

Phân lập gene mã hóa enzyme *cis*-prenyltransferase 4 (CPT4) từ cây Cao su (*Hevea brasiliensis*)

Nguyễn Huỳnh Cẩm Tú^{1,*}, Phạm Thị Mỹ Bình¹, Trần Thanh Lượng¹,
Trần Thanh², Nguyễn Thị Hồng Thương¹

¹Khoa Sinh học và Công nghệ sinh học, Trường Đại học Khoa Học Tự Nhiên, ĐHQGHCM, Việt Nam
²Bộ môn Giống, Viện Nghiên cứu Cao su Việt Nam

Nhận ngày 16 tháng 8 năm 2017

Chỉnh sửa ngày 20 tháng 9 năm 2017; Chấp nhận đăng ngày 10 tháng 10 năm 2017

Tóm tắt: *Cis*-prenyltransferase (CPT) là enzyme xúc tác phản ứng trùng ngưng liên tiếp các đơn vị isopentenyl diphosphate (IPP) lên các chất nhận allylic diphosphate để tổng hợp các *cis*-prenyldiphosphate với chiều dài chuỗi khác nhau, từ neryl diphosphate đến cao su thiên nhiên. Ngoài hai trình tự *HRT1* (*Hevea rubber transferase 1*) và *HRT2* (*Hevea rubber transferase 1*) mã hóa cho các *cis*-prenyltransferase tham gia trong quá trình tổng hợp cao su thiên nhiên đã được phân lập và nghiên cứu, các gene mã hóa các CPT còn lại trong họ enzyme *cis*-prenyltransferase vẫn chưa được khảo sát chức năng. Kết quả tra cứu ban đầu trên cơ sở dữ liệu hệ phiên mã của cây cao su *Hevea brasiliensis* bằng phần mềm TBLASTX với trình tự truy vấn *HRT2* cho ra 6 “hit” trình tự. Trong số này có một “hit” chứa trình tự mã hóa protein tương đồng 75% với *HRT2* và được đặt tên là *HbCPT4*. Gene *HbCPT4* được dự đoán có 3 exon và 2 intron. Trình tự mã hóa (CDS) *HbCPT4* được phân lập thành công từ mẫu cDNA của mủ cao su *Hevea brasiliensis* RRIV 209, được giải trình tự và so sánh với trình tự dự đoán *in silico*. Trình tự CDS *HbCPT4* phân lập từ thực nghiệm và trình tự CDS *HbCPT4* được dự đoán *in silico* khác nhau 7 nucleotide, tuy nhiên hai trình tự polypeptide được suy ra từ hai trình tự CDS này thì hoàn toàn giống nhau. Protein *HbCPT4* được dự đoán có 296 amino acid, có trọng lượng phân tử khoảng 34 kDa, pI khoảng 8,19 và chứa 5 vùng bảo tồn (vùng I–V) liên quan đến hoạt tính xúc tác và khả năng liên kết cơ chất của các enzyme họ *cis*-prenyltransferase (CPT).

Từ khóa: *Cis*-prenyltransferase, isoprenoid, *Hevea brasiliensis*, mủ cao su.

1. Mở đầu

Isoprenoid (terpenoid) từ thực vật là nhóm hợp chất tự nhiên đa dạng về mặt cấu trúc và có nhiều ứng dụng trong đời sống hàng ngày của con người ở các lĩnh vực khác nhau như dược phẩm, thực phẩm, mỹ phẩm và gần đây chúng

được khai thác để sản xuất nhiên liệu sinh học [1]. Ở thực vật, isoprenoid được tổng hợp từ isopentenyl diphosphate (IPP) và dimethylallyl diphosphate (DMAPP) thông qua con đường methylerythritol 4-phosphate (MEP) diễn ra ở lục thể hoặc con đường mevalonate (MEV) diễn ra ở tế bào chất [2].

Cis-prenyltransferase (CPT) là một trong các enzyme quan trọng tham gia vào con đường sinh tổng hợp isoprenoid. *Cis*-prenyltransferase

* Tác giả liên hệ. ĐT.: 84-1265221859.

Email: camtu2601@gmail.com

<https://doi.org/10.25073/2588-1140/vnunst.4562>

xúc tác phản ứng trùng ngưng liên tiếp các IPP lên các chất nhận “allylic diphosphate” để tạo ra sản phẩm là các *cis*-prenyldiphosphate chuỗi ngắn như neryl diphosphate (NPP), (*Z,Z*)-farnesyl diphosphate (*(Z,Z)*-FPP), nerylneryl diphosphate (NNPP), hoặc các *cis*-polyisoprenoid chuỗi dài như dolichol, polyprenol hay cao su thiên nhiên. Vai trò của các CPT trong vi khuẩn, nấm men và động vật có vú đã được nghiên cứu khá kỹ, tuy nhiên, những hiểu biết về CPT ở thực vật vẫn còn hạn chế [2].

Cây Cao su (*Hevea brasiliensis*) là cây thân gỗ thuộc họ Euphorbiaceae [3], có năng suất mủ cao và cho cao su thiên nhiên với các đặc tính vượt trội. Cao su tự nhiên là một *cis*-polyisoprenoid được tổng hợp từ các đơn vị isoprene (C5) dưới sự xúc tác của *cis*-prenyltransferase (CPT) hay còn gọi là rubber transferase. Hiện nay, chỉ có hai gene *HRT1* (*Hevea rubber transferase 1*) và *HRT2* (*Hevea rubber transferase 2*) mã hóa cho *cis*-prenyltransferase liên quan đến quá trình sinh tổng hợp cao su thiên nhiên đã được phân lập và nghiên cứu [4]. Gần đây, bộ gene và hệ phiên mã của cây Cao su (*Hevea brasiliensis*) đã được giải trình tự và các nhà nghiên cứu đang dần hoàn thiện cơ sở dữ liệu bộ gene của loài này [3, 5, 6]. Điều này mở ra cơ hội dự đoán và nghiên cứu chức năng các gene mã hóa cho các protein quy định các tính trạng quan trọng ở loài thực vật này, làm cơ sở để xây dựng các phương pháp chọn tạo giống cây cao su một cách hiệu quả. Trong nghiên cứu này, chúng tôi sử dụng *HRT2* làm trình tự truy vấn để tra cứu cơ sở dữ liệu hệ phiên mã (TSA database) của cây Cao su (*Hevea brasiliensis*) bằng phần mềm tin sinh học và xác định được thêm ít nhất 8 trình tự mã hóa *CPT* tương đồng cao với *HRT2* (đặt tên là *HbCPT1-8*). Ngoại trừ *HRT1* và *HRT2*, chức năng của các gene mã hóa các *CPT* còn lại vẫn chưa được làm sáng tỏ. Trong nỗ lực phân tích chức năng của các gene *CPT* mới dự đoán, chúng tôi đã phân lập được một trình tự mã hóa *HbCPT* hoàn chỉnh có trình tự giống với trình tự *HbCPT4* được dự đoán ở

trên từ mủ của giống cao su *Hevea brasiliensis* RRIV 209.

2. Vật liệu và phương pháp

2.1. Vật liệu

Mủ của cây cao su (*Hevea brasiliensis*) RRIV 209 14 năm tuổi được trồng tại Viện Nghiên cứu Cao su Việt Nam. Vector pJET1.2/blunt (Thermo Scientific) mang gene kháng ampicillin. Chủng vi khuẩn *E. coli* TOP10 (Invitrogen) được sử dụng để nhân bản plasmid.

2.2. Phương pháp

2.2.1. Dự đoán các trình tự gene *HbCPT* mới từ cây cao su (*Hevea brasiliensis*) bằng các công cụ tin sinh học

Sử dụng trình tự *HRT2* (mã số Genbank: AB064661.2) làm trình tự truy vấn và công cụ tìm kiếm TBLASTX để tra cứu cơ sở dữ liệu hệ phiên mã của giống *Hevea brasiliensis* (taxid:3981) trên NCBI (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3575267/>) nhằm xác định các “hit” chứa trình tự mã hóa protein tương đồng với *HRT2*. Sau đó, phần mềm “Expasy Translate” được sử dụng để chuyển đổi trình tự nucleotide sang trình tự amino acid tương ứng theo mã bộ ba, làm cơ sở để dự đoán các “khung đọc mở” khả hữu trên trình tự “hit” có được. Giới hạn của trình tự mã hóa (CDS) được xác định bằng cách dò tìm vị trí codon khởi đầu (ATG) và codon kết thúc (TAG, TAA và TGA) trên các khung đọc mở dự đoán ở trên. Sau đó, trình tự polypeptide dự đoán dựa trên các CDS đã xác định ở bước trên được tra cứu trở lại trên cơ sở dữ liệu “non-redundant protein sequence” của NCBI để xác nhận sự tương đồng của trình tự polypeptide đó với *HRT2* và các trình tự *CPT* đã được công bố và lưu trữ trong cơ sở dữ liệu. Trình tự CDS được sử dụng làm trình tự truy vấn để xác định trình tự genomic tương ứng của

nó trong bản đồ bộ gene của *Hevea brasiliensis*. Cấu trúc gene *HbCPT* được dự đoán sơ bộ bằng công cụ FGENESH (www.softberry.com) và sau đó được hiệu chỉnh thủ công bằng cách đối sánh trình tự genomic với trình tự CDS dự đoán ở trên.

2.2.2. Tách chiết RNA tổng số và tổng hợp cDNA từ mô cao su

Mẫu mô là sự tổ hợp mô của 9 cây cao su khác nhau nhằm tăng tính đại diện cho mẫu, sau đó được trữ trong nitơ lỏng và được tách chiết RNA tổng số theo hướng dẫn của EZ-10 Spin column Plant RNA Mini-preps Kit (BioBasic). Quá trình tổng hợp cDNA được thực hiện theo hướng dẫn của Reverse transcriptase Aid First Strand Kit (Thermo Fisher Scientific).

2.2.3. Phân lập trình tự mã hóa protein *HbCPT4* từ cây cao su *Hevea brasiliensis*

Trình tự mã hóa protein *HbCPT4* được phân lập từ cDNA của mô cây cao su *H. brasiliensis* RRIV 209 bằng PCR sử dụng Phusion DNA polymerase (Thermo Scientific) và cặp mồi đặc hiệu (XhoI-A-*HbCPT4*-F: 5'-CTCGAGAATGGAAATATATACGGGTCA G-3' và BamHI-A-*HbCPT4*-R: 5'-GGATCCATTTCAAATATTCCTTGTGCTTC-3'). Chu kỳ nhiệt cho phản ứng PCR: 98°C/30 giây, 5 chu kỳ gồm 98°C/10 giây, 56°C/20 giây, 72°C/30 giây; 32 chu kỳ gồm 98°C/10 giây, 63°C/20 giây, 72°C/30 giây; 72°C/5 phút và giữ ở 10°C trong 15 phút. Sản phẩm PCR được điện di trên gel agarose 1% và đoạn CDS *HbCPT4* được tinh sạch từ gel bằng GeneJET Gel Extraction Kit (Thermo Scientific). Phân đoạn sau tinh sạch được chèn vào vector pJET1.2/blunt (Thermo Scientific) bằng phản ứng nối được xúc tác bởi T4 DNA ligase tạo thành plasmid pJET-*HbCPT4*. Hỗn hợp phản ứng nối được biến nạp vào tế bào khả nạp *E. coli* TOP10 bằng phương pháp hóa biến nạp. Các thể biến nạp *E. coli* TOP10/pJET-*HbCPT4* được sàng lọc bằng phản ứng PCR khuôn lạc với mồi xuôi bắt cặp đặc hiệu với trình tự trên vector (pJET1.2_F) và mồi ngược bắt cặp đặc hiệu với trình tự gene *HbCPT4* (XhoI-A-

HbCPT4-F). Khuẩn lạc cho kết quả PCR khuôn lạc dương tính được lựa chọn để nuôi tăng sinh khối và tách chiết plasmid bằng GeneJET Plasmid Miniprep Kit (Thermo Scientific). Plasmid pJET-*HbCPT4* được cắt kiểm tra kích thước bằng phản ứng cắt với enzyme cắt giới hạn *Bam*HI và *Xho*I. Trình tự CDS *HbCPT4* trên plasmid pJET-*HbCPT4* được gửi giải mã với mồi pJET1.2_F.

2.2.4. Phân tích và hiệu chỉnh trình tự mã hóa protein *HbCPT4*

Trình tự mã hóa protein *HbCPT4* trên plasmid tái tổ hợp pJET-*HbCPT4* được so sánh với trình tự mã hóa protein *HbCPT4* dự đoán *in silico* bằng phần mềm ClustalW (<http://www.geneome.jp/tools/clustalw/>). Trình tự CDS được dịch mã thành trình tự protein *HbCPT4* tương ứng bằng phần mềm “ExPasy Translate” (<http://web.expasy.org/translate/>). Các thông số khác của protein *HbCPT4* như giá trị pI, khối lượng phân tử (Mr) được dự đoán thông qua phần mềm “Compute pI/MW tool” (http://web.expasy.org/compute_pi/).

3. Kết quả thảo luận

3.1. Sử dụng các công cụ tin sinh để tìm và dự đoán cấu trúc gene *HbCPT*

Kết quả TBLASTX cơ sở dữ liệu hệ phiên mã (TSA database) của cao su *Hevea brasiliensis* (taxid:3981) với trình tự truy vấn là *HRT2* (mã số Genbank: AB064661.2) để tìm ra các “hit” trình tự có độ tương đồng cao với *HRT2* đã xác định được 6 contig, trong đó có contig10897 với giá trị E-value là 2e-11. Sử dụng kết hợp các phương pháp phân tích tin sinh học như mô tả ở mục 2.2.1, chúng tôi xác định được một gene bao gồm 3 exon và 2 intron (Hình 1), dự đoán mã hóa protein có 296 amino acid và tương đồng 75% với trình tự protein *HRT2*. Gene này được đặt tên là *HbCPT4*.

>HbCPT4

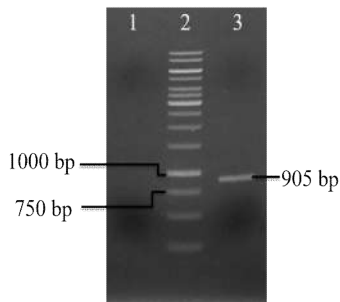
ATGGAAATATATACGGGTCAGAGGCCAAGTGTGTTTAAACTTTTTGGGAAATTTATGAGAAAAGGGTTATATC
GCATCCTAACCCAAGGTCCCATTCTACTCATCTTGCCTTCATATTGGATGGAAACCGGAGGTTTGCTAAGAAG
CATAAAATGAACGAAGCAGAAGGTTACAAGGCAGGATATTTAGCTCTTCTGAAAACACTAACTTATTGCTATG
AGTTGGGAGTGAGGTACGTAACCATTTATGCCTTTAGCATTGATAATTTTCGAAGGCAACCTCAGGAGGTTTCAG
TGCGTAATGAATCTTATGATGGAGAAGATTGAAGAGATTATTGTGGAAGAAAGTATCATGAATGCATATGATG
TTGGCGTACGTATTGTGGGTAACCTGAAGCTTTTAGATGAGCCAATCAGGATTGCAGCAGAAAAAATTATGAG
GGCTACTGCCAATAATCCAGGTTTGTGCTTCTCATTGCTATAGCCTATAGTTCAACTGATGAGATCGTGCATG
CTGTAGAAGAATCCTCTAAAGATAAATTGAACTCCAATGAAGTTTGCACAATGGGATTGAAGCTGAACAAGA
ATTTAAGGAGGCAAATGGAAGTGGAAACAGTGTGATTCTGTACAGAAGACGGAGTCATATTCTGGAATAAAT
CTTGTAGACCTTGAGAAAAACACCTACGTAATCCTTATCCTGATGTCCTGATTTCGAACCTTCTGGGTTGAGCCG
TCTAAGTAACCTACCTACTTTGGCAGACTAGCAATTGCATACATGTATTCTCCTTTTGCACCTGTGGCCAGAGATTG
GTCTCGGACACTTGGTATGGACAGTAATGGACTTCCAACGTCATCATTCTTATTTGAAGAAGCACAAGGAATAT
TTGAAATAA

Hình 1. Cấu trúc dự đoán của gene *HbCPT4* (các exon được gạch dưới).

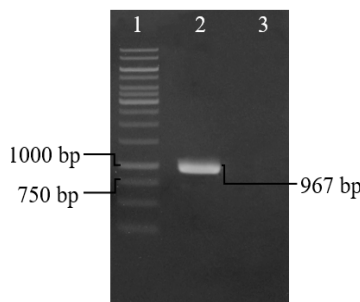
3.2. Phân lập trình tự mã hóa protein *HbCPT4* từ cây Cao su (*H. Brasiliensis*)

cDNA được tổng hợp từ RNA tổng số chiết từ mủ cao su *H. brasiliensis* RRIV 209 được sử dụng làm khuôn cho các phản ứng PCR với cặp mồi đặc hiệu với gene *HbCPT4*. Kết quả điện di sản phẩm PCR từ khuôn cDNA cho thấy có sự xuất hiện một vạch lớn hơn 750 bp và nhỏ hơn 1000 bp khi so sánh với thang DNA chuẩn

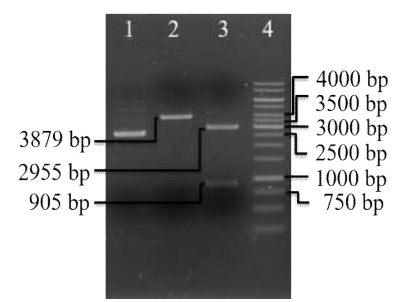
(Hình 2, giếng 2) tương ứng với kích thước dự đoán 905 bp của đoạn trình tự mã hóa *HbCPT4* (CDS *HbCPT4*) (Hình 2, giếng 3). Ngược lại, đối chứng âm không cho vạch có kích thước tương ứng (Hình 2, giếng 1). Đoạn CDS *HbCPT4* sau đó được tinh sạch từ gel agarose và được nối vào vector pJET1.2/blunt. Hỗn hợp phản ứng nối được biến nạp vào tế bào khả nạp *E.coli* TOP10.



Hình 2. Sản phẩm PCR nhân bản đoạn CDS *HbCPT4*. 1. Đối chứng âm, 2. Thang chuẩn DNA 1 kb, 3. Sản phẩm PCR.



Hình 3. Kết quả sàng lọc dòng tế bào *E. coli* TOP10/pJET-*HbCPT4* bằng PCR khuẩn lạc. 1. Thang chuẩn DNA 1kb, 2. Sản phẩm PCR khuẩn lạc, 3. Đối chứng âm.



Hình 4. Kết quả kiểm tra kích thước plasmid pJET-*HbCPT4*. 1. Plasmid pJET-*HbCPT4*, 2. Plasmid pJET-*HbCPT4* được cắt bởi *Bam*HI, 3. Plasmid pJET-*HbCPT4* được cắt bởi *Bam*HI và *Xho*I, 4. Thang chuẩn DNA 1 kb.

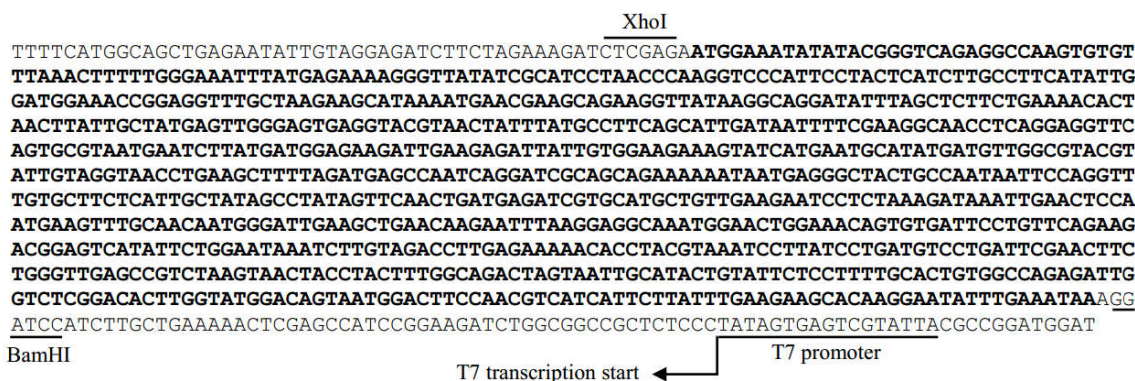
Kết quả sàng lọc các dòng tế bào *E.coli* TOP10 mang plasmid pJET-*HbCPT4* bằng phản ứng PCR khuẩn lạc với cặp mồi pJET1.2_F và *Xho*I-A-*HbCPT4*-F (Hình 3) cho thấy sản phẩm nhân bản ở phản ứng có sử dụng

DNA khuẩn lạc làm khuôn có một vạch tương ứng với kích thước dự đoán của sản phẩm là 967 bp (Hình 3, giếng 2). Trong khi đó, đối chứng âm không xuất hiện vạch có cùng kích thước như trên (Hình 3, giếng 3). Khuẩn lạc cho

kết quả PCR dương tính này được chọn nuôi nhân sinh khối để tách chiết plasmid pJET-HbCPT4. Plasmid pJET-HbCPT4 được cắt với *Bam*HI và *Xho*I, sản phẩm phản ứng cắt được điện di trên gel agarose 1%. Kết quả được mô tả ở hình 4 cho thấy ở giếng 2 xuất hiện 1 vạch nằm giữa hai vạch 3000 bp và 4000 bp khi so sánh với thang DNA chuẩn (Hình 4, giếng 4), phù hợp với kích thước dự đoán của plasmid pJET-HbCPT4 là 3879 bp. Ở giếng 3 của hình 4 xuất hiện 2 vạch: một vạch ứng với kích thước của CDS *HbCPT4* được chèn vào là 905 bp, vạch còn lại tương ứng với kích thước của vector pJET1.2/blunt là 2955 bp sau khi cắt bỏ đoạn DNA ở giữa vị trí nhận biết của hai enzyme cắt giới hạn *Bam*HI và *Xho*I.

3.3. Phân tích trình tự mã hóa *HbCPT4* được phân lập từ cây cao su *Hevea brasiliensis*

Plasmid tái tổ hợp pJET-HbCPT4 được giải trình tự với môi pJET1.2_F và kết quả giải trình tự được trình bày ở hình 5. Kết quả sắp giống cột trình tự CDS *HbCPT4* hiện diện trong plasmid pJET-HbCPT4 và trình tự CDS *HbCPT4* được dự đoán bằng phân tích tin sinh học cho thấy trình tự CDS *HbCPT4* dự đoán khác với trình tự CDS *HbCPT4* được phân lập ở 7 nucleotide có vị trí 250, 274, 420, 435, 519, 630 và 765 tính từ mã khởi đầu ATG. Tuy nhiên, hai trình tự protein được suy ra từ 2 trình tự này thì giống nhau 100%. Từ kết quả giải trình tự trên, chúng tôi đã hiệu chỉnh lại trình tự của CDS *HbCPT4*.



Hình 5. Kết quả giải trình tự plasmid pJET-HbCPT4 với môi pJET1.2_F. Đoạn trình tự được in đậm là trình tự mã hóa *HbCPT4* hoàn chỉnh được tạo dòng vào vector pJET1.2/blunt. Mã khởi đầu ATG, mã kết thúc TAA.

Theo kết quả dự đoán bởi phần mềm tính giá trị điểm đẳng điện pI và khối lượng phân tử MW dựa trên trình tự của chuỗi polypeptide, protein HbCPT4 có giá trị pI khoảng 8,19 và MW khoảng 34 kDa. Kết quả sắp giống cột trình tự protein HbCPT4 với các trình tự CPT đã được khảo sát chức năng ở các loài khác nhau như *Escherichia coli* (EcUPPS), *Saccharomyces cerevisiae* (ScRER2), *Arabidopsis thaliana* (AtCPT1), *Solanum lycopersicum* (SlCPT1 và SlCPT3) và *Hevea brasiliensis* (HRT1 và HRT2) (Hình 6) cho thấy HbCPT4 vẫn duy trì 5 vùng bảo tồn: (vùng I-V, bao gồm amino acid Asp41 (D41) ở vùng I,

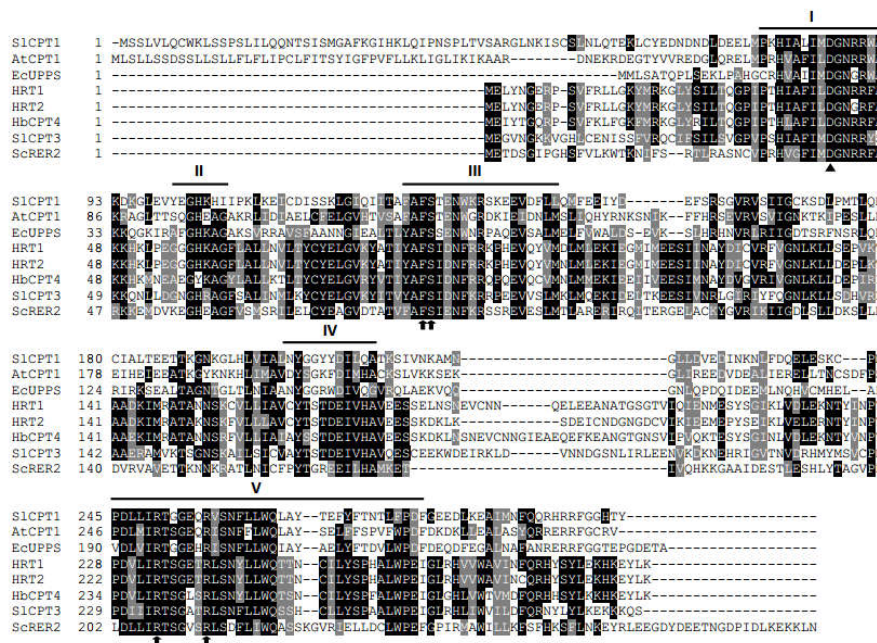
Phe85 (F85) và Ser86 (S86) ở vùng III, Arg239 (R239) và Arg245 (R245) ở vùng V) có vai trò quan trọng đối với hoạt tính xúc tác và khả năng liên kết cơ chất của các enzyme họ *cis*-prenyltransferase [2].

4. Kết luận

Kết hợp các phân tích *in silico* và kỹ thuật sinh học phân tử, chúng tôi đã phân lập được từ mô của cây cao su (*Hevea brasiliensis*) RRIV 209 một gene mã hóa chuỗi polypeptide tương đồng 75% với HRT2 một *cis*-prenyltransferase

tham gia vào quá trình sinh tổng hợp cao su thiên nhiên, và đặt tên là *HbCPT4*. Gene *HbCPT4* có 3 exon và 2 intron, và mã hóa cho protein được dự đoán có 296 amino acid với khối lượng phân tử khoảng 34 kDa, giá trị điểm

đẳng điện pI khoảng 8,19 và chứa 5 vùng bảo tồn (vùng I-V) có vai trò quan trọng đối với hoạt tính xúc tác cũng như khả năng liên kết cơ chất của các enzyme họ *cis*-prenyltransferase (CPT).



Hình 6. So sánh trình tự protein HbCPT4 với các trình tự *cis*-prenyltransferase đã được khảo sát chức năng ở *E. coli* (EcUPPS), *S. cerevisiae* (ScRER2), *A. thaliana* (AtCPT1), *S. lycopersicum* (S1CPT1 và S1CPT3) và *H. brasiliensis* (HRT1 và HRT2). Đường ngang màu đen là năm vùng bảo tồn (vùng I-V) ở các CPT. Hình tam giác tương ứng với vị trí aspartate bảo tồn có liên quan đến hoạt tính xúc tác. Mũi tên chỉ các amino acid liên quan đến khả năng liên kết cơ chất.

Tài liệu tham khảo

[1] Tholl D., Biosynthesis and biological functions of terpenoids in plants, *Advances in Biochemical Engineering/Biotechnology*, 148 (2015) 63.
 [2] Akhtar T.A., Matsuba Y., Schauvinhold I., Yu G., Lees H.A., Klein S.E., Pichersk E., The tomato *cis*-prenyltransferase gene family, *Plant Journal*, 73 (2013) 640.
 [3] Rahman A. Y. A., Usharraj A. O., Misra B. B., Thottathil G. P., Jayasekaran K., Feng Y., Hou S., Ong S. Y., Ng F. L., Lee L. S., Tan H. S., Sakaff M. K. L. M., Teh B. S., Khoo B. F., Badai S. S., Aziz N. A., Yuryev A., Knudsen B., Laporte A. D., Mchunu N. P., Yu Q., Langston B. J., Freitas T. A. K., Young A. G., Chen R., Wang L., Najimudin N., Saito J. A., Alam M., Draft genome sequence of the rubber tree *Hevea brasiliensis*, *BMC Genomics*, 14, 75 (2013) 1471.

[4] Asawateratanakul K., Zhang Y. W., Wititsuwannakul D., Wititsuwannakul R., Takahashi S., Rattanapittayaporn A., Koyama T., Molecular cloning, expression and characterization of cDNA encoding *cis*-prenyltransferases from *Hevea brasiliensis*, *European Journal of Biochemistry*, 270, 23 (2003) 4671.
 [5] Lau N. S., Makita Y., Kawashima M., Taylor T. D., Kondo S., Othman A. S., Chien A. C. S. C., Matsui M., The rubber tree genome shows expansion of gene family associated with rubber biosynthesis, *Scientific Report*, 6, 28594 (2016) 1.
 [6] Makita Y., Ng K. K., Singham G. V., Kawashima M., Hirakawa H., Sato S., Othman A. S., Matsui M., Large-scale collection of full-length cDNA and transcriptome analysis in *Hevea brasiliensis*, *DNA Research*, 24, 2 (2017) 159.

Isolation of Gene Encoding *cis*-prenyltransferase 4 (CPT4) from Rubber Tree (*Hevea brasiliensis*)

Nguyen Huynh Cam Tu¹, Pham Thi My Binh¹, Tran Thanh Luong¹,
Tran Thanh², Nguyen Thi Hong Thuong¹

¹Faculty of Biology and Biotechnology, VNU HCM University of Science, Vietnam

²Breeding Division, Rubber Research Institute of Vietnam

Abstract: *Cis*-prenyltransferase (CPT) catalyzes consecutive *cis*-condensation reactions of isopentenyl diphosphate (IPP) with allylic diphosphate acceptors to synthesize *cis*-prenyldiphosphates with a certain chain length. With the exception of HRT1 and HRT2 which are shown to be a part of rubber biosynthetic machinery, the biochemical functions of other CPTs remain unknown. By using *HRT2* as a query sequence for TBLASTX searches against the recently released transcriptome database of *Hevea brasiliensis*, we identified six contigs, one of which contains a sequence encoding for a protein that shares 75% identity to a previously reported *cis*-prenyltransferase, HRT2. This coding sequence, designated as *HbCPT4*, was successfully isolated from the latex of *Hevea brasiliensis* RRIV 209 plants and sequenced. A comparison of the full-length sequences of the *HbCPT4* CDS obtained by PCR and the *HbCPT4* CDS predicted by bioinformatics analysis revealed seven nucleotides that were different, but none of them leading to the changes in the amino acid sequence. The deduced protein sequence has 296 amino acid residues, a predicted mass of 34 kDa, and a pI of 8.19. It also contains five conserved regions (I-V regions) related to catalytic activity and substrate binding of *cis*-prenyltransferases.

Keywords: *Cis*-prenyltransferase, isoprenoid, *Hevea brasiliensis*, latex.