

DOI: 10.59715/pntjmp.2.1.9

## Quy trình giải trình tự Sanger chẩn đoán các biến thể đa hình đơn Nucleotide trên gen CLDN-1 liên quan đến bệnh viêm da cơ địa

Lê Dương Hoàng Huy<sup>1,2</sup>, Nguyễn Đoàn Huỳnh Anh Phúc<sup>1</sup>, Nguyễn Hữu Ngọc Tuấn<sup>1,3</sup>

<sup>1</sup>Bộ môn Hóa Sinh - Sinh Học Phân Tử Y Học, Trường Đại học Y khoa Phạm Ngọc Thạch

<sup>2</sup>Trung Tâm Nghiên Cứu Y Sinh - Đơn vị Dịch Vụ, Trường Đại học Y khoa Phạm Ngọc Thạch

<sup>3</sup>Trung Tâm Nghiên Cứu Y Sinh - Đơn vị Phân Tử, Trường Đại học Y khoa Phạm Ngọc Thạch

### Tóm tắt

**Giới thiệu:** Tổn thương hàng rào bảo vệ da trong viêm da cơ địa (VDCĐ) có liên quan đến yếu tố di truyền, trong đó các biến thể đa hình đơn nucleotide (SNP) trên gen CLDN-1 mã hóa protein claudin-1 được mô tả có liên quan đến bệnh sinh của VDCĐ. Việc xác định chính xác các biến thể này góp phần thực hiện các nghiên cứu nhằm nâng cao hiệu quả quản lý bệnh nhân VDCĐ, trong đó kỹ thuật giải trình tự Sanger được xem là tiêu chuẩn vàng trong xác định các biến thể.

**Mục tiêu:** Xây dựng quy trình giải trình tự Sanger và khảo sát bốn biến thể rs17501010, rs9290927, rs9290929 và rs893051 trên gen CLDN-1 liên quan đến VDCĐ.

**Đối tượng và phương pháp nghiên cứu:** Thiết kế 4 cặp đoạn mồi đặc hiệu cho các biến thể, tối ưu hóa nhiệt độ bắt cặp của phản ứng PCR đa mồi chứa cả 4 biến thể và tối ưu hóa phản ứng giải trình tự Sanger trên hệ thống Applied Biosystems 3500 Series Genetic Analyzer của hãng Thermo Fisher. Áp dụng toàn bộ quy trình giải trình tự Sanger đã tối ưu lên 12 mẫu máu của bệnh nhân VDCĐ nhằm đánh giá thông số kỹ thuật và đặc điểm của các biến thể.

**Kết quả:** Xây dựng thành công quy trình giải trình tự Sanger bao gồm: thiết kế được bốn cặp đoạn mồi khuếch đại đặt hiệu bốn biến thể quan tâm, tối ưu hóa được nồng độ DNA, primer và chu trình nhiệt của phản ứng PCR đa mồi chứa bốn biến thể quan tâm, tối ưu hóa nồng độ DNA đầu vào của phản ứng giải trình tự. Khảo sát được 12/12 mẫu DNA từ bệnh nhân VDCĐ với tất cả kết quả giải trình tự có vị trí nucleotide nằm trong vùng QVB cao.

**Kết luận:** Quy trình có thể đưa vào ứng dụng để xác định biến thể đa hình đơn nucleotide. Đồng thời là các kết quả bước đầu, hứa hẹn ứng dụng kỹ thuật sinh học phân tử để khảo sát các đặc điểm di truyền ở bệnh viêm da cơ địa. Từ đó đưa ra các chiến lược phù hợp để tiếp cận và quản lý bệnh VDCĐ ở Việt Nam.

**Từ khóa:** Giải trình tự Sanger, bệnh viêm da cơ địa, CLDN-1, protein claudin-1.

### Abstract

**Sanger sequencing process of single nucleotide polymorphisms variants in CLDN-1 genes associated with atopic dermatitis**

**Introduction:** Damage to the skin barrier in atopic dermatitis (AD) is related to genetic factors, in which a single nucleotide polymorphism (SNP) variant on the gene CLDN-1 encoding the claudin-1 protein has been described as having related

Ngày nhận bài:

20/11/2022

Ngày phản biện:

20/12/2022

Ngày đăng bài:

20/01/2023

Tác giả liên hệ:

Lê Dương Hoàng Huy

Email: huyldh@pnt.

edu.vn

ĐT: 0783639199

to the pathogenesis of AD. Accurate identification of these variants contributes to the implementation of studies to improve the management of patients with AD, in which Sanger sequencing is considered the gold standard in identifying variants.

**Objective:** Develop Sanger sequencing and investigate variants rs17501010, rs9290927, rs893051 and rs9290929 on related CLDN-1 gene related atopic dermatitis.

**Materials and methods:** Design 4 pairs of primers specific for variants, optimize the pairing temperature of multiplex PCR containing all 4 variants, and optimize Sanger sequencing on Applied Biosystems 3500 Series Genetic Analyzer by Thermo Fisher. Applying the entire optimized Sanger sequencing procedure to 12 blood samples of AD patients to evaluate the specifications and characteristics of the variants.

**Results:** Successfully built Sanger sequencing included: designing four pairs of amplifying primers that signal the four variants of interest, optimizing the concentration of DNA, primer and thermal cycling of the multi - primed PCR reaction containing four variants of interest, optimizing the input DNA concentration of the sequencing reaction. Surveyed 12/12 DNA samples from AD patients with all sequencing results having nucleotide positions in the high QVB region.

**Conclusion:** The procedure is applicable to the identification of single nucleotide polymorphisms. At the same time, these are initial results, promising to apply molecular biology techniques to investigate genetic characteristics in atopic dermatitis. From there, provide appropriate strategies to approach and manage AD in Vietnam.

**Keywords:** Sanger sequencing, atopic dermatitis, CLDN-1, claudin-1 protein.

## 1. ĐẶT VẤN ĐỀ

Viêm da cơ địa (VDCĐ) là một bệnh viêm da mãn tính thường gặp, chiếm 15 - 30% ở trẻ em, 2 - 10% ở người lớn [1]. Cơ chế bệnh sinh của VDCĐ phức tạp, là kết quả của sự tương tác giữa yếu tố di truyền và yếu tố môi trường. Lâm sàng của bệnh đa dạng, thay đổi theo tuổi, tái đi tái lại nhiều lần và điều trị bệnh nhân VDCĐ mức độ trung bình đến nặng còn nhiều khó khăn, thách thức. Trong đó, lớp hàng rào thứ hai bên dưới lớp sừng (lớp hàng rào thứ nhất) là các điểm nối kín (tight junctions) với chức năng khóa kín và kiểm soát con đường vận chuyển giữa hai tế bào thượng bì góp phần quan trọng vào cơ chế bệnh sinh của VDCĐ [2]. Điểm nối kín cấu tạo bởi nhiều protein xuyên màng, giá đỡ và điều hòa khác nhau, trong đó claudin-1 với sự phân bố đa dạng là thành phần được nghiên cứu nhiều nhất [3 - 5]. Gen CLDN1 mã hóa protein claudin-1 có nhiều biến thể di truyền đa hình đơn nucleotid (SNP) cụ thể là rs17501010, rs9290927, rs893051,

rs9290929 đã được nhiều nghiên cứu mô tả mối liên quan với VDCĐ [6 - 9]. Thông tin các SNP này sẽ góp phần vào tối ưu hóa quy trình chẩn đoán, hướng dẫn điều trị và các kế hoạch can thiệp cộng đồng. Từ đó, đặt ra nhu cầu về một công cụ chẩn đoán có độ tin cậy cao nhằm xác định giá trị lâm sàng của các SNPs, cũng như hoạt động chẩn đoán thường quy. Vì vậy, chúng tôi tiến hành xây dựng quy trình chẩn đoán các SNPs trên gen CLDN1 với các mục tiêu sau:

1. Thiết kế các đoạn mồi sử dụng trong quá trình giải trình tự DNA bằng phương pháp Sanger chẩn đoán các biến thể rs17501010, rs9290927, rs893051 và rs9290929.
2. Tối ưu hóa phản ứng PCR đa mồi khuếch đại các đoạn DNA chứa các biến thể nêu trên.
3. Tối ưu lượng DNA đầu vào phản ứng giải trình tự Sanger từ sản phẩm PCR.
4. Đánh giá các thông số kỹ thuật của quy trình vừa xây dựng trên 12 mẫu DNA của người.

## 2. ĐỐI TƯỢNG VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Thiết kế nghiên cứu: nghiên cứu thực nghiệm.

Thời gian và địa điểm nghiên cứu: từ tháng 01/08/2022 đến tháng 15/11/2022, tại Trung tâm Nghiên cứu Y sinh, trường Đại học Y khoa Phạm Ngọc Thạch.

Phương pháp nghiên cứu:

Đối tượng nghiên cứu: Các đặc điểm kỹ thuật của thí nghiệm chẩn đoán các SNP bằng kỹ thuật giải trình tự Sanger, bao gồm: trình tự mỗi khuếch đại đoạn nucleotide có chứa các SNP; chu trình nhiệt phản ứng PCR đa môi khuếch đại đoạn gen cần giải trình tự; lượng DNA đầu vào của phản ứng giải trình tự Sanger, chỉ số chất lượng của Sanger gồm LOR và QVB. Mẫu máu dùng để xây dựng quy trình được lấy từ những người tình nguyện đang làm việc tại Trung tâm Nghiên cứu Y sinh, mẫu dùng để đánh giá quy trình lấy từ bệnh nhân được chẩn đoán VDCĐ theo tiêu chuẩn của Hanifin và Rajka (1980) tại bệnh viện Da liễu TPHCM.

Các bước tiến hành nghiên cứu

1. Lấy máu tĩnh mạch người tình nguyện, ly trích DNA (TopPure Blood DNA Extraction Kit - ABT, Việt Nam) có sử dụng thêm dung dịch ly giải hồng cầu ACK (Ammonium Chloride, Potassium Bicarbonate). Tiêu chí chất lượng cần đạt: nồng độ DNA thu được > 500 ng/ml; độ tinh khiết OD<sub>260/280</sub> từ 1,8 - 2,0.

2. Sử dụng thông tin từ cơ sở dữ liệu của Trung tâm Thông tin Công nghệ Sinh học Quốc gia - NCBI. Trích xuất ra đoạn trình tự gen quan tâm trên bộ gen người (phiên bản GRCh38), sau đó sử dụng nền tảng Primer - BLAST của cơ sở dữ liệu của NCBI để thiết kế hai cặp đoạn môi cho mỗi biến thể rs17501010, rs9290927, rs9290929 và rs893051 trên gen CLDN-1 với các tiêu chí: (1) độ dài môi: 15 - 30 nucleotide, (2) nồng độ GC trong khoảng 40 - 60%, (3) nhiệt độ nóng chảy trong khoảng 57 - 62 độ C, (4) chiều dài đoạn khuếch đại trong khoảng 400 - 1000bp.

3. Xác định tổ hợp các cặp môi từ các cặp môi vừa thiết kế ở mục 2 và chu trình nhiệt bằng phản ứng PCR đa môi (AmpliTaq Gold™ 360

kit, ThermoFisher, Hoa Kỳ). Tổ hợp môi và chu trình nhiệt được xem là tối ưu khi: (1) có 4 sản phẩm PCR có độ dài ứng với độ dài đoạn gen được khuếch đại, (2) sản phẩm PCR trên gel điện di (LabChip GX Touch HT Nucleic Acid Analyzer) có đỉnh rõ, (3) không có sản phẩm phụ hoặc tín hiệu nhiễu (đỉnh phụ).

4. Tinh sạch sản phẩm PCR (ExoSAP-IT PCRproduct Cleanup Reagent, Thermo, Hoa Kỳ).

5. Xác định nồng độ DNA từ PCR đa môi của từng SNP để đưa vào giải trình tự với hệ thống SeqStudio Genetic Analyzer 3500 bằng sinh phẩm BigDye Terminator v3.1 Cycle Sequencing Kit (ThermoFisher, Hoa Kỳ).

6. Tinh sạch sản phẩm bước 5 bằng ethanol và tái huyền phù với dung dịch Hi-Di™ Formamide. Điện di mao quản bằng SeqStudio Genetic Analyzer 3500.

7. Sử dụng phần mềm Sequencing Analysis Software v6.0 và Variant Reporter™ Software v2.0 để đánh giá chỉ số QVB (Quality value of Base - chỉ số ước tính độ chính xác của mỗi nucleotid base được gọi ra) và LOR (Length of Read - khoảng chiều dài dài nhất đọc được không bị gián đoạn của nucleotid). Kết quả được xem là đạt khi: (1) tỉ lệ LOR so với chiều dài thiết kế  $\geq 90\%$ , (2) SNP nằm trong vùng có QVB  $\geq 20$ .

8. Đánh giá lại toàn bộ quy trình với các điều kiện thử nghiệm đã tối ưu từ bước 1 đến bước 7 trên 12 mẫu DNA của bệnh nhân VDCĐ.

Xử lý số liệu: Số liệu được xử lý bằng phần mềm Microsoft Excel, phần mềm Sequencing Analysis Software v6.0 và phần mềm Variant Reporter™ Software v2.0.

Đạo đức trong nghiên cứu: Chứng nhận chấp thuận đạo đức trong nghiên cứu y sinh, số 693/TĐHYKPNT-HĐĐĐ ngày 01/07/2022 của Trường Đại học Y khoa Phạm Ngọc Thạch.

## 3. KẾT QUẢ

### 1. Trình tự các đoạn môi được thiết kế để xác định các SNP

Trình tự các cặp môi thiết kế được trình bày trong bảng 2, thỏa các tiêu chí ban đầu. Sau khi tiến hành các phản ứng tiếp sau, cặp môi nào không đặc hiệu sẽ được thiết kế lại.

**Bảng 1.** Trình tự và đặc điểm các cặp môi đã thiết kế cho 4 SNP quan tâm

STT	Tên biến thể	Tên cặp môi	Trình tự (5' - 3')	Độ dài (nu)	Tm	GC%	Độ dài đoạn gen khuếch đại
1	rs17501010	CLDN1_10F01	CAAAAGTAGGGCACCTCCCAG	20	60.04	60	592
		CLDN1_10R01	AGAATGGTTTCATGGGCTGCT	20	59.97	50	
2		CLDN1_10F02	AAGGCAGAGAGAAGCAGCAG	20	60.04	55	633
		CLDN1_10R02	ACTCACAGATACCCCAACCCT	20	59.58	55	
3		CLDN1_27F01	GACTCAACAAGCCATCCCCA	20	59.96	55	647
		CLDN1_27R01	TGGCCTGTTCCCTGGACAATG	20	60.25	55	
4	rs9290927	CLDN1_27F02	CTGGCCCTGTGATAGGTGTC	20	59.62	60	748
		CLDN1_27R02	TAGCATGATGGGTGAGGTGC	20	59.62	55	
5		CLDN1_29F01	CTTGTCGACTGATGCTTGG	20	59.83	55	604
		CLDN1_29R01	ACAGCACTAGGGTTTGAGCC	20	59.96	55	
6	rs9290929	CLDN1_29F02	AAGAGGCAGCATGGCAGATT	20	60.03	50	467
		CLDN1_29R02	CAGCACTAGGGTTTGAGCCA	20	59.96	55	
7		CLDN1_51F01	GGCCTTCACTCCCTCAACTC	20	60.04	60	490
		CLDN1_51R01	TTGGGCTTCATTCTCGCCTT	20	59.96	55	
8	rs893051	CLDN1_51F02	CCCCAAATCTCGGAATGCCT	20	60.11	55	548
		CLDN1_51R02	GCAGCTGTTGGGCTTCATTC	20	60.11	55	

## 2. Tối ưu hóa phản ứng PCR đa môi khuếch đại các đoạn gen chứa các SNP

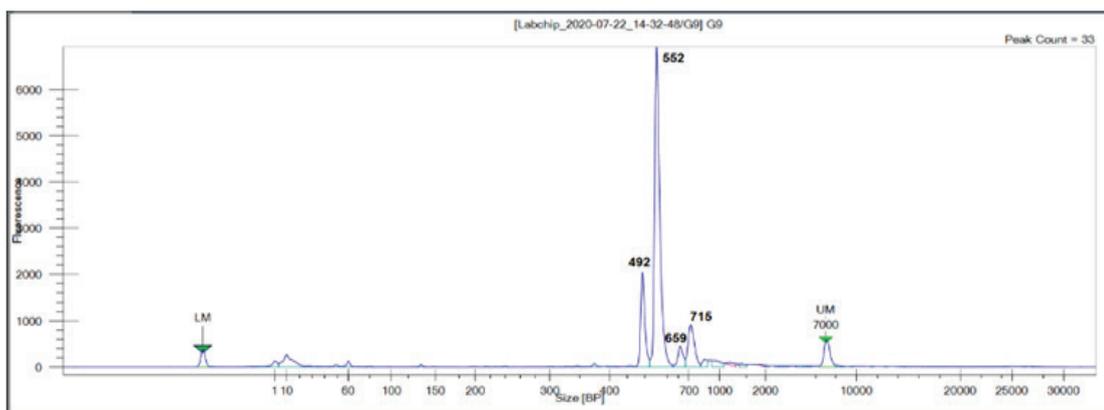
Thực hiện các phản ứng, thay đổi nhiệt độ bắt cặp theo gradient nhiệt độ tăng dần từ nhiệt độ lý thuyết. Từ các dãy gradient nhiệt độ đó chọn ra nhiệt độ tối ưu và điện di các sản phẩm tương ứng bên hình 1.



**Hình 1.** Kết quả điện di gel sản phẩm PCR khuếch đại các cặp đoạn môi đã thiết kế với nhiệt độ tối ưu

Nhận xét: Trong mẻ chạy, với nhiệt độ tối ưu chọn ra được từ dãy gradient, các sản phẩm đều rõ, không có sản phẩm phụ, hiện diện đúng kích thước đã thiết kế. Có thể sử dụng các nhiệt độ này để thực hiện các phản ứng PCR đơn môi.

Thực hiện 16 phản ứng là 16 tổ hợp của 4 cặp đoạn môi, thay đổi các cặp đoạn môi và chu trình nhiệt. Kết quả điện di trên labchip sản phẩm PCR khuếch đại của 1 tổ hợp biến thể được minh họa ở hình 1. Tổng hợp các điều kiện tối ưu hóa PCR của các cặp môi được liệt kê trong bảng 2.



**Hình 2:** Kết quả điện di sản phẩm PCR khuếch đại các biến thể rs17501010, rs9290927, rs9290929 và rs893051 trên gen CLDN-1 bằng các cặp môi CLDN\_10\_01; CLDN\_27\_02; CLDN\_29\_02; CLDN\_51\_02

Nhận xét: Trong mẻ chạy gồm 16 tổ hợp, mẫu có hai điểm đánh dấu marker gồm điểm đánh dấu dưới (LM, Lower Marker) và điểm đánh dấu trên (Upper Marker) rõ ràng. Các đỉnh sản phẩm rõ, có 4 đỉnh rõ tương ứng với chiều dài khuếch đại đã thiết kế, không có đỉnh phụ.

**Bảng 2.** Tổ hợp các cặp môi và chu trình nhiệt trong phản ứng PCR đa môi

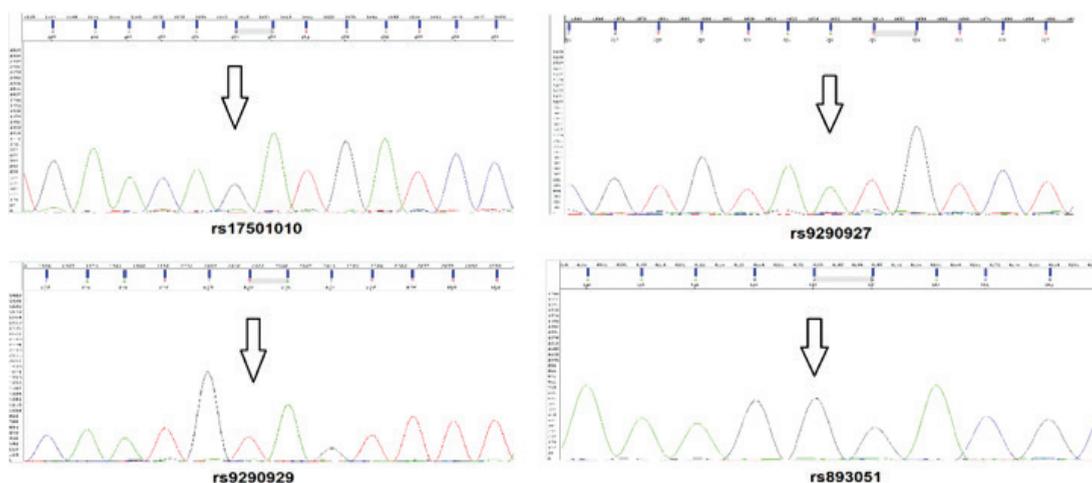
STT	Tên cặp môi	Lượng DNA đầu vào PCR đa môi	Chu trình nhiệt			
			Hoạt hóa thành phần của mix	95	10 phút	01 chu kỳ
1	CLDN1_10_01	50 ng	Biến tính	95	30 giây	35 chu kỳ
2	CLDN1_27_02		Bắt cặp	57	30 giây	
3	CLDN1_29_02		Nối dài	72	60 giây	
4	CLDN1_51_02		Nối dài cuối	72	7 phút	01 chu kỳ

**3. Tối ưu lượng DNA đầu vào phản ứng giải trình tự Sanger từ sản phẩm PCR.**

Thực hiện giải trình tự với các với sản phẩm PCR đa môi, có thay đổi lượng DNA đầu vào với từng SNP để được. Kết quả lượng sản phẩm PCR đầu vào tối ưu cho phản ứng giải trình tự được thể hiện qua bảng 3. Một số hình ảnh giải trình tự được thể hiện ở các hình 2.

**Bảng 3.** Trình tự đoạn môi và lượng DNA tối ưu của từng SNP

STT	Tên SNP	Trình tự đoạn môi sử dụng để giải trình tự	Lượng DNA đầu vào của mỗi phản ứng giải trình tự
1	rs17501010	CAAAGTAGGGCACCTCCCAG	10 - 12 ng
2	rs9290927	TAGCATGATGGGTGAGGTGC	5 - 6 ng
3	rs9290929	CAGCACTAGGGTTTGAGCCA	5 - 6 ng
4	rs893051	GCAGCTGTTGGGCTTCATTC	5 - 6 ng



**Hình 3:** Kết quả giải trình tự Sanger bốn biến thể rs17501010, rs9290927, rs893051 và rs9290929 trên gen CLDN-1 với nồng độ DNA đầu vào đã tối ưu hóa. (Mũi tên chỉ vị trí của SNP, nucleotide tại vị trí này hiện diện rõ, không quan tâm là nucleotide gì).

Nhận xét: Đây là mẫu DNA của người tình nguyện, hình ảnh các đỉnh tín hiệu rõ ràng, không có tín hiệu nhiễu. Tại vị trí của biến thể đa hình (vị trí mũi tên), có đỉnh tín hiệu rõ ràng, không có tín hiệu nhiễu làm sai lệch kết quả. QVB đều  $\geq 20$ .

**4. Đánh giá quy trình giải trình tự trên 12 mẫu DNA bệnh nhân VDCĐ**

Thực hiện giải trình tự với các thông số thí nghiệm tối ưu đã xác định để phân tích 4 biến thể quan tâm trên mẫu DNA của 12 mẫu DNA, dựa trên các tiêu chí đã mô tả. Kết quả các chỉ số trung bình được thể hiện qua bảng 4, các chỉ số của từng mẫu được thể hiện chi tiết qua bảng 5 trong phần phụ lục.

**Bảng 4.** Kết quả đánh giá quy trình giải trình tự Sanger trên 12 mẫu

Tên SNP	LOR so với chiều dài đoạn thiết kế	Trung bình vị trí SNP có QVB $\geq 20$	Trung bình vị trí bắt đầu có QVB $\geq 20$	Trung bình vị trí kết thúc có QVB $\geq 20$	Đánh giá
rs17501010	95,15%	470,92	104,62	559,92	Đạt
rs9290927	96,35%	392,38	102,46	696,00	Đạt
rs9290929	93,79%	172,23	85,92	433,46	Đạt
rs893051	94,71%	353,31	51,15	518,31	Đạt

Nhận xét: Quy trình tối ưu hóa sau khi áp dụng lên các mẫu DNA của bệnh nhân viêm da cơ địa đều cho kết quả giải trình tự đạt chất lượng, tại vị trí các biến thể quan tâm, SNP được xác định rõ.

#### 4. BÀN LUẬN

Trong các kỹ thuật có thể phát hiện các biến đổi đa hình đơn nucleotide, giải trình tự Sanger được xem là tiêu chuẩn vàng. Kỹ thuật này cung cấp một kết quả tin cậy về trình tự của chính đoạn gen hoặc đoạn trình tự đang quan tâm. Đối với những biến thể đa hình đơn nucleotide mới hoặc vừa được mô tả vai trò trên y văn thì việc lựa chọn một kỹ thuật có độ chính xác cao được xem là thiết yếu. Điều này vừa có ý nghĩa về mặt khoa học, vừa được xem là thước đo để phát triển và đánh giá các kỹ thuật khác như qPCR.

Nghiên cứu đã tự thiết kế các cặp mồi cho mỗi SNP quan tâm, có tính đặc hiệu cao thể hiện qua kết quả PCR và giải trình tự. Ưu điểm của việc sử dụng đoạn mồi tự thiết kế là có thể điều chỉnh chiều dài đoạn khuếch đại theo nhu cầu, từ đó có thể tạo ra các phản ứng PCR đa mồi. Việc thiết kế này tạo điều kiện cho nhóm nghiên cứu tối ưu phản ứng PCR đa mồi so với sử dụng đoạn mồi có sẵn từ các nghiên cứu khác, kết quả là nhóm chủ động thiết kế các sản phẩm lệch nhau từ 50 đến 100 nucleotid, điều này sẽ tạo điều kiện cho việc đánh giá các sản phẩm khi thực hiện PCR đa mồi. Mặt khác, các đoạn mồi tự thiết kế sau khi đã kiểm tra chất lượng và xây dựng đúng quy trình chuẩn có thể được sử dụng mà không lo ngại vấn đề bản quyền so với các đoạn mồi đã được công bố trước. Cuối cùng, một quy trình thiết kế mồi chung, với các bước hướng dẫn cụ thể, có thể sử dụng để áp dụng cho các vùng gen hay trình tự quan tâm về sau.

Trong nghiên cứu này, ngoài việc kiểm tra tính đặc hiệu của từng cặp mồi trên gel agarose, với mỗi SNP nhóm thiết kế hai cặp đoạn mồi cho mỗi SNP, với 4 SNP riêng biệt nhóm tạo được

16 tổ hợp các cặp đoạn mồi, để đánh giá kết quả của 16 phản ứng PCR đa mồi một cách chính xác và đặc hiệu hơn, nhóm đã sử dụng hệ thống LabChip GX Touch HT Nucleic Acid Analyzer. Kết quả dựa vào đỉnh của biểu đồ để có thể đánh giá sản phẩm của phản ứng PCR đa mồi.

Trong quá trình giải trình tự Sanger trên hệ thống Applied Biosystems 3500 Series Genetic Analyzer của hãng Thermo Fisher, chất lượng của sản phẩm giải trình tự được đánh giá thông qua thông số QVB hay còn gọi là Phred quality score. Đây là điểm chất lượng, là thước đo chất lượng của việc xác định các nucleobase được tạo ra bằng cách giải trình tự DNA tự động. QVB  $\geq 20$  có nghĩa là khả năng nucleotide này được đọc không chính xác là 1%. Chúng tôi đã sử dụng thêm chỉ số QVB (Quality value of Base - chỉ số ước tính độ chính xác của mỗi nucleotid base được gọi ra) và LOR (Length of Read - khoảng chiều dài dài nhất đọc được không bị gián đoạn của nucleotid) làm thước đo cho chất lượng chuỗi tổng thể. Việc kết hợp nhiều tiêu chuẩn để đánh giá chất lượng kết quả giải trình tự được xem là thiết yếu. Trong nghiên cứu này ở tất cả các mẫu, chỉ số QVB  $\geq 20$  đều kéo dài và chứa cả SNP quan tâm. Chiều dài LOR luôn được kéo dài hơn 90%.

Áp dụng toàn bộ quy trình sau khi chuẩn hóa lên 12 mẫu DNA của bệnh nhân viêm da cơ địa, kết quả về các thông số kỹ thuật đều đạt các tiêu chí đã đặt ra và không có sự khác biệt với các tiêu chuẩn mà nhà sản xuất khuyến cáo. Cụ thể, đối với chiều dài đoạn khuếch đại, số lượng nucleotide có QVB  $\geq 20$  luôn bao phủ được vùng chứa các SNP cần được gọi, phần nucleotide có chất lượng QVB thấp thường tập

trung ở đoạn đầu và đoạn cuối của phần giải trình tự, không có tình trạng nằm rải rác, không gây mất tính liên tục của chiều dài đoạn trình tự, minh chứng của điều này là chỉ số LOR của tất cả các mẫu đều > 90%. Điều này phù hợp với nguyên tắc giải trình tự Sanger và khuyến cáo nhà sản xuất [10, 11]. Về cường độ tín hiệu thì các nucleotide có chất lượng QVB cao đều có cường độ tín hiệu khá tương đồng nhau, dao động trong khoảng 600 - 1300. Kết hợp với việc tinh sạch mẫu tốt để khử các tín hiệu nhiễu, từ đó làm tăng chất lượng của mỗi nucleotide được xác định. Tín hiệu tại ngay vị trí SNP quan tâm luôn rõ ràng, nếu là đồng hợp thì có 1 loại tín hiệu duy nhất, nếu là dị hợp thì hai đỉnh có độ cao tương đồng, đỉnh đồng vị trí.

Hiện chưa thấy thông tin nào về các biến thể này tại Việt Nam cũng như trong khu vực, Trong tương lai, việc ứng dụng kỹ thuật này sẽ hỗ trợ thu thập các thông tin di truyền dạng SNP trên bệnh nhân VDCĐ nói riêng và các bệnh lý khác.

## 5. KẾT LUẬN

Đã xây dựng và tối ưu quy trình giải trình tự Sanger để xác định biến thể đa hình đơn nucleotide: rs17501010, rs9290927, rs893051 và rs9290929 trên gen CLDN-1. Bước đầu áp dụng quy trình trên người tình nguyện, làm cơ sở cho các khảo sát với cỡ mẫu đại diện giúp tiếp cận và quản lý ở phương diện phân tử bệnh nhân VDCĐ ở Việt Nam.

## TÀI LIỆU THAM KHẢO

1. McAleer, M., G. O'Regan, and A. Irvine, Atopic dermatitis. *Dermatology*. Ed. Bologna JL, Schaffer JV, Cerroni L. 2018, Elsevier.
2. Kubo, A., K. Nagao, and M.J.T.J.o.c.i. Amagai, Epidermal barrier dysfunction and cutaneous sensitization in atopic diseases. 2012. 122(2): p. 440-447.
3. Bergmann, S., et al., Claudin-1 decrease impacts epidermal barrier function in atopic dermatitis lesions dose-dependently. 2020. 10(1): p. 1-12.
4. Furuse, M., et al., Claudin-based tight junctions are crucial for the mammalian epidermal barrier a lesson from claudin-1-deficient mice. 2002. 156(6): p. 1099-1111.
5. Tokumasu, R., et al., Dose-dependent role of claudin-1 in vivo in orchestrating features of atopic dermatitis. 2016. 113(28): p. E4061-E4068.
6. Yu, H.-S., et al., Claudin-1 polymorphism modifies the effect of mold exposure on the development of atopic dermatitis and production of IgE. 2015. 135(3): p. 827-830. e5.
7. Asad, S., et al., The tight junction gene Claudin-1 is associated with atopic dermatitis among Ethiopians. 2016. 30(11): p. 1939-1941.
8. De Benedetto, A., et al., Tight junction defects in patients with atopic dermatitis. 2011. 127(3): p. 773-786. e7.
9. De Benedetto, A., et al., Reductions in claudin-1 may enhance susceptibility to herpes simplex virus 1 infections in atopic dermatitis. 2011. 128(1): p. 242-246. e5.
10. Phred - Quality Base Calling. Retrieved 2011. 02: p. 24.
11. Ledergerber, C. and C. Dessimoz, Base-calling for next-generation sequencing platforms. *Briefings in bioinformatics*, 2011. 12(5): p. 489-497.

**PHỤ LỤC**

**Bảng 5:** Kết quả chi tiết đánh giá quy trình giải trình tự Sanger trên 12 mẫu

Tên SNP	Số TT mẫu	LOR so với chiều dài đoạn thiết kế	Vị trí SNP có QVB $\geq 20$	Vị trí bắt đầu có QVB $\geq 20$	Vị trí kết thúc có QVB $\geq 20$	Đánh giá
rs17501010	ADE01	95,27%	477	103	560	Đạt
	ADE02	96,28%	458	115	546	Đạt
	ADE03	92,91%	456	122	546	Đạt
	ADE04	92,57%	474	102	565	Đạt
	ADE05	95,61%	472	102	562	Đạt
	ADE06	95,27%	473	99	561	Đạt
	ADE07	95,44%	475	100	564	Đạt
	ADE08	95,61%	475	101	564	Đạt
	ADE09	95,78%	472	99	561	Đạt
	ADE10	95,27%	475	102	561	Đạt
	ADE11	96,62%	473	99	564	Đạt
	ADE12	95,44%	470	96	562	Đạt
rs9290927	ADE01	95,99%	392	83	718	Đạt
	ADE02	96,12%	391	106	583	Đạt
	ADE03	96,52%	394	124	656	Đạt
	ADE04	95,59%	387	96	715	Đạt
	ADE05	96,26%	392	107	708	Đạt
	ADE06	96,39%	393	101	721	Đạt
	ADE07	96,39%	393	100	710	Đạt
	ADE08	96,66%	395	107	656	Đạt
	ADE09	96,52%	394	103	710	Đạt
	ADE10	96,79%	395	102	724	Đạt
	ADE11	96,66%	395	102	723	Đạt
			96,39%	392	104	708

Tên SNP	Số TT mẫu	LOR so với chiều dài đoạn thiết kế	Vị trí SNP có QVB $\geq 20$	Vị trí bắt đầu có QVB $\geq 20$	Vị trí kết thúc có QVB $\geq 20$	Đánh giá
rs9290929	ADE01	93,36%	174	83	435	Đạt
	ADE02	93,79%	174	84	435	Đạt
	ADE03	94,86%	178	90	439	Đạt
	ADE04	93,58%	172	83	436	Đạt
	ADE05	94,22%	175	86	436	Đạt
	ADE06	94,00%	175	87	436	Đạt
	ADE07	94,22%	175	86	436	Đạt
	ADE08	91,86%	145	56	406	Đạt
	ADE09	92,93%	170	81	431	Đạt
	ADE10	94,22%	176	85	437	Đạt
	ADE11	93,79%	174	85	435	Đạt
	ADE12	94,22%	175	124	436	Đạt
rs893051	ADE01	94,34%	347	16	517	Đạt
	ADE02	94,53%	407	49	518	Đạt
	ADE03	95,62%	354	34	524	Đạt
	ADE04	94,71%	349	25	519	Đạt
	ADE05	94,53%	348	75	518	Đạt
	ADE06	94,71%	349	81	519	Đạt
	ADE07	94,34%	347	27	517	Đạt
	ADE08	95,07%	351	82	521	Đạt
	ADE09	94,53%	348	75	518	Đạt
	ADE10	94,34%	347	39	517	Đạt
	ADE11	94,89%	350	51	520	Đạt
	ADE12	94,71%	349	80	510	Đạt