BỘ GIÁO DỤC VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Nguyễn Văn Tụng

NGHIÊN CỨU ĐỘT BIẾN VÀ ĐA HÌNH CỦA MỘT SỐ GEN Ở TRẺ EM MẮC BỆNH TEO ĐƯỜNG MẬT BẨM SINH

TÓM TẮT LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

NGÀNH: CÔNG NGHỆ SINH HỌC

Mã số: 9420201

Hà Nội - 2024

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

Người hướng dẫn khoa học:

1. Người hướng dẫn 1: GS.TS. Nguyễn Huy Hoàng – Viện Nghiên cứu hệ Gen – Viện Hàn lâm Khoa học và Công nghệ Việt Nam
2. Người hướng dẫn 2: TS. Nguyễn Thị Kim Liên – Viện Nghiên cứu hệ Gen – Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Phản biện 1: GS.TS. Trần Huy Thịnh

Phản biện 2: PGS.TS. Vũ Chí Dũng

Phản biện 3: GS.TS. Nguyễn Duy Bắc

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi 9 giờ 00, ngày 24 tháng 12 năm 2024

Có thể tìm hiểu luận án tại:

1. Thư viện Học viện Khoa học và Công nghệ

2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

Teo đường mật bẩm sinh (biliary atresia – BA) là bệnh lý đặc trưng bởi quá trình viêm, xơ hóa, phá hủy toàn bộ đường mật trong và ngoài gan dẫn đến sự cản trở lưu thông của mật. Sự khiếm khuyết của các gen liên quan đến sự hình thành ống mật và hệ thống miễn dich là một yếu tố gây ra bệnh teo đường mật bẩm sinh, do đó việc xác định các biến đổi di truyền liên quan đến bệnh sẽ là một dấu hiệu có thể hỗ trợ chẩn đoán trong thời gian ngắn, cải thiện kết quả phẫu thuật.

Ở Việt Nam, các nghiên cứu sâu về di truyền các gen có liên quan còn khá hạn chế. Xuất phát từ những thực tiễn đó, đề tài "*Nghiên cứu đột biến và đa hình của một số gen ở trẻ em mắc bệnh teo đường mật bẩm sinh"* được thực hiện nhằm xác định các biến đổi di truyền trên các gen liên quan đến bệnh teo đường mật bẩm sinh. Các biến đổi trên các gen liên quan đến bệnh bao gồm cả đột biến và đa hình đơn nucleotide sẽ được kiểm chứng lại trên số lượng lớn bệnh nhân nhằm tìm ra mối liên hệ kiểu gen và kiểu hình.

Mục tiêu nghiên cứu

- Giải trình tự toàn bộ vùng gen mã hóa của một số bệnh nhân mắc teo đường mật bẩm sinh từ đó xác định được các đột biến gen liên quan đến bệnh teo đường mật bẩm sinh.

- Phân tích mối liên quan giữa các đa hình nucleotide đơn với nguy cơ mắc bệnh teo đường mật bẩm sinh ở người Việt Nam

Nội dung nghiên cứu

- Thu thập mẫu máu và tách chiết DNA tổng số từ mẫu máu của bệnh nhân teo đường mật bẩm sinh và các mẫu đối chứng.

- Giải trình tự toàn bộ vùng mã hóa của một số bệnh nhân mắc bệnh teo đường mật bẩm sinh điển hình.

- Phân tích số liệu, sàng lọc và tìm đột biến trên các gen liên quan đến bệnh teo đường mật bẩm sinh ở các bệnh nhân được giải trình tự toàn bộ vùng mã hóa.

- Phân tích mối liên quan giữa các đa hình nucleotide đơn với nguy cơ mắc bệnh teo đường mật bẩm sinh trên số lượng lớn bệnh nhân và đối chứng.

Những đóng góp mới của đề tài

1. Đã phát hiện 02 đột biến mới dị hợp tử kép c.412G>A (p.Gly138Arg) trên gen *FAH* và c.2225A>G (p.Tyr742Cys) trên gen *ERCC4* ở hai bệnh nhân là chị em ruột trong một gia đình. Phát hiện 01 đột biến dịch khung c.50_51insG (p.Gly17Glyfs77*) trên gen *KRT18* và 02 đột biến c.314C>A (p.Ser105∗), c.2975C>T (p.Pro992Leu) đều ở dạng dị hợp tử trên gen *ATP7B* ở 01 bệnh nhân mắc đồng thời hai bệnh teo đường mật bẩm sinh và Wilson, trong đó đột biến c.50_51insG (p.Gly17Glyfs77*) trên gen *KRT18* có tiềm năng liên quan đến bệnh teo đường mật bẩm sinh.

2. Đánh giá mối liên quan giữa các đa hình đơn nucleotide rs2287622 trên gen *ABCB11*, đa hình rs927344 trên gen *ABCC2*, đa hình rs1815930 trên gen *MYO5B* với nguy cơ mắc bệnh teo đường mật bẩm sinh ở bệnh nhân người Việt Nam.

Chương 1. TỔNG QUAN

1.1. Bệnh teo đường mật bẩm sinh

Teo đường mật bẩm sinh là bệnh lý hiếm gặp của gan và đường mật, xảy ra trong quá trình phát triển tạo ống của đường mật trong thời kỳ tạo phôi. Bệnh khởi phát ở thời kỳ sơ sinh, được đặc trưng bởi sự khiếm khuyết của hệ thống ống mật, dẫn đến tắc nghẽn dòng chảy mật, gây ứ mật. Tần xuất mắc bệnh khác nhau theo từng khu vực trên thế giới, thậm chí giữa các khu vực trong cùng một đất nước. Bệnh teo đường mật bẩm sinh xuất hiện ở châu Á và châu Phi nhiều hơn châu Âu, ở nữ nhiều hơn ở nam.

1.1.1. Triệu chứng lâm sàng và phân loại bệnh

Triệu chứng lâm sàng chủ yếu của bệnh teo đường mật bẩm sinh là tăng bilirubin, vàng da, phân bạc màu và gan lớn. Trẻ mắc bệnh có thể có vẻ ngoài bình thường khi mới sinh. Về biểu hiện lâm sàng, thường khó phân biệt trẻ bị teo đường mật bẩm sinh với trẻ bị vàng da sinh lý hoặc một dạng ứ mật khác.

Về mặt lâm sàng, trẻ sơ sinh bị teo đường mật bẩm sinh thường có gan to và lách to khi trẻ được 4-8 tuần tuổi và các xét nghiệm thường cho thấy tăng ALT, AST, bilirubin toàn phần và bilirubin trực tiếp, trong khi GGT (Gamma Glutamyl transferase) và nồng độ phosphatase kiềm thường tăng lên đến mức trên 1000 U/dl.

Theo giải phẫu bệnh dựa trên ảnh chụp đường mật cản quang, bệnh teo đường mật bẩm sinh có thể chia thành 3 dạng chính:

- Type I, bệnh nhân bị teo ở phần ống mật chủ (chiếm 5%).
- Type II là các bệnh nhân bị teo lên đến ống gan chung, có thể kết hợp với nang đường mật (chiếm 2%).
- Type III là các bệnh nhân teo đường mật ngoài gan và ống mật đến vùng rốn gan (chiếm tới 90%).

Dựa theo các triệu chứng lâm sàng, bệnh teo đường mật bẩm sinh có thể chia thành ba loại như sau:

- Teo mật bẩm sinh không kèm các dị tật bất thường khác: chiếm khoảng 70 - 85% trường hợp.
- Teo mật kết hợp với các dị dạng khác (hội chứng BASM – biliary atresia splenic malformation): chiếm khoảng 10 – 15% trường hợp.
- Teo mật có nang: teo đường mật ngoài gan đặc trưng bởi sự teo hoặc vắng mặt do viêm gây hủy hoại đường mật.

1.1.2. Nguyên nhân gây bệnh teo đường mật bẩm sinh

Hầu hết các nhà nghiên cứu làm việc trong lĩnh vực này đều tin rằng teo đường mật bẩm sinh không phải là căn bệnh có một nguyên nhân duy nhất mà là sự kết hợp của nhiều kiểu hình khác nhau có chung các đặc điểm lâm sàng nhất định. Teo đường mật bẩm sinh có thể là một bệnh đa yếu tố, trong đó sự tương tác giữa môi trường và di truyền là cơ sở sinh bệnh học của căn bệnh này. Một số giả thuyết về nguyên nhân gây bệnh được đưa ra như: các rối loạn về gen di truyền, cơ chế tự miễn, các bất thường về miễn dịch, do nhiễm độc và nhiễm virus.

1.2. Vai trò của yếu tố di truyền trong cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh

Cơ sở di truyền của bệnh teo đường mật bẩm sinh rất phức tạp. Người ta thấy rằng bệnh có thể di truyền theo kiểu trội hoặc kiểu lặn nhưng nhiều khả năng là do tình trạng đa gen, tính không đồng nhất về gen và các biểu hiện lâm sàng khác nhau [54]. Ngày càng có nhiều bằng chứng cho thấy sự liên quan giữa bệnh teo đường mật bẩm sinh với các yếu tố di truyền đặc biệt là đột biến xảy ra trong những gen quy định sự phát triển của ống mật. Nghiên cứu sinh bệnh học của teo đường mật bẩm sinh cho thấy các khuyết tật di truyền phôi, các bất thường tiền sản, bất thường thai nhi do di truyền, các yếu tố nhạy cảm với bệnh đóng vai trò nhất định [55].

Thông thường, mỗi gen không chỉ đóng một vai trò trong sinh bệnh học của bệnh teo đường mật bẩm sinh vì phần lớn chúng góp phần vào nhiều quá trình sinh học khác nhau. Dựa trên các đặc điểm bệnh lý của teo đường mật bẩm sinh và chức năng của gen, các biến thể di truyền có tiềm năng liên quan đến teo đường mật bẩm sinh có thể được phân loại thành 4 con đường: sự phát triển gan mật, quá trình xơ hóa, viêm, sự bất thường vi lông mao (bệnh lý tiêm mao). Việc xác định và phân tích biểu hiện của gen liên tham gia vào các quá trình sinh bệnh học khác nhau của teo đường mật bẩm sinh

sẽ cung cấp những hiểu biết rõ ràng hơn về cơ chế bệnh sinh của căn bệnh này.

Thông thường, mỗi gen không chỉ đóng một vai trò trong sinh bệnh học của bệnh teo đường mật bẩm sinh vì phần lớn chúng góp phần vào nhiều quá trình sinh học khác nhau. Dựa trên các đặc điểm bệnh lý của teo đường mật bẩm sinh và chức năng của gen, các biến thể di truyền có tiềm năng liên quan đến teo đường mật bẩm sinh có thể được phân loại thành 4 con đường: sự phát triển gan mật, quá trình xơ hóa, viêm, sự bất thường vi lông mao (bệnh lý tiêm mao). Việc xác định và phân tích biểu hiện của gen liên tham gia vào các quá trình sinh bệnh học khác nhau của teo đường mật bẩm sinh sẽ cung cấp những hiểu biết rõ ràng hơn về cơ chế bệnh sinh của căn bệnh này.

1.3. Tình hình nghiên cứu bệnh teo đường mật bẩm sinh

1.3.1. Tình hình nghiên cứu trên thế giới

Teo đường mật bẩm sinh được biết đến là một bệnh di truyền có tính không đồng nhất cao, gây nhiều khó khăn trong chẩn đoán và điều trị. Việc không được chẩn đoán chính xác và điều trị sớm có thể dẫn đến kết quả điều trị không thành công hoặc dẫn đến sự tử vong ở trẻ. Vì vậy, đã có nhiều nghiên cứu được tiến hành trên thế giới nhằm xác định chính xác nguyên nhân di truyền của bệnh để có định hướng điều trị và tư vấn di truyền cho gia đình bệnh nhân.

1.3.2. Tình hình nghiên cứu ở Việt Nam

Việt Nam, tỷ lệ mắc bệnh khá cao ước tính 1:2400 trẻ đẻ ra sống [72]. Khó khăn ở Việt Nam là hiện nay các nghiên cứu sâu về di truyền các gen có liên quan còn khá hạn chế. Cho đến nay, mới chỉ có các nghiên cứu về lâm sàng trên các bệnh nhân mắc teo đường mật bẩm sinh được điều trị tại các thành phố lớn là Hà Nội và thành phố Hồ Chí Minh.

Chương 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Đối tượng nghiên cứu

Các mẫu máu nghiên cứu là từ các bệnh nhân bị mắc bệnh teo đường mật bẩm sinh và người không bị mắc bệnh làm đối chứng (266 bệnh nhân và 250 mẫu đối chứng).

2.2. Hóa chất và trang thiết bị

Các hóa chất, máy móc, thiết bị phụ vụ cho nghiên cứu thuộc Viện Nghiên cứu hệ Gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

2.3. Phương pháp nghiên cứu

2.3.1. Tách chiết DNA tổng số

Trong nghiên cứu này, DNA tổng số của các đối tượng tham gia nghiên cứu được tách chiết bằng bộ kit DNAamp Blood Mini của hãng QIAGEN.

2.3.2. Giải trình tự toàn bộ vùng gen mã hóa

Phương pháp giải trình tự vùng gen mã hóa gồm 2 bước chính: tạo thư viện DNA và giải trình tự bằng máy Illumina.

2.3.2.1. Tạo thư viện DNA

Thư viện DNA được thiết lập bằng bộ kit Agilent SureSelect Target Enrichment.

2.3.2.2. Giải trình tự bằng Illumina NextSeq 500

Thư viện đạt chất lượng được giải trình tự bằng thiết bị Illumina NextSeq 500 tại Viện Nghiên cứu hệ gen.

2.3.3. Xác định và chú giải biến thể

Sau khi giải trình tự trên máy Illumina, chất lượng trình tự được kiểm tra bằng phần mềm FastQC. Dữ liệu trình tự được sắp xếp và so sánh với ngân hàng gen người (GRCh37) bằng phần mềm BWA 0.7.10. Công cụ Picard được sử dụng để xử lý dữ liệu sau khi sắp hàng. Biến thể được phát hiện bằng phần mềm GATK v3.4. Phần mềm SnpEff được sử dụng để đánh giá ảnh hưởng của biến thể đến trình tự và chức năng của protein.

2.3.4. Sàng lọc biến thể

Các biến thể được sàng lọc để tìm ra đột biến gen có tiềm năng gây bệnh. Biến thể có tần số alen MAF > 1% được coi là SNP và được đánh giá trên số lượng lớn bệnh nhân và đối chứng để tìm ra mối liên quan đến nguy cơ mắc bệnh.

2.3.5. Xác định kiểu gen của đa hình rs2287622 bằng kỹ thuật RFLP

Đoạn gen *ABCB11* kích thước 333 bp có chứa đa hình rs2287622 được khuếch đại bằng phản ứng PCR sử dụng cặp mồi đặc hiệu. Enzyme *Hae*III được sử dụng để xác định kiểu gen tại vị trí chứa SNP.

2.3.6. Xác định kiểu gen của đa hình bằng kỹ thuật ARMS-PCR

Trong nghiên cứu này, kiểu gen của các SNP rs927344 trên gen *ABCC2* và rs1815930 trên gen *MYO5B* được xác định bằng phương pháp ARMS-PCR.

2.3.7. Phân tích thống kê

Kiểm định Chi-Square (χ^2) được sử dụng kiểm tra trạng thái cân bằng HardyWeinberg (HWE) của quần thể. Kiểm định được coi là có ý nghĩa khi giá trị p>0,05. Đồng thời, kiểm định Chi-Squares cũng được sử dụng để đánh giá mối liên quan giữa kiểu gen và kiểu alen của ba đa hình nucleotide đơn *ABCC2* rs927344 *và MYO5B* rs1815930 *ABCB11* rs2287622 với khả năng mắc bệnh teo đường mật bẩm sinhtrên 3 mô hình cộng gộp (additive model), trội (dominant model), lặn (recessive model) và được ước tính bằng chỉ số OR (odds ratio) với khoảng tin cậy 95%. Các kiểm định được coi là có ý nghĩa thống kê khi giá trị p<0,05.

2.3.8. Giải trình tự Sanger

Trình tự các đoạn gen chứa đột biến được xác định bằng máy giải trình tự ABI3500 sử dụng bộ kit BigDye® Terminator v3.1 Cycle Sequencing.

Chương 3. KẾT QUẢ

3.1. Thu thập mẫu và tách chiết DNA tổng số

Mẫu máu của các đối tượng tham gia nghiên cứu được thu thập bởi các bác sĩ tại Bệnh viện Nhi Trung ương. DNA tổng số của người tham gia nghiên cứu được kiểm tra bằng điện di trên gel agarose. Kết quả cho thấy DNA không bị đứt gãy, có nồng độ và độ tinh sạch cao. Trong số 266 bệnh nhân tham gia nghiên cứu, có 05 bệnh nhân được lựa chọn để giải trình tự toàn bộ vùng gen mã hóa.

3.2. Giải trình tự toàn bộ vùng mã hóa

3.2.1. Kết quả tạo thư viện DNA

Thư viện DNA của 05 bệnh nhân được kiểm tra chất lượng bằng máy Agilent Technologies 2100 Bioanalyzer. Các thư viện có chất lượng tốt, nồng độ cao, đạt yêu cầu cho bước nghiên cứu tiếp theo với phần lớn đoạn DNA sau khi được phân mảnh có kích thước nằm trong khoảng từ 250 – 500 bp.

3.2.2. Kết quả giải trình tự toàn bộ vùng gen mã hóa

Dữ liệu được tạo ra từ máy giải trình tự thế hệ mới có dung lượng cao, chất lượng tốt, tỉ lệ nucleotide có điểm chất lượng lớn hơn 20 và 30 cao (Q20 trên 96%, Q30 trên 91%). Thống kê chất lượng giải trình tự của bệnh nhân mắc bệnh được thể hiện rõ hơn ở bảng 3.4.

Bảng 3.4. Thông tin đọc trình tự toàn bộ vùng mã hóa của 05 bệnh nhân teo đường mật bẩm sinh

	BA01	**BA02**	**BA03**	**BA04**	**BA05**
Tổng số nucleotide	9.933 Mb	9.247 Mb	9.693 Mb	10.973 Mb	10.797 Mb
Tổng số đoạn đọc	66.561.946	62.004.790	64.870.694	73.536.620	72.227.972

	BA01	BA02	BA03	BA04	BA05
%GC	51,9	51,8	52,1	52,3	50,7
Q20(%)	96,7	96,8	96,7	97,0	98,1
Q30(%)	91,5	91,7	91,5	92,2	
Độ dài đoạn đọc trung bình (bp)	149,24	149,15	149,43	149,22	148,77
Độ sâu bao phủ trung bình (X)	164,3	152,9	160,3	181,5	157,6

3.3. Kết quả xử lý số liệu và chú giải biến thể

3.3.1. Sắp hàng dữ liệu với hệ gen tham chiếu

Trình tự toàn bộ vùng gen mã hóa của 05 bệnh nhân teo đường mật bẩm sinh được sắp hàng với hệ gen tham chiếu GRCh37 bằng công cụ BWA. Có thể nhận thấy, kết quả sắp hàng dữ liệu giải trình tự của bệnh nhân với hệ gen tham chiếu tốt (tỉ lệ sắp hàng thành công đạt 99,8-99,9%), tỉ lệ đoạn dữ liệu được sắp hàng nằm trong vùng exone cao (70-79%).

3.3.2. Xác định và chú giải biến thể

Từ kết quả so sánh, quy trình xử lý số liệu sử dụng công cụ Picard và GATK được sử dụng để phát hiện biến thể trong nghiên cứu này. Ngoài ra, ảnh hưởng của biến thể đến trình tự axit amin được phân tích bằng phần mềm SnpEff. Kết quả phân tích được thể hiện ở bảng 3.6.

Bảng 3.6: Kết quả phân tích và dự đoán biến thể trong toàn bộ vùng mã hóa của 05 bệnh nhân mắc bệnh teo đường mật bẩm sinh

Thông số	Mẫu BA01	Mẫu BA02	Mẫu BA03	Mẫu BA04	Mẫu BA05
Tổng số biến thể thay thế	**23.295**	**23.743**	**23.413**	**23.508**	**23.774**
Biến thể đồng nghĩa	11.870	12.033	11.958	12.012	12.036
Biến thể sai nghĩa	11.271	11.559	11.317	11.368	11.579
Thêm bộ ba kết thúc	114	112	99	89	119
Mất bộ ba kết thúc	40	39	39	39	40
Tổng số indel	**724**	**711**	**738**	**776**	**712**
Biến thể frameshift	325	318	348	350	317
Thêm đoạn ngắn	182	174	175	196	187
Mất đoạn ngắn	217	219	215	230	208

3.4. Sàng lọc biến thể

Sau quá trình sàng lọc, những biến thể nằm trong vùng exone của các gen liên quan đến bệnh teo đường mật bẩm sinh được giữ lại. Danh sách những gen được sử dụng để sàng lọc đột biến được trình bày ở phụ lục 2. Những biến thể này được chia làm hai loại: (i) đột biến – là những biến thể có tần số alen ≤1%; (ii) đa hình nucleotide đơn – là những biến thể có tần số

alen >1% (tần số alen MAF được tính theo tần số suất hiện trong cơ sở dữ liệu 1000 Genome phase 3).

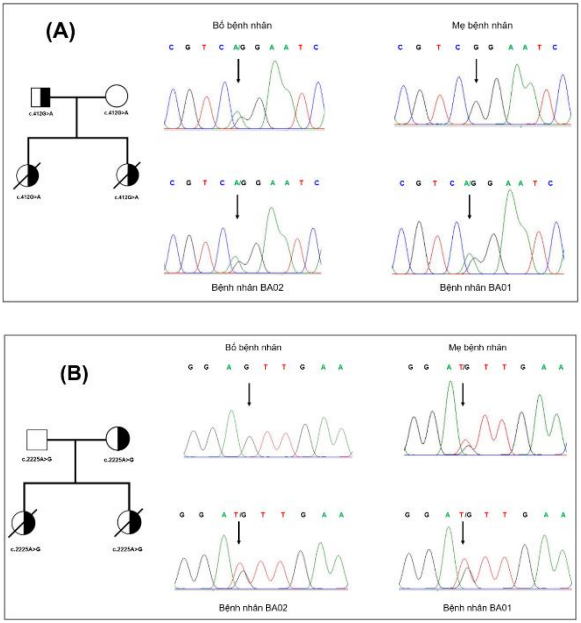
Số lượng biến thể xảy ra trong số 133 gen liên quan đến bệnh teo đường mật bẩm sinh là rất lớn. Tuy nhiên, phần lớn các đột biến này nằm trong vùng intron của gen, có rất ít đột biến nằm trong vùng exone (số lượng đột biến trong vùng exone ở bệnh nhân BA01 là 14, ở bệnh nhân BA02 là 12, ở bệnh nhân BA03 là 12, ở bệnh nhân BA04 là 08 và ở bệnh nhân BA05 là 10 đột biến).

3.5. Đánh giá ảnh hưởng của các đột biến

Các đột biến tiềm năng được đánh giá là có ảnh hưởng đến bệnh teo đường mật bẩm sinh nếu cả bốn phần mềm SIFT, PolyPhen2, Mutation Taster, Provean đánh giá là gây bệnh. Như vậy, kết quả phân tích bằng các phần mềm tin sinh học cho thấy bệnh nhân BA01 và BA02 mang 02 đột biến dạng dị hợp tử có khả năng gây bệnh là c.412G>A trên gen *FAH* và c.2225A>G trên gen *ERCC4*; bệnh nhân BA03 và BA04 mang 01 đột biến dạng dị hợp tử có khả năng gây bệnh c.368A>C trên gen *CYP27A1*; bệnh nhân BA05 mang 01 đột biến dạng đồng hợp tử c.50_51insG (p.Gly17Glyfs77 ∗) trên gen *KRT18* có tiềm năng gây bệnh.

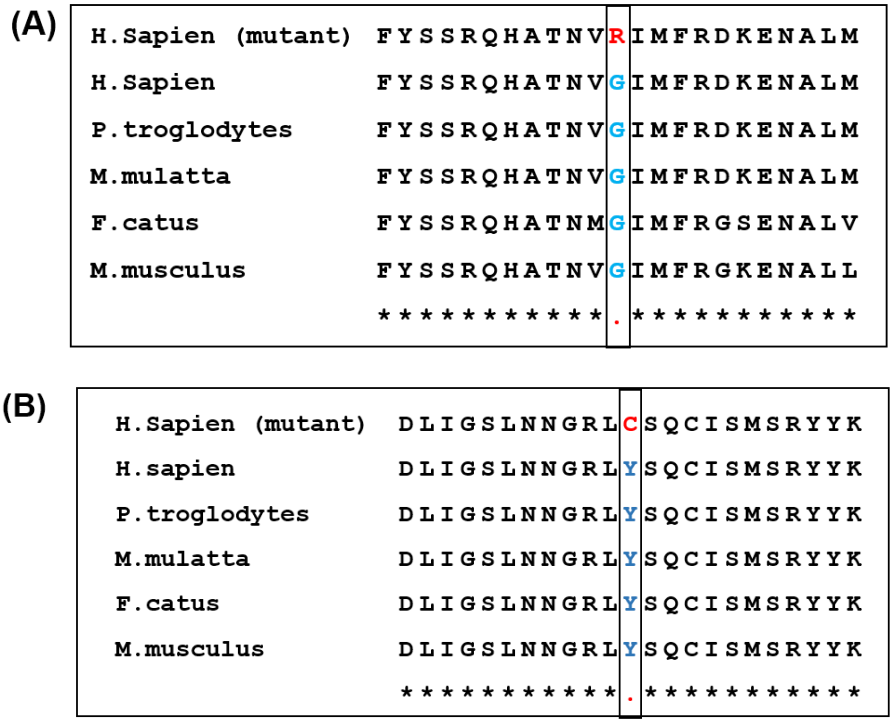
3.5.1. Bệnh nhân BA01 và BA02

Hai bệnh nhân BA01 và BA02 là chị em ruột trong một gia đình. Hai chị em mang đột biến dị hợp tử kép (double heterozygous) trên hai gen *FAH* và *ERCC4*. Trong đó, cả hai bệnh nhân di truyền đột biến c.412G>A (p.Gly138Arg) trên gen *FAH* từ người bố và đột biến c.2225A>G (p.Tyr742Cys) trên gen *ERCC4* từ người mẹ (Hình 3.3).



Hình 3.3: Đột biến dị hợp tử kép trên bệnh nhân BA01 và BA02
(A) Sơ đồ phả hệ đột biến c.412G>A trên gen FAH; (B) Sơ đồ phả hệ đột biến c.2225A>G trên gen ERCC4

Ngoài ra, kết quả so sánh trình tự axit amin xung quanh vị trí đột biến giữa các loài bao gồm thể đột biến, loài người (*H.sapien*), tinh tinh (*P.troglodytes*), khỉ vàng (*M.mulatta*), mèo (*F.catus*) và chuột nhắt (*M.musculus*) cho thấy Gly138 trên protein mã hóa bởi gen *FAH* và Tyr742 trên protein mã hóa bởi gen *ERCC4* là những axit amin có tính bảo thủ cao giữa các loài khác nhau (Hình 3.4).



Hình 3.4:So sánh trình tự axit amin xung quanh vị trí đột biến giữa các loài (A) Vị trí đột biến p.Gly138Arg trên protein FAH; (B) Vị trí đột biến p.Tyr742Cys trên gen ERCC4

3.5.2. Bệnh nhân BA03 và BA04

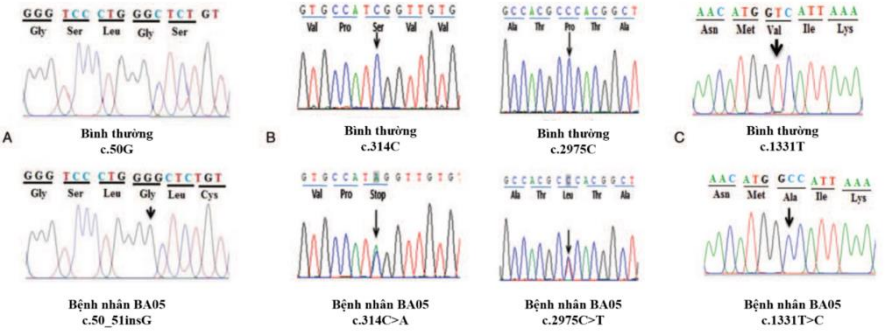
Kết quả phân tích dữ liệu WES cho thấy hai bệnh nhân BA03 và BA04 mang đột biến c.368A>C dạng dị hợp tử trên gen *CYP27A1*. Tuy nhiên, kết quả giải trình tự bằng Sanger không phát hiện đột biến này. Do đó, chúng tôi không xác định được đột biến gây bệnh trên hai bệnh nhân BA03 và BA04.

3.5.3. Bệnh nhân BA05

Kết quả phân tích cho thấy bệnh nhân BA05 mang đột biến ở dạng đồng hợp tử c.50_51insG (p.Gly17Glyfs77 ∗) trên gen *KRT18* (Hình 3.77).

Việc thêm một nucleotide Guanine (G) vào vị trí 51 dẫn đến hiện tượng dịch khung, làm protein bị cắt ngắn sau vị trí đột biến 77 axit amin.

Bệnh nhân BA05 đồng thời được chẩn đoán mắc bệnh Wilson, do đó, chúng tôi tiến hành mở rộng phạm vi sàng lọc đột biến. Kết quả phát hiện bệnh nhân mang hai đột biến c.314C>A (p.Ser105∗) và c.2975C>T (p.Pro992Leu) đều ở dạng dị hợp tử trên gen *ATP7B*. Hai đột biến này đều được ghi nhận là gây bệnh trên cơ sở dữ liệu ClinVar và được báo cáo trên ngân hàng dữ liệu dbSNP với mã số lần lượt là rs753236073 và rs201038679.



Hình 3.7: Đột biến trên bệnh nhân BA05

(A) đột biến đồng hợp tử c.50_51insG trên gen KRT18; (B) đột biến dị hợp tử phức c.314C>A và c.2975C>T trên gen ATP7B; (C) đột biến đồng hợp tử c.1331T>C trên gen ABCB11

Ngoài ra, bệnh nhân BA05 mang biến thể c.1331T>C (p.Val444Ala) trên gen *ABCB11* ở dạng đồng hợp tử. Đây là điểm đa hình đơn nucleotide được ghi nhận trên cơ sở dữ liệu dbSNP với mã số rs2287622. Ảnh hưởng của SNP này đến độ mẫn cảm bệnh teo đường mật bẩm sinh sẽ được đánh giá trong nghiên cứu này.

3.6. Ảnh hưởng của các đa hình nucleotide đến mức độ mẫn cảm bệnh teo đường mật bẩm sinh

3.6.1. Ảnh hưởng của đa hình rs2287622

Trong nghiên cứu này, kiểu gen của đa hình rs2287622 được xác định bằng kỹ thuật RFLP. Thống kê kiểu gen và tần suất alen của đa hình rs2287622 được thể hiện trong bảng 3.11

Bảng 3.11: Thống kê kiểu gen và tần suất alen đa hình ABCB11 rs2287622

	Kiểu alen			Tần số alen (%)		HWE p-value	HWE
	TT	TC	CC	T	C		
Nhóm bệnh	56	130	80	45,5	54,5	0,81	+
Nhóm đối chứng	89	115	46	58,6	41,4	0.138	+
Tổng số	145	245	126	51,84	48,16	0,06	+

Kiểu alen ở cả nhóm bệnh và nhóm đối chứng đều tuân theo cân bằng di truyền Hardy – Weinberg (p>0,05) (bảng 12), đồng thời, tần số alen C ở nhóm bệnh (54,5%) cao hơn nhiều so với tần số alen C ở nhóm đối chứng (33%).

Sự liên quan giữa alen và khả năng mắc bệnh teo đường mật được đánh giá bằng phép kiểm nghiệm Chi-square trên cả ba mô hình: cộng gộp, trội và mô hình lặn (Bảng 3.12).

Bảng 3.12: Kết quả phân tích sự liên quan giữa ABCB11 rs2287622 và nguy cơ mắc bệnh teo đường mật bẩm sinh

Kiểu gen	Nhóm bệnh	Nhóm đối chứng	OR	95% CI	Giá trị P
Mô hình cộng gộp					
TT	56 (21,05%)	89 (35,60%)	1		
TC	130 (48,87%)	115 (46,00%)	1,79	1,18-2,73	0,006

CC	80 (30,08%)	46 (18,40%)	2,76	1,69-4,53	0,0001
Mô hình trội					
TT	56 (21,05%)	89 (35,60%)	1		
TC + CC	210 (78,95%)	161 (64,40%)	13,13	8,44-20,39	0,0001
Mô hình lặn					
TT + TC	186 (69,92%)	204 (81,6%)	1		
CC	80 (30,08%)	46 (18,40%)	1,91	1,26-2,88	0,002
Alen					
T	242 (45,49%)	293 (58,60%)	1		
C	290 (54,51%)	207 (41,40%)	1,69	1,33-2,17	0,0001

Ghi chú: OR: Tỉ số odds ratio; 95%CI: Khoảng tin cậy 95%.

Có sự khác biệt đáng kể giữa tần số kiểu gen CC ở nhóm bệnh và nhóm đối chứng. Sự khác nhau này có liên quan đến việc làm tăng nguy cơ mắc bệnh teo đường mật so với kiểu gen TT (OR = 5,14, 95% CI: 2,82 – 9,39, p<0,01). Sự xuất hiện kiểu gen TC cũng làm tăng nguy cơ mắc bệnh teo đường mật, nhưng ít hơn so với kiểu gen CC (OR = 2,88, 95% CI: 1,81 - 4,59, p<0,01).

Ở mô hình trội, kiểu gen TC và CC làm tăng nguy cơ mắc bệnh teo đường mật bẩm sinh so với kiểu gen TT (OR = 3,46 95% CI: 2,24 – 5,35, p<0,01). Tương tự như vậy, ở mô hình lặn, kiểu gen CC cũng làm tăng nguy cơ mắc bệnh khi so với kiểu gen TT+TC (OR = 2,79, 95% CI: 1,63 – 4,79, p<0,01). Xét tổng thể, sự xuất hiện của alen C làm tăng nguy cơ mắc bệnh teo đường mật bẩm (OR = 2,47, 95% CI: 1,84 – 3,32, p<0,01) so với alen T.

3.6.2. Ảnh hưởng của đa hình rs927344

Trong nghiên cứu này, kiểu gen của SNP rs927344 trên gen *ABCC2* được xác định bằng phương pháp ARMS-PCR. Kiểm định Chi bình phương cho thấy sự phân bố của đa hình này được xác định tuân theo định luật cân bằng di truyền HardyWeinberg với giá trị p>0,05 (bảng 3.13).

Bảng 3.13: Thống kê kiểu gen và tần suất alen đa hình ABCC2 rs927344

	Kiểu alen			Tần suất alen		HWE	HWE
	AA	AT	TT	A	T	giá trị p	
Nhóm bệnh	5	40	221	0,09	0,91	0.451	+
Nhóm đối chứng	3	39	208	0,12	0,88	0,587	+
Tổng số	8	79	429	0,10	0,90	0,056	+

Ghi chú: HWE: Cân bằng di truyền HardyWeinberg; "+": Tuân theo cân bằng di truyền HardyWeinberg

Sự liên quan giữa alen và khả năng mắc bệnh teo đường mật bẩm sinh được đánh giá bằng phép kiểm nghiệm Chi-square (bảng 3.14). Kết quả cho thấy ở cả ba mô hình tỉ lệ kiểu gen và kiểu alen ở đa hình rs927344 không liên quan đến nguy cơ mắc bệnh ở quần thể nghiên cứu.

Bảng 3.14: Kết quả phân tích sự liên quan giữa ABCC2 rs927344 và nguy cơ mắc bệnh teo đường mật bẩm sinh

Kiểu gen	Nhóm bệnh	Nhóm đối chứng	OR	95% CI	Giá trị P
Mô hình cộng gộp					
AA	5 (1,88%)	3(1,20%)	1		
AT	40 (15,04%)	39(15,60%)	0,61	0,14 – 2,75	0,53
TT	221 (83,08%)	208(83,20%)	0,64	0,15 – 2,70	0,54

Mô hình trội					
AA	5 (1,88%)	3 (1,20%)	1		
AT + TT	261 (98,12%)	247 (98,80%)	0,63	0,15 - 2,68	0,54
Mô hình lặn					
AA +AT	45 (16,92%)	42 (16,80%)	1		
TT	221 (83,08%)	208 (83,20%)	0,99	0,63 - 1,57	0,97
Alen					
A	50 (9,4%)	45 (9,00%)	1		
T	482 (90,6%)	455 (91,00%)	0,95	0,63 - 1,45	0,82

Ghi chú: OR: Tỉ số odds ratio; 95%CI: Khoảng tin cậy 95%.

Kết quả phân tích cho thấy ở cả ba mô hình, giá trị p đều lớn hơn 0,05. Do đó, các phép kiểm định không có ý nghĩa thống kê. Như vậy, tỉ lệ kiểu gen và kiểu alen ở đa hình rs927344 không liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh ở quần thể nghiên cứu.

3.6.3. Ảnh hưởng của đa hình rs1815930

Thống kê kiểu gen và tần suất alen của đa hình rs1815930 ở tất cả các mẫu được thể hiện trong bảng 3.15.

Bảng 3.15: Thống kê kiểu gen và tần suất alen đa hình MYO5B rs1815930

	Kiểu alen			**Tần suất alen**		**HWE**	**HWE**
	AA	**AG**	**GG**	**A**	**G**	**giá trị p**	
Nhóm bệnh	2	39	225	0,08	0,92	0.829	+
Nhóm đối chứng	4	33	213	0,08	0,92	0,051	+
Tổng số	6	72	438	0,08	0,92	0,128	+

Ghi chú: HWE: Cân bằng di truyền HardyWeinberg; "+": Tuân theo cân bằng di truyền HardyWeinberg

Sự tương quan giữa alen và khả năng mắc bệnh teo đường mật bẩm sinh được đánh giá bằng phép kiểm nghiệm Chi bình phương trên cả ba mô hình: cộng gộp, trội và mô hình lặn (bảng 3.16).

Bảng 3.16: Kết quả phân tích sự liên quan giữa MYO5B rs1815930 và nguy cơ mắc bệnh teo đường mật bẩm sinh

Kiểu gen	Nhóm bệnh	Nhóm đối chứng	OR	95% CI	Giá trị P
Mô hình cộng gộp					
AA	2 (0,75%)	4 (1,60%)	1		
AG	39 (14,66%)	33 (13,20%)	2,36	0,41 – 13,73	0,33
GG	225 (84,59%)	213 (85,20%)	2,11	0,38 - 11,65	0,39
Mô hình trội					
AA	2 (0,75%)	4 (1,60%)	1		
AG + GG	264 (99,25%)	246 (98,40%)	2,15	0,39 - 11,82	0,38
Mô hình lặn					
AA + AG	41 (15,41%)	37 (14,80%)	1		
GG	225 (84,59%)	213 (85.20%)	0,95	0,59 - 1,54	0,85
Alen					
A	43 (8,08%)	41 (8,20%)	1		
G	489 (91,92%)	259 (91,80%)	1,02	0,65 - 1,59	0,95

Ghi chú: OR: Tỉ số odds ratio; 95%CI: Khoảng tin cậy 95%.

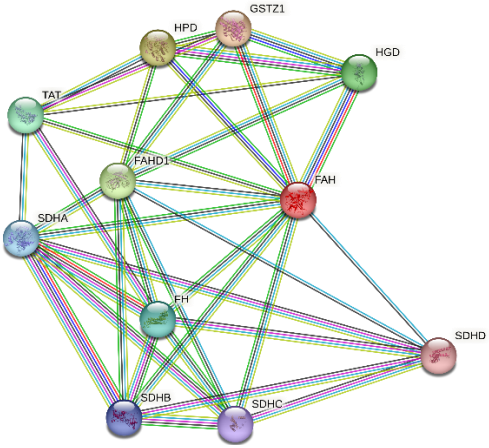
Kết quả phân tích cho thấy ở cả ba mô hình, giá trị p đều lớn hơn 0,05. Do đó, các phép kiểm định không có ý nghĩa thống kê. Như vậy, tỉ lệ kiểu gen và kiểu alen ở đa hình rs1815930 không liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh ở quần thể nghiên cứu.

Chương 4. THẢO LUẬN

4.1. Ảnh hưởng của các đột biến đến bệnh teo đường mật bẩm sinh

Đối với ở 05 bệnh nhân được lựa chọn để giải trình tự toàn bộ vùng mã hóa trong nghiên cứu này, các gen khác có thể liên quan đến rối loạn phát triển ống mật như *CFTR, JAG1, ZIC3, CFC1, INV, MIF, VEGF* và *IFN-* γ và những gen được cho là liên quan chặt chẽ đến bệnh teo đường mật bẩm sinh như *GPC1, ICAM1, ITGB2, NOTCH1, NOTCH2, NOTCH3, ZIC3, FOXA2, PKD1L1, ADD3, XPNPEP1* đều không phát hiện đột biến hoặc mang những đột biến đã được xác định là lành tính trên cơ sở dữ liệu ClinVar.

Bệnh nhân BA01 và BA02 mang đột biến dị hợp tử kép c.412G>A (p.Gly138Arg) trên gen *FAH* và c.2225A>G (p.Tyr742Cys) trên gen *ERCC4*. Gen *FAH* nằm trên nhiễm sắc thể 15, chứa 14 exon và mã hóa cho 420 axit amin. Phân tích bằng cơ sở dữ liệu mạng tương tác protein STRING v.11 cho thấy protein mã hóa bởi gen *FAH* tương tác mạnh với các protein mã hóa bởi các gen *GSTZ1* (score: 0,986), *HPD* (score: 0,92) và *FAHD1* (0,980) (Hình 4.1). Đây là những gen đóng vai trò quan trọng trong con đường chuyển hóa tyrosine. Việc tyrosine không được phân hủy đúng cách dẫn đến sự tích tụ bất thường của tyrosine và các chất chuyển hóa của nó trong gan, thận có khả năng dẫn đến các hội chứng liên quan đến gan mật như xơ gan hay ung thư biểu mô gan.



Hình 4.1: Tương tác giữa protein FAH và các protein khác

Cơ sở dữ liệu đột biến gen ở người HGMD ghi nhận 69 đột biến trên gen *ERCC4*, trong đó chủ yếu là đột biến dạng Sai nghĩa, đột biến vô nghĩa và đột biến điểm cắt nối mRNA. Đột biến trên gen *ERCC4* ở bệnh nhân teo đường mật bẩm sinh cũng được Sangkhathat và đồng tác giả phát hiện khi sử dụng phương pháp giải trình tự toàn bộ vùng gen mã hóa để tìm kiếm các đột biến gen trong 20 bệnh nhân được chẩn đoán mắc bệnh teo đường mật ở Thái Lan. Mặc dù kết quả này còn nhiều tranh cãi nhưng nghiên cứu này đã mở ra một cách tiếp cận mới để tìm ra nhiều gen khác có thể liên quan đến bệnh teo đường mật bẩm sinh.

Bệnh nhân BA05 mang đột biến c.50_51insG (p.Gly17Glyfs77∗) trên gen *KRT18*. Keratin là một họ lớn các protein sợi trung gian, là thành phần quan trọng của bộ xương tế bào. Ở động vật có vú, sự tổng hợp protein Keratin là đặc điểm nổi bật của tế bào biểu mô, bao gồm tế bào gan và tế bào đường mật. Hệ thống mật bao gồm một mạng lưới các ống dẫn trong gan vận chuyển mật do tế bào gan tiết ra đến các ống dẫn lớn ngoài gan tiếp giáp với túi mật và ruột. Năm 2020, Babu và các đồng tác giả nghiên cứu về mức độ biểu hiện gen và sự tích lũy beta-amyloid ở xung quanh ống mật trên bệnh

nhân teo đường mật bẩm sinh. Kết quả của nghiên cứu cho thấy họ protein keratin bao gồm *KRT8, KRT18, KRT19* có sự tăng cường biểu hiện ở các bệnh nhân teo đường mật bẩm sinh. Điều này cho thấy ảnh hưởng của gen *KRT18* đối với sự hình thành bệnh teo đường mật bẩm sinh.

4.2. Sự liên quan giữa các đa hình nucleotide đơn đến bệnh teo đường mật bẩm sinh

Nhiều nghiên cứu về ảnh hưởng của đa hình nucleotide rs2287622 trên gen *ABCB11* đến nguy cơ mắc các hội chứng liên quan đến gan mật, ứ mật trong gan đã được thực hiện trên nhiều nhóm đối tượng khác nhau. Nghiên cứu trên bệnh nhân người Hán ở Hồ Nam mắc bệnh viêm gan C, Lei và đồng tác giả phát hiện thấy kiểu gen CC và TC của đa hình nucleotide đơn rs2287622 có liên quan đến nguy cơ mắc viêm gan C cao hơn kểu gen TT. Nghiên cứu về đa hình rs2287622 trên nhóm bệnh nhân viêm gan C người Ai Cập, Besheer và đồng tác giả nhận thấy mối liên quan giữa kiểu gen CC và alen C với tình trạng nồng độ axit mật cao ở các bệnh nhân này [108]. Theo đó, kiểu gen đồng hợp tử CC và alen C ở nhóm bệnh nhân có tình trạng xơ gan cao hơn so với nhóm không bị xơ gan, hơn nữa, kiểu gen CC và alen C cũng xuất hiện ở nhóm bệnh nhân bị xơ hóa tiến triển nhiều hơn ở những bệnh nhân bị xơ hóa sớm. Hơn nữa, nồng độ axit mật quan sát được ở bệnh nhân viên gan mạn tính có kiểu gen CC cao hơn nhiều so với những người có kiểu gen TC hoặc TT. Khảo sát đa hình nucleotide đơn rs2287622 trong gen *ABCB11* trên 266 bệnh nhân và 250 đối chứng cho thấy kiểu gen CC và TC khác biệt đáng kể giữa bệnh nhân teo đường mật bẩm sinh và người khỏe mạnh ($P < 0,01$), và alen C có liên quan đến việc tăng nguy cơ mắc bệnh (OR = 2,47; 95%CI: 1,84–3,32; $P < 0,01$). Tần số kiểu gen TC + CC của biến thể p.Val444Ala (c.1331T> C, rs2287622) trong gen *ABCB11* khác biệt có ý nghĩa ở bệnh nhân Việt Nam mắc bệnh teo đường mật bẩm sinh. Biến

thể này có thể liên quan mật thiết đến tình trạng ứ mật và dẫn đến những tổn thương nghiêm trọng đường mật gây tắc nghẽn đường mật cho người bệnh.

Ngoài những nghiên cứu về đột biến, đánh giá mối liên quan giữa các đa hình nucleotide đơn với nguy cơ mắc bệnh cũng được nhiều nhà khoa học quan tâm. *ABCC2* và *MYO5B* là những gen đã được nhóm nhà khoa học Thái Lan nghiên cứu và phát hiện có những biến thể có liên quan đến bệnh teo đường mật bẩm sinh. Các biến thể trên gen *ABCC2* và *MYO5B* đều có đặc điểm chung là dẫn đến tình trạng ứ mật và tăng bilirubin. Gen *ABCC2* mã hóa protein MRP2 (multidrug resistance protein 2). Các biến thể trên gen *ABCC2* làm rối loạn chức năng MRP2 dẫn đến làm tăng bilirubin trong máu. Đồng thời, nguy cơ ứ mật trong thai kỳ có liên quan đến đa hình nucleotide đơn của *MRP2*. Biến thể trên gen *ABCC2* được cho là liên quan đến tình trạng ứ mật ở bệnh nhân mắc hội chứng Dubin-Johnson. Trong nghiên cứu này, chúng tôi không tìm thấy mối liên quan giữa đa hình nucleotide đơn rs927344 trên gen *ABCC2* với nguy cơ mắc bệnh teo đường mật bẩm sinh trên bệnh nhân người Việt Nam.

Các biến đổi trên gen *MYO5B* được phát hiện trên một số bệnh nhân ứ mật trong gan tiến triển gia đình (Progressive familial intrahepatic cholestasis-PFIC). Ngoài ra, các biến thể trong gen *MYO5B* cũng được cho là có liên quan đến dị thường đường mật dẫn đến những triệu chứng lâm sàng của bệnh teo đường mật bẩm sinh như vàng da, ứ mật trong gan. Đột biến *MYO5B* được báo cáo là ảnh hưởng đến chức năng gan mật, làm tăng nồng độ axit mật trong huyết thanh và dẫn đến ứ mật ở những bệnh nhân teo đường mật bẩm sinh. Trong nghiên cứu này, chúng tôi không tìm thấy mối liên quan giữa đa hình nucleotide đơn rs1815930 trên gen *MYO5B* với nguy cơ mắc bệnh teo đường mật bẩm sinh trên bệnh nhân người Việt Nam.

KẾT LUẬN VÀ KIẾN NGHỊ

Kết luận

1. Phát hiện 02 đột biến dị hợp tử kép mới c.412G>A (p.Gly138Arg) trên gen *FAH* và c.2225A>G (p.Tyr742Cys) trên gen *ERCC4* ở hai bệnh nhân là chị em ruột trong một gia đình.

2. Phát hiện 01 đột biến dịch khung c.50_51insG (p.Gly17Glyfs77*) trên gen *KRT18* và 02 đột biến c.314C>A (p.Ser105∗), c.2975C>T (p.Pro992Leu) đều ở dạng dị hợp tử trên gen *ATP7B* ở 01 bệnh nhân mắc đồng thời hai bệnh teo đường mật bẩm sinh và Wilson, trong đó đột biến c.50_51insG (p.Gly17Glyfs77*) trên gen *KRT18* có tiềm năng liên quan đến bệnh teo đường mật bẩm sinh.

3. Đã đánh giá mối liên quan giữa các đa hình đơn nucleotide rs2287622 trên gen *ABCB11*, đa hình rs927344 trên gen *ABCC2*, đa hình rs1815930 trên gen *MYO5B* với nguy cơ mắc bệnh teo đường mật bẩm sinh ở bệnh nhân người Việt Nam. Trong đó:

- Đa hình nucleotide đơn rs2287622 trên gen *ABCB11* có mối liên quan với bệnh teo đường mật bẩm sinh. Cụ thể, người mang kiểu gen TC và CC có nguy cơ mắc bệnh cao hơn so với người mang kiểu gen TT.

- Đa hình rs927344 trên gen *ABCC2* và đa hình rs1815930 trên gen *MYO5B* không có mối liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh ở người Việt Nam.

Kiến nghị

Tiếp tục nghiên cứu sâu hơn về ảnh hưởng của những đột biến này đến chức năng protein, từ đó làm rõ cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh.

Mặt khác, có thể mở rộng phạm vi nghiên cứu nhằm phân tích thêm những điểm đa hình nucleotide đơn khác có liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh ở người Việt Nam.

DANH MỤC CÁC BÀI BÁO ĐÃ XUẤT BẢN LIÊN QUAN ĐẾN LUẬN ÁN

1. **Nguyễn Văn Tụng**, Nguyễn Thị Kim Liên, Nguyễn Thị Phương Mai, Nguyễn Phạm Anh Hoa, Nguyễn Huy Hoàng. Khảo sát mối liên quan của *ABCC2* rs927344 và *MYO5B* rs1815930 đối với bệnh teo đường mật bẩm sinh ở người Việt Nam. *Hội nghị Công nghệ sinh học toàn quốc 2020*.724-729.

2. **Van Tung Nguyen**, Lien Nguyen Thi Kim, Lan Nguyen Ngoc, Mai Nguyen Thi Phuong, Yen Pham Thi Hai, Hoa Nguyen Pham Anh, Hoang Nguyen Huy. The role of p.Val444Ala variant in the *ABCB11* gene and susceptibility to biliary atresia in Vietnamese patients. *Medicine.* 2021; 100(47) p e28011.

3. Hoa Nguyen Pham Anh, Lien Nguyen Thi Kim, **Tung Nguyen Van**, Lan Nguyen Ngoc, Mai Nguyen Thi Phuong, Huong Nguyen Thi Mai, Thach Hoang Ngoc, Hoang Nguyen Huy. Biliary atresia combined Wilson disease identified by whole exome sequencing in Vietnamese patient with severe liver failure. *Medicine*. 2022;101(2):p e28547