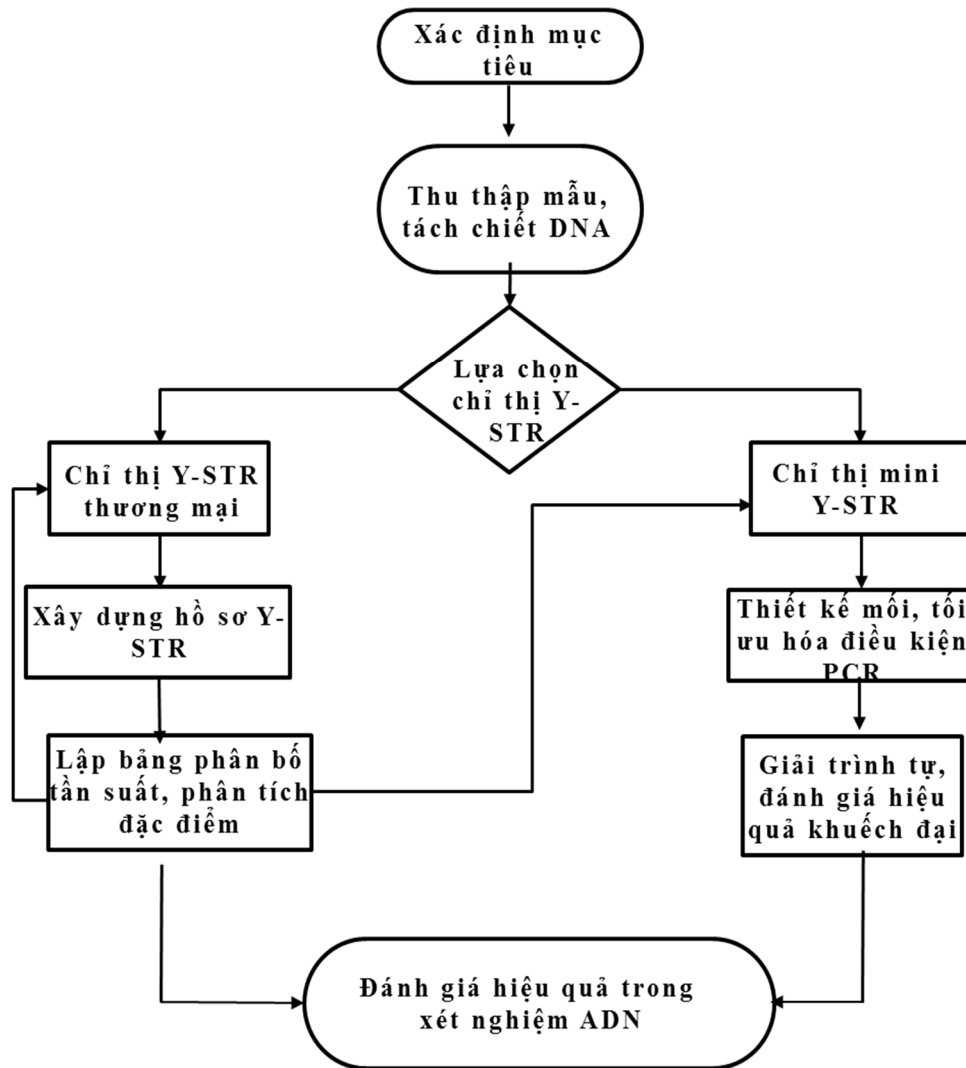
$$Probability\ of\ Relationship = (CRI \times Pr / [CRI \times Pr + (1 - Pr)]) \times 100$$

Trong đó CRI chính là LR, pr (prior probability) là xác suất tiên nghiệm là một phương pháp dự đoán khả năng một biến cố sẽ xuất hiện theo một cách nhất định dựa trên tỉ lệ phần trăm, chứ không phải được xác định từ dữ liệu lịch sử. Pr thường được giả thiết là 0.5.

2.4. Đạo đức trong nghiên cứu

- Mẫu chỉ dùng cho mục đích nghiên cứu, không dùng cho bất kỳ mục đích nào khác.
- Kết quả và thông tin mẫu hoàn toàn được bảo mật.
- Nghiên cứu được hội đồng khoa học Viện Pháp y quốc gia thông qua.
- Thông tin mẫu và các loại mẫu sử dụng trong nghiên cứu được sự cho phép của hội đồng khoa học.
- Nghiên cứu được tiến hành tại Khoa Y sinh học – Viện Pháp y quốc gia trong thời gian 4 năm từ tháng 1/2016 đến tháng 12/2020.

2.5. Sơ đồ thí nghiệm



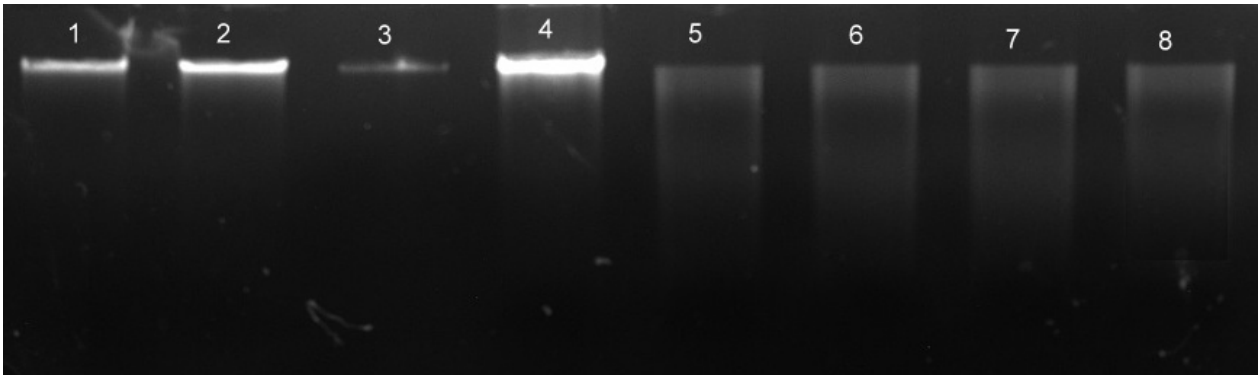
Hình 2.2. Sơ đồ thí nghiệm trong nghiên cứu

CHƯƠNG III. KẾT QUẢ VÀ BÀN LUẬN

3.1. Kết quả tách chiết ADN tổng số từ các mẫu nghiên cứu

3.1.1. Mẫu tách chiết theo phương pháp Chelex

Để tách chiết ADN đạt hiệu suất và độ tinh sạch cao thì việc lựa chọn một phương pháp tách chiết phù hợp với đối tượng nghiên cứu là rất quan trọng. Chất lượng ADN thu được sau tách chiết và tinh sạch ảnh hưởng đến kết quả quá trình nghiên cứu. Trong nghiên cứu này với những mẫu người sống có nồng độ ADN cao bao gồm: mẫu máu, tóc, niêm mạc miệng chúng tôi sử dụng phương pháp tách chiết bằng Chelex® 100 – là phương pháp đơn giản, nhanh và có hiệu quả cao. Phương pháp tách bằng hạt chelex có ưu điểm là ngăn chặn sự phân hủy ADN từ các enzym phân hủy (ADNase) và từ các chất gây nhiễm tiềm ẩn như Mg²⁺ có thể ức chế các phân tích sau. Tuy nhiên phương pháp có hạn chế là các bước gia nhiệt làm biến tính chuỗi xoắn kép và kết quả là ADN sợi đơn kém ổn định hơn trong quá trình bảo quản, nồng độ ADN thu được sau tách chiết không cao và không loại bỏ được hoàn toàn các chất gây nhiễm hoặc ức chế phản ứng PCR nên phù hợp với phân tích sử dụng các kit thương mại có độ nhạy lớn.



Hình 3. 1. Ảnh điện di ADN tổng số tách chiết bằng Chelex® 100 trên gel agarose 0.8%

Giếng 1,2: mẫu máu, giếng 3,4: mẫu niêm mạc miệng, giếng 5,6,7,8: mẫu tóc.

ADN tổng số sau tách chiết được kiểm tra trên gel agarose 0.8% và đo nồng độ trên máy Quantus™ Fluorometer (Promega). Đây là phương pháp đo nồng độ ADN dựa trên tín hiệu huỳnh quang phát ra, có độ nhạy và độ chính xác cao hơn so với phương pháp đo OD truyền thống. Nồng độ ADN thu được từ các loại mẫu niêm mạc, tóc, máu tách theo phương pháp Chelex® 100 dao động trong khoảng từ 2,70 ng/μl đến 10,05 ng/μl. Như vậy, các mẫu ADN tổng số này đều đạt tiêu chuẩn về số lượng và chất lượng để sử dụng làm ADN đầu vào cho nghiên cứu 29 chỉ thị Y-STR sử dụng bộ kit PPY23 và YPlus.

3.1.2. Mẫu tách chiết bằng kit thương mại

Với các mẫu có chất lượng kém, mẫu đã bị phân hủy nhiều như mẫu xương hài cốt, mẫu giám định án có lượng ADN ít, đã bị phân hủy đứt gãy nhiều dưới tác động của các yếu tố môi trường. Ngoài ra những mẫu này còn dễ nhiễm từ nhiều nguồn, có chứa các chất ức chế phản ứng PCR. Do đó việc tách chiết bằng Chelex[®] 100 là không phù hợp. Thay vào đó chúng tôi sử dụng QIAamp[®] DNA micro kit và QIAamp[®] DNA investigator kit – là phương pháp tách chiết bằng cột silica cho hiệu quả tách chiết ADN cao hơn đồng thời loại bỏ được những chất gây nhiễm khác. Tuy vậy phương pháp này cũng có nhược điểm là cần nhiều hóa chất, thời gian tách chiết hơn so với phương pháp Chelex[®] 100, cần có máy ly tâm, khó có thể thực hiện tự động hóa do đó khi áp dụng trên số mẫu lớn đòi hỏi tiêu tốn thời gian, công sức nhiều hơn.

Nồng độ ADN thu được sau tách chiết rất thấp nên khó áp dụng phương pháp định lượng thông thường như đo OD hoặc điện di gel để kiểm tra hiệu suất tách chiết. Thay vào đó chúng tôi sử dụng phương pháp định lượng bằng Trio DNA Quantification Kits (hãng Thermo Fisher Scientific – Mỹ) trên hệ thống máy Real time PCR. Phương pháp có độ nhạy, độ đặc hiệu cao cho phép đánh giá được nồng độ từng loại ADN kích thước dài, ngắn và ADN trên NST Y có trong mẫu:

- ADN đích có kích thước nhỏ: xác định nồng độ đoạn ADN có kích thước 80 bp.
- ADN đích có kích thước lớn: xác định nồng độ đoạn ADN có kích thước 214 bp, là chỉ số để đánh giá mức độ phân hủy khi so sánh với tỉ lệ ADN có kích thước nhỏ
- ADN trên NST Y: xác định nồng độ các đoạn ADN kích thước 75 bp trên NST Y, cho phép định lượng thành phần ADN bộ gen nam giới và đặc biệt hữu ích trong việc đánh giá các mẫu lẫn nhiều nguồn ADN từ cả nam và nữ.

Kết quả Realtime PCR của mẫu xương, răng hài cốt ký minh họa cho thấy hàm lượng ADN kích thước lớn (T. Large autosomal) là rất ít trong khi hàm lượng ADN kích thước nhỏ (T. Small autosomal) lại cao hơn. Chỉ số phân hủy (DI-Degradation Index) được tính bằng tỷ lệ giữa hàm lượng ADN kích thước nhỏ/ ADN kích thước lớn đều lớn hơn 1 chứng tỏ mẫu hài cốt có ADN bị phân hủy. Những mẫu như trên sẽ được xếp vào nhóm mẫu đã phân hủy và được dùng trong nghiên cứu các chỉ thị mini Y-STR.

Bảng 3. 1. Kết quả realtime PCR của mẫu xương, răng ký hiệu R1-R5.

T.Large autosomal: ADN kích thước lớn, T.Small autosomal: ADN kích thước nhỏ, T.Y: hàm lượng ADN trên NST Y.

Ký hiệu mẫu	ADN đích	Nồng độ ADN ng/ μ l	Chỉ số phân hủy (Degration index)
M1H1KA	T.Large Autosomal	0.0572	1.376
	T.Small Autosomal	0.0787	
	T.Y	0.0721	
M3H1KA	T.Large Autosomal	0.0542	1.897
	T.Small Autosomal	0.1028	
	T.Y	0.0974	
M13H1KA	T.Large Autosomal	0.0124	1.633
	T.Small Autosomal	0.0202	
	T.Y	0.0196	
AS_151B5	T.Large Autosomal	0.0544	1.740
	T.Small Autosomal	0.0947	
	T.Y	0.0575	
AS_172B5	T.Large Autosomal	0.0665	1.514
	T.Small Autosomal	0.1007	
	T.Y	0.0975	

3.2. Kết quả khảo sát 29 chỉ thị Y-STR từ 2 bộ kit PPY23 và Yfiler Plus

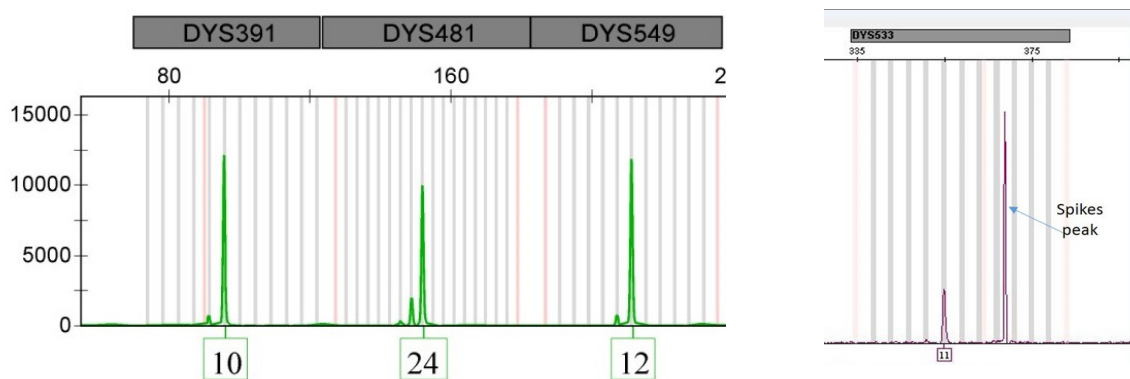
3.2.1. Xây dựng hồ sơ Y-STR

Việc lập hồ sơ Y-STR của 400 mẫu nghiên cứu được thực hiện dựa trên nguyên lý PCR đa mồi gắn huỳnh quang và điện di mao quản tuân theo hướng dẫn của nhà sản xuất. Dữ liệu được đọc kết quả bằng phần mềm GeneMapper® ID-X 1.3. Trong quá trình đọc phân tích dữ liệu cần lưu ý đến một số vấn đề hay gặp phải để tránh đọc sai các đỉnh (peak).

Phổ biến hay gặp nhất là hiện tượng đỉnh lặp (stutter peak). Dạng đỉnh lặp được tìm thấy trong hầu hết mọi bản peak, thường gặp với các locus có cấu trúc lặp 4 nucleotit (tetranucleotide). Đỉnh lặp là các đỉnh nhỏ xuất hiện ngay trước hoặc sau một đỉnh thực sự. Nguyên nhân là do trong quá trình khuếch đại PCR, enzyme polymerase có thể nhận nhầm vị trí gắn mồi khi sao chép một chuỗi ADN, thường trượt về phía trước hoặc lùi về sau bốn cặp nucleotit. Kết quả là một số lượng nhỏ các bản sao đoạn ADN được lặp lại với kích thước nhỏ hơn hoặc lớn hơn đoạn thực sự được khuếch đại một hoặc hai đơn vị lặp. Thông

thường, các alen có số lượng đơn vị lặp lại nhiều hơn sẽ biểu hiện phần trăm stutter peak cao hơn. Các locus lặp lại dạng trinucleotide cũng thường có stutter cao hơn so với các locus có độ dài lặp lại dài hơn. DYS481 là một dạng lặp lại trinucleotide lặp lại và có dạng stutter peak cao (Hình 3.2). Stutter peak được chấp nhận khi có chiều cao đỉnh nhỏ hơn 10% khi so với đỉnh của peak chính. Khi đọc dữ liệu đặc biệt là trong các trường hợp mẫu lần, mẫu có nồng độ ADN thấp, tín hiệu đỉnh không cao thì cần phân biệt đỉnh thực sự và peak lặp để đưa ra kết luận chính xác hoặc thực hiện lại phản ứng PCR [74].

Trong nghiên cứu này chúng tôi nhận thấy khi phân tích với bộ kit PPY23 hiện tượng stutter peak thường gặp ở các locus DYS391, DYS481, DYS635, DYS390, DYS458. Với bộ kit YPlus stutter peak thường gặp ở các locus DYS392, DYS518, DYS481, DYS449. Hiện tượng đỉnh gai thường ít gặp hơn với cả 2 bộ kit. Theo hướng dẫn của nhà sản xuất chúng tôi giữ nguyên các bước và thành phần hóa chất trong chu trình PCR và điện di mao quản tuy nhiên đã tối ưu lại nồng độ ADN đầu vào cho phản ứng PCR để hạn chế tối đa các peak giả. Theo hướng dẫn của nhà sản xuất 2 bộ kit cho phép khuếch đại ADN khuôn với nồng độ 0.5 ng/ μ l trở lên nhưng nếu quá nhiều ADN khuôn có thể gây nên một số vấn đề khi thực hiện PCR và đọc dữ liệu như đỉnh peak quá cao, tạo nhiều peak giả, chiều cao các đỉnh peak không cân bằng... Trong nghiên cứu khảo sát với hai bộ kit chúng tôi nhận thấy nồng độ ADN tối ưu trong khoảng 0.5 – 10 ng/ μ l.



Hình 3. 2. Dạng đỉnh lặp xuất hiện ở locus DYS481 khi phân tích với bộ PPY23 (mẫu phân tích ký hiệu ID101) và Dạng đỉnh gai xuất hiện ở locus DYS533 khi phân tích với bộ Yfiler Plus (mẫu phân tích ký hiệu ID56)

Đỉnh gai (spikes peak) cũng là một dạng đỉnh giả gặp trong quá trình phiên giải mẫu. Đỉnh gai thường thể hiện một hình dạng đỉnh mỏng, hẹp bất thường. Nguyên nhân xuất

hiện có thể do sự dao động của điện áp hoặc sự hiện diện của bọt khí trong mao quản khi điện di. Đỉnh gai cũng có thể được gây ra bởi các tinh thể trong polymer chạy máy. Đỉnh gai, không giống như các dạng peak giả khác, thường được nhìn thấy ở cùng một vị trí trong tất cả các kênh màu. Sự xuất hiện của các đỉnh gai có thể được giảm thiểu bằng cách tuân theo đúng hướng dẫn bảo quản mẫu và thuốc thử của nhà sản xuất.

Dữ liệu được coi có đủ độ chính xác và dùng trong kết luận loại trừ khi đỉnh peak trên 150 RFUs (RFU: relative fluorescence units). Những trường hợp không đáp ứng đủ yêu cầu như đỉnh thấp, mẫu bị nhiễm trong quá trình thao tác (peak lên nhiều hơn 2 ở các locus thông thường và nhiều hơn 3 ở hai locus *DYS385ab* và *DYF387S1ab*) sẽ không đưa vào dữ liệu thống kê chung để đảm bảo tính chính xác cho kết quả thống kê. Một dữ liệu về hồ sơ Y-STR hoàn chỉnh phải bao gồm đủ các locus, mỗi locus chỉ gồm 1 alen, các đỉnh nằm trong khoảng đơn vị cho phép, các alen nằm giới hạn của thang alen chuẩn. Như vậy với việc tuân theo đúng hướng dẫn về việc phiên giải dữ liệu Y-STR do Ủy ban ADN của Hiệp hội Di truyền Pháp y Quốc tế (ISFN) ban hành, chúng tôi đã xây dựng hồ sơ dữ liệu STR hoàn chỉnh của 400 mẫu nghiên cứu [28]. Công cụ “Data file validator” trên website YHRD hỗ trợ hiệu chỉnh, chỉ ra các alen nằm ngoài giới hạn được công bố giúp ích đáng kể cho việc xem xét thẩm định lại dữ liệu.

3.2.2. Bảng phân bố tần suất alen thuộc 29 chỉ thị Y-STR

Tần suất của từng alen trong các locus Y-STR được thống kê và tính toán trên phần mềm Arlequin 1.3 để lập thành bảng tần suất phân bố các alen của từng locus Y-STR. Bảng dữ liệu tần suất alen của từng locus cũng được so sánh với dữ liệu chung đã được tập hợp từ số liệu được công bố trên trang web YHRD (www.yhrd.org) để thấy được sự tương đồng cũng như sai khác về tần suất phân bố alen của 29 locus Y-STR trong nghiên cứu. Với mỗi locus alen có tần suất phân bố lớn nhất được in đậm trong bảng thống kê. Các alen có tần số thấp là các alen có tần số $\leq 5/2N$, trong đó N là kích thước mẫu. Như vậy, đối với nghiên cứu này thì các alen có tần số ≤ 0.0125 được coi là các alen có tần số thấp, ít xuất hiện trong quần thể.

Bảng 3. 2. Tần suất phân bố chung của 25 locus Y-STR thuộc hai bộ kit PPY23 và YPlus (4 chỉ thị *DYS385a/b* và *DYF387S1* được thống kê riêng)

DYS576			DYS389 I			DYS448			DYS389 II			DYS19		
Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus

14	0.000	0.005	11	0.085	0.085	16	0.010		25	0.000	0.005	12		0.005
15	0.030	0.015	12	0.310	0.350	17	0.010	0.045	27	0.050	0.095	13	0.045	0.015
16	0.080	0.060	13	0.385	0.385	18	0.425	0.530	28	0.295	0.260	14	0.120	0.135
17	0.160	0.185	14	0.215	0.180	18.2	0.010		29	0.355	0.340	15	0.415	0.450
18	0.335	0.375	15	0.005		19	0.235	0.215	30	0.225	0.230	16	0.355	0.330
19	0.265	0.265				19.2	0.005		31	0.060	0.055	17	0.065	0.065
20	0.110	0.080				20	0.180	0.155	32	0.015	0.010			
21	0.015	0.015				21	0.115	0.055	33		0.005			
22	0.005					22	0.010							
DYS391			DYS481			DYS549			DYS533			DYS438		
sAlen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus
6	0.020		19		0.005	10	0.005		9	0.005		8	0.005	0.005
8		0.005	21	0.020	0.020	11	0.175		10	0.340	0.390	9	0.005	0.005
9	0.010	0.025	22	0.085	0.130	12	0.560		11	0.445	0.435	10	0.840	0.825
10	0.630	0.615	23	0.330	0.405	13	0.235		12	0.190	0.165	11	0.120	0.145
11	0.325	0.330	24	0.295	0.170	14	0.015		13	0.015	0.010	12	0.030	0.015
12	0.015	0.025	25	0.140	0.115	15	0.010		14	0.005		14		0.005
			26	0.095	0.105									
			27	0.030	0.035									
			28	0.005	0.015									
DYS437			DYS570			DYS635			DYS390			DYS439		
Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus
12	0.000	0.005	14	0.000	0.025	18	0.005	0.005	22	0.025	0.035	10	0.035	0.025
13	0.005	0.000	15	0.040	0.045	19	0.050	0.030	23	0.255	0.180	11	0.255	0.260
14	0.730	0.730	16	0.330	0.310	20	0.130	0.110	24	0.360	0.385	12	0.530	0.520
15	0.255	0.250	17	0.230	0.175	21	0.365	0.405	25	0.305	0.365	13	0.160	0.160
16	0.010	0.015	18	0.135	0.135	22	0.195	0.205	26	0.055	0.035	14	0.015	0.030
			19	0.145	0.140	23	0.190	0.195				15	0.005	0.005
			20	0.075	0.120	24	0.050	0.040						
			21	0.025	0.030	25	0.015	0.010						
			22	0.015	0.010									
			23	0.005	0.010									
DYS449			DYS518			DYS393			DYS458			DYS627		
Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus
23		0.005	32	0.005		10	0.000	0.005	12		0.010	13		0.005
25		0.040	33	0.005		11	0.000	0.025	13	0.005	0.005	14		
26		0.160	34	0.015		12	0.375	0.310	14	0.025	0.015	15		
27		0.080	35	0.060		13	0.180	0.165	15	0.120	0.135	16		0.005
28		0.090	36	0.120		14	0.400	0.465	16	0.170	0.190	17		0.005
29		0.105	37	0.120		15	0.045	0.030	17	0.195	0.200	18		0.065
30		0.095	38	0.095					18	0.310	0.295	19		0.080
31		0.145	39	0.105					19	0.140	0.085	20		0.230
32		0.125	40	0.150					20	0.025	0.045	21		0.230
33		0.050	41	0.145					21		0.015	22		0.235
34		0.040	42	0.105					22	0.010	0.000	23		0.100
35		0.040	43	0.065					23		0.005	24		0.045
36		0.025	44	0.005										
			45	0.005										
DYS456			YGATAH4			DYS643			DYS460			DYS392		
Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	PPY23	YPlus
12		0.005	10	0.130	0.125	8	0.015		9		0.110	10		0.010
13	0.090	0.060	11	0.365	0.425	9	0.075		10		0.645	11	0.100	0.045
14	0.180	0.165	12	0.465	0.425	10	0.190		11		0.190	12	0.035	0.030
15	0.525	0.570	13	0.040	0.025	11	0.290		12		0.050	13	0.655	0.700
16	0.160	0.155				12	0.350		13		0.005	14	0.200	0.205
17	0.025	0.040				13	0.080					15	0.010	0.010
18	0.020	0.005												

3.2.2.1. Locus *DYS576***Bảng 3. 3.** Tần suất phân bố các alen thuộc locus *DYS576*

DYS576

ALEN	14	15	16	17	18	19	20	21	22
PPY23		0.030	0.080	0.160	0.335	0.265	0.110	0.015	0.005
YPlus	0.005	0.015	0.060	0.185	0.375	0.265	0.080	0.015	

Locus *DYS576* có tất cả 9 alen được phát hiện trong 400 mẫu nghiên cứu trong đó mỗi bộ kit phát hiện được 7 alen. Alen 14 và 22 là các alen hiếm, mỗi alen chỉ gặp ở 01 cá thể trong quần thể. Alen số 18 có tần suất cao nhất với tần suất khoảng 0.335 – 0.375 khá tương đồng với dữ liệu thống kê chung. Tiếp theo là alen 19, 17 lần lượt có tần suất lớn thứ 2, 3 trong nghiên cứu. Tuy nhiên, trong dữ liệu công bố chung tần suất 2 alen này lại có sự phân bố ngược lại so với trong nghiên cứu.

3.2.2.2. Locus *DYS389I*

Bảng 3. 4. Tần suất phân bố các alen thuộc locus *DYS389I*

DYS389 I					
ALEN	11	12	13	14	15
PPY23	0.085	0.310	0.385	0.215	0.005
YPlus	0.085	0.350	0.385	0.180	

Locus *DYS389I* có tất cả 5 alen được phát hiện trong đó alen 15 là alen hiếm chỉ bắt gặp ở 1 cá thể trong quần thể khi phân tích với bộ kit PPY23. Alen 13 có tần suất cao nhất tần suất với cả 2 bộ kit đều là khoảng 0.385, tiếp theo là alen 12 với tần suất khoảng 0.31-0.35. Sự phân bố các alen trong locus *DYS389I* tương đối giống với dữ liệu thống kê chung.

3.2.3.3. Locus *DYS448*

Bảng 3. 5. Tần suất phân bố các alen thuộc locus *DYS448*

DYS448									
ALEN	16	17	18	18.2	19	19.2	20	21	22
PPY23	0.010	0.010	0.425	0.010	0.235	0.005	0.180	0.115	0.010
YPlus		0.045	0.530		0.215		0.155	0.055	

Locus *DYS448* có tất cả 9 alen được phát hiện trong đó các alen 16, 18.2, 19.2, 22 là các alen tương đối hiếm chỉ gặp ở 1-2 cá thể khi phân tích với bộ kit PPY23. Ở cả 2 bộ kit alen 18 là alen có tần suất cao nhất tần suất khoảng 0.435-0.53, tiếp theo là alen 19, và 20. Sự phân bố các alen ở locus *DYS448* có sự sai khác lớn so với dữ liệu

thống kê chung khi ở đó alen 19, 20 mới là alen có tần suất cao nhất, alen 18 chỉ có tần suất khoảng 0.17-0.18.

3.2.2.4. Locus *DYS389II*

Bảng 3. 6. Tần suất phân bố các alen thuộc locus *DYS389II*

DYS389 II								
ALEN	25	27	28	29	30	31	32	33
PPY23		0.050	0.295	0.355	0.225	0.060	0.015	
YPLUS	0.005	0.095	0.260	0.340	0.230	0.055	0.01	0.005

Với locus *DYS389II* có tất cả 8 alen được phát hiện trong đó alen số 25, 27, 33 là alen tương đối hiếm. Alen 29, 28, 30 lần lượt là các alen có tần suất cao nhất và tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.5. Locus *DYS19*

Bảng 3. 7. Tần suất phân bố các alen thuộc locus *DYS19*

DYS19						
ALEN	12	13	14	15	16	17
PPY23		0.045	0.120	0.415	0.355	0.065
YPlus	0.005	0.015	0.135	0.450	0.330	0.065

DYS19 là locus đầu tiên được khám phá và khuếch đại bằng phương pháp PCR [75]. Khi phân tích với locus *DYS19* có tất cả 6 alen được quan sát thấy trong quần thể nghiên cứu trong đó alen 12 là alen hiếm chỉ bắt gặp ở 1 cá thể khi phân tích với bộ kit YPlus. Alen số 15 có tần suất cao nhất (0.415-0.45), tiếp theo là alen 16, các alen còn lại có tần suất khá thấp. Ngược lại, trong dữ liệu thống kê chung alen có tần suất alen cao nhất là 14 cho thấy sự khác biệt trong sự phân bố các alen thuộc locus *DYS19* ở quần thể người Việt Nam.

3.2.2.6. Locus *DYS391*

Bảng 3. 8. Tần suất phân bố các alen thuộc locus *DYS391*

DYS391

ALEN	6	8	9	10	11	12
PPY23	0.020		0.010	0.630	0.325	0.015
YPlus		0.005	0.025	0.615	0.330	0.025

Với locus *DYS391* có tất cả 6 alen được phát hiện trong đó alen số 6, 8 là alen hiếm. Mỗi alen chỉ quan sát thấy ở 1 cá thể với mỗi bộ kit phân tích Alen số 10 là alen chiếm đa số trong quần thể với tần suất khoảng 0.615-0.63 và phản ánh sự tương đồng với dữ liệu thống kê chung trên YHRD. Đây cũng là locus được lựa chọn để có trong tập hợp STR phân tích quan hệ huyết thống trực hệ cha/mẹ-con của hầu hết các bộ kit hiện nay như PowerPlex® Fusion System (Promega), GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific)... như một chỉ thị hỗ trợ giúp phân biệt giữa mẫu nữ giới và nam giới.

3.2.2.7. Locus *DYS481*

Bảng 3. 9. Tần suất phân bố các alen thuộc locus *DYS481*

DYS481									
ALEN	19	21	22	23	24	25	26	27	28
PPY23		0.020	0.085	0.330	0.295	0.140	0.095	0.030	0.005
YPlus	0.005	0.020	0.130	0.405	0.170	0.115	0.105	0.035	0.015

Đây được coi là 1 trong những locus có số lượng alen nhiều nhất khi phân tích với cả 2 bộ kit. Locus *DYS481* có tất cả 9 alen được phát hiện trong 400 mẫu nghiên cứu trong đó alen 19 là alen hiếm, chỉ bắt gặp ở 1 cá thể. Alen số 23 có tần suất cao nhất (khoảng 0.33-0.40), tiếp theo là alen 24. So sánh với dữ liệu phân bố chung trên YHRD cho thấy có sự khác biệt nhỏ khi ở đó alen 22 có tần suất lớn thứ 2 với khoảng > 0.2 nhưng ở quần thể người Việt Nam là alen 24 mới là alen có tần suất lớn thứ 2 trong khi alen 22 có tần suất ít hơn hẳn (chỉ khoảng 0.08 - 0.13).

3.2.2.8. Locus *DYS549*

Bảng 3. 10. Tần suất phân bố các alen thuộc locus *DYS549*

DYS549						
ALEN	10	11	12	13	14	15

PPY23	0.005	0.175	0.56	0.235	0.015	0.01
-------	-------	-------	-------------	-------	-------	------

Locus DYS549 chỉ có trong bộ kit PPY23 với 6 alen được phát hiện trong đó alen 12 có tần suất cao nhất. Alen 10, 15 là các alen hiếm. Sự phân bố các alen trong locus này tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.9. Locus DYS533

Bảng 3. 11. Tần suất phân bố các alen thuộc locus DYS533

DYS533						
ALEN	9	10	11	12	13	14
PPY23	0.005	0.340	0.445	0.190	0.015	0.005
YPlus		0.390	0.435	0.165	0.010	

Khi phân tích với locus DYS533 phát hiện được tất cả 6 alen trong đó có alen số 9, 14 là hai alen hiếm chỉ bắt gặp ở 01 cá thể khi phân tích với bộ kit PPY23. Alen 11 có tần suất cao nhất khoảng 0.43 tương tự như trong dữ liệu công bố chung YHRD, tiếp theo là alen 10. Tuy nhiên trên YHRD với locus DYS533 alen có tần suất cao thứ 2 lại là alen số 12 với tần suất khoảng 0.4 trong khi ở quần thể người Việt Nam alen này có tần suất chỉ khoảng 0.18 cho thấy sự khác biệt trong sự phân bố các alen giữa quần thể người Việt Nam và dữ liệu chung trên thế giới.

3.2.2.10. Locus DYS438

Bảng 3. 12. Tần suất phân bố các alen thuộc locus DYS438

DYS438						
ALEN	8	9	10	11	12	14
PPY23	0.005	0.005	0.840	0.120	0.030	
YPlus	0.005	0.005	0.825	0.145	0.015	0.005

Với alen DYS438 có tất cả 6 alen được quan sát thấy trong quần thể nghiên cứu với alen số 8, 9 và 14 là các alen hiếm. Alen số 10 chiếm ưu thế rất lớn trong quần thể nghiên cứu với tần suất khoảng 0.85-0.84. Trên dữ liệu YHRD alen này cũng là alen chiếm ưu thế nhưng tần suất phân bố chỉ khoảng 0.45.

3.2.2.11. Locus DYS437

Bảng 3. 13. Tần suất phân bố các alen thuộc locus DYS437

DYS437					
ALEN	12	13	14	15	16
PPY23		0.005	0.730	0.255	0.010
YPlus	0.005		0.730	0.25	0.015

Đây được coi là một trong những alen có số alen ít nhất được khảo sát trong nghiên cứu với chỉ 5 alen được quan sát thấy, trong đó 2 alen số 12, 13 là các alen hiếm. Alen số 14 với tần suất phân bố 0.73 ở cả 2 bộ kit là alen chiếm ưu thế tối đa, tiếp theo là alen 15, 16. Sự phân bố các alen trên locus DYS437 hoàn toàn tương đồng với dữ liệu chung YHRD.

3.2.2.12. Locus DYS570

Bảng 3. 14. Tần suất phân bố các alen thuộc locus DYS570

DYS570										
ALEN	14	15	16	17	18	19	20	21	22	23
PPY23		0.040	0.330	0.230	0.135	0.145	0.075	0.025	0.015	0.005
YPlus	0.025	0.045	0.310	0.175	0.135	0.140	0.120	0.030	0.010	0.010

DYS570 là locus chung của cả 2 bộ kit có số lượng alen nhiều nhất với 10 alen được quan sát thấy trong nghiên cứu. Các alen có tần suất phân bố lớn nhất lần lượt là alen số 16, 17, 18. Sự phân bố này hơi khác so với dữ liệu chung YHRD khi alen số 17, 18 mới là các alen chiếm phân bố ưu thế, alen số 16 chỉ có tần suất khoảng 0.16.

3.2.2.13. Locus DYS635

Bảng 3. 15. Tần suất phân bố các alen thuộc locus DYS635

DYS635								
ALEN	18	19	20	21	22	23	24	25
PPY23	0.005	0.050	0.130	0.365	0.195	0.190	0.050	0.015
YPlus	0.005	0.030	0.110	0.405	0.205	0.195	0.040	0.010

Với locus DYS635 có tất cả 8 alen được quan sát thấy trong quần thể nghiên cứu. Trong đó alen 18 là alen hiếm, alen 21 là alen có tần suất phân bố lớn nhất khoảng 0.365-0.40 khá tương đồng so với dữ liệu YHRD. Các alen 22 và 23 có tần suất xấp xỉ nhau là các alen có tần suất phân bố lớn thứ 2 trong nghiên cứu. Ngược lại trong dữ liệu YHRD alen số 22 có tần suất khá thấp so với 2 alen còn lại.

3.2.2.14. Locus DYS390

Bảng 3. 16. Tần suất phân bố các alen thuộc locus DYS390

DYS390					
ALEN	22	23	24	25	26
PPY23	0.025	0.255	0.360	0.305	0.055
YPlus	0.035	0.180	0.385	0.365	0.035

Locus DYS390 có tất cả 5 alen được phát hiện trong đó alen 24 là alen có tần suất cao nhất tần suất ở cả 2 bộ kit (khoảng 0.36-0.385), tiếp theo là alen 23, 25 với tần suất gần bằng. Sự phân bố các alen này tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.15. Locus DYS439

Bảng 3. 17. Tần suất phân bố các alen thuộc locus DYS439

DYS439						
ALEN	10	11	12	13	14	15
PPY23	0.035	0.255	0.530	0.160	0.015	0.005
YPlus	0.025	0.260	0.520	0.160	0.030	0.005

Locus DYS439 có tất cả 6 alen được phát hiện trong đó alen 15 là alen hiếm. Alen 12 là alen có tần suất cao nhất tần suất ở cả 2 bộ kit (khoảng 0.52), tiếp theo là alen 11 có tần suất khoảng 0.25. Sự phân bố các alen này tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.16. Locus DYS392

Bảng 3. 18. Tần suất phân bố các alen thuộc locus DYS392

DYS392						
ALEN	10	11	12	13	14	15
PPY23		0.100	0.035	0.655	0.200	0.010
YPlus	0.010	0.045	0.030	0.700	0.205	0.010

Với alen DYS392 có tất cả 6 alen được quan sát thấy trong quần thể nghiên cứu. Alen số 13 chiếm ưu thế rất lớn trong quần thể nghiên cứu với tần suất khoảng 0.655-0.70, tiếp theo là alen số 12 với tần suất khoảng 0.30, các alen khác có tần suất khá thấp >0.20. Trái ngược lại trên dữ liệu YHRD alen 11 và 13 mới là các alen có tần suất cao nhất, alen số 12 lại có tần suất phân bố rất thấp dưới 0.10.

3.2.2.17. Locus DYS643

Bảng 3. 19. Tần suất phân bố các alen thuộc locus DYS643

DYS643						
ALEN	8	9	10	11	12	13
PPY23	0.015	0.075	0.190	0.290	0.350	0.08

Cùng với locus DYS549 đây là locus thứ 2 chỉ có mặt trong bộ kit PPY23 mà không có trong bộ kit YPlus. Có tất cả 6 alen được quan sát thấy với alen 12 có tần suất lớn nhất. Sự phân bố các alen trong locus này cũng có sự khá biệt đáng kể so với dữ liệu thống kê chung khi ở YHRD alen số 10 mới là alen có tần suất phân bố cao nhất, alen số 12 lại chỉ có tần suất <0.2.

3.2.2.18. Locus DYS393

Bảng 3. 20. Tần suất phân bố các alen thuộc locus DYS393

DYS393						
ALEN	10	11	12	13	14	15
PPY23			0.375	0.180	0.400	0.045
YPlus	0.005	0.025	0.310	0.165	0.465	0.03

Số lượng các alen được quan sát thấy trên locus DYS393 khá ít với chỉ khoảng 4 alen phổ biến và 2 alen hiếm số 10, 11. Có sự khác biệt về phân bố alen trên locus DYS393

giữa quần thể nam giới Việt Nam với dữ liệu chung trên YHRD. Trong nghiên cứu này alen số 14 có tần suất lớn nhất (khoảng 0.40-0.46), tiếp theo là alen số 12. Ngược lại trên YHRD alen số 13 mới là alen có tần suất lớn nhất, alen số 14 có tần suất < 0.2.

3.2.2.19. Locus *DYS458*

Bảng 3. 21. Tần suất phân bố các alen thuộc locus *DYS458*

DYS458						
ALEN	12	13	14	15	16	17
PPY23		0.005	0.025	0.120	0.170	0.195
YPlus	0.010	0.005	0.015	0.135	0.190	0.200
ALEN	18	19	20	21	22	23
PPY23	0.310	0.140	0.025		0.010	
YPlus	0.295	0.085	0.045	0.015		0.005

Đây được coi là locus chung của cả 2 bộ kit có số lượng alen nhiều nhất với tổng cộng 12 alen được quan sát thấy. Alen 12, 13, 22, 23 là các alen tương đối hiếm gặp. Alen 18 là alen có tần suất phân bố lớn nhất, tiếp theo là các alen 17, 16. Trên YHRD alen 17, 16 mới là các alen có tần suất phân bố lớn nhất và lớn hơn cả alen 18. Đây cũng là khác biệt của quần thể người Việt Nam so với dữ liệu chung.

3.2.2.20. Locus *DYS456*

Bảng 3. 22. Tần suất phân bố các alen thuộc locus *DYS456*

DYS456							
ALEN	12	13	14	15	16	17	18
PPY23		0.090	0.180	0.525	0.160	0.025	0.020
YPLUS	0.005	0.060	0.165	0.570	0.155	0.040	0.005

Với locus *DYS456* có tất cả 7 alen được phát hiện trong đó alen số 12 là alen tương đối hiếm. Alen 15 là alen chiếm ưu thế hơn hẳn so với các alen khác. Sự phân bố các alen thuộc locus *DYS456* trong quần thể người Việt Nam tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.21. Locus *YGATA-H4*

Bảng 3. 23. Tần suất phân bố các alen thuộc locus YGATA-H4

YGATAH4				
ALEN	10	11	12	13
PPY23	0.130	0.365	0.465	0.040
YPLUS	0.125	0.425	0.425	0.025

Locus YGATA-H4 có tất cả 4 alen được phát hiện trong đó alen 12 là alen có tần suất cao nhất tần suất ở cả 2 bộ kit (khoảng 0.425-0.465), tiếp theo là alen 11 với tần suất gần bằng. Sự phân bố các alen này tương đối giống với dữ liệu thống kê chung trên YHRD.

3.2.2.22. Locus *DYS460*

Bảng 3. 24. Tần suất phân bố các alen thuộc locus *DYS460*

DYS460					
ALEN	9	10	11	12	13
YPL	0.110	0.645	0.190	0.050	0.005

Đây là 1 trong số 6 locus mới được bổ sung vào bộ YPlus mà các bộ kit trước đây không có. Locus này có alen tương đối ít với chỉ 5 locus được quan sát thấy. Trong đó alen số 10 chiếm ưu thế lớn với tần suất 0.645, các locus khác có tần suất thấp hơn hẳn < 0.20. Điều này hơi khác biệt so với dữ liệu chung khi alen số 9 và 11 cũng có tần suất tương đối lớn và alen số 10 chiếm ưu thế với tần suất chỉ khoảng 0.40.

3.2.2.23. Locus *DYS627*

Bảng 3. 25. Tần suất phân bố các alen thuộc locus *DYS627*

DYS627										
ALEN	13	16	17	18	19	20	21	22	23	24
YPlus	0.005	0.005	0.005	0.065	0.080	0.230	0.230	0.235	0.100	0.045

Đây cũng locus mới xuất hiện trong bộ kit YPlus với 10 locus được quan sát thấy trong quần thể người Kinh, Việt Nam. Alen 13,16, 17 là các alen hiếm gặp. Các alen 20, 21, 22 là các alen chiếm tần suất cao nhất. Sự phân bố các alen này cũng tương đồng với dữ liệu chung trên YHRD.

3.2.2.24. Locus *DYS518*

Bảng 3. 26. Tần suất phân bố các alen thuộc locus DYS518

DYS518							
ALEN	32	33	34	35	36	37	38
YPLUS	0.005	0.005	0.015	0.060	0.120	0.120	0.095
ALEN	39	40	41	42	43	44	45
YPLUS	0.105	0.150	0.145	0.105	0.065	0.005	0.005

Với tất cả 14 alen được quan sát thấy trong 200 mẫu đây được coi là locus có số lượng alen nhiều nhất trong nghiên cứu. Điều đáng lưu ý đây cũng là alen mới xuất hiện trong bộ kit YPlus. Các alen 32, 33, 44, 45 tương đối hiếm gặp trong quần thể. Alen 40, 41 chiếm tần suất lớn nhất nhưng không có sự chênh lệch đáng kể so với các alen còn lại như 36, 37, 38. Sự phân bố này hơi khác biệt so với dữ liệu chung khi alen 37, 38, 39 mới là các alen có tần suất cao nhất.

3.2.2.25. Locus DYS449

Bảng 3. 27. Tần suất phân bố các alen thuộc locus DYS449

DYS449							
ALEN	23	25	26	27	28	29	30
YPlus	0.005	0.040	0.160	0.080	0.090	0.105	0.095
ALEN	31	32	33	34	35	36	
YPlus	0.145	0.125	0.05	0.04	0.04	0.025	

Locus DYS449 mới được bổ sung trong bộ kit YPlus bao gồm 13 locus được quan sát thấy trong 200 mẫu nghiên cứu. Trong đó alen 31, 32 lần lượt là các alen chiếm tần suất lớn nhất tương tự như trong dữ liệu công bố chung. Các alen còn lại 26, 27, 28, 29 có tần suất chênh lệch nhau không lớn.

3.2.2.26. Locus DYS385a/b

Với locus DYS385 có 16 alen được quan sát thấy trên 2 locus DYS385a và DYS385b trong đó alen 13 có tần suất cao nhất, tiếp đến là alen 18, alen 12. Sự phân bố này khác so với thống kê chung trên YHRD với alen 14, 13, 11 lần lượt là các alen có tần suất phân bố cao nhất

Bảng 3. 28. Tần suất phân bố các alen thuộc locus DYS385a/b

DYS385								
ALEN	9	10	11	12	12.3	13	14	15
PPY23	0.0025		0.0525	0.1175	0.0025	0.2675	0.0575	0.0500
YPLUS			0.0300	0.1100		0.3000	0.0475	0.0875
ALEN	16	17	18	19	20	21	22	23
PPY23	0.0400	0.0500	0.1475	0.1050	0.0725	0.0300	0.0050	
YPLUS	0.0450	0.0400	0.1275	0.1225	0.0675	0.0150	0.0050	0.0025

Tuy nhiên với DYS385 do cùng 1 cặp môi nhưng khuếch đại đồng thời 2 vị trí khác nhau thường được thống kê theo tập hợp 2 locus DYS385a và b như bảng 3.3. Kết quả thống kê cho thấy có 55 tổ hợp khác nhau được tạo ra bởi 2 locus DYS385a và b trong đó tổ hợp 13,18 chiếm tần suất nhiều nhất.

Bảng 3. 29. Tần suất phân bố các tổ hợp alen thuộc locus DYS385a/b

Haplo type	DYS385 PPY23	DYS385 YPlus	Haplo type	DYS385 PPY23	DYS385 YPlus	Haplo Type	DYS385 PPY23	DYS385 YPlus
9,17	0.005		12,19	0.035	0.035	14,19	0.020	0.025
11,11	0.015	0.005	12,20	0.040	0.055	14,20	0.010	
11,12		0.015	12,21	0.010		14,21	0.015	
11,13	0.005		13,13	0.065	0.080	15,15		0.010
11,14	0.005	0.005	13,14	0.015	0.015	15,16	0.005	0.010
11,15	0.005		13,15		0.015	15,17		0.010
11,17	0.005	0.005	13,16	0.005	0.015	15,18	0.030	0.005
11,18	0.035	0.020	13,17	0.050	0.045	15,19	0.020	0.060
11,19	0.010		13,18	0.150	0.185	15,20	0.015	0.025
11,20	0.010	0.005	13,19	0.120	0.105	15,21	0.015	0.015
12,12	0.025	0.010	13,20	0.040	0.035	15,22		0.005
12,12.3	0.005		13,21	0.005	0.010	15,23		0.005
12,13	0.005	0.010	13,22	0.010	0.005	16,17	0.005	
12,14	0.005	0.015	14,14		0.005	16,18	0.005	
12,15		0.005	14,15	0.010		16,19	0.005	0.015
12,16	0.015	0.015	14,16	0.005	0.015	16,20	0.020	0.015
12,17	0.020	0.020	14,17	0.005		16,21	0.015	0.005
12,18	0.050	0.030	14,18	0.025	0.010	17,20	0.010	

3.2.2.27. Locus *DYF387S1*

Xét trên cả 2 locus là *DYF387S1a* và *b* có tất cả 8 alen được quan sát thấy trong đó alen 37, 38 là các alen có tần suất cao nhất. Sự phân bố các alen trong locus này cũng tương đồng với dữ liệu thống kê chung trên YHRD. Theo tổ hợp 2 locus có tất cả 28 haplotype khác nhau trong đó tổ hợp 37,38 chiếm tần suất nhiều nhất.

Bảng 3. 30. Tần suất phân bố các alen thuộc locus *DYF387S1*

DYF387S1								
ALEN	34	35	36	37	38	39	40	41
YPlus	0.0175	0.0425	0.175	0.245	0.2725	0.155	0.0725	0.02

Bảng 3. 31. Tần suất phân bố các tổ hợp alen thuộc locus *DYF387S1*

Haplotype	Tần số	Haplotype	Tần số	Haplotype	Tần số
34,34	0.010	36,36	0.040	37,40	0.015
34,36	0.005	36,37	0.050	37,41	0.010
34,37	0.005	36,38	0.100	38,38	0.095
34,40	0.005	36,39	0.070	38,39	0.040
35,36	0.005	36,40	0.025	38,40	0.030
35,37	0.010	36,41	0.015	39,39	0.040
35,38	0.045	37,37	0.090	39,40	0.020
35,39	0.015	37,38	0.140	39,41	0.005
35,40	0.010	37,39	0.080	40,40	0.020
				40,41	0.005

Nhận xét: Như vậy từ việc lập hồ sơ Y-STR của 400 mẫu nghiên cứu với 2 bộ kit PPY23 và YPlus chúng tôi đã lập được bảng phân bố tần suất alen của 29 locus Y-STR. Kết quả so sánh cho thấy có 16/29 locus có sự phân bố tần suất alen tương đối giống với dữ liệu chung trên YHRD. Ngược lại có 13/29 locus có sự phân bố khác so với dữ liệu chung. Ngoài ra thống kê cũng cho thấy có 41 alen hiếm/ 29 locus. Điều này cho thấy sự sai khác trong cấu trúc di truyền của quần thể người Kinh, Việt Nam khi so sánh với các quần thể khác. Bảng phân bố tần suất tổng hợp của 29 locus Y-STR được thiết lập trong

bảng 3.2 là bảng dữ liệu hoàn chỉnh nhất phản ánh sự phân bố các alen thuộc Y-STR trong quần thể người Kinh, Việt Nam.

Việc lập bảng tần suất phân bố alen có ý nghĩa rất quan trọng trong giám định ADN giúp chúng ta nhận biết các alen phổ biến, alen hiếm trong từng locus, tính toán chỉ số có quan hệ họ hàng (kinship index) giữa các hồ sơ ADN nhằm xác định huyết thống, định danh cá thể mất tích trong thiên tai thảm họa, truy nguyên cá thể đặc biệt là trong điều tra hình sự. Với locus trên NST Y do đặc điểm bảo thủ là chỉ di truyền theo dòng cha, không có tái tổ hợp qua các thế hệ nên khi so sánh hồ sơ ADN Y-STR giữa Q (mẫu hiện trường) và K (mẫu nghi phạm đã biết) sẽ có 03 trường hợp xảy ra: (1) loại trừ: nếu hồ sơ ADN từ Q và K là khác nhau, (2) không kết luận được: khi thu được dữ liệu ADN không đầy đủ hoặc không rõ ràng, (3): không loại trừ được khi hồ sơ ADN giữa Q và K là trùng khớp. Trường hợp (3) rất phức tạp vì theo quy luật di truyền và loại trừ đột biến xảy ra tất cả những cá thể họ hàng nam của nghi phạm (ông nội, chú, bố, anh - em trai nghi phạm) đều có chung hồ sơ Y-STR. Khi đó việc ước lượng xác suất trùng khớp hồ sơ Y-STR trong quần thể là điều cần thiết để đưa ra kết luận cuối cùng cho vụ án.

Số lượng cá thể không có quan hệ họ hàng trong cơ sở dữ liệu khảo sát càng lớn (nghĩa là mẫu số trong tính toán phương pháp đếm), số liệu thống kê cho ước tính tần số khớp ngẫu nhiên sẽ càng có độ chính xác cao. Việc khảo sát tần suất haplotype Y-STR và tần suất từng locus trong quần thể sẽ giúp ích cho việc tính toán tỉ suất tương đồng (LR) và đưa ra được xác suất giữa 2 hồ sơ có bao nhiêu % xác suất có quan hệ huyết thống theo dòng cha. Từ đó giúp ích đáng kể cho việc đưa ra kết luận cho vụ việc. Sở dĩ phải sử dụng phương pháp tính LR vì để trả lời câu hỏi rằng: Liệu có một người nào khác không phải là thủ phạm đã để lại dấu vết ở hiện trường? Nếu kiểu gen của thủ phạm và dấu vết tại hiện trường hoàn toàn trùng khớp nhau, khi đó LR sẽ được tính xem khả năng thủ phạm để lại dấu vết gấp bao nhiêu lần so với dân số quốc gia. Với tầm quan trọng đó, mỗi trung tâm, phòng thí nghiệm được khuyến cáo nên xây dựng bảng tần suất phân bố các alen cho các chỉ thị Y-STR đang sử dụng khi áp dụng vào việc đưa ra kết luận trong thực tế để đảm bảo độ chính xác cao nhất [76, 77].

Dựa vào bảng tần suất phân bố alen của 29 chỉ thị Y-STR và sử dụng công cụ Kinship Analysis trên YHRD.com chúng tôi đã tính toán được được tỷ số tương đồng (likelihood

ratio – LR) và xác suất có quan hệ theo dòng cha từ các hồ sơ Y-STR. Ví dụ minh họa: trong vụ án hiếp dâm khi phân tích hồ Y-STR sử dụng bộ PPY23 thu từ mẫu nghi ngờ tình dục trên ga trải giường (ký hiệu V218A) cho thấy trùng khớp với nghi phạm Nguyễn Trung K (ký hiệu V218B). Sử dụng bảng tần suất phân bố từ nghiên cứu trên chúng tôi tính được chỉ số LR là 23.300 với độ tin cậy là 99.9978 %. Chỉ số LR này còn được hiểu là khả năng bắt gặp hồ sơ Y-STR thu từ mẫu hiện trường có quan hệ huyết thống theo dòng cha với mẫu thu từ nghi phạm cao hơn 23300 lần so với khả năng hai mẫu không liên quan.

Bảng 3.32: Hồ sơ Y-STR của 2 mẫu so sánh từ nghi phạm và dấu vết hiện trường

Ký hiệu	DYS 576	DYS 481	DYS 549	DYS 533	DYS 570	DYS 456	DYS 389I	DYS 390	DYS 389II	DYS 458	DYS 19
V218A	17	22	12	10	16	16	13	24	30	18	16
V218B	17	22	12	10	16	16	13	24	30	18	16
Ký hiệu	DYS 393	DYS 391	DYS 439	DYS 635	DYS 392	GAT AH4	DYS 437	DYS 643	DYS 438	DYS 448	DYS 385
V218A	14	10	12	22	13	11	14	12	10	18	13, 18
V218B	14	10	12	22	13	11	14	12	10	18	13, 18

Ngoài ra, bảng tần suất còn hỗ trợ trong việc tính toán xác suất có quan hệ huyết thống theo dòng cha trong trường hợp có đột biến xảy ra. Ví dụ minh họa trong trường hợp dưới đây: hai mẫu ký hiệu là CH01 và CT01 nghi ngờ có quan hệ chú ruột – cháu trai. Kết quả phân tích hồ sơ Y-STR với bộ PPY23 cho thấy sai khác tại một locus DYS393 (chênh lệch nhau 1 alen). Với giả thiết xảy ra đột biến qua 1 thế hệ tỷ suất tương đồng LR giữa mối quan hệ huyết thống so với mối quan hệ không cùng huyết thống là 0.1667. Với tỷ lệ LR có thể kết luận giữa hai mẫu không có quan hệ huyết thống theo dòng ch

Bảng 3.33: Hồ sơ Y-STR của 2 mẫu so sánh từ chú và cháu trai

Ký hiệu	DYS 576	DYS 481	DYS 549	DYS 533	DYS 570	DYS 456	DYS 389I	DYS 390	DYS 389II	DYS 458	DYS 19
CH01	17	27	13	11	17	13	12	25	28	18	16
CT01	17	27	13	11	18	13	12	25	28	18	16
Ký hiệu	DYS 393	DYS 391	DYS 439	DYS 635	DYS 392	GAT AH4	DYS 437	DYS 643	DYS 438	DYS 448	DYS 385
CH01	12	10	12	20	13	12	14	11	10	19	12, 20
CT01	13	10	12	20	13	12	14	11	10	19	12, 20

Tuy nhiên cũng với hai hồ sơ Y-STR nêu trên giả thiết xảy ra đột biến tại locus DYS570 thay vì DYS393 sẽ cho kết quả LR là 108. Điều đó có nghĩa là hai mẫu có xác

suất có quan hệ huyết thống theo dòng cha lên tới 99.539%. Điều này có thể được hiểu là do tần số đột biến ở các locus khác nhau là khác nhau. Locus DYS393 có tần suất đột biến rất thấp là 0.00122 trong khi DYS570 lại được coi là locus có tốc độ đột biến nhanh với tần suất là 0.00922 nên xác suất xảy ra sai khác giữa hai cá thể cùng dòng cha là cao hơn so với các locus khác (tần suất đột biến của các locus được thống kê trên YHRD.org).

Bảng tần suất phân bố Y-STR được xây dựng trong nghiên cứu này là bảng có số lượng mẫu nghiên cứu cùng như số locus nhiều nhất được khảo sát trong quần thể người Việt Nam cho đến nay với số lượng mẫu và phương pháp nghiên cứu đạt độ tin cậy cao. Các công bố trước đây đối với quần thể người Việt Nam chủ yếu được thực hiện với 13-17 Y-STR trên số lượng mẫu tương đối nhỏ dưới 200 mẫu. Việc xây dựng bảng tần suất của những locus Y-STR cũng mới là điều cần thiết, đáp ứng nhu cầu tính toán chỉ số hình sự liên quan đến nam giới trong quần thể người Việt Nam.

Trên dữ liệu chung YHRD có thể thấy số lượng haplotype được công bố sử dụng 2 bộ kit PPY23 và YPlus từ các nghiên cứu với các quần thể trên toàn thế giới là khá lớn. Ngược lại với quần thể người Việt Nam trước nghiên cứu này số lượng haplotype được công bố còn rất hạn chế (Bảng 3.32) chỉ với 45 haplotype từ bộ PPY23 và 47 haplotype với bộ YPlus được đóng góp thông một số nghiên cứu nhỏ lẻ trên quần thể người Việt Nam.

(https://yhrd.org/pages/resources/national_databases).

Bảng 3. 34. Thống kê số lượng Haplotype Y-STR ở một số nước trên thế giới được công bố trên YHRD

Tên nước	Bộ PPY23	Bộ YPlus
Ấn Độ	1950 Haplotypes	558 Haplotypes
Trung quốc	28502 Haplotypes	51515 Haplotypes
Nhật Bản	323 Haplotypes	1299 Haplotypes
Hàn Quốc	1309 Haplotypes	1040 Haplotypes
Thái Lan	497 Haplotypes	140 Haplotypes
Đức	2821 Haplotypes	522 Haplotypes
Nigeria	81 Haplotypes	358 Haplotypes
Thụy Sĩ	1349 Haplotypes	756 Haplotypes
Anh	3798 Haplotypes	144 Haplotypes

Tên nước	Bộ PPY23	Bộ YPlus
Mỹ	5717 Haplotypes	2122 Haplotypes
Việt Nam	45 Haplotypes	47 Haplotypes

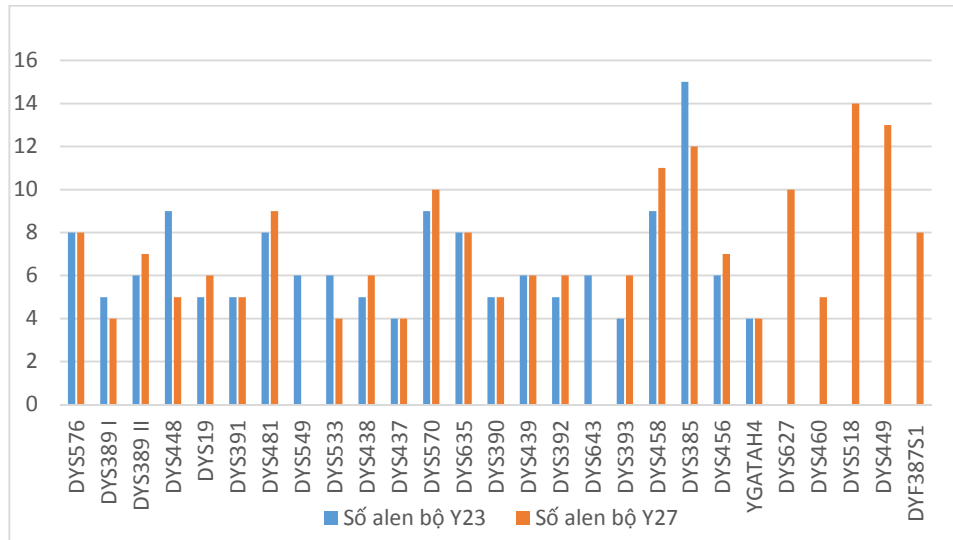
Với nghiên cứu này chúng tôi đã đóng góp thêm 200 hapotype sử dụng bộ PPY23 vào dữ liệu chung YHRD với mã số truy cập là YA004576 nâng tổng số hapotype trên dữ liệu YHRD lên 245 với bộ PPY23 (Hình 3.3).

Population samples						
Population Sample	Minimal	PowerPlex Y	Yfiler	PowerPlex Y23	Yfiler Plus	Maximal
Hanoi, Vietnam [Vietnamese]	48 Haplotypes	48 Haplotypes	48 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Vietnamese]	209 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Ho-Chi-Minh City, Vietnam [Kinh]	45 Haplotypes	45 Haplotypes	45 Haplotypes	45 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Kinh]	123 Haplotypes	123 Haplotypes	123 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Mong]	59 Haplotypes	59 Haplotypes	59 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Tay]	21 Haplotypes	21 Haplotypes	21 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Nuong]	7 Haplotypes	7 Haplotypes	7 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
Vietnam [Hoi]	7 Haplotypes	7 Haplotypes	7 Haplotypes	0 Haplotypes	0 Haplotypes	0 Haplotypes
North Vietnam, Vietnam [Kinh]	200 Haplotypes	200 Haplotypes	200 Haplotypes	200 Haplotypes	0 Haplotypes	0 Haplotypes
Lao Cai, Vietnam [Bouyei]	29 Haplotypes	29 Haplotypes	29 Haplotypes	0 Haplotypes	29 Haplotypes	0 Haplotypes
Ha Giang, Vietnam [Bouyei]	18 Haplotypes	18 Haplotypes	18 Haplotypes	0 Haplotypes	18 Haplotypes	0 Haplotypes
(Overall)	766 Haplotypes	557 Haplotypes	557 Haplotypes	245 Haplotypes	47 Haplotypes	0 Haplotypes

Hình 3.3. Số liệu thống kê số haplotype đóng góp từ quần thể người Việt Nam trên YHRD

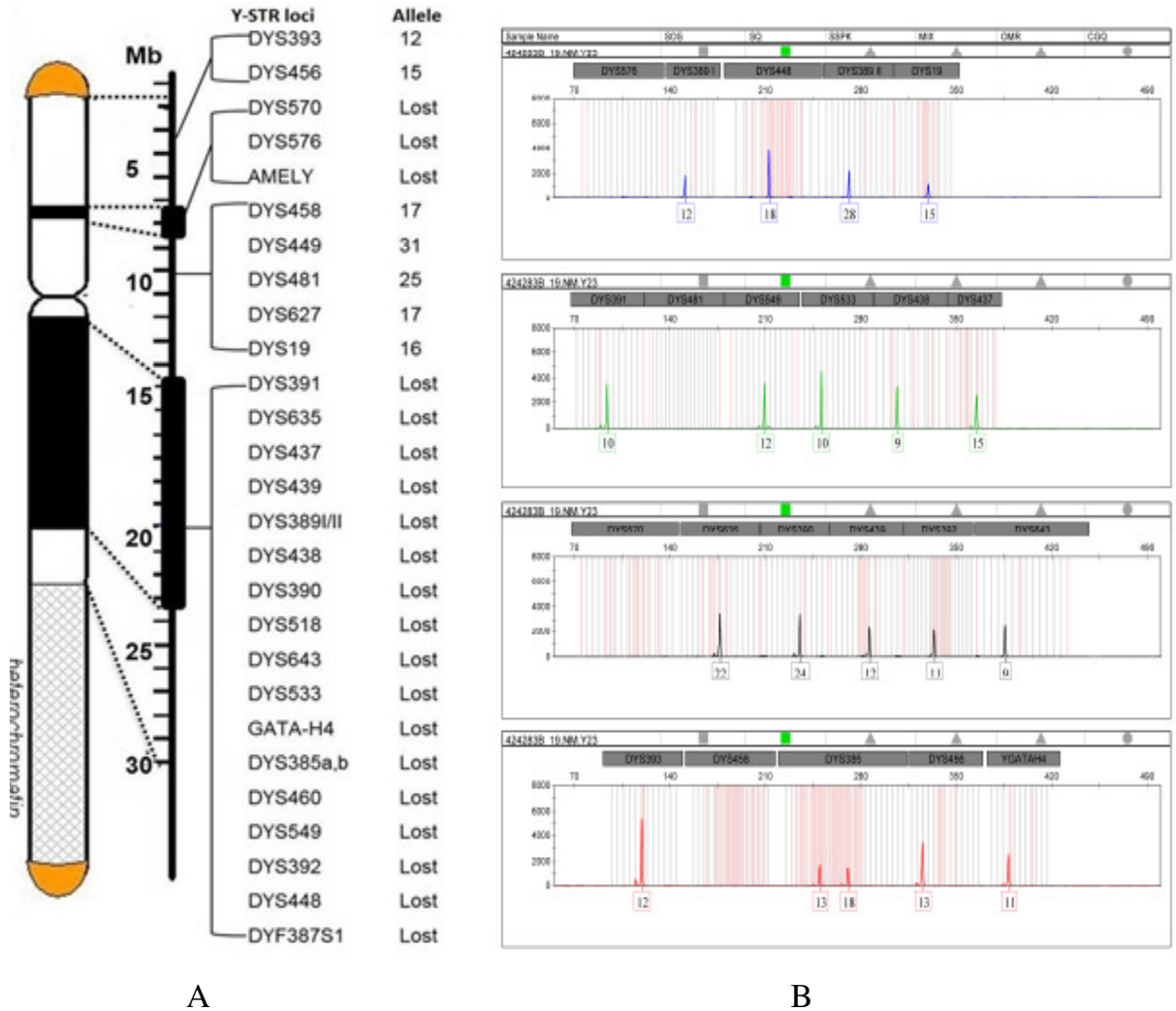
3.2.3. *Đánh giá đặc điểm các alen thuộc 29 chỉ thị Y-STR*

Với 29 chỉ thị Y-STR được khảo sát có tất cả 209 alen được quan sát thấy. Số lượng alen dao động từ 4 đến 15 alen với trung bình là 7 alen / 1 locus. Trong đó với bộ kit PPY23 có tổng số 144 alen, bộ kit YPlus là 183 alen. Với riêng tổ hợp gồm 2 alen, DYS385ab có 55 tổ hợp alen khác nhau, DYS387S1ab có 28 tổ hợp alen khác nhau được quan sát thấy. Với cả hai bộ kit nói chung, DYS458, DYS570 là các locus có số lượng alen nhiều nhất với khoảng 10 alen cho mỗi locus. Ngược lại locus DYS437 tiếp theo là DYS438, DYS393, DYS390 có số alen ít nhất (4-5 alen/locus). Riêng bộ kit YPlus có DYS518, DYS449 là các locus mới được bổ sung thêm so với bộ PPY23 đồng thời cũng là các locus có số lượng alen nhiều nhất (DYS518: 14 alen, DYS449: 13 alen).



Hình 3. 4. Số alen của 29 locus Y-STR dựa trên 2 bộ kit PPY23 và YPlus

Tổng số alen quan sát thấy trong nghiên cứu là 209 alen, nhiều hơn đáng kể so với các công bố trước đây với quần thể người Việt Nam. Trong nghiên cứu khảo sát trên 205 nam giới sử dụng bộ kit AmpFLSTR™ Yfiler™ PCR Amplification Kit gồm 17 locus, tác giả Toàn T.T ghi nhận 99 alen [78]. Số alen trong nghiên cứu cũng ở mức tương đương khi so sánh với quần thể khác. Ví dụ như trong nghiên cứu khảo sát trên 581 nam giới dân tộc Hán thuộc tỉnh Thiểm Tây – Trung Quốc sử dụng bộ kit PPY23 quan sát thấy 237 alen [79]. Nghiên cứu khảo sát trên 436 mẫu người nam giới thuộc các tiểu Vương quốc Ả rập Thống nhất (UAE) sử dụng bộ YPlus cho thấy có 202 alen khác nhau [80].



Hình 3. 5. Các trường hợp mất locus trong hồ sơ Y-STR (A) vị trí mất 18/29 locus Y-STR trên bản đồ NST Y (B) hồ sơ Y-STR của trường hợp mất 4/23 locus Y-STR.

Những trường hợp alen bị lệch thang, mất alen (null alen), alen biến thể đòi hỏi người phân tích dữ liệu phải có kinh nghiệm để tránh nhầm lẫn giữa các tình huống. Ví dụ như biến thể alen có thể bị nhầm với trường hợp lệch thang. Để chắc chắn nên kiểm chứng lại bằng phương pháp giải trình tự với những locus xuất hiện biến thể. Trong nghiên cứu này trên tất cả locus đều không phát hiện thấy trường hợp nằm ngoài thang alen (off ladder), phát hiện thấy 02 biến thể alen tại các locus DYS448 (alen 18.2, 19.2), 1 biến thể tại locus DYS385b (alen 12.3).

Với trường hợp mất alen trong một hoặc nhiều locus cần phân biệt là do thực sự bị mất alen hay không lên đủ alen do mẫu có chất lượng ADN kém. Cụ thể, trong quá trình khảo sát 400 mẫu nghiên cứu chúng tôi đã phát hiện hai trường hợp mất alen tại một số

locus. Với trường hợp đầu tiên là mẫu nam giới phát hiện mất 18/23 locus Y-STR khi phân tích với bộ kit PPY23. Phân tích xác nhận lại bằng bộ kit YPlus cũng cho kết quả tương tự. Dựa vào việc xác định vị trí các locus bị mất trên bản đồ NST Y như trong hình 3.5A nhận thấy đây là trường hợp mất đoạn lớn trên NST Y. Việc phân tích thêm hồ sơ các STR trên NST X nhận thấy có 1 số locus ở dạng dị hợp tử. Điều này dẫn đến giả thiết có 2 NST X và mẫu nam giới này có thể cặp NST giới tính ở dạng XXY (hội chứng Klinefelter). Trong quá trình nghiên cứu đề tài dạng đột biến này khá hiếm gặp và chỉ có 1 trường hợp tương tự được công bố trên thế giới [81]. Do hạn chế của đề tài chúng tôi chưa có điều kiện làm nhiễm sắc thể đồ của mẫu nghiên cứu để chứng minh giả thiết nêu trên.

Trường hợp còn lại là có null alen tại 4 locus DYS481, DYS576, DYS570, DYS458 (hình 3.5B). Việc mất alen của 4 locus nằm kế tiếp nhau trên cánh ngắn của NST Y cùng với mất đoạn AMEL Y khi phân tích bằng bộ kit PowerPlex® Fusion System (hãng Promega) gợi ý cho thấy nguyên nhân là do mất đoạn trên cánh ngắn NST Y chứ không phải do đột biến tại vị trí gắn môi của các locus Y-STR. Để đảm bảo việc tính toán tần suất alen được chính xác chúng tôi không đưa 2 trường hợp này vào bảng dữ liệu chung.

3.2.4. Độ đa hình của 29 chỉ thị Y-STR

Từ tần suất phân bố các alen chúng tôi đã tính toán được độ đa hình hay là độ đa dạng gen (GD - gene diversity) của từng locus Y-STR trong nghiên cứu. Bảng so sánh độ lệch chuẩn giữa 2 bộ kit cho thấy không có sự sai khác nhau nhiều về giá trị GD giữa 2 bộ kit PPY23 và YPlus (độ lệch chuẩn đều < 0.05) do đó giá trị GD cuối cùng sẽ được tính là trung bình của 2 bộ kit (Bảng 3.35).

Bảng 3. 35. Độ đa dạng gen của 29 locus Y-STR trong nghiên cứu và so sánh với giá trị trung bình của quần thể người châu Á

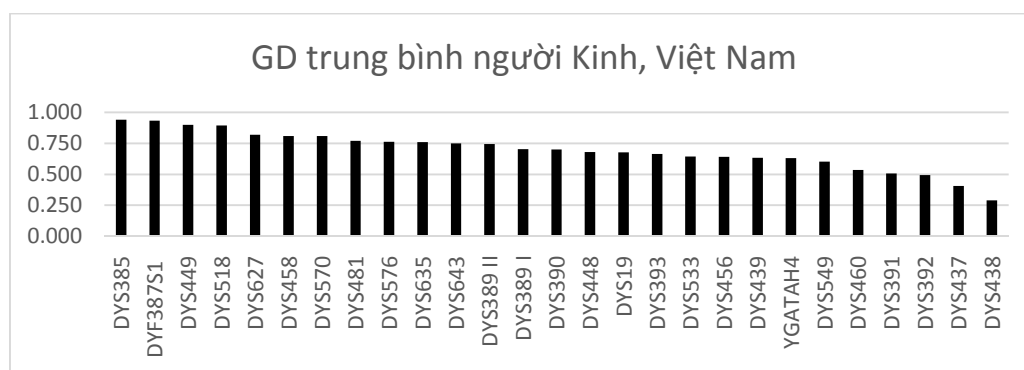
STT	Tên alen	PPY23 GD	YPlus GD	Độ lệch chuẩn		GD trung bình	GD châu Á	Độ lệch chuẩn
1	DYS576	0.776	0.748	0.01		0.762	0.7996	0.02
2	DYS389 I	0.713	0.693	0.01		0.703	0.6564	0.02
3	DYS389 II	0.734	0.751	0.01		0.742	0.6585	0.04
4	DYS448	0.713	0.647	0.03		0.680	0.7650	0.04
5	DYS19	0.685	0.669	0.01		0.677	0.6920	0.01

STT	Tên alen	PPY23 GD	YPlus GD	Độ lệch chuẩn		GD trung bình	GD châu Á	Độ lệch chuẩn
6	DYS391	0.499	0.514	0.01		0.507	0.4136	0.05
7	DYS481	0.771	0.768	0.00		0.769	0.8416	0.04
8	DYS549	0.603				0.603	0.6425	0.02
9	DYS533	0.653	0.635	0.01		0.644	0.6265	0.01
10	DYS438	0.280	0.300	0.01		0.290	0.5707	0.14
11	DYS437	0.404	0.406	0.00		0.405	0.4891	0.04
12	DYS570	0.795	0.821	0.01		0.808	0.8313	0.01
13	DYS635	0.774	0.745	0.01		0.760	0.7700	0.01
14	DYS390	0.712	0.687	0.01		0.700	0.7325	0.02
15	DYS439	0.630	0.638	0.00		0.634	0.6823	0.02
16	DYS392	0.522	0.467	0.03		0.495	0.7434	0.12
17	DYS643	0.749				0.749	0.7510	0.00
18	DYS393	0.668	0.662	0.00		0.665	0.6597	0.00
19	DYS458	0.806	0.813	0.00		0.809	0.8275	0.01
20	DYS385	0.947	0.936	0.01		0.941	0.9741	0.02
21	DYS456	0.661	0.622	0.02		0.641	0.6093	0.02
22	YGATAH4	0.635	0.626	0.00		0.630	0.6345	0.00
23	DYS627		0.820			0.820	0.8120	0.00
24	DYS460		0.536			0.536	0.6750	0.07
25	DYS518		0.893			0.893	0.8670	0.01
26	DYS449		0.900			0.900	0.8820	0.01
27	DYF387S1		0.934			0.934	0.9450	0.01

Kết quả cho thấy hai locus với tổ hợp 2 alen là *DYS385a/b* và *DYS387S1* là các locus có độ đa dạng di truyền cao nhất (GD lần lượt 0.941, 0.934). Với các locus đơn alen có chung ở hai bộ kit *DYS458*, *DYS570* được coi là các locus có độ đa hình cao nhất (GD trung bình lần lượt là 0.809, 0.808) đồng thời cũng là các locus có số alen nhiều nhất (9 - 10 alen). Ngược locus *DYS438* là locus có độ đa hình thấp nhất (GD khoảng 0.28) với chỉ 4 - 5 alen phát hiện được trong quần thể người Việt Nam, tiếp theo là *DYS437* (GD: 0.490),

DYS392 (GD: 0.495), DYS391 (GD:0.507). Đây được coi là các locus có độ đa hình thấp nhất nhỏ hơn 0.5.

Các nghiên cứu trước đây về quần thể người Kinh Việt Nam cũng cho thấy DYS437, DYS438 cũng là các locus có giá trị GD thấp nhất [63,66,78]. Kết quả này cũng khá tương đồng với phân tích toàn cầu về 23 locus Y-STR cho thấy các locus DYS391, DYS437 và DYS438 có giá trị GD thấp hơn so với các locus Y-STR khác trong các quần thể người châu Á dựa theo nghiên cứu khảo sát tần suất 23 Y-STR từ 129 quần thể thuộc 51 quốc gia của Josephine Purps và cs [82]. Kết quả so sánh độ đa hình 23 locus Y-STR với một số nước khác trong khu vực cũng cho thấy sự tương đồng. Cụ thể trong nghiên cứu trên quần thể người Philippine cho thấy DYS391 và DYS438 có độ đa hình thấp nhất (dưới 34%), 2 markers DYS385 và DYS387S1 có độ đa hình cao nhất [83]. Trong khảo sát 29 locus Y-STR với quần thể người Thượng Hải (Trung Quốc) cho thấy DYS449 là locus có tần số cao nhất (GD: 0.8696), ngược lại DYS438 có tần suất thấp nhất (GD: 0.4186) tiếp theo là DYS437 [84]. Độ đa hình của 6 locus mới trong bộ YPlus được so sánh với thống kê trên quần thể người châu Á tại Mỹ do nhà sản xuất cung cấp cũng cho thấy tương quan trong giá trị [69].



Hình 3. 6. Xếp hạng độ đa hình các chỉ thị Y-STR theo giá trị từ cao đến thấp

Tuy nhiên độ đa hình cũng có sự chênh lệch đáng kể so với dữ liệu chung ở 4 locus là DYS392, DYS391, DYS438 và DYS460. Trong đó locus DYS392, DYS438, DYS460 trong quần thể người Việt Nam có độ đa hình thấp hơn khá nhiều so với các nước cùng khu vực, ngược lại DYS391 lại có độ đa hình cao hơn. Những khác biệt này phản ánh đặc trưng và sự phân tách vốn gen trong quần thể người Kinh, Việt Nam so với các quần thể khác.

Kết quả sắp xếp theo độ đa hình của 29 locus Y-STR trong hình 3.6 cho thấy số locus có độ đa hình từ 0.7-1 chiếm tỷ lệ nhiều nhất với 55.2% (bao gồm 16/29 locus với 5 locus GD > 0.9, 4 locus GD nằm trong khoảng 0.8- 0.9, 7 locus GD trong khoảng 0.7-0.8), 9 locus có GD trong khoảng 0.5 – 0.6 chiếm 31% và chỉ có 4 locus có GD \leq 0.5. Qua đó cho thấy hầu hết các locus trong nghiên cứu đều có độ đa hình khá cao, có tiềm năng lớn trong việc phân biệt cá thể nam giới trong quần thể người Kinh, Việt Nam được khảo sát.

3.2.5. Độ đa dạng haplotype (HD) và khả năng phân biệt

Với riêng từng bộ kit khi phân tích 200 mẫu nghiên cứu chúng tôi quan sát thấy 200 haplotype khác nhau, nghĩa là mỗi haplotype đều là duy nhất trong quần thể khảo sát. Do đó độ đa dạng haplotype (haplotype diversity - HD) là 1.00. Khả năng phân biệt (discriminating capacity - DC) được tính là tỷ số của số haplotype khác nhau được quan sát thấy trong quần thể/ tổng số mẫu nghiên cứu. Do các haplotype đều là duy nhất nên khả năng phân biệt cá thể là 100% với cả 2 bộ kit PPY23 và YPlus. Điều này phản ánh tiềm năng cao của 29 locus Y-STR trong việc phân biệt các cá thể nam không có liên quan huyết thống theo dòng cha. Kết quả này cho thấy tiềm năng ứng dụng lớn của tập hợp Y-STR trong phân biệt cá thể nam thuộc quần thể người Việt Nam. Kết quả này cũng phù hợp với nghiên cứu của Toan T.T khi khảo sát trên quần thể người Kinh, Việt Nam với tổ hợp 17 Y-STR trong bộ YPlus cũng cho khả năng phân biệt 100% [66,78]. Cũng trên quần thể người Việt Nam trong nghiên cứu của Koji Dewa khảo sát trên 119 mẫu với tổ hợp 13 Y-STR chỉ cho khả năng phân biệt chỉ 95% [64]. Điều này cho thấy việc tăng số lượng Y-STR phân tích sẽ góp phần làm tăng độ đa dạng haplotype và khả năng phân biệt.

Nghiên cứu khảo sát với bộ kit PPY23 trên 129 quần thể khác nhau trên thế giới cũng cho thấy các quần thể châu Á có khả năng phân biệt cao nhất (> 97%), tiếp theo là Châu Âu và Châu Mỹ Latinh (DC khoảng 96%) và cuối cùng là Châu Phi (DC khoảng 85%) [82].

Tuy nhiên khi so sánh với một số quần thể khác trong cùng khu vực giá trị HD lại không đạt được tuyệt đối. Khảo sát trên 255 nam giới người Hán ở Bắc Kinh sử dụng bộ PPY23 cho thấy HD là 0.99996 với khả năng phân biệt là hơn 99.6% [85]. Khảo sát trên 142 nam giới thuộc các dân tộc khác nhau người Nigeria cho thấy HD là 0.9998 (± 0.0009) với DC là 98.5% [86]. Điều này có thể giải thích do số lượng mẫu khảo sát trong các công

bố lớn hơn nên xác suất bắt gặp 2 haplotype trùng nhau cũng cao hơn hoặc do đặc điểm di truyền khác nhau giữa các quần thể.

3.3. Kết quả nghiên cứu một số chỉ thị mini Y-STR mới

Do yêu cầu trong các bộ kit thương mại phải khuếch đại đồng thời nhiều locus Y-STR khác nhau nên để phân biệt được trong quá trình điện di mao quản đòi hỏi các locus phải có kích thước khác nhau. Kết quả là một số locus Y-STR có kích thước khá lớn (>200 bp) rất khó để phân tích thành công với các mẫu có ADN đã bị phân hủy, ADN bị đứt gãy nhiều. Trên thế giới nhận thấy tầm quan trọng của STR nói chung trong việc phân tích các mẫu đã bị phân hủy nhiều phòng thí nghiệm đã nghiên cứu và phát triển các bộ kit gồm các mini STR như bộ kit AGCU Mini-STR Amplification Kit (AGCU ScienTech Incorporation, China), AmpFLSTR Minifiler™ kit (Applied Biosystem) gồm 8 STR trong nhân và chỉ thị Amelogenin xác định giới tính [87-89]. Tuy nhiên đây hầu hết là các STR trong nhân, đến nay vẫn chưa có bộ kit riêng nào về các mini Y-STR được thương mại hóa. Các mini Y-STR với đặc điểm là có kích thước ngắn dưới 200 bp phù hợp với các loại mẫu sinh học đã phân hủy có ADN bị đứt gãy nhiều.

Hiện nay, tại Việt Nam số lượng các công trình nghiên cứu về locus Y-STR còn rất hạn chế và cũng chưa có nghiên cứu về thiết kế các cặp mồi khuếch đại locus mini Y-STR nhằm ứng dụng phân tích ADN cho các mẫu sinh học đã phân hủy. Vì vậy, để tạo điều kiện thuận lợi cho công tác giám định ADN cũng như giám định hài cốt tại Viện Pháp y quốc gia đạt được kết quả tốt bên cạnh 29 locus Y-STR có trong 2 bộ kit nêu trên chúng tôi thử nghiên cứu các chỉ thị mini Y-STR có kích thước ngắn hơn và có tiềm năng ứng dụng trong giám định ADN với quần thể người Việt Nam.

3.3.1. Lựa chọn chỉ thị mini Y-STR

Trong nghiên cứu này chúng tôi lựa chọn các locus có độ đa hình cao trong quần thể người Việt Nam, ngoài ra còn có kích thước lớn trong các bộ kit thương mại và cố gắng giảm tối thiểu kích thước các locus Y-STR để đạt được tỷ lệ phân tích thành công cao với các mẫu đã bị phân hủy. Trong số 29 chỉ thị Y-STR thuộc hai bộ kit chúng tôi lựa chọn các chỉ thị có độ đa hình cao > 0.6, có kích thước > 200 bp và có khả năng rút gọn kích thước vùng lõi lặp (vùng lõi lặp ở dạng đơn giản gồm 3 – 5 nucleotit, số lần lặp của lõi < 30). Dựa

trên các tiêu chí đó chúng tôi đã lựa chọn 5 chỉ thị Y-STR là DYS481, DYS643, DYS533, DYS19, Y-GATA-H4. Đây đều là các chỉ thị nằm trong bộ kit PPY23.

Bên cạnh đó chúng tôi cũng thiết kế môi để khuếch đại thêm 1 chỉ thị DYS460 nằm trong bộ YPlus và 4 chỉ thị mini Y-STR là: DYS505, DYS522, DYS508, DYS388 là những locus chưa có trong kit thương mại nhưng theo một số nghiên cứu đã được công bố cho thấy có độ đa hình cao, có tiềm năng ứng dụng trong phân tích Y-STR..

Bảng 3. 36. Độ đa dạng của các mini Y-STR mới trong một số quần thể

Độ đa dạng gen	DYS508	DYS522	DYS388	DYS505
Bắc Bò Đào Nha [90]	0.497	0.63		
Nhật Bản [91]	0.406	0.58		
Trung Quốc [92]		0.723	0.495	
Tunisia [93]			0.59	
Thái Lan [94]		0.654		0.6692

3.3.2. Tối ưu hoá phản ứng PCR

Phản ứng PCR được khảo sát với 10 cặp môi trong 10 phản ứng PCR riêng rẽ. Nhiệt độ gắn môi của từng locus dao động từ 56 – 58°C. Chu trình phản ứng PCR được tối ưu hóa với các điều kiện bao gồm: nồng độ môi, nhiệt độ gắn môi, thời gian kéo dài chuỗi. Sản phẩm sau PCR được kiểm tra trên gel Polyacrylamide 6%. Hình ảnh điện di cho thấy các băng sản phẩm PCR rõ nét, không có băng phụ và có kích thước đúng như các công bố trước đó. Kích thước các locus dao động từ 91-213bp. Điều này cho thấy các cặp môi được thiết kế có tính khả dụng cho việc khuếch đại các locus Y-STR mong muốn.

Phản ứng PCR được khảo sát trong 10 phản ứng PCR riêng rẽ. Chu trình phản ứng PCR được tối ưu hóa với các điều kiện bao gồm: nồng độ môi, nhiệt độ gắn môi, thời gian kéo dài chuỗi. Sản phẩm sau PCR được kiểm tra trên gel Polyacrylamide 6%. Cụ thể:

Nhiệt độ gắn môi được lựa chọn tối ưu trong khoảng nhiệt độ từ 55 - 66°C sử dụng công nghệ gradient nhiệt 2 chiều trên máy PCR effendorf mới giúp tối ưu hoá phản ứng PCR được nhanh chóng. Kết quả cho thấy 8/10 locus lên band rõ nét ở nhiệt độ gắn môi 58°C. Riêng locus DYS19 và DYS522 nhiệt độ gắn môi được hạ xuống khoảng 56°C để cho kết quả band điện di rõ nét và không có band phụ.

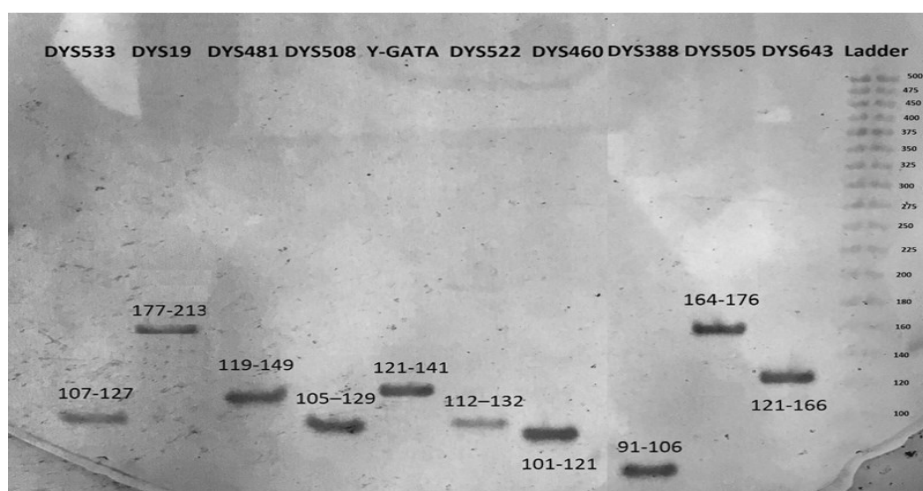
Nồng độ môi xuôi và môi ngược cho mỗi phản ứng PCR được tối ưu ở 4 điều kiện: 5, 10, 15, 20 ppmol. Kết quả thống kê trong Bảng 3.37 cho thấy với tất cả các nồng độ môi đều cho kết quả band lên rõ nhưng ở nồng độ môi 15-20 ppmol có hiện tượng dư môi sau

PCR. Nồng độ môi dư có thể ảnh hưởng đến kết quả giải trình tự gen. Hơn thế quá nhiều môi sẽ làm tăng khả năng liên kết không đặc hiệu hoặc thậm chí bắt cặp lẫn nhau trong phản ứng multiplex PCR ở bước tiếp theo. Với nồng độ 5 ppmol băng ADN lên không được rõ khi so sánh với nồng độ 10 ppmol. Thêm vào đó do các băng có kích thước khá gần nhau nên để đảm bảo hiệu quả phân biệt trong quá trình điện di chúng tôi lựa chọn nồng độ môi tối ưu cho phản ứng khuếch đại 10 locus gen là 10 ppmol.

Bảng 3. 37. Kết quả tối ưu hóa nồng độ môi khuếch đại 10 locus Y-STR

Nồng độ môi	Kết quả điện di gel	Dư môi
5	Lên band rõ	Không
10	Lên band rõ	Không
15	Lên band rõ	Có
20	Lên band rõ	Có

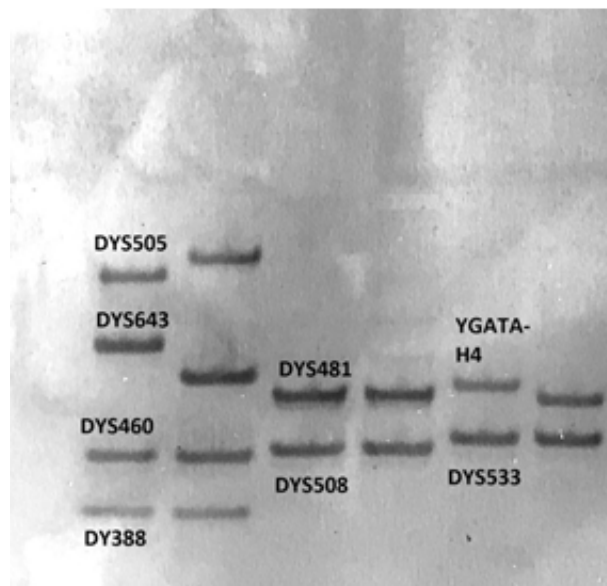
Thời gian kéo dài chuỗi được tối ưu ở 3 điều kiện khác nhau: 30s, 60s và 90s. Kết quả cho thấy ở cả 3 điều kiện về thời gian kéo dài chuỗi đều cho kết quả band rõ nét, không có band phụ. Do các locus có kích thước gần nhau nên thời gian kéo dài chuỗi tối ưu được lựa chọn là 60s.



Hình 3. 7. Ảnh điện di sản phẩm sau PCR với 10 locus Y-STR trên gel Polyacrylamide 6% (Sử dụng thang chuẩn ILS 500 (Promega - Mỹ)).

Với điều kiện phản ứng PCR được tối ưu cho kết quả điện di là các băng sản phẩm PCR rõ nét, không có băng phụ và có kích thước đúng như các công bố trước đó. Kích thước các locus dao động từ 91-213 bp (Hình 3.7). Điều này cho thấy các cặp môi được thiết kế có tính khả dụng cho việc khuếch đại các locus Y-STR mong muốn.

Việc thực hiện các phản ứng PCR đơn lẻ với cùng một mẫu nghiên cứu đòi hỏi tiêu tốn nhiều hóa chất, thời gian và công sức. Phản ứng PCR đa môi được thiết kế dựa trên việc trộn nhiều cặp môi vào trong cùng 1 phản ứng nhằm khuếch đại đồng thời nhiều locus Y-STR là cải tiến phù hợp hơn. Nguyên tắc để lựa chọn các locus Y-STR khuếch đại cùng nhau là các locus phải có nhiệt độ gắn môi tương đồng nhau, ngược lại sản phẩm PCR phải có kích thước chênh lệch nhau để đảm bảo sự phân tách trong quá trình điện di kiểm tra.



Hình 3. 8. Ảnh điện di trên gel polyacrylamide các phản ứng PCR đa môi

Trong nghiên cứu này do các locus Y-STR đều có kích thước trong khoảng 100 - 200 bp, kích thước chênh lệch nhau không nhiều nên số các locus trong phản ứng PCR đa môi không vượt quá 4 locus. Chúng tôi đã tối ưu nồng độ môi, nhiệt độ gắn môi và khuếch đại thành công 8 trong số 10 locus Y-STR thông qua 3 phản ứng PCR đa môi khác nhau, cụ thể:

- Phản ứng multiplex 1: gồm các locus DYS505, DYS643, DYS460, DYS388
- Phản ứng multiplex 2: gồm các locus DYS481, DYS508
- Phản ứng multiplex 3: gồm các locus Y-GATA H4, DYS533
- Riêng 2 locus DYS19 và DYS522 do nhiệt độ gắn môi khác nhau nên không thể ghép chung trong phản ứng PCR đa môi.

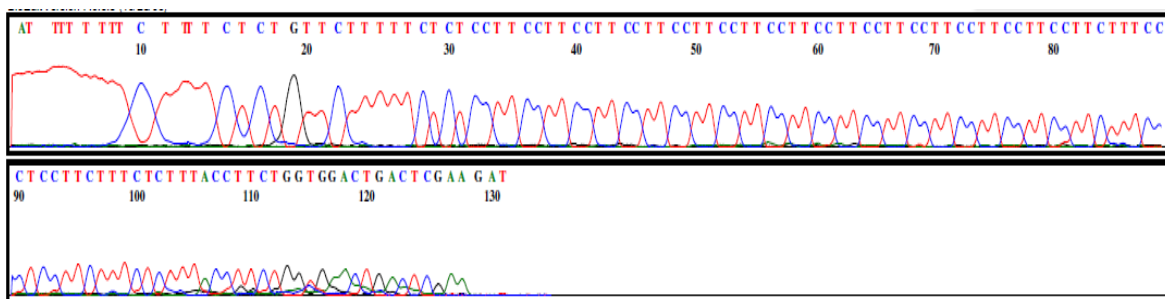
Phản ứng multiplex PCR với nhiều cặp môi cùng cũng là cơ sở để tạo nên các bộ kit STR và Y-STR thương mại hiện nay có thể phân tích lên tới 27 locus thông qua một phản

ứng PCR duy nhất. Trong nghiên cứu này với kỹ thuật PCR đa môi và điện di gel polyacrylamide vẫn đòi hỏi tiêu tốn khá nhiều thời gian, công sức và chưa đạt độ nhạy cao. Việc thiết kế các môi có gắn dấu huỳnh quang (fluorescence dye) kết hợp với phương pháp điện di mao quản sẽ là giải pháp tối ưu cho việc phân tích đồng thời nhiều STR và làm tăng khả năng phát hiện các locus STR ngay cả ở nồng độ ADN thấp. Đây cũng là hạn chế và là mục tiêu tiếp theo mà đề tài cần giải quyết.

3.3.3. Giải trình tự xác định cấu trúc lặp các mini Y-STR

Để kiểm tra chính xác sản phẩm sau PCR có phải là các locus Y-STR mong muốn, sản phẩm PCR được giải trình tự gen. Sau khi tinh sạch sản phẩm sau PCR các mẫu được chạy cycle sequencing PCR với Bigdye Terminator Kit, tinh sạch sau phản ứng Bigdye và giải trình tự trên máy ABI 3500, capillary 50 cm và pop 7. Theo hướng dẫn của nhà sản xuất với thời gian kéo dài chuỗi là 60°C trong 4 phút. Với chu kỳ nhiệt này, phòng thí nghiệm chúng tôi đã sử dụng cho các cặp môi nhân đoạn trên 400bp đã cho kết quả giải trình tự đẹp.

Tuy nhiên trong nghiên cứu do các đoạn cần giải trình tự có kích thước dưới 200bp nên chúng tôi đã điều chỉnh giảm một nửa thời gian kéo dài chuỗi trong chu trình nhiệt chạy phản ứng PCR Bigdye để rút ngắn tổng thời gian chạy phản ứng Bigdye, vì thế đã tiết kiệm đáng kể thời gian cho phản ứng giải trình tự tiếp theo. Kết quả giải trình tự trên máy ABI 3500 với chương trình chạy của hãng cho trình tự có các đỉnh sóng rõ ràng.

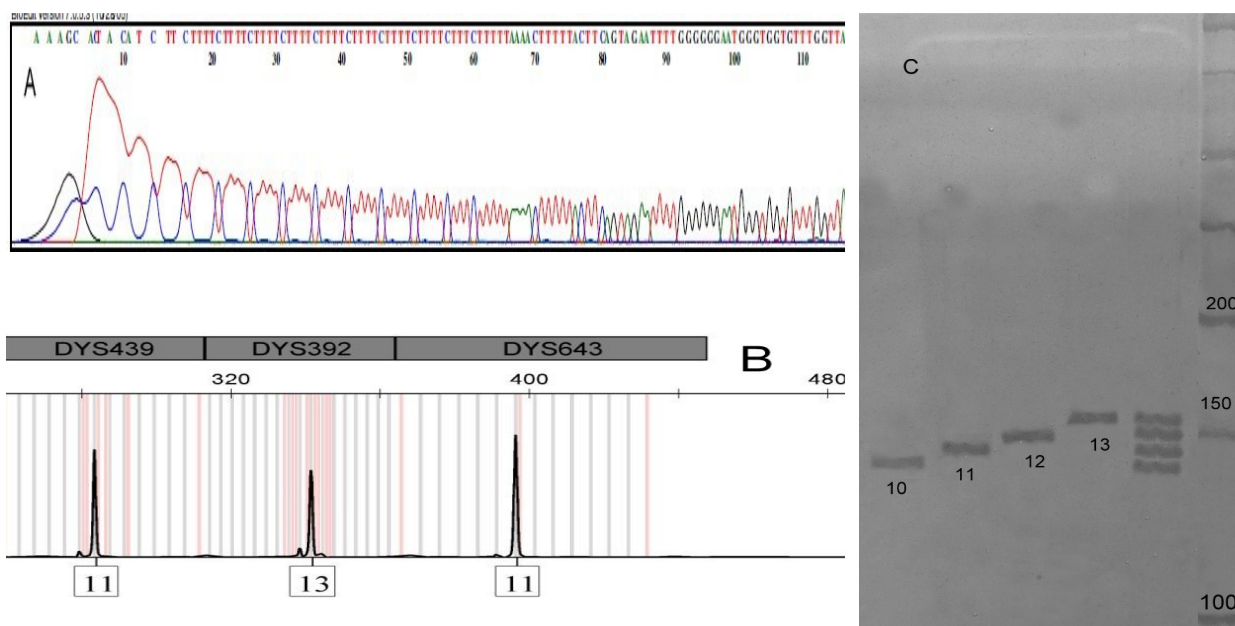


Hình 3. 9. Kết quả giải trình tự locus DYS505

Kết quả giải trình tự locus DYS505 cho thấy kích thước đoạn khoảng hơn 170 bp và cấu trúc lặp TCCT.

là alen 15 (hình 3.9). Đây cũng là cơ sở để xây dựng thang alen cho các locus mới chưa có trong các bộ kit thương mại.

Minh họa với locus DYS643, với các mẫu đã được xác định rõ alen bằng bộ kit PPY23 chúng tôi đã lựa chọn các mẫu có alen khác nhau và thực hiện phản ứng PCR bằng phương pháp mini Y-STR. Tiếp đó chúng tôi thực hiện giải trình tự để kiểm định lại số alen giữa phương pháp phân tích bằng kit PPY23 và mini Y-STR có trùng khớp nhau không. Việc đếm số lần lặp của cấu trúc lõi bằng phương pháp giải trình tự cũng cho kết quả trùng khớp với việc phân tích bằng kit thương mại như minh họa với locus DYS643 (Hình 3.13A). Trộn lẫn sản phẩm PCR và điện di gel các mẫu có số alen khác nhau lần lượt là 10, 11, 12, 13 cho phép xây dựng thang alen chuẩn cho locus mini DYS643 (hình 3.13C). Thang alen này sẽ làm cơ sở để xác định số alen chính xác mà không cần phải giải trình tự lại trong các phân tích sau này đồng thời cũng là cơ sở để phân biệt sự sai khác alen giữa các mẫu.



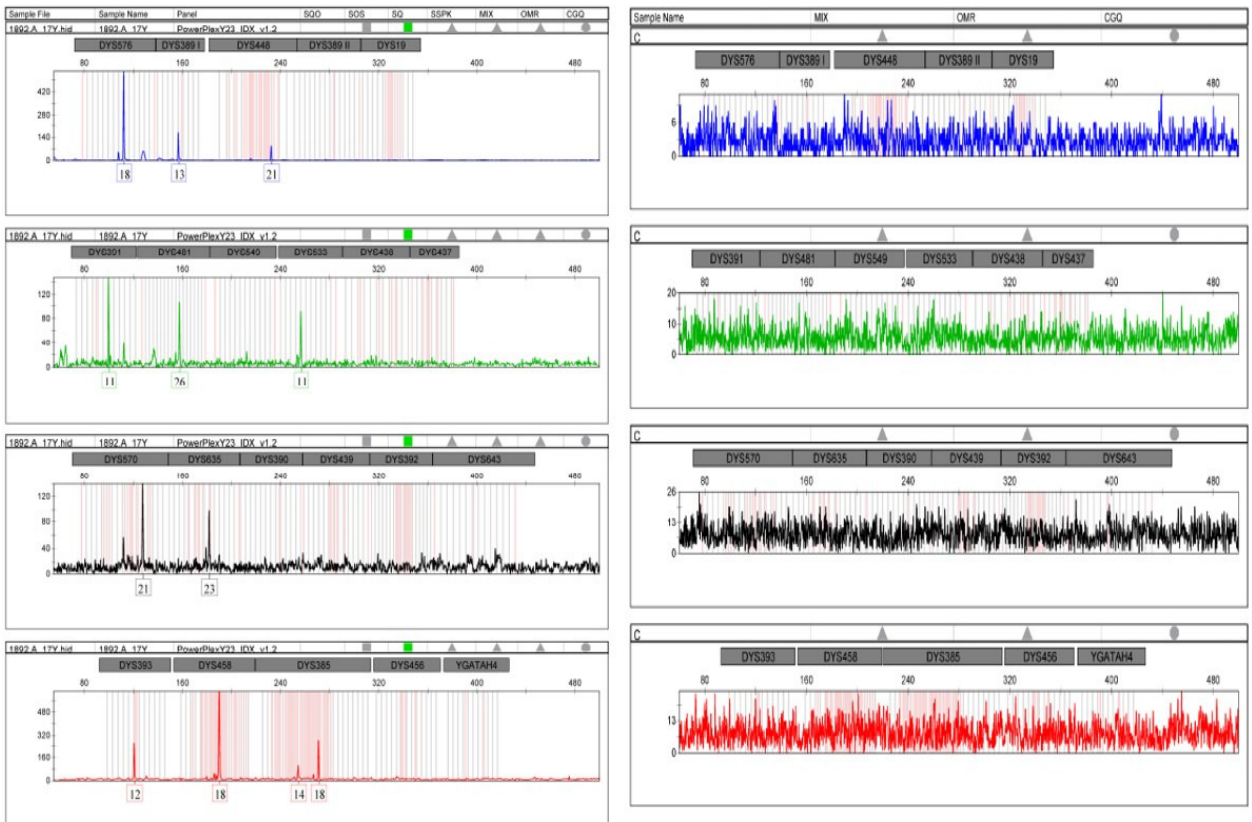
Hình 3. 13. Kết quả tạo thang chuẩn mini Y-STR với chỉ thị DYS643

A. Kết quả GTT với số lần lặp CTTT là 11, B. Kết quả phân tích sử dụng bộ PPY23 với số alen là 11, C: Thang chuẩn chỉ thị DYS643 với 4 alen 10, 11,12, 13

3.3.4. Tỷ lệ khuếch đại thành công các mini Y-STR

Dựa vào phương pháp PCR đa môi với 10 cặp môi nêu trên chúng tôi đã ứng dụng trong việc phân tích 30 mẫu đã phân hủy là những mẫu xương và răng hài cốt liệt sỹ thu thập từ các nghĩa trang. Những mẫu này đã được phân tích sử dụng bộ kit PPY23 và YPlus nhưng không cho được kết quả hoặc kết quả có các đỉnh tín hiệu lên rất thấp không đủ độ tin cậy để đọc.

Với các mẫu răng, xương đã bị phân hủy tỷ lệ khuếch đại thành công của 10 mini Y-STR dao động từ 44 % đến 82%. Cụ thể: có 2 locus DYS533 và DYS481 đạt tỷ lệ thành công cao nhất (lần lượt là 80%, 76.7%) với kích thước locus chỉ khoảng 100 - 150 bp. Tiếp theo là các locus DYS388, DYS508, DYS460, YGATA-H4 đạt tỷ lệ thành công từ 66.7 – 73.3% với khoảng kích thước khuếch đại dao động từ 91 – 141 bp. Hai locus DYS522, DYS505 với kích thước khoảng 112 – 176 bp chỉ đạt tỷ lệ thành công 53,3% và 56,7%. Riêng hai locus có kích thước lớn là DYS19 (177-213 bp) và DYS643 (121 – 166 bp) đạt tỷ lệ khuếch đại khá thấp, lần lượt là 43.3 % và 50%.



Hình 3. 14. Hồ sơ Y-STR của mẫu hài cốt sử dụng bộ kit PPY23

Tỷ lệ khuếch đại với mẫu hài cốt đã phân hủy ở các mini Y-STR có sự chênh lệch nhau. Điều này có thể được lý giải do locus có kích thước càng lớn (ví dụ locus DYS19 ~

200bp) thì tỷ lệ khuếch đại thành công càng thấp. Đặc biệt là với các mẫu hài cốt có ADN đã bị phân hủy, đứt gãy nhiều. Trong nghiên cứu của Myung Jin Park trên các mẫu đã bị phân hủy thành công với locus DYS19 cũng chỉ đạt 23,3%, YGATA-H4 là 33,3% [45].

Bảng 3. 38. Bảng thống kê tỷ lệ khuếch đại thành công 10 locus mini Y-STR trên 30 mẫu đã bị phân hủy

Locus	Kích thước trong PPY23	Kích thước trong YPlus	Kích thước dùng mini Y-STR	Tỷ lệ khuếch đại thành công
DYS533	245-285	338-379	107-127	80.0%
DYS19	312-352	184-224	177-213	43.3%
DYS481	139-184	207-252	119-149	76.7%
DYS643	368-423		121-166	50.0%
Y-GATA-H4	374-414	236-264	121-141	70.0%
DYS460		79-109	101-121	60.0%
DYS505			164-176	56.7%
DYS508			105-129	73.3%
DYS388			91-106	66.7%
DYS522			112-132	53.3%

Ngoài ra một yếu tố khác ảnh hưởng đến hiệu suất PCR là chất lượng mẫu nghiên cứu. Trong nghiên cứu với những mẫu hài cốt lâu năm đã phân hủy nặng (chỉ số phân hủy ở mức cao), ADN bị đứt gãy nhiều thì việc khuếch đại hết các locus Y-STR cũng gặp khó khăn. Đặc biệt trong điều kiện khí hậu Việt Nam nóng ẩm, nhiều loại vi sinh vật trong lòng đất ảnh hưởng lớn đến chất lượng mẫu xương.

Có thể thấy việc ứng dụng mini Y-STR giải quyết trong nhiều tình huống đặc biệt là trong công tác giám định mẫu phân hủy. Một số nghiên cứu khác cũng chỉ ra việc tối ưu lại kích thước các locus Y-STR giúp ích đáng kể trong việc phân tích mẫu đã bị phân hủy. Điển hình trong nghiên cứu của Y. Gu và cs cho thấy với việc giảm kích thước 3 locus DYS385a/b, DYS389 II, and DYS392 cho phép khuếch đại thành công mẫu có ADN ở nồng độ rất thấp chỉ từ 0.25 ng [46]. Trong nghiên cứu khác của Park MJ cho thấy mini Y-

STR tự thiết kế cho tín hiệu phân tích tốt hơn so với sử dụng bộ kit thương mại có sẵn khi phân tích trên mẫu xương 50 năm tuổi và các mẫu đã bị xử lý phân hủy bằng enzyme [45].

3.4. Hiệu quả của các chỉ thị Y-STR trong xét nghiệm ADN

3.4.1. Hiệu quả trong phân tích mẫu đã bị phân huỷ

Sau khi nghiên cứu thành công kỹ thuật khuếch đại các mini Y-STR với các cặp môi tương ứng, chúng tôi đã thử ứng dụng vào một số trường hợp giám định cụ thể tại Khoa Y - Sinh học, Viện Pháp y quốc gia.

3.4.1.1. Trong công tác giám định án

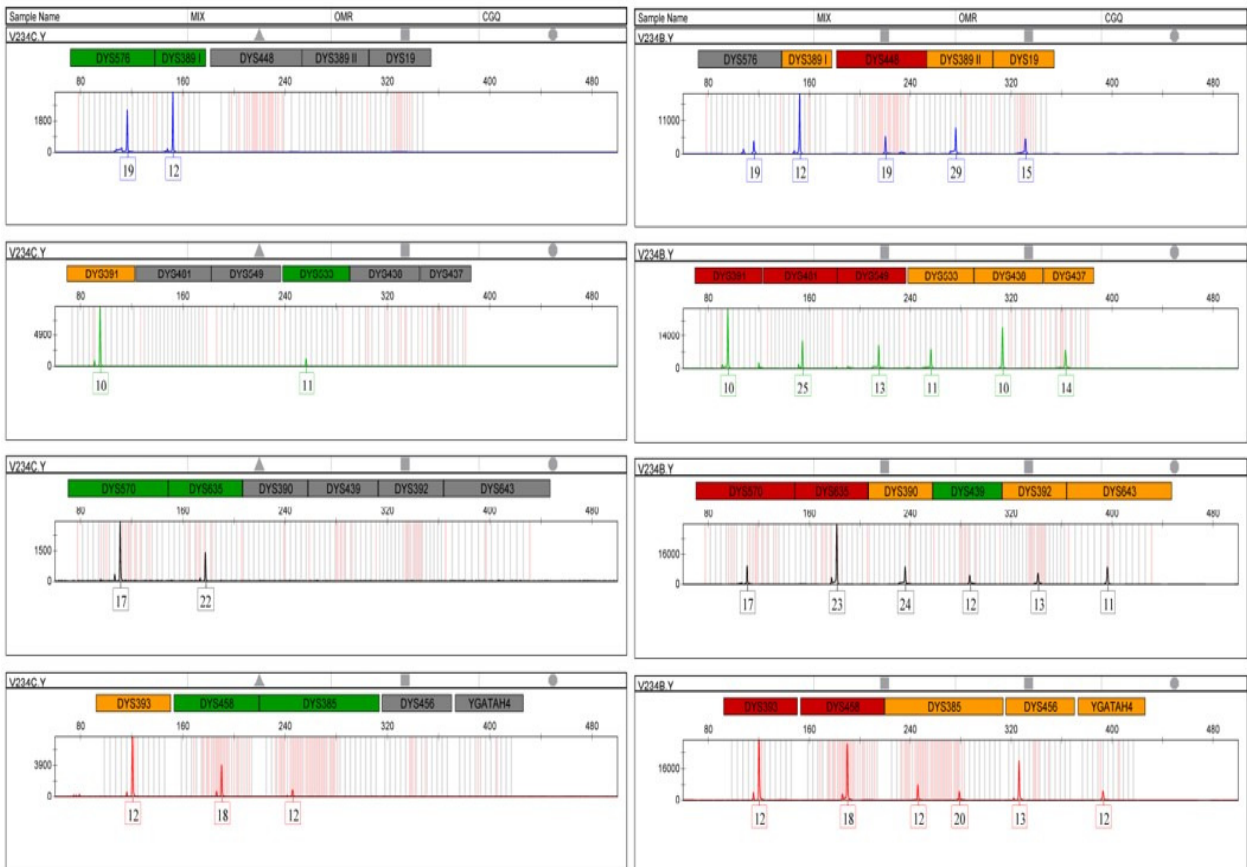
Hiện nay mỗi năm tại Khoa Y sinh học đang thực hiện nhiệm vụ giám định vụ án từ quyết định trưng cầu giám định của Cơ quan công an. Rất nhiều mẫu thu thập tại hiện trường vụ án được thu sau thời điểm xảy ra vụ án một thời gian khiến chất lượng ADN trên mẫu vật không còn nguyên vẹn hoặc lượng ADN thu thập trên mẫu rất ít (mẫu ở dạng vi vết). Điều này gây khó khăn rất nhiều cho công tác giám định và đưa ra kết luận. Với hiệu quả trong nghiên cứu mini Y-STR chúng tôi đã ứng dụng trong thực tế giám định và thu được các kết quả nhất định có thể minh họa qua trường hợp cụ thể sau:

Trường hợp 1. Cơ quan cảnh sát điều tra Công an tỉnh Sơn La trưng cầu giám định ADN tại Viện Pháp y quốc gia vụ án ký hiệu V234: Tử thi chết chưa rõ nguyên nhân. Tại nương của gia đình anh Thào A Thào thuộc xã Chiềng Sơn, huyện Mộc Châu, tỉnh Sơn La phát hiện một tử thi nam giới chết trong tư thế nằm sấp, vùng đầu của nạn nhân có nhiều vết thương. Cơ quan cảnh sát điều tra công an tỉnh Sơn La đã phối hợp cùng các đơn vị chức năng tiến hành khám nghiệm hiện trường, khám nghiệm tử thi, tiến hành điều tra xác minh ban đầu tử thi là Xênh A V (Năm sinh: 1990; Trú tại: Phù Yên, Sơn La). Quá trình điều tra, Cơ quan CSĐT Công an tỉnh Sơn La đã tiến hành thu mẫu con dao cách hiện trường gây án 20m và thu mẫu của Lò Văn D (Năm sinh: 1994) là nghi phạm trong vụ án. Cơ quan CSĐT trưng cầu giám định ADN tìm trên cán của con dao có ADN trùng nghi phạm hay không?

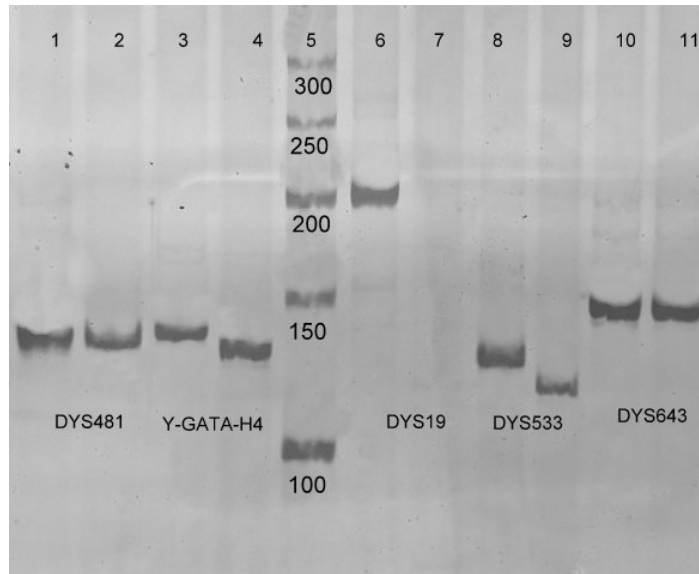
Khi tiến hành phân tích mẫu phết phần cán con dao thu được hồ sơ STR có lẫn từ nhiều người và lên thiếu peak, chúng tôi tiến hành phân tích hồ sơ Y-STR sử dụng bộ kit PPY23. Kết quả cho thấy mẫu cán dao thu được 9/23 locus và có 8/9 locus đều trùng với hồ sơ Y-STR của nghi phạm. Thực tế nghi ngờ đây là mẫu các tế bào niêm mạc da và mồ hôi của nghi phạm

đỉnh trên cán dao, hung khí lại được thu thập sau 3 ngày kể từ khi xảy ra vụ việc nên nồng độ ADN thu được rất thấp. Điều này cũng được thể hiện rõ qua cường độ tín hiệu peak của mẫu cán dao rất thấp so với mẫu so sánh từ nghi phạm. So sánh dữ liệu ADN thu từ mẫu hiện trường với dữ liệu 400 mẫu trong nghiên cứu cho thấy trùng khớp với 1 hồ sơ. Do đó câu hỏi được đặt ra là với 1 chỉ thị Y-STR khác nhau khi so sánh hồ sơ Y-STR giữa 2 mẫu chưa đủ để đưa ra kết luận hai hồ sơ này là khác nhau hay do đột biến?

Vì vậy để có cơ sở khẳng định, chúng tôi tiến hành phân tích thêm 5 mini STR có trong bộ PPY23 bằng phương pháp PCR đơn mồi. Kết quả phân tích cho thấy có sự sai khác rõ ở 2 locus Y-GATA H4 và DYS643, và DYS19. 3 locus DYS533, DYS481 trùng nhau. Tại vị trí locus DYS19 với mẫu hiện trường không lên được băng do nồng độ ADN kém. Điều này cũng đủ cho phép chúng tôi rút ra kết luận nghi phạm ADN thu được trên cao không trùng với ADN của nghi phạm. Trong trường hợp này phân tích thêm chỉ thị Y-STR đã giúp loại trừ nghi phạm liên quan vụ án.



Hình 3. 15. So sánh hồ sơ Y-STR từ mẫu hiện trường (trái) và mẫu nghi phạm (phải) vụ án kí hiệu V234

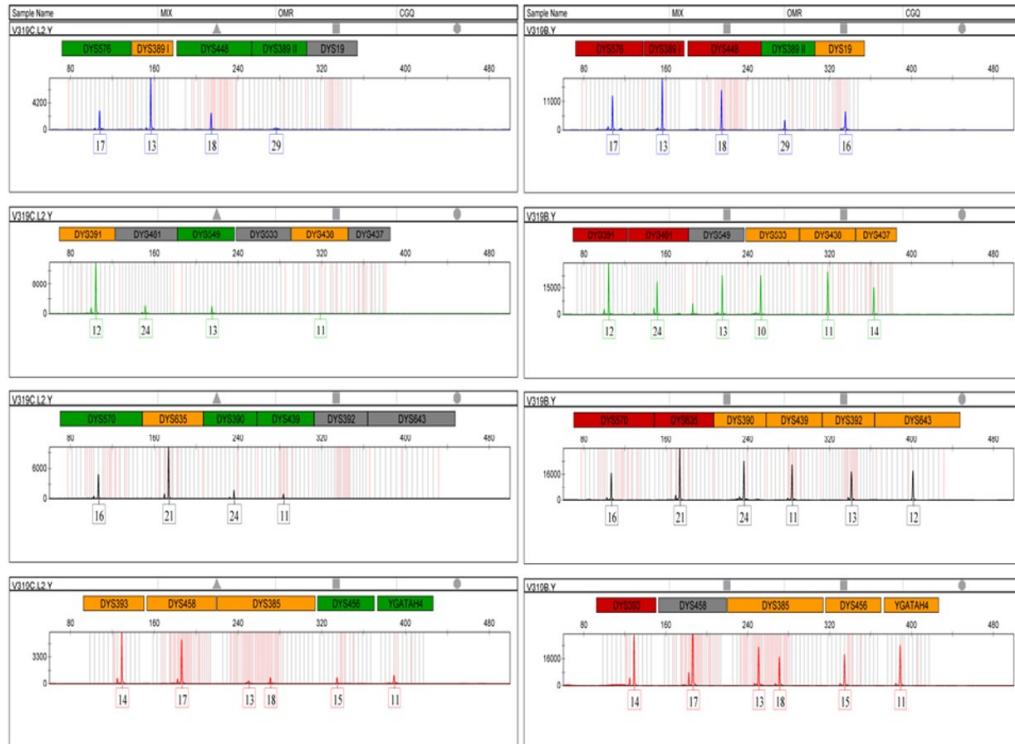


Hình 3. 16. So sánh mini STR từ mẫu hiện trường và mẫu nghi phạm vụ án kiện V213
Giếng 1,3,7,9,11: mẫu từ hiện trường, giếng 2,4,6,8,10: mẫu từ nghi phạm, 5: ladder

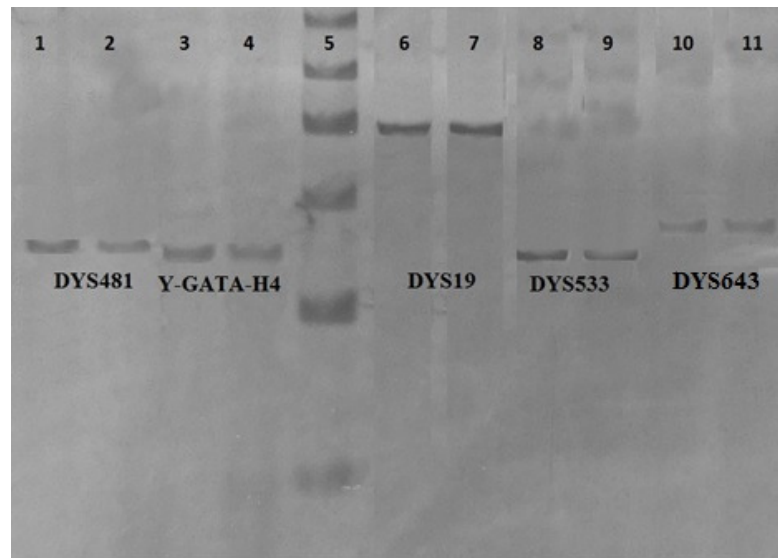
Trường hợp 2: Cơ quan cảnh sát điều tra Công an huyện Sốp Cộp, tỉnh Sơn La trung cầu Viện Pháp y quốc gia vụ án ký hiệu V319: Chết người không rõ nguyên nhân:

Chị Giàng Thị D (Năm sinh: 2001; trú tại xã Lậm Lạch, huyện Sốp Cộp, Sơn La) đi dự đám cưới ở bản khác trong cùng xã đến tối không thấy về nhà nên gia đình tổ chức đi tìm kiếm; Sau 2 ngày tìm kiếm gia đình phát hiện thi thể của Giàng Thị D trên đồi cách nơi có đám cưới 4 km. Gia đình trình báo chính quyền. Các cơ quan chức năng vào cuộc. Quá trình khám nghiệm tử thi của Giàng Thị D để xác định nguyên nhân cái chết và đồng thời thu mẫu dịch âm đạo của Giàng Thị D trung cầu giám định ADN mẫu dịch âm đạo của Giàng Thị D có tinh dịch không? Nếu có tinh dịch trong âm đạo của Giàng Thị D trùng với 3 mẫu nghi phạm (Cơ quan cảnh sát điều tra Công an huyện Sốp Cộp sử dụng nghiệp vụ tìm ra được 3 nghi phạm có khả năng liên quan đến vụ án).

Tiến hành tách chiết ADN từ mẫu dịch âm đạo và phân tích hồ sơ STR trên NST thường chỉ thu được ADN của nạn nhân. Điều này được lí giải là do nồng độ ADN trong dịch âm đạo của nạn nhân là rất nhiều và lấn át hoàn toàn ADN nam giới của nghi phạm để lại (nếu có). Khi đó phân tích chỉ thị Y-STR - là những chỉ thị chỉ có ở nam giới sẽ là phương pháp hữu ích. Kết quả phân tích hồ sơ Y-STR từ mẫu dịch âm đạo sử dụng bộ PPY23 phát hiện được 18/23 locus và đều trùng với 1 nghi phạm.



Hình 3. 17. So sánh hồ sơ Y-STR giữa mẫu nạn nhân (trái) và mẫu nghi phạm (phải) vụ án ký hiệu V319



Hình 3. 18. So sánh mini STR từ mẫu nạn nhân và mẫu nghi phạm vụ án ký hiệu V319

Giếng 1,3,6,8,10: mẫu từ nghi phạm; Giếng 2,4,7,9,11: mẫu từ nạn nhân, giếng 5: ladder

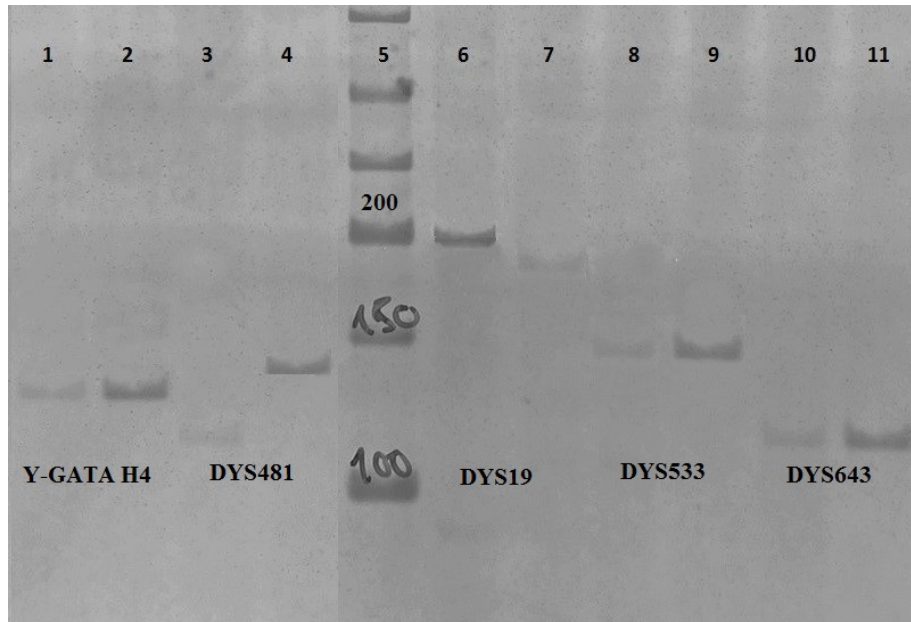
Do cả 3 đều trú tại cùng xã nên có khả năng có quan hệ họ hàng theo dòng cha hoặc hồ sơ Y-STR trùng khớp nhau nhiều nên để đảm bảo kết quả chính xác chúng tôi tiến hành

phân tích thêm 5 chỉ thị mini Y-STR có trong nghiên cứu. Kết quả phân tích với cặp mỗi đơn cho thấy trùng khớp nhau hoàn toàn giữa hai mẫu ở cả 5 chỉ thị Y-STR là DYS19, DYS643, DYS481, DYS533 và Y-GATA H4, Kết quả này cùng với tình hình thực tế của vụ án đủ cho phép chúng tôi đưa ra kết luận ADN thu từ mẫu thu dịch âm đạo nạn nhân trùng khớp với ADN của nghi phạm.

3.4.1.2. Trong công tác giám định danh tính hài cốt liệt sỹ

Trong công tác giám định danh tính hài cốt liệt sỹ; hài cốt – là những mẫu có ADN bị phân huỷ, đứt gãy nhiều việc ứng dụng mini Y-STR cũng đem lại nhiều hiệu quả. Thông thường để thực hiện giám định hài cốt sẽ dựa vào phân tích và so sánh trình tự các vùng siêu biến HV1, HV2 trên ADN ty thể nhằm xác định mối quan hệ huyết thống theo dòng mẹ. Tuy nhiên trong một số trường hợp khi phân tích vùng HV1, HV2 trên ADN ty thể cho thấy có sự sai khác rất nhỏ tại 1-2 vị trí. Sự sai khác này nhất có thể do sự khác nhau về huyết thống giữa 2 cá thể hoặc do đột biến ty thể, sự amin hoá do mẫu tiếp xúc lâu với điều kiện môi trường không thuận lợi. Với những trường hợp này sẽ không đủ cơ sở để kết luận mối quan hệ huyết thống theo dòng mẹ. Hoặc có trường hợp chỉ có mẫu so sánh theo dòng cha (con trai, cháu trai...) thì không thể áp dụng phương pháp phân tích ADN ty thể. Khi đó việc phân tích thêm Y-STR sẽ cung cấp dữ liệu bổ sung có giá trị quan trọng trong việc đưa ra kết luận cuối cùng.

Trường hợp điển hình là việc phân tích mẫu hài cốt ký hiệu NCC3035. Kết quả phân tích vùng HV1 trên ADN ty thể giữa mẫu hài cốt và mẫu thân nhân so sánh cho thấy có sự sai khác tại đúng một vị trí trên vùng HV1. Trình tự vùng HV2 giữa 2 mẫu hài cốt và thân nhân cũng chỉ khác nhau tại 1 vị trí. Với kết quả này rất khó để đưa ra kết luận cuối cùng bởi theo khuyến cáo giữa hai mẫu phải khác nhau từ vị trí 2 nucleotit trở lên mới được phép kết luận hai mẫu không có quan hệ huyết thống theo dòng mẹ. Sự sai khác tại 1-2 vị trí nucleotit có thể do đột biến, do mẫu hài cốt bị khử amin hoặc do sự sai khác thật sự giữa hai mẫu cần so sánh. Vì mẫu hài cốt và mẫu thân nhân được cho là 2 anh em trai ruột cùng cha cùng mẹ nên chúng tôi đã thử phân tích thêm 5 locus mini Y-STR trong đề tài nghiên cứu. Kết quả cho thấy có sự sai khác giữa 2 mẫu tại locus DYS19 và DYS481. Điều này giúp chúng tôi có thêm cơ sở để kết luận giữa mẫu hài cốt NCC3035 và mẫu thân nhân ký



Hình 3. 19. So sánh mini STR từ mẫu mẫu hài cốt và mẫu thân nhân

Giếng 1,3,7,8,10: mẫu hài cốt, giếng 2,4,6,9,11: mẫu thân nhân, 5: ladder

3.4.1.3. Hiệu quả của 2 bộ kit trong phân tích mẫu đã phân hủy

Bên cạnh các mini Y-STR, hai bộ kit PPY23 và YPlus cũng đều gồm một số các alen vừa có kích thước nhỏ vừa có độ đa dạng gen cao như locus DYS576 (kích thước khoảng 80 - 130 bp, số alen: 6, GD = 0,73), locus DYS570 (kích thước khoảng 100 - 160bp, số alen: 8, GD = 0,83), DYS458 (kích thước khoảng 110 - 160 bp, số alen: 9, GD = 0,80) sẽ là các locus gen lý tưởng cho việc phân tích bổ sung trong các trường hợp giám định ADN với mẫu đã bị phân hủy nhiều, ADN đã bị đứt gãy thành các đoạn nhỏ như trong mẫu xương hài cốt, mẫu sau thảm họa, thiên tai...

Trong nghiên cứu này để đánh khả năng phân tích với mẫu đã bị phân huỷ của 2 bộ kit chúng tôi lọc riêng dữ liệu hồ sơ ADN từ 400 mẫu nghiên cứu khảo sát với 2 bộ kit gồm các Y-STR có kích thước < 220 bp thuộc 2 bộ kit và tính toán các chỉ số di truyền liên quan. Bộ PPY23 gồm 8 locus có kích thước < 220 bp (DYS576, DYS389I, DYS391, DYS481, DYS570, DYS635, DYS393 và DYS458). Bộ YPlus gồm 10 locus có kích thước < 220 bp (DYS576, DYS389I, DYS458, DYS456, DYS390, DYS570, DYS437, DYS393, DYS439 và DYS460). Kết quả trong Bảng 3.39 cho thấy bộ PPY23 có haplotype giống nhau ở 2, 3 cá thể trong khi bộ YPlus có haplotype giống nhau ở 2, 3, 4 và 7 cá thể. Số haplotype quan sát được trong bộ PPY23 và YPlus lần lượt là 193, 185 haplotypes (trên tổng số 200 mẫu

phân tích). Khả năng phân biệt của bộ PPY23 là 97% cũng nhiều hơn đáng kể so với bộ YPlus là 93%.

Bảng 3. 41. So sánh các chỉ số di truyền thu được từ tập hợp các chỉ thị Y-STR có kích thước < 220 bp trong 2 bộ PPY23 và YPlus

Số haplotype	Bộ haplotype PPY23	Bộ haplotype YPlus
Chỉ xuất hiện ở 1 cá thể	188	178
Xuất hiện ở 2 cá thể	3	4
Xuất hiện ở 3 cá thể	2	1
Xuất hiện ở 4 cá thể	0	1
Xuất hiện ở 5 cá thể	0	0
Xuất hiện ở 6 cá thể	0	0
Xuất hiện ở 7 cá thể	0	1
Xuất hiện ở 8 -10 cá thể	0	0
Tổng số locus	8	10
Tổng số haplotype	193	185
HD	0.99954	0.99829
DC	97%	93%

Có thể nhận thấy dù có số lượng chỉ thị Y-STR ít hơn nhưng tập hợp Y-STR kích thước < 220 bp trong bộ PPY23 lại có khả năng phân biệt tốt hơn so với tập hợp trong bộ YPlus. Điều này được lý giải do các locus có kích thước nhỏ nhưng có độ đa hình cao trong bộ PPY23 là DYS481 (GD: 0.77), DYS635 (GD: 0.76). Hai locus này đã có hiệu quả làm đáng kể khả năng phân biệt cho bộ kit. Ngược lại tập hợp trong bộ YPlus tuy có số lượng locus kích thước nhỏ nhiều hơn nhưng lại chủ yếu là các locus có độ đa hình thấp (GD khoảng 0.5 – 0.6) nên độ đa dạng haplotype và khả năng phân biệt chung không được cao.

Kết quả này cũng phù hợp với các công bố trước đây cho thấy tiềm năng cao của bộ PPY23 trong việc phân biệt các mẫu đã bị phân huỷ [25]. Bên cạnh đó, hiệu quả của bộ kit YPlus trong việc phân tích các mẫu xương lâu năm, nạn nhân sau thảm họa cũng đã được khảo sát trong nghiên cứu của Ambers và cs [96]. Trong nghiên cứu này các mẫu xương thu thập từ Nội chiến Hoa Kỳ (1861-1865), kỷ nguyên đào vàng cuối thế kỷ XIX ở

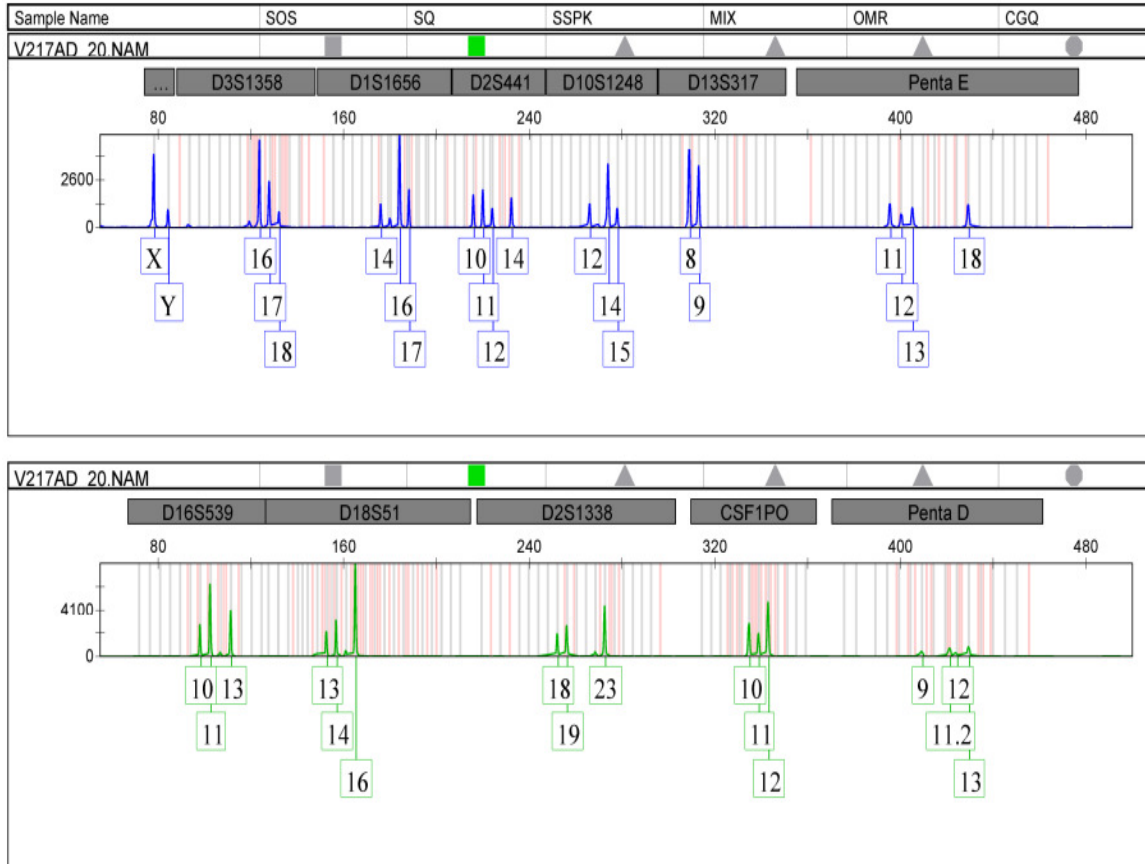
Deadwood, Hoa Kỳ (1874-1877), Chiến tranh Bảy năm (1756-1763), một địa điểm khảo cổ thế kỷ XVII ở Raspenava, Bohemia (Cộng hòa Séc) và Thế chiến II (1939-1945) được phân tích với cả 2 bộ kit Yfiler và YPlus. Kết quả 22 trong số 24 mẫu xương thu được số lượng alen lớn hơn khi sử dụng bộ YPlus cho thấy tiềm năng bộ kit ứng dụng trong giám định các mẫu xương hài cốt đã bị phân huỷ. Bởi vậy, việc kết hợp phân tích Y-STR với STR trong nhân là xu hướng giải quyết nhiều trường hợp mẫu đã bị phân huỷ, đặc biệt là xác định danh tính hài cốt [97].

3.4.3. Hiệu quả sử dụng Y-STR trong phân tích mẫu lẫn (mixture sample)

Tại Việt Nam, phân tích ADN đã bắt được ứng dụng từ những năm 1990 và dần trở thành công cụ đắc lực phục vụ công tác điều tra, phá án. ADN được coi là bằng chứng không thể chối cãi cũng như có độ chính xác cao nhất tại các phiên tòa. Đặc biệt hiện nay các vụ án có xu hướng ngày càng phức tạp do một số nguyên nhân như: cách thức phạm tội, xóa dấu vết và che giấu hiện trường tinh vi...Số lượng các vụ án cũng không ngừng tăng đòi hỏi việc sử dụng phương pháp phân tích cho kết quả nhanh chóng, chính xác. Với đặc thù đó phân tích ADN trong điều tra hình sự đang ngày càng được áp dụng phổ biến tại Việt Nam cũng như trên thế giới.

Hiện nay tại các trung tâm giám định, Viện nghiên cứu phân tích ADN phục vụ khoa học hình sự chủ yếu tập trung vào lập hồ sơ ADN dựa trên các STR trên NST thường vào so sánh khớp nối các hồ sơ này giữa mẫu nghi phạm và mẫu thu từ hiện trường vụ án. Tuy nhiên trong trường hợp mẫu thu tại hiện trường là mẫu lẫn (mixture sample) có sự đóng góp ADN của nhiều hơn một người thì tại mỗi vị trí locus phân tích thường sẽ thu được nhiều hơn 2 alen. Bởi vậy việc so sánh khớp nối trong hồ sơ ADN sẽ gặp nhiều khó khăn và khó xác định nguồn gốc các alen từ những người tham gia đóng góp trong hồ sơ ADN.

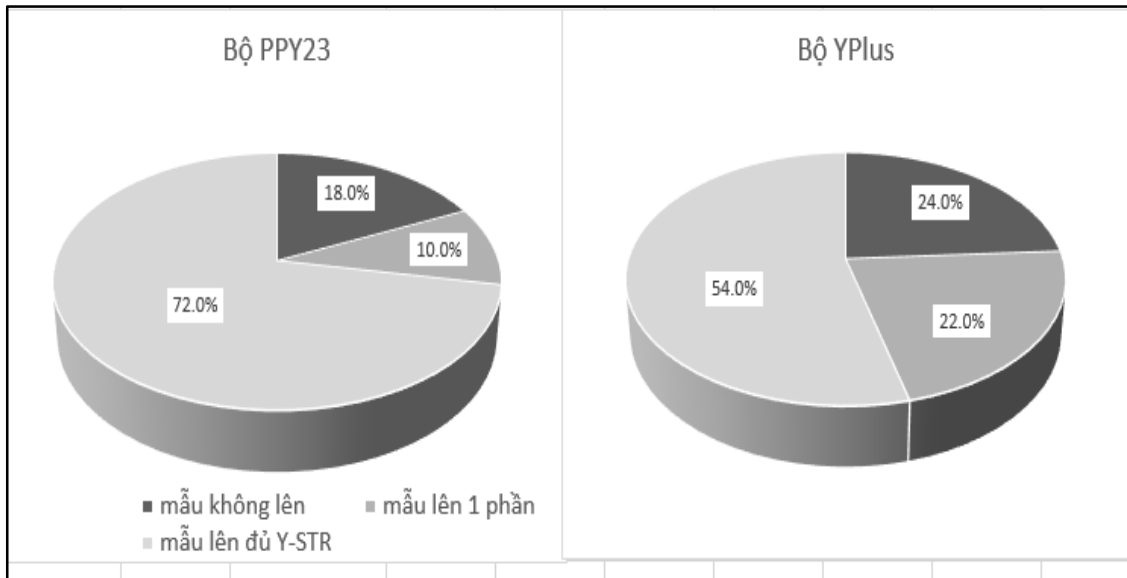
Một trong những dạng mẫu lẫn thường gặp nhất là mẫu thu từ dịch âm đạo của nạn nhân nữ trong các vụ án hiếp dâm, tấn công tình dục. Trong đó thường rất khó thu được hồ sơ ADN hoàn chỉnh của các STR trong nhân từ nghi phạm nam do nồng độ ADN của nạn nhân nữ thường lẫn át, cao hơn nhiều lần. Khi đó việc phân tích thêm chỉ thị Y-STR đặc trưng cho nam giới có khả năng cung cấp thêm thông tin, củng cố độ tin cậy của hồ sơ ADN.



Hình 3. 20. Hồ sơ STR phân tích từ mẫu dịch âm đạo có dạng mẫu lẫn nhiều nguồn ADN

Trong nghiên cứu này, chúng tôi đã thử phân tích 50 mẫu lẫn với cả 2 bộ kit PPY23 và YPlus nhằm đánh giá khả năng tạo hồ sơ Y-STR hoàn chỉnh của cả 2 bộ kit. Những mẫu này đều là mẫu dịch trong âm đạo của nạn nhân nữ thu thập từ vụ giám định án hiếp dâm tại Khoa Y sinh học, Viện Pháp y quốc gia. Phân tích các STR trong nhân bằng bộ kit Powerplex® fusion system (Promega) trong những mẫu này đều cho thấy hồ sơ ADN thu được chủ yếu ở dạng mẫu lẫn giữa nam và nữ nên không thể được sử dụng để truy nguyên cá thể.

Chúng tôi đã tiến hành phân tích các mẫu trên với hai bộ kit Y-STR nhằm thu thập hồ sơ STR từ nam giới. Kết quả cho thấy với bộ kit PPY23 số mẫu lên được hồ sơ Y-STR đầy đủ là 36 mẫu chiếm tỷ lệ nhiều nhất là 72% tổng số mẫu nghiên cứu, số mẫu lên được một phần là 5 mẫu chiếm 10%, mẫu không lên được dữ liệu Y-STR chiếm 18%. Với bộ kit YPlus số mẫu lên được dữ liệu hoàn chỉnh chiếm đa số với 27 mẫu (54%), số mẫu lên một phần dữ liệu là 11 mẫu chiếm 22%, còn lại là 12 mẫu không lên được dữ liệu chiếm 24%.



Hình 3. 21. Đồ thị so sánh khả năng tạo hồ sơ Y-STR từ mẫu lần từ bộ PPY23 và YPlus

Thứ nhất là do số lượng locus Y-STR trong bộ kit YPlus nhiều hơn bộ PPY23 nên về lý thuyết tỷ lệ khuếch đại thành công tất cả các Y-STR sẽ thấp hơn. Hơn nữa, phần lớn các locus mới được bổ sung trong bộ YPlus đều có kích thước lớn nên sẽ khó khuếch đại thành công đặc biệt là những mẫu có đã bị phân huỷ ADN bị đứt gãy nhiều hoặc mẫu có nồng độ ADN thấp. Một số nghiên cứu trước đó cũng chứng minh với mẫu có nồng độ ADN nam giới < 0.5 ng bộ PPY23 có độ nhạy và khả năng khuếch đại thành công cao hơn so với bộ YPlus [34].

Một giải thích khác là do các mẫu ADN có nguồn gốc từ hiện trường vụ án thường có chứa các chất ức chế phản ứng PCR như axit tannic và axit humic. PPY23 và YPlus cho thấy độ nhạy tương tự nhưng PPY23 cho thấy khả năng chịu axit humic cao hơn YPlus theo nghiên cứu của Ju Yeon Jung và cs [98]. Do đó, PPY23 sẽ là một bộ phù hợp hơn để phân tích các mẫu thu thập từ hiện trường vụ án – là dạng mẫu có thể dễ dàng bị ô nhiễm bởi điều kiện môi trường. Tuy nhiên cũng cần lưu ý rằng bộ YPlus cho phép khuếch đại đồng thời 27 locus và có nhiều locus RM Y-STR hơn, trong khi bộ PPY23 chỉ cho phép khuếch đại 23 locus. Do đó bộ YPlus có ưu điểm hơn trong việc phân biệt giữa các cá thể với độ đa hình cao của các locus Y-STR mới được thêm vào trong YPlus so với PPY23.

Khả năng phân tích mẫu lần của 2 bộ kit PPY23, YPlus đã được chứng minh trong một số công bố trước đây. Trong nghiên cứu kiểm định bộ kit Y23 của Jessica Tokarz đã chứng minh bộ PPY23 có khả năng phân tích thành công hồ sơ Y-STR từ mẫu lần nam/nữ

với tỷ lệ ADN nam: ADN nữ là 1: 19. Nghiên cứu cũng chứng minh bộ kit có khả năng tạo hồ sơ Y-STR từ mẫu lần 2 nam giới với tỷ lệ ADN nam: ADN nam lần lượt là 1:1, 1:2, 1:5, 2:1, and 5:1 [99]. Jonelle M. Thompson và cs đã chứng minh bộ PPY23 có khả năng tạo hồ sơ Y-STR từ mẫu lần có nồng độ ADN nữ giới cao gấp 24 lần ADN nam giới [25]. Bộ kit YPlus cũng chứng minh có khả năng tạo hồ sơ Y-STR hoàn chỉnh từ dạng mẫu lần có nồng độ ADN nữ giới rất cao lên tới 1 μ g. Độ nhạy của bộ kit cũng khá cao cho phép phát hiện Y-STR khi nồng độ ADN nam giới trong mẫu chỉ khoảng 125 pg [100]. Bộ kit cũng cho thấy khả năng phát hiện hồ sơ Y-STR nam giới trong những trường hợp xét nghiệm định tính phát hiện tinh dịch/tinh trùng trong mẫu dịch thu âm đạo cho kết quả âm tính [101]. Đây là những trường hợp không có hiện tượng xuất tinh khi xâm hại tình dục nhưng vẫn có tế bào biểu mô nam giới để lại trong mẫu dịch âm đạo với lượng rất nhỏ.

Như vậy với việc bổ sung dữ liệu Y-TR đã giúp bổ sung thêm thông tin, tăng tỷ lệ thành công trong các trường hợp có mẫu lần, hỗ trợ đáng kể cho việc đưa ra kết luận cho vụ án. Tại Khoa Y sinh học từ năm 2016 chúng tôi đã đưa việc bổ sung dữ liệu Y-STR trong công tác giám định án và thu được những kết quả nhất định. Với những hồ sơ Y-STR lên đủ hoàn toàn chúng tôi có thể dựa vào đó để đối chiếu với mẫu so sánh từ nghi phạm để củng cố kết quả, đưa ra kết luận cuối cùng. Với những hồ sơ Y-STR lên được một phần cũng có hiệu quả trong việc loại trừ nghi không phải nghi phạm thực sự của vụ án.

3.4.1. Hiệu quả trong việc tăng khả năng phân biệt giữa các cá thể

So với các bộ kit trước đây mà phổ biến nhất là bộ Yfiler gồm 17 chỉ thị Y-STR, trong bộ kit PPY23 và YPlus đã bổ sung thêm 12 locus như trong bảng 3.42. Các locus mới được bổ sung đã đem lại những ưu thế đáng kể cho 2 bộ kit trên. Đây cũng là những locus lần đầu tiên được khảo sát trong quần thể người Kinh, Việt Nam khi so sánh với các nghiên cứu trước đó.

Việc bổ sung thêm 2 locus đa bản sao trong bộ kit YPlus và PPY23 (là locus có nhiều hơn 1 phiên bản STR ở những vị trí khác nhau trên NST Y) so với các bộ kit trước đó. DYF387S1a/b và DYSS385a/b chứng tỏ là các locus có độ đa dạng gen cao góp phần đáng kể trong việc tăng khả năng phân biệt của bộ kit này.

Bảng 3. 42. Danh sách 12 locus mới có trong bộ kit PPY23 và YPlus

Tên alen	Bộ PPY23	Bộ YPlus	Số alen	GD	Tần số đột biến
DYS481	✓	✓	9	0.769	0.0046
DYS533	✓	✓	6	0.644	0.0029
DYS549	✓		6	0.603	0.0041
DYS570	✓	✓	10	0.808	0.0092
DYS576	✓	✓	6	0.762	0.0121
DYS643	✓		9	0.749	0.0010
DYS449		✓	13	0.900	0.0102
DYS460		✓	5	0.536	0.0044
DYS518		✓	14	0.893	0.0119
DYS627		✓	10	0.820	0.0147
DYF387S1a/b		✓	28	0.934	0.0079

Ngoài ra trong số 29 chỉ thị Y-STR có 7 Y-STR là các locus có tốc độ đột biến nhanh (rapidly mutating Y-STR, gọi tắt là RM Y-STR). Bộ kit PPY23 bổ sung 02 RM Y-STR là DYS570, DYS576; bộ kit YPlus bổ sung cả 7 RM Y-STR. Đây được hiểu là các locus Y-STR có tần suất đột biến cao hơn mức trung bình ($> 1 \times 10^{-3}$ theo như nhà sản xuất kit công bố). Các RM Y-STR có vai trò quan trọng trong việc phân biệt giữa các cá thể nam có quan hệ huyết thống gần gũi (ví dụ: anh trai - em trai, bố - con trai). Theo một nghiên cứu, khoảng 50% cặp bố - con trai, 60% cặp anh - em trai có thể được phân biệt bởi tổ hợp gồm 13 RM Y-STRs trong khi con số này lần lượt chỉ khoảng 7,7% và 8% nếu dùng tổ hợp 17 locus trong bộ Yfiler [43]. Đây được coi là một cải tiến lớn trong hai bộ kit. Kết quả thống kê từ bảng 3.34 cho thấy 7 locus RM Y-STR mới được bổ sung cũng đều là các locus có số lượng alen nhiều, độ đa dạng gen cao. Như vậy có thể thấy các locus mới được bổ sung vừa làm tăng khả năng phân biệt các cá thể nam trong quần thể vừa có ý nghĩa lớn trong việc phân biệt các cá thể nam có thể quan hệ họ hàng gần.

Hai bộ kit mới phát triển cũng được tối ưu hoá điều kiện phản ứng PCR bao gồm đệm, enzyme hot-start Taq ADN polymerase để khuếch đại các mẫu ADN sau tách chiết và cũng phù hợp với kiểu khuếch đại ADN trực tiếp từ nhiều loại chất nền, bao gồm mẫu

máu và niêm mạc miệng thu trên giấy FTA (GECare / Whatman) cũng như các mẫu niêm mạc miệng thu trên tấm bông chuyên dụng [34].

Để đánh giá hiệu quả của các locus mới được bổ sung trong 2 bộ kit PPY23 và YPlus chúng tôi cũng lập bảng so sánh giá trị HD, DC với các tổ hợp Y-STR khác nhau thường được dùng trong phân tích hình sự. Các tổ hợp bao gồm:

- Bộ haplotype tối thiểu (MHT - European Minimal Haplotype) bao gồm 9 locus: DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392 và DYS393.

- Bộ haplotype theo khuyến cáo của Hiệp hội Các nhà khoa học làm việc trong lĩnh vực phân tích ADN (SWGAM): bao gồm 9 locus của bộ MHT và 2 locus: DYS438, DYS439.

- Bộ AmpFLSTR™ Yfiler™ PCR Amplification Kit (gọi tắt là Yfiler, hãng ThermoFisher) được phát triển vào năm 2004 gồm 11 locus theo khuyến cáo của SWGDAM kể trên và 6 locus mới có độ đa hình cao là: DYS437, DYS448, DYS456, DYS458, Y GATA H4, DYS635.

- Bộ PPY23 và bộ YPlus: bao gồm 17 locus nêu trên và thêm các locus mới.

Bảng 3. 43. So sánh số lượng haplotype quan sát được và các chỉ số thống kê thu được từ tập hợp các marker Y-STR trong các bộ haplotype từ dữ liệu 200 mẫu với bộ YPlus

Số haplotype	Bộ haplotype	Bộ haplotype	Bộ Yfiler	Bộ
Chỉ xuất hiện ở 1 cá thể	154	168	188	200
Xuất hiện ở 2 cá thể	15	11	3	0
Xuất hiện ở 3 cá thể	1	2	2	0
Xuất hiện ở 4 cá thể	2	1	0	0
Xuất hiện ở 5 cá thể	1	0	0	0
Xuất hiện ở 6 cá thể	0	0	0	0
Tổng số locus	9	11	17	27
Tổng số haplotype quan sát	173	182	193	200
HD	0.9979	0.9988	0.9995	1.000
DC	87%	91%	97%	100%

Bảng 3. 44. So sánh số lượng haplotype quan sát được và các chỉ số thống kê thu được từ tập hợp các marker Y-STR trong các bộ haplotype từ dữ liệu 200 mẫu với bộ PPY23

Số haplotype quan sát được	Bộ haplotype tối thiểu	Bộ haplotype SWGDAM	Bộ Yfiler	Bộ PPY23
Chỉ xuất hiện ở 1 cá thể	153	169	192	200
Xuất hiện ở 2 cá thể	11	11	4	0
Xuất hiện ở 3 cá thể	5	0	0	0
Xuất hiện ở 4 cá thể	0	1	0	0
Xuất hiện ở 5 cá thể	2	1	0	0
Xuất hiện ở 6 cá thể	0	0	0	0
Tổng số locus	9	11	17	23
Tổng số haplotype quan sát	171	182	196	200
HD	0.9977	0.9986	0.9958	1.000
DC	86%	91%	98%	100%

Kết quả so sánh trong Bảng 3.43 và 3.44 cho thấy việc gia tăng số lượng các marker Y-STR trong phân tích (từ 9 marker ở bộ Haplotype Y cơ bản lên 23-27 marker ở bộ YPlus) sẽ làm gia tăng đáng kể số haplotype độc nhất (chỉ xuất hiện ở 1 cá thể nam duy nhất trong quần thể nghiên cứu), giảm số lượng các haplotype xuất hiện ở 2 người trở lên. Qua đó làm tăng tổng số haplotype quan sát được từ 171 -173 haplotype (Bộ 9 locus) -> 193 - 196 haplotype (Bộ 17 locus) lên 200 haplotype (Bộ 23 và 27 locus). Từ đó độ đa dạng haplotype và khả năng phân biệt cũng được gia tăng. Cụ thể giá trị DC tăng từ 86% ở bộ MHT-> 98% ở bộ Yfiler lên 100% ở bộ PPY23 và YPlus.

Từ các kết quả của nghiên cứu chúng tôi cũng đã mạnh dạn áp dụng việc phân tích các chỉ thị Y-STR từ hai bộ kit PPY23 và Yfiler Plus trong các công tác giám định tại Khoa Y sinh học và đạt được nhiều kết quả. Trong giám định huyết thống, các chỉ thị Y-STR được sử dụng giám định trong những trường hợp không còn mẫu người bố so sánh, xác định mối quan hệ huyết thống giữa những người cùng dòng cha như anh – em trai, ông

nội/chú/ bác – cháu trai. Đặc biệt với một số trường hợp xét nghiệm huyết thống cha – con mà có sự sai khác tại 1 hoặc 2 locus - nghi ngờ do đột biến hoặc lỗi phân tích, chúng tôi cũng áp dụng phân tích mở rộng bộ Y-STR để đảm bảo đưa ra kết luận chính xác nhất.

3.4.4. Ứng dụng chỉ thị Y-STR để so sánh khoảng cách di truyền giữa các quần thể

Khoảng cách di truyền là thước đo sự khác biệt di truyền giữa các loài hoặc giữa các quần thể trong một loài, có thể là đo khoảng cách thời gian từ tổ tiên chung hay mức độ của sự phân hóa. Các quần thể càng có nhiều alen tương tự có khoảng cách di truyền càng nhỏ. Điều này chỉ ra rằng chúng có liên quan chặt chẽ và có tổ tiên chung gần đây. Khoảng cách di truyền rất hữu ích cho việc tái cấu trúc lịch sử của quần thể, tìm hiểu được nguồn gốc chung của các quần thể phân tách.

Thước đo khoảng cách di truyền thường được sử dụng là chỉ số cố định (R_{st} – chỉ số khoảng cách di truyền quần thể theo cặp). R_{st} liên quan trực tiếp đến Phương sai (variance) về tần số alen giữa các quần thể và ngược lại với mức độ tương đồng giữa các cá thể trong quần thể. Nếu R_{st} nhỏ, điều đó có nghĩa là tần số alen giữa các quần thể rất giống nhau hay nói cách khác hai quần thể là giống nhau về mặt di truyền (rất nhỏ hoặc không có sự đa dạng di truyền giữa hai quần thể); nếu giá trị R_{st} càng lớn có nghĩa là tần số alen giữa các quần thể rất khác nhau và hai quần thể là khác biệt về mặt di truyền (đa dạng di truyền tối đa giữa hai quần thể). Do đó dựa vào bảng phân bố tần suất alen trong quần thể có thể tính toán được giá trị R_{st} để ước lượng khoảng cách di truyền giữa các quần thể khảo sát.

Do kiểu di truyền đặc trưng của nam giới và bản chất đơn bội của nhiễm sắc thể Y, các chỉ thị trên NST Y có độ nhạy cao hơn so với các chỉ thị trên NST thường trong việc đánh giá sự trôi dạt di truyền (genetic drift), khoảng cách di truyền và hiệu ứng sáng lập. Y-STR được sử dụng rộng rãi để đánh giá cấu trúc hoặc sự khác biệt trong quần thể bằng cách kiểm tra khoảng cách di truyền theo cặp. Việc phân cụm hoặc chia tách các quần thể có thể được hình dung rõ bằng cách sử dụng phân tích tỷ lệ đa chiều (MDS) dựa trên sự giống nhau của khoảng cách di truyền theo cặp giữa các quần thể. Các nghiên cứu trước đây cho thấy một số bộ haplotype, chẳng hạn như bộ haplotype tối thiểu, bộ Yfiler, bộ

PowerPlex Y23 và bộ YPlus, có thể được sử dụng để đánh giá khoảng cách di truyền giữa các quần thể [103-106].

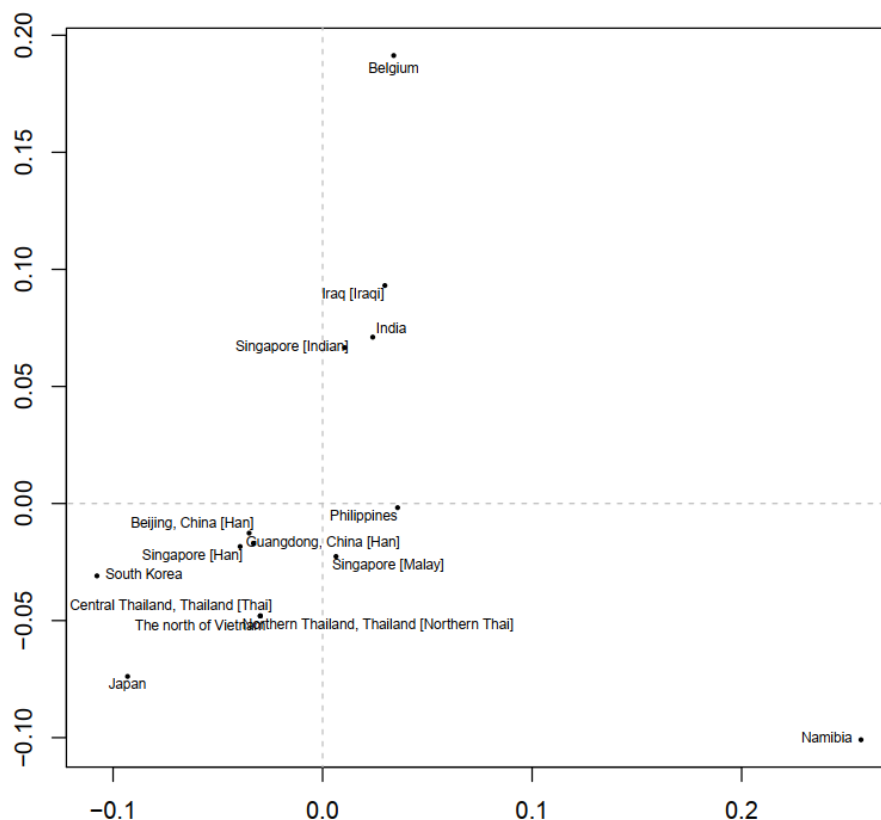
Trong nghiên cứu này công cụ online AMOVA trên www.yhrd.org được sử dụng để tính toán khoảng cách di truyền dựa vào so sánh theo cặp (giá trị R_{st}) và xác suất kết hợp (P value) giữa quần thể người Việt Nam trong nghiên cứu với số liệu đã được công bố trước đây của 13 quần thể khác trên thế giới. Danh sách các quần thể so sánh được liệt kê trong bảng 3.43. Công cụ cũng cho phép thiết lập sơ đồ minh họa khoảng cách giữa các quần thể trong không gian 2 chiều (MDS) (hình 3.22).

Kết quả trong bảng 3.43 cho thấy quần thể người Kinh (Việt Nam) trong nghiên cứu có ý nghĩa khác biệt nhất so với quần thể người Châu Phi và châu Âu (Namibia và Bỉ) với $R_{st} > 0.2$, $P < 0.001$. Ngược lại, quần thể người Kinh, Việt Nam có khoảng cách di truyền khá gần với quần thể người Thái Lan và người Hán (Trung Quốc) hơn so với các quần thể người châu Á khác. Trong đó khoảng cách di truyền giữa người Kinh (Việt Nam) và người Thái Lan ($R_{st} = 0.0085$, $P = 0.0157$) là gần gũi nhất.

Khoảng cách di truyền gần gũi giữa quần thể người Kinh, Việt Nam và người Thái Lan cũng đã được đề cập trong một số nghiên cứu. Cụ thể dự án hệ gen người Việt Nam công bố năm 2019 cho thấy người Kinh và Thái có sự tương đồng cao trong bộ gen và mối quan hệ tiến hóa chặt chẽ. Nó củng cố giả thuyết khoa học rằng những người từ Châu Phi định cư ở Đông Nam Á và sau đó di cư sâu vào lục địa từ phía nam đến phía bắc [107]. Nghiên cứu khảo sát 23 locus Y-STR trên 235 nam giới người Thái Lan của tác giả K. Vongpaisarnsin và cs (2015) cũng cho thấy khoảng cách di truyền gần gũi với quần thể người Việt Nam hơn so với các quần thể khác trong khu vực như: Singapore, Malaysia, Philipine, Trung Quốc [108].

Population	The north of Vietnam	Beijing, China [Han]	Guangdong, China [Han]	Iraq [Iraqi]	Singapore [Han]	Singapore [Indian]	Singapore [Malay]	Central Thailand Thailand [Thai]	Belgium	India	Japan	Namibia	Philippines	South Korea
The north of Vietnam	-	0	0	0	0	0	0	0.0157	0	0	0	0	0	0
Beijing, China [Han]	0.0418	-	0.0386	0	0.0637	0	0	0	0	0	0	0	0	0
Guangdong, China [Han]	0.0292	0.0031	-	0	0.0565	0	0	0	0	0	0	0	0	0
Iraq [Iraqi]	0.1849	0.1163	0.1276	-	0	0.004	0	0	0	0.0005	0	0	0	0
Singapore [Han]	0.0385	0.0058	0.0052	0.1674	-	0	0.0001	0.0005	0	0	0	0	0	0
Singapore [Indian]	0.1331	0.076	0.0881	0.0204	0.1176	-	0	0	0	0.0003	0	0	0	0
Singapore [Malay]	0.0656	0.0513	0.0374	0.1131	0.0531	0.076	-	0.0001	0	0	0	0	0	0
Central Thailand Thailand [Thai]	0.0065	0.032	0.0195	0.16	0.0225	0.1075	0.0298	-	0	0	0	0	0	0
Belgium	0.2941	0.2027	0.2049	0.0961	0.263	0.1356	0.2089	0.2681	-	0	0	0	0	0
India	0.136	0.0838	0.0899	0.0192	0.1235	0.0174	0.087	0.1199	0.1032	-	0	0	0	0
Japan	0.0836	0.0849	0.0808	0.2234	0.0917	0.1662	0.099	0.0807	0.2866	0.17	-	0	0	0
Namibia	0.3256	0.2853	0.2734	0.2724	0.3075	0.2986	0.2455	0.2869	0.4163	0.2611	0.3285	-	0	0
Philippines	0.1069	0.0665	0.0569	0.1192	0.0768	0.0828	0.0329	0.0818	0.1507	0.0939	0.1457	0.286	-	0
South Korea	0.0799	0.0558	0.0673	0.2193	0.0713	0.1609	0.1271	0.086	0.2505	0.1705	0.0444	0.3956	0.1506	-

Bảng 3. 45. Kết quả phân tích khoảng cách di truyền theo cặp dựa trên các giá trị Rst giữa quần thể Việt Nam trong nghiên cứu với các quần thể được chọn. Giá trị P được hiển thị phía trên đường chéo, giá trị rst phía dưới đường chéo



Hình 3. 22. Sơ đồ biểu thị khoảng cách di truyền giữa các quần thể trong không gian 2 chiều

Quan trọng hơn, so sánh khoảng cách di truyền giữa tập hợp bộ haplotype Y cơ bản (MHT) của 400 mẫu từ nghiên cứu này và các quần thể Việt Nam khác đã được công bố trên YHRD cho thấy sự chênh lệch giá trị Rst không đáng kể. Điều này gợi ý khoảng cách tối thiểu giữa các nhóm mẫu và chỉ ra các quần thể dân tộc trong Việt Nam nói chung tương đối đồng, ít có sai khác trong cấu trúc di truyền.

Bảng 3.46. Khoảng cách di truyền giữa các quần thể người Việt Nam

Population	The north of Vietnam	Ho-Chi-Minh City, Vietnam [Kinh]	Vietnam [Kinh]	Hanoi, Vietnam [Vietnamese]	Vietnam [Vietnamese]
The north of Vietnam	-	0.3075	0.1634	0.0331	0.6759
Ho-Chi-Minh City, Vietnam [Kinh]	0.0015	-	0.0276	0.1955	0.4549
Vietnam [Kinh] [66]	0.0029	0.0253	-	0.0574	0.0399
Hanoi, Vietnam [Vietnamese][109]	0.0193	0.0101	0.0163	-	0.0336
Vietnam [Vietnamese] [110]	-0.0016	-0.0019	0.0099	0.0207	-

Qua các phân tích trên có thể nói tiềm năng ứng dụng của các chỉ thị Y-STR là rất lớn góp phần bổ sung dữ liệu trong nhiều loại xét nghiệm ADN khác nhau. Kết quả của

luận án với tổng cộng 29 Y-STR thuộc 2 bộ kit thương mại phổ biến và 10 chỉ thị mini Y-STR đã được nghiên cứu, khảo sát đã chỉ rõ những chỉ thị Y-STR có độ đa hình và hiệu quả khuếch đại cao áp dụng trên nhiều loài là cơ sở để lựa chọn và phát triển bộ chỉ thị phân tích riêng cho nam giới người Việt Nam.

KẾT LUẬN VÀ KIẾN NGHỊ

KẾT LUẬN

Từ những kết quả nghiên cứu trong luận án chúng tôi xin đưa ra một số kết luận như sau:

1. Trong nghiên cứu, chúng tôi đã khảo sát đặc điểm của 29 chỉ thị Y-STR trong quần thể nam giới phân bố ở miền Bắc, Việt Nam. Cụ thể:

+ Bảng tần suất phản ánh những đặc trưng trong cấu trúc di truyền của quần thể nam giới người Kinh, Việt Nam khi so sánh với các quần thể khác và có ý nghĩa trong tính toán chỉ số quan hệ huyết thống theo dòng cha.

+ Kết quả khảo sát cũng cho thấy một số chỉ thị có độ đa hình cao > 0.7 (chiếm 55.2% tổng số locus khảo sát) hứa hẹn là các locus có độ đa hình cao, có tiềm năng lớn trong việc phân biệt các cá thể nam trong quần thể người Kinh, Việt Nam.

+ Kết quả khảo sát sử dụng hai bộ kit Y-STR phổ biến hiện nay là PowerPlex[®]Y23 system (Promega) và bộ Yfiler[™] Plus PCR Amplification Kit (Thermo Fisher Scientific) cho khả năng phân biệt 100% các cá thể trong tập hợp 400 mẫu nghiên cứu.

2. Xây dựng thành công quy trình khuếch đại 10 mini Y-STR là những chỉ thị có kích thước nhỏ dưới 200bp, độ đa hình cao. Kết quả phân tích với các mẫu đã bị phân hủy cho thấy tỷ lệ khuếch đại của 10 mini Y-STR dao động từ 44 % đến 82% sẽ góp phần bổ sung dữ liệu khi kết quả phân tích với chỉ thị Y-STR thông thường không thành công.

3. Kết quả khảo sát trên các mẫu thực tế cho thấy các 29 chỉ thị Y-STR và 10 mini Y-STR có tiềm năng cao trong nhiều hướng ứng dụng:

+ Trong xét nghiệm huyết thống theo dòng cha, phân tích mẫu có chất lượng kém, mẫu đã bị phân hủy, mẫu lẫn từ nhiều nguồn ADN. Tùy vào nhu cầu phân tích có thể lựa chọn loại chỉ thị Y-STR phù hợp.

+ Nghiên cứu cấu trúc di truyền quần thể: kết quả so sánh tần suất, độ phân bố 29 chỉ thị Y-STR trong nghiên cứu cho thấy cấu trúc di truyền khá đồng nhất giữa các nhóm dân tộc trong quần thể người Việt Nam và có khoảng cách di truyền gần gũi nhất với quần thể người Thái Lan.

KIẾN NGHỊ

Nhóm nghiên cứu cũng xin đưa ra một số kiến nghị để nghiên cứu được hoàn thiện trong thời gian tới:

1. Tiến hành khảo sát thêm tần suất phân bố alen các chỉ thị Y-STR với các nhóm dân tộc khác phổ biến trong quần thể người Việt Nam, đặc biệt là tần suất các locus có độ đa hình cao.
2. Phát triển phương pháp khuếch đại các mini Y-STR và các chỉ thị Y-STR có độ đa hình cao bằng môi gắn huỳnh quang và điện đi mao quản để tạo nên bộ kit riêng cho người Việt Nam.

**DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ
LIÊN QUAN ĐẾN LUẬN ÁN TIẾN SĨ**

1. **Hà Hữu Hảo**, Lê Tuấn Anh, Chu Thị Thủy, Đinh Thị Lan, Hoàng Thái Thanh, Nguyễn Thị Phương Liên, Phạm Thị Trà (2018), “Nghiên cứu tính đa hình của 6 locus Y - STR mới trong quần thể dân tộc Kinh tại Việt Nam”, *Tạp chí Y học thực hành*, số 7 (1073), 6-9.
2. **Hà Hữu Hảo**, Nguyễn Đức Nhựt, Lê Tuấn Anh, Chu Thị Thủy, Đinh Thị Lan, Hoàng Thái Thanh, Nguyễn Thị Phương Liên, Chu Hoàng Hà, Lê Văn Sơn (2018), “Đặc điểm di truyền của 27 Locus Y-STR trong quần thể người kinh, Việt Nam sử dụng bộ YPlus PCR Amplification Kit”, *Báo cáo khoa học, Hội nghị Khoa học Công nghệ sinh học toàn quốc 2018*, 564-569.
3. **Hao Huu Ha**, Trang Hong Nguyen, Linh Huyen Tran, Hanh Thi Hong Nguyen, Ha Hoang & Hoang Ha Chu (2019), “Genetic characteristics of 23 Y-chromosomal STRs in the Kinh population in Northern Vietnam”, *International Journal of Legal Medicine*, volume 133: 1403–1404.
4. **Hà Hữu Hảo**, Lê Tuấn Anh, Chu Thị Thủy, Đinh Thị Lan, Hoàng Thái Thanh, Nguyễn Thị Phương Liên (2019), “Nghiên cứu và ứng dụng các mini y-str trong giám định các mẫu đã bị phân hủy tại Viện pháp y quốc gia”, *Tạp chí Y học thực hành*, số 11 (1128), 6-9.
5. **Hao Huu Ha**, Nhu Nguyen Duc, Anh Le Tuan, Thuy Thi Chu, Lan Dinh Thi, Lien Nguyen Thi Phuong, Thanh Hoang Thai (2020), “A case study on sexual assault involving an individual with a large 24 Y chromosome deletion and two X chromosomes genotype”, *The Asian sciences network newsletter*, Issue 10: 24-25.

TÀI LIỆU THAM KHẢO

1. A. Jamieson, *Forensic biology: its past, present and future*, Biologist (London), 2000, 47 (2), 69-73.
2. A.J. Jeffreys, V. Wilson, S.L. Thein, *Individual-specific "fingerprints" of Human DNA*, Nature, 1985, 316, 76-79.
3. M. Kowalczyk, D. Szewczuk, M. Gryzińska, A. Jakubczak, et al., *Molecular markers used in forensic genetics*, Med. Sci. Law, 2018, 58 (4), 201-209.
4. A. Grover, P.C. Sharma, *Development and use of molecular markers: past and present. Crit Rev Biotechnol*, Crit. Rev. Biotechnol., 2016, 36 (2), 290–302.
5. Y. Nakamura, M. Leppert, P. O'Connell, R. Wolff, et al., *Variable number of tandem repeat (VNTR) markers for human gene mapping*, Science, 1987, 235 (4796), 1616-22.
6. A. Moller, and B. Brinkmann, *PCR-VNTRs (PCR-Variable Number of Tandem Repeats) in forensic science*, Cell Mol. Biol. (Noisy-le-grand), 1995, 41 (5), 715-724.
7. B. Sobrino, and A. Carracedo, *SNP typing in forensic genetics: a review*, Methods Mol. Biol., 2005, 297, 107-126.
8. D. Tautz, *Notes on the definition and nomenclature of tandemly repetitive DNA sequences*, EXS, 1993, 67, 21-28.
9. J.M. Butler, *Genetics and Genomics of Core Short Tandem Repeat Loci Used in Human Identity Testing*, J. Forensic Sci., 2006, 51 (2), 253-65.
10. S. Subramanian, R.K. Mishra, L. Sing, *Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions*, Genome Biol., 4 (2003), p. R13.
11. H. Ellegren, *Heterogeneous mutation processes in human microsatellite DNA sequences*, Nat. Genet., 2000, 24 (4), 400- 402.
12. C. Joel, L. Michael, T. Charles , *Forensic DNA tests and Hardy-Weinberg equilibrium*, Science, 1991, 253, 1039–1041.
13. P.C. Sharma, A. Grover, G. Kahl, *Mining microsatellites in eukaryotic genomes*, Trends Biotechnol, 2007, 25 (11), 490-498.
14. L. Roewer, J.T. Epplen, *Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work*, Forensic Science International, 1992, 53 (2), 163-71.
15. E.K. Hanson, and J. Ballantyne, *Comprehensive annotated STR physical map of the human Y chromosome: Forensic implications*, Leg. Med. (Tokyo), 2006, 8 (2), 110-120.

16. P. de Knijff, M. Kayser, A. Caglia, D. Corach, et al., *Chromosome Y microsatellites: population genetic and evolutionary aspects*, *Int. J. Legal. Med.*, 1997, 110 (3), 134-149.
17. M.A. Jobling & C. Tyler-Smith, *The human Y chromosome: An evolutionary marker comes of age*, *Nature Reviews Genetics*, 2003, 4 (8), 598–612.
18. K.S. Jeong, H. Shin, S.J. Lee, H.S. Kim, et al., *Genetic characteristics of Y-chromosome short tandem repeat haplotypes from cigarette butt samples presumed to be smoked by North Korean men*, *Genes Genomics*, 2018, 40 (8), 819-824.
19. T.M. Diegoli, *Forensic typing of short tandem repeat markers on the X and Y chromosomes*, *Forensic Sci. Int. Genet.*, 2015, 18, 140-51.
20. M. Kayser, *Forensic use of Y-chromosome DNA: a general overview*, *Hum. Genet.*, 2017, 136 (5), 621-635.
21. A. Hall, J. Ballantyne, *The development of an 18-locus Y-STR system for forensic casework*, *Anal. Bioanal. Chem.*, 2003, 376 (8), 1234–1246 .
22. E.K. Hanson, J. Ballantyne, *An ultra-high discrimination y chromosome short tandem repeat multiplex DNA typing system*, *PloS One*, 2007, 2 (8), e688.
23. B.E. Krenke, L. Viculis, M.L. Richard, M. Prinz, et al., *Validation of male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex*, *Forensic Sci. Int.*, 2005, 151, 111–124.
24. J.J. Mulero, C.W. Chang, L.M. Calandro, R.L. Green, et al., *Development and validation of the AmpFI STR Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system*, *J. Forensic Sci.*, 2006, 51 (1), 64–75.
25. J.M. Thompson, M.M. Ewing, W.E. Frank, J.J. Pogemiller, et al., *Developmental validation of the PowerPlex(R) Y23 System: a single multiplex Y-STR analysis system for casework and database samples*, *Forensic Sci. Int. Genet.*, 2013, 7 (2), 240-250.
26. S. Gopinath, C. Zhong, V. Nguyen, J. Ge, et al., *Developmental validation of the Yfiler Plus PCR amplification kit: an enhanced Y-STR multiplex for casework and database applications*, *Forensic Sci. Int. Genet.*, 2016, 24, 164–175.
27. L. Roewer, *Y chromosome STR typing in crime casework*, *Forensic Sci. Med. Pathol.*, 2009, 5 (2), 77-84.
28. L. Gusmão, J.M. Butler, A. Carracedo, P. Gill, et al., *DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis*, *Int. J. Legal. Med.*, 2006, 120 (4), 191-200.

29. P. Gill, C. Brenner, B. Brinkmann, B. Budowle, et al., *DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs*, *Forensic Sci. Int.*, 2001, 124 (1), 5-10.
30. I. Pickrahn, E. Müller, W. Zahrer, B. Dunkelmann, et al., *Yfiler® Plus amplification kit validation and calculation of forensic parameters for two Austrian populations*, *Forensic Sci. Int. Genet.*, 2016, 21, 90-94.
31. J. Purps, M. Geppert, M. Nagy, L. Roewer *Validation of a combined autosomal/Y-chromosomal STR approach for analyzing typical biological stains in sexual-assault cases*, *Forensic Sci. Int. Genet.*, 2015, 19, 238-242.
32. M. Prinz, K. Boll, H. Baum, B. Shale, *Multiplexing of Y chromosome specific STRs and performance for mixed samples*, *Forensic Sci. Int.*, 1997, 85 (3), 209-218.
33. I. Sibille, C. Duverneuil, G. Lorin de la Grandmaison, K. Guerrouache, et al., *Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa*, *Forensic Sci. Int.*, 2002, 125 (2-3), 212-216.
34. B. Ferreira-Silva, M. Fonseca-Cardoso, M.J. Porto, T. Magalhães, et al., *A Comparison Among Three Multiplex Y-STR Profiling Kits for Sexual Assault Cases*, *J. Forensic Sci.*, 2018, 63 (6), 1836-1840.
35. J. Butler, *Fundamentals of Forensic DNA Typing 1st Edition*, Academic Press, 2009, San Diego.
36. C.W. Dieffenbach, T.M. Lowe, G.S. Dveksler. *General concepts of PCR primer design*, *PCR Methods and Applications*, 1993, 3 (3), S30 – S37.
37. D. Sint, L. Raso, M. Traugott, *Advances in multiplex PCR: balancing primer efficiencies and improving detection success*, *Methods Ecol. Evol.*, 2012, 3 (5), 898-905.
38. J.M. Butler, *Constructing multiplex STR assays*, *Methods in Molecular Biology*, 2005, 297, 53 –66
39. E.B. Annelise, W.B. Harvey, *DNA separations by slab gel and capillary electrophoresis*, *Separation Purification Methods*, 1995, 24, 1-118.
40. E. Guichoux, L.Lagache , S. Wagner et al, *Current trends in microsatellite genotyping*. *Molecular Ecology Resources*, 2011, 11, 591–611.
41. G.W. Slater, C. Desruisseaux, S.J. Hubert, J.F. Mercier, et al., *Theory of DNA electrophoresis: A look at some current challenges* , *Electrophoresis*, 2000, 21(18), 3873 – 3887.

42. P. Grubwieser, R. Mühlmann, H. Niederstätter, et al., *Unusual variant alleles in commonly used short tandem repeat loci*, *International Journal of Legal Medicine*, 2005, 119(3), 164 – 166.
43. A.E. Decker, M.C. Kline, J.W. Redman, T.M. Reid, et al., *Analysis of mutations in father-son pairs with 17 Y-STR loci*, *Forensic. Sci. Int. Genet.*, 2008, 2 (3), e31-35.
44. J.M. Butler, Y. Shen, B.R. McCord, *The development of reduced size STR amplicons as tools for analysis of degraded DNA*, *J. Forensic. Sci.*, 2003, 48 (5), 1054-1064.
45. M.J Park, H.Y Lee, U. Chung, S.C Kang, et al., *Y-STR analysis of degraded DNA using reduced-size amplicons*, *International Journal of Legal Medicine*, 2007, 121(2), 152–157.
46. Y. Gu, Z. Zhang, D. Wu, L. Zha, et al., *Analysis of degraded casework DNA by redesigning a mini Y-STR multiplex.* *Forensic Science International: Genetics Supplement Series*, 2017, 6, 184-186.
47. L. Roewer, P.J. Croucher, S. Willuweit, T.T. Lu, et al., *Signature of recent historical events in the European Y-chromosomal STR haplotype distribution*, *Hum. Genet.*, 2005, 116 (4), 279 – 291.
48. J.J. Mulero, C.W. Chang, L.M. Calandro, R.L. Green, et al., *Development and validation of the AmpFlSTR Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system*, *J. Forensic Sci.*, 2006, 51 (1), 64-75.
49. C. Nuñez, M. Baeta, M. Fernández, M. Zarrabeitia, et al., *Highly discriminatory capacity of the PowerPlex(®) Y23 System for the study of isolated populations*, *Forensic Sci. Int. Genet.*, 2015, 17, 104-107.
50. T. Jain, P. Shrivastava, D.D. Bansal, H.R. Dash, et al., *PowerPlex Y23 System: A Fast, Sensitive and Reliable Y-STR Multiplex System for Forensic and Population Genetic Purpose*. *J. Mol. Biomark. Diagn.*, 2016, 7, 3.
51. M.M. Andersen, H.S. Mogensen, P.S. Eriksen, N. Morling, *Yfiler® Plus population samples and dilution series: stutters, analytic thresholds, and drop-out probabilities*, *Int. J. Legal Med.*, 2017, 131 (6), 1503-1511.
52. M. Bodner, I. Bastisch, J.M. Butler, R. Fimmers, et al., *Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER)*, *Forensic Sci. Int. Genet.*, 2016, 24, 97-102.

53. National Research Council (US) Committee on DNA Forensic Science: An Update. *The Evaluation of Forensic DNA Evidence* (chapter 4). Washington (DC), National Academies Press (US), 1996.
54. Scientific Working Group on DNA Analysis Methods. *SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories*, Approved 01/09/2014.
55. L. Roewer, M. Krawczak, S. Willuweit, M. Nagy, et al., *Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes*, *Forensic Sci. Int.*, 2001, 118, 106-13.
56. L. Roewer, *Y-chromosome short tandem repeats in forensics—Sexing, profiling, and matching male DNA*, *WIREs Forensic Sci.*, 2019, 1, e1336.
57. A.E. Kofi, H.M. Hakim, H.O. Khan, S.A. Ismail, et al., *Population data of 23 Y chromosome STR loci for the five major human subpopulations of Ghana*, *Int. J. Legal Med.*, 2019, [Epub ahead of print].
58. Atif DNAan, K. Kasim, A. Rakha, A. Noor, et al., *Population data of 23 Y STRs from Manchu population of Liaoning Province, Northeast China*, *Int. J. Legal Med.*, 2019, 133 (3), 785-788.
59. R. Bai, Y. Liu, J. Zhang, M. Shi, et al., *Analysis of 27 Y-chromosomal STR haplotypes in a Han population of Henan province, Central China*, *Int. J. Legal Med.*, 2016, 130 (5), 1191-4.
60. S.K. Lim, Y. Xue, E.J. Parkin, C. Tyler-Smith, *Variation of 52 new Y-STR loci in the Y Chromosome Consortium worldwide panel of 76 diverse individuals*, *Int. J. Legal Med.*, 2007, 121(2), 124-127.
61. M.E. D'Amato, V.B. Bajic, S. Davison, *Design and validation of a highly discriminatory 10-locus Y-chromosome STR multiplex system*, *Forensic Sci. Int. Genet.*, 2011, 5(2), 122-125.
62. F. Calafell, M.H.D. Larmuseau, *The Y chromosome as the most popular marker in genetic genealogy benefits interdisciplinary research*, *Hum. Genet.*, 2017, 136 (5), 559-573.
63. X. Qian, J. Hou, Z. Wang, Y. Ye, et al., *Next Generation Sequencing Plus (NGS+) with Y-chromosomal Markers for Forensic Pedigree Searches*. *Scientific Reports*, 2017, 7 (11324).
64. K. Dewa, N.Q. Tuyen, S. Rand, C. Hohoff, et al., *13 Y-chromosomal STRs in a Vietnamese population*, *International Congress Series*, 2003, 1239, 315-317.

65. V. L. Loi, *Tính toán tần suất haplotype hệ 17 Y-STR phục vụ công tác giám định tại Viện Khoa học hình sự*, Luận án thạc sĩ khoa học sinh học, Viện sinh thái và tài nguyên sinh vật, Đại học Thái Nguyên, 2013.
66. F. Miranda-Barros, C. Romanini, L.A. Pérez, N.D. Nhu, et al., *Y Chromosome STR haplotypes in different ethnic groups of Vietnam*, Forensic Sci. Int. Genet., 2016, 22, e18-e20.
67. P.S. Walsh, D.A. Metzger, R. Higuchi. *Chelex® 100 as a Medium for Simple Extraction of DNA for PCR-Based Typing from Forensic Material*, Biotechniques, 1991, 10(4), 506-513.
68. Promega, Corp., *PowerPlex® Y23 System for Use on the Applied Biosystems® Genetic Analyzers*. Technical manual, Revised 2019.
69. Thermo Fisher Scientific., *Yfiler™ Plus PCR Amplification Kit User Guide*.
70. M. Nei, *Molecular Evolutionary Genetics*, Columbia University Press, 1987 New York.
71. C. Núñez, M. Baeta, N. Ibarbia, U. Ortueta, S. Jiménez-Moreno, et al., *17 to 23: A novel complementary mini Y-STR panel to extend the Y-STR databases from 17 to 23 markers for forensic purposes*. Electrophoresis, 2017, 38(7), 1016-1021.
72. H.C. Park, E.J. Lee, Y.H. Nam. *Genetic variation for three Y-STR loci: DYS390, DYS518, and DYS643*. Int J Legal Med, 2019;133(4), 1007-1013.
73. P.M. Schneider, S. Meuser, W. Waiyawuth, et al. *Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations*. Forensic Sci Int, 1998 12;97(1), 61-70.
74. Leclair, B., et al, *Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples*. Journal of Forensic Sciences, 2004, 49, 968– 980.
75. L. Roewer et al. *Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts*. Hum. Genet, 1992, 89, 389–394.
76. I. W. Evett, B. S. Weir., *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists*. Sunderland, MA: Sinauer Associates, 1998.
77. W. K. Fung, Y. Q. Hu, *Statistical DNA Forensics: Theory, Methods and Computation*, New Jersey: John Wiley & Sons, 2008.
78. T.T. Toàn, H.Q. Khanh, V.L. Lợi, N.Q. Vinh, *Nghiên cứu đa dạng Haplotype Y- STR của quần thể người Kinh - Việt Nam*. Tạp chí Khoa học ĐHQGHN: Khoa học Tự nhiên và Công nghệ, 2017, 33, 212-218.

79. Jie Zhaoa, Xiaoye Jinb, Zhen Songa, Sanping Song. *Genetic distributions and diversity analyses of 23 Y-STR loci in Xi'an Han population*. Legal Medicine, 2019, 41, 101635.
80. T. Zeyad, A. Adam., R. Alghafri., R. Iratni., *Study of 27 Y-STR markers in United Arab Emirates population, Forensic Science International, 2020, Reports 2.*
81. Y. Posada, M.E. Sánchez, J.D. Granda, et al. *Case report: Male phenotype with incomplete Y chromosome and X chromosome double dosage*. Forensic Sci. Int. Genet. Suppl. Ser, 2017, 6, e573-574.
82. J. Purps, S. Siegert, S. Willuweit et al., *A Global Analysis of Y-chromosomal Haplotype Diversity for 23 STR Loci*. Forensic Sci Int Genet, 2014, 12(100), 12-23.
83. F. Miranda-Barros, A. Gorostiza, E. F. Carvalh, A. González-Martín, L. Gusmão., *Y-STR haplotype background of Philippines: Comparison with other Southeast Asian populations*. Forensic Sci. Int. Genet. Suppl. Ser, 2015, e428-e429.
84. Y. Zhou, C. Shao, L. Li, Y. Zhang et al., *Genetic Analysis of 29 Y-STR Loci in the Chinese Han Population from Shanghai*. Forensic Sci Int Genet, 2018, 32, e1-e4.
85. S. Hao, X. Zhang, Y. Liu, D. Lu., *Genetic diversities of 23 Y-Chromosome short tandem repeat loci in a han population in the Beijing Region*. J Forensic Sci Me, 2018, 4, 111-114.
86. B. Martinez, L. Catelli, M. Romero, V.O. Okolie, S.O. Keshinro, et al., *Forensic evaluation of 27 y-str haplotypes in a population sample from Nigeria*. Forensic Sci Int Genet, 2017, 6, e289-e291.
87. Y.Chen, X.Yang, X.Pan, P. Li, Y.Tang & C. Liu., *Validation of a 6 Dye Mini STR kit with 17 loci*, Australian Journal of Forensic Sciences, 2019.
88. V. Lopes, L. Andrade, M. Carvalho, A. Serra, et al., *Mini-STRs: A powerful tool to identify genetic profiles in samples with small amounts of DNA*. Forensic Sci Int Genet Supplement Series 2, 20099, 121–122.
89. C.M. Constantinescu, L.E. Barbarii, C.B. Iancu, A. Constantinescu et al., *Challenging DNA samples solved with MiniSTR analysis*. Brief overview. Rom J Leg Med, 2012, 20, 51-56.
90. T. Martins, M.F. Pinheiro., *Y-miniSTR: New application for compromised samples (population data)*. Forensic Sci Int Genet: Genetics Supplement Series 1, 2008, 222–223.
91. H Asamura, H Sakai, M Ota, H Fukushima., *MiniY-STR quadruplex systems with short amplicon lengths for analysis of degraded DNA samples*, Forensic Sci Int Genet, 2007, 1(1):56-61

92. L. Yanfang, Y. Tingting, M. Shuyan, J. Xiaoye, L. Qiong. *Forensic characteristics and genetic affinity analyses of Xinjiang Mongolian group using a novel six fluorescent dye-labeled typing system including 41 Y-STRs and 3 Y-InDels*, Mol Genet Genomic Med, 2020, 8:e1097.
93. F. Makki-Rmida, A.Kammoun 1, N.Mahfoudh., et al , Genetic diversity and haplotype structure of 21 Y-STRs, including nine noncore loci, in South Tunisian Population: Forensic relevance, Electrophoresis 2015 Dec;36(23):2908-13
94. T. Sathirapatyaa , P. Sukawutthiyaa , K. Vongpaisarnsin., *Massively parallel sequencing of 24 Y-STR loci in Thai population*, Forensic Sci Int Genet: Genetics Supplement Series 6, 2017, e310–e313.
95. H. Koyama, M.Iwasa, T. Tsuchimochi, Y. Maeno et al., *Utility of Y-STR haplotype and mtDNA sequence in personal identification of human remains*. Am J Forensic Med Pathol, 2002, 23(2), 181-185.
96. A. Ambers, J. Votrubova, D. Vanek, A. Sajantila., *Improved Y-STR Typing for Disaster Victim Identification, Missing Persons Investigations, and Historical Human Skeletal Remains*. Int J Legal Med, 2018, 132,6,1545-1553.
97. J.Olasz, V. Seidenberg, S. Hummel, Z. Szentirmay. *DNA profiling of Hungarian King Béla III and other skeletal remains originating from the Royal Basilica of Székesfehérvá*. Archaeological and Anthropological Sciences,2019, 11, 1345–1357.
98. J.Y. Jung, J.H. Park, Y.L. Oh, H.S. Kwon., et al. *Forensic genetic study of 29 Y-STRs in Korean population*. Legal Medicine, 2016, 23,17–20.
99. J. Tokarz, S. Foster, S.Milne, P. Staton et al., *Validation of PowerPlex® Y23 PCR Amplification System for Forensic Casework*. Marshall University, Forensic science graduate program.
100. S.Gopinath, C. Zhonga, V. Nguyen, J. Ge et al. *Developmental validation of the Yfiler® Plus PCR Amplification Kit: An enhanced Y-STR multiplex for casework and database applications*. Forensic Sci Int Genet, 2016, 24, 164-175.
101. J. Henry, L. Scandrett., *Assessment of the Yfiler® Plus PCR amplification kit for the detection of male DNA in semen-negative sexual assault cases*. Science & Justice, 2019,59, 480–485.
102. T. Gao, L. Yuna, Y. Gua, W. Hea, Z. Wang, Y.Hou., *Phylogenetic analysis and forensic characteristics of 12 populations using 23 Y-STR loci*. Forensic Sci. Int. Genet, 2015, 19, 130–133.

103. Z. Wang, W. Du, G. He, J. Liu, Y. Hou., *Forensic characteristics and phylogenetic analysis of Hubei Han population in central China using 17 Y-STR loci*. Forensic Sci. Int. Genet, 2017, 29, e4–e8.
104. M. Singh, A. Sarkar, M.R. Nandineni., *A comprehensive portrait of Y-STR diversity of Indian populations and comparison with 129 worldwide populations*. Sci. Rep, 2018, 8, 2–8.
105. G.L He, P.Y Chen, X. Zou, X. Chen, F. Song, J. Yan, Y.P. Hou., *Genetic polymorphism investigation of the Chinese Yi minority using PowerPlex® Y23 STR amplification system*. Int. J. Leg. Med, 2017, 131, 663–666.
106. H. Wang, J. Mao, Y. Xia, X. Bai, W. Zhu, D. Peng, W. Liang., *Genetic polymorphisms of 17 Y-chromosomal STRs in the Chengdu Han population of China*. Int. J. Leg. Med, 2017, 131, 967–968.
107. L.S.Vinh, T.T. Kien, B.T.P. Hoa, L.T.T. Huong., *A Vietnamese human genetic variation database*. Human Mutation, 2019, 1–12.
108. K. Vongpaisarnsin, P. Saengkaetrakul, A. Boonlert, N. Tansrisawad., *Gene and haplotype diversity of 23 Y-chromosomal short tandem repeat loci in the central Thai population*, Forensic Sci Int Genet, 2015, 14, 191-193.
109. KD. Kwak, HJ. Jin., DJ. Shin, JM. Kim, L. Roewer, M. Krawczak, C.Tyler Smith, W. Kim., *Y-chromosomal STR haplotypes and their applications to forensic and population studies in east Asia*, Int J Legal Med, 2005, 119(4), 195-201
110. R. Lessig, S. Willuweit, M.Krawczak, FC.Wu, et al., *Asian online Y-STR Haplotype Reference Database*, Leg Med (Tokyo) 5 Suppl 1, S160-3.
111. JS Buckleton., CM Triggs., SJ Walsh. *Forensic DNA evidence interpretation*, 2005, 1st ed., CRC press.

PHẦN PHỤ LỤC

PHỤ LỤC 1

Thông tin các chỉ thị Y-STR sử dụng trong nghiên cứu, ^a trong bộ PPY23, ^b trong bộ YPlus, ^c tần số đột biến được lấy theo thống kê trên website: yhrd.org

TT	Y-STR	Cấu trúc lặp	Khoảng alen	Kích thước trong bộ Y23	Kích thước trong bộ Yplus (bp)	Tần số đột biến
1	DYS19	TAGA	9 - 19	312-352	184-224	0.00212
2	DYS385a/ b	GAAA	7 - 28	223-307	225-302	0.00286
3	DYS389I	(TCTG) (TCTA)	9 - 17	147-179	147-179	0.00244
4	DYS389II	(TCTG) (TCTA)	24 - 35	259-303	265-309	0.00464
5	DYS390	(TCTA) (TCTG)	17 - 29	207-255	144-192	0.00203
6	DYS391	TCTA	5 - 16	86-130	353-398	0.00240
7	DYS392	TAT	4 - 20	314-362	274-322	0.00052
8	DYS393	AGAT	7 - 18	101-145	90-135	0.00122
9	DYS437	TCTA	11-18 ^a 10-18 ^b	344-380	178-210	0.00132
10	DYS438	TTTTTC	6-16	293-343	207-258	0.00035
11	DYS439	AGAT	6-17	263-307	150-195	0.00515
12	DYS448	AGAGAT	14-24	196-256	278-338	0.00137
13	DYS456	AGAT	11-23 ^a 10-24 ^b	316-364	76-134	0.00430
14	DYS458	GAAA	10-24 ^a 11-24 ^b	159-215	120-171	0.00660
15	DYS635	(TGTA) (TCTA)	15-28 ^a 15-30 ^b	150-202	191-252	0.00424
16	Y-GATA- H4	TAGA	8-18 ^a 8-15 ^b	374-414	236-264	0.00301
17	DYS481	CTT	17-32	139-184	207-252	0.00458

TT	Y-STR	Cấu trúc lặp	Khoảng alen	Kích thước trong bộ Y23	Kích thước trong bộ Yplus (bp)	Tần số đột biến
18	DYS533	ATCT	7-17	245-285	338-379	0.00294
19	DYS549	GATA	7-17	198-238		0.00407
20	DYS570	TTTC	10-25 ^a 10-26 ^b	90-150	98-163	0.00922
21	DYS576	AAAG	11-23 ^a 10-25 ^b	97-145	72-133	0.0121
22	DYS643	CTTTT	6-17	368-423		0.00103
23	DYS449	TTTC	22-42		325-398	0.0102
24	DYS460	ATAG	6-14		79-109	0.00441
25	DYS518	AAAG	28-48		332-401	0.0119
26	DYS627	AAAG	11-28		324-387	0.0147
27	DYF387S 1a/b	RAAG	29-44		264-319	0.00791
28	DYS505	TCCT	9-15			0.003
29	DYS508	TATC	8-15			0.0031
30	DYS522	GATA	8-17			<10 ⁻⁵
31	DYS388	ATT	10-18			0.000378

PHỤ LỤC 2

Danh sách vị trí thu thập 30 mẫu hài cốt đã bị phân huỷ

2.1. Mẫu thu thập tại nghĩa trang hữu nghị Việt Lào, huyện Anh Sơn, tỉnh Nghệ An

STT	Ký hiệu mẫu	Vị trí quy tập mộ
1	AS_51B5	Nghĩa trang Nam Phu Nốc Cốc, bản Phu Nốc Cốc, huyện Mường Khăm, tỉnh Xiêng Khoảng
2	AS_111B5	Nghĩa trang Noọng Pừ, bản Hối 1, huyện Mường Khùn, tỉnh Xiêng Khoảng
3	AS_151B5	Nghĩa trang 1 Thăm Tàu, bản Phu Pha Xay, huyện Phả Xay, tỉnh Xiêng Khoảng
4	AS_161B5	Nghĩa trang 1 Thăm Tàu, bản Phu Pha Xay, huyện Phả Xay, tỉnh Xiêng Khoảng.
5	AS_221B5	Nghĩa trang Bản Xao 2, bản Xao 2, huyện Mường Pệc, tỉnh Xiêng Khoảng
6	AS_72B5	Nghĩa trang Na Hong, bản Thà Viêng, huyện Thà Thôm, tỉnh Xiêng Khoảng.
7	AS_82B5	Suối Nậm Hán, bản Nậm Hán, huyện Mường Khăm, Xiêng Khoảng
8	AS_172B5	Nghĩa trang bản Hòm, bản Hữu Kiệm, huyện Kỳ Sơn, tỉnh Nghệ An
9	AS_233B5	Nghĩa trang Thăm Khúm, Pả Chay, bản Na Xả La, Noọng Pét, huyện Mường pệc, tỉnh Xiêng Khoảng.
10	AS_263B5	Nghĩa trang Phu Khí Mốt, bản Phôn, huyện Phả Xay, tỉnh Xiêng Khoảng
11	AS_303B5	Nghĩa trang 1 Đông sân bay Loong Chặng, bản Loong Chặng, huyện Xây Xôm Bun, tỉnh Viêng Chăn.
12	AS_343B5	Nghĩa trang 1 Đông sân bay Loong Chặng, bản Loong Chặng, huyện Xây Xôm Bun, tỉnh Viêng Chăn
13	AS_373B5	Nghĩa trang 1 Đông sân bay Loong Chặng, bản Loong Chặng, huyện Xây Xôm Bun, tỉnh Viêng Chăn
14	AS_383B5	Nghĩa trang 1 Đông sân bay Loong Chặng, bản Loong Chặng, huyện Xây Xôm Bun, tỉnh Viêng Chăn
15	AS_403B5	Nghĩa trang Phu Thuông, bản Na Đác, huyện Mường Khùn, tỉnh Xiêng Khoảng

2.2. Tại nghĩa trang liệt sỹ huyện Định Hóa, tỉnh Thái Nguyên

STT	Ký hiệu	Vị trí quy tập mộ
1	MIH1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên

2	M3H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
3	M5H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
4	M6H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
5	M8H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
6	M9H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
7	M10H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
8	M11H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
9	M13H1KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
10	M18H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
11	M19H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
12	M20H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
13	M21H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
14	M22H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên
15	M29H2KA	NTLS Huyện Định Hóa, tỉnh Thái Nguyên

PHỤ LỤC 3: Danh sách các quần thể so sánh khoảng cách di truyền

Population	Accession number
Kinh (Northern Vietnam)	YA004576
Han (Beijing PRC)	YA004160, YA003197
Han (Guangdong PRC)	YA004066
Han (Singapore)	YA003270
Indian (Singapore)	YA003271, YA004368
Malaysian (Singapore)	YA003272, YA004367
Thai (Central Thailand)	YA003923, YA003414

PHỤ LỤC 4: Nồng độ ADN tổng số 400 mẫu tách chiết bằng phương pháp Chelex 100 (nồng độ đo bằng máy Quantus Fluorometer, hãng Promega - Mỹ).

Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số	Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số	Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số
ID01	Tóc	5.3	ID51	Tóc	2.7	ID101	Tóc	4.2
ID02	Tóc	4.2	ID52	Tóc	3.2	ID102	Tóc	3.5
ID03	Tóc	5.8	ID53	Tóc	3.6	ID103	Tóc	3.6
ID04	Tóc	6.5	ID54	Tóc	5.6	ID104	Tóc	6.2
ID05	Tóc	3.6	ID55	Tóc	4.2	ID105	Tóc	5.3
ID06	Tóc	5.6	ID56	Tóc	3.5	ID106	Tóc	4.2
ID07	Tóc	4.2	ID57	Tóc	3.6	ID107	Tóc	5.8
ID08	Tóc	3.5	ID58	Tóc	6.2	ID108	Tóc	6.5
ID09	Tóc	3.6	ID59	Tóc	5.3	ID109	Tóc	3.6
ID10	Tóc	6.2	ID60	Tóc	4.2	ID110	Tóc	5.6
ID11	Tóc	5.3	ID61	Tóc	5.8	ID111	Tóc	4.2
ID12	Tóc	4.2	ID62	Tóc	6.5	ID112	Tóc	3.5
ID13	Tóc	5.8	ID63	Tóc	4.5	ID113	Tóc	3.6
ID14	Tóc	3.6	ID64	Tóc	3.8	ID114	Tóc	5.8
ID15	Tóc	5.8	ID65	Tóc	3.9	ID115	Tóc	6.5
ID16	Tóc	6.5	ID66	Tóc	7.2	ID116	Tóc	4.5
ID17	Tóc	4.5	ID67	Tóc	5.6	ID117	Tóc	3.8
ID18	Tóc	3.8	ID68	Tóc	6.3	ID118	Tóc	3.9
ID19	Tóc	3.9	ID69	Tóc	6.2	ID119	Tóc	7.2
ID20	Tóc	7.2	ID70	Tóc	6.4	ID120	Tóc	5.6
ID21	Tóc	5.6	ID71	Tóc	3.8	ID121	Tóc	6.3
ID22	Tóc	6.3	ID72	Tóc	4.5	ID122	Tóc	4.5
ID23	Tóc	3.8	ID73	Tóc	2.9	ID123	Tóc	3.8
ID24	Tóc	3.9	ID74	Tóc	6.3	ID124	Tóc	3.9
ID25	Tóc	7.2	ID75	Tóc	5.8	ID125	Tóc	7.2
ID26	Tóc	5.6	ID76	Tóc	9.3	ID126	Tóc	6.2
ID27	Tóc	6.3	ID77	Tóc	8.3	ID127	Tóc	5.3
ID28	Tóc	4.5	ID78	Tóc	7.5	ID128	Tóc	4.2
ID29	Tóc	3.8	ID79	Tóc	6.5	ID129	Tóc	5.8
ID30	Tóc	3.9	ID80	Tóc	8.5	ID130	Tóc	6.5
ID31	Tóc	7.2	ID81	Tóc	4.6	ID131	Tóc	4.5
ID32	Tóc	5.8	ID82	Tóc	5.6	ID132	Tóc	4.5
ID33	Tóc	6.5	ID83	Tóc	7.5	ID133	Tóc	2.9
ID34	Tóc	4.5	ID84	Tóc	4.6	ID134	Tóc	6.3
ID35	Tóc	4.5	ID85	Tóc	8.5	ID135	Tóc	5.8
ID36	Tóc	2.9	ID86	Tóc	4.5	ID136	Tóc	9.3
ID37	Tóc	6.3	ID87	Tóc	7.5	ID137	Tóc	8.3
ID38	Tóc	5.8	ID88	Tóc	6.5	ID138	Tóc	5.6
ID39	Tóc	9.3	ID89	Tóc	8.8	ID139	Tóc	7.5
ID40	Tóc	8.3	ID90	Tóc	8.9	ID140	Tóc	4.6

ID41	Tóc	5.6	ID91	Tóc	7.5	ID141	Tóc	8.5
ID42	Tóc	9.3	ID92	Tóc	8.5	ID142	Tóc	4.5
ID43	Tóc	8.3	ID93	Tóc	4.6	ID143	Tóc	7.5
ID44	Tóc	5.6	ID94	Tóc	5.6	ID144	Tóc	6.5
ID45	Tóc	7.5	ID95	Tóc	7.5	ID145	Tóc	6.5
ID46	Tóc	4.6	ID96	Tóc	5.6	ID146	Tóc	6.5
ID47	Tóc	8.5	ID97	Tóc	5.6	ID147	Tóc	8.5
ID48	Tóc	4.5	ID98	Tóc	7.5	ID148	Tóc	4.6
ID49	Tóc	7.5	ID99	Tóc	4.6	ID149	Tóc	5.6
ID50	Tóc	6.5	ID100	Tóc	8.5	ID150	Tóc	7.5
Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số	Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số	Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số
ID151	Niêm mạc	7.5	ID201	Niêm mạc	8.9	ID251	Niêm mạc	5.8
ID152	Niêm mạc	8.7	ID202	Niêm mạc	7.5	ID252	Niêm mạc	6.7
ID153	Niêm mạc	4.6	ID203	Niêm mạc	9.2	ID253	Niêm mạc	8.9
ID154	Niêm mạc	5.8	ID204	Niêm mạc	4.5	ID254	Niêm mạc	7.5
ID155	Niêm mạc	6.7	ID205	Niêm mạc	10.5	ID255	Niêm mạc	9.2
ID156	Niêm mạc	8.9	ID206	Niêm mạc	7.5	ID256	Niêm mạc	4.5
ID157	Niêm mạc	5.8	ID207	Niêm mạc	5.8	ID257	Niêm mạc	8.6
ID158	Niêm mạc	6.7	ID208	Niêm mạc	6.7	ID258	Niêm mạc	7.5
ID159	Niêm mạc	10.5	ID209	Niêm mạc	8.9	ID259	Niêm mạc	5.8
ID160	Niêm mạc	7.5	ID210	Niêm mạc	7.5	ID260	Niêm mạc	6.7
ID161	Niêm mạc	9.2	ID211	Niêm mạc	9.2	ID261	Niêm mạc	8.9
ID162	Niêm mạc	4.5	ID212	Niêm mạc	4.5	ID262	Niêm mạc	7.5
ID163	Niêm mạc	8.6	ID213	Niêm mạc	8.6	ID263	Niêm mạc	9.2
ID164	Niêm mạc	7.5	ID214	Niêm mạc	7.5	ID264	Niêm mạc	4.5
ID165	Niêm mạc	9.3	ID215	Niêm mạc	5.8	ID265	Niêm mạc	8.6
ID166	Niêm mạc	8.5	ID216	Niêm mạc	6.7	ID266	Niêm mạc	10.5
ID167	Niêm mạc	6.2	ID217	Niêm mạc	8.9	ID267	Niêm mạc	7.5
ID168	Niêm mạc	8.6	ID218	Niêm mạc	7.5	ID268	Niêm mạc	9.2
ID169	Niêm mạc	4.6	ID219	Niêm mạc	9.2	ID269	Niêm mạc	4.5
ID170	Niêm mạc	7.5	ID220	Niêm mạc	4.5	ID270	Niêm mạc	8.6
ID171	Niêm mạc	6.5	ID221	Niêm mạc	8.6	ID271	Niêm mạc	7.5
ID172	Niêm mạc	8.5	ID222	Niêm mạc	7.5	ID272	Niêm mạc	9.3
ID173	Niêm mạc	9.5	ID223	Niêm mạc	7.5	ID273	Niêm mạc	8.5
ID174	Niêm mạc	8.6	ID224	Niêm mạc	9.2	ID274	Niêm mạc	6.2
ID175	Niêm mạc	4.6	ID225	Niêm mạc	9.3	ID275	Niêm mạc	8.6
ID176	Niêm mạc	7.8	ID226	Niêm mạc	8.5	ID276	Niêm mạc	4.6
ID177	Niêm mạc	9.8	ID227	Niêm mạc	6.2	ID277	Niêm mạc	8.5
ID178	Niêm mạc	9.5	ID228	Niêm mạc	8.6	ID278	Niêm mạc	6.2

ID179	Niêm mạc	9.3	ID229	Niêm mạc	10.5	ID279	Niêm mạc	8.6
ID180	Niêm mạc	8.6	ID230	Niêm mạc	8.5	ID280	Niêm mạc	4.6
ID181	Niêm mạc	7.6	ID231	Niêm mạc	6.2	ID281	Niêm mạc	7.5
ID182	Niêm mạc	8.5	ID232	Niêm mạc	8.6	ID282	Niêm mạc	6.5
ID183	Niêm mạc	8.2	ID233	Niêm mạc	4.6	ID283	Niêm mạc	8.5
ID184	Niêm mạc	8.4	ID234	Niêm mạc	7.5	ID284	Niêm mạc	9.5
ID185	Niêm mạc	8.6	ID235	Niêm mạc	6.5	ID285	Niêm mạc	8.6
ID186	Niêm mạc	7.6	ID236	Niêm mạc	6.2	ID286	Niêm mạc	8.6
ID187	Niêm mạc	7.8	ID237	Niêm mạc	8.6	ID287	Niêm mạc	4.6
ID188	Niêm mạc	10.5	ID238	Niêm mạc	4.6	ID288	Niêm mạc	7.8
ID189	Niêm mạc	5.6	ID239	Niêm mạc	7.5	ID289	Niêm mạc	9.8
ID190	Niêm mạc	7.6	ID240	Niêm mạc	6.5	ID290	Niêm mạc	9.5
ID191	Niêm mạc	8.6	ID241	Niêm mạc	8.5	ID291	Niêm mạc	9.3
ID192	Niêm mạc	8.7	ID242	Niêm mạc	9.5	ID292	Niêm mạc	8.6
ID193	Niêm mạc	7.5	ID243	Niêm mạc	8.6	ID293	Niêm mạc	7.6
ID194	Niêm mạc	7.5	ID244	Niêm mạc	8.6	ID294	Niêm mạc	8.5
ID195	Niêm mạc	7.1	ID245	Niêm mạc	4.6	ID295	Niêm mạc	8.2
ID196	Niêm mạc	7.4	ID246	Niêm mạc	7.6	ID296	Niêm mạc	5.6
ID197	Niêm mạc	7.3	ID247	Niêm mạc	7.8	ID297	Niêm mạc	7.6
ID198	Niêm mạc	7.5	ID248	Niêm mạc	4.6	ID298	Niêm mạc	8.6
ID199	Niêm mạc	7.9	ID249	Niêm mạc	5.6	ID299	Niêm mạc	8.7
ID200	Niêm mạc	7.6	ID250	Niêm mạc	7.6	ID300	Niêm mạc	7.5
Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số	Ký hiệu mẫu	Loại mẫu	Nồng độ ADN tổng số			
ID301	Máu	2.8	ID351	Máu	3.6			
ID302	Máu	3.6	ID352	Máu	3.4			
ID303	Máu	3.4	ID353	Máu	3.5			
ID304	Máu	3.5	ID354	Máu	3.4			
ID305	Máu	3.8	ID355	Máu	5.6			
ID306	Máu	5.6	ID356	Máu	3.6			
ID307	Máu	3.4	ID357	Máu	3.8			
ID308	Máu	5.6	ID358	Máu	6.1			
ID309	Máu	3.6	ID359	Máu	5.3			
ID310	Máu	3.8	ID360	Máu	5.2			
ID311	Máu	6.1	ID361	Máu	4.5			
ID312	Máu	5.3	ID362	Máu	4.6			
ID313	Máu	5.2	ID363	Máu	3.5			
ID314	Máu	4.5	ID364	Máu	3.8			
ID315	Máu	4.6	ID365	Máu	5.6			
ID316	Máu	5.2	ID366	Máu	3.4			

ID317	Máu	3.6	ID367	Máu	5.2
ID318	Máu	3.5	ID368	Máu	3.6
ID319	Máu	3.4	ID369	Máu	3.5
ID320	Máu	3.8	ID370	Máu	3.4
ID321	Máu	3.7	ID371	Máu	3.8
ID322	Máu	4.5	ID372	Máu	3.7
ID323	Máu	6.2	ID373	Máu	4.5
ID324	Máu	2.7	ID374	Máu	2.9
ID325	Máu	6.5	ID375	Máu	6.3
ID326	Máu	6.4	ID376	Máu	6.5
ID327	Máu	6.5	ID377	Máu	2.7
ID328	Máu	6.1	ID378	Máu	6.5
ID329	Máu	5.3	ID379	Máu	6.1
ID330	Máu	5.2	ID380	Máu	5.3
ID331	Máu	3.5	ID381	Máu	5.2
ID332	Máu	3.4	ID382	Máu	3.5
ID333	Máu	3.2	ID383	Máu	3.4
ID334	Máu	3.5	ID384	Máu	3.2
ID335	Máu	4.2	ID385	Máu	2.7
ID336	Máu	4.6	ID386	Máu	6.4
ID337	Máu	2.7	ID387	Máu	6.5
ID338	Máu	4.6	ID388	Máu	2.9
ID339	Máu	4.2	ID389	Máu	5.3
ID340	Máu	4.1	ID390	Máu	4.6
ID341	Máu	4.8	ID391	Máu	5.2
ID342	Máu	5.2	ID392	Máu	4.6
ID343	Máu	3.1	ID393	Máu	4.2
ID344	Máu	3.5	ID394	Máu	4.1
ID345	Máu	3.8	ID395	Máu	4.8
ID346	Máu	3.5	ID396	Máu	5.2
ID347	Máu	6.2	ID397	Máu	3.1
ID348	Máu	6.4	ID398	Máu	2.8
ID349	Máu	5.2	ID399	Máu	3.8
ID350	Máu	4.6	ID400	Máu	3.5

PHỤ LỤC 5: Nồng độ ADN tổng số 30 mẫu tách chiết bằng kit (nồng độ đo sử dụng kit Trio ADN Quantification Kits, hãng Thermo Fisher Scientific – Mỹ trên hệ thống máy Real time PCR)

Ký hiệu mẫu	ADN đích	Nồng độ ADN	Chỉ số phân hủy	Ký hiệu mẫu	ADN đích	Nồng độ ADN	Chỉ số phân hủy
M1H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0572 0.0787 0.0721	1.376	AS_51B5	T.Large Autosomal T.Small Autosomal T.Y	0.047852 0.07685 0.0721	1.606
M3H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0542 0.1028 0.0974	1.897	AS_111B5	T.Large Autosomal T.Small Autosomal T.Y	0.0635 0.1223 0.0675	1.926
M5H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0718 0.0921 0.0886	1.282	AS_151B5	T.Large Autosomal T.Small Autosomal T.Y	0.0544 0.0947 0.0575	1.740
M6H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0695 0.1279 0.0559	1.839	AS_161B5	T.Large Autosomal T.Small Autosomal T.Y	0.0862 0.1335 0.0559	1.549
M8H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0539 0.0740 0.0576	1.373	AS_221B5	T.Large Autosomal T.Small Autosomal T.Y	0.0537 0.0633 0.0576	1.181
M9H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0504 0.0511 0.0579	1.014	AS_72B5	T.Large Autosomal T.Small Autosomal T.Y	0.0635 0.0856 0.0579	1.348
M10H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0666 0.1007 0.0975	1.513	AS_82B5	T.Large Autosomal T.Small Autosomal T.Y	0.0688 0.1123 0.0975	1.631
M11H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0647 0.0909 0.0789	1.406	AS_172B5	T.Large Autosomal T.Small Autosomal T.Y	0.0665 0.1007 0.0975	1.514
M13H1KA	T.Large Autosomal T.Small Autosomal T.Y	0.0124 0.0202 0.0196	1.633	AS_233B5	T.Large Autosomal T.Small Autosomal T.Y	0.03546 0.056321 0.0196	1.588
M18H2KA	T.Large Autosomal T.Small Autosomal T.Y	0.0495 0.0559 0.0468	1.129	AS_263B5	T.Large Autosomal T.Small Autosomal T.Y	0.0465 0.0565 0.0468	1.214
M19H2KA	T.Large Autosomal T.Small Autosomal T.Y	0.0572 0.0787 0.0721	1.376	AS_303B5	T.Large Autosomal T.Small Autosomal T.Y	0.0634 0.0734 0.0721	1.157
M20H2KA	T.Large Autosomal T.Small Autosomal T.Y	0.0635 0.0808 0.0821	1.272	AS_343B5	T.Large Autosomal T.Small Autosomal T.Y	0.0962 0.1533 0.0821	1.593
M21H2KA	T.Large Autosomal T.Small Autosomal T.Y	0.0671 0.0798 0.0824	1.190	AS_373B5	T.Large Autosomal T.Small Autosomal T.Y	0.0863 0.0963 0.0824	1.116
M22H2KA	T.Large Autosomal T.Small Autosomal T.Y	0.0542 0.1028 0.0975	1.897	AS_383B5	T.Large Autosomal T.Small Autosomal T.Y	0.0334 0.0532 0.0975	1.593
M29H2KA	T.Large Autosomal	0.0679	1.396	AS_403B5	T.Large Autosomal	0.0862	1.117

	T.Small Autosomal	0.0947			T.Small Autosomal	0.0963	
	T.Y	0.0575			T.Y	0.0575	

PHỤ LỤC 6: Hồ sơ Y-STR của 200 mẫu sử dụng bộ kit PowerPlex® Y23 System

STT	DYS 576	DYS 389 I	DYS 448	DYS 389 II	DYS 19	DYS 391	DYS 481	DYS 549	DYS 533	DYS 438	DYS 437	DYS 570	DYS 635	DYS 390	DYS 439	DYS 392	DYS 643	DYS 393	DYS 458	DYS 385a	DYS 385b	DYS 456	YGA TAH4
1	19	12	19.2	28	16	11	24	12	11	10	14	19	22	26	12	13	10	12	17	12	12	14	13
2	19	11	18	27	15	10	23	12	10	10	14	16	22	25	12	13	12	14	18	14	18	14	10
3	19	12	19	29	15	10	25	13	11	10	14	18	23	25	12	13	11	12	18	12	18	13	12
4	16	13	20	29	15	11	25	12	12	10	14	16	21	23	10	11	9	14	16	11	18	15	12
5	19	12	19	29	15	10	25	13	11	10	14	17	23	24	12	13	11	12	18	12	20	13	12
6	18	12	19	29	15	10	25	13	12	10	14	17	23	26	12	13	11	12	18	12	20	13	12
7	18	12	20	28	16	10	24	12	12	10	15	20	22	23	11	13	11	12	17	15	18	15	12
8	20	13	18	29	16	11	23	12	10	10	14	17	23	24	13	13	12	14	18	13	18	15	11
9	19	12	19	29	15	10	25	13	11	10	14	17	23	25	12	13	11	12	20	12	20	13	12
10	15	14	21	30	15	10	25	13	12	10	14	16	21	23	12	11	8	15	15	11	19	14	11
11	18	13	18.2	28	15	10	24	13	10	11	14	19	21	25	14	13	11	14	17	13	18	15	11
12	17	14	21	31	16	10	23	11	11	10	15	18	21	24	11	13	11	12	18	16	21	13	13
13	18	14	18	30	16	11	23	11	10	10	14	15	22	24	12	13	12	14	18	14	19	15	11
14	16	13	20	29	15	11	24	12	12	10	14	17	21	23	10	11	9	14	15	11	18	15	12
15	19	12	20	28	14	10	24	12	11	11	15	16	20	24	11	14	12	12	19	13	17	15	11
16	19	12	19	29	15	10	25	13	11	10	14	17	25	25	12	13	11	12	18	12	20	13	12
17	18	13	21	28	16	10	26	12	11	10	15	19	23	26	12	13	10	12	17	13	19	15	12
18	18	14	21	29	16	10	22	11	12	10	14	18	20	24	13	11	9	15	17	11	18	15	11
19	19	14	22	29	15	10	24	12	12	10	14	18	21	24	12	11	9	15	17	11	18	15	11
20	18	12	20	28	14	10	23	12	11	11	14	18	20	24	12	14	11	12	17	14	17	15	12
21	18	13	18	29	16	11	23	11	10	10	14	16	21	24	12	13	12	14	18	13	19	15	11
22	17	11	18	28	16	11	23	12	10	10	14	15	21	25	12	13	12	14	17	13	18	14	10
23	17	12	20	28	14	10	23	13	11	11	14	19	20	23	12	14	11	12	17	13	19	15	12
24	19	14	20	30	15	10	23	13	11	11	15	17	20	25	14	14	12	12	18	13	17	15	12
25	18	12	19	28	16	10	24	13	11	10	14	19	22	26	13	13	10	12	22	14	19	16	11
26	20	13	18	29	16	11	24	11	10	10	14	16	21	24	12	13	12	15	18	13	18	15	11

27	17	13	18	30	17	10	23	12	10	10	14	15	22	23	12	13	12	14	17	14	18	16	11
28	19	14	19	30	13	6	24	12	11	12	14	19	22	24	11	14	10	14	17	16	20	17	10
29	21	12	20	28	15	11	25	12	11	10	14	19	21	23	11	13	11	12	16	12	16	14	11
30	17	13	18	29	17	11	22	12	10	10	14	16	22	24	12	13	11	14	17	13	19	15	11
31	18	13	20	29	14	10	23	12	11	11	14	18	20	24	12	14	11	12	17	12	13	16	12
32	19	13	18	29	15	10	23	12	10	10	14	16	22	25	12	13	12	15	15	16	21	15	11
33	18	12	20	28	15	10	24	13	12	10	15	17	20	23	12	13	12	13	17	12	12	15	12
34	18	14	18	30	15	10	23	12	11	10	14	16	21	24	11	13	13	14	15	15	20	15	11
35	17	13	18	29	15	11	21	12	10	10	14	16	22	25	13	13	13	14	15	14	21	15	11
36	17	13	21	28	16	10	26	11	11	10	15	19	22	25	11	13	10	12	17	13	18	15	12
37	18	13	18	29	16	10	24	12	10	10	14	16	22	24	13	13	12	14	17	13	18	15	11
38	20	14	18	30	15	10	23	12	12	10	14	16	21	25	12	13	13	14	15	15	20	16	11
39	18	12	19	30	14	10	25	12	12	11	15	17	19	24	11	11	10	13	17	15	16	14	12
40	17	14	22	31	16	10	24	12	12	10	14	18	21	23	11	11	9	13	16	11	17	15	11
41	18	12	20	27	14	10	23	12	11	11	15	17	20	25	13	14	12	12	18	13	17	15	12
42	18	13	21	28	16	10	25	11	11	10	15	19	23	25	12	13	10	12	18	13	18	15	12
43	18	12	20	28	15	10	24	13	12	10	15	17	21	23	13	13	12	13	17	13	19	16	12
44	19	13	18	29	17	11	22	12	10	10	14	16	23	24	12	13	12	13	17	13	19	15	11
45	18	14	18	31	15	10	23	13	10	10	14	18	23	24	12	13	11	14	15	12	17	15	12
46	20	11	18	27	16	11	23	12	10	10	14	16	21	25	12	13	12	14	18	13	18	14	10
47	17	11	20	27	14	10	26	13	11	10	14	19	23	24	13	13	11	12	19	13	19	15	12
48	20	12	20	28	14	10	24	13	11	11	15	17	20	25	13	14	13	12	19	13	17	16	12
49	19	14	19	32	15	10	25	12	12	10	15	16	20	26	11	12	12	12	18	12	16	15	12
50	17	13	18	29	17	11	22	12	10	10	14	16	23	24	11	13	12	13	18	13	19	15	11
51	20	13	18	29	16	11	23	11	10	10	14	16	20	24	12	13	12	14	16	13	14	15	11
52	18	13	18	29	16	10	23	13	10	10	14	15	22	24	12	13	12	13	18	13	20	15	11
53	18	11	18	27	15	10	23	12	10	10	14	16	21	25	12	13	12	14	18	13	18	14	10
54	20	14	18	30	15	10	23	12	11	10	14	16	21	25	11	13	13	15	16	15	19	16	11
55	18	11	18	28	16	11	23	13	10	10	14	16	21	25	12	13	12	14	18	13	17	14	10
56	16	12	20	28	15	11	24	12	12	10	14	16	21	23	10	11	9	14	15	11	18	15	12

57	15	12	16	28	15	10	27	13	11	10	16	19	22	23	11	14	11	13	18	14	15	18	12
58	18	12	20	28	15	11	23	15	11	11	14	18	20	24	12	14	12	12	19	12	17	15	11
59	19	12	17	29	16	10	24	11	10	10	14	16	22	25	12	13	12	14	19	14	18	15	10
60	17	14	18	30	16	11	22	12	10	10	14	16	22	23	12	13	12	14	18	13	18	15	11
61	19	13	18	29	16	10	24	12	10	10	14	16	22	23	12	13	12	13	20	13	17	15	11
62	17	14	19	30	15	10	24	13	12	9	14	17	18	22	13	11	9	12	18	16	20	15	12
63	16	14	19	29	14	10	25	12	12	10	14	18	23	23	11	14	10	13	15	11	11	17	12
64	16	11	18	29	16	11	23	12	10	10	14	16	23	25	12	13	12	13	18	13	18	14	10
65	20	14	19	30	13	6	24	13	12	12	14	19	22	24	11	14	10	14	16	17	20	17	10
66	20	13	21	28	16	10	26	11	11	10	15	20	23	25	12	13	10	12	18	13	19	15	12
67	18	11	18	27	16	11	22	12	10	10	14	16	21	24	12	13	12	14	18	13	18	14	10
68	18	13	18	28	15	10	23	12	10	10	14	16	21	25	12	13	12	13	18	17	20	16	11
69	18	13	21	30	15	10	28	12	12	10	14	16	20	23	11	11	8	14	13	11	20	15	11
70	18	14	21	31	15	11	23	12	11	10	15	20	21	24	11	13	11	12	19	14	21	14	12
71	20	13	18	29	16	11	23	12	10	10	14	16	23	24	12	13	12	14	19	13	18	15	11
72	18	13	18.2	28	15	10	24	14	10	11	14	19	21	25	13	13	11	14	16	13	18	15	11
73	17	13	21	28	16	10	26	11	11	10	15	20	23	25	12	13	10	12	16	13	19	15	12
74	18	13	18	29	15	10	25	13	11	11	14	17	23	23	13	14	12	14	18	12	17	14	12
75	17	13	21	28	16	10	26	11	11	10	15	19	22	25	12	13	10	12	19	13	18	15	12
76	19	13	17	31	16	11	23	11	10	10	14	16	22	24	12	13	12	14	17	12	19	15	11
77	18	13	20	29	15	11	24	12	12	10	13	16	21	23	10	11	9	14	16	12	18	15	12
78	21	12	19	28	17	10	24	13	11	10	14	17	21	23	11	13	10	12	18	12	19	14	12
79	18	13	18	30	15	10	23	11	10	10	14	17	21	24	12	13	12	14	18	13	19	15	12
80	16	13	18	29	15	11	21	12	10	10	14	17	21	25	12	13	13	14	16	15	18	15	12
81	18	12	19	28	15	10	25	13	11	10	14	17	23	25	12	13	11	12	18	12	21	13	12
82	17	15	18	32	16	11	23	11	10	10	14	16	23	24	12	13	12	14	19	12	18	15	11
83	19	12	20	27	16	10	24	11	12	10	15	20	21	23	11	13	12	12	17	15	18	15	12
84	18	12	20	28	16	10	24	12	12	10	15	22	24	23	12	13	11	12	17	15	19	14	12
85	18	14	19	29	17	10	23	12	11	10	14	17	23	24	12	13	10	12	19	12	19	14	13
86	18	13	18	29	16	11	23	11	10	10	14	15	21	24	12	13	12	14	19	13	19	15	11

87	18	14	19	30	14	12	22	12	11	11	14	18	22	23	12	14	9	13	15	12	12	16	13
88	17	11	19	28	16	11	24	12	10	10	14	16	21	25	12	13	12	14	18	12	18	14	10
89	19	12	19	29	15	10	25	13	11	10	14	17	23	26	12	13	11	12	18	12	21	13	12
90	19	13	21	28	16	10	26	11	11	10	15	19	23	25	12	13	10	12	17	14	18	15	12
91	17	12	18	29	15	11	25	12	11	10	14	18	19	23	12	13	11	14	15	13	14	16	12
92	20	12	19	29	15	10	26	13	11	10	14	18	23	25	12	13	11	12	18	12	19	13	12
93	18	14	18	30	15	11	21	12	10	10	14	18	22	24	12	13	10	14	17	13	18	14	10
94	18	12	18	28	15	11	27	12	11	10	15	21	20	22	13	14	11	13	14	13	13	16	11
95	18	11	18	28	16	11	23	13	10	10	14	17	21	25	11	14	12	14	19	13	18	14	10
96	20	14	18	31	15	11	23	13	10	11	14	16	21	24	11	13	11	14	16	13	18	15	10
97	18	13	21	28	16	10	27	12	11	10	15	19	23	25	12	13	10	12	18	13	20	15	12
98	19	14	18	31	15	10	23	12	10	11	14	16	20	24	11	13	12	14	15	14	19	15	10
99	18	12	19	30	17	10	24	14	11	10	15	18	21	24	11	13	10	12	19	13	19	14	12
100	19	13	21	28	16	10	26	12	11	10	15	19	23	24	12	13	10	12	18	13	19	15	12
101	17	12	20	27	17	10	22	12	11	10	15	20	24	24	11	13	10	12	19	13	21	15	12
102	18	13	18	29	16	10	23	11	10	10	14	16	21	24	12	13	12	14	18	13	20	15	11
103	19	14	18	31	15	11	24	12	10	10	14	15	20	25	12	13	12	14	15	15	21	15	11
104	16	13	18	29	14	10	26	12	12	10	14	19	19	23	11	14	11	13	16	11	14	15	12
105	18	12	20	28	13	10	23	13	11	11	15	17	20	26	13	14	12	12	19	13	16	15	12
106	18	12	20	28	16	10	24	12	13	10	15	20	21	23	12	14	11	12	17	15	18	15	12
107	19	12	16	30	15	11	25	12	11	10	14	17	19	23	12	14	11	13	16	13	13	17	12
108	19	13	18	30	16	11	23	11	10	10	14	16	21	24	12	13	12	14	19	13	19	15	11
109	20	13	19	30	15	11	24	11	11	11	14	18	24	25	11	11	9	14	19	11	13	16	12
110	17	13	18	28	16	11	23	11	10	10	14	16	24	24	11	13	13	14	17	13	17	15	11
111	18	11	18	28	16	11	23	13	10	10	14	16	21	25	12	13	12	14	18	13	18	14	10
112	18	13	18	29	15	11	24	12	10	10	14	16	22	24	13	13	12	14	17	13	19	15	11
113	19	12	20	28	14	10	23	13	11	11	15	18	20	25	11	15	11	12	19	13	19	16	12
114	17	12	20	28	14	10	23	15	11	11	15	17	20	23	12	14	11	12	18	15	19	15	12
115	19	12	21	27	14	11	24	12	11	11	15	17	20	24	11	14	12	12	16	13	19	15	12
116	18	13	21	28	16	10	26	11	11	10	14	23	23	25	12	13	10	12	17	13	18	15	12

117	17	12	18	30	15	10	24	12	12	10	14	19	24	23	11	14	11	13	15	12	14	16	12
118	16	14	19	30	14	10	24	12	11	10	14	19	19	23	12	14	11	13	16	13	14	15	12
119	19	14	19	29	17	10	24	12	11	10	14	18	23	24	12	13	10	12	19	12	19	14	13
120	18	13	19	30	16	11	23	11	10	10	14	16	21	24	13	13	12	14	19	13	19	15	11
121	20	13	20	29	14	10	22	12	11	11	15	19	20	24	12	14	10	12	19	13	17	15	12
122	16	11	18	28	15	10	24	12	12	10	14	21	21	23	12	14	11	13	15	13	13	16	12
123	18	11	18	28	15	11	23	12	10	10	14	16	21	25	12	13	12	14	18	13	18	14	10
124	20	13	18	28	15	11	23	12	10	10	14	16	21	24	12	13	13	14	16	15	21	15	11
125	16	13	20	29	16	11	24	12	12	10	14	16	21	23	10	11	9	14	16	11	18	15	12
126	19	13	19	29	15	10	22	12	11	10	15	16	19	25	13	13	11	12	19	13	13	15	12
127	20	12	19	29	15	10	25	13	11	10	14	17	22	25	12	13	11	12	18	12	20	13	12
128	19	14	21	29	16	10	26	11	11	10	15	20	23	25	12	13	10	12	18	13	19	15	12
129	18	11	18	28	16	10	24	12	10	10	14	16	22	25	12	13	12	14	18	13	18	14	10
130	19	12	18	29	15	11	23	12	12	10	14	16	25	25	13	13	10	12	18	14	21	18	11
131	15	13	18	28	14	10	25	12	12	10	15	22	21	22	12	14	11	13	14	13	13	15	12
132	21	12	20	28	14	10	22	13	11	10	15	17	20	23	13	12	11	12	20	11	15	16	13
133	18	12	21	29	16	10	23	11	11	10	15	19	21	24	12	13	11	12	16	13	18	15	12
134	19	13	18	29	16	11	23	12	10	10	14	17	22	24	13	13	12	14	18	13	18	15	11
135	18	13	20	30	17	10	25	12	11	10	14	17	21	23	12	14	10	12	17	12	12.3	15	11
136	19	13	18	29	15	10	23	12	10	10	14	16	21	25	11	13	12	15	15	16	21	16	11
137	20	12	19	29	13	10	22	13	13	10	15	19	22	25	12	12	9	13	17	12	12	14	12
138	20	13	19	29	15	10	22	12	11	10	15	16	19	24	13	13	11	12	18	12	12	15	11
139	18	14	19	29	17	10	24	12	11	10	14	18	24	24	12	13	10	12	18	13	20	14	13
140	18	14	20	30	16	10	27	12	11	8	15	20	21	24	12	14	11	12	17	13	20	16	12
141	16	14	18	31	16	12	23	11	10	10	14	16	25	24	12	13	12	14	18	12	18	15	11
142	17	14	18	29	15	10	23	12	11	10	14	16	21	24	11	13	13	14	15	15	19	16	11
143	20	13	18	30	16	11	23	12	10	10	14	17	22	24	12	12	12	13	18	13	19	14	11
144	17	12	19	30	14	10	24	12	12	11	14	18	19	24	11	11	10	13	16	14	16	14	11
145	18	11	18	28	16	10	24	12	10	10	14	16	21	25	13	13	12	14	18	12	18	14	10
146	19	12	19	29	15	10	25	13	11	10	14	17	22	25	11	13	11	12	18	12	19	13	12

147	16	12	18	28	16	10	24	12	11	10	14	19	20	23	11	14	11	13	17	13	13	16	12
148	17	14	19	29	17	10	24	13	11	10	14	17	23	24	11	13	10	12	18	12	18	14	12
149	19	14	19	30	13	6	24	12	12	12	14	19	22	24	11	14	10	14	16	16	20	16	10
150	17	12	19	28	17	10	24	12	11	10	14	20	22	26	13	13	10	12	22	14	19	14	11
151	19	13	18	30	16	11	23	11	9	10	14	16	21	24	11	13	12	14	18	13	20	15	11
152	18	12	21	28	16	10	23	12	11	10	15	20	22	23	13	13	13	12	17	13	22	16	12
153	15	13	20	29	15	10	24	12	13	10	14	17	21	23	10	11	9	14	16	11	11	15	12
154	15	13	18	29	14	10	25	12	11	10	15	21	21	22	12	14	11	13	14	13	13	15	12
155	19	12	19	29	16	12	25	13	11	10	14	17	21	25	11	13	10	12	20	12	18	14	11
156	17	13	19	30	15	11	23	12	10	10	14	17	23	25	12	13	12	14	16	15	21	16	11
157	18	14	18	30	16	10	24	12	11	10	14	16	21	24	11	13	13	14	15	16	19	16	11
158	19	12	19	29	15	10	25	13	11	10	14	17	24	25	12	13	11	12	18	12	19	13	12
159	17	13	18	30	15	11	21	12	10	10	14	17	21	25	12	13	13	14	16	15	18	15	12
160	18	12	18	29	15	11	26	12	11	10	14	18	19	23	11	14	11	13	15	13	13	16	12
161	17	13	20	29	15	10	24	13	11	10	15	20	21	24	11	13	12	13	16	13	22	15	11
162	16	13	20	30	13	10	24	12	12	10	14	18	21	23	13	11	13	13	16	16	17	16	11
163	19	12	18	28	15	11	24	13	11	10	14	18	21	23	12	14	11	13	14	13	13	15	12
164	19	13	18	30	16	11	23	11	10	10	14	15	21	24	13	13	12	14	19	13	19	15	11
165	19	13	18	29	15	11	25	13	11	11	14	16	21	23	13	13	12	14	17	12	17	14	12
166	19	13	18	30	16	11	23	10	10	10	14	16	21	25	12	13	12	14	18	13	19	15	11
167	19	13	19	29	16	10	27	13	11	10	15	18	21	23	11	13	11	12	18	14	20	15	12
168	16	13	18	29	14	10	24	11	11	10	15	19	21	23	12	14	11	13	15	13	13	15	12
169	20	12	19	29	15	10	25	12	11	10	14	17	23	25	12	13	11	12	18	12	20	13	12
170	18	14	19	31	13	6	24	12	12	12	14	19	22	24	11	14	10	14	15	16	20	18	10
171	19	13	18	30	16	11	22	12	11	12	14	21	21	23	12	13	11	13	18	9	17	15	11
172	18	14	21	30	15	10	26	13	14	10	14	17	21	22	10	11	8	15	15	11	19	15	11
173	18	12	19	29	15	10	25	13	11	10	14	17	23	26	12	13	11	13	19	12	20	13	12
174	22	14	18	30	15	10	24	12	10	10	14	17	21	25	12	13	13	14	16	16	18	15	11
175	19	13	18	30	15	11	23	12	10	10	14	16	24	24	11	15	12	14	16	14	20	15	10
176	19	12	19	28	15	10	25	14	11	10	14	17	23	25	12	13	11	12	18	13	20	13	12

177	18	13	21	28	16	10	26	12	12	10	15	19	23	25	12	13	10	12	18	13	18	15	12
178	17	14	18	30	16	11	23	11	10	10	14	17	22	24	12	13	12	14	17	13	18	15	11
179	17	14	20	32	15	10	26	12	11	10	14	20	22	23	11	12	12	14	14	14	15	13	10
180	19	14	19	30	13	9	26	12	12	12	14	22	22	24	12	14	10	14	16	15	20	16	10
181	20	12	19	29	15	10	26	12	11	10	14	17	23	25	12	13	11	12	18	12	20	13	11
182	19	12	20	28	14	10	22	13	11	10	15	17	20	23	13	12	11	12	20	11	11	18	13
183	17	13	18	30	15	11	23	11	10	10	14	16	21	24	13	13	12	14	16	13	18	15	11
184	16	12	18	29	15	10	24	12	12	10	14	21	21	23	13	14	11	13	15	13	13	16	12
185	19	12	19	28	16	10	24	13	11	10	15	20	22	26	14	13	10	12	17	12	18	14	12
186	20	12	18	28	16	10	23	12	12	10	14	16	23	25	13	13	10	12	17	13	20	16	11
187	19	13	18	29	15	10	23	11	10	10	16	18	24	24	12	13	13	14	18	13	18	15	11
188	18	11	18	27	16	11	23	12	10	10	14	16	21	25	11	13	12	14	17	13	18	14	10
189	19	13	20	29	14	10	22	12	11	11	15	18	20	24	12	14	11	12	19	13	17	15	12
190	19	14	18	30	15	10	23	11	10	10	14	18	23	24	12	13	12	13	16	14	18	16	11
191	19	14	21	31	15	10	27	11	12	10	14	16	22	25	12	11	12	13	16	13	13	17	12
192	18	12	19	28	16	10	24	13	11	10	15	20	23	26	15	13	10	12	17	12	18	14	12
193	18	13	18	29	16	11	24	12	10	10	14	15	24	24	13	13	12	14	19	13	19	15	11
194	19	12	19	28	13	11	22	12	11	10	15	16	19	23	12	12	11	12	18	12	16	15	11
195	17	13	21	30	15	10	26	13	12	10	14	16	21	23	11	11	9	15	16	11	20	16	11
196	18	14	18	30	15	9	24	12	10	10	14	16	21	24	11	13	13	14	15	15	18	16	11
197	15	13	18	31	14	10	24	12	11	10	15	19	20	23	12	14	11	13	16	13	13	15	12
198	18	11	18	28	16	10	23	12	10	10	14	17	22	25	12	13	12	14	19	13	18	13	10
199	18	13	18	29	15	10	23	12	10	10	14	16	21	23	11	13	12	14	19	13	17	15	12
200	16	13	20	29	15	11	24	12	12	10	14	16	21	23	11	11	9	14	16	11	18	15	12

PHỤ LỤC 5: Hồ sơ Y-STR của 200 mẫu sử dụng bộ Yfiler™ Plus PCR Amplification Kit

STT	DYS 576	DYS 389I	DYS 635	DYS 389II	DYS 627	DYS 460	DYS 458	DYS 19	YGA TAH4	DYS 448	DYS 391	DYS 456	DYS 390	DYS 438	DYS 392	DYS 518	DYS 570	DYS 437	DYS 385	DYS 449	DYS 393	DYS 439	DYS 481	DYF 387S1	DYS 533
1	18	11	21	27	21	10	18	16	10	18	11	14	25	10	13	40	16	14	13, 19	25	14	12	26	37, 38	11
2	18	14	21	30	21	12	18	16	12	21	10	15	24	10	11	35	18	14	12, 18	30	15	12	24	34, 37	12
3	17	12	21	29	19	10	18	15	13	20	9	14	24	10	13	39	18	14	14, 19	28	12	11	24	36, 36	11
4	19	13	21	31	19	10	17	15	11	18	10	15	24	10	13	41	21	14	12, 20	30	14	12	23	34, 34	11
5	17	14	21	29	20	10	16	15	11	18	10	14	24	10	13	38	19	14	11, 17	28	14	11	23	36, 38	10
6	18	11	21	28	21	10	20	16	10	18	11	14	25	10	13	41	16	14	13, 18	23	14	12	23	38, 38	11
7	19	13	21	30	20	10	18	16	12	18	11	15	25	10	13	42	16	14	13, 19	25	14	12	23	36, 39	10
8	18	11	21	27	21	10	18	16	10	18	11	14	25	10	13	40	16	14	13, 19	25	14	12	23	37, 38	10
9	17	12	22	29	23	10	18	15	11	19	10	14	25	10	13	36	20	14	12, 12	31	12	12	24	36, 37	12
10	18	13	24	28	22	10	15	16	13	21	10	15	24	10	13	40	20	15	13, 18	34	12	11	26	40, 40	11
11	19	13	23	28	21	10	18	16	12	21	10	15	25	10	13	41	19	14	13, 18	35	12	12	26	39, 40	11
12	18	12	20	28	18	9	17	13	12	20	10	15	23	11	14	36	19	14	13, 19	36	12	11	23	35, 38	11
13	17	12	21	27	20	10	18	16	11	20	11	14	25	10	13	41	20	14	13, 19	27	13	11	26	36, 39	11
14	17	13	25	30	21	12	18	16	11	18	10	15	24	10	13	41	17	14	14, 16	27	14	12	23	37, 39	10
15	18	12	23	28	24	10	20	17	11	19	10	14	26	10	13	37	20	14	14, 19	31	12	13	24	36, 36	11
16	17	13	21	30	20	10	17	15	11	18	11	15	24	10	13	43	16	14	13, 19	26	14	13	23	37, 39	11
17	18	13	23	28	21	10	18	16	12	21	10	15	25	10	13	39	19	15	13, 19	35	13	12	26	40, 40	11
18	19	12	21	28	18	9	17	14	12	20	10	15	24	11	14	36	17	15	13, 18	32	11	14	23	35, 38	11
19	17	11	21	28	21	10	19	16	10	18	10	14	25	10	13	40	16	14	13, 18	26	14	12	23	37, 38	11
20	18	12	23	28	20	11	19	16	12	19	10	15	26	10	13	37	21	15	12, 18	32	12	14	24	37, 38	11
21	20	12	20	28	18	10	18	14	12	20	10	15	25	11	14	35	16	15	13, 17	34	12	13	23	36, 38	11
22	19	13	21	29	22	10	16	15	12	18	10	15	23	10	13	39	17	14	13, 13	28	12	12	25	36, 36	11
23	20	12	22	29	20	10	18	15	12	19	11	13	24	10	13	37	17	14	12, 20	31	12	12	25	36, 38	11
24	19	12	23	29	20	10	18	15	12	19	10	13	25	10	13	36	17	14	12, 20	31	12	12	25	36, 38	11
25	20	12	21	28	22	10	17	16	12	19	10	17	25	10	14	38	16	16	12, 18	34	11	11	25	39, 41	13
26	18	12	20	27	18	9	17	14	11	20	11	15	24	11	14	36	21	15	13, 19	32	12	12	24	36, 38	11
27	18	13	22	30	22	10	17	15	11	18	10	17	25	10	13	40	16	14	15, 15	28	14	14	22	36, 37	11
28	16	13	22	29	21	10	18	16	10	18	11	15	24	10	13	43	14	14	13, 18	26	14	12	23	37, 38	10
29	17	13	25	29	23	10	18	15	10	18	10	15	24	10	13	37	17	14	14, 19	29	14	11	23	38, 38	10
30	17	13	22	29	20	10	18	16	13	18	11	15	24	10	13	41	17	14	13, 19	27	14	13	22	36, 39	10
31	18	12	20	28	22	10	17	15	12	19	11	16	23	10	12	36	18	15	13, 16	34	12	15	22	36, 40	11

32	19	13	23	29	22	10	16	15	11	18	10	16	25	10	13	42	17	14	16, 19	30	14	12	25	37, 37	10
33	18	13	21	29	24	10	15	15	11	18	10	15	25	10	13	38	15	14	15, 21	29	14	12	23	36, 37	10
34	18	13	21	31	22	10	18	15	10	18	11	15	25	10	13	40	16	14	16, 21	32	14	12	23	37, 37	10
35	18	12	20	28	18	9	18	14	12	20	10	15	24	11	14	34	19	14	13, 18	35	12	12	21	35, 40	11
36	17	14	22	30	21	12	15	15	11	20	10	15	22	10	11	38	16	14	11, 18	29	14	12	26	39, 39	12
37	18	12	22	28	21	10	17	15	11	18	11	15	24	10	13	40	15	14	13, 18	26	14	12	22	37, 39	10
38	19	14	22	29	20	10	23	16	11	18	11	15	24	10	13	43	16	14	13, 18	27	14	12	22	36, 39	10
39	18	13	23	28	22	10	17	16	12	21	10	15	25	10	13	40	20	15	13, 13	36	12	12	26	41, 41	11
40	18	14	21	29	22	10	16	15	11	18	10	17	24	10	14	43	16	14	16, 20	28	14	12	23	38, 39	10
41	20	13	23	28	24	11	16	16	12	21	10	15	25	10	13	40	20	15	13, 18	35	12	12	21	36, 38	10
42	19	13	21	29	20	11	20	16	11	18	10	15	23	10	13	40	16	14	13, 18	26	14	12	26	39, 39	11
43	17	12	21	25	20	10	19	16	11	20	11	14	24	10	13	40	20	14	12, 19	28	13	11	24	38, 38	10
44	19	12	23	28	20	10	18	15	12	19	10	13	25	10	13	36	17	14	12, 20	32	12	12	26	36, 38	11
45	17	12	21	30	23	10	15	15	12	19	10	16	23	11	14	38	19	14	12, 14	28	13	12	25	36, 38	11
46	20	13	23	29	21	11	16	15	11	18	10	16	25	10	13	40	17	14	15, 20	28	14	12	25	36, 39	12
47	18	13	23	28	21	10	17	16	12	21	10	15	25	10	13	40	19	15	13, 18	36	12	13	22	36, 37	10
48	19	12	19	30	18	12	16	14	12	19	10	15	24	11	11	37	18	15	15, 16	25	13	11	26	39, 39	11
49	19	12	23	27	20	10	18	15	12	19	10	13	25	10	13	37	17	14	12, 20	32	12	12	24	36, 36	12
50	18	12	23	28	21	9	20	17	12	20	10	15	24	10	13	38	20	15	13, 22	33	12	11	26	36, 38	11
51	18	12	21	28	21	9	18	15	11	19	10	16	24	10	13	39	17	15	13, 21	33	12	12	22	36, 41	11
52	17	13	21	29	20	11	15	14	12	18	10	15	23	10	14	35	20	15	13, 14	31	13	12	28	38, 38	11
53	18	14	18	30	20	11	14	14	12	18	10	15	22	10	14	34	20	15	13, 13	32	13	12	24	36, 39	12
54	21	14	23	31	21	10	18	16	11	18	11	15	24	10	13	41	16	14	13, 20	26	14	12	25	36, 38	11
55	17	13	21	28	19	9	17	15	11	19	10	15	24	10	14	42	18	15	13, 13	34	12	13	22	36, 40	11
56	20	12	22	27	22	9	18	15	12	19	10	16	23	11	12	37	15	15	13, 16	32	12	11	28	37, 38	11
57	19	13	21	30	20	10	17	16	11	18	11	15	24	10	13	43	16	14	13, 17	25	14	13	22	35, 38	11
58	18	13	22	29	19	10	17	16	11	18	11	15	24	10	13	39	15	14	13, 19	27	14	12	23	36, 41	10
59	17	14	20	30	22	11	21	15	11	19	10	15	24	10	10	34	21	15	15, 16	31	12	11	23	37, 38	10
60	17	12	21	28	20	10	20	16	12	18	10	14	25	10	13	37	20	14	12, 19	32	12	11	27	37, 37	12
61	18	14	22	30	21	10	16	13	10	20	10	17	24	12	14	35	19	14	16, 20	30	14	11	23	35, 39	11
62	18	12	19	29	22	10	15	15	12	18	10	16	23	10	14	39	18	14	13, 13	30	13	11	24	37, 37	12
63	20	12	21	27	22	9	19	15	12	19	10	16	23	11	12	37	14	16	12, 16	32	12	11	25	38, 38	11
64	16	13	21	29	21	11	16	15	12	20	11	15	23	10	11	38	16	14	11, 18	31	14	10	22	35, 38	12
65	19	14	21	30	23	10	19	16	12	20	10	16	24	10	14	40	20	15	13, 20	30	12	13	24	36, 38	12
66	18	12	22	28	22	10	16	15	12	18	10	15	24	10	13	37	14	15	13, 15	30	13	12	26	36, 40	11

67	19	14	21	30	22	11	15	15	12	18	10	17	25	10	13	41	16	14	16, 19	29	14	11	24	39, 39	12
68	18	13	24	30	21	11	16	14	12	19	10	15	23	11	14	36	18	14	11, 12	28	13	13	23	37, 37	11
69	18	13	23	29	21	11	17	14	12	19	11	15	24	12	13	39	17	15	12, 14	31	13	12	22	38, 38	12
70	18	12	22	29	19	10	18	15	10	18	10	15	24	10	13	39	16	14	14, 16	27	15	10	19	35, 36	12
71	18	14	22	29	19	11	16	16	11	18	11	15	24	10	13	41	15	14	13, 18	27	14	12	24	37, 38	10
72	17	12	22	28	23	10	16	15	12	18	10	15	24	10	13	36	14	15	13, 16	31	13	11	23	36, 39	10
73	20	12	23	29	20	10	18	14	12	19	10	13	25	10	13	35	18	14	12, 19	31	12	12	24	39, 39	12
74	19	13	20	31	22	10	16	15	10	18	10	15	24	11	13	39	16	14	14, 18	26	14	11	25	36, 39	11
75	20	12	20	28	16	9	19	14	12	20	10	15	25	11	15	35	16	15	13, 17	32	12	13	23	36, 38	10
76	21	14	21	30	21	11	15	15	11	18	10	16	25	10	13	43	16	14	15, 20	29	14	11	23	35, 38	11
77	18	13	22	29	22	10	19	16	11	18	10	15	25	10	13	41	16	14	13, 20	26	15	12	23	37, 37	11
78	20	12	23	29	20	10	18	15	12	19	10	13	25	10	13	35	18	14	12, 19	31	12	12	26	36, 39	11
79	18	13	23	28	22	10	17	16	12	20	10	15	25	10	13	39	22	15	13, 18	35	12	13	26	37, 41	11
80	18	12	20	28	21	9	18	14	12	20	11	15	23	11	14	40	17	15	15, 17	32	12	12	23	35, 39	11
81	17	12	22	28	19	9	17	15	12	20	10	14	23	10	13	38	21	15	15, 19	33	12	12	25	37, 41	12
82	18	11	21	27	22	10	18	16	10	18	11	14	25	11	13	40	16	14	13, 18	26	14	13	23	37, 38	10
83	14	13	24	30	20	12	15	14	12	19	10	18	23	10	14	35	18	14	11, 12	28	13	11	22	36, 36	13
84	17	13	21	29	21	11	14	14	12	18	10	15	22	10	14	36	21	15	13, 13	32	13	12	25	35, 39	11
85	19	13	21	30	21	10	18	16	11	18	10	15	24	10	13	42	18	14	13, 18	26	14	12	23	37, 38	10
86	19	12	23	30	20	10	16	17	11	18	10	17	24	14	13	39	18	14	15, 19	28	14	12	22	36, 37	10
87	19	12	24	29	20	10	18	15	12	19	10	14	25	10	13	36	17	14	12, 20	31	12	12	25	36, 38	11
88	17	12	22	28	22	10	21	16	11	19	10	15	26	10	13	36	20	14	14, 19	31	12	12	24	36, 36	11
89	18	12	20	28	17	10	19	14	11	20	10	15	24	11	14	33	18	15	15, 18	30	12	11	23	35, 38	11
90	19	13	21	29	20	10	17	16	11	18	11	15	25	10	13	40	16	14	13, 19	26	14	12	23	37, 39	11
91	18	11	21	28	19	10	18	15	10	18	11	14	25	10	13	42	16	14	13, 18	26	14	12	23	37, 38	10
92	17	12	20	28	22	12	15	15	10	18	10	15	23	10	14	38	19	14	13, 14	31	13	11	24	36, 40	11
93	18	13	23	29	21	10	17	15	12	18	11	16	24	10	13	40	17	14	16, 20	29	14	12	23	37, 37	10
94	18	11	22	28	21	10	18	16	10	18	11	13	25	10	13	40	16	14	13, 18	26	14	12	23	38, 38	10
95	18	11	23	28	20	10	18	16	10	18	12	14	25	10	13	41	17	14	12, 18	26	14	13	24	37, 37	10
96	18	13	22	30	18	10	17	16	11	18	10	15	24	10	13	43	18	14	13, 18	27	13	12	22	39, 39	11
97	19	13	22	30	22	10	17	15	11	18	10	17	25	10	13	40	16	14	13, 14	28	14	14	22	37, 37	10
98	17	13	21	29	20	11	19	15	11	18	9	15	23	10	13	40	16	14	12, 18	26	14	12	22	38, 38	10
99	19	11	21	27	22	10	18	16	10	18	11	14	25	10	13	42	16	14	13, 18	26	14	12	23	37, 38	10
100	20	14	22	30	23	10	16	13	10	19	9	16	24	12	14	35	19	14	15, 23	30	14	11	24	38, 38	12
101	18	14	22	30	22	11	15	16	11	18	10	16	24	10	13	41	16	14	16, 19	29	14	11	23	37, 39	12

102	16	12	23	27	24	10	17	17	11	18	10	15	23	10	13	32	19	15	13, 20	36	13	13	27	36, 36	12
103	19	13	22	29	22	11	16	15	11	18	10	15	25	10	14	42	16	14	15, 19	29	14	13	23	37, 38	10
104	17	12	21	30	20	10	19	17	12	18	10	14	24	10	13	38	19	15	14, 14	32	12	11	24	36, 38	11
105	17	13	23	30	23	11	16	15	11	18	11	16	25	10	13	41	17	14	15, 15	30	14	13	23	35, 37	10
106	19	14	21	30	20	11	15	15	11	18	11	16	25	10	13	41	15	14	15, 20	29	14	12	24	36, 37	10
107	18	12	21	29	18	10	17	14	12	20	11	15	24	11	14	36	18	15	13, 18	33	12	12	23	38, 38	10
108	17	13	20	29	22	11	18	16	11	21	10	15	24	10	11	37	17	14	11, 18	31	15	11	24	34, 36	12
109	16	13	21	30	21	11	15	16	12	18	10	16	24	10	14	37	20	14	13, 13	32	13	13	26	37, 38	12
110	18	13	21	29	20	9	16	15	11	20	10	15	24	10	13	40	19	15	13, 18	31	13	11	24	37, 39	12
111	16	12	20	28	21	11	17	14	12	18	10	15	24	10	13	38	19	16	13, 13	32	13	13	24	36, 40	11
112	19	13	21	29	19	11	19	16	11	18	10	15	23	10	13	39	16	14	13, 13	27	14	12	23	38, 38	10
113	19	11	21	29	21	10	19	16	10	19	11	14	25	10	13	42	16	14	14, 18	25	14	12	23	37, 38	10
114	19	14	23	31	21	10	16	14	12	18	10	15	24	10	13	42	16	14	15, 20	28	14	12	24	37, 38	10
115	17	14	21	31	21	10	16	15	10	18	10	15	24	10	12	38	19	14	15, 17	27	14	12	22	36, 38	11
116	16	12	21	30	20	10	15	15	12	17	10	15	22	10	13	38	20	14	12, 13	30	13	12	26	38, 38	12
117	18	12	22	27	20	10	17	14	11	18	11	15	23	10	13	41	15	14	13, 19	26	14	12	22	37, 39	10
118	19	13	22	29	22	10	16	15	11	18	10	15	25	11	14	43	16	14	15, 19	29	14	13	23	37, 38	10
119	18	13	23	28	22	10	17	16	12	20	10	15	25	10	13	39	20	15	13, 18	36	12	12	27	40, 40	11
120	19	14	21	30	23	10	16	16	10	18	11	14	24	10	13	42	18	14	13, 18	26	14	12	23	36, 36	10
121	19	13	20	29	21	10	12	15	12	18	11	15	25	10	13	43	18	14	15, 22	29	15	11	23	38, 38	10
122	19	13	21	29	22	10	16	16	11	18	10	15	26	10	13	42	17	14	13, 18	27	14	12	23	37, 37	10
123	18	13	20	29	21	10	12	15	12	18	11	15	25	10	13	43	17	14	15, 21	29	14	11	23	38, 38	10
124	17	13	23	29	21	10	18	17	11	18	11	15	24	10	13	39	16	15	13, 17	25	13	12	22	38, 39	10
125	15	12	21	29	22	11	15	15	12	19	10	16	23	10	14	37	19	14	13, 13	30	13	13	26	36, 41	12
126	15	13	23	29	20	11	16	15	12	20	11	15	23	10	11	38	16	14	11, 18	33	14	10	24	37, 39	12
127	18	12	19	29	22	10	15	15	12	18	12	16	23	10	14	37	19	14	13, 13	32	14	11	27	38, 40	11
128	18	11	21	28	21	10	19	16	10	18	10	14	25	10	13	41	16	14	13, 17	26	14	12	23	37, 38	10
129	19	13	23	28	21	9	18	16	12	21	10	15	25	10	13	41	19	15	13, 19	35	12	13	26	40, 40	11
130	19	12	20	27	22	10	17	15	12	19	11	15	23	10	12	37	16	15	12, 13	35	12	12	22	37, 39	12
131	16	13	22	30	21	10	19	15	11	18	10	14	24	11	14	41	17	14	13, 20	25	12	13	24	37, 37	10
132	18	12	22	28	19	11	18	16	11	18	10	15	25	10	13	41	18	14	13, 18	33	14	12	23	38, 39	10
133	18	12	22	28	20	10	17	14	11	17	11	15	24	10	13	41	19	14	13, 21	27	14	12	22	37, 39	11
134	16	14	21	32	20	10	18	14	11	20	10	15	23	11	10	37	20	15	12, 17	29	14	11	27	37, 40	11
135	19	13	19	29	24	10	16	15	12	19	10	15	24	10	13	37	19	15	12, 16	32	12	12	22	38, 39	12
136	18	12	24	28	20	9	17	16	12	20	10	14	23	10	13	37	19	15	15, 19	32	12	12	24	37, 40	12

137	18	11	21	27	23	10	18	16	10	18	11	14	25	10	13	42	16	14	13, 18	26	14	12	24	37, 39	10
138	19	13	22	31	20	9	16	15	11	18	11	15	25	10	13	41	17	14	13, 19	26	14	13	23	38, 38	11
139	19	12	20	28	18	9	19	14	12	20	10	15	24	11	14	35	17	15	13, 18	31	11	13	23	35, 38	11
140	19	12	20	29	22	9	16	16	12	20	10	16	22	10	13	37	17	14	11, 20	32	12	11	25	37, 39	11
141	18	14	20	32	24	9	18	15	12	19	10	14	26	10	12	36	16	15	12, 16	31	12	11	25	35, 40	12
142	19	13	22	30	18	10	17	16	11	18	11	15	24	10	13	40	16	14	13, 17	28	14	12	23	38, 39	10
143	19	12	23	29	21	10	18	15	12	19	10	13	25	10	13	36	18	14	12, 20	31	12	12	25	36, 38	11
144	19	14	21	30	23	12	16	15	11	18	10	16	23	10	13	40	16	14	15, 19	29	14	11	23	37, 37	11
145	17	12	21	28	20	11	16	15	12	18	10	15	23	10	14	35	20	14	13, 13	31	13	11	24	35, 38	11
146	20	13	21	30	23	10	15	15	11	17	11	15	25	10	13	39	16	14	14, 16	30	13	12	23	36, 37	10
147	18	13	22	29	22	12	18	16	11	17	11	15	24	10	13	41	17	14	13, 18	28	14	13	23	38, 39	10
148	18	13	21	29	22	12	21	16	12	18	11	16	24	10	13	39	19	14	12, 15	31	12	12	22	37, 39	10
149	18	14	21	30	24	10	18	16	11	18	10	16	25	10	14	41	16	14	15, 19	32	15	14	22	37, 38	10
150	17	11	21	27	23	10	18	15	10	18	12	14	25	11	13	42	16	14	13, 18	26	14	12	23	38, 38	10
151	15	12	24	28	22	12	17	15	11	19	11	15	25	11	11	39	23	14	15, 19	27	13	11	25	38, 38	12
152	19	12	23	29	20	10	18	15	12	19	10	13	25	10	13	36	17	14	12, 20	30	12	12	25	36, 38	11
153	18	13	22	30	19	10	16	16	11	18	11	15	24	10	13	40	16	14	18, 19	27	14	13	23	37, 38	10
154	17	14	22	30	19	11	18	16	11	18	11	15	24	10	13	42	17	14	13, 18	26	14	12	23	36, 39	10
155	19	12	21	29	22	10	20	17	12	19	11	14	23	10	14	36	18	14	13, 19	31	12	12	24	37, 38	10
156	18	14	22	30	21	10	13	17	11	17	11	15	24	10	13	40	16	14	13, 19	27	14	13	21	37, 37	10
157	17	11	21	27	21	10	18	17	11	17	11	14	25	10	13	42	16	14	13, 18	26	14	13	23	37, 38	10
158	20	13	23	31	19	10	18	15	11	18	10	15	23	10	13	41	23	14	12, 20	32	14	12	23	34, 34	10
159	19	14	22	29	21	11	15	15	11	18	10	15	25	10	13	41	16	14	15, 19	29	14	11	23	36, 38	11
160	20	13	22	29	20	10	17	17	11	19	10	15	24	10	13	41	15	14	13, 19	26	14	12	23	37, 38	10
161	19	12	22	29	22	10	16	15	12	19	10	13	25	10	13	36	18	14	12, 19	31	12	12	25	36, 38	11
162	20	14	21	30	22	11	15	15	11	18	10	16	24	10	13	42	16	14	15, 19	29	14	11	23	37, 38	11
163	18	13	20	30	19	10	18	14	12	20	10	16	24	11	14	37	19	14	12, 14	33	12	12	23	35, 37	11
164	17	13	20	29	20	11	14	14	12	18	10	15	23	10	14	36	20	15	13, 13	32	14	12	24	36, 39	11
165	21	13	23	29	21	10	16	15	11	18	11	15	25	10	14	45	16	14	15, 19	28	14	14	23	37, 38	10
166	16	12	21	30	13	11	15	15	11	18	10	15	23	10	14	36	20	14	13, 13	30	13	12	24	37, 39	11
167	18	14	21	29	22	10	16	15	12	18	10	15	24	10	13	39	19	14	12, 17	30	14	11	23	36, 38	10
168	19	13	21	30	20	10	18	16	11	18	11	15	24	10	13	41	16	14	13, 19	26	14	12	23	36, 39	10
169	18	12	24	29	21	10	18	16	11	18	10	15	25	10	13	38	16	14	13, 20	34	12	13	23	38, 38	12
170	19	11	21	27	20	10	18	16	10	18	11	14	25	10	13	42	16	14	13, 18	26	14	12	24	37, 38	10
171	17	13	22	29	22	11	19	16	11	18	11	15	23	10	15	41	16	14	13, 17	26	14	12	23	38, 39	10

172	17	13	21	29	23	10	15	12	11	17	11	14	24	9	13	40	16	14	13, 19	26	14	11	23	37, 38	10
173	17	13	21	28	22	10	15	15	11	18	11	15	25	10	13	38	17	14	15, 21	29	14	12	23	36, 37	10
174	18	13	21	29	23	10	18	16	11	18	11	16	26	8	13	42	16	14	13, 18	29	13	12	23	37, 37	10
175	19	12	23	28	20	10	18	15	12	19	10	13	25	10	13	37	17	14	12, 20	31	12	12	25	36, 37	11
176	19	14	21	30	23	10	15	15	11	19	11	12	25	10	13	42	15	14	13, 15	34	12	13	23	37, 37	10
177	19	13	21	30	21	11	17	17	11	17	12	15	24	10	13	43	18	14	13, 18	26	14	11	23	39, 39	10
178	18	12	21	27	23	10	16	15	11	18	11	16	25	11	13	41	14	14	13, 15	29	10	12	23	38, 39	10
179	16	12	21	28	22	10	17	16	12	18	10	15	23	11	13	37	17	15	12, 19	30	11	11	27	34, 40	11
180	18	12	19	29	23	10	15	15	12	18	11	16	23	10	14	37	18	14	13, 13	31	13	12	25	38, 40	11
181	18	14	21	30	22	11	15	15	11	18	10	16	24	10	13	42	16	14	15, 19	30	14	11	23	37, 37	11
182	17	14	20	31	20	11	16	15	12	18	10	15	24	10	13	38	18	14	12, 17	29	14	11	23	38, 40	11
183	16	14	21	29	21	10	16	15	11	18	10	14	24	10	13	38	19	14	12, 17	28	14	11	23	36, 39	10
184	18	11	23	27	24	10	17	15	12	21	10	15	25	10	13	40	20	15	13, 17	35	12	12	26	39, 40	11
185	18	13	21	28	24	10	17	17	11	17	12	15	24	10	13	42	18	14	13, 18	26	14	11	23	37, 39	10
186	18	14	23	31	21	11	15	14	11	19	11	16	24	11	14	36	18	14	11, 12	30	13	11	23	38, 40	12
187	18	12	19	29	23	10	17	15	13	18	10	17	23	10	14	37	19	14	13, 13	31	13	13	25	38, 40	11
188	18	14	23	31	18	11	17	15	13	20	10	15	25	11	11	44	19	14	11, 14	33	13	10	23	38, 38	11
189	19	12	22	28	23	10	18	15	12	19	9	14	25	10	13	35	18	15	12, 20	33	12	11	26	36, 39	11
190	18	13	21	28	22	9	18	16	12	19	10	16	22	10	13	39	17	15	11, 11	32	12	12	24	37, 40	12
191	18	13	21	30	19	10	18	16	11	18	8	14	22	10	13	40	17	14	13, 19	27	14	12	22	37, 39	10
192	16	13	21	29	20	9	20	15	10	20	11	15	23	10	11	38	20	12	13, 20	32	14	10	28	38, 40	11
193	19	13	22	29	22	10	16	15	11	18	11	15	26	10	13	39	16	14	15, 20	28	14	12	22	37, 38	10
194	19	12	24	29	20	10	18	15	12	19	10	13	25	10	13	36	17	14	12, 18	31	12	12	25	36, 37	12
195	18	14	21	30	22	10	15	15	11	19	10	16	24	10	13	43	22	14	14, 19	29	12	11	23	37, 37	11
196	20	12	21	28	18	9	20	14	12	19	9	15	24	11	14	36	17	15	13, 18	31	11	13	23	35, 38	11
197	17	14	23	31	20	10	20	17	11	18	11	15	24	10	13	43	16	14	12, 12	26	14	12	23	39, 40	10
198	18	13	23	29	23	10	17	16	12	21	10	15	25	10	13	40	20	15	13, 17	34	12	12	27	39, 39	11
199	19	11	23	27	19	10	19	14	12	20	10	15	25	10	13	39	19	15	13, 19	31	12	11	26	39, 40	11
200	18	12	21	28	18	13	17	15	12	20	10	15	24	11	14	36	20	15	12, 19	33	13	11	21	37, 37	11