



(12) **BẢN MÔ TẢ GIẢI PHÁP HỮU ÍCH THUỘC BẰNG ĐỘC QUYỀN  
GIẢI PHÁP HỮU ÍCH**

(19) **Cộng hòa xã hội chủ nghĩa Việt Nam (VN)  
CỤC SỞ HỮU TRÍ TUỆ**

(11)



**2-0003787**

(51) **C12Q 1/6869**  
2024.01

(13) **Y**

---

(21) 2-2024-00479

(22) 08/04/2022

(67) 1-2022-02210

(45) 25/11/2024 440

(43) 27/06/2022 411

(73) Trường Đại học Dược Hà Nội (VN)

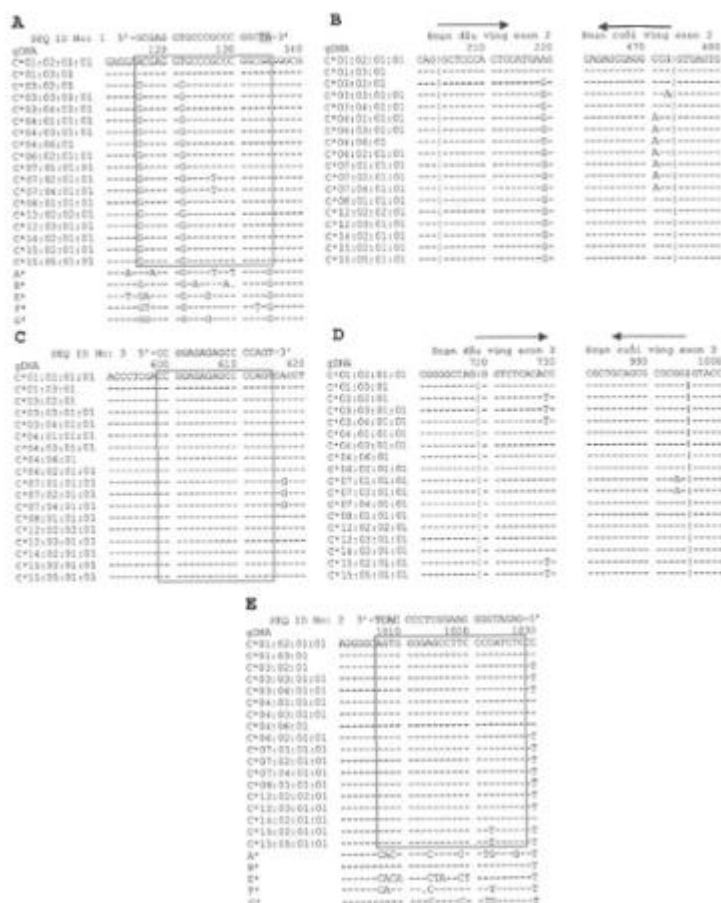
13-15 Lê Thánh Tông, phường Phan Chu Trinh, quận Hoàn Kiếm, thành phố Hà Nội

(72) Phùng Thanh Hương (VN); Phạm Trần Thu Hà (VN); Trần Quang Bình (VN).

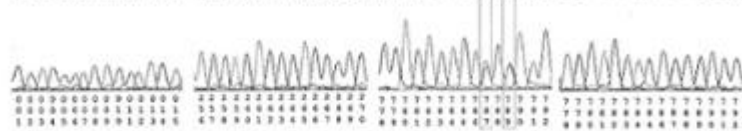
---

(54) **QUY TRÌNH GIẢI TRÌNH TỰ VÙNG ĐA HÌNH EXON 2 - EXON 3 CỦA GEN  
HLA-C**

(57) Sáng chế đề cập đến quy trình giải trình tự vùng đa hình exon 2 - exon 3 của gen *HLA-C* bao gồm bước (i) khuếch đại vùng exon 2- exon 3 của gen *HLA-C* bằng cặp mồi đặc hiệu; và (ii) giải trình tự sử dụng 2 mồi xuôi. Quy trình theo sáng chế dễ thực hiện, đảm bảo độ chính xác, giúp giảm thiểu chi phí, thời gian xử lý kết quả, tạo điều kiện thuận lợi để áp dụng trong các nghiên cứu về di truyền.



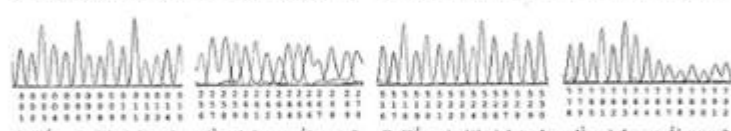
00TCCACTCCATGA ACCAGAGCGAGGCCG 00TCTCACATCTTCC C0CTGCA0C0C0C0G



A. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.1

B. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.1

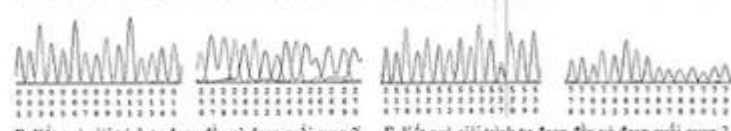
00TCCACTCCATGA AC CAGAGC0AG0E C0 00TCTCACATCTTCC C0CTGCA0C0C0C0G



C. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.2

D. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.2

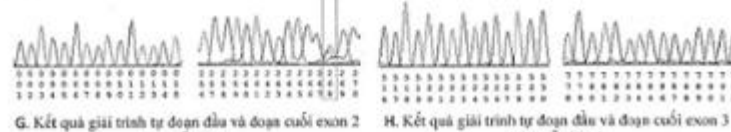
00TCCACTCCATGA ACCAGAGC0AG0C C0 00TCTCACATCTTCC C0CTGCA0C0C0C0G



E. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.3

F. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.3

00TCCACTCCATGA AC CAGAGC0AG0C0C0 00TCTCACATCTTCC C0CTGCA0C0C0C0G



G. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.4

H. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.4

### Lĩnh vực kỹ thuật được đề cập

Sáng chế thuộc lĩnh vực sinh học phân tử và xét nghiệm gen, cụ thể là đề cập đến một quy trình giải trình tự vùng đa hình exon 2 - exon 3 của gen *HLA-C* để thực hiện, đảm bảo độ chính xác và một môi được thiết kế mới để sử dụng cho quy trình này.

### Tình trạng kỹ thuật của sáng chế

*HLA* là siêu họ gen mã hóa phức hợp hòa hợp tổ chức chính (Major Histocompatibility Complex – MHC) hay còn gọi là kháng nguyên bạch cầu người (*Human Leukocyte Antigen - HLA*). Siêu họ gen *HLA* nằm trên nhánh ngắn của nhiễm sắc thể số 6 (6p21.1 – 6p21.3), có chiều dài khoảng 4 Mbp. Mặc dù chỉ chiếm khoảng 0,1% toàn bộ chiều dài bộ gen người, nhưng siêu họ gen này chứa tới 0,6% tổng số gen người đã được xác định. Có thể nói, *HLA* là hệ thống gen có mật độ gen dày đặc nhất trong bộ gen người [9], [11]. Siêu họ gen *HLA* được chia làm 3 lớp, lớp I, lớp II và lớp III. Mỗi lớp lại được phân thành nhiều typ khác nhau. *HLA* lớp I với 3 typ chính *HLA-A*, *HLA-B* và *HLA-C* chứa các gen mã hóa nhiều protein tham gia vào hoạt động miễn dịch của cơ thể [9]. Siêu họ gen *HLA* không chỉ là hệ thống gen có mật độ gen dày đặc nhất trong bộ gen người và còn là hệ thống gen đa hình nhất trong bộ gen người. Tính đến năm 2020, cơ sở dữ liệu HLA IPD-IMGT/HLA đã ghi nhận mỗi typ *HLA* có thể có vài nghìn alen *HLA* đa hình khác nhau [26] và số lượng các alen *HLA* mới được phát hiện mỗi năm vẫn liên tục tăng lên cho thấy sự đa hình rất phức tạp của siêu họ gen này [30].

Gen *HLA-C* là gen *HLA* lớp I, typ C, nằm trên nhiễm sắc thể số 6, tại vị trí 6p22.33, có chiều dài > 4000 bp. Cho tới nay, đã có 5140 alen đa hình của gen *HLA-C* được phát hiện [26]. Các phân tử *HLA-C* được biểu hiện bởi gen *HLA-C* có nhiều vai trò khác nhau trong nhau trong hoạt động của hệ miễn dịch [29]. Phân tử *HLA-C* là một yếu tố quyết định chính đối với hoạt động của tế bào diệt tự

nhân (natural killer cells – NK cells) [20], [22]. Năm 1983, nghiên cứu đầu tiên của McIntyre JA và cộng sự đã chứng minh tầm quan trọng của phân tử HLA-C đối với sự phát triển của thai nhi và nhau thai [21]. Ngoài ra, phân tử HLA-C còn được phát hiện có liên quan đến nhiều bệnh, như nhiễm virus, ung thư, các bệnh tự miễn [27], [28], và có liên quan tới sự gia tăng nguy cơ thải ghép [10], [17], [18]. Bên cạnh đó, phân tử HLA-C cũng liên quan tới nguy cơ gặp phản ứng có hại khi sử dụng các thuốc điều trị Basedow, vảy nến và tăng acid uric máu [19], [25] và liên quan tới mức độ đáp ứng điều trị của một số thuốc khác [13].

Do vai trò quan trọng của phân tử HLA trong hoạt động của hệ miễn dịch nên các gen *HLA* ngày càng được quan tâm nghiên cứu, với nhiều công bố và sáng chế về quy trình xác định kiểu gen *HLA* lớp I được ra đời. Các công bố và sáng chế này đề xuất các quy trình giải trình tự toàn bộ chiều dài gen *HLA-C* (> 4000 bp) và 2 gen *HLA* lớp I khác là *HLA-A* và *HLA-B* nhằm nghiên cứu thông tin di truyền của gen *HLA* lớp I [12], [29], hoặc ứng dụng trong phẫu thuật ghép tạng [1], [4], [5], [6]... Hạn chế của những công bố và sáng chế này là thực hiện phức tạp, phải sử dụng từ hàng chục cặp môi đến xây dựng thư viện hàng trăm cặp môi và thực hiện hàng chục phản ứng PCR giải trình tự 2 chiều, rất tốn kém về thời gian và chi phí cũng như quá trình xử lý kết quả đòi hỏi nguồn nhân lực có trình độ cao, đặc biệt với các sáng chế sử dụng công nghệ giải trình tự gen thế hệ mới [4], [5].

Vùng đa hình exon 2 – exon 3 của gen *HLA* nói chung và gen *HLA-C* nói riêng, tập trung hầu hết các đa hình đơn nucleotit đặc trưng của từng alen *HLA-C*. Trình tự vùng đa hình này quyết định trình tự acid amin của protein HLA-C được biểu hiện. Vì vậy, việc giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C* có ý nghĩa quan trọng trên lâm sàng, và các nghiên cứu gen dược học, gen bệnh học. Tuy nhiên, những công bố và sáng chế dễ thực hiện nhất để giải trình tự vùng exon 2 – exon 3 của gen *HLA-C* trước đây cũng phải sử dụng tối thiểu từ 4 – 6 môi và thực hiện từ 5 – 7 phản ứng, trong đó bao gồm 4 phản ứng giải trình tự cho 2 chiều của chuỗi ADN [2], [23].

Về mặt lý thuyết, các hệ thống máy giải trình tự Sanger tự động hay cả hệ thống giải trình tự gen thế hệ mới có thể giải trình tự những đoạn ADN có kích thước  $> 950$  bp, tuy nhiên kết quả giải trình tự trên thực tế thường chỉ đảm bảo tín hiệu ổn định nhất trong khoảng  $\sim 400 - 500$  bp, và khoảng 50 bp đầu tiên kể từ nucleotit cuối cùng của mỗi thì tín hiệu giải trình tự có thể chưa ổn định. Do đó, hạn chế chính của những quy trình giải trình tự đã công bố là việc thiết kế các cặp mồi có vị trí gắn trên khuôn ADN quá gần hoặc quá xa vùng exon cần giải trình tự dẫn tới tín hiệu giải trình tự đoạn đầu và đoạn cuối exon bị nhiễu, và cách khắc phục là thực hiện giải trình tự 2 chiều, do đó, làm tăng số lượng mồi và số lượng phản ứng cần thực hiện. Ví dụ như quy trình của Peterson [23], vị trí mồi xuôi giải trình tự vùng exon 3 chỉ cách đoạn trình tự đầu của exon này khoảng 30 nucleotit; trong khi đó vị trí mồi ngược giải trình tự vùng exon 3 cách đoạn trình tự cuối của exon này 40 nucleotit. Mặt khác, khoảng cách giữa mồi xuôi giải trình tự vùng exon 2 và mồi xuôi giải trình tự vùng exon 3 lại quá xa, lên tới 573 bp, gây khó khăn khi xử lý kết quả giải trình tự của tổng hợp của các mồi. Một quy trình khác do Delfilo [14] xây dựng cũng có hạn chế như đã đề cập ở trên khi có mồi xuôi giải trình tự vùng exon 2 chỉ cách nucleotit đầu tiên của exon này 20 nucleotit và mồi ngược giải trình tự vùng exon 2 cũng chỉ cách nucleotit cuối cùng của exon này 40 nucleotit.

Vì vậy, việc lựa chọn và thiết kế mồi có vị trí gắn trên khuôn ADN có vai trò quan trọng trong việc đảm bảo độ ổn định của tín hiệu giải trình tự ở cả đoạn đầu và đoạn cuối vùng cần giải trình tự, nhờ vậy, giúp giảm bớt số mồi cần sử dụng và số phản ứng giải trình tự cần thực hiện. Hơn nữa, với việc lựa chọn được mồi có vị trí gắn trên khuôn ADN phù hợp, dẫn tới không cần phải giải trình tự 2 chiều, giúp giảm bớt thời gian xử lý kết quả giải trình tự, đặc biệt với các mẫu dị hợp tử của gen *HLA*.

Tại Việt Nam mới có một công bố về quy trình giải trình tự vùng exon 2 – exon 3 của gen *HLA-A* [8] nhưng chưa có công bố hay sáng chế nào về quy trình giải trình tự gen *HLA-C*.

Do đó, vẫn cần một quy trình giải trình tự vùng đa hình exon 2 - exon 3 của gen *HLA-C* với số mẫu giảm thiểu và ít phản ứng hơn để dễ dàng tiến hành, giúp giảm thiểu thời gian xử lý dữ liệu nhưng vẫn đảm bảo độ chính xác, nhờ đó giảm thiểu chi phí, từ đó tạo điều kiện thuận lợi để áp dụng trong lâm sàng và các nghiên cứu gen dược học, gen bệnh học.

### **Bản chất kỹ thuật của sáng chế**

Sáng chế nhằm giải quyết vấn đề nêu trên, cụ thể là sáng chế đề xuất một quy trình giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C* có thể áp dụng trong lâm sàng cũng như các nghiên cứu gen dược học, gen bệnh học với độ chính xác cao, giảm thiểu số mẫu và số phản ứng phải thực hiện, nhờ đó giảm được chi phí xét nghiệm và thời gian xử lý kết quả.

Sáng chế đề cập đến quy trình giải trình tự gồm 2 bước:

- (i) thực hiện phản ứng PCR khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* với cặp mẫu đặc hiệu gồm mỗi xuôi có trình tự nêu trong SEQ ID No: 1 và mỗi ngược có trình tự nêu trong SEQ ID No: 2 theo chiều 5' đến 3'. Sản phẩm PCR thu được tinh sạch bằng bộ kit thương mại và được sử dụng làm khuôn cho các phản ứng giải trình tự ở bước 2.
- (ii) thực hiện giải trình tự với 2 mẫu xuôi có trình tự nêu trong SEQ ID No: 1 và SEQ ID No: 3 bằng hệ thống giải trình tự tự động.

Bộ mẫu theo sáng chế bao gồm 03 mẫu được sử dụng trong quy trình giải trình tự gồm 2 bước để xác định trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C*, trong đó, cặp mẫu SEQ ID No: 1 và SEQ ID No: 2 được sử dụng ở bước 1; hai mẫu xuôi SEQ ID No: 1 và SEQ ID No: 3 được sử dụng ở bước 2.

Bước 1 của quy trình theo sáng chế sử dụng cặp mẫu để khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* bao gồm mỗi xuôi có trình tự nêu trong SEQ ID No: 1 và mỗi ngược có trình tự nêu trong SEQ ID No: 2, trong đó, mỗi xuôi SEQ ID No: 1 có vị trí gắn mẫu trên khuôn ADN từ nucleotit 116 – 135 và mỗi ngược SEQ ID No: 2 có vị trí gắn mẫu trên khuôn ADN từ nucleotit 1007 – 1028. Mỗi xuôi SEQ ID No: 1 có một nucleotit (t) không bắt cặp bổ sung với

khuôn ADN của gen *HLA-C* và các gen *HLA* typ I khác tại vị trí nucleotit thứ 2 kể từ nucleotit tận cùng đầu 3'; và mỗi ngược SEQ ID No: 2 có các nucleotit ở đầu 3' (từ nucleotit 1007 – 1010) bắt cặp bổ sung với khuôn ADN của gen *HLA-C* mà không bắt cặp bổ sung với khuôn ADN của các gen *HLA* lớp I khác, đảm bảo độ đặc hiệu của cặp mỗi với gen *HLA-C*.

Bước 2 của quy trình theo sáng chế sử dụng mỗi xuôi SEQ ID No: 1 có vị trí gắn mỗi tại vùng intron 1 của gen *HLA-C* cho một phản ứng giải trình tự. Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến nucleotit đầu tiên của vùng exon 2 của gen *HLA-C* là 88 bp (từ nucleotit 116 – 204). Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến nucleotit cuối cùng của vùng exon 2 của gen *HLA-C* là 357 bp (từ nucleotit 116 – 473). Mỗi xuôi SEQ ID No: 3 có vị trí gắn mỗi tại vùng intron 2 của gen *HLA-C*, từ nucleotit 599 – 615, được sử dụng cho phản ứng giải trình tự còn lại ở bước 2. Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 3 đến nucleotit đầu tiên của vùng exon 3 của gen *HLA-C* là 121 bp (từ nucleotit 599 – 720). Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 3 đến nucleotit cuối cùng của vùng exon 3 của gen *HLA-C* là 396 bp (từ nucleotit 599 – 995).

Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến mỗi xuôi SEQ ID No: 3 là 483 bp (từ nucleotit 116 – 599).

### **Mô tả vắn tắt các hình vẽ**

Hình 1 là sơ đồ quy trình giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C* theo sáng chế.

Hình 2 thể hiện vị trí gắn các mỗi trên khuôn ADN của gen *HLA-C* trong quy trình theo sáng chế.

Hình 3 là kết quả giải trình tự vùng đa hình exon 2 – exon 3 gen *HLA-C* của 04 mẫu ADN sử dụng quy trình giải trình tự theo sáng chế.

### Mô tả chi tiết sáng chế

Sáng chế đề cập đến quy trình giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C* bao gồm 2 bước và sử dụng 03 môi, trong đó gen *HLA-C* là một trong những gen có tính đa hình cao nhất trong bộ gen người, với 5140 alen *HLA-C* và 1540 vị trí biến thể đã được phát hiện theo cơ sở dữ liệu của gen *HLA* năm 2020 [26]. Tại Việt Nam, đã có 18 alen *HLA-C* được phát hiện với 3 alen có tần suất cao nhất lần lượt là *HLA-C\*01:02* (16,5%), *HLA-C\*08:01* (15,6%), *HLA-C\*07:02* (14,7%) [16]. Các môi được sử dụng trong quy trình có trình tự nêu trong SEQ ID No: 1 - 3. Trình tự nucleotit của alen tham chiếu *HLA-C\*01:02:01:01* được sử dụng để thiết kế các môi, được nêu trong SEQ ID No: 4 [30]. Quy trình theo sáng chế gồm 2 bước: (i) thực hiện phản ứng PCR khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* với cặp môi SEQ ID No: 1/SEQ ID No: 2. Sản phẩm PCR thu được được tinh sạch để làm khuôn cho hai phản ứng giải trình tự ở bước (ii). Bước (ii) thực hiện 2 phản ứng giải trình tự với 2 môi xuôi SEQ ID No: 1 và SEQ ID No: 3.

Trình tự các gen *HLA-C* có những đặc điểm được mô tả trong Hình 2. **(A)**. Vị trí gắn của môi xuôi SEQ ID No: 1. **(B)**. Đoạn trình tự đầu và đoạn trình tự cuối vùng exon 2. **(C)**. Vị trí gắn của môi xuôi SEQ ID No: 3. **(D)**. Đoạn trình tự đầu và đoạn trình tự cuối vùng exon 3. **(E)**. Vị trí gắn của môi ngược SEQ ID No: 2. Các nucleotit đặc hiệu của môi với gen *HLA-C* được in đậm. Khung vuông thể hiện vị trí gắn của môi trên ADN.

Gen *HLA-C* nói chung có sự tương đồng với các gen *HLA* lớp I khác là *HLA-A*, *HLA-B*, *HLA-E*, *HLA-F*, *HLA-G*. Gen *HLA-C* phân biệt với các gen *HLA* lớp I khác tại nucleotit **a** ở vị trí 135 (nằm ở intron 1) và 2 – 4 nucleotit **ca-ca** ở vị trí 1007 – 1010 (nằm ở intron 3).

Vùng intron 1 trải dài từ nucleotit 74 – 203, vùng exon 2 trải dài từ nucleotit 204 – 473, vùng intron 3 trải dài từ nucleotit 474 – 719, vùng exon 3 trải dài từ nucleotit 720 – 995.

Vùng exon 2 – exon 3 là vùng tập trung phần lớn các đa hình đơn nucleotit đặc trưng của gen *HLA-C*. Trình tự ADN exon 2 và exon 3 của gen *HLA-C* quyết định tính đa hình của các protein *HLA-C* được biểu hiện. Các protein *HLA-C* có nhiều vai trò khác nhau trong hoạt động của hệ miễn dịch [29]. Vì vậy, xác định trình tự ADN exon 2 và exon 3 của gen *HLA-C* có ý nghĩa quan trọng trong các nghiên cứu về mối liên quan giữa các alen đa hình của gen này với nguy cơ phản ứng có hại của thuốc [19], [25], khả năng đáp ứng điều trị với thuốc [13], và các bệnh tự miễn phổ biến [27], [28].

Như đã trình bày ở trên, điểm khác biệt của quy trình được đề xuất trong sáng chế này là chỉ sử dụng 03 môi với 03 phản ứng, ít hơn nhiều so với các công bố về quy trình giải trình tự gen *HLA-C* (sử dụng 4 – 6 môi và cần thực hiện 5 – 7 phản ứng [14], [15], [23], [29]) và những sáng chế về quy trình giải trình tự gen *HLA-C* khác (sử dụng từ hàng chục môi và hàng chục phản ứng [2], [3] đến thư viện hàng trăm môi để giải trình tự toàn bộ gen *HLA-C* bằng công nghệ giải trình tự gen thế hệ mới [1], [4], [5], [6], [7]). Sáng chế đề xuất một quy trình gồm 2 bước với sơ đồ tổng quát thể hiện trong Hình 1:

- (i) thực hiện phản ứng PCR khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* với cặp môi đặc hiệu gồm môi xuôi có trình tự SEQ ID No: 1 và môi ngược có trình tự SEQ ID No: 2 theo chiều 5' đến 3'. Sản phẩm PCR bước 1 có kích thước 912 bp được tinh sạch bằng bộ kit thương mại và sử dụng làm khuôn cho các phản ứng giải trình tự ở bước 2.
- (ii) thực hiện giải trình tự với 2 môi xuôi có trình tự SEQ ID No: 1 và SEQ ID No: 3 bằng hệ thống giải trình tự tự động.

Một điểm khác biệt nữa của quy trình được đề xuất trong sáng chế này so với các sáng chế khác là sử dụng cặp môi có trình tự nêu trong SEQ ID No: 1 và SEQ ID No: 2 có trình tự như hai môi HLACB1F và HLACB1R theo công bố của Phạm và cộng sự (công bố về quy trình PCR phát hiện đặc hiệu alen *HLA-C\*03:02*, không phải công bố về quy trình giải trình tự gen *HLA-C*) [24] để khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C*. Lựa chọn cặp môi SEQ ID No: 1/ SEQ

ID No: 2 để thực hiện phản ứng PCR khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* bước 1 do những đặc điểm sau:

- Mỗi xuôi SEQ ID No: 1 có nucleotit **a** tận cùng đầu 3' không bắt cặp bổ sung với khuôn ADN của các gen *HLA* lớp I khác, đồng thời được bổ sung nucleotit (**t**) không bắt cặp bổ sung với khuôn ADN của các gen *HLA* tại vị trí nucleotit thứ 2 kể từ nucleotit tận cùng đầu 3'.
- Bên cạnh đó, mỗi ngược SEQ ID No: 2 có các nucleotit ở đầu 3' (từ nucleotit 1007 – 1010) bắt cặp bổ sung với khuôn ADN của gen *HLA-C* mà không bắt cặp bổ sung với khuôn ADN của các gen *HLA* lớp I khác.

Những đặc điểm trên của mỗi xuôi SEQ ID No: 1 và mỗi ngược SEQ ID No: 2 đảm bảo độ đặc hiệu với gen *HLA-C*, đồng thời tránh tạo ra các sản phẩm PCR không đặc hiệu là vùng exon 2 – exon 3 của các gen *HLA-A*, *HLA-B*, *HLA-E*, *HLA-F*, *HLA-G* (Hình 2). Trình tự của cặp mỗi SEQ ID No: 1/ SEQ ID No: 2 được trình bày ở Bảng 1 (các nucleotit của mỗi xuôi SEQ ID No: 1 và mỗi ngược SEQ ID No: 2 đặc hiệu với gen *HLA-C* được in đậm).

**Bảng 1. Trình tự của cặp mỗi sử dụng để khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* ở bước 1**

| Mỗi          | Trình tự (chiều 5'-3')                 |
|--------------|----------------------------------------|
| SEQ ID No: 1 | 5'-gcgaggtgcc <b>cgcccggcta</b> -3'    |
| SEQ ID No: 2 | 5'-gagatgggga <b>aggctcccca ct</b> -3' |

Một điểm nổi bật khác của quy trình được đề xuất trong sáng chế này là bước 2 chỉ thực hiện 2 phản ứng giải trình tự với 2 mỗi xuôi, thay vì phải thực hiện giải trình tự cả 2 chiều với  $\geq 4$  mỗi và  $\geq 4$  phản ứng [1], [2], [3], [4], [5], [6], [7], [15], [29]. Trình tự của 2 mỗi xuôi SEQ ID No: 1 và SEQ ID No: 3 được trình bày ở Bảng 2.

**Bảng 2. Trình tự 2 mỗi xuôi sử dụng để giải trình tự ở bước 2**

| Mỗi          | Trình tự (chiều 5'-3')              |
|--------------|-------------------------------------|
| SEQ ID No: 1 | 5'-gcgaggtgcc <b>cgcccggcta</b> -3' |

|              |                          |
|--------------|--------------------------|
| SEQ ID No: 3 | 5'-ccggagagag cccaggt-3' |
|--------------|--------------------------|

Mỗi xuôi SEQ ID No: 1 sử dụng cho một phản ứng giải trình tự ở bước 2 có vị trí gắn môi tại vùng intron 1 của gen *HLA-C*. Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến nucleotit đầu tiên của vùng exon 2 của gen *HLA-C* là 88 bp (từ nucleotit 116 – 204). Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến nucleotit cuối cùng của vùng exon 2 của gen *HLA-C* là 357 bp (từ nucleotit 116 – 473) (Hình 2).

Mỗi xuôi SEQ ID No: 3, là môi được thiết kế mới, sử dụng cho một phản ứng giải trình tự ở bước 2 có vị trí gắn môi tại vùng intron 2 của gen *HLA-C*, từ nucleotit 599 – 615. Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 3 đến nucleotit đầu tiên của vùng exon 3 của gen *HLA-C* là 121 bp (từ nucleotit 599 – 720). Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 3 đến nucleotit cuối cùng của vùng exon 3 của gen *HLA-C* là 396 bp (từ nucleotit 599 – 995) (Hình 2).

Chiều dài đoạn giải trình tự từ mỗi xuôi SEQ ID No: 1 đến mỗi xuôi SEQ ID No: 3 là 483 bp (từ nucleotit 116 – 599) (Hình 2).

Như vậy, chiều dài tất cả các đoạn giải trình tự của 2 môi xuôi SEQ ID No: 1 và SEQ ID No: 3 đều nhỏ hơn 500 bp đảm bảo sự ổn định của tín hiệu giải trình tự cả đoạn cuối của exon 2 và exon 3 cũng như thuận lợi khi ghép kết quả giải trình tự của 2 môi xuôi để xử lý kết quả kiểu gen. Bên cạnh đó, khoảng cách từ hai môi xuôi SEQ ID No: 1 và SEQ ID No: 3 đến đoạn đầu của exon 2 và exon 3 đều ~ 100 bp giúp kết quả giải trình tự đoạn đầu các exon không bị nhiễu. Hai môi xuôi SEQ ID No: 1 và SEQ ID No: 3 có vị trí gắn trên khuôn ADN phù hợp giúp đảm bảo tín hiệu giải trình tự đoạn đầu các exon và không bỏ sót các vị trí dị hợp tử do giảm mức độ phân giải của các tín hiệu giải trình tự ở đoạn cuối các exon. Nhờ vậy, thay vì phải sử dụng nhiều môi và thực hiện nhiều phản ứng để giải trình tự 2 chiều như các công bố và sáng chế khác, quy trình được đề xuất trong sáng chế này đã giảm cả số môi sử dụng và số phản ứng cần thực hiện để giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C*.

### Ví dụ thực hiện sáng chế

Ví dụ 1: giải trình tự mẫu ADN có kiểu gen dị hợp tử *HLA-C\*01:02/C\*15:05* bằng quy trình được đề xuất trong sáng chế. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 được thể hiện ở Hình 3 (A). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 được thể hiện ở Hình 3 (B).

Ví dụ 2: giải trình tự mẫu ADN có kiểu gen đồng hợp tử *HLA-C\*03:02/C\*03:02* bằng quy trình được đề xuất trong sáng chế. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 được thể hiện ở Hình 3 (C). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 được thể hiện ở Hình 3 (D).

Ví dụ 3: giải trình tự mẫu ADN có kiểu gen dị hợp tử *HLA-C\*03:02/C\*03:04* bằng quy trình được đề xuất trong sáng chế. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 được thể hiện ở Hình 3 (E). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 được thể hiện ở Hình 3 (F).

Ví dụ 4: giải trình tự mẫu ADN có kiểu gen dị hợp tử *HLA-C\*01:02/C\*04:03* bằng quy trình được đề xuất trong sáng chế. Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 được thể hiện ở Hình 3 (G). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 được thể hiện ở Hình 3 (H).

- No.1 có kiểu gen dị hợp tử *HLA-C\*01:02/15:05*. (A). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.1 (B). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.1.
- No.2 có kiểu gen đồng hợp tử *HLA-C\*03:02/03:02*. (C). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.2 (D). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.2.
- No.3 có kiểu gen dị hợp tử *HLA-C\*03:02/03:04*. (E). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.3 (F). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.3.
- No.4 có kiểu gen dị hợp tử *HLA-C\*01:02/04:03*. (G). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 2 của mẫu No.4 (H). Kết quả giải trình tự đoạn đầu và đoạn cuối exon 3 của mẫu No.4.

## Tài liệu tham khảo

### Sáng chế

1. Thomas George NIETOJ, Joanne Dwan STOCKTON, Andrew David BEGGS, *Methods, compositions and kits for hla typing*, WO2021191634A1, 2021-09-30.
2. Caroline Elizabeth Garey, James Leushner, *Method and kit for amplification sequencing and typing of classical HLA class I genes*, US5910413A, 1999-06-08.
3. William Hildebrand, Mary Ellexson, Pierre Chretien, R. Duthie, *Class I sequence based typing of HLA-A, -B, and -C alleles by direct DNA sequencing*, US20030017458A1, 2003-01-23.
4. Malali GOWDA, *Super hla typing method and kit thereof*, WO2019229649A1, 2019-12-05.
5. Fumihiko Matsuda, Shuji Kawaguchi, Masakazu Shimizu, Koichiro Higasa, *PCR primer set for HLA gene, and sequencing method using same*, US20190367983A1, 2019-12-05.
6. Zhong Zheng, Liao Kuanzhen, Deng Jia, Liu Xian, Wang Liping, *For the primer of HLA gene high-resolution genotypings, kit and method*, CN108048546A, 2018-05-18.
7. Li Jian, Chen Shiping, Zhang Xiandong, Liu Ying, Zhang Caifen, Liu Tao, Zhao Meiru, *New pcr sequencing method and use thereof in hla genotyping*, KR101709826B1, 2017-02-24.

### Bài báo khoa học

8. Phạm Trần Thu Hà Trần Quang Bình, Nguyễn Hải Hà, Phùng Thanh Hương (2020), "Quy trình giải trình tự gen HLA-A bằng phương pháp Sanger", Tạp chí Y Dược học, 2, pp. 165-170.

9. Althaf M. M., El Kossi M., et al. (2017), "Human leukocyte antigen typing and crossmatch: A comprehensive review", *World J Transplant*, 7(6), pp. 339-348.
10. Apps Richard, Qi Ying, et al. (2013), "Influence of HLA-C expression level on HIV control", 340(6128), pp. 87-91.
11. Berlingerio Michele, Bonchi Francesco, et al. (2009), "Mining Clinical, Immunological, and Genetic Data of Solid Organ Transplantation", 224, pp. 211-236.
12. Cereb N., Maye P., et al. (1995), "Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles", *Tissue Antigens*, 45(1), pp. 1-11.
13. Dand Nick, Duckworth Michael, et al. (2019), "<em>HLA-C\*06:02</em> genotype is a predictive biomarker of biologic treatment response in psoriasis", *Journal of Allergy and Clinical Immunology*, 143(6), pp. 2120-2130.
14. Delfino L., Morabito A., et al. (2003), "HLA-C sequence based typing: nucleotide analysis from exon 1 through exon 8. Identification of a new allele: Cw\*0718", *Tissue Antigens*, 62(5), pp. 418-25.
15. Delfino L., Morabito A., et al. (1998), "HLA-C high resolution typing: analysis of exons 2 and 3 by sequence based typing and detection of polymorphisms in exons 1-5 by sequence specific primers", *Tissue Antigens*, 52(3), pp. 251-9.
16. Hoa B. K., Hang N. T., et al. (2008), "HLA-A, -B, -C, DRB1 and DQB1 alleles and haplotypes in the Kinh population in Vietnam", *Tissue Antigens*, 71(2), pp. 127-34.
17. Israeli Moshe, Roelen Dave L, et al. (2014), "Association between CTL precursor frequency to HLA-C mismatches and HLA-C antigen cell surface expression", 5, pp. 547.

18. Kulkarni Smita, Qi Ying, et al. (2013), "Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease", 110(51), pp. 20705-20710.
19. Li X., Jin S., et al. (2019), "Association of HLA-C\*03:02 with methimazole-induced liver injury in Graves' disease patients", *Biomed Pharmacother*, 117, pp. 109095.
20. Ljunggren Hans-Gustaf, Kärre Klas %J Immunology today (1990), "In search of the 'missing self': MHC molecules and NK cell recognition", 11, pp. 237-244.
21. McINTYRE JOHN A, FAUL W PAGE %J American journal of reproductive immunology (1983), "Recurrent spontaneous abortion in human pregnancy: results of immunogenetical, cellular, and humoral studies", 4(4), pp. 165-170.
22. Parham Peter (2008), *The genetic and evolutionary balances in human NK cell receptor diversity*, Seminars in immunology, Elsevier, pp. 311-316.
23. Peterson Trevor, Bielawny Thomas, et al. (2014), "Diversity and Frequencies of HLA Class I and Class II Genes of an East African Population", *Open Journal of Genetics*, 04, pp. 99-124.
24. Pham T. T. H., Tran Q. B., et al. (2021), "A Novel Allele-Specific PCR Protocol for the Detection of the HLA-C\*03:02 Allele, a Pharmacogenetic Marker, in Vietnamese Kinh People", *Appl Clin Genet*, 14, pp. 27-35.
25. Pham Tran Thu Ha, TranQuang Binh, et al. "Allopurinol-induced severe cutaneous adverse reactions in Vietnamese: the role of HLA alleles and other risk factors", 0(0), pp. null.
26. Robinson J., Barker D. J., et al. (2020), "IPD-IMGT/HLA Database", *Nucleic Acids Res*, 48(D1), pp. D948-D955.
27. Shiina T., Inoko H., et al. (2004), "An update of the HLA genomic region, locus information and disease associations: 2004", *Tissue Antigens*, 64(6), pp. 631-49.

28. Siegel R. J., Bridges S. L., Jr., et al. (2019), "HLA-C: An Accomplice in Rheumatic Diseases", *ACR Open Rheumatol*, 1(9), pp. 571-579.
29. Turner S., Ellexson M. E., et al. (1998), "Sequence-based typing provides a new look at HLA-C diversity", *J Immunol*, 161(3), pp. 1406-13.

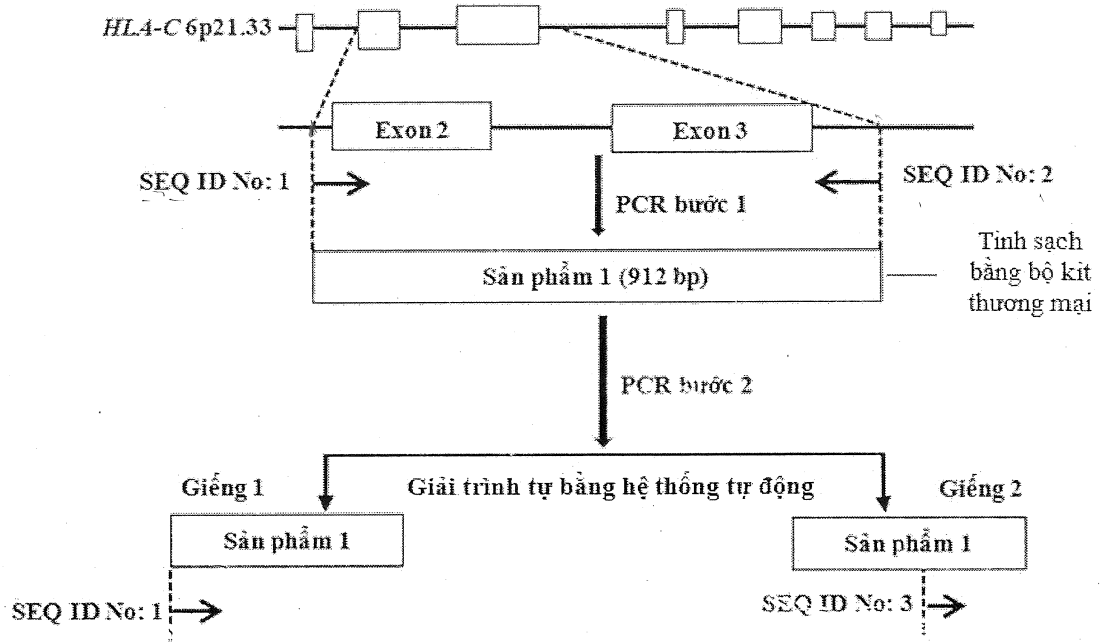
***Trang web***

30. The IPD-IMGT/HLA Allele Query Tool , “The IPD-IMGT/HLA Database”, Truy cập 27/03/2022, từ <https://www.ebi.ac.uk/ipd/imgt/hla/alleles/allele/?accession=HLA00401>.

**YÊU CẦU BẢO HỘ**

1. Quy trình giải trình tự vùng đa hình exon 2 – exon 3 của gen *HLA-C* bao gồm 2 bước:

- (i) thực hiện phản ứng PCR khuếch đại đặc hiệu vùng exon 2 – exon 3 của gen *HLA-C* với cặp mồi đặc hiệu gồm mồi xuôi có trình tự nêu trong SEQ ID No: 1 và mồi ngược có trình tự nêu trong SEQ ID No: 2 theo chiều 5' đến 3' để thu được sản phẩm PCR, tinh sạch sản phẩm PCR thu được bằng bộ kit thương mại và sử dụng làm khuôn cho các phản ứng giải trình tự ở bước (ii);
- (ii) thực hiện giải trình tự với 2 mồi xuôi có trình tự nêu trong SEQ ID No: 1 và SEQ ID No: 3 bằng hệ thống giải trình tự tự động.



Hình 1

**A**

SEQ ID No: 1 5'-GCGAG GTGCCCCGCC GGCTA-3'

| gDNA          | 120         | 130               | 140        |
|---------------|-------------|-------------------|------------|
| C*01:02:01:01 | GAGGAACGAG  | GTGCCCCGCC        | GGCGAGGGCG |
| C*01:03:01    | -----       | -----             | -----      |
| C*03:02:01    | -----G----- | -----G-----       | -----      |
| C*03:03:01:01 | -----G----- | -----G-----       | -----      |
| C*03:04:01:01 | -----G----- | -----G-----       | -----      |
| C*04:01:01:01 | -----G----- | -----G-----       | -----      |
| C*04:03:01:01 | -----G----- | -----G-----       | -----      |
| C*04:06:01    | -----G----- | -----G-----       | -----      |
| C*06:02:01:01 | -----G----- | -----G-----       | -----      |
| C*07:01:01:01 | -----G----- | -----G-----       | -----      |
| C*07:02:01:01 | -----G----- | -----G-----T----- | -----      |
| C*07:04:01:01 | -----G----- | -----G-----T----- | -----      |
| C*08:01:01:01 | -----G----- | -----G-----       | -----      |
| C*12:02:02:01 | -----G----- | -----G-----       | -----      |
| C*12:03:01:01 | -----G----- | -----G-----       | -----      |
| C*14:02:01:01 | -----G----- | -----G-----       | -----      |
| C*15:02:01:01 | -----G----- | -----G-----       | -----      |
| C*15:05:01:01 | -----G----- | -----G-----       | -----      |
| A*            | ---A---A--- | ---G---T---T---   | ---G---    |
| B*            | -----G----- | ---G-A-----A---   | ---G---    |
| E*            | ---T-GA---  | ---G---G-----     | ---G---    |
| F*            | -----GT---  | ---G-----         | ---T-G---  |
| G*            | -----GG---  | ---G---G-----     | ---G---    |

**B**

| Đoạn đầu vùng exon 2 |             | Đoạn cuối vùng exon 2 |            |             |
|----------------------|-------------|-----------------------|------------|-------------|
| gDNA                 | 210         | 220                   | 470        | 480         |
| C*01:02:01:01        | CAG GCTCCCA | CTCCATGAAG            | CAGAGCGAGG | CCG GTGAGTG |
| C*01:03:01           | ----- ----- | -----                 | -----      | -----       |
| C*03:02:01           | ----- ----- | -----G-               | -----      | -----       |
| C*03:03:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*03:04:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*04:01:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*04:03:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*04:06:01           | ----- ----- | -----G-               | -----      | -----       |
| C*06:02:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*07:01:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*07:02:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*07:04:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*08:01:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*12:02:02:01        | ----- ----- | -----G-               | -----      | -----       |
| C*12:03:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*14:02:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*15:02:01:01        | ----- ----- | -----G-               | -----      | -----       |
| C*15:05:01:01        | ----- ----- | -----G-               | -----      | -----       |

**C**

SEQ ID No: 3 5'-CC GGAGAGAGCC CCAGT-3'

| gDNA          | 600        | 610        | 620        |
|---------------|------------|------------|------------|
| C*01:02:01:01 | ACCCTCGACC | GGAGAGAGCC | CCAGTCACCT |
| C*01:03:01    | -----      | -----      | -----      |
| C*03:02:01    | -----      | -----      | -----      |
| C*03:03:01:01 | -----      | -----      | -----      |
| C*03:04:01:01 | -----      | -----      | -----      |
| C*04:01:01:01 | -----      | -----      | -----      |
| C*04:03:01:01 | -----      | -----      | -----      |
| C*04:06:01    | -----      | -----      | -----      |
| C*06:02:01:01 | -----      | -----      | -----      |
| C*07:01:01:01 | -----      | -----      | -----G---  |
| C*07:02:01:01 | -----      | -----      | -----G---  |
| C*07:04:01:01 | -----      | -----      | -----G---  |
| C*08:01:01:01 | -----      | -----      | -----      |
| C*12:02:02:01 | -----      | -----      | -----      |
| C*12:03:01:01 | -----      | -----      | -----      |
| C*14:02:01:01 | -----      | -----      | -----      |
| C*15:02:01:01 | -----      | -----      | -----      |
| C*15:05:01:01 | -----      | -----      | -----      |

**D**

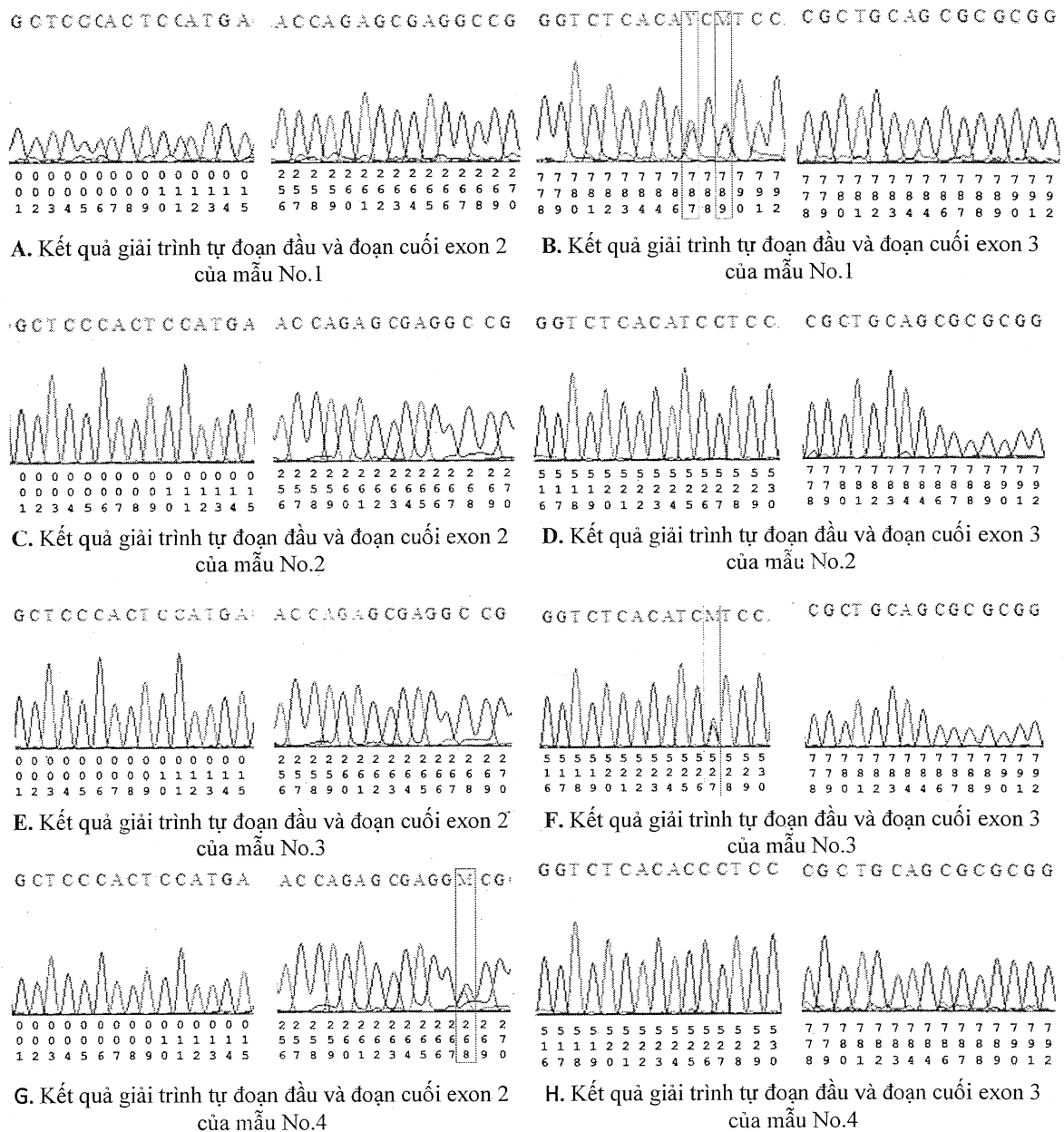
| Đoạn đầu vùng exon 3 |             | Đoạn cuối vùng exon 3 |            |             |
|----------------------|-------------|-----------------------|------------|-------------|
| gDNA                 | 720         | 730                   | 990        | 1000        |
| C*01:02:01:01        | CGGGGCCAG G | GTCTCACACC            | CGCTGCAGCG | CGCGG GTACC |
| C*01:03:01           | ----- ----- | -----T-               | -----      | -----       |
| C*03:02:01           | ----- ----- | -----T-               | -----      | -----       |
| C*03:03:01:01        | ----- ----- | -----T-               | -----      | -----       |
| C*03:04:01:01        | ----- ----- | -----T-               | -----      | -----       |
| C*04:01:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*04:03:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*04:06:01           | ----- ----- | -----                 | -----      | -----       |
| C*06:02:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*07:01:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*07:02:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*07:04:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*08:01:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*12:02:02:01        | ----- ----- | -----                 | -----      | -----       |
| C*12:03:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*14:02:01:01        | ----- ----- | -----                 | -----      | -----       |
| C*15:02:01:01        | ----- ----- | -----T-               | -----      | -----       |
| C*15:05:01:01        | ----- ----- | -----T-               | -----      | -----       |

**E**

SEQ ID No: 2 3'-TCAC CCCTCGGAAG GGGTAGAG-5'

| gDNA          | 1010       | 1020        | 1030        |
|---------------|------------|-------------|-------------|
| C*01:02:01:01 | AGGGGCAGTG | GGGAGCCTTC  | CCCATCTCC   |
| C*01:03:01    | -----      | -----       | -----       |
| C*03:02:01    | -----      | -----       | -----T      |
| C*03:03:01:01 | -----      | -----       | -----T      |
| C*03:04:01:01 | -----      | -----       | -----T      |
| C*04:01:01:01 | -----      | -----       | -----       |
| C*04:03:01:01 | -----      | -----       | -----       |
| C*04:06:01    | -----      | -----       | -----       |
| C*06:02:01:01 | -----      | -----       | -----T      |
| C*07:01:01:01 | -----      | -----       | -----T      |
| C*07:02:01:01 | -----      | -----       | -----T      |
| C*07:04:01:01 | -----      | -----       | -----T      |
| C*08:01:01:01 | -----      | -----       | -----T      |
| C*12:02:02:01 | -----      | -----       | -----T      |
| C*12:03:01:01 | -----      | -----       | -----T      |
| C*14:02:01:01 | -----      | -----       | -----T      |
| C*15:02:01:01 | -----      | -----       | -----T      |
| C*15:05:01:01 | -----      | -----       | -----T      |
| A*            | -----CAC-  | ---C---C-   | -TG---G---T |
| B*            | -----      | -----       | -----T      |
| E*            | -----CACA  | ---CTA---CT | -----T      |
| F*            | -----CA--  | ---C-----   | ---T-----T  |
| G*            | -----      | ---C---C-   | -TG-----T   |

Hình 2



Hình 3

## Danh mục trình tự

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <110>      | Trường Đại học Dược Hà Nội.                                                                                                                                                                                                                                                                                                                                                                                                                                                    |    |
| <120>      | Quy trình giải trình tự vùng đa hình exon 2 - exon 3 của gen <i>HLA-C</i>                                                                                                                                                                                                                                                                                                                                                                                                      |    |
| <160>      | 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <210>      | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <211>      | 20                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
| <212>      | Oligonucleotit                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
| <213>      | Trình tự nhân tạo                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <220>      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
| <223>      | Mỗi xuôi đặc hiệu alen, bắt cặp bổ sung với sợi ADN chiều 5'-3' của gen <i>HLA-C</i> tại vị trí nucleotit 116 - 135, có một nucleotit không bắt cặp bổ sung với khuôn ADN, thay g bằng t vào vị trí ngay cạnh nucleotit tận cùng đầu 3' để tăng tính đặc hiệu của mỗi với gen <i>HLA-C</i> , cùng với mỗi ngược SEQ ID No: 2 tạo thành sản phẩm PCR có kích thước 912 bp. Ngoài ra, mỗi xuôi này còn được sử dụng ở bước 2 để giải trình tự vùng exon 2 của gen <i>HLA-C</i> . |    |
| <400>      | 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| gcgaggtgcc | cgcccggtta                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 20 |
| <210>      | 2                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <211>      | 22                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
| <212>      | Oligonucleotit                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
| <213>      | Trình tự nhân tạo                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <220>      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
| <223>      | Mỗi ngược, bắt cặp bổ sung với sợi ADN chiều 3'-5' của gen <i>HLA-C</i> tại vị trí nucleotit 1007 - 1028, cùng với mỗi xuôi SEQ ID No: 1 tạo thành sản phẩm PCR có kích thước 912 bp.                                                                                                                                                                                                                                                                                          |    |
| <400>      | 2                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| gagatgggga | aggctcccca ct                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 22 |
| <210>      | 3                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <211>      | 17                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
| <212>      | Oligonucleotit                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
| <213>      | Trình tự nhân tạo                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <220>      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |    |
| <223>      | Mỗi xuôi bắt cặp bổ sung với sợi ADN chiều 5'-3' của gen <i>HLA-C</i> tại vị trí nucleotit 599 - 615, sử dụng để giải trình tự vùng exon 3 của gen <i>HLA-C</i> .                                                                                                                                                                                                                                                                                                              |    |
| <400>      | 3                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| CCGGAGAGAG | CCCCAGT                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 17 |
| <210>      | 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
| <211>      | 4304                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |    |
| <212>      | ADN                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |    |

|             |                                                                                                                                                                                                                                      |            |            |             |             |      |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|-------------|-------------|------|
| <213>       | Homo sapiens.                                                                                                                                                                                                                        |            |            |             |             |      |
| <300>       |                                                                                                                                                                                                                                      |            |            |             |             |      |
| <301>       | S.Suzuki, S.Ranade, S.Ito                                                                                                                                                                                                            |            |            |             |             |      |
| <302>       | Reference Grade Characterization of Polymorphisms in Full-Length HLA Class I and II Genes With Short-Read Sequencing on the ION PGM System and Long-Reads Generated by Single Molecule, Real-Time Sequencing on the PacBio Platform. |            |            |             |             |      |
| <303>       | Frontiers in Immunology                                                                                                                                                                                                              |            |            |             |             |      |
| <304>       | 9                                                                                                                                                                                                                                    |            |            |             |             |      |
| <305>       | 2294                                                                                                                                                                                                                                 |            |            |             |             |      |
| <306>       | Ấn bản trực tuyến                                                                                                                                                                                                                    |            |            |             |             |      |
| <307>       | 2018-10-04                                                                                                                                                                                                                           |            |            |             |             |      |
| <308>       | 3037930                                                                                                                                                                                                                              |            |            |             |             |      |
| <309>       | 2022-03-27                                                                                                                                                                                                                           |            |            |             |             |      |
| <400>       | 4                                                                                                                                                                                                                                    |            |            |             |             |      |
| ttatttttgc  | ggatgtagtt                                                                                                                                                                                                                           | taatattacc | tgaggtaagg | taaggcaaag  | agtgggagggc | 60   |
| agggagtcca  | gttcaggggac                                                                                                                                                                                                                          | ggggattcca | ggagaagtga | aggggaaggg  | gctggggcgca | 120  |
| gcctgggggt  | ctctccctgg                                                                                                                                                                                                                           | tttccacaga | cagatccttg | gccaggactc  | aggcacacag  | 180  |
| tgtgacaaaag | atgcttggtg                                                                                                                                                                                                                           | taggagaaga | gggatcagga | cgaagtccca  | gggtcccgggc | 240  |
| ggggttctca  | gggtctcagg                                                                                                                                                                                                                           | ctccaagggc | cgtgtctgca | ctggggagggc | gccgcgttga  | 300  |
| ggattctcca  | ctcccctgag                                                                                                                                                                                                                           | tttcaattct | tctcccaacc | tgcgtcgggt  | ccttcttctct | 360  |
| gaatactcat  | gacgcgtccc                                                                                                                                                                                                                           | caattcccac | tcccattggg | tgtcgggttc  | tagagaagcc  | 420  |
| aatcagcgtc  | tccgcagtc                                                                                                                                                                                                                            | cggttctaaa | gtcccagtc  | accacccgg   | actcagattc  | 480  |
| tcccagacg   | ccgagatgcg                                                                                                                                                                                                                           | ggtcatggcg | ccccgaacc  | tcctcctgct  | gctctcggga  | 540  |
| gccctggccc  | tgaccgagac                                                                                                                                                                                                                           | ctgggcctgt | gagtgcgggg | ttgggagggg  | aacggcctct  | 600  |
| gcggagagga  | acgaggtgcc                                                                                                                                                                                                                           | cgcccggcga | gggcgcagga | cccggggagc  | cgcgcagggg  | 660  |
| ggaggggtcg  | gcgggtctca                                                                                                                                                                                                                           | gcccctcctc | gccccaggc  | tcccactcca  | tgaagtattt  | 720  |
| cttcacatcc  | gtgtcccggc                                                                                                                                                                                                                           | ctggccgcgg | agagccccgc | ttcatctcag  | tgggctacgt  | 780  |
| ggacgacacg  | cagttcgtgc                                                                                                                                                                                                                           | ggttcgacag | cgacgcgcgc | agtcgagag   | gggagccgcg  | 840  |
| ggcgccgtgg  | gtggagcagg                                                                                                                                                                                                                           | aggggcccga | gtattgggac | cggyagacac  | agaagtacaa  | 900  |
| gcgcraggca  | cagactgacc                                                                                                                                                                                                                           | gagtgagcct | gcggaacctg | cgcggctact  | acaaccagag  | 960  |
| cgagggccgt  | gagtgacccc                                                                                                                                                                                                                           | ggcccggggc | gcaggtcacg | accctcctc   | atccccacg   | 1020 |
| gacggcccgg  | gtcgcccca                                                                                                                                                                                                                            | gtctcccgtt | ctgagatcca | ccccgaggct  | gcggaacccg  | 1080 |
| cccagacct   | cgaccggaga                                                                                                                                                                                                                           | gagccccagt | cacctttacc | cggtttcatt  | ttcagtttag  | 1140 |
| gccaaaatcc  | ccgcgggttg                                                                                                                                                                                                                           | gtcgggactg | gggcggggct | cgggggacgg  | ggctgaccac  | 1200 |
| gggggcgggg  | ccagggtctc                                                                                                                                                                                                                           | acaccctcca | gtggatgtgt | ggctgcgacc  | tggggcccga  | 1260 |
| cgggcgcctc  | ctccgcgggt                                                                                                                                                                                                                           | atgaccagta | cgcctacgac | ggcaaggatt  | acatcgccct  | 1320 |
| gaacgaggac  | ctgcgtcctt                                                                                                                                                                                                                           | ggaccgcgc  | ggacaccgcg | gtcagatca   | cccagcgcaa  | 1380 |
| gtgggagggc  | gcccgtgagg                                                                                                                                                                                                                           | cggagcagcg | gagagcctac | ctggagggca  | cgtgcgtgga  | 1440 |
| gtggctccgc  | agatacctgg                                                                                                                                                                                                                           | agaacgggaa | ggagacgctg | cagcgcgcgcg | gtaccagggg  | 1500 |
| cagtggggag  | ccttcccat                                                                                                                                                                                                                            | ctcccgtaga | tctcccggca | tggcctccca  | cgaggagggg  | 1560 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aggaaaatgg gatcagcgct agaatatcgc cctcccttga atggagaatg ggatgagttt  | 1620 |
| tcctgagttt cctctgaggg cccctctgc tctctaggac aattaaggga tgaagtcctt   | 1680 |
| gaggaaatgg aggggaagac agtccctgga atactgatca ggggtcccct ttgaccactt  | 1740 |
| tgaccactgc agcagctgtg gtcaggctgc tgacctttct ctcaggcctt gttctctgcc  | 1800 |
| tcacgctcaa tgtgtttgaa ggtttgatcc cagcttttct gagtccttcg gcctccactc  | 1860 |
| aggtcaggac cagaagtcgc tgttcctccc tcagagacta gaactttcca atgaatagga  | 1920 |
| gattatccca ggtgcctgtg tccaggctgg cgtctgggtt ctgtgcccc ttccccaccc   | 1980 |
| cagggtgtcct gtccattctc aggatggtca catgggcgct gttggagtgt cgcaagagag | 2040 |
| atacaaagtg tctgaatttt ctgactcttc cgtcagaac acccaaagac acacgtgacc   | 2100 |
| caccatcccg tctctgacca tgaggccacc ctgaggtgct gggccctggg cttctaccct  | 2160 |
| gcggagatca cactgacctg gcagtgggat ggggaggacc aaactcagga caccgagctt  | 2220 |
| gtggagacca ggccagcagg agatggaacc ttccagaagt gggcagctgt gatggtgcct  | 2280 |
| tctggagaag agcagagata cacgtgccat gtgcagcagc aggggctgcc ggagcccctc  | 2340 |
| accctgagat ggggtaagga gggggatgag ggggtgatgtg tcttctcagg gaaagcagaa | 2400 |
| gtcctggagc ccttcagcng ggtcagggtt gaggcttggg ggtcagggcc cctcaccctc  | 2460 |
| ccctcccttc ccagagccgt cttcccagcc caccatcccc atcgtgggca tcgttgetgg  | 2520 |
| cctggctgtc ctggctgtcc tagctgtcct aggagctgtg gtggctgttg tgatgtgtag  | 2580 |
| gaggaagagc tcaggtaggg aaggggtgag gagtggggtc tgggttttct tgttccactg  | 2640 |
| ggagtttcaa gcccaggtg gaagtgtgcc ccacctcggt actggaagca ccattccacac  | 2700 |
| atgggccatc ccagcctggg accctgtgtg ccagcactta ctctgttgtg aagcacatga  | 2760 |
| caatgaagga cagatgtatc accttgatga ttatggtgtt ggggtccttg attccagcat  | 2820 |
| tcattgagtca ggggaaggtc cctgctaagg acagacctta ggagggcagt tgcttcagaa | 2880 |
| cccacagctg ctttccccgt gtttcctgat cctgccctgg gtctgcagtc atagttctgg  | 2940 |
| aaacttctct tgggtccaag actaggaggt tcccctaaga tcgcatggcc ctgcctcctc  | 3000 |
| cctgtccctt cacagggcat tttcttccca cagggtgaaa aggagggagc tgctctcagg  | 3060 |
| ctgcgtgtaa gtgatggcgg tgggcgtgtg gaggagctca cccaccccat aattcctctt  | 3120 |
| gtcccacatc tcctgcgggc tctgaccagg tctttttttt tgttctaccc cagccagcaa  | 3180 |
| cagtgccagc ggtctgatg agtctctcat cgttgtaaa ggtgagattc tggggagctg    | 3240 |
| aagtggctcg ggggtgggca gagggaaaag gcctaggtaa tggggatcct ttgattggga  | 3300 |
| cgtttcgaat gtgtggtgag ctgttcagag tgatcact tactatgact gactgaatt     | 3360 |
| tgttcatgac tattgtgttc tctagcctga gacagctgcc tctgtgggac tgagatgcag  | 3420 |
| gattttctca cacctctcct ttgtgacttc aagagcctct ggcattctct tctgcaaagg  | 3480 |
| catctgaatg tgtctgcgtt cctgttagca taatgtgagg aggtggagag acagcccacc  | 3540 |
| cccggtlcca ccgtgaacccc tgtcccaca ctgacctgtg ttccctcccc gatcalttt   | 3600 |
| cctgttccag agaagtgggc tggatgtctc catctctgtc tcaacttcat ggtgcgtga   | 3660 |
| gctgcaactt cttacttccc taatgaagtt aagaacctga atataaattt gttttctcaa  | 3720 |
| atatttgcta tgaaggggtg atggattaat taaataagtc aattcctgga agttgagaga  | 3780 |
| gcaaataaag aactgagaac cttccagaat ccgcatgttc gctgtgctga gctgttgca   | 3840 |
| gggtggggtg ggggaaggct tgaggagacg agtgtggacg gggcctgtgc ctagtgtgtg  | 3900 |
| ttcagttctt catgggcttt atgtggtcag tctcagctg ggtcaccttc actgctccat   | 3960 |
| tgtccttgtc ccttcagtgg aaacttgtcc agcgggagct gtgaccacag aggtccacac  | 4020 |
| atcgcccagg gcagccctg cacacgggag tccctgtgct tctgagaca aattttcaga    | 4080 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cccagtcagc tcctgccctc cttctagggc tcctcttctg ctttgggtctc ctgccctctc | 4140 |
| tccttccct gattccagtg atcttcgtgc tgactccaat cccaactcat gaatctaaag   | 4200 |
| cagagcctaa tttagattta tatttgtttg taaaattggg tccatagtct agaattgttc  | 4260 |
| cttctgaag agagaaacct gattgtgtgc tgcagtgtgc gggg                    | 4304 |