

**BỘ GIÁO DỤC VÀ ĐÀO TẠO      VIỆN HÀN LÂM KHOA HỌC  
VÀ CÔNG NGHỆ VIỆT NAM  
VIỆN SINH THÁI VÀ TÀI NGUYÊN SINH VẬT**

**LƯƠNG THỊ PHÚC**

**KHẢO SÁT VÀ XÂY DỰNG CƠ SỞ DỮ LIỆU  
TẦN SUẤT CÁC ALEN Ở 15 LOCUS GEN  
BẰNG HỆ IDENTIFILER TỪ QUẦN THỂ NGƯỜI  
DÂN TỘC NÙNG ỨNG DỤNG TRONG GIÁM ĐỊNH ADN**

**LUẬN VĂN THẠC SỸ SINH HỌC**

**HÀ NỘI - NĂM 2015**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO      VIỆN HÀN LÂM KHOA HỌC  
VÀ CÔNG NGHỆ VIỆT NAM  
VIỆN SINH THÁI VÀ TÀI NGUYÊN SINH VẬT**

**LƯƠNG THỊ PHÚC**

**KHẢO SÁT VÀ XÂY DỰNG CƠ SỞ DỮ LIỆU  
TẦN SUẤT CÁC ALLEN Ở 15 LOCUS GEN  
BẰNG HỆ IDENTIFILER TỪ QUẦN THỂ NGƯỜI  
DÂN TỘC NÙNG ỨNG DỤNG TRONG GIÁM ĐỊNH ADN**

**Chuyên ngành: Sinh học thực nghiệm**

**Mã số: 60 42 01 14**

**LUẬN VĂN THẠC SĨ SINH HỌC**

**Người hướng dẫn: PGS, TS. Nguyễn Văn Hà**

**HÀ NỘI - NĂM 2015**

## LỜI CẢM ƠN

Lời đầu tiên, tôi xin được bày tỏ lòng biết ơn sâu sắc tới PGS. TS Nguyễn Văn Hà - Phó giám đốc phụ trách Trung tâm Giám định Sinh học Pháp lý - Viện Khoa học hình sự - Bộ Công An là người thầy đã dẫn dắt, tạo mọi điều kiện giúp đỡ để tôi hoàn thành luận văn này. Đồng thời, tôi xin chân thành cảm ơn Trung tá Trịnh Tuấn Toàn – Phó Giám đốc Trung tâm Giám định Sinh học Pháp lý – Viện Khoa học hình sự và các thành viên trong dự án Xây dựng tàng thư gen tội phạm Quốc gia đã giúp đỡ tôi rất nhiều trong quá trình làm luận văn.

Tôi cũng gửi lời cảm ơn tới các cán bộ thuộc Trung tâm Giám định Sinh học Pháp lý – Viện Khoa học hình sự đã tạo điều kiện giúp đỡ về kỹ thuật và cơ sở vật chất cho tôi trong quá trình làm luận văn.

Để hoàn thành luận văn này tôi xin được gửi lời cảm ơn chân thành tới các giảng viên, CBCNV Viện Sinh thái và Tài nguyên Sinh vật - Viện Hàn lâm Khoa học Công nghệ Việt Nam đã tận tình truyền đạt kiến thức và giúp đỡ tôi trong suốt quá trình học tập.

Cuối cùng tôi xin bày tỏ lòng biết ơn đến gia đình, bạn bè, đồng nghiệp đã động viên, giúp đỡ tôi trong suốt quá trình hoàn thành luận văn thạc sĩ này.

*Hà nội ngày tháng năm 2015*

**Học viên**

**Lương Thị Phúc**

## MỤC LỤC

<b>MỞ ĐẦU</b> .....	1
<b>Chương 1. TỔNG QUAN TÀI LIỆU</b> .....	5
1.1. Các phương pháp truy nguyên cá thể dùng trong Khoa học hình sự .....	5
1.3. Mười lăm (15) locus gen hệ Identifiler trên các nhiễm sắc thể .....	8
1.3.1 Các tiêu chuẩn cho locus STR dùng trong giám định ADN .....	8
1.3.2. Ý nghĩa của cơ sở dữ liệu tần suất alen của các locus STR.....	10
1.3.3. Cơ sở khoa học và thực tiễn của việc nghiên cứu đề tài .....	11
1.4. Tình hình nghiên cứu và phát triển các bộ kit dùng trong giám định ADN nhân tế bào.....	12
1.4.1. Tình hình nghiên cứu trên thế giới.....	12
1.4.2. Tình hình nghiên cứu trong nước.....	13
1.5. Dân tộc Nùng sinh sống trên lãnh thổ Việt Nam.....	15
<b>Chương 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU</b> .....	20
2.1. Vật liệu nghiên cứu .....	20
2.1.1. Đối tượng nghiên cứu .....	20
2.1.2. Hóa chất, thiết bị và dụng cụ nghiên cứu.....	20
2.2. Phương pháp nghiên cứu .....	22
2.2.1. Quy trình phân tích ADN .....	22
2.2.2. Xử lý thống kê số liệu và tính tần suất các locus gen .....	23
<b>Chương 3. KẾT QUẢ NGHIÊN CỨU VÀ BÀN LUẬN</b> .....	25
3.1. Phân tích kiểu gen của 15 locus gen từ 120 mẫu thu thập.....	25
3.1.1 Kết quả thu, bảo quản mẫu và tách chiết ADN.....	25
3.1.2 Kết quả thực hiện phản ứng PCR, điện di và phân tích kiểu gen .....	25
3.2. Tính tần suất của 15 locus gen từ 120 mẫu thu thập được .....	30
3.2.1. Kết quả tính tần suất các alen.....	30

3.2.2. So sánh chỉ số kiểm định $\chi^2$ giữa tính toán ( $\chi^2_{tt}$ ) với lý thuyết ( $\chi^2_{lt}$ ) và đánh giá sự phù hợp với định luật Hardy - Weinberg.....	30
3.2.3. So sánh tần suất alen của người Nùng với một số dân tộc .....	31
3.3. Một số ví dụ về ứng dụng kết quả của đề tài trong công tác giám định ADN tại Viện Khoa học hình sự - Bộ Công an .....	47
<b>KẾT LUẬN</b> .....	49
<b>KIẾN NGHỊ</b> .....	50
<b>TÀI LIỆU THAM KHẢO</b> .....	51
<b>PHỤ LỤC</b>	

## DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

ADN	- Axit Deoxyribo Nucleic
ARN	- Axit ribonucleic
A	- Adenine
T	- Thymine
G	- Guanine
C	- Cytosine
NST	- Nhiễm sắc thể
VNTR	- Variable Number of Tandem Repeat - Các trình tự lặp ngắn
STR	- Short Tandem Repeat - Các trình tự lặp ngắn
ID	- Identifiler/Identify definition
PCR	- Polymerase Chain Reaction - Phản ứng chuỗi trùng hợp
FTA	- Tên riêng của một vật mang thu mẫu máu (dạng thẻ) trong khoa học hình sự
bp	- Base pair
KHHS	- Khoa học hình sự

## DANH MỤC HÌNH VẼ VÀ BẢNG BIỂU

### Hình vẽ

Hình 3.1: Hình ảnh kiểu gen dạng peak từ mẫu ký hiệu N14.....	27
Hình 3.2: Hình ảnh kiểu gen dạng peak từ mẫu ký hiệu N40.....	28

### Bảng biểu

Bảng 3.2: So sánh $\chi^2_{TT}$ và $\chi^2_{LT}$ .....	30
Bảng 3.3: Tần suất alen (%) của locus D8S1179.....	31
Bảng 3.4: Tần suất alen (%) của locus D21S11.....	32
Bảng 3.5: Tần suất alen (%) của locus D7S820.....	33
Bảng 3.6: Tần suất alen (%) của locus CSF1PO .....	34
Bảng 3.7: Tần suất alen (%) của locus D3S1358.....	35
Bảng 3.8: Tần suất alen (%) của locus THO1 .....	36
Bảng 3.9: Tần suất alen (%) của locus D13S317.....	37
Bảng 3.10: Tần suất alen (%) của locus D16S539.....	38
Bảng 3.11: Tần suất alen (%) của locus D2S1338.....	39
Bảng 3.12: Tần suất alen (%) của locus D19S433.....	40
Bảng 3.13: Tần suất alen (%) của locus vWA .....	41
Bảng 3.14: Tần suất alen (%) của locus TPOX .....	42
Bảng 3.15: Tần suất alen (%) của locus D18S51.....	43
Bảng 3.16: Tần suất alen (%) của locus D5S818.....	44
Bảng 3.17: Tần suất alen (%) của locus FGA.....	45

## MỞ ĐẦU

### 1. Tính cấp thiết của đề tài

Viện Khoa học hình sự là một đơn vị đầu ngành của lực lượng kỹ thuật hình sự, một trung tâm khoa học của ngành Công an Việt Nam. Một trong những nhiệm vụ quan trọng của Viện Khoa học hình sự là công tác giám định phục vụ tố tụng hình sự, tố tụng dân sự và các yêu cầu cá nhân, trong đó nổi bật là lĩnh vực giám định ADN.

Năm 2006 Viện Khoa học hình sự đã đưa vào ứng dụng hệ Identifiler (gồm 15 locus: D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, D5S818 và FGA) trong giám định ADN.

Trong giám định ADN hình sự đòi hỏi bắt buộc phải tính xác suất một người ngẫu nhiên trong quần thể có cấu trúc di truyền trùng lặp với ADN của các mẫu vật giám định. Để tính toán được xác suất này thì phải có dữ liệu tần suất của từng alen [5].

Theo lý thuyết di truyền học, mỗi quần thể người (dân tộc) khác nhau có những đặc điểm di truyền đặc trưng, thể hiện bằng sự phân bố tần suất alen trong mỗi quần thể là khác nhau và không thể áp dụng cơ sở dữ liệu của quần thể này cho một quần thể khác. Do đó, bắt buộc phải tiến hành khảo sát tần suất các alen của các locus dùng trong giám định ADN hình sự đối với mỗi dân tộc để đảm bảo tính khoa học, chính xác và khách quan trong kết luận giám định. Việc khảo sát tần suất các alen của các locus đang sử dụng trong giám định ADN đối với các dân tộc trên toàn lãnh thổ Việt Nam là một việc làm mang tính cấp bách.

Việt Nam có 54 dân tộc, trong đó dân tộc Kinh chiếm tỉ lệ lớn nhất (gần 90%) [1], còn lại là các dân tộc thiểu số. Tính đến năm 2014, Viện Khoa học



hình sự đã công bố bảng tần suất các alen hệ Identifiler đối với các dân tộc Kinh, H'mông, Tày và Khmer. Do điều kiện về cơ sở vật chất cũng như con người chưa đáp ứng được nên chưa thể tiến hành khảo sát tần suất alen cho tất cả các dân tộc.

Theo Tổng điều tra dân số và nhà ở năm 2009, người Nùng ở Việt Nam có dân số 968.800 người, là dân tộc có dân số đứng thứ 7 tại Việt Nam, có mặt trên tất cả 63 tỉnh, thành phố. Người Nùng cư trú tập trung tại các tỉnh: Lạng Sơn (314.295 người, chiếm 42,9% dân số toàn tỉnh và 32,4% tổng số người Nùng tại Việt Nam), Cao Bằng (157.607 người, chiếm 31,1% dân số toàn tỉnh và 16,3% tổng số người Nùng tại Việt Nam), Bắc Giang (76.354 người), Đắk Lắk (71.461 người), Hà Giang (71.338 người), Thái Nguyên (63.816 người), Bắc Kạn (27.505 người), Đắk Nông (27.333 người), Lào Cai (25.591 người), Lâm Đồng (24.526 người), Bình Phước (23.198 người)...

Người Nùng chủ yếu phân bố ở các tỉnh đông bắc Bắc Bộ như Lạng Sơn, Cao Bằng, Bắc Kạn, Thái Nguyên, Bắc Giang, Tuyên Quang v.v (chiếm tới 84%). Đây là các tỉnh gần với biên giới Trung Quốc, là những điểm nóng về tình hình An ninh và trật tự an toàn xã hội. Hiện nay, số lượng vụ án có đối tượng là người dân tộc Nùng có xu hướng ngày càng tăng (số lượng vụ việc trong các năm gần đây được trưng cầu giám định AND tại Viện khoa học hình sự cụ thể như sau: Năm 2013 số vụ là 25 vụ, năm 2014 là 32 vụ, 6 tháng đầu năm của năm 2015 là 20 vụ). Tuy nhiên, ở Việt Nam cho đến nay chưa có một nghiên cứu hoàn chỉnh nào được công bố về khảo sát sự phân bố tần suất các alen của các gen hình sự hệ Identifiler (gồm 15 gen) đối với quần thể người dân tộc Nùng.

Yêu cầu thực tế đặt ra đòi hỏi phải có tần suất các alen của các locus gen hệ Identifiler để dùng trong tính toán, đưa ra kết luận trong các vụ án có mẫu của đối tượng là người dân tộc Nùng ngày càng cấp bách. Đề tài “*Khảo*

sát và xây dựng cơ sở dữ liệu tần suất các alen của 15 locus gen hệ Identifiler từ quần thể người dân tộc Nùng ứng dụng trong giám định ADN” được hoàn thành sẽ đáp ứng yêu cầu cấp thiết của giám định ADN và đóng góp dữ liệu vào hệ thống gen hình sự quốc tế. Kết quả của đề tài là cơ sở khoa học vững chắc để phân tích, đánh giá và kết luận giám định đối với những vụ án liên quan đến người thuộc dân tộc Nùng và là điều kiện để xây dựng cơ sở pháp lý cho quy trình giám định ADN.

## **2. Mục tiêu của đề tài**

Khảo sát và xây dựng cơ sở dữ liệu tần suất các alen của 15 locus gen (ADN) hệ Identifiler từ dân tộc Nùng phục vụ cho công tác giám định ADN trong các vụ án liên quan đến người dân tộc Nùng. Trên cơ sở đó, tính tần suất xuất hiện của mỗi alen trong từng locus của dân tộc Nùng, làm cơ sở khoa học để phân tích, đánh giá và rút ra kết luận giám định xác định quan hệ huyết thống hoặc truy nguyên cá thể là người dân tộc Nùng.

## **3. Đối tượng và phạm vi nghiên cứu**

Nghiên cứu tần suất các alen theo hệ Identifiler của 120 cá thể người Nùng không có quan hệ họ hàng, sống rải rác trên các tỉnh trung du và miền núi phía Bắc.

## **4. Nội dung và phương pháp nghiên cứu**

Để thực hiện mục tiêu nghiên cứu đề tài sẽ bao gồm một số nội dung chính với các vấn đề cần giải quyết sau:

- Thu mẫu máu của 120 cá thể người Nùng Theo nguyên tắc thu ngẫu nhiên từ những người không có quan hệ họ hàng huyết thống trực hệ ở các tỉnh khác nhau (*có thông tin cá nhân*)
- Tách chiết ADN bằng phương pháp tách chiết vô cơ, sử dụng chelex 100 (hãng Bio – Rad, Mỹ).