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

Chuyên ngành: Công nghệ sinh học
Mã số: 9.42.02.01

TÓM TẮT LUẬN ÁN TIẾN SĨ CÔNG NGHỆ SINH HỌC

Hà Nội - 2023

Công trình được hoàn thành tại: Học viện Khoa học và Công nghệ - Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Người hướng dẫn khoa học 1: GS.TS. Chu Hoàng Hà

Người hướng dẫn khoa học 2: PGS.TS. Lê Văn Sơn

Phản biện 1: …

Phản biện 2: …

Phản biện 3: ….

Luận án sẽ được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện, họp tại Học viện Khoa học và Công nghệ - Viện Hàn lâm Khoa học và Công nghệ Việt Nam

vào hồi … giờ ...', ngày … tháng … năm 202...

Có thể tìm hiểu luận án tại:

- Thư viện Học viện Khoa học và Công nghệ
- Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tính cấp thiết của luận án

Bắt đầu từ năm 1953 với sự kiện Watson và Crick phát hiện ra cấu trúc xoắn kép của ADN cho đến nay, công nghệ sinh học phân tử đã có những bước phát triển vượt bậc và trở thành công cụ hữu hiệu cho nhiều ngành khoa học khác nhau như y học, nông nghiệp, khảo cổ học, pháp y… Đặc biệt, trong lĩnh vực giám định pháp y, khoa học hình sự công nghệ ADN cũng đóng vai trò quan trọng để giải quyết các vụ việc như điều tra xác định danh tính tội phạm, xét nghiệm các mối quan hệ huyết thống, tìm người mất tích...

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 …ADN là bằng chứng có tính pháp lý cao và có giá trị quyết định trong các phiên toà. Hơn thế nữa, trong những vụ thiên tai, thảm hoạ 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.

Việc 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 là điều quan trọng nhằm ứng dụng phân tích ADN trong khoa học điều tra, pháp y. Trong số đó, 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 đã 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 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 (NST) thường, việc nghiên cứu ứng dụng các locus STR trên NST 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 qua các thế hệ do đó giữa các cá thể nam giới cùng dòng cha sẽ có hồ sơ Y-STR giống nhau. Khi việc sử dụng các STR trên NST thường không thể đáp ứng được yêu cầu giám định, phân tích thêm hồ sơ Y-STR sẽ giúp giải quyết nhiều trường hợp, đặc biệt là giám định ADN đối với các mẫu từ nam giới như xác định mối quan hệ theo dòng cha (anh - em trai, chú – cháu…), các vụ án hiếp dâm với nghi phạm nam giới, lập phả hệ di truyền theo dòng cha…

Hiện nay 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) với 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***"

2. Mục tiêu nghiên cứu của luận án

- Lựa chọn và nghiên cứu 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.

- Nghiên cứu một số chỉ thị có kích thước nhỏ (mini STR) trên NST Y.

- Khảo sát tiềm năng 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).

3. Các nội dung nghiên cứu chính của luận án

Để đạt được mục tiêu của đề tài, chúng tôi đã thực hiện các nội dung nghiên cứu chính sau:

- Lựa chọn đối tượng nghiên cứu, tiến hành thu thập mẫu và tách chiết ADN.

- 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.

- 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.

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

CHƯƠNG 1. TỔNG QUAN

1.1. Phân tích ADN trong giám định pháp y

1.2. Các loại chỉ thị phân tử (marker)

1.3. Tổng quan chỉ thị STR

1.3.1. Đặc điểm trong bộ gen

1.3.2. Phân loại STR

1.4. STR trên nhiễm sắc thể giới tính Y

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)

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

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

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

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

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

1.6.3. Phân tích sử dụng chiến lược mini STR

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

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

1.8.1. Tình hình nghiên cứu trên thế giới

1.8.2. Tình hình nghiên cứu tại Việt Nam

CHƯƠNG 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

- Nghiên cứu khảo sát đa hình locus Y-STR thuộc bộ kit PowerPlex® Y23 System và Yfiler Plus: 400 mẫu nam giới dân tộc Kinh được thu thập ở miền bắc Việt Nam, không có quan hệ huyết thống với nhau. Mẫu được thu bao gồm: mẫu tóc, niêm mạc miệng hoặc mẫu máu

- Nghiên cứu trên mẫu đã bị phân huỷ: 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).

- Nghiên cứu trên mẫu lẫn (mixture sample): 50 mẫu dịch âm đạo thu từ nạn nhân nữ trong những vụ án hiếp dâm được cung cấp bởi cơ quan công an các tỉnh, thành phố trưng cầu giám định ADN tại Khoa Y sinh học – Viện Pháp y quốc gia.

2.2. Phương pháp nghiên cứu

2.2.1. Phương pháp tách chiết ADN

Đề tài sử dụng 3 phương pháp tách chiết ADN tùy thuộc từng loại mẫu và mục đích nghiên cứu:

- Phương pháp tách chiết ADN bằng Chelex® 100: Đây là phương pháp tách chiết sử dụng đối với các mẫu thông thường như: máu, lông, tóc và cả mẫu mô chưa bị phân hủy.

- Phương pháp tách chiết ADN bằng Qiamp® ADN 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 và khó tách chiết ADN.

- Phương pháp tách chiết ADN bằng QIAamp® ADN investigator kit của hãng QIAgen – Đức: Phương pháp này được sử dụng đối với các mẫu xương, răng hài cốt đã bị phân huỷ, có nồng độ ADN thấp, ADN bị đứt gãy nhiều.

2.2.2. Phương pháp định lượng ADN

- Với mẫu tách chiết theo phương pháp Chelex® 100: Đo nồng độ ADN sau tách chiết bằng máy Quantus Fluorometer (Promega) theo hướng dẫn của nhà sản xuất,

- Với mẫu tách chiết sử dụng kit: sau khi tách chiết ADN 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).

2.2.3. Phương pháp điện di gel

Chạy điện di phát hiện sản phẩm PCR trên gel Polyacrylamide 6%.

2.2.4. Phương pháp PCR

- Phân tích 29 chỉ thị Y-STR: theo bộ kit PowerPlex® Y23 System (hãng Promega - Mỹ) và Yfiler™ Plus PCR Amplification Kit (Thermo Fisher Scientific –Mỹ), thành phần phản ứng PCR với hai bộ kit tuân theo đúng hướng dẫn của nhà sản xuất.

- Phân tích 10 chỉ thị mini Y-STR: thiết kế cặp mồi và tối ưu chu trình phản ứng PCR

2.2.5. Phương pháp điện di mao quản

- Phân tích 29 chỉ thị Y-STR từ 2 bộ kit PowerPlex® Y23 System và YPlus theo hướng dẫn của nhà sản xuất

- Sản phẩm sau PCR được trộn với thang nội chuẩn và Hi-Di Formamide, được tra vào giếng trên đĩa 96 giếng, sốc ở nhiệt độ 95^{o}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

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

2.2.6. Phương pháp giải trình tự gen

Theo phương pháp giải trình tự gen Sanger với các bước chính: tinh sạch sản phẩm sau PCR, Thực hiện phản ứng PCR với BigDye Terminator v3.1 Cycle Sequencing Kit , tinh sạch sau chạy Bigdye, phân tích trên hệ thống máy ABI 3500Genetic Analyzer (Applied Biosystems, Thermo Scientific, Mỹ).

2.3. Phân tích và xử lý số liệu

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:

$$GD = n*(1- \Sigma\ pi2)/(n-1).$$

Haplotype diversity (HD) theo công thức: $HD = n*(1- \Sigma\ pi_{th}^2)/(n-1)$

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

Trong đó n là số mẫu, pi 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 webiste http://yhrd.org để tính toán khoảng cách di truyền quần thể (Rst) và giá trị xác suất kết hợp (P value) giữa quần thể nghiên cứu và các quần thể lân cận.

CHƯƠNG 3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Kết quả tách chiết ADN từ các mẫu 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 (DNase) 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. Mẫu sau tách chiết được định lượng bằng phương pháp đo độ hấp thụ trên máy Quantus Fluorometer (hãng Promega - Mỹ). Đây là phương pháp đo 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. Giá trị trung bình đo được là 2,70 ng/µl đến 10,05 ng/µl. 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 PowerPlex® Y23 System và YPlus..

Với các mẫu có chất lượng kém, mẫu đã bị phân huỷ nhiều như mẫu xương hài cốt, mẫu giám định án nồng độ ADN thu được sau tách chiết khá thấp nên khó áp dụng phương pháp điện di gel để kiểm tra. 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.

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	ADN đích	Nồng độ ADN	Chỉ số phân hủy
R1	T.Large Autosomal T.Small Autosomal T.Y	0.0572 0.0787 0.0721	1.376
R2	T.Large Autosomal T.Small Autosomal T.Y	0.0542 0.1028 0.0974	1.897
R3	T.Large Autosomal T.Small Autosomal T.Y	0.0124 0.0202 0.0196	1.633
R4	T.Large Autosomal T.Small Autosomal T.Y	0.0544 0.0947 0.0575	1.740
R5	T.Large Autosomal T.Small Autosomal T.Y	0.0665 0.1007 0.0975	1.513

Kết quả Realtime PCR của mẫu xương, răng hài cốt ký hiệu R1-R5 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 autsomal) lại cao hơn. Chỉ số phân huỷ (DI-Degradation Index) được tính bằng tỷ lệ giữa hàm lượng ADN kích thước nhỏ/ ADN kích thước lớn là >1 chứng tỏ mẫu hài cốt có ADN bị phân huỷ. Những mẫu như trên sẽ được xếp vào nhóm mẫu dùng trong nghiên cứu đánh giá hiệu quả phân tích Y-STR trên các mẫu đã phân huỷ nhiều.

3.2. Kết quả nghiên cứu 29 chỉ thị Y-STR từ 2 bộ kit PowerPlex® Y23 System và Yfiler Plus

3.2.1. Lựa chọn các chỉ thị Y-STR dùng trong nghiên cứu

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, 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 PowerPlex® Y23 System) và bộ Yfiler™ Plus PCR Amplification Kit (hãng Thermo Fisher Scientific, gọi tắt là YPlus).

3.2.2. Xây dựng hồ sơ Y-STR

Dựa theo hướng dẫn của nhà cung cấp chúng tôi đã thu được dữ liệu Y-STR từ 400 mẫu nghiên cứu đáp ứng các yêu cầu về hồ sơ Y-STR hoàn chỉnh. 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) hay nhận nhầm đỉnh giả. 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.

3.2.3. Bảng khảo sát tần suất 29 locus Y-STR thuộc bộ kit PowerPlex® Y23 System và YPlus

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.

Nhận xét: Như vậy từ việc lập hồ sơ Y-STR của 400 mẫu nghiên cứu với 2 bộ kit PowerPlex® Y23 System và YPlus chúng tôi đã lập được bảng phân bố tần suất alen riêng 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 3.2: Tần suất phân bố 29 locus Y-STR thuộc hai bộ kit PowerPlex® Y23 System (PPY23) và YPlus

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												

Bảng 3.3. Tần suất phân bố alen locus DYS385a/b và DYF387S1

DYS385a/b						**DYF387S1**	
Alen	PPY23	YPlus	Alen	PPY23	YPlus	Alen	YPlus
9, 17	0.005		13,20	0.040	0.035	34,34	0.010
11,11	0.015	0.005	13,21	0.005	0.010	34,36	0.005
11,12		0.015	13,22	0.010	0.005	34,37	0.005
11,13	0.005		14,14		0.005	34,40	0.005
11,14	0.005	0.005	14,15	0.010		35,36	0.005
11,15	0.005		14,16	0.005	0.015	35,37	0.010
11,17	0.005	0.005	14,17	0.005		35,38	0.045
11,18	0.035	0.020	14,18	0.025	0.010	35,39	0.015
11,19	0.010		14,19	0.020	0.025	35,40	0.010
11,20	0.010	0.005	14,20	0.010		36,36	0.040
12,12	0.025	0.010	14,21	0.015		36,37	0.050
12,12.3	0.005		15,15		0.010	36,38	0.100
12,13	0.005	0.010	15,16	0.005	0.010	36,39	0.070
12,14	0.005	0.015	15,17		0.010	36,40	0.025
12,15		0.005	15,18	0.030	0.005	36,41	0.015
12,16	0.015	0.015	15,19	0.020	0.060	37,37	0.090
12,17	0.020	0.020	15,20	0.015	0.025	**37,38**	**0.140**
12,18	0.050	0.030	15,21	0.015	0.015	37,39	0.080
12,19	0.035	0.035	15,22		0.005	37,40	0.015
12,20	0.040	0.055	15,23		0.005	37,41	0.010
12,21	0.010		16,17	0.005		38,38	0.095
13,13	16.000	0.080	16,18	0.005		38,39	0.040
13,14	3.000	0.015	16,19	0.005	0.015	38,40	0.030
13,15		0.015	16,20	0.020	0.015	39,39	0.040
13,16	0.005	0.015	16,21	0.015	0.005	39,40	0.020
13,17	0.050	0.045	17,20	0.010		39,41	0.005
13,18	**0.150**	**0.185**	18,19		0.005	40,40	0.020
13,19	0.120	0.105				40,41	0.005

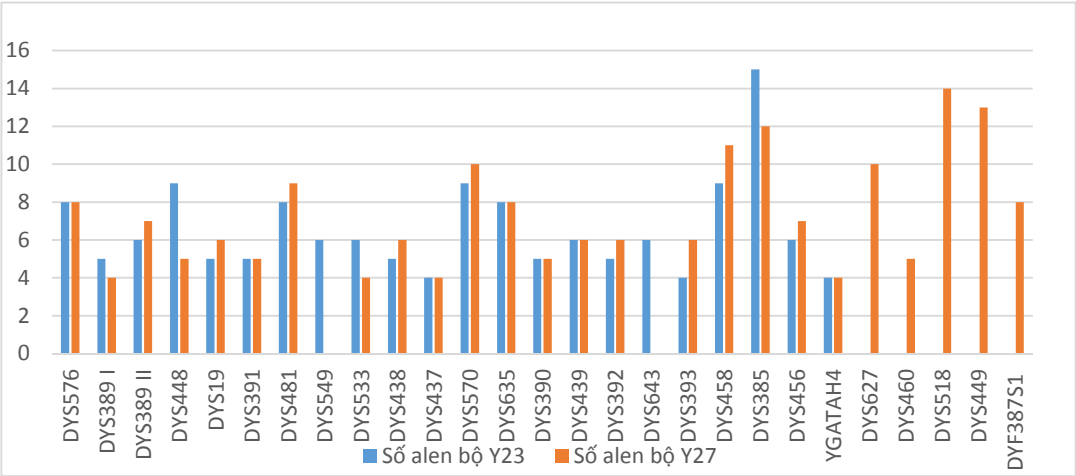
Bảng tần suất phân bố Y-STR được xây dựng trong nghiên cứu này là bảng có số lượng mẫu nghiên cứu cùng như số locus nhiều nhất được khảo sát trong quần thể người Việt Nam cho đến nay với số lượng mẫu và phương pháp nghiên cứu đạt độ tin cậy cao. Các công bố trước đây đối với quần thể người Việt Nam chủ 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 PowerPlex® Y23 System 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 chỉ với 45 haplotype từ bộ PowerPlex® Y23 System và 47 halotype với bộ YPlus được đóng góp thông một số nghiên cứu nhỏ lẻ trên quần thể người Việt Nam. Với nghiên cứu này chúng tôi đã đóng góp thêm 200 haptype sử dụng bộ PowerPlex® Y23 System vào dữ liệu chung YHRD với mã số truy cập là YA004576 nâng tổng số hapotype trên dữ liệu YHRD ở quần thể người Việt Nam lên 245 với bộ PowerPlex® Y23 System.

3.2 4. Đánh giá đặc điểm các alen thuộc 29 chỉ thị Y-STR

Với từng locus số lượng alen dao động từ 4 đến 15 alen với trung bình là 7 alen / 1 locus. Trong đó với bộ kit PowerPlex® Y23 System có tổng số 144 alen, bộ kit YPlus là 183 alen. Với riêng tổ hợp gồm 2 alen, DYS385a/b có 55 tổ hợp alen khác nhau, DYF387S1a/b 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 PowerPlex® Y23 Systemđồng thời cũng là các locus có số lượng alen nhiều nhất (DYS518: 14 alen, DYS449: 13 alen).

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 của tác giả Toàn T.T ghi nhận 99 alen. Số alen trong nghiên cứu cũng ở mức tương đương khi so sánh với quần thể khác trên thế giới. 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. 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.

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 locus DYS448 (alen 18.2, 19.2), 1 biến thể tại locus DYS385b (alen 12.3). Trong quá trình khảo sát cũng quan sát thấy 2 trường hợp mất alen (còn gọi là null alen). Với trường hợp đầu tiên phát hiện mất 18/23 locus Y-STR khi phân tích với bộ kit PowerPlex® Y23 System. Trường hợp còn lại là có null alen tại 4 locus DYS481, DYS576, DYS570, DYS458. 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 nhận thấy đây là trường hợp mất đoạn lớn trên cả cánh dài và cánh ngắn của NST Y.



Hình 3.1: Số alen của 29 locus Y-STR dựa trên 2 bộ kit PowerPlex® Y23 System và YPlus

3.2.5. 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 dạng gen (GD - gene diversity) của từng locus Y-STR trong nghiên cứu. Kết quả cho thấy hai locus với tổ hợp 2 alen là DYS385a/b và DYS387S1a/b là các locus có độ da dạng di truyền cao nhất (GD lần lượt 0.941, 0.934). Với các locus đơn alencó 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. 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 Ádựa theo nghiên cứu khảo sát tần suất 23 Y-STR từ 129 quần thể thuộc 51 quốc gia của Josephine Purps và cs. 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. 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.

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.

Bảng 3.4 . 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 locus	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
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 sắp xếp theo độ đa hình của 29 locus Y-STR trong hình 3.2 cho thấy số locus có độ đa hình từ 0.7-1 chiếm tỷ lệ nhiều nhất với 84.2% (bao gồm 16/29 locus với 5 locus có giá trị GD> 0.9, 4 locus có 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 ≤ 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 các cá thể nam giới trong quần thể người Kinh, Việt Nam được khảo sát.

3.2.6. 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 cá 2 bộ kit PowerPlex® Y23 System 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 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 Yfiler cũng cho khả năng phân biệt 100%. 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%. Điều này một lần nữa 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 PowerPlex® Y23 System 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%).

3.3. Nghiên cứu một số chi thi mini Y-STR

3.3.1. Lựa chọn locus mini Y-STR

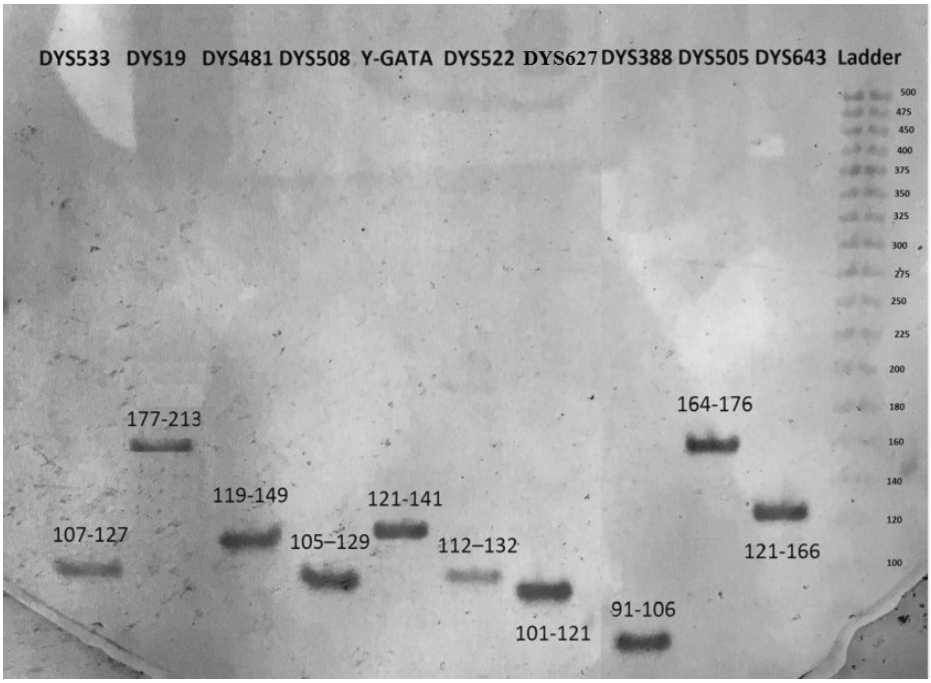
Do yêu cầu trong các bộ kít thương mại phải khuếch đại đồng thời nhiều locus Y-STR khác nhau nên để phân biệt được trong quá trình điện di mao quản đòi hỏi các locus phải có kích thước khác nhau. Kết quả là một số locus Y-STR có kích thước khá lớn (>200 bp) rất khó để phân tích thành công với các mẫu có ADN đã bị phân hủy, ADN bị đứt gãy nhiều. 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 huỷ 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 locus Y-STR mới có tiềm năng ứng dụng trong giám định ADN với quần thể người Việt Nam.

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. 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 PowerPlex® Y23 System. Bên cạnh đó chúng tôi cũng thiết kế mồi 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..

3.3.2. Tối ưu hoá điều kiện phản ứng PCR

Phản ứng PCR được khảo sát với 10 cặp mồi trong 10 phản ứng PCR riêng rẽ. Nhiệt độ gắn mồi của từng locus dao động từ 56 – 58 ^{o}C. Chu trình phản ứng PCR được tối ưu hóa với các điều kiện bao gồm: nồng độ mồi, nhiệt độ gắn mồi, thời gian kéo dài chuỗi. Sản phẩm sau PCR được kiểm tra trên gel Polyacryamide 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 mồi được thiết kế có tính khả dụng cho việc khuếch đại các locus Y-STR mong muốn.

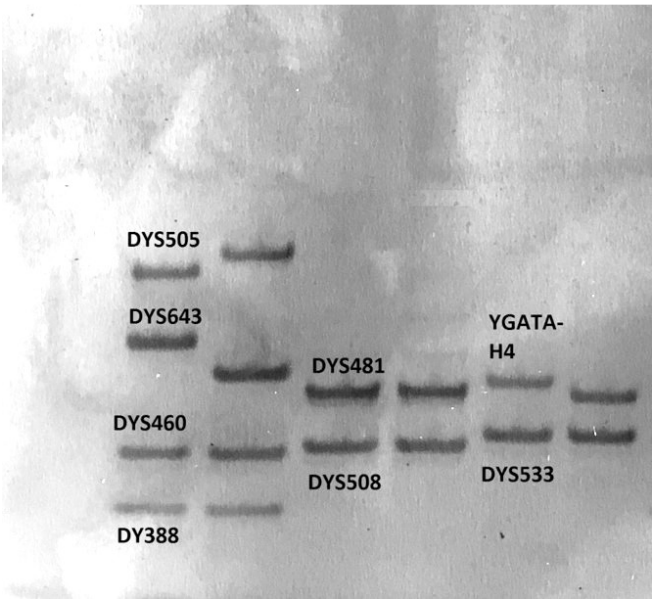
Phản ứng PCR được khảo sát trong 10 phản ứng PCR riêng rẽ. Chu trình phản ứng PCR được tối ưu hóa với các điều kiện bao gồm: nồng độ mồi, nhiệt độ gắn mồi, thời gian kéo dài chuỗi. Sản phẩm sau PCR được kiểm tra trên gel Polyacryamide 6%.



Hình 3.2: Ảnh điện di sản phẩm sau PCR với 10 locus Y-STR trên gel Polyacrylamide 6% (Sử dụng thang chuẩn ILS 500 (Promega – Mỹ))

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 mồi được thiết kế dựa trên việc trộn nhiều cặp mồi 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 mồi 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. 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 mồi không vượt quá 4 locus.

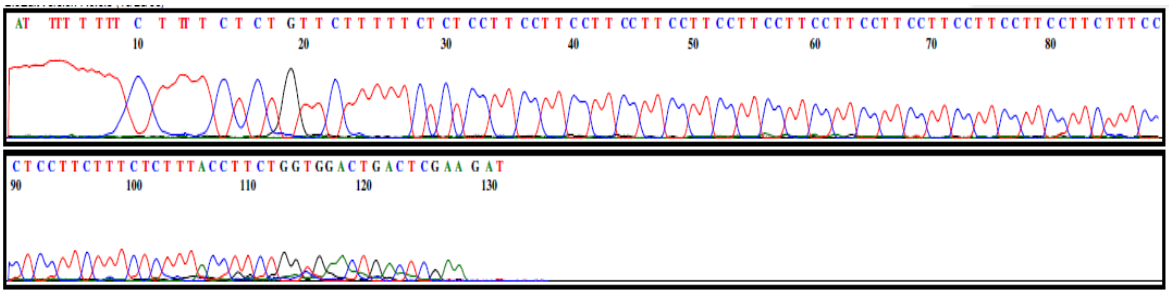
Chúng tôi đã tối ưu nồng độ mồi, nhiệt độ gắn mồi và khuếch đại thành công 8 trong số 10 locus Y-STR thông qua 3 phản ứng PCR đa mồi khác nhau. Phản ứng multiplex PCR với nhiều cặp mồi cùng cũng là cơ sở để tạo nên các bộ kit STR và Y-STR thương mại hiện nay có thể phân tích lên tới 27 locus thông qua một phản ứng PCR duy nhất



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

3.3.3. Giải trình tự xác định cấu trúc lặp của các locus Y-STR mới

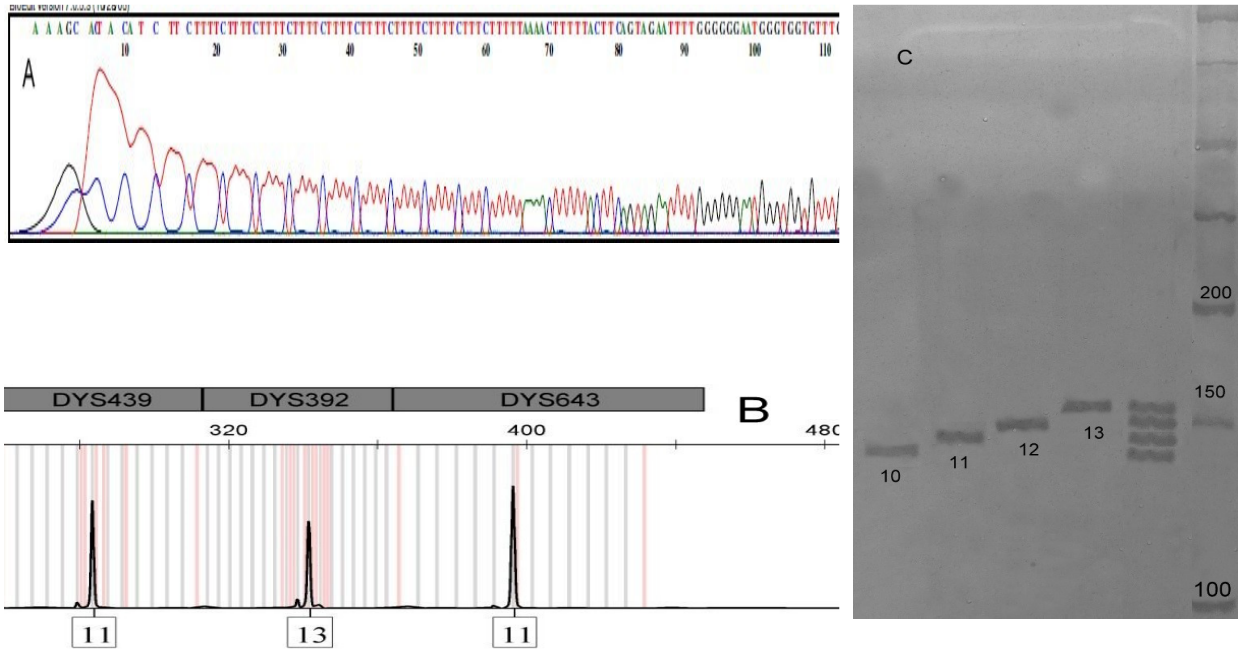
Kết quả giải trình tự trên máy ABI 3500, Pop7, Cappilarry 50 cm với chương trình chạy của hãng cho trình tự có các đỉnh sóng rõ ràng. Việc giải trình tự cũng cho phép xác định rõ motif lặp, số lần lặp đặc trưng của từng locus Y-STR. Ví dụ locus DYS505 có cấu trúc lặp là TCCT với số lần lặp là 15 hay chính là alen 15 (hình 3.4).



Hình 3.4. 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. Việc giải trình tự cũng cho phép xác định rõ motif lặp, số lần lặp đặc trưng của từng locus Y-STR. Ví dụ locus DYS505 có cấu trúc lặp là TCCT với số lần lặp là 15 hay chính là alen 15 (hình 3.4). Đây cũng là cơ sở để xây dựng thang alen cho các locus mới chưa có trong các bộ kit thương mại. 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 locus DYS643 (Hình 3.5). Trong nghiên cứu này với các mẫu đã được xác định rõ alen bằng chúng tôi đã xây dựng thang alen của chỉ thị DYS643 (đã có trong bộ PowerPlex® Y23 System) dựa vào phương pháp mini Y-STR. 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 3.5. Kết quả tạo thang chuẩn mini Y-STR với chỉ thị DYS643

A. Kết quả GTT với số lần lặp CTTTT là 11, B. Kết quả phân tích sử dụng bộ PowerPlex® Y23 System với số alen là 11, C: Thang chuẩn chỉ thị DYS643 với 4 alen 10, 11,12, 13

3.3.4. Kết quả phân tích 10 mini Y-STR trên mẫu phân hủy

Dựa vào phương pháp PCR đa mồi với 10 cặp mồi nêu trên chúng tôi đã ứng dụng trong việc phân tích 50 mẫu nghiên cứu đã bị phân hủy là những mẫu xương và răng hài cốt liệt sỹ thu thập từ các nghĩa trang. 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%.

Bảng 3. 5. 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%

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.

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. Dựa theo kết quả nghiên cứu có thể thấy có khoảng ít nhất 9 mẫu phân tích không thành công khi sử dụng mini STR. 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.

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 mồi tương ứng, chúng tôi đã thử ứng dụng vào một số trường hợp giám định cụ thể tại 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. Những trường hợp này khi xét nghiệm so sánh giữa hồ sơ mẫu thu tại hiện trưởng hoặc mẫu thu của nạn nhân có hồ sơ Y-STR bằng 2 bộ PowerPlex® Y23 System và YPlus không lên đủ ở tất cả các locus, đặc biệt là các locus có kích thước lớn nên không đủ cơ sở để so sánh với mẫu nghi phạm. Khi đó việc phân tích thêm các mini Y-STR sẽ giúp bổ sung dữ liệu vào hồ sơ Y-STR, cho phép đưa ra kết luận mẫu thu tại hiện trường (nạn nhân) là trùng/khác với mẫu thu từ nghi phạm. Đây là biện pháp có hiệu quả trong một số vụ án đã được giám định.

3.4.1.2. Trong công tác giám định danh tính 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. 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 mini Y-STR sẽ cung cấp dữ liệu bổ sung có giá trị quan trọng trong việc đưa ra kết luận cuối cùng.

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 PowerPlex® Y23 System và YPlus cũng đều gồm một số các alen vừa có kích thước nhỏ vừa có độ đa dạng gen 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… Để đánh giá tiềm năng phân tích với mẫu đã bị phân huỷ chúng tôi lọc riêng dữ liệu tập hợp các Y-STR có kích thước < 220 bp thuộc 2 bộ kit khi phân tích với 400 mẫu nghiên cứu và tính toán các chỉ số di truyền liên quan.

Bảng 3.6. So sánh các chỉ số di truyền thu được từ tập hợp các marker Y-STR có kích thước < 220 bp trong 2 bộ PowerPlex® Y23 System (PPY23) và YPlus

Số haplotype	Bộ haplotype	Bộ haplotype
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 cá thể	0	0
Xuất hiện ở 9 cá thể	0	0
Xuất hiện ở 10 cá thể	0	0
Tổng số locus	8	10
Tổng số haplotype quan sát	193	185
HD	0.99954	0.99829
DC	97%	93%

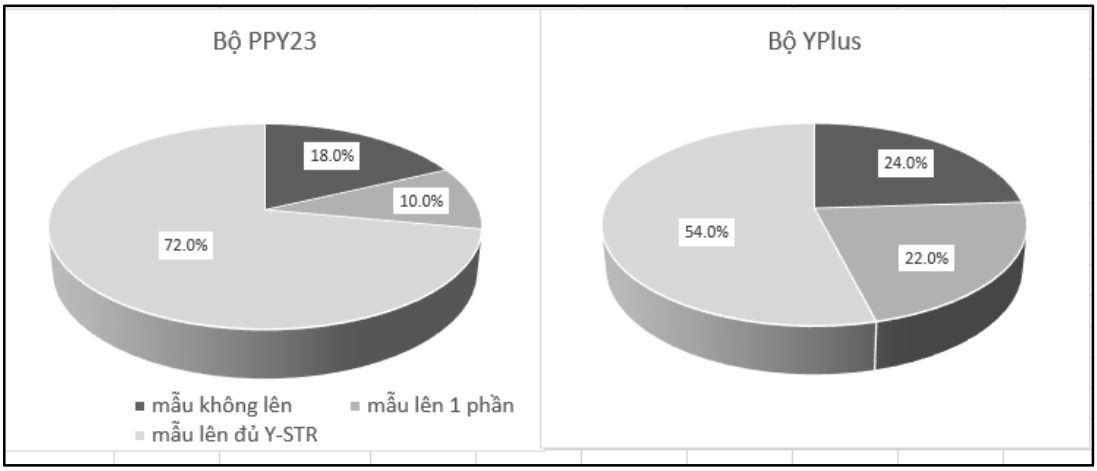
Bộ PowerPlex® Y23 System gồm 8 locus có kích thước < 220 bp (DYS576, DYS389I, DYS391, DYS481, DYS570, DYS635, DYS393 và DYS458). Bộ YPlus gồm 10 locus có kích thước < 220 bp (DYS576, DYS389I, DYS458, DYS456, DYS390, DYS570, DYS437, DYS393, DYS439 và DYS460). Kết quả trong Bảng 3.5 cho thấy với tập hợp Y-STR thuộc bộ PowerPlex® Y23 System xuất hiện 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 PowerPlex® Y23 System và YPlus lần lượt là 193, 185 haplotypes (trên tổng số 200 mẫu

phân tích). Khả năng phân biệt của bộ PowerPlex® Y23 System là 97% cũng nhiều hơn đáng kể so với bộ YPlus là 93%. Như vậy dù có số lượng Y-STR ít hơn nhưng tập hợp Y-STR kích thước < 220 bp trong bộ PowerPlex® Y23 System lại có khả năng phân biệt tốt hơn so với tập hợp trong bộ YPlus. Điều này được lý giải do trong số các locus có kích thước ngắn thuộc bộ PowerPlex® Y23 System có 2 locus có độ đa hình cao là DYS481 (GD: 0.77) và DYS635 giúp tăng đáng kể khả năng phân biệt.

Kết quả phân tích khả năng tạo lập hồ sơ Y-STR cũng gợi ý việc thiết kế và lựa chọn các Y-STR trong bộ kit thương mại để đảm bảo vừa có khả năng phân tích mẫu có ADN bị phân huỷ, ADN nồng độ thấp vừa có khả năng tạo haplotype có độ đa dạng cao. Có thể nói bên cạnh các STR trong nhân, Y-STR là các chỉ thị lý tưởng trong việc xác định quan hệ huyết thống, định danh cá thể, điều tra hình sự.

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

Trong nghiên cứu này, chúng tôi đã thử phân tích 50 mẫu lẫn với cả 2 bộ kit PowerPlex® Y23 System 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/tấn công tình dục 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.



Hình 3.6. Đồ thị so sánh khả năng tạo hồ sơ Y-STR sử dụng bộ PowerPlex® Y23 System và Yplus

Kết quả cho thấy việc sử dụng bộ kit PowerPlex® Y23 System cho số lượng hồ sơ Y-STR hoàn chỉnh nhiều hơn bộ kit YPlus. Ngược lại số mẫu chỉ lên được 1 phần và không lên được tất cả các locus Y-STR của bộ kit YPlus lại nhiều hơn so với bộ PowerPlex® Y23 System. Có một số nguyên nhân lý giải kết quả. Thứ nhất, do số lượng locus Y-STR trong bộ kit YPlus nhiều hơn bộ PowerPlex® Y23 System nên về lý thuyết tỷ lệ khuếch đại thành công tất cả các Y-STR sẽ thấp hơn. Thứ hai, 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. Ngoài ra 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. PowerPlex® Y23 System và YPlus cho thấy độ nhạy tương tự nhưng PowerPlex® Y23 System cho thấy khả năng chịu axit humic cao hơn YPlus theo nghiên cứu trước đó. Do đó, PowerPlex® Y23 System 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.

Như vậy với việc phân tích thêm dữ liệu Y-STR đã giúp bổ sung thông tin, tăng tỷ giải quyết thành công trong các trường hợp có mẫu lẫn, hỗ trợ đáng kể cho việc đưa ra kết luận cho vụ án.

3.4.3. 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 PowerPlex® Y23 System và YPlus đã bổ sung thêm 12 locus. 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 PowerPlex® Y23 System (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. 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 PowerPlex® Y23 System bổ sung 02 RM Y-STR là DYS570, DYS576; bộ kit YPlus bổ sung cả 7 RM Y-STR.. 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). Để đánh giá hiệu quả của các locus mới được bổ sung trong 2 bộ kit PowerPlex® Y23 System 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 (SWGDAM): 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 PowerPlex® Y23 System và bộ YPlus: bao gồm 17 locus nêu trên và thêm các locus mới.

Bảng 3. 7. 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 PowerPlex® Y23 System (PPY23)

Số haplotype quan sát được	Bộ haplotype tối thiểu (9 locus)	Bộ haplotype SWGDAM (11 locus)	Bộ Yfiler (17 locus)	Bộ PPY23 (23 locus)
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
DC	86%	91%	98%	100%

Kết quả so sánh 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 PowerPlex® Y23 System và YPlus.

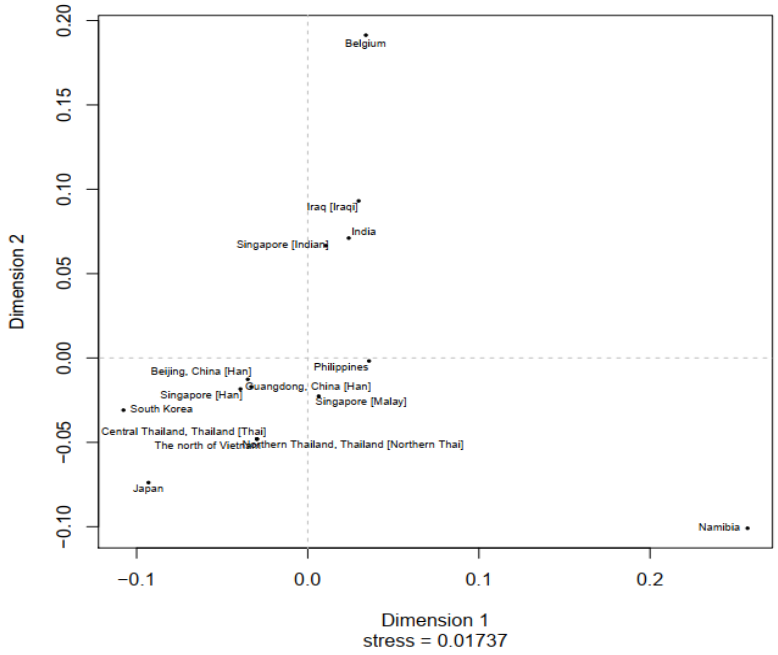
Từ các kết quả của nghiên cứu chúng tôi cũng đã mạnh dạn áp dụng việc phân tích các chỉ thị Y-STR từ hai bộ kit PowerPlex® Y23 System và Yfiler Plus trong các công tác giám định tại Khoa Y sinh học và đạt được nhiều kết quả. Trong giám định huyết thống, các chỉ thị Y-STR được sử dụng giám định trong 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. 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) 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. 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 Rst) 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 18 quần thể khác trên thế giới.

Giá trị Rst càng nhỏ và P-value càng lớn thể hiện khoảng cách di truyền gần và không có ý nghĩa khác biệt giữa các quần thể. Kết quả 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 Rst > 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 (Rst =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. 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 trong không gian đa chiều nhờ công cụ AMOVA (hình 3.7).



Hình 3.7. 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 tương đồng, ít có sai khác trong cấu trúc di truyền.

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.

CHƯƠNG IV. KẾT LUẬN VÀ KIẾN NGHỊ

Kết luận

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) cũng 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 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 mồi gắn huỳnh quang và điện đi mao quản để tạo nên bộ kit riêng cho người Việt Nam.

NHỮNG 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 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, DYF387S1, 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 lẫn giữa nam và nữ, giám định mẫu lâu năm, mẫu đã bị phân huỷ. 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.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

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 Nhự, 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 Yfiler Plus PCR Amplication 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.