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

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Hà Nội – 2024

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

LUẬN ÁN TIẾN SĨ SINH HỌC ỨNG DỤNG

Ngành: Công nghệ sinh học

Mã số: 9420201

Xác nhận của Học viện Khoa học và Công nghệ	Người hướng dẫn 1	Người hướng dẫn 2
	GS. TS. Nguyễn Huy Hoàng	TS. Nguyễn Thị Kim Liên

Hà Nội - 2024

LỜI CAM ĐOAN

Tôi xin cam đoan luận án "Nghiên cứu đột biến và đa hình của một số gen ở trẻ em mắc bệnh teo đường mật bẩm sinh" là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Nghiên cứu được thực hiện trong khuôn khổ đề tài Quỹ phát triển Khoa học và Công nghệ Quốc gia Nafosted mã số 108.02-2018.305 "Nghiên cứu vai trò của biến đổi di truyền trên các gen liên quan ở bệnh teo đường mật bẩm sinh bằng giải trình tự toàn bộ vùng mã hóa" do GS.TS Nguyễn Huy Hoàng làm chủ nhiệm. Luận án sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được công bố chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 28 tháng 10 năm 2024

Tác giả luận án

Nguyễn Văn Tụng

LỜI CẢM ƠN

Tôi xin được bày tỏ lòng kính trọng và biết ơn sâu sắc tới GS.TS. Nguyễn Huy Hoàng và TS. Nguyễn Thị Kim Liên đã tận tình hướng dẫn, tạo mọi điều kiện thuận lợi cho tôi trong suốt quá trình học tập, nghiên cứu cũng như thực hiện luận án.

Tôi cũng xin chân thành cảm ơn ban Giám đốc, tập thể cán bộ Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã tạo điều kiện thuận lợi nhất cho tôi trong quá trình học tập và nghiên cứu tại đây.

Tôi xin chân thành cảm ơn Ban lãnh đạo Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã giúp đỡ và tạo điều kiện thuận lợi cho tôi trong quá trình thực hiện luận án.

Tôi xin trân trọng cảm ơn sự ủng hộ và tham gia nhiệt tình của các cán bộ, đồng nghiệp của phòng Hệ gen học chức năng, Viện Nghiên cứu hệ gen cùng các bác sĩ tại khoa Gan mật, khoa Di truyền và Sinh học Phân tử, Bệnh viện Nhi Trung Ương đã giúp đỡ, đồng hành cùng tôi trong quá trình triển khai thí nghiệm, phân tích kết quả và hoàn thiện luận án.

Tôi xin trân trọng cảm ơn phòng Quản lý tổng hợp Viện Nghiên cứu hệ gen đã nhiệt tình giúp đỡ và tạo điều kiện thuận lợi về mọi mặt, hỗ trợ tôi thực hiện các thủ tục cần thiết để tôi có thể hoàn thành chương trình học tập và nghiên cứu của nghiên cứu sinh.

Tôi muốn bày tỏ lòng biết ơn sâu sắc tới gia đình và người thân đã ở bên tôi chăm sóc, động viên, giúp tôi yên tâm học tập và nghiên cứu.

Tác giả luận án

Nguyễn Văn Tụng

MỤC LỤC

LỜI CAM ĐOAN i
LỜI CẢM ƠN ii
DANH MỤC CÁC TỪ VIẾT TẮT vi
DANH MỤC BẢNG vii
DANH MỤC HÌNH ix
MỞ ĐẦU 1
Chương 1. TỔNG QUAN 4
1.1. Bệnh teo đường mật bẩm sinh 4
1.1.1. Triệu chứng lâm sàng và phân loại bệnh 5
1.1.1.1. Triệu chứng lâm sàng 5
1.1.1.2. Phân loại bệnh 7
1.1.2. Nguyên nhân gây bệnh teo đường mật bẩm sinh 8
1.2. Vai trò của yếu tố di truyền trong cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh 13
1.2.1. Gen *ABCB1*1 20
1.2.2. Gen *ABCC2* 21
1.2.3. Gen *ADD3* và *XPNPEP1* 22
1.2.4. Gen *GPC1* 23
1.2.5. Gen *ARF6* 24
1.2.6. Gen *MIR499* 25
1.2.7. Gen *ITGB2* và *ICAM1* 25
1.2.8. Nhóm gen liên quan đến quá trình xơ hóa 26
1.2.9. Một số gen liên quan đến sự hình thành lông mao bất thường 27
1.3. Tình hình nghiên cứu bệnh teo đường mật bẩm sinh 28
1.3.1. Tình hình nghiên cứu trên thế giới 28
1.3.2. Tình hình nghiên cứu ở Việt Nam 31

Chương 2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU 33
2.1. Đối tượng nghiên cứu..... 33
2.1.1. Tiêu chuẩn lựa chọn mẫu 33
2.1.1.1 Mẫu bệnh..... 33
2.1.1.2. Mẫu đối chứng 34
2.1.1.3. Tiêu chuẩn loại trừ 34
2.1.2. Cỡ mẫu nghiên cứu 34
2.2. Hóa chất và trang thiết bị 35
2.3. Phương pháp nghiên cứu..... 35
2.3.1. Tách chiết DNA tổng số..... 36
2.3.2. Giải trình tự toàn bộ vùng gen mã hóa..... 37
2.3.2.1. Tạo thư viện DNA..... 37
2.3.2.2. Giải trình tự bằng Illumina NextSeq 500..... 38
2.3.3. Xác định và chú giải biến thể..... 38
2.3.4. Sàng lọc biến thể 42
2.3.5. Xác định kiểu gen của đa hình rs2287622 bằng kỹ thuật RFLP 44
2.3.6. Xác định kiểu gen của đa hình bằng kỹ thuật ARMS-PCR..... 45
2.3.7. Giải trình tự Sanger..... 48
2.3.8. Phân tích thống kê..... 48
Chương 3. KẾT QUẢ 49
3.1. Thu thập mẫu và tách chiết DNA tổng số..... 49
3.2. Giải trình tự toàn bộ vùng mã hóa 52
3.2.1. Kết quả tạo thư viện DNA 52
3.2.2. Kết quả giải trình tự toàn bộ vùng gen mã hóa..... 55
3.3. Kết quả xử lý số liệu và chú giải biến thể..... 58
3.3.1. Sắp hàng dữ liệu với hệ gen tham chiếu 58

3.3.2. Xác định và chú giải biến thể 59
3.4. Sàng lọc biến thể 60
3.5. Đánh giá ảnh hưởng của các đột biến 65
3.5.1. Bệnh nhân BA01 và BA02 69
3.5.2. Bệnh nhân BA03 và BA04 71
3.5.3. Bệnh nhân BA05 72
3.6. Ảnh hưởng của các đa hình nucleotide đến mức độ mẫn cảm bệnh teo đường mật bẩm sinh 74
3.6.1. Ảnh hưởng của đa hình rs2287622 74
3.6.2. Ảnh hưởng của đa hình rs927344 78
3.6.3. Ảnh hưởng của đa hình rs1815930 80
Chương 4. THẢO LUẬN 83
4.1. Ảnh hưởng của các đột biến đến bệnh teo đường mật bẩm sinh 83
4.2. Sự liên quan giữa các đa hình nucleotide đơn đến bệnh teo đường mật bẩm sinh 89
KẾT LUẬN VÀ KIẾN NGHỊ 96
DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ LIÊN QUAN ĐẾN LUẬN ÁN 98
TÀI LIỆU THAM KHẢO 99
PHỤ LỤC ..... a

DANH MỤC CÁC TỪ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
ARMS-PCR	Amplification refractory mutation system polymerase chain reaction	PCR đặc hiệu allele tích hợp phân tích điểm đột biến
BA	Biliary Atresia	Teo đường mật bẩm sinh
BASM	Biliary atresia splenic malformation	Teo đường mật kết hợp với các dị dạng khác
dbSNP	The single nucleotide polymorphismdatabase	Cơ sở dữ liệu đa hình đơn nucleotide
HWE	Hardy – Weinberg Equilibrium	Cân bằng Hardy-Weinberg
NCBI	National Center for Biotechnology Information	Trung tâm thông tin Công nghệ sinh học Hoa Kỳ
NGS	Next generation sequencing	Giải trình tự thế hệ mới
RFLP	Restriction fragment length polymorphisms	Đa hình chiều dài đoạn cắt giới hạn
SNP	Single nucleotide polymorphisms	Đa hình đơn nucleotide
WES	Whole exome sequencing	Giải trình tự toàn bộ vùng mã hóa
WGS	Whole genome sequencing	Giải trình tự toàn bộ hệ gen

DANH MỤC BẢNG

Bảng 1.1: Các biến thể gen và vai trò của chúng trong cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh 17

Bảng 2.1: Thành phần của phản ứng cắt enzyme *Hae*III..... 45

Bảng 2.2: Chu trình nhiệt của phản ứng ARMS-PCR 47

Bảng 2.3: Số lượng và kích thước sản phẩm PCR tương ứng với các kiểu gen của *ABCC2* rs927344 và *MYO5B* rs1815930 47

Bảng 2.4: Thành phần phản ứng PCR giải trình tự Sanger 48

Bảng 3.1: Thông tin lâm sàng của người tham gia nghiên cứu 49

Bảng 3.2: Thông tin các mẫu bệnh nhân được lựa chọn để giải trình tự WES 50

Bảng 3.3: Kích thước và nồng độ thư viện DNA 53

Bảng 3.4: Thông tin đọc trình tự toàn bộ vùng mã hóa của 05 bệnh nhân teo đường mật bẩm sinh 58

Bảng 3.5: Kết quả sắp hàng các mẫu giải trình tự với hệ gen tham chiếu..... 58

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 60

Bảng 3.7: Kết quả sàng lọc đột biến xảy ra trong những gen liên quan đến bệnh teo đường mật bẩm sinh 61

Bảng 3.8: Đột biến xảy ra trong gen liên quan đến bệnh teo đường mật bẩm sinh của các bệnh nhân..... 62

Bảng 3.9: Kết quả đánh giá ảnh hưởng của các điểm đột biến bằng công cụ tin sinh 66

Bảng 3.10: Các đột biến có tiềm năng gây bệnh..... 69

Bảng 3.11: Thống kê kiểu gen và tần suất alen đa hình *ABCB11* rs2287622 74

Bảng 3.12: Kết quả phân tích sự liên quan giữa *ABCB11* rs2287622 và nguy cơ mắc bệnh teo đường mật bẩm sinh..... 75

Bảng 3.13: Thống kê kiểu gen và tần suất alen đa hình *ABCC2* rs927344 78

Bảng 3.14: Kết quả phân tích sự liên quan giữa *ABCC2* rs927344 và nguy cơ mắc bệnh teo đường mật bẩm sinh..... 79

Bảng 3.15: Thống kê kiểu gen và tần suất alen đa hình *MYO5B* rs1815930 . 80

Bảng 3.16: Kết quả phân tích sự liên quan giữa *MYO5B* rs1815930 và nguy cơ mắc bệnh teo đường mật bẩm sinh..... 81

Bảng 4.1: Các biến thể trong gen *MYO5B* được phát hiện ở bệnh nhân teo đường mật bẩm sinh..... 94

DANH MỤC HÌNH

Hình 1.1: Tương tác giữa các cơ chế sinh bệnh của teo đường mật bẩm sinh . 9

Hình 1.2: Mối liên hệ giữa các con đường và cơ chế gây bệnh teo đường mật bẩm sinh [56]..... 14

Hình 1.3: Cấu trúc protein MRP2 mã hóa bởi gen *ABCC2* 22

Hình 2.1: Sơ đồ nghiên cứu 36

Hình 2.2: Quy trình xử lý số liệu bằng các công cụ tin sinh học..... 40

Hình 2.3: Quy trình sàng lọc biến thể từ dữ liệu WES 42

Hình 2.4: Nguyên lý của phản ứng ARMS-PCR..... 46

Hình 3.1: Kết quả kiểm tra chất lượng thư viện DNA..... 55

Hình 3.2: Chất lượng giải trình tự 05 bệnh nhân mắc teo đường mật bẩm sinh 57

Hình 3.3: Đột biến dị hợp tử kép trên bệnh nhân BA01 và BA02 70

Hình 3.4: So sánh trình tự axit amin xung quanh vị trí đột biến giữa các loài71

Hình 3.5: Kết quả giải trình tự Sanger cho thấy bệnh nhân BA03 và BA04 không mang đột biến c.368A>C dạng dị hợp tử trên gen *CYP27A1* 72

Hình 3.6: Đột biến c.50_51insG xảy ra ở amino axit thứ 17 (màu đỏ) làm protein bị cắt ngắn sau vị trí đột biến 77 amino axit (phần màu xanh) 73

Hình 3.7: Đột biến trên bệnh nhân BA05 73

Hình 3.8: Phân tích di truyền đa hình nucleotide đơn rs2287622 ở 05 bệnh nhân teo đường mật bẩm sinh được lựa chọn để giải trình tự WES..... 77

Hình 4.1: Tương tác giữa protein *FAH* và các protein khác..... 86

Hình 4.2: Tương tác giữa protein mã hóa bởi *ABCC2* với các protein khác.. 93

MỞ ĐẦU

Teo đường mật bẩm sinh (biliary atresia – BA) là bệnh lý hiếm gặp 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. Trẻ mắc bệnh có thể có vẻ ngoài bình thường khi mới sinh. Da và mắt bệnh nhân teo đường mật bẩm sinh có màu vàng do lượng sắc tố mật trong máu cao. Thông thường, đây là dấu hiệu đầu tiên của bệnh teo đường mật bẩm sinh. Phương pháp chẩn đoán bệnh hiện nay chủ yếu dựa trên lâm sàng và cận lâm sàng như chỉ số bilirubin, sinh thiết gan... và phẫu thuật để chẩn đoán xác định. Bệnh có thể điều trị bằng phẫu thuật trong giai đoạn mới sinh. Tuy nhiên, nếu không được phát hiện và phẫu thuật sớm bệnh sẽ phát triển thành xơ gan mật thứ phát dẫn đến tử vong khi 1 tuổi.

Cho đến nay những hiểu biết về cơ chế bệnh nguyên và bệnh sinh của các tổn thương gan và hệ thống đường mật trong teo đường mật bẩm sinh vẫn chưa được làm rõ. 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. Trong đó, rối loạn về gen di truyền bao gồm đột biến và đa hình nucleotide đơn được nhiều nhà khoa học trên thế giới quan tâm. Việc xác định được mối liên quan giữa các đa hình nucleotide đơn với nguy cơ mắc bệnh teo đường mật bẩm sinh giúp các bác sĩ chẩn đoán sớm từ đó đưa ra phương pháp điều trị phù hợp. Tuy nhiên, hiện nay các nghiên cứu ở Việt Nam chủ yếu tập trung vào yếu tố lâm sàng, các phương pháp tăng cường hiệu quả điều trị bệnh. Nghiên cứu về yếu tố di truyền ở bệnh teo đường mật bẩm sinh còn khá hạn chế.

Trên thế giới, các nhà khoa học đã tìm ra mối quan hệ giữa các đột biến gen và đa hình đơn nucleotide trên nhiều gen liên quan đến bệnh teo đường mật bẩm sinh. Với số lượng và kích thước gen lớn, việc nghiên cứu từng gen riêng lẻ đòi hỏi nhiều thời gian, chi phí cao. Giải trình tự toàn bộ vùng mã hóa và toàn bộ hệ gen bằng công nghệ giải trình tự thế hệ mới cho phép xác định biến thể xảy ra đồng thời trên nhiều gen, qua đó rút ngắn quá trình phân tích đã trở thành phương pháp thay thế hiệu quả khi tiến hành các nghiên cứu di truyền. Hàng chục nghìn biến thể gen có thể được xác định trong mỗi vùng mã hóa trong nhiều bệnh phức tạp như: tim mạch, thần kinh, chuyển hóa…. 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 các con đường: sự phát triển gan mật, quá trình xơ hóa, viêm, sự bất thường vi lông mao (bệnh lý tiêm mao). Do đó, ngoài việc sử dụng phương pháp giải trình tự toàn bộ vùng gen mã hóa để xác định các đột biến gen gây bệnh, trong nội dung luận án này còn thực hiện phân tích mối liên quan giữa ba đa hình nucleotide đơn 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. Các gen này tham gia vào một con đường gây bệnh teo đường mật bẩm sinh khác nhau, đồng thời, chưa có nghiên cứu cứu nào trên thế giới đánh giá mối liên quan của chúng đến bệnh teo đường mật bẩm sinh.

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.

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. Hệ thống đường mật được tạo nên từ túi thừa gan của ruột trước vào tuần thứ 4 của phôi và biệt hóa thành 2 thành phần đầu và đuôi. Túi mật, ống túi mật và ống mật chủ được tạo nên từ phần đầu. Còn đường mật trong gan cũng như phần đường mật ngoài gan còn lại được tạo nên từ phần đuôi. 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. Bệnh có thể điều trị bằng phẫu thuật trong giai đoạn mới sinh. Tuy nhiên, nếu không được phát hiện và phẫu thuật sớm bệnh sẽ phát triển thành xơ gan mật thứ phát dẫn đến tử vong khi 1 tuổi. Teo đường mật bẩm sinh là nguyên nhân gây tử vong phổ biến nhất trong những hội chứng về gan mật ở trẻ em [1]. Phương pháp điều trị duy nhất hiện nay là phẫu thuật Kasai mở thông ruột để cải thiện dòng mật. Tuy nhiên có một tỷ lệ cao bệnh nhân ngay cả khi khôi phục được lưu lượng dòng mật vẫn phát triển tiền trạng viêm và xơ tiến triển đường mật trong gan dẫn đến xơ gan thứ phát. Độ tuổi thực hiện phẫu thuật Kasai là một yếu tố quan trọng ảnh hưởng đến kết quả. Những bệnh nhân đã nhận được phẫu thuật Kasai trong vòng 60 ngày đầu đời ít có khả năng phải ghép gan hơn.

Tần suất 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. Ở Anh và Pháp tỉ lệ mắc từ 1/17.000-1/19.000 trẻ sinh sống, ở Hoa Kỳ là 1/15.000 trẻ [2–4]. Ở Hà Lan, tỉ lệ mắc ở nông thôn cao hơn 68% so với ở thành thị [1]. Tần suất mắc bệnh ở các nước Châu Á và Tây Thái Bình Dương khá cao, ở Nhật Bản là 1/10.000 trẻ và ở Đài Loan là 1/5.000 trẻ [5]. Đặc biệt, theo một nghiên cứu năm 2016, bệnh teo đường mật bẩm sinh ở Trung Quốc có tỉ lệ mắc ít hơn so với các nước phương Tây [6]. Lý do cho điều này vẫn chưa rõ ràng nhưng gợi ý rằng các yếu tố môi trường hoặc di truyền có thể góp phần gây ra hội chứng teo đường mật bẩm sinh. Tỉ lệ mắc bệnh cao hơn cũng được thấy ở những Hawaii, người Inuits và người Mỹ bản địa [7]. Tuy nhiên, nguyên nhân dẫn đến sự khác biệt về tỷ lệ mắc bệnh giữa các khu vực trên thế giới vẫn chưa được

làm rõ. Một nghiên cứu gần đây cho thấy sự khác biệt về tỉ lệ mắc bệnh teo đường mật bẩm sinh có thể tương quan với tần suất của haplotype kháng nguyên bạch cầu (HLA) ở người [8]. Bệnh teo đường mật bẩm sinh thường xuất hiện riêng lẻ, rất ít trường hợp được báo có có tính chất gia đình. Bệnh không đồng nhất về mặt di truyền, các mối liên hệ gia đình là rất hiếm. Phân tích từ 12 tài liệu khoa học với 35 cặp song sinh, Xu và đồng tác giả tổng hợp được 36 bệnh teo đường mật bẩm sinh (trên tổng số 70 đối tượng) cho thấy teo đường mật bẩm sinh ở các cặp song sinh là rất hiếm gặp và tình trạng bệnh ở các cặp sinh đôi cũng không giống nhau, tuy nhiên nghiên cứu về bệnh trên các trường hợp này cung cấp nhiều thông tin hữu ích về căn bệnh có nguyên nhân phức tạp [9].

1.1.1. Triệu chứng lâm sàng và phân loại bệnh

1.1.1.1. Triệu chứng lâm sàng

Triệu chứng lâm sàng chủ yếu của bệnh teo đường mật bẩm sinh là tăng bilirubin, vàng da, phân bạc màu và gan lớn [10]. Trẻ mắc bệnh có thể có vẻ ngoài bình thường khi mới sinh. Da và mắt bệnh nhân teo đường mật bẩm sinh có màu vàng do lượng sắc tố mật trong máu cao. Thông thường, đây là dấu hiệu đầu tiên của bệnh teo đường mật bẩm sinh. Ngoài ra, nước tiểu có màu sẫn cũng là một đặc điểm lâm sàng khác. Một số bilirubin tăng trong máu được thận lọc và loại bỏ qua nước tiểu. Các 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. Nguyên nhân là bởi tình trạng vàng da và ứ mật có thể xuất hiện ngay từ thời kỳ sau sinh rất giống với với tình trạng vàng da sinh lý thường gặp ở hơn một nửa số trẻ sơ sinh. Tuy nhiên, ở một số trẻ các triệu chứng vàng da và phân bạc màu xuất hiện muộn hơn về sau. Người ta nhận thấy ở nhiều bệnh nhi, phân su có màu vàng hay vàng nhạt, nước tiểu trở nên đậm màu, gan lớn là do hiện tượng ứ mật. Một điểm đáng lưu ý nữa là phần lớn trẻ bị teo đường mật bẩm sinh có sự phát triển cân nặng và thể chất hoàn toàn bình thường trong những tháng đầu, thậm chí cho đến lúc phẫu thuật. Tuy nhiên, vẫn có một số ít trẻ có tình trạng suy dinh dưỡng, thiếu máu hay chậm phát triển. Triệu chứng kèm theo bao gồm sự giảm tỷ lệ prothrombin do tình trạng kém hấp thu vitamin K, chảy máu nội sọ, chảy máu ngoài da, bilirubin máu tăng, chủ yếu là bilirubin

trực tiếp, nồng độ axit mật, nồng độ lipo-protein X (Lp-X), nồng độ gamma-glutamyl transpetidase trong máu tăng.

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. Hình ảnh siêu âm có thể cho thấy túi mật không có hoặc không có túi mật. Việc phát hiện và chẩn đoán bệnh muộn thường làm giảm hiệu quả của phẫu thuật Kasai, đặc biệt là với trẻ sơ sinh trên 45 ngày tuổi.

Cho đến nay, siêu âm gan mật vẫn được xem là phương tiện chẩn đoán hình ảnh được sử dụng đầu tiên ở một bệnh nhi được nghi ngờ teo đường mật bẩm sinh. Siêu âm cần được làm ở hai thời điểm, lúc trẻ đói và sau khi trẻ bú. Mục đích của siêu âm ở hai thời điểm là nếu có túi mật và không teo đường mật thì sau khi trẻ bú, túi mật sẽ tăng co bóp theo đáp ứng sinh lý bình thường. Các dấu hiệu gián tiếp đánh giá teo đường mật bằng siêu âm có thể thấy: túi mật kích thước nhỏ (đôi khi không nhìn thấy túi mật khi trẻ nhịn bú hoặc chỉ nhìn thấy di tích túi mật), túi mật thành dày, bờ không đều. Kích thước túi mật không thay đổi trước và sau khi bú. Trường hợp túi mật có kích thước nhỏ hơn bình thường và sự thay đổi kích thước trước và sau bú dưới 30% cũng có thể loại trừ teo đường mật.

Một phương pháp được sử dụng rộng rãi ở các nước châu Âu là chụp nhấp nháy gan-mật (chụp SPECT) sử dụng đồng vị phóng xạ Tc99 nhằm phát hiện sự đào thải các chất đồng vị phóng xạ qua đường mật xuống ruột. Bình thường chất đồng vị phóng xạ được tế bào gan hấp thu, bài tiết và đào thải xuống theo đường mật. Khi bị teo đường mật, chất đồng vị phóng xạ không được bài tiết xuống ruột do đó sẽ cho hình ảnh dương tính sau 24 h. Sự có mặt của đồng vị phóng xạ ở ruột non sau khi sử dụng thuốc cho phép loại trừ được bệnh teo đường mật bẩm sinh. Tuy nhiên, trong những trường hợp hội chứng mật đặc hay trong trường hợp thiểu sản đường mật, xét nghiệm hình ảnh này ít giúp được chẩn đoán phân biệt, độ đặc hiệu của phương pháp giao động từ 40 – 100% và có giá thành cao.

Chụp đường mật trong mổ là tiêu chuẩn vàng để chẩn đoán teo mật bẩm sinh với các bệnh nhân nghi ngờ trong khi các thăm dò khác chưa rõ ràng nên

được mổ thăm dò. Chụp trong mổ nhằm xác định chất màu có chảy bình thường vào ống tiêu hóa hay không.

Sinh thiết vẫn được xem là yếu tố quyết định trong chẩn đoán những trường hợp teo đường mật bẩm sinh khó, teo đường mật ngoài gan. Hiệu quả chẩn đoán của phương pháp đạt 96 – 98%, phụ thuộc nhiều vào mảnh sinh thiết lấy được. Mẫu sinh thiết ở giai đoạn sớm của bệnh có thể nhầm với viêm gan sơ sinh, ngược lại khi ở giai đoạn muộn tổn thương quá nặng rất khó để chẩn đoán đúng. Sinh thiết thường được chỉ định trong những trường hợp bệnh nhi có túi mật. Việc sinh thiết được tiến hành qua đường mổ thường, qua đường nội soi ổ bụng hay thậm chí qua da. Các đặc điểm giải phẫu bệnh của teo đường mật bẩm sinh bao gồm: gan ứ mật xanh thẫm, bề mặt sần sùi, thô ráp, hình ảnh ứ mật, xơ hoá quanh khoảng cửa, vặn xoắn các bè gan.

1.1.1.2. Phân loại bệnh

Hiện nay, có hai cách chính để phân loại các thể teo đường mật bẩm sinh là dựa trên lâm sàng và dựa vào giải phẫu bệnh. 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 [11]:

- Teo đường mật bẩm sinh không kèm các dị tật bất thường khác: chiếm khoảng 70 - 85% trường hợp. Ở nhóm này, trẻ thường vàng da tiến triển dần trong 2 tháng đầu sau sinh kết hợp với phân bạc màu. Teo đường mật bẩm sinh không đi kèm các dị tất bất thường khác được cho là kết quả của sự kết hợp giữa các tác động từ môi trường và khuynh hướng di truyền [12].

- Teo đường mật kết hợp với các dị dạng khác (hội chứng BASM – biliary atresia splenic malformation): chiếm khoảng 10 – 15% trường hợp. Các dị dạng

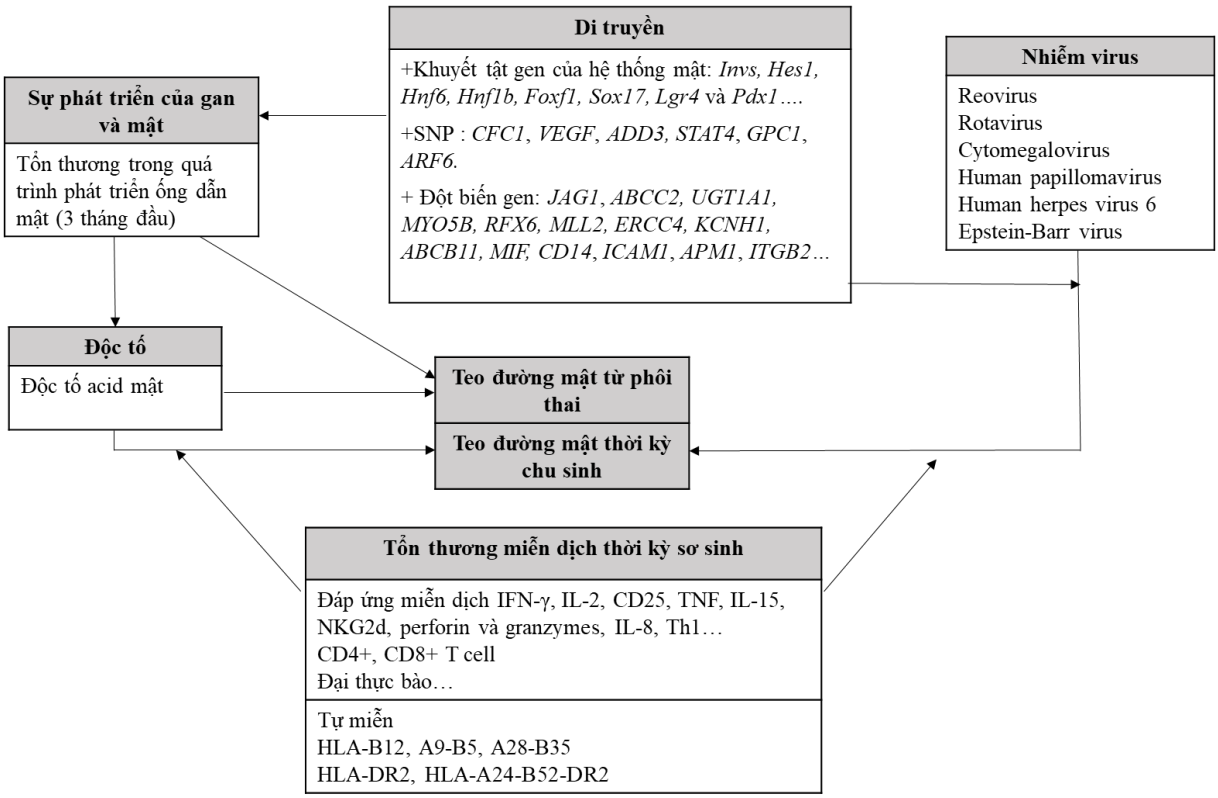
bao gồm: bất thường lách, đảo ngược phủ tạng, bất thường xoay, gián đoạn tĩnh mạch chủ dưới và bất thường về tim. Tiên lượng sau phẫu thuật không tốt so với nhóm không có dị tật. Teo đường mật kết hợp với các dị dạng khác có thể được hình thành do hình thái khiếm khuyết của ống mật trong quá trình phát triển phôi thai và các bất thường về nhiễm sắc thể [13].

- Teo đường mật có nang: teo đường mật ngoài gan đặc trưng bởi sự teo hoặc vắng mặt do viêm gây hủy hoại đường mật. Tuy nhiên, khoảng 10% các trường hợp có sự hình thành các nang đường mật có thể dẫn đến chẩn đoán nhầm với nang ống mật chủ. Những trường hợp teo đường mật có nang có tiên lượng sau phẫu thuật tốt hơn so với nhóm teo đường mật trong gan.

1.1.2. Nguyên nhân gây bệnh teo đường mật bẩm sinh

Cho đến nay những hiểu biết về cơ chế bệnh nguyên và bệnh sinh của các tổn thương gan và hệ thống đường mật trong teo đường mật bẩm sinh vẫn chưa được làm rõ. 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. Ngoài ra, người ta còn đưa ra giả thuyết rằng một số yếu tố ngoại sinh vẫn chưa được biết đến, có thể virus (ví dụ như cytomegalovirus) hoặc chất độc (biliatresone) có thể là căn nguyên của căn bệnh này [14,15] (Hình 1.1). Biliatresone gây ra sự phá hủy hệ thống mật ngoài gan ở cá ngựa vằn. Chất độc này cũng làm mất lông mao ở các tế bào đường mật ngoài gan của chuột sơ sinh trong môi trường nuôi cấy, cho thấy rằng bệnh lý lông mao do độc tố gây ra góp phần vào cơ chế bệnh sinh của teo đường mật bẩm sinh. Một số nghiên cứu về tỷ lệ mắc bệnh teo đường mật bẩm sinh ở một số nhóm người chẳng hạn như Châu Á và Polynesia, cho thấy rằng các yếu tố di truyền và môi trường có thể là nguyên nhân gây ra bệnh. Các nghiên cứu di truyền gần đây đã phát hiện ra mối liên hệ giữa vàng da ứ mật và các khuynh hướng di truyền trong cả DNA hạt nhân và DNA ty thể [16,17]. Có giả thuyết cho rằng đột biến soma cũng được coi là một nguyên nhân gây ra bệnh teo đường mật bẩm sinh [18].

Virus được cho là liên quan đến nguyên nhân của bệnh teo đường mật. Kể từ khi được Benjamin Landing đề cập vào năm 1974, nhiều loại virus đã được phát hiện trực tiếp ở gan bị tổn thương và tàn dư mật hoặc gián tiếp do sự hiện diện của các dấu hiệu huyết thanh nhiễm trùng ở những bệnh nhân bị teo đường mật, bao gồm human papillomavirus, human herpes virus 6, Epstein–Barr virus, reovirus, rotavirus…. Ngoài ra, đã có báo cáo cho rằng sự tương tác giữa hệ vi sinh vật đường ruột với rối loạn chuyển hóa tryptophan và axit mật có thể làm nặng thêm tổng thương gan do teo đường mật bẩm sinh [19].



Hình 1.1: Tương tác giữa các cơ chế sinh bệnh của teo đường mật bẩm sinh

Một số nghiên cứu đã phát hiện sự có mặt của virus viêm gan C ở các bệnh nhân teo đường mật bẩm sinh [20,21]. Đối với cytomegalovirus, Tarr và đồng tác giả phát hiện 5 trong số 23 bệnh nhân bị teo mật bẩm sinh dương tính với cytomegalovirus [22]. Tương tự như vậy, Oliveira và đồng tác giả đã phát hiện IgM dương tính với cytomegalovirus ở 28,5% bệnh nhân EHBA hoặc u nang mật [23]. Reovirus, đặc biệt là reovirus type 3 đã được nhắc đến trong căn nguyên gây teo đường mật bẩm sinh. Tyler và đồng tác giả (1998) đã tìm thấy gen của reovirus L1 trong mô gan của 11 trong số 20 bệnh nhân teo đường mật bẩm sinh [24]. Trong nghiên cứu của mình, Rauschenfels và các đồng tác giả

đã kiểm tra các mẫu sinh thiết gan của 74 bệnh nhân teo đường mật bẩm sinh, và xác định DNA/RNA của virus ở 30 bệnh nhân trong đó có 21 trường hợp phát hiện DNA / RNA của virus reovirus [25]. Nghiên cứu của Riepenhoff-Talty (1996) đã phát hiện được RNA của rotavirus nhóm C trong 50% tổng số mẫu gan của bệnh nhân teo đường mật bẩm sinh [26]. Tuy nhiên, một số nghiên cứu khác lại cho thấy không có mối liên hệ giữa việc nhiễm reovirus 3 và BA hoặc không phát hiện được RNA của rotavirus (nhóm A, B, C) từ mẫu gan của bệnh nhân BA [27,28]. Nghiên cứu gần đây cũng chỉ ra không có mối liên quan giữa teo đường mật bẩm sinh và các virus viêm gan A, B, C [29].

Một nguyên nhân nữa cũng được giả thiết gây ra teo đường mật bẩm sinh là do các tổ thương của hệ miễn dịch bẩm sinh và miễn dịch thích ứng. Hệ thống miễn dịch bẩm sinh là tuyến phòng thủ đầu tiên chống lại các tác nhân gây bệnh bao gồm các đại thực bào, tế bào đuôi gai và bạch cầu trung tính. Các tế bào đuôi gai và đại thực bào đóng vai trò là cầu nối giữa hệ thống miễn dịch bẩm sinh và miễn dịch thích ứng. Rối loạn hệ miễn dịch được giả thiết có nguyên nhân sâu sa từ nhiễm khuẩn, virus và các chất độc hại dẫn đến các tế bào trình diện kháng nguyên ở biểu mô đường mật trình diện kháng nguyên lạ, kích hoạt đại thực bào và tế bào lympho về vị trí viêm gây tổn thương biểu mô đường mật. Sự biểu hiện bất thường của kháng nguyên bạch cầu người (HLA) – DR trong biểu mô đường mật ở bệnh nhân mất máu là bằng chứng khác về sự tham gia của quá trình miễn dịch, vì sự hiện diện của nó cho thấy rằng các tế bào này hoạt động như các tế bào trình kháng nguyên và kích hoạt trực tiếp tế bào lympho T [30]. Hàm lượng phân tử kết dính nội bào (ICAM - 1) và phân tử kết dính tế bào mạch máu (VCAM – 1) cao cũng làm tăng mức độ trầm trọng của bệnh. Mức ICAM-1 và VCAM-1 cao có liên quan đến bệnh gan tiến triển [31]. Cơ chế này cho thấy rằng các phân tử bám dính có thể đóng một vai trò đáng kể trong phản ứng viêm ở nhu động mật, có thể bởi sự lưu giữ và kích hoạt của bạch cầu lưu hành.

Đáp ứng miễn dịch tế bào chủ yếu trong teo đường mật bẩm sinh bao gồm sự kích hoạt các tế bào T CD4+, CD8+ trong vùng sản xuất cytokine Th1 (IL-2, IFN-γ) và đại thực bào tiết ra TNF-α [34,35]. Những tế bào lympho này xâm nhập vào trong biểu mô ống mật dẫn đến sự thoái hóa ống mật [36]. Việc sản xuất và hoạt động của cytokine thường được định vị trong cơ quan đích,

tuy nhiên khi sự sản sinh cytokine tăng lên có thể được lưu thông trong huyết tương.

Một giả thuyết có thể giải thích cho sự tiến triển của tổn thương đường mật ở các bệnh nhân teo đường mật bẩm sinh là một cuộc tấn công qua trung gian miễn dịch trên biểu mô đường mật. Sau khi nhiễm virus, các tế bào biểu mô ống mật bị tổn thương có thể biểu hiện kháng nguyên tự miễn và phản ứng tự động tạo ra Th1 trung gian và sản xuất tế bào B của kháng thể tấn công ống mật. Ngoài ra, các protein virus có thể có cấu trúc tương tự như protein trong biểu mô ống mật do đó tạo ra sự miễn dịch tự miễn dựa trên mô phỏng phân tử. Bằng chứng mạnh nhất cho giả thuyết này đã thu được trên chuột, nhưng ở người chỉ có bằng chứng gián tiếp cho vai trò của tự miễn trong sinh bệnh teo đường mật bẩm sinh [37].

Bản chất tiến triển của tổn thương gan ở bệnh nhân teo đường mật sự hiện diện của các tế bào lympho trong gan, và liên kết với một số loại HLA nhất định được cho là tự miễn [12]. Đối với kháng nguyên bạch cầu người (Human Leukocyte Antigen-HLA), Silveira và đồng tác giả [38] tìm thấy một tỷ lệ cao của HLA-B12 và kiểu gen đơn bội A9-B5 và A28-B35 ở trẻ em bị teo đường mật, đặc biệt là ở những người không có dị tật liên quan. Yuasa và đồng tác giả [39] phát hiện sự kết hợp của kiểu gen đơn bội HLA-DR2 và HLA-A24-B52-DR2 với teo đường mật chu sinh. Kết quả này có thể chỉ ra rằng một hoặc nhiều gen gần locus HLA đóng vai trò trong sinh bệnh học của teo đường mật, hoặc HLA- DR2 trên bề mặt của đường mật có thể được liên kết trực tiếp với cơ chế sinh lý bệnh của bệnh này. Hơn nữa, trẻ em bị teo ống dẫn mật ngoài (extrahepatic bile duct atresia – EHBA) có thể bị viêm gan tự miễn trong giai đoạn sau ghép [40].

Ngoài ra, yếu tố di truyền cũng đóng vai trò quan trọng trong việc hình thành bệnh teo đường mật bẩm sinh. Nhiều đột biến gen liên quan đến bệnh teo đường mật bẩm sinh đã được phát hiện trên các gen như *ADD3, EFEMP1, MMP7* và *ARF6.* Với sự phát triển của công nghệ GWAS và giải trình tự gen thế hệ mới, một số biến thể số lượng bản sao (CNV) và đa hình nucleotide đơn đã được phát hiện ở bệnh nhân teo đường mật bẩm sinh và có liên quan đến cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh.

Sự bất hoạt di truyền của các yếu tố hạt nhân tế bào gan (HNF), như HNF-1β [41] và HNF6 [42] được xem là gây ra các dị thường hình thái trong ống dẫn mật và trong túi mật. HNF-1β được kết hợp với hàng nghìn ống dẫn mật nội tạng, trong khi HNF6 có liên quan đến dị dạng ống tấm và sự hiện diện của u nang nội tạng [43]. Vai trò của các yếu tố biểu sinh trong sinh bệnh học của bệnh teo đường mật cũng đã được đánh giá bởi Zhang và đồng tác giả [44]. Trong nghiên cứu này, các tác giả tiến hành điều tra sự biểu hiện gen ở trẻ em bị teo đường mật phôi chu sinh hoặc phôi. Các tác giả tìm thấy sự gia tăng các gen liên quan đến các yếu tố điều tiết nhiễm sắc tố (*SMARCA-1, HDAC3* và *RYBP*). Các gen này ảnh hưởng đến quá trình biểu sinh, vì vậy các tác giả suy đoán chúng đóng vai trò tiềm năng trong sinh bệnh tắc nghẽn đường mật.

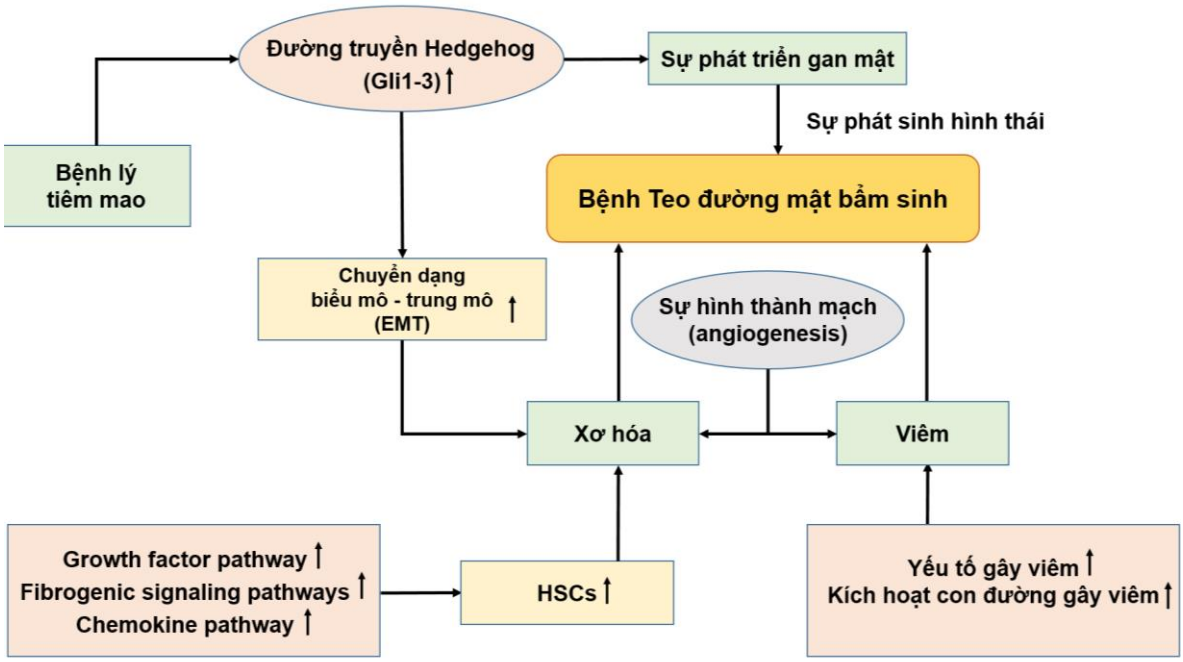
Nghiên cứu trong nhiều năm qua cho thấy sự biến đổi gen liên quan đến bệnh như đột biến trên các gen *MIF* (migration inhibitory factor), *CD14, ICAM1* (intercellular adhesion molecule-1, *APM1* (adiponectin) và *ITGB2* (*CD18*). Nghiên cứu mới đây của Sangkhathat và đồng tác giả (2018) đã xác định được đột biến trên các gen *JAG1, ABCC2, UGT1A1, MYO5B, RFX6, MLL2, ERCC4, KCNH1, ABCB11* trên 20 bệnh nhân mắc teo đường mật bẩm sinh bằng phương pháp giải trình tự toàn bộ vùng mã hóa [45]. Bên cạnh đó, đa hình trên một số gen như: *CFC1, VEGF, ADD3, STAT4, GPC1, ARF6* cũng được xác định là làm tăng nguy cơ mắc bệnh teo đường mật bẩm sinh [46–50].

Sự xuất hiện của các đột biến liên quan đến gen inversin (inv) trong nhiễm sắc thể số 4, được biết đến là nguyên nhân gây ra sự phát triển bất thường của hệ thống gan mật ở chuột cũng đã được nghiên cứu ở bệnh nhân teo đường mật bẩm sinh nhưng không có đột biến nào được xác định là có ảnh hưởng rõ ràng [51]. Tương tự như vậy, một số các nghiên cứu đã chỉ ra mối liên quan giữa một số alen HLA và bệnh teo đường mật bẩm sinh [52] trong khi những nghiên cứu khác không xác nhận những phát hiện này [53]. Như vậy, mặc dù có bằng chứng về ảnh hưởng của một số bất thường di truyền trong cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh nhưng vai trò chính xác của chúng trong bệnh này vẫn chưa rõ ràng và vẫn đang được nghiên cứu.

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. Teo đường mật có liên quan đến một số hội chứng bẩm sinh, dị thường nhiễm sắc thể và đa hình gen ở một số quần thể cụ thể. 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]. Các gen được xác định liên quan đến sự phát triển ống dẫn mật trong quá trình phát triển phôi như: *Invs*, *Hes1*, *Hnf6*, *Hnf1b*, *Foxf1*, *Sox17*, *Lgr4* và *Pdx1*.

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). Mối liên hệ giữa các con đường gây bệnh được thể hiện ở Hình 1.2, các biến thể di truyền tiềm năng có liên quan đến bệnh teo đường mật bẩm sinh được thể hiện ở Bảng 1.1. 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. Nghiên cứu trên cỡ mẫu lớn hoặc trên mô hình động vật là cần thiết để chứng minh vai trò và tầm quan trọng của các gen đến sinh bệnh học của teo đường mật bẩm sinh



Hình 1.2: Mối liên hệ giữa các con đường và cơ chế gây bệnh teo đường mật bẩm sinh [56]

Trong giai đoạn tạo phôi, hình thái bất thường của gan mật có thể dẫn đến bệnh teo đường mật bẩm sinh đi kèm các hội chứng bất thường khác. Teo đường mật bẩm sinh có nang – một loại teo đường mật bẩm sinh không phổ biến cũng có thấy hình thái ống mật bất thường như một đặc điểm bệnh lý của bệnh. Ba gen nhạy cảm tham gia vào quá trình phát triển và hình thành ống mật được xác minh bằng nghiên cứu trên mô hình động vật là *ADD3, ARF6* và *GPC1*. Tuy nhiên, đối với bệnh teo đường mật bẩm sinh không đi kèm các hội chứng khác, tác động của quá trình phát triển gan mật đến bệnh là chưa rõ ràng.

Teo đường mật bẩm sinh là một bệnh có đặc trưng bởi sự tăng cường biểu hiện của các yếu tố gây viêm trong gan của bệnh nhân teo đường mật bẩm sinh có sự gia tăng các phân tử bám dính viêm và cytokine có thể đóng vai trò là yếu tố dự báo khả năng sống sót tự nhiên của gan sau phẫu thuật Kasai [57,58]. Tuy nhiên, ngay cả với phẫu thuật Kasai, gan vẫn có thể tiếp tục bị viên dai dẳng dẫn đến xơ gan. Trong sinh thiết gan của bệnh nhân teo đường mật bẩm sinh, chúng ta có thể phát hiện một số thay đổi trong tế bào lympho. Các tế bào viêm khác, bao gồm bạch cầu ái toan, tế bào plasma và đại thực bào cũng có mặt [59]. Một số alen của đa hình trên các gen liên quan đến khía cạnh khác nhau của con đường viêm đã được chứng minh là có liên quan đến nguy cơ mắc bệnh. Trong những năm gần đây, sự điều hòa của các gen gây viêm,

như *MMP7, HMGB1, MIF* và các gen khác đã được phát hiện trong nhiều nghiên cứu. Tuy nhiên, sự biểu hiện cao của các gen này có thể là nguyên nhân hoặc hậu quả của tổn thương gan ứ mật mạn tính ở bệnh nhân teo đường mật bẩm sinh. Đa hình nucleotide đơn trong một số gen như *MIF, ITGB2, ICAM1* cung cấp bằng chứng mạnh mẽ rằng những gen này là gen nhạy cảm với teo đường mật bẩm sinh và sự rối loạn điều hòa của chúng dẫn đến kiểu hình viêm của bệnh nhân teo đường mật bẩm sinh.Do đó, việc điều hòa các gen tiền viêm có thể đóng một vai trò quan trọng trong sinh bệnh học teo đường mật bẩm sinh.

Sự hình thành mạch không chỉ thúc đẩy quá trình xơ hóa gan mà còn đưa nhiều chất dinh dưỡng đến đường mật để kích thích sự tăng sinh của ống dẫn. Xơ hóa là một đặc điểm chính của gan ở bệnh nhân teo đường mật bẩm sinh. Giống như tình trạng viêm, xơ hóa cũng có thể tồn tại trong gan ở bệnh nhân teo đường mật bẩm sinh ngay cả sau phẫu thuật Kasai và tình trạng này có thể dẫn đến xơ gan. Kích hoạt tế bào hình sao ở gan (HSC) đóng vai trò quan trọng trong tiến trình xơ hóa. Ngoài HSC, quá trình chuyển dạng biểu mô – trung mô (Epithelial Mesenchymal Transformation - EMT) cũng góp phần gây ra tình trạng xơ hóa gan thông qua việc tạo ra các nguyên bào sợi cơ sản xuất collagen. Sự tham gia của quá trình truyền tín hiệu yếu tố tăng trưởng, các con đường truyền tín hiệu Fibrogenic và các con đường chemokine đều được phát hiện là có liên quan đến việc kích hoạt HSC. Các đa hình nucleotide di truyền có thể dẫn đến sự hình thành mạch và xơ hóa liên quan đến teo đường mật bẩm sinh.

Lông mao sơ cấp đại diện cho một cấu trúc dưới tế bào được bảo tồn tiến hóa tồn tại ở hầu hết các loại tế bào trong cơ thể con người. Lông mao sơ cấp phối hợp nhiều con đường truyền tín hiệu khác nhau, bao gồm cả những con đường được điều chỉnh bởi Hh, thụ thể kết hợp protein G (GPCR), WNT, tyrosine kinase thụ thể (RTK) và TGFβ/protein hình thái xương (BMP), các yếu tố liên quan đến việc kiểm soát quá trình phát triển, độ dẻo của mô và chức năng của các cơ quan. Bằng cách cảm nhận những thay đổi trong môi trường ngoại bào, lông mao sơ cấp phối hợp các tầng tín hiệu được khuếch đại tiếp theo trong toàn bộ tế bào. Lông mao nằm trên màng đỉnh của lòng trong của biểu mô đường mật và có chức năng sinh lý như là chất thúc đẩy dòng mật và cảm biến các thành phần mật và độ thẩm thấu. Các nhà nghiên cứu nhận thấy

rằng, lông mao nguyên phát bị giảm có chọn lọc trong các tế bào đường mật ngoài gan của bệnh nhân teo đường mật bẩm sinh. Các giả thuyết cho rằng đột biến gen liên quan đến sự phát triển lông mao có thể dẫn đến sự phát triển kiểu hình teo đường mật bẩm sinh theo hai cơ chế sinh học khác nhau. Đầu tiên, người ta cho rằng cấu trúc và chức năng của lông mao tế bào mật bị khiếm khuyết có thể tự nó gây ra rối loạn điều hòa đường Hh, thúc đẩy quá trình sửa chữa mô bị rối loạn chức năng và dẫn đến viêm gan và xơ hóa [60]. Lông mao không thể thiếu trong con đường tín hiệu Hh, tương tác với các thành phần khác nhau tại các điểm khác nhau của tầng tín hiệu để điều chỉnh quá trình tái tạo và sửa chữa gan. Thứ hai, có khả năng lông mao bị khiếm khuyết sẽ làm tổn hại chức năng bảo vệ của các tế bào đường mật cholangiocytes ở trẻ sơ sinh chưa trưởng thành trước sự tác động của axit mật, dẫn đến tổn thương gan mãn tính, cũng có thể kích hoạt tín hiệu Hh. Lông mao tế bào đường mật ngắn hơn, sai hướng hoặc ít phong phú hơn thường được quan sát thấy trong một số nghiên cứu ở cả bệnh nhân teo đường mật bẩm sinh có hội chứng và không có hội chứng [61] trong khi đó hoạt động của con đường tín hiệu Hh được chứng minh là có liên quan đến khả năng sống sót không bị bệnh vàng da của bệnh nhân teo đường mật bẩm sinh trong một nghiên cứu khác [62].

Bảng 1.1: Các biến thể gen và vai trò của chúng trong cơ chế bệnh sinh của bệnh teo đường mật bẩm sinh

Gen	Vị trí	Dân tộc	dbSNP	Loại biến thể	Alen	Minor Alen	MAF	Sự phát triển gan mật	Xơ hóa	Viêm	Bệnh lý tiêm mao	Tài liệu tham khảo
ADD3	10q25.1-q25.2	Trung Quốc	rs17095355	Vùng intron	C/T	T	0.494–0.551	X	X	-	-	[63,64]
GPC1	2q27.3	Người da trắng	CNV(deletion)	-	-	-	-	X	-	X	-	[65,66]
		Trung Quốc	rs2292832	Vùng intron	T/C	C	0.223	X	-	X	-	[67]
			rs3828336	Vùng intron	C/T	T	0.06	X	-	X	-	
		Trung Quốc	rs6707262	5′upstream	T/C	C	0.438	X	-	X	-	[64]
			rs6750380	5′upstream	T/C	C	0.434	X	-	X	-	
ARF6	14q21.3	Người da trắng	rs3126184	3′-UTR	T/C	T	0.2857	X	-	-	-	[68]
MIR499	20q11.22	Trung Quốc	rs3746444	Vùng intron	A/G	G	0.258	-	X	X	-	[69]
		Ai Cập					0.46	-	X	X	-	[70]
ITGB2	21q22.3	Trung Quốc	rs1160263	3′-UTR	G/T	T	0.191	-	-	X	-	[71]
ICAM1	19p13.2	Thổ Nhĩ Kỳ	rs1799969	Sai nghĩa	G/A	A	0.632	-	-	X	-	[72]
MIF	22q11.23	Thổ Nhĩ Kỳ	rs755622	Chưa xác định	G/C	C	0.611	-	X	X	-	[73]

Gen	Vị trí	Dân tộc	dbSNP	Loại biến thể	Alen	Minor Alen	MAF	Sự phát triển gan mật	Xơ hóa	Viêm	Bệnh lý tiêm mao	Tài liệu tham khảo
		Ai Cập	rs5844572	Chưa xác định	insATTC	-	-	-	X	X	-	[74]
VEGFA	6p12	Trung Quốc	rs3025039	3′-UTR	C/T	T	0.089–0.154	-	X	X	-	[75,76]
PDGFA	7p22	Trung Quốc	rs9690350	Vùng intron	G/C	C	0.37	-	X	X	-	[77]
EFEMP1	2p16.1	Trung Quốc, Châu Âu	rs10865291	Vùng intron	G/A	A	0.43	-	X	-	-	[78]
CD14	5q31.3	Trung Quốc	rs2569190	Vùng intron	T/C	C	0.03	X	-	X	-	[79]
NOTCH2	1p12	Trung Quốc	rs835576	3′ UTR	A/G	G	0.42	X	-	-	-	[79]
ADIPOQ	3q27.3	Thái Lan	rs1501299	Vùng intron	G/T	T	0.189	X	-	X	-	[80]
IL18*	11q22.2-q22.3	Trung Quốc	rs187238	Promoter	C/G	G	0.14074	-	-	X	-	[81]
			rs1946518	Promoter	T/G	G	0.408969	-	-	X	-	[81]
PKD1L1*	7p12.3	Đa chủng tộc	rs139293796	Vùng exon	G/A	A	0.00275	X	-	-	X	[82]
			rs148011149	Vùng exon	A/G	G	0.000288	X	-	-	X	[82]
			rs776420484	Vùng exon	A/G	G	3.00E-04	X	-	-	X	[82]

Gen	Vị trí	Dân tộc	dbSNP	Loại biến thể	Alen	Minor Alen	MAF	Sự phát triển gan mật	Xơ hóa	Viêm	Bệnh lý tiêm mao	Tài liệu tham khảo
			rs139858574	Vùng exon	G/A	A	0.000308	X	-	-	X	[82]
			rs528302390	Vùng exon	AG/-	-	0.000912	X	-	-	X	[82]
			rs143005953	Vùng exon	G/A	A	0.006978	X	-	-	X	[82]
			rs140456142	Vùng exon	G/A	A	0.001154	X	-	-	X	[82]
			rs770832954	Vùng exon	C/T	T	4.00E-05	X	-	-	X	[82]
			rs752673990	Vùng exon	G/C	C	4.00E-05	X	-	-	X	[82]
PKHD1	6p12.3-p12.2	Trung Quốc	rs137852950	Sai nghĩa	A/C	C	1.00E-04	-	-	-	X	[83]
			rs139127465	Sai nghĩa	A/G	G	1.90E-04	-	-	-	X	[83]

**: MAF được lấy từ cơ sở dữ liệu NCBI*

1.2.1. Gen ABCB11

Gen *ABCB11* nằm trên nhánh dài của nhiễm sắc thể số 2, tại vị trí q31.1. Gen có độ dài 110.54 kb, gồm 28 exon. Gen mã hóa cho protein BSEP có chức năng bơm bài xuất muối mật, vận chuyển muối mật ra khỏi gan. Muối mật là một thành phần của mật, được sử dụng để tiêu hóa chất béo. Muối mật được sản xuất bởi các tế bào gan và sau đó được vận chuyển ra khỏi tế bào để tạo ra mật. Sự giải phóng muối mật từ các tế bào gan là rất quan trọng đối với việc tiết mật bình thường. Cùng với các protein đa kháng thuốc như MDR1 (gen *ABCB1*), MDR3 (gen *ABCB4*) và MRP2 (gen *ABCC2*), bơm bài xuất muối mật (BSEP, gen *ABCB11*) là những chất vận chuyển dòng mật ra ngoài quan trọng nhất. Các biến thể di truyền di truyền xảy ra trên các chất vận chuyển họ ABC này có thể điều chỉnh biểu hiện của chất vận chuyển và có thể liên quan đến nhiều bệnh gan ứ mật di truyền và mắc phải. Ứ mât, tăng bilirubin máu xảy ra khi có sự rối loạn cân bằng nội môi của bilirubin và axit mật do các protein tham gia vào quá trình tổng hợp, hấp thu, giải độc và vận chuyển axit mật và bilirubin. Siêu họ của các chất vận chuyển cassette gắn ATP ở gan (ABC) đóng một vai trò quan trọng trong việc xử lý axit mật, bilirubin và nhiều loại thuốc. Những khiếm khuyết trong tổng hợp hoặc chức năng của BSEP dẫn đến giảm bài tiết muối mật và tích tụ muối mật. Tầm quan trọng của BSEP được thể hiện qua tình trạng ứ mật di truyền gây ra bởi đột biến gen ở *ABCB11* từ mức độ ứ mật dai dẳng và nghiêm trọng cần ghép gan (như đã thấy ở ứ mật trọng gan tiến triển có tính chất gia đình) đến các dạng ứ mật nhẹ hơn (như tình trạng ứ mật trong gan tái phát lành tính - benign recurrent intrahepatic cholestasis (BRIC)). Các biến thể trên gen này gây ảnh hưởng đến hoạt động của protein, làm ức chế quá trình vẫn chuyển muối mật ra khỏi gan dẫn đến tình trạng ứ mật trong gan. Các biến thể trong gen *ABCB11* được biết là có liên quan đến teo đường mật bẩm sinh ngoài gan ở bệnh nhân Thái Lan [45]. Tuy nhiên, cho đến nay, chưa có nghiên cứu nào được thực hiện để đánh giá vai trò của biến thể này đối với tính mẫn cảm với bệnh teo đường mật bẩm sinh.

Đột biến gen *ABCB11* có thể gây ứ mật trong gan tái phát lành tính loại 2 (Benign Recurrent Intrahepatic Cholestasis - BRIC2). Người bệnh đôi khi bị suy giảm bài tiết mật dẫn đến ngứa dữ dội và vàng da vàng mắt. Đôi khi, người mắc BRIC2 sau đó được chẩn đoán mắc một bệnh lý nghiêm trọng hơn được

gọi là ứ mật trong gan gia đình tiến triển loại 2 khi các triệu chứng trở nên tệ hơn. Đột biến gen *ABCB11* có thể làm giảm 40–50% quá trình vận chuyển muối mật, do đó muối mật tích tụ trong gan đến mức gây ra các triệu chứng của BRIC2. Người ta chưa rõ nguyên nhân vì sao các giai đoạn bắt đầu hoặc kết thúc.

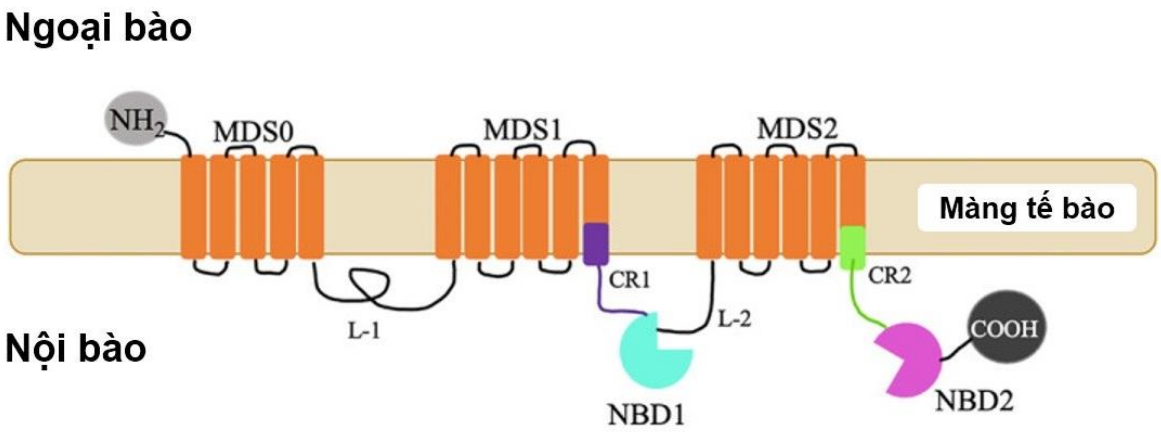
Hơn 100 đột biến gen *ABCB11* gây ra một dạng bệnh gan nghiêm trọng thường dẫn đến suy gan, bệnh được gọi là ứ mật trong gan gia đình tiến triển loại 2 (Progressive Familial Intrahepatic Cholestasis type 2 - PFIC2). Trong bệnh này, đột biến gen *ABCB11* làm giảm 70–100% hoạt động vận chuyển muối mật ra khỏi gan. Do đó, muối mật tích tụ trong tế bào gan đến mức gây ra các dấu hiệu và triệu chứng liên quan. Đột biến dẫn đến một phiên bản protein ngắn và không hoạt động hoặc hoàn toàn không sản xuất protein (xu hướng liên quan đến bệnh gan nặng xuất hiện sớm hơn trong đời). Người không có protein BSEP hoạt động dường như cũng gặp nguy cơ cao phát triển một loại ung thư gan được gọi là ung thư biểu mô tế bào gan (hepatocellular carcinoma).

Cùng với gen *ABCB4*, các biến thể di truyền ở gen *ABCB11* được cho là có liên quan đến khoảng 20% bệnh nhân bị ứ mật nặng trong gan khi mang thai, khi các hormone thai kỳ thúc đẩy sự phát triển của tình trạng ứ mật, điển hình là trong tháng thứ hai của thai kỳ khi nồng độ estrogen cao nhất [84,85]. Phụ nữ mang đột biến gen *ABCB11* có nguy cơ phát triển tình trạng ứ mật trong gan thai kỳ (intrahepatic cholestasis of pregnancy). Người bệnh thường bị giảm bài tiết mật (ứ mật) và ngứa trong ba tháng cuối của thai kỳ, truy nhiên những biểu hiện này sẽ biến mất sau khi em bé chào đời. Với những phụ nữ mang đột biến trên gen *ABCB11* dạng dị hợp tử, họ vẫn còn đủ chức năng BSEP để tiết đủ mật trong phần lớn thời gian. Nồng độ estrogen và progesterone cao trong thai kỳ càng làm suy yếu chức năng của BSEP, dẫn đến suy giảm bài tiết mật và các dấu hiệu của ứ mật trong gan khi mang thai. Tuy nhiên, nhiều yếu tố có thể góp phần vào nguy cơ phát triển bệnh lý phức tạp này.

1.2.2. Gen ABCC2

Gen *ABCC2* mã hóa protein MRP2 (multidrug resistance protein 2). Protein này được biểu hiện ở phần ống của tế bào gan, có chức năng là chất vận

chuyển anion hữu cơ đa đặc hiệu và bilirubin ra khỏi tế bào gan và vào mật. Về mặt cấu trúc, protein MRP2 có ba miền trải dài màng (MSD-0, MSD-1 và MSD-2), hai miền liên kết nucleotide (NBF-1 và NBF-2) và các vòng tế bào chất (L- 1, L- 2) (Hình 1.3).



Hình 1.3: Cấu trúc protein MRP2 mã hóa bởi gen ABCC2

Miền mở rộng màng MSD0 gồm 5 vòng xoắn xuyên màng ở vùng đầu NH_2, cùng với 2 miền mở rộng màng khác MSD1 và MSD2 với 6 chuỗi xoắn xuyên màng và 2 miền liên kết nucleotide (NBD1 và NDB2). Các vòng tế bào chất L-1 và L2. Các vùng kết nối CR1 và CR2 [86].

Protein MRP2 có chung cấu trúc với các protein họ MRP khác, nhưng có một miền trải dài màng bổ sung (MSD-0) và một trình tự đặc biệt trong vùng COOH. Các trình tự đặc biệt trong vùng MSD-0 và L-0 này tương ứng là cần thiết cho hoạt động MRP2 và vận chuyển màng huyết tương.

Một số đột biến làm ảnh hưởng đến chức năng của protein MRP2 được mô tả là nguyên nhân gây ra hội chứng Dubin-Johnson, có liên quan đến tình trạng tăng bilirubin trong máu.

1.2.3. Gen ADD3 và XPNPEP1

Gen *XPNPEP1* (X-Prolyl Aminopeptidase 1), nằm ở vị trí 10q25.1 và mã hóa XPNPEP1, góp phần làm thoái hóa bradykinin và chất P [87]. Các sản phẩm thoái hóa như vậy tham gia vào phản ứng viêm ở nhiều cơ quan và bệnh tật. Gen *ADD3* (Adducin 3), nằm ở 10q25.1-q25.2 và mã hóa Gamma-adducin, thúc đẩy sự kết nối và bám dính tế bào - tế bào trong tế bào gan và tế bào biểu mô đường mật, điều chỉnh sự di chuyển và bám dính của tế bào cũng như tham

gia vào quá trình co bóp của tế bào mật như một protein liên kết với màng-tế bào. *XPNPEP1* đã được chứng minh là biểu hiện ở tất cả các mô được kiểm tra cho đến nay, bao gồm tuyến tụy, tim, cơ, thận, gan, phổi và não.

Tsai và đồng tác giả nhận thấy *ADD3* có mức độ biểu hiện cao hơn ở gan của bệnh nhân teo đường mật bẩm sinh so với người thường. Tang và đồng tác giả sử dụng phương pháp phân tích *ADD3a* để tìm hiểu cơ chế bệnh sinh teo đường mật bẩm sinh ở cá ngựa vằn. Kết quả cho thấy sự bất hoạt của *ADD3* gây ra các khiếm khuyết trong gan và làm giảm chức năng đường mật trong khi bất hoạt *XPNPEP1* không dẫn tới những phản ứng như vậy. Các kết quả tương tự cũng được quan sát thấy ở các đột biến đồng hợp tử trên gen *ADD3a* và người ta cũng cho rằng *ADD3* có thể dẫn đến teo đường mật bẩm sinh bằng cách ảnh hưởng đến con đường tín hiệu Hedgehog [88]. Hedgehog là một trong số các con đường truyền tín hiệu được sử dụng để liên lạc giữa các tế bào trong quá trình phát triển [89] và đóng một vai trò quan trọng trong quá trình hình thành hình thái phôi thai. Đã có báo cáo chỉ ra rằng teo đường mật bẩm sinh có liên quan đến hoạt động của Hedgehog, tác động sau đó có thể kích thích sự chuyển dạng trung mô trong biểu mô ống mật (EMT) và dẫn đến dị tật ống mật. Các nhà khoa học đã chứng minh rằng các protein phiên mã hạt nhân phản ứng Hh, Gli1-3, được biểu hiện mạnh hơn trong gan của bệnh nhân teo đường mật bẩm sinh, điều đó có nghĩa là hoạt động truyền tín hiệu Hh ở bệnh nhân teo đường mật bẩm sinh tăng cao. Gli-3 có mối liên quan với EMT, nghĩa là tín hiệu Hh tham gia vào quá trình xơ hóa ở teo đường mật bẩm sinh. Kết hợp lại với nhau, các nghiên cứu được mô tả ở trên cung cấp bằng chứng mạnh mẽ về sự liên quan của *ADD3* trong cơ chế bệnh sinh của teo đường mật bẩm sinh.

1.2.4. Gen GPC1

Gen *GPC1* nằm ở vị trí 2q37.3 trên nhiễm sắc thể 2 mã hóa glypican 1, một loại protein điều chỉnh các con đường truyền tín hiệu khác nhau bao gồm các yếu tố tăng trưởng nguyên bào sợi (FGF), yếu tố tăng trưởng nội mô mạch máu A (*VEGFA*), yếu tố tăng trưởng biến đổi β (TGF-β), WNT và Hedgehog để điều chỉnh các phản ứng viêm trong tín hiệu nội bào [90]. Glypicans đóng vai trò quan trọng trong việc làm trung gian cho sự phát triển và phân nhánh của ống dẫn trong việc phát triển thận ở chuột [91] và khí quản ở chi *Drosophila*

[92]. Các quá trình này tương tự như sự phát triển đường mật ở cá ngựa vằn và do đó có thể có chung các đặc điểm điều hòa phân tử với sự phát triển đường mật của động vật có vú.

Hơn nữa, kết quả thu được bằng cách tạo ra các khiếm khuyết trên gen *GPC1* ở mô hình cá ngựa vằn, cho thấy rằng hình thái bất thường ở các ống mật trong và ngoài gan của các phôi cá ngựa vằn này đã dẫn đến giảm sự hấp thu của túi mật, đồng thời ngăn chặn con đường Hedgehog [47]. Những phát hiện như vậy cho thấy rõ *GPC1* có liên quan đến cơ chế bệnh sinh của teo đường mật bẩm sinh thông qua con đường truyền tín hiệu Hh. Các biến thể liên quan đến teo đường mật bẩm sinh trong gen GPC1 không chỉ là CNV. Vào năm 2016, kết quả từ một nghiên cứu được thực hiện với 134 bệnh nhân teo đường mật bẩm sinh và 618 đối chứng trong một đoàn hệ Trung Quốc chỉ ra rằng rs2292832 có liên quan đáng kể đến nguy cơ mắc teo đường mật bẩm sinh, trong khi rs3828336 dường như chỉ có tác động không nhiều đến nguy cơ mắc bệnh teo đường mật bẩm sinh [67]. Hơn nữa, vào năm 2020, Bai và động tác giả tìm thấy hai đa hình nucleotide đơn là rs6707262 và rs6750380 trong gen *GPC1* có liên quan đáng kể với bệnh teo đường mật bẩm sinh trong một đoàn hệ Trung Quốc khác [64]. Như vậy, thông qua khả năng tạo ra hình thái bất thường của đường mật, *GPC1* đóng một vai trò quan trọng trong sinh bệnh học của teo đường mật bẩm sinh.

1.2.5. Gen ARF6

Năm 2016, Ningappa và đồng tác giả phát hiện ra rằng mức độ biểu hiện thấp của *ARF6* dẫn đến gan nhỏ lại và các đường mật trong gan tương đối hẹp [50]. Hơn nữa, sự biểu hiện của gen này bị giảm ở các mẫu teo đường mật bẩm sinh. Gen này nằm trên 14q21.3, mã hóa ARF6 với tư cách là một protein gắn với GTP, có liên quan đến việc vận chuyển protein điều chỉnh việc tái cấu trúc khung tế bào [87]. Gen *ARF6* tham gia vào quá trình phát triển tế bào gan của thai nhi. Sau khi gen *ARF6* bị phá hủy trong phôi, có sự hình thành bất thường của dây gan. Hai đa hình nucleotide đơn rs3126184 và rs10140366 trên gen *ARF6* được phát hiện là có liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh [68]. Hình thái bất thường đường mật có thể xảy ra khi đáp ứng với biểu hiện *ARF6* thấp, là nguyên nhân hình thành bệnh teo đường mật bẩm sinh.

1.2.6. Gen MIR499

Năm 2016, Shan và đồng tác giả là nhóm nghiên cứu đầu tiên cho thấy đa hình nucleotide rs3746444 trong *MIR499* (MicroRNA-499) có liên quan đến bệnh teo đường mật bẩm sinh và tương quan với sự tăng biểu hiện của các yếu tố gây viêm (IL-6, TNF-α) trong huyết thanh và gan của bệnh nhân teo đường mật bẩm sinh. Alen G của đa hình rs3746444 có liên quan đến mức độ viêm. Nghiên cứu trên một nhóm bệnh nhân người Ai Cập với 100 bệnh nhân teo đường mật bẩm sinh, 100 trường hợp mắc bệnh gan ứ mật không phải teo đường mật bẩm sinh và 100 người đối chứng khỏe mạnh, Gawish và đồng tác giả chỉ ra rằng rs3746444 có tương quan với mức độ xơ hóa và kết quả sau khi phẫu thuật Kasai ở bệnh nhân teo đường mật bẩm sinh [70]. Gen *MIR499,* nằm ở 20q11.22, mã hóa các micro RNA đóng vai trò như một loại RNA không mã hóa liên quan đến việc điều hòa biểu hiện gen sau phiên mã ở sinh vật đa bào, chủ yếu liên quan đến việc ảnh hưởng đến sự ổn định và dịch mã của mRNA. Những miRNA này góp phần vào sự hình thành và chức năng của gan và ống mật.

1.2.7. Gen ITGB2 và ICAM1

Năm 2013, Zheng và cộng sự. đưa ra bằng chứng đầu tiên cho thấy đa hình nucleotide rs1160263 trong gen *ITGB2* có liên quan chặt chẽ với sinh bệnh học teo đường mật bẩm sinh. Họ cũng phát hiện *ITGB2* được biểu hiện quá mức ở gan của bệnh nhân teo đường mật bẩm sinh thông qua sinh thiết gan [71]. Nằm ở vị trí 21q22.3, gen này mã hóa chuỗi integrin beta, Integrinβ2 (CD18), kết hợp với một số chuỗi alpha khác nhau để tạo ra các dị vòng integrin khác nhau và được biểu hiện ở nhiều loại bạch cầu, tham gia vào quá trình kết dính và di chuyển bạch cầu và hỗ trợ apoptosis của đại thực bào và thực bào bạch cầu trung tính. Integrinβ2 (CD18), là tiểu đơn vị của kháng nguyên liên quan đến tế bào lympho-1 (LFA-1) và đại thực bào-1 (MAC-1), làm trung gian cho hoạt động ái lực cao trong nhiều loại tế bào biểu hiện các phân tử bám dính giữa các tế bào (ICAM, đặc biệt là ICAM-1). Gen *ICAM1* được biểu hiện quá mức ở các ống mật ngoài và trong gan của bệnh nhân teo đường mật bẩm sinh.

ICAM1 nằm ở 19p13.2 và là gen mã hóa glycoprotein bề mặt tế bào (ICAM1/CD56) thường được biểu hiện ở tế bào nội mô và hệ thống miễn dịch

[87]. Bằng cách này, *ICAM1* chủ yếu tham gia vào các con đường truyền tín hiệu miễn dịch và viêm bao gồm tín hiệu interferon và interleukin [93]. Alen A của SNP rs1799969 trong *ICAM1* có mối liên quan với nguy cơ mắc bệnh teo đường mật bẩm sinh [72].

1.2.8. Nhóm gen liên quan đến quá trình xơ hóa

Theo nghiên cứu trên quần thể người Hán, gen *VEGFA* được phát hiện là có sự tăng cường biểu hiện trong gan của bệnh nhân teo đường mật bẩm sinh và đa hình nucleotide đơn rs3025039 trong gen này có sự liên quan đến sự phát triển của bệnh [75]. Liu và các đồng tác giả cũng phát hiện được 2 đa hình nucleotide đơn rs10434 và rs3025039 có liên quan đáng kể đến nguy cơ mắc teo đường mật bẩm sinh thông qua phân tích Haplotype [76]. Gen *VEGFA* nằm ở vị trí 6p12 và mã hóa cho protein tham gia vào quá trình tạo mạch và làm trung gian tính thấm của mạch máu, cũng như tham gia vào quá trình viêm thông qua việc thúc đẩy sự di chuyển và cảm ứng sự kích hoạt của bạch cầu đơn nhân. Hơn nữa, sự giải phóng *VEGFA* bởi các tế bào đường mật có thể hoạt động như một tín hiệu liên kết sự phát triển của ống và động mạch trong gan do đó thúc đẩy quá trình hình thành mạch quanh đường mật. Đáng chú ý nhất là tính đa hình rs3025039, đặc biệt là alen C có liên quan đến teo đường mật bẩm sinh và có thể làm tăng mức độ nhạy cảm đối với căn bệnh này. Theo đó, thông qua khả năng thúc đẩy sự hình thành mạch và tình trạng viêm, *VEGFA* được cho là gen nhạy cảm với teo đường mật bẩm sinh.

Tình trạng viêm và xơ hóa có tương quan với nồng độ TGF-β1 và PDGF trong gan của bệnh nhân teo đường mật bẩm sinh. Năm 2020, Liu và đồng tác giả đã chứng minh rằng alen G trong rs9690350 của gen *PDGFA* làm tăng đáng kể tính nhạy cảm với bệnh teo đường mật bẩm sinh ở trẻ sơ sinh [77]. *PDGFA* nằm ở 7p22 và mã hóa cho protein yếu tố tăng trưởng đóng một vai trò thiết yếu trong việc điều hòa sự phát triển của phôi, tăng sinh tế bào, di chuyển tế và sống sót của tế bào. Nghiên cứu của Yang và đồng tác giả cho thấy sự biểu hiện vượt mức của PDGFR với trình trạng xơ hóa ở gan của chuột [94]. Thông qua việc thúc đẩy quá trình viêm và xơ hóa, *PDGFA* có thể tham gia vào các cơ chế góp phần tạo gây nên teo đường mật bẩm sinh.

Gen *EFEMP1* nằm ở 2p16.1 và mã hóa một thành viên thuộc họ fibulin của glycoprotein ma trận ngoại. Protein ma trận ngoại bào này liên quan đến quá trình tái tạo mô, hình thành cơ quan và có thể đóng vai trò trong sự kết dính và di chuyển tế bào. Chen và đồng tác giả đã chỉ ra rằng *EFEMP1* có thể là một gen tiềm năng về tính nhạy cảm với teo đường mật bẩm sinh khi nghiên cứu trên 343 bệnh nhân teo đường mật bẩm sinh và 1.716 người khỏe mạnh [78]. Nghiên cứu này phát hiện ra rằng biểu hiện của gen này tăng rõ rệt ở bệnh nhân teo đường mật bẩm sinh và 3 đa hình nucleotide đơn rs10865291, rs6761893 và rs727878 có mối liên quan cao đến nguy cơ mắc bệnh. Gen *EFEMP1* được biểu hiện cao ở các tế bào đường mật bổ sung và tế bào cơ trơn mạch máu ở bệnh nhân teo đường mật bẩm sinh cũng như trong các bệnh ứ mật khác, trong khi nó chỉ được biểu hiện ở mức độ bình thương trong các tế bào cơ trơn mạch máu ở những người khỏe mạnh.

1.2.9. Một số gen liên quan đến sự hình thành lông mao bất thường

Gen *MAN1A2* nằm ở 1p12, mã hóa MAN1A2 và được biểu hiện chủ yếu ở nhau thai và tinh hoàn. Nghiên cứu của So và đồng tác giả năm 2020 cho thấy giảm mức độ biểu hiện của gen *MAN1A2* thông qua knockdown dẫn đến sự hình thành mạng lưới đường mật kém, rối loạn chức năng đường mật trong túi Kupffer, rối loạn điều hòa trong biểu hiện của EGF, TGF và con đường Hedgehog cùng với sự giảm biểu hiện của các gen liên quan đến sự phát triển lông mao [95]. Nghiên cứu này cũng phát hiện thấy việc loại bỏ đồng thời cả hai gen *MAN1A2* và *ARF6* gây ra những bất thường trong sự phát triển của ống mật, trong khi không có tác động nào được phát hiện nếu chỉ loại bỏ một trong hai gen. Như đã trinh bày ở trên, gen *ARF6* được xác định là gen nhạy cảm với teo đường mật bẩm sinh, những tác động tổng hợp của *ARF6* với *MAN1A2* cho thấy *MAN1A2* có thể ảnh hưởng đến sự phát triển vi lông mao và tín hiệu EGFR cũng như điều chỉnh sự hình thành mạng lưới đường mật trong gan bằng cách tương tác với *ARF6*. Tuy nhiên, cho đến nay, không có biến thể nào trong gen này được tìm thấy ở các bệnh nhân teo đường mật bẩm sinh.

Gen *PKHD1* nằm trên 6p12.3-p12.2 và mã hóa fibrocystin có thể hoạt động trong quá trình biệt hóa ống góp (collecting - duct) và đường mật cũng như tham gia vào quá trình điều hòa sự tăng sinh tế bào đường mật [96]. Các

biến thể trong gen này có liên quan đến một dạng bệnh thận đa nang nghiêm trọng và trong một số trường hợp là bệnh teo đường mật bẩm sinh trong gan. Bệnh lý này được đặc trưng bởi sự giãn nở của các ống góp và xơ hóa đường mật do dị tật tấm ống. Dị tật tấm ống động mạch cũng là một đặc điểm được công nhận của teo đường mật bẩm sinh dựa trên kiểm tra mô học, điều này cho thấy cơ chế bệnh sinh tương tự như các bệnh lý lông mao khác.

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. Gen *JAG1* nằm trên nhiễm sắc thể 20 mã hóa protein Jagged1 là một trong năm protein bề mặt tế bào tương tác với bốn thụ thể trong con đường tín hiệu Notch ở động vật có vú. Một khi tương tác JAG1-NOTCH (thụ thể-phối tử) diễn ra, nó sẽ khởi động một loạt các phản ứng truyền tín hiệu (Notch signal) ảnh hưởng đến các chức năng của tế bào. Tín hiệu Notch kiểm soát cách một số loại tế bào phát triển trong phôi thai đặc biệt là tế bào liên quan đến tim, gan, mắt và cột sống. Gen *JAG1* được cho là có liên quan đến cơ chế bệnh sinh của teo đường mật bẩm sinh với tần suất cao của đa hình nucleotide đơn trên gen *JAG1* ở các bệnh nhân. Đột biến trong gen *JAG1* dẫn đến những khiếm khuyết trong protein Jagged-1 ở màng tế bào sẽ ngăn cản sự tương tác của nó với protein Notch và ngăn chặn tín hiệu của tế bào. Thiếu tín hiệu Notch gây ra những sai sót trong quá trình phát triển dẫn đến thiếu hoặc hẹp ống dẫn mật trong gan, dị tật tim, các đặc điểm trên khuôn mặt đặc biệt và những thay đổi ở các bộ phận khác của cơ thể. Nghiên cứu của Kohsaka và các đồng tác giả trên 102 bệnh nhân BA thể teo đường mật ngoài gan đã xác định được 9 đột biến sai nghĩa trên gen *JAG1* có liên quan đến mức độ trầm trọng của bệnh ở các bệnh nhân này [97]. Năm 2008, Davit-Spraul và đồng tác giả đã phát hiện được biến đổi dị hợp tử

c.433G>A (p.Ala145Thr) trên exon 5 của gen *CFC1* ở 5 bệnh nhân teo đường mật bẩm sinh.

Adducin là các protein bao gồm các tiểu đơn vị khác nhau được gọi là adducin alpha, beta và gramma. Các protein ADD tham gia vào quá trình truyền tín hiệu, di chuyển tế bào và kết dính giữa các tế bào [98]. Trong đó, gramma – adducin được mã hóa bởi gen *ADD3* nằm trên nhiễm sắc thể số 10 thường được biểu hiện trong các tế bào biểu mô gan và đường mật. Protein này đóng vai trò quan trọng trong việc tái cấu trúc màng tế bào. Có nhiều nghiên cứu về ảnh hưởng của các biến thể trên gen *ADD3* tới bệnh teo đường mật bẩm sinh được thực hiện. Tuy nhiên, các phát hiện về mức độ biểu hiện gen *ADD3* hay tác động của đột biến trên gen này tới bệnh nhân teo đường mật bẩm sinh là chưa nhất quán. Thực hiện nghiên cứu trên 305 bệnh nhân và 571 đối chứng, Garcia-Barcelo và đồng tác giả xác định SNP rs2501577 là một locus nhạy cảm với bệnh teo đường mật bẩm sinh, cụ thể là làm tăng nguy cơ mắc bệnh ở quần thể người Thái Lan [63]. Tuy nhiên, Tsai và đồng tác giả lại không tìm thấy sự gia tăng nguy cơ mắc bệnh đối với locus này trong nghiên cứu về người da trắng Caucasians [48]. Vivian Tang và các đồng tác giả nhận thấy rằng biểu hiện *ADD3* giảm ở các bệnh nhân teo đường mật bẩm sinh gây ra sự kích hoạt đường truyền tín hiệu Hedgehog (Hedgehog signaling pathway), hậu quả là đường mật bị dị tật [88]. Mặt khác, nghiên cứu này của Tang cũng chứng minh rằng dự giảm điều hòa đồng thời, chứ không phải riêng lẻ của gen *ADD3* và *GPC1* đã tạo ra sự suy giảm chức năng hấp thu và các khiếm khuyết ở đường mật. Teng-Fei Li và đồng tác giả nghiên cứu đa hình nucleotide đơn rs2501577 trên gen *ADD3* cho thấy rằng mặc dù cơ chế tác động của *ADD3* đến bệnh teo đường mật bẩm sinh hiện chưa rõ ràng nhưng SNP này có liên quan đáng kể đến bệnh teo đường mật bẩm sinh [99].

Sử dụng phương pháp TaqMan assay, nghiên cứu của Lee và đồng tác giả [100] đã xác định được ba đa hình (-2578 A>C, -634 G>C, +936 C>T) trên gen *VEGF* và bốn đa hình (-1615 C>T, -183 C>T, +874 A>T, +2197 A>G) trên gen *IFNG* ở bệnh nhân người Đài Loan có liên quan đến sự mẫn cảm với bệnh teo đường mật bẩm sinh.

Năm 2010, Leyva-Vega và đồng tác giả nghiên cứu trên 5 bệnh nhân teo đường mật bẩm sinh và đã xác định được các đột biến mất đoạn lớn ở dạng dị hợp tử trên nhiễm sắc thể 2q37.3 bao gồm *GPC1* có thể là các ứng viên cho sự mẫn cảm với bệnh teo đường mật bẩm sinh [66]. Trong một nghiên cứu về đột biến CNV với 61 trường hợp mắc teo đường mật bẩm sinh và 5.088 đối chứng trong đoàn hệ người Mỹ, Cui và đồng tác giả [47] đã xác định được sự liên quan của gen *GPC1* và sự mẫn cảm với bệnh teo đường mật bẩm sinh. Đa hình rs17095355 C>T trên vùng xen kẽ giữa hai gen *ADD3* và *XPNPEP1* có liên quan đến bệnh teo đường mật bẩm sinh trên bệnh nhân người Thái Lan [101], người Trung Quốc [46]. Nghiên cứu của Zeng và đồng tác giả năm 2014 cho thấy hai đa hình rs17095355 C>T và rs10509906 G>C trên gen *ADD3* có liên quan đến sự mẫn cảm với bệnh của 134 bệnh nhân teo đường mật bẩm sinh người Trung Quốc [49]. Ningappa và đồng tác giả [50] tìm thấy mối liên quan giữa các đa hình nucleotide đơn trên gen *ARF6* liên quan đến bệnh teo đường mật bẩm sinh trong nhóm bệnh nhi da trắng ở Mỹ. Năm 2018, Laochareonsuk và các đồng tác giả cũng cho thấy hai đa hình rs17095355 và rs2501577 làm tăng nguy cơ mắc bệnh teo đường mật bẩm sinh ở các bệnh nhân người Thái Lan [45].

Sử dụng phương pháp giải trình tự vùng gen mã hóa, Lam và đồng tác giả đã phát hiện hai biến thể hiếm gặp rs137852950 và rs139127465 trên gen *PKHD1* trên các bệnh nhân teo đường mật bẩm sinh người Trung Quốc và Việt Nam [83]. Mặc dù cho đến nay không có thêm biến thể nào trên gen này được phát hiện nhưng *PKHD1* vẫn được cho là đóng vai trò quan trọng trong sinh bệnh học của teo đường mật bẩm sinh thông qua khả năng ảnh hưởng đến sự phát triển của vi lông mao. Năm 2019, Berauer và cộng sự. tìm thấy 9 biến thể hiếm (rs139293796, rs148011149, rs776420484, rs139858574, rs528302390, rs143005953, rs140456142, rs770832954, rs752673990) trong gen *PKD1L1* ở bệnh nhân teo đường mật bẩm sinh có đi kèm các dị tật khác (BASM) [82]. Tần số alen nhỏ (MAF) của các biến thể này là <0,05. Nghiên cứu này cũng chỉ ra rằng *PKD1L1* được biểu hiện mạnh mẽ ở tế bào đường mật so với mô gan. Thông qua ảnh hưởng đến bệnh lý lông mao, sự phát triển của gan mật và ứ mật, *PKD1L1* được coi là gen ứng cử viên mới cho teo đường mật bẩm sinh.

Một nghiên cứu trên nhóm bệnh nhân người Thổ Nhĩ Kỳ phát hiện rằng alen C của đa hình nucleotide đơn rs755622 trên gen *MIF* có thể có mối liên quan đến tính nhạy cảm của bệnh teo đường mật bẩm sinh [73]. Gen *MIF* mã hóa yếu tố ức chế di chuyển đại thực bào, đóng vai trò là yếu tố tiền viêm tham gia vào quá trình tự miễn dịch để đáp ứng với mầm bệnh. Môt nghiên cứu khác được tiến hành trên bệnh nhân người Ai Cập năm 2017 đã không tìm thấy mối liên quan nào giữa hai đa hình nucleotide đơn rs755622 và rs5844572 với nguy cơ mắc bệnh teo đường mật bẩm sinh. Tuy nhiên nghiên cứu này phát hiện thấy sự liên quan giữa rs5844572 với mức độ xơ hóa và tốc độ tiến triển của bệnh teo đường mật bẩm sinh [74]. Những kết quả khác biệt về ảnh hưởng của các đa hình nucleotide trên gen *MIF* đến bệnh có thể do các nghiên cứu được thực hiện trên những quần thể khác nhau. Tuy nhiên, điều đó cũng cho thấy rằng *MIF* có thể là gen tiềm năng về tính nhạy cảm di truyền của bệnh teo đường mật bẩm sinh.

1.3.2. Tình hình nghiên cứu ở Việt Nam

Ở Việt Nam, tỷ lệ mắc bệnh khá cao ước tính 1:2400 trẻ đẻ ra sống [102]. Mỗi năm tại Bệnh viện Nhi trung ương điều trị cho khoảng 70 bệnh nhân nhi mắc mới với teo đường mật bẩm sinh. Thời gian vàng để phẫu thuật là trước 8 tuần. Do đó, biện pháp hữu hiệu hiện nay là chẩn đoán xác định teo đường mật bẩm sinh khi tiến hành phẫu thuật. Các triệu chứng lâm sàng và cận lâm sàng của bệnh teo đường mật bẩm sinh dễ bị chẩn đoán nhầm với thiểu sản đường mật hoặc nang ống mật chủ. Các nghiên cứu trên bệnh nhân teo đường mật bẩm sinh ở Việt Nam chủ yếu tập trung vào các yếu tố lâm sàng, phương pháp tăng cường hiệu quả điều trị bệnh [58]. Nhiều nghiên cứu nhằm cải thiện lưu lượng dẫn mật, tăng khả năng sống sót sau khi thực hiện phẫu thuật Kasai, giảm chi phí ghép gan được thực hiện [103]. Năm 2017, Max Bolun Liu khảo sát trên 287 bệnh nhân teo đường mật bẩm sinh người Việt Nam với độ tuổi trung bình khi được chẩn đoán là 2,3 tháng tuổi, trong đó 149 bệnh nhân được điều trị không dùng thủ thuật Kasai, 138 bệnh nhân thực hiện phẫu thuật Kasai. Kết quả cho thấy nhóm được điều trị không dùng thủ thuật Kasai có tỷ lệ sống sót sau 1 và 2 năm lần lượt là 52% và 28% (tỷ lệ sống trung bình là 6,6 tháng); ngược lại, nhóm được điều trị bằng thủ thuật Kasai có tỷ lệ sống sót sau 1 và 2

năm không cần cấy ghép là 84% và 71%. Đặc biệt, không có bệnh nhân nào được điều trị bằng ghép gan do thiếu chương trình ghép gan ở Việt Nam.

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 [104,105]. Năm 2020, Kien và đồng tác giả thực hiện nghiên cứu trên 41 bệnh nhân nhi mắc teo đường mật bẩm sinh từ đó xác định 28 đột biến (17 đột biến liên kết NST X, 6 đột biến de novo và 5 đột biến đồng hợp tử) trong 25 gen [106].

Năm 2021, Lam và đồng tác giả giải trình tự toàn bộ vùng gen mã hóa của 89 bệnh nhân teo đường mật bẩm sinh Hồng Kông và Việt Nam, tập trung vào phân tích và sàng lọc những biến thể hiếm gặp, làm thay đổi chức năng protein trên những gen được biểu hiện ở mô gan/đường mật. Nhóm nghiên cứu đã phát hiện được nhiều biến thể de novo phân bố trong 239 gen, phần lớn là những gen liên quan đến vi lông mao [83]. Điểm đặc biệt là trong số 239 gen phát hiện đột biến, không có gen nào DC báo cáo là có liên quan đến bệnh teo đường mật bẩm sinh DC nghiên cứu trước đó [107]. Nhóm nghiên cứu này cũng phát hiện ra rằng việc loại bỏ các gen này trong mô hình tế bào người và cá ngựa vằn có thể dẫn đến những khiếm khuyết ở đường mật. Kết quả nghiên cứu này cho thấy đột biến gen vi lông mao có thể góp phần vào sự phát triển của kiểu hình teo đường mật bẩm sinh.

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 gia đình. Các bệnh nhân này được chẩn đoán mắc bệnh teo đường mật bẩm sinh bởi bác sĩ khoa Gan mật, Bệnh viện Nhi Trung ương, Hà Nội, Việt Nam. Trong nghiên cứu này, mẫu máu của 516 đối tượng được thu thập bởi các bác sĩ, bao gồm 266 trẻ mắc bệnh teo đường mật bẩm sinh và 250 mẫu đối chứng (trẻ không mắc các bệnh về gan mật). Mỗi đối tượng nghiên cứu được lấy 2ml máu toàn phần. Mẫu máu được bảo quản trong ống chống đông EDTA và giữ ở -20^{o}C.

Trong số 266 bệnh nhân nhi mắc teo đường mật bẩm sinh, có 05 bệnh nhân được chọn lọc để giải trình tự toàn bộ vùng gen mã hóa, trong đó có 02 cặp anh/chị em ruột trong cùng một gia đình.

Việc nghiên cứu trên bệnh nhân được tuân thủ đúng y đức theo tiêu chuẩn của tuyên bố Helsinki của Hội đồng Y khoa thế giới và được Hội đồng đạo đức trong nghiên cứu y sinh học Viện Nghiên cứu hệ gen thông qua và đồng ý cho phép lấy mẫu, nghiên cứu trên người theo quyết định số 6/QD-NCHG năm 2019.

2.1.1. Tiêu chuẩn lựa chọn mẫu

2.1.1.1 Mẫu bệnh

Mẫu bệnh nhân được lựa chọn theo các tiêu chí sau:

- Trẻ trong độ tuổi từ 5 ngày tuổi đến 11 tuổi được chẩn đoán là mắc bệnh teo đường mật bẩm sinh, đã được phẫu thuật Kasai

- Có các biểu hiện vàng da, phân bạc màu, ứ mật

- Transaminase tăng

- Siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no.

- Người nhà bệnh nhân đồng ý tham gia nghiên cứu

2.1.1.2. Mẫu đối chứng

- Trẻ trong độ tuổi từ 5 ngày tuổi đến 11 tuổi không mắc teo đường mật bẩm sinh

- Không mắc các bệnh liên quan đến gan mật, không có tình trạng ứ mật

- Người nhà bệnh nhân đồng ý tham gia nghiên cứu

2.1.1.3. Tiêu chuẩn loại trừ

- Trẻ mắc các hội chứng khác liên quan đến gan mật

- Người nhà bệnh nhân không đồng ý tham gia nghiên cứu.

2.1.2. Cỡ mẫu nghiên cứu

Cỡ mẫu cần thiết cho nghiên cứu được tính dựa trên tần số xuất hiện của các đa hình nucleotide đơn, được xác định theo công thức sau:

$$n = \frac{4C}{(lnOR)^2 p(1-p)}$$

Trong đó:

n: số lượng bệnh nhân cần thu thập

C: hằng số liên quan đến các sai số loại I và loại II. Lấy giá trị α = 0,05 và β = 0,20 thì giá trị C = 7,85

OR: tỉ số nguy cơ

p: tần số xuất hiện đa hình gen

Trong nghiên cứu này chúng tôi phân tích các đa hinh nucleotide đơn rs2287622 trên gen *ABCB11,* rs927344 trên gen *ABCC2* và rs1815930 trên gen *MYO5B* đến nguy cơ mắc bệnh. Trên thế giới chưa có nghiên cứu nào đánh giá mối liên quan của ba đa hình nucleotide đơn này đến nguy cơ mắc bệnh teo đường mật bẩm sinh. Do đó, để đảm bảo cỡ mẫu phù hợp, có thể bao phủ được cả ba đa hình trên, chúng tôi lựa chọn tỉ số nguy cơ OR và tần số xuất hiện alen đa hình p theo nghiên cứu của Besheer và đồng tác giả về đánh giá mối liên quan của đa hình rs2287622 trên gen *ABCB11* với tình trạng xơ hóa tiến triển và nồng độ axit mật [108] với tỉ số nguy cơ OR = 2,58 và tần số xuất hiện alen đa hình p = 0,2. Các chỉ số OR và p được lấy theo nghiên cứu này có giá trị

nhỏ, đảm bảo cỡ mẫu lớn và bao phủ được cả ba đa hình nucleotide đơn cần nghiên cứu. Thay các giá trị vào công thức:

$$n = \frac{4 \times 7{,}85}{(ln2.58)^2 \times 0{,}2 \times 0{,}8} = 219{,}58$$

n = 219,58, làm tròn thành 220. Ngoài ra, để tránh các sai số trong quá trình thu thập mẫu và phân tích, nghiên cứu này sẽ tiến hành lấy nhiều hơn 20% so với cỡ mẫu tối thiểu. Như vậy, số mẫu bệnh nhân cần thu thập là 266 bệnh nhân và nhóm đối chứng là 250 người.

2.2. Hóa chất và trang thiết bị

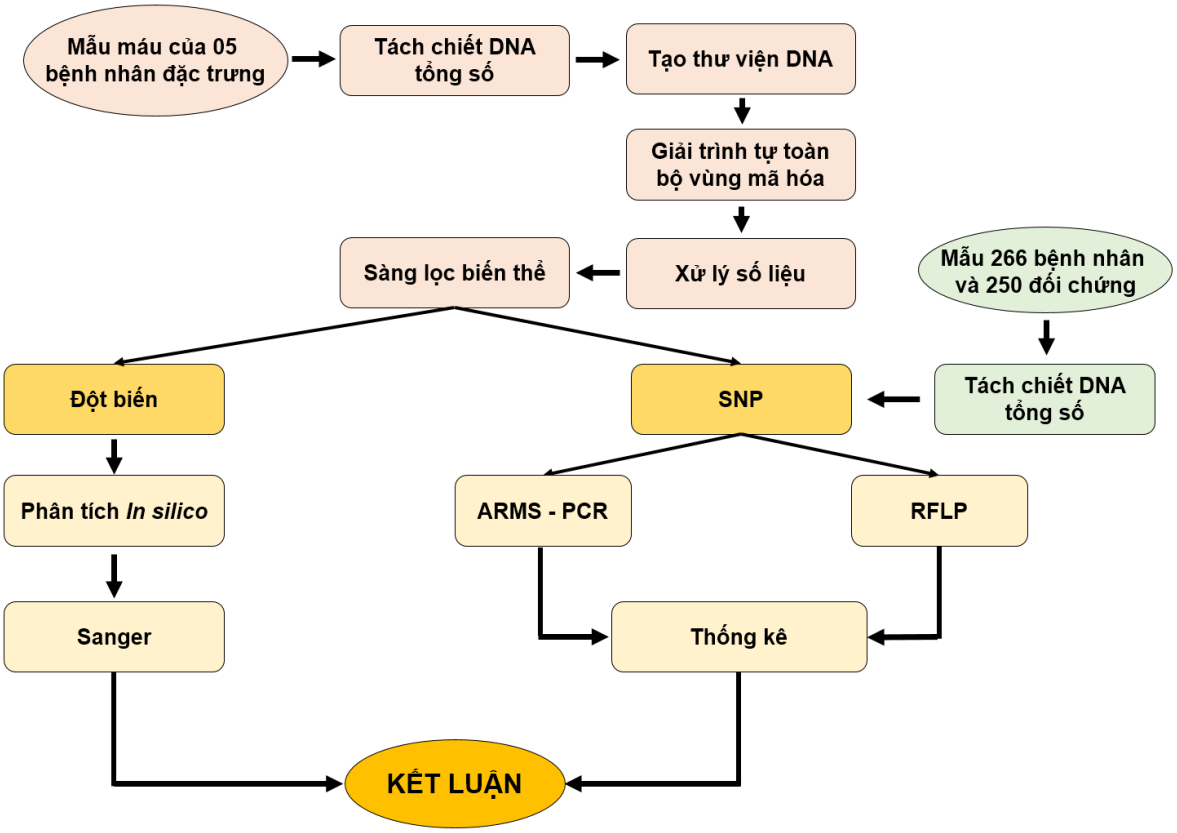
Các hóa chất chính được sử dụng trong nghiên cứu bao gồm: Kit tách chiết DNA tổng số từ mẫu máu QIAamp DNA Mini Kit (QIAGEN, Đức), bộ kit tạo thư viện giải trình tự gen thế hệ mới Kit Agilent SureSelect Target Enrichment (Agilent, Mỹ) Tag DNA polymerase (Thermo Fisher Scientific™), dNTP 10X mix (Thermo Fisher Scientific™), marker DNA 1kb (Thermo Fisher Scientific™), Marker DNA 100 bp (Thermo Fisher Scientific™), agarose (Merk, Mỹ), enzyme *Hae*III (Thermo Fisher Scientific™).

Các thiết bị chính được sử dụng trong nghiên cứu bao gồm: Máy giải trình tự gen thế hệ mới NextSeq 500 (Illumina, Mỹ), máy giải trình tự ABI3500 Genetic Analyzer (Thermo Fisher Scientific, Mỹ), máy ly tâm (Eppendorf, Đức), Máy soi gel và chụp ảnh DigiDoc-It® Imaging System (Ultra-violet production, Mỹ), máy PCR (Eppendorf, Đức), máy đo huỳnh quang Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Mỹ), cân kỹ thuật (Ohaus, Mỹ).

2.3. Phương pháp nghiên cứu

Để thực hiện các mục tiêu đã đề ra, nghiên cứu được tiến hành theo sơ đồ như Hình 2.1. Trong nghiên cứu này, mẫu máu toàn phần của 266 bệnh nhân nhi mắc bệnh teo đường mật bẩm sinh và 250 người không mắc bệnh được thu thập bởi các bác sĩ bệnh viện Nhi Trung ương. DNA tổng số của người tham gia nghiên cứu được tách chiết bằng kit QIAgen. Trong số 266 bệnh nhân mắc bệnh, 05 bệnh nhân có kiểu hình đặc trưng được lựa chọn để giải trình tự toàn bộ vùng gen mã hóa. Dữ liệu giải trình tự của các bệnh nhân này được phân tích để tìm kiếm đột biến và sàng lọc các điểm đa hình đơn nucleotide từ đó

đánh giá ảnh hưởng đến khả năng gây bệnh và mối liên quan với mức độ mẫn cảm bệnh của các biến thể tìm được.



Hình 2.1: Sơ đồ 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 QIAamp DNA Mini của hãng QIAGEN, quy trình tách chiết bao gồm các bước chính như sau:

Hút 20 µl proteinase K (20 mg/ml) vào mỗi ống Eppendorf sau đó bổ sung 200 µl máu và 200 µl AL buffer. Hỗn hợp được trộn đều bằng máy vortex trong 15 giây rồi ủ ở 56^{o}C trong 20 phút. Thêm 200 µl ethanol 100% và trộn đều sau đó chuyển hỗn hợp lên cột tách chiết và li tâm với tốc độ 8000 rpm trong 1 phút. Chuyển cột sang ống mới và bổ sung 500 µl AW1, li tâm với tốc độ 8000 rpm trong 1 phút. Chuyển cột sang ống thu mới, bổ sung 500 µl AW2 và li tâm ở tốc độ 12000 rpm trong 3 phút. Loại bỏ dịch qua cột, tiếp tục li tâm ở tốc độ 12000 rpm trong 1 phút để làm khô cột. Chuyển cột sang ống Eppendorf mới và bổ sung 200 µl AE vào giữa cột. Ủ ở nhiệt độ phòng trong 2 phút rồi li tâm ở 8000 rpm trong 1 phút để thu được DNA.

DNA tổng số được điện di kiểm tra chất lượng trên gel agarose 0,8% và đo quang phổ để xác định nồng độ và độ tinh sạch và bảo quản ở -20^{o}C.

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. Các giai đoạn chính của việc chuẩn bị thư viện NGS là: (i) phân mảnh DNA thành các đoạn nhỏ có kích thước mong muốn, phù hợp với thiết bị giải trình tự; (ii) gắn các oligonucleotide cần thiết cho quá trình đọc trình tự bao gồm các index để phân biệt DNA từng mẫu với nhau và adapter để gắn các đoạn DNA vào bề mặt flowcell, (iii) khuếch đại vùng mã hóa; (iv) kiểm tra chất lượng sản phẩm thư viện trước khi giải trình tự bằng máy 2100 Bioanalyzer.

Chất lượng của thư viện DNA bao gồm 2 thông số chính: sự phân bố kích thước và nồng độ thư viện. Kích thước các đoạn DNA đích trong thư viện là một tham số quan trọng trong việc xây dựng thư viện NGS. Có ba phương pháp phân mảnh DNA thường đươc sử dụng là phương pháp vật lý, hóa học và sử dụng enzyme. Trong đó phương pháp vật lý (siêu âm) hoặc enzyme (ví dụ, các loại cocktail endonuclease không đặc hiệu và phản ứng phân đoạn transposase) là những phương pháp phổ biến nhất [109]. Nghiên cứu này phân mảnh DNA bằng cách sử dụng phương pháp vật lý, thiết bị Covaris (Covaris, Woburn, MA) được thực hiện để thu được các đoạn DNA trong phạm vi 100–5000 bp, trong khi Covaris g-TUBE được sử dụng cho phạm vi 6–20 Kbp cần thiết cho việc tạo thư viện mate-pair.

Nồng độ thư viện là một yếu tố quan trọng ảnh hưởng đến quá trình giải trình tự. Nồng độ thư viện DNA cần được xác định bằng các phương pháp đo huỳnh quang dành cho DNA sợi kép (dsDNA), tránh các phương pháp xác định nồng độ DNA sợi đơn (ssDNA). Đồng thời, quá trình xác định nồng độ thư viện DNA không nên sử dụng các phương pháp quang phổ (NanoDrop) hay UV vì những phương pháp này đo nồng độ axit nucleic trong mẫu do đó RNA và các tạp chất lẫn trong mẫu có thể làm sai lệch kết quả. Qubit® Fluorometer (Life Technologies) và Quantus™ Fluorometer (Promega Corporation) là

những phương pháp đo nồng độ DNA sợi kép được tất cả các hệ thống giải trình tự NGS khuyến cáo sử dụng. Hãng Illumina khuyến cáo nên sử dụng thuốc nhuộm QuantiFluor® dsDNA của Promega để đo nồng độ DNA chính xác hơn khi chuẩn bị thư viện bằng kit DNA Nextera Rapid Capture. Thuốc nhuộm QuantiFluor® dành riêng cho DNA sợi kép giúp hạn chế sự ảnh hưởng của DNA sợi đơn, RNA, protein và các hợp chất gây nhiễu.

2.3.2.2. Giải trình tự bằng Illumina NextSeq 500

Thư viện DNA đạt chất lượng được giải trình tự bằng thiết bị Illumina NextSeq 500. Đây là hệ thống giải trình tự DNA thế hệ mới sử dụng công nghệ đọc trình tự theo nguyên lý tổng hợp (sequencing by synthesis – SBS) tương tự như phương pháp Sanger. Trong đó DNA polymerase tổng hợp chuỗi DNA hình thành bằng cách sử dụng dNTP gắn vào đầu 3’ của chuỗi DNA đang tổng hợp theo nguyên tắc bổ sung. Tuy nhiên, đối với giải trình tự thế hệ mới, thay vì giải trình tự một đoạn đơn lẻ, kỹ thuật này cho phép giải trình tự với một lượng lớn các đoạn DNA khác nhau song song tại cùng một thời điểm, từ đó tiết kiệm thời gian và cho lượng dữ liệu đầu ra vô cùng lớn so với phương pháp Sanger cũ. Quá trình giải trình tự bao gồm hai bước chính là tạo cụm (tạo cluster) và giải trình tự. Trong quá trình tạo cụm, mỗi sợi DNA được giữ lại trên bề mặt thiết bị giải trình tự (flowcell) bằng adapter đã được gắn trước đó. Mỗi sợi DNA sau khi được khuếch đại sẽ tạo thành một cụm DNA có trình tự giống hệt nhau để sử dụng cho quá trình giải trình tự. Ở bước giải trình tự, dNTP có gắn các tín hiệu huỳnh quang tương ứng với 4 loại nucleotide, Tín hiệu huỳnh quang của từng nulceotide được ghi lại trong quá trình tổng hợp theo nguyên tắc bổ sung trên sợi DNA khuôn.

2.3.3. Xác định và chú giải biến thể

Quá trình phân tích, xác định và chú giải biến thể được thực hiện bằng các phần mềm cài đặt trên hệ thống máy tính hiệu năng cao tại Viện Nghiên cứu hệ gen, Viện Hàn lâm Khoa học và Công nghệ Việt Nam với hệ điều hành Centos 7.0.

Dữ liệu giải trình tự được lưu trữ dưới dạng file Fastq. Trong lĩnh vực giải trình tự gen, dữ liệu dạng Fastq đã phát triển như một định dạng tệp tiêu chuẩn để lưu trữ và chia sẻ dữ liệu đọc trình tự. File Fastq cung cấp nhiều thông

tin hơn so với định dạng Fasta, cho phép lưu trữ điểm chất lượng dạng số được liên kết với mỗi nucleotide trong một trình tự. Định dạng dữ liệu Fastq ban đầu được viện nghiên cứu Wellcome Trust Sanger sử dụng để mô tả thông tin về đoạn trình tự DNA và chất lượng khi đọc trình tự của đoạn DNA đó là một số nguyên được mã hóa dưới dạng mã ASCII hiện nay đã trở thành định dạng dữ liệu chuẩn đầu ra của các thiết bị giải trình tự thế hệ mới [110].

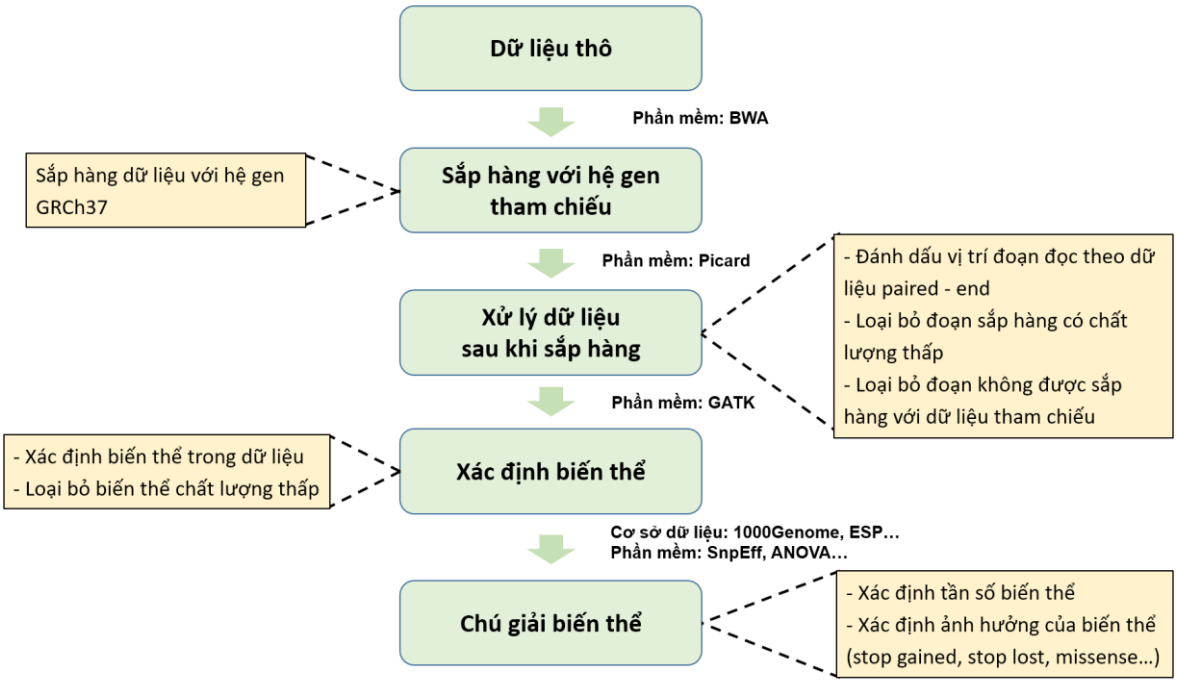
Với các thiết bị giải trình tự thế hệ mới, độ chính xác của mỗi nucleotide được thể hiện qua giá trị Q, được tính theo công thức:

$$Q = -10\log_{10}\frac{p}{1-p}$$

Trong đó giá trị p là xác suất nucleotide đó bị đọc sai. Như vậy, giả sử Q=20 thì xác suất nucleotide đó bị thiết bị đọc sai là 1%. Như vậy, giá trị Q càng cao thì chất lượng đọc trình tự càng tốt. Trong giải trình tự toàn bộ vùng mã hóa hoặc giải trình tự toàn bộ hệ gen để phát hiện và sàng lọc đột biến, hai giá trị quan trọng để đánh giá chất lượng đọc trình tự là Q20 (số lượng nucdeotide có điểm chất lượng lớn hơn 20) và Q30 (số lượng nucdeotide có điểm chất lượng lớn hơn 20).

Chất lượng của dữ liệu giải trình tự toàn bộ vùng mã hóa được đánh giá bằng phần mềm FastQC. Phần mềm này thống kê và đánh giá chất lượng của dữ liệu dựa trên một số tiêu chí quan trọng như thông tin tổng quan về dữ liệu như thông tin về máy giải trình tự, tổng số đoạn trình tự, và trung bình tỉ lệ nucleotide loại GC của toàn bộ dữ liệu…

Quy trình phân tích và xử lý số liệu bằng các phần mềm tinh sinh học được mô tả như hình (Hình 2.2).



Hình 2.2: Quy trình xử lý số liệu bằng các công cụ tin sinh học

Dữ liệu trình tự được sắp xếp và so sánh với ngân hàng gen người (GRCh37) bằng phần mềm BWA 0.7.10 [111].). Phần mềm BWA sử dụng thuật toán BWT có hiệu quả tính toán cao hơn, đòi hỏi ít bộ nhớ hơn do đó những phần mềm sử dụng thuật toán này ngày càng được phát triển rộng rãi [112]. BWA là một bộ công cụ có chức năng sắp hàng các đoạn trình tự với một hệ gen tham chiếu lớn. BWA hỗ trợ phân tích dữ liệu từ nhiều nền tảng giải trình tự khác nhau như Illumina hay AB SOLiD. Đánh giá trên cả dữ liệu mô phỏng và dữ liệu thực tế cho thấy công cụ này nhanh hơn một số phần mềm khác từ 10 đến 20 lần trong khi vẫn đạt được độ chính xác tương tự.

Kết quả sắp hàng được lưu trữ dưới dạng SAM (Sequence Alignment Map). Để cải thiện hiệu suất và giảm bộ nhớ lưu trữ, tập tin SAM nén thành định dạng BAM mà vẫn giữ đầy đủ thông tin như file SAM ban đầu. File SAM/BAM được tiền xử lý nhằm giảm thời gian phân tích, tăng độ chính xác cho bước gọi đột biến. Theo thống kê, công cụ BWA tạo ra hơn 15% sai sót trong bước sắp hàng đối với những biến thể dạng chèn hoặc xóa. Nếu không có bước hiệu chỉnh dữ liệu, những sai lệch này sẽ tạo ra những biến thể giả, ảnh hưởng tới độ tin cậy của kết quả. Do đó, việc xử lý dữ liệu này nhằm mục đích loại bỏ những đoạn trình tự không được sắp hàng hoặc những được

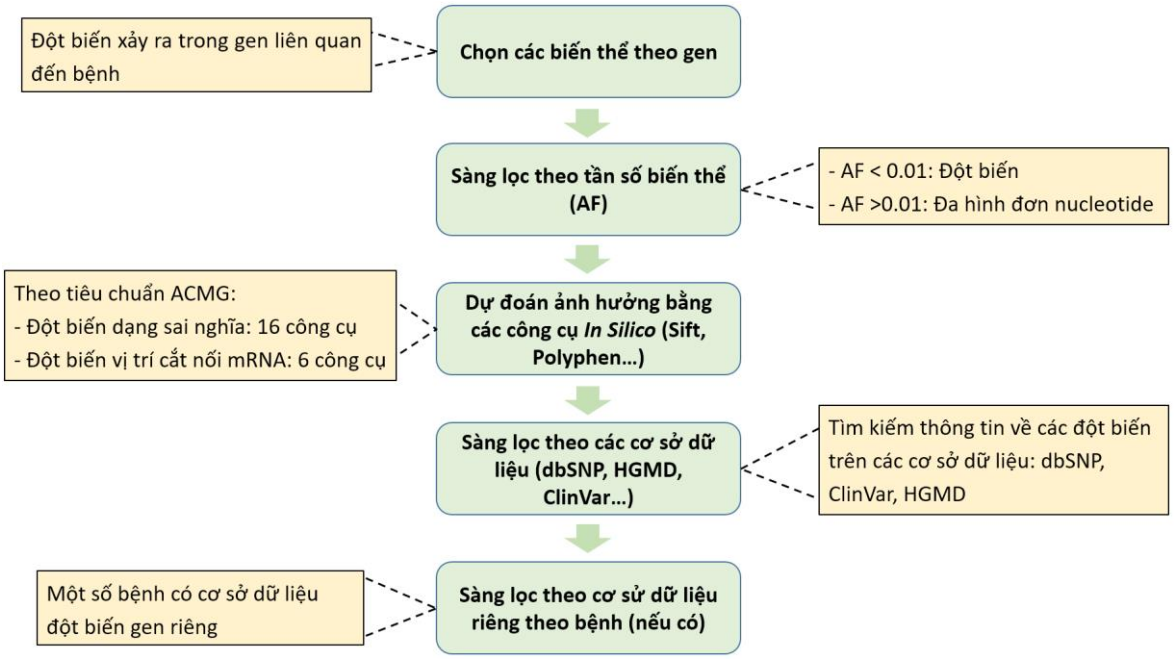
sắp hàng có độ tin cậy thấp qua đó giảm tỉ lệ dương tính giả trong dữ liệu đặc biệt là với những biến thể dạng chèn hoặc xóa. Ngoài ra, bước phân tích này cũng có ý nghĩa chuyên sâu về mặt tính toán, góp phần làm giảm tỉ lệ sai sót khi xác định biến thể và giảm thời gian cần thiết để phân tích dữ liệu. Trong nghiên cứu này, file SAM/BAM được xử lý bằng công cụ Picard nhằm hiệu chỉnh chất lượng của từng nucleotide, sắp xếp, loại bỏ những đoạn lặp….

Biến thể được phát hiện bằng phần mềm Genome Analysis Toolkit v3.4 [113]. GATK được chứng minh là tốt nhất đối với dữ liệu được tạo ra từ thiết bị Illumina [114]. Bộ công cụ GATK được sử dụng để gọi biến thể của đa hình nucleotide đơn (SNPs) và các biến thể chèn, xóa đoạn nhỏ (indels). GATK Best Practice là quy trình phát hiện biến thể từ dữ liệu NGS được sử dụng rộng rãi do quy trình này thể hiện kết quả tốt với độ chính xác cao trên dữ liệu tạo ra từ thiết bị giải trình tự Illumina. Liu và cộng sự đã so sánh hiệu suất của một số quy trình phát hiện biến thể khác nhau trên nhiều loại dữ liệu [115]. Kết quả cho thấy GATK có kết quả tốt nhất trên dữ liệu WES, với cả dữ liệu thực và dữ liệu mô phỏng

Cuối cùng, các biến thể được chú giải và phân loại bằng các phần mềm như ANNOVAR[116], SnpEff [117]. ANNOVAR là công cụ phân tích ảnh hưởng của các biến thể di truyền đến trình tự và chức năng protein từ dữ liệu giải trình tự gen thế hệ mới một cách hiệu quả, nhanh chóng. Đồng thời, công cụ này cung cấp quy trình sàng lọc cách linh hoạt giúp xác định được những biến thể tiềm năng gây bệnh. SnpEff v4.1 là công cụ chú thích và dự báo ảnh hưởng của các biến thể gen (như thay đổi axit amin). Dữ liệu đầu vào của công cụ này là các biến thể được dự đoán (SNPs, chèn, xóa), là kết quả của giải trình tự, và có định dạng VCF (Variant Call Format). Trong dữ liệu đầu ra, SnpEff sẽ phân tích các biến đầu vào để chú giải và tính toán các tác động mà các biến thể có thể tạo ra trên gen. SnpEff đưa ra các kết quả như sau: kiểu gen và các điểm bị ảnh hưởng bởi biến thể; vị trí của các biến thể; ảnh hưởng của biến thể đến quá trình tổng hợp protein; so sánh với các dữ liệu khác để tìm các biến thể đã biết.

2.3.4. Sàng lọc biến thể

Sau khi phân tích, các biến thể được sàng lọc theo các quy trình như trình bày ở Hình 2.3.



Hình 2.3: Quy trình sàng lọc biến thể từ dữ liệu WES

Các tiêu chí chính để sàng lọc và tìm kiếm các biến thể có liên quan đến bệnh teo đường mật bẩm sinh được áp dụng là: (1) biến thể xảy ra trên các gen đã được công bố là liên quan đến bệnh teo đường mật bẩm sinh hoặc những hội chứng liên quan đến gan mật, (2) Biến thể có tần số alen MAF > 1% được coi là đa hình đơn nucleotide (SNP). Những biến thể này được khảo sát trên số lượng lớn mẫu bệnh nhân và đối chứng từ đó đánh giá mức độ liên quan của SNP với nguy cơ mắc bệnh teo đường mật bẩm sinh, (3) Biến thể có tần suất alen tham khảo dựa trên bộ cơ sở dữ liệu 1000 hệ gen người (1000 Genomes Project) MAF < 1% được coi là đột biến. Các đột biến này được đánh giá mức độ ảnh hưởng đến chức năng protein. Các đột biến thêm bộ ba kết thúc dẫn đến cắt ngắn protein (đột biến vô nghĩa), mất bộ ba kết thúc dẫn đến kéo dài protein, thêm/mất nucleotide dẫn đến dịch khung protein và thay đổi vị trí trượt gen có ảnh hưởng lớn đến chức năng của protein được lựa chọn. Các đột biến sai nghĩa được đánh giá sơ bộ khả năng gây hại đến chức năng protein qua công cụ SIFT [118], PolyPhen-2 [119], PROVEAN [120] và phần mềm Mutation Taster [121].

SIFT là công cụ sử dụng trình tự tương đồng để dự đoán liệu sự thay thế amino axit có ảnh hưởng đến chức năng protein hay không và do đó có khả năng làm thay đổi kiểu hình không. Thuật toán SIFT hoạt động chỉ dựa vào trình tự nucleotide mà không cần cấu trúc protein nên ưu điểm lớn nhất là có thể dự đoán nhanh với một số lượng lớn biến thể. Thuật toán của phần mềm cho rằng các amino axit quan trọng sẽ được bảo tồn trong họ protein và do đó những thay đổi ở các vị trí được bảo tồn tốt có xu hướng ảnh hưởng đến chức năng protein và thường được dự đoán là gây hại. Phần mềm SIFT đánh giá ảnh hưởng của đột biến theo thang điểm từ 0 đến 1. Điểm số SIFT nhỏ hơn 0,05 được gọi là gây hại và ngược lại sẽ là lành tính

Polyphen-2 là phần mềm dự đoán tác động của một biến thể sai nghĩa dựa trên trình tự amino axit, thông tin phát sinh gen và thông tin về cấu trúc của protein. Polyphen-2 cũng xem xét các protein tương đồng để xem liệu đột biến sai nghĩa cần phân tích có được quan sát thấy ở các protein khác cùng họ hay không. Phần mềm PolyPhen-2 đánh giá ảnh hưởng của đột biến theo thang điểm từ 0 đến 1 với điểm số gần với mức 1 thì coi là có hại. Các đột biến có điểm đánh giá PolyPhen-2 trong khoảng 0,957 đến 1 được cho là có hại (D – damaging); thang điểm trong khoảng 0,453–0,956 là có thể gây hại (P – possibly damaging) và các đột biến có điểm đánh giá trong khoảng 0–0,452 là lành tính (B – benign).

PROVEAN (Protein Variation Effect Analyzer) là một công cụ dự đoán tác động của các thay đổi trong trình tự amino axit đến chức năng của protein. Công cụ này đặc biệt hữu ích để sàng lọc các biến thể không đồng nghĩa, biến thể dạng chèn, xóa nucleotide được dự đoán là có tác động quan trọng đến protein. Tương tự như SIFT, PROVEAN là một công cụ dự đoán chỉ dựa vào trình tự nucleotide nên có thể dự đoán nhanh với một số lượng lớn biến thể.

Mutation Taster là một ứng dụng để đánh giá khả năng gây bệnh của các biến thể trong trình tự DNA. Phần mềm này thực hiện nhiều đánh giá in silico để ước tính tác động của biến thể lên gen và protein. Mutation Taster đánh giá ảnh hưởng của đột biến ở cả mức độ DNA và protein do đó phần có thể xử lý nhiều loại biến thể phức tạp như (i) những biến thể dạng đồng nghĩa, không làm thay đổi amino axit, (ii) biến thể dạng chèn, xóa (indel) nucleotide, (iii)

biến thể dạng thay thế một hoặc nhiều nucleotide, ảnh hưởng đến một hoặc nhiều amino axit, (iv) tính bảo tồn của amino axit bị thay đổi trên các loài khác nhau, (v) khả năng mất các miền protein chức năng, thay đổi chiều dài protein hay các tác động đến điểm cắt nối mRNA. Các kết quả phân tích cuối cùng được đánh giá và phân loại dựa theo phân lớn Naive Bayes để quyết định xem tác động kết hợp của chúng có gây hại cho protein hay không.

Ngoài việc sử dụng các công cụ tin sinh học để đánh giá mức độ gây bệnh của các biến thể, những biến thể được báo cáo là lành tính trên cơ sở dữ liệu ClinVar cũng bị loại bỏ.

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

Phương pháp RFLP (Restriction Fragment Length Polymorphisms – đa hình chiều dài đoạn cắt giới hạn) là kỹ thuật nghiên cứu tính đa hình chiều dài của các đoạn DNA dựa trên điểm cắt các enzyme giới hạn. Nguyên tắc của kỹ thuật này dựa trên độ đặc hiệu của các enzyme cắt giới hạn đối với vị trí nhận biết của chúng trên DNA. Trong phương pháp RFLP, DNA mẫu được cắt thành các đoạn nhỏ bằng cách sử dụng các enzyme cắt giới hạn, và sau đó các đoạn DNA nhỏ tạo thành được phân tách dựa theo kích thước bằng kỹ thuật điện di trên gel. Mặc dù ngày nay kỹ thuật RFLP đã trở nên lỗi thời do bị thay thế bởi công nghệ giải trình tự nhưng đây vẫn là là công nghệ nghiên cứu đa hình DNA đầu tiên đủ rẻ để có thể được ứng dụng một cách rộng rãi.

Trong nghiên cứu này, các đoạn mồi cho phản ứng khuếch đại gen được thiết kế bằng phần mềm Primer-Blast. Đoạn gen *ABCB11* kích thước 333 bp có chứa đa hình rs2287622 được khuếch đại bằng phản ứng PCR sử dụng cặp mồi:

- Mồi xuôi: 5′-CACACAGACACCGAGTATCAACACA-3′
- Mồi ngược: 5′-CAGGACAGTCTCAATGTATGCTACACCT-3′

Chu trình nhiệt của phản ứng: 95°C – 12 phút, 35 chu kỳ (95°C – 45 giây, 64°C – 45 giây, 72°C – 45 giây), kéo dài ở 72°C trong 8 phút và giữ ở 4°C.

Enzyme *Hae*III có vị trí cắt đặc hiệu 5’-GGCC-3’ được sử dụng để xác định kiểu gen tại vị trí chứa đa hình.

Thành phần của phản ứng cắt enzyme được trình bày ở Bảng 2.1

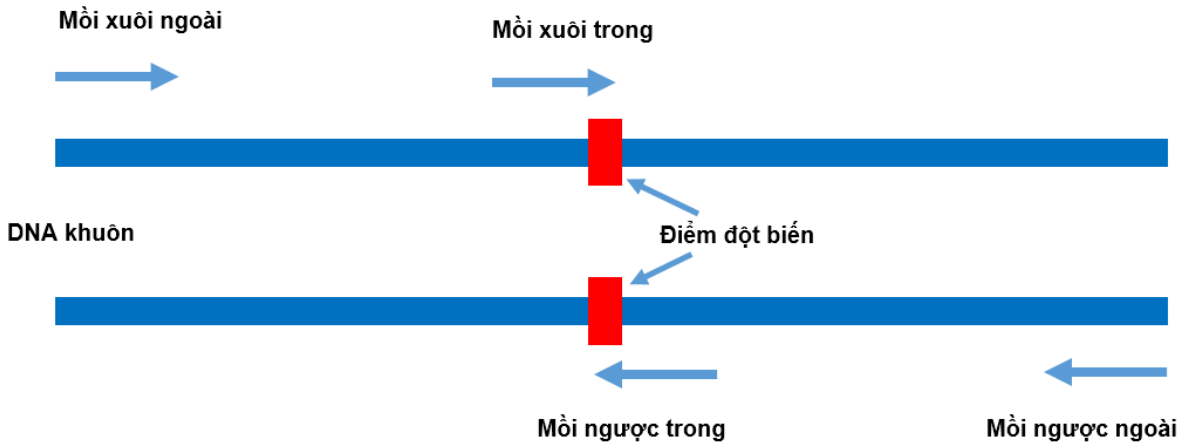
Bảng 2.1: Thành phần của phản ứng cắt enzyme HaeIII

dH_2O	15 µl
10x Buffer R	2 µl
Enzyme *Hae*III	1 µl
Sản phẩm PCR	2 µl
Ủ ở 37^oC trong 2h	

Theo tính toán, nếu alen tại vị trí đa hình là đồng hợp tử TT thì kết quả trên gel điện di sẽ hiện 1 băng với kích thước 333 bp do không có điểm cắt đặc hiệu của enzyme *Hae*III; nếu alen tại vị trí đa hình là đồng hợp tử CC thì đoạn gen sẽ bị cắt hoàn toàn thành 2 đoạn nhỏ có kích thước 206 bp và 127 bp; nếu alen tại vị trí đa hình là dị hợp tử TC thì trên hình ảnh điện di sẽ có cả 3 băng với các kích thước 333 bp, 206 bp và 127 bp.

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

Kỹ thuật ARMS-PCR (Amplification Refractory Mutation System Polymerase Chain Reaction) được sử dụng trong xác định kiểu gen của đa hình nucleotide với bốn đoạn mồi trong một phản ứng PCR duy nhất. Phản ứng này dùng một mồi đặc hiệu có trình tự đầu 3’ bổ sung với alen đột biến và 1 mồi ngược lại có trình tự đặc hiệu với alen thường (Hình 2.4). Sự có mặt của đột biến được thể hiện bằng sản phẩm DNA khuếch đại với các kích thước khác nhau đã biết trước và cho phép quan sát dễ dàng dưới gel agarose khi điện di. ARMS-PCR là một quy trình được sử dụng dễ dàng và rộng rãi trong các phòng thí nghiệm lâm sàng chẩn đoán phân tử [122]. Việc sử dụng ARMS-PCR đáp ứng được kỳ vọng của nghiên cứu bộ gen hiện đại và cho phép nghiên cứu SNP một cách nhanh chóng, đáng tin cậy và chi phí thấp, tuy nhiên, bước tối ưu hóa chu trình nhiệt của phản ứng có thể rất phức tạp và tốn thời gian. Một nhược điểm khác của ARMS – PCR là không phát hiện được biến thể mất/thêm đoạn lớn và các bất thường nhiễm sắc thể.



Hình 2.4: Nguyên lý của phản ứng 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 [123]. Vùng gen *ABCC2* chứa rs927344 được khuếch đại bằng cặp mồi đặc hiệu:

ABCC2-F1: 5'- GTCATTAAATGGTGAATTTGTATCCA -3'

ABCC2-R1: 5'- GTGCACCAAGGAATTAGAGTTCA -3'

ABCC2-F2: 5'- GTGTGGATTCCCTTGGGTTT -3'

ABCC2-R2: 5'- GGGCCAGGAGCCATAAGT -3'

Vùng gen *MYO5B* chứa rs1815930 được khuếch đại bằng cặp mồi đặc hiệu:

MYO5B-F1: 5'- AGATGAGCCTTGATTGAGCTGATTG -3'

MYO5B-R1: 5'- GCCCACTGACTACTCAGGAGGACTT -3'

MYO5B-F2: 5'- CCAATCTATGGACAAGATGTCATCTCTG -3'

MYO5B-R2: 5'- CCCATGTTTTGGCCACTGTATGT -3'

Các cặp mồi này được thiết kế bằng phần mềm Primer1 [124] và tổng hợp bởi công ty Sinh hóa Phù Sa (Cần Thơ, Việt Nam). Phản ứng ARMS-PCR của cả hai đa hình được thực hiện với chu trình nhiệt đã được tối ưu hóa như bảng Bảng 2.2

Bảng 2.2: Chu trình nhiệt của phản ứng ARMS-PCR

95°C	12 phút	
95°C	45 phút	35 chu kỳ
64°C	45 giây	
72°C	45 giây	
72°C	8 phút	

Sản phẩm PCR được điện di kiểm tra trên gel agarose 1%. Kiểu gen đa hình được xác định dựa vào số băng và kích thước sản phẩm PCR thu được (Bảng 2.3).

Bảng 2.3: Số lượng và kích thước sản phẩm PCR tương ứng với các kiểu gen của ABCC2 rs927344 và MYO5B rs1815930

Kiểu gen	**Số lượng băng sản phẩm PCR**	**Chiều dài sản phẩm PCR (bp)**
Gen *ABCC2*		
AA	2	401, 282
AT	3	401, 282, 157
TT	2	401, 157
Gen *MYO5B*		
AA	2	349, 246
AG	3	349, 246, 154
GG	2	349, 154

2.3.7. 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. Thành phần phản ứng được trình bày ở Bảng 2.4.

Bảng 2.4: Thành phần phản ứng PCR giải trình tự Sanger

Thành phần	**Thể tích**
Nước khử ion	9 µl
BigDye Terminator Master Mix 2.5X	4 µl
BigDye Terminator 5X sequencing buffer	2 µl
Mồi (1,7 pmol/µl)	1 µl
Sản phẩm PCR	4 µl
Tổng	**20 µl**

Tinh sạch sản phẩm PCR sử dụng Bead dành riêng cho giải trình tự Sanger. Bổ sung 6 µl HiDi vào mỗi mẫu sản phẩm PCR tinh sạch và biến tính ở 98^{o}C trong thời gian 2 phút trên máy PCR sau đó làm lạnh trên đá trong 5 phút rồi đưa vào máy giải trình tự ABI 3500 Genetic Analyzer.

2.3.8. 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 rs2287622 trên gen *ABCB11,* rs927344 trên gen *ABCC2* và rs1815930 trên gen *MYO5B* với khả năng mắc bệnh teo đường mật bẩm sinh. Mối liên quan giữa kiểu gen của đa hình với bệnh được xem xét ở 3 mô hình khác nhau, bao gồm 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.

Chương 3. KẾT QUẢ

3.1. Thu thập mẫu và tách chiết DNA tổng số

Đối tượng tham gia nghiên cứu được chẩn đoán là mắc bệnh teo đường mật bẩm sinh. Tất cả các đối tượng tham gia nghiên cứu đều được cung cấp đầy đủ thông tin, mục tiêu nghiên cứu và tham gia nghiên cứu tự nguyện theo đúng quy định về đạo đức trong nghiên cứu y sinh học của tổ chức trong nước và trên thế giới.

Mẫu máu của các đối tượng tham gia nghiên cứu được thu thập bởi các bác sĩ tại Bệnh viện Nhi Trung ương. Mỗi đối tượng được lấy 2 ml máu toàn phần được bảo quản trong ống chống đông EDTA và giữ ở -20^{o}C. Thông tin lâm sàng của người tham gia nghiên cứu được trình bày ở Bảng 3.1

Bảng 3.1: Thông tin lâm sàng của người tham gia nghiên cứu

Thông tin lâm sàng	**Mẫu bệnh**	**Mẫu đối chứng**
Số lượng mẫu	266	250
Type I (bệnh nhân)	5	-
Type II (bệnh nhân)	1	-
Type III (bệnh nhân)	260	-
Tuổi chẩn đoán (tháng)	$2{,}70 \pm 1{,}25$	$2{,}64 \pm 0{,}78$
Bilirubin trực tiếp (mg/dL)	$5{,}68 \pm 1{,}69$	$6{,}30 \pm 1{,}21$
Protein (g/L)	$55{,}92 \pm 4{,}42$	$58{,}23 \pm 5{,}06$
Albumin (g/L)	$36{,}4 \pm 7{,}87$	$38{,}05 \pm 3{,}10$
Aspartate aminotransferase (AST) (U/L)	$217{,}43 \pm 135{,}71$	$240{,}99 \pm 162{,}85$
Alanine aminotransferase (ALT) (U/L)	$129{,}32 \pm 88{,}02$	$141{,}02 \pm 83{,}53$

Thông tin lâm sàng	Mẫu bệnh	Mẫu đối chứng
Gamma glutamyltransferase (GGT) (IU/L)	604,25 ± 467,81	606,02 ± 440,84
Alkaline phosphatase (ALP) (U/L)	579,34 ± 306,14	744,52 ± 224,90

DNA tổng số của tất cả bệnh nhân và đối chứng được kiểm tra bằng phương pháp điện di trên gel agarose 0,8% và đo OD ở bước sóng 260 và 280 nm để xác định nồng độ và độ tinh sạch. DNA tổng số thu được ở cả mẫu bệnh nhân và mẫu đối chứng đều có nồng độ cao, không bị đứt gãy và có độ tinh sạch cao thể hiện qua các dải bằng sáng rõ. Nồng độ các mẫu đều lớn hơn 20ng/µl và giá trị A260/280 của các mẫu đo bằng máy Nanodrop đều nằm trong khoảng 1,8 – 2.

Trong số 266 bệnh nhân nhi mắc teo đường mật bẩm sinh, có 05 bệnh nhân thuộc 03 gia đình được chọn lọc để giải trình tự toàn bộ vùng gen mã hóa (Bảng 3.2). Trong đó, bệnh nhân BA01 và BA02 là chị em ruột, bệnh nhân BA03 và BA04 là chị em ruột, bệnh nhân BA05 không có quan hệ huyết thống với các bệnh nhân khác trong nghiên cứu này.

Bảng 3.2: Thông tin các mẫu bệnh nhân được lựa chọn để giải trình tự WES

STT	Kí hiệu	Giới tính	Tuổi phẫu thuật Kasai
1	BA01	Nữ	2 tháng
2	BA02	Nữ	Không phẫu thuật
3	BA03	Nữ	32 ngày
4	BA04	Nam	2,5 tháng
5	BA05	Nữ	2,5 tháng

Tất cả bệnh nhân đều được chẩn đoán là mắc bệnh teo đường mật bẩm sinh với những triệu chứng đặc trưng như tình trạng vàng da, ứ mật, phân nhạt màu, transaminase tăng, siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no.

Bệnh nhân BA01 xuất hiện vàng da tăng dần từ 20 ngày tuổi, phân nhạt màu liên tục từ khi hết phân su. Trẻ được khám và chẩn đoán teo đường mật bẩm sinh tại khoa gan mật. Tại đây, bệnh nhân được làm xét nghiệm máu thấy tình trạng vàng da ứ mật, transaminase tăng, siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no. Bệnh nhân BA01 được phẫu thuật Kasai vào thời điểm 2 tháng tuổi. Trong quá trình phẫu thuật trẻ được chụp đường mật và sinh thiết gan. Chẩn đoán xác định sau khi phẫu thuật là teo đường mật bẩm sinh. Sau phẫu thuật, bệnh nhân diễn biến bệnh không tốt, còn vàng da và phân nhạt màu. Bệnh nhân mất lúc 17 tháng tuổi.

Bệnh nhân BA02 xuất hiện vàng da tăng dần từ 1 tháng tuổi, phân nhạt màu liên tục từ 2 tuần tuổi. Bệnh nhân được khám và chẩn đoán teo đường mật bẩm sinh tại khoa gan mật. Tại đây, bệnh nhân được làm xét nghiệm máu thấy tình trạng vàng da ứ mật, transaminase tăng, siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no. Bệnh nhân không thực hiện phẫu thuật Kasai. Sau đó, bệnh nhân có hiện tượng vàng da tăng, bụng cổ trướng, rối loạn đông máu, xuất huyết tiêu hóa. Bệnh nhân mất lúc 11 tháng tuổi.

Bệnh nhân BA03 xuất hiện vàng da tăng dần từ 7 ngày tuổi, kèm phân nhạt màu liên tục. Bệnh nhân được khám và chẩn đoán teo đường mật bẩm sinh tại khoa Gan mật. Tại đây, bệnh nhân được làm xét nghiệm máu thấy tình trạng vàng da ứ mật, Transaminase tăng, siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no. Bệnh nhân BA03 được phẫu thuật Kasai vào thời điểm 32 ngày tuổi. Trong quá trình phẫu thuật trẻ được chụp đường mật và sinh thiết gan. Chẩn đoán xác định sau khi phẫu thuật là teo đường mật bẩm sinh. Sau phẫu thuật, bệnh nhân giảm vàng da dần, phân vàng. Bệnh nhân có 3 đợt xuất huyết tiêu hóa, nội soi dạ dày thực quản thấy giãn tĩnh mạch thực quản độ 3. Hiện tại bệnh nhân BA03 tiếp tục được theo dõi và điều trị tại khoa Gan mật.

Bệnh nhân BA04 xuất hiện vàng da tăng dần từ 5 ngày tuổi, phân nhạt màu liên tục từ 1 tháng 20 ngày tuổi. Bệnh nhân được khám và chẩn đoán teo đường mật bẩm sinh tại khoa gan mật. Tại khoa trẻ được làm xét nghiệm máu thấy tình trạng vàng da ứ mật, transaminase tăng, siêu âm gan mật có chuẩn bị thấy kích thước túi mật bé, ít thay đổi khi đói và sau ăn no. Bệnh nhân phẫu thuật Kasai vào thời điểm 2,5 tháng tuổi. Trong quá trình phẫu thuật, trẻ được chụp đường mật và sinh thiết gan. Chẩn đoán xác định sau khi phẫu thuật là teo đường mật bẩm sinh. Sau phẫu thuật, trẻ giảm vàng da dần, phân vàng hơn. Hiện tại trẻ tiếp tục được theo dõi và điều trị tại khoa gan mật.

Bệnh nhân BA05 được chẩn đoán mắc bệnh teo đường mật bẩm sinh lúc 2,5 tháng tuổi với các triệu chứng vàng da ứ mật, phân bạc màu, gan lách to, siêu âm túi mật kích thước nhỏ (12 × 3 mm). Sinh thiết mô gan cho thấy các tổn thương gan khác nhau bao gồm ứ mật trong gan. Bệnh nhân được phẫu thuật Kasai và được đánh giá nằm trong nhóm có tiên lượng tốt sau điều trị. Bệnh nhân không có triệu chứng vàng da 4 tháng sau phẫu thuật và duy trì ổn định trong 3 năm đầu đời. Khi lên 8 tuổi, bệnh nhi đột ngột có những diễn biến bất thường với các triệu chứng vàng da, suy gan cấp kèm theo tán huyết. Bệnh nhân vào viện trong tình trạng vàng da, nước tiểu sẫm màu, gan trái to 4 cm, lách to độ II. Đánh giá mô học của gan cho thấy sự tích tụ đồng trong mô gan. Bệnh nhân được chẩn đoán mắc bệnh Wilson. Bệnh nhân được điều trị theo phác đồ dành cho bệnh nhân Wilson và các phương pháp điều trị hỗ trợ cho bệnh nhân mắc bệnh gan giai đoạn cuối. Tuy nhiên, các triệu chứng của suy gan ngày càng xấu đi với các đợt hôn mê gan xảy ra liên tiếp trên nền bệnh gan mãn tính với các đợt tan máu nặng. Do cả teo đường mật bẩm sinh và Wilson đều là những căn bệnh hiếm gặp nên bệnh nhân BA05 được lựa chọn để giải trình tự WES nhằm tìm ra những biến thể mới.

3.2. Giải trình tự toàn bộ vùng mã hóa

3.2.1. Kết quả tạo thư viện DNA

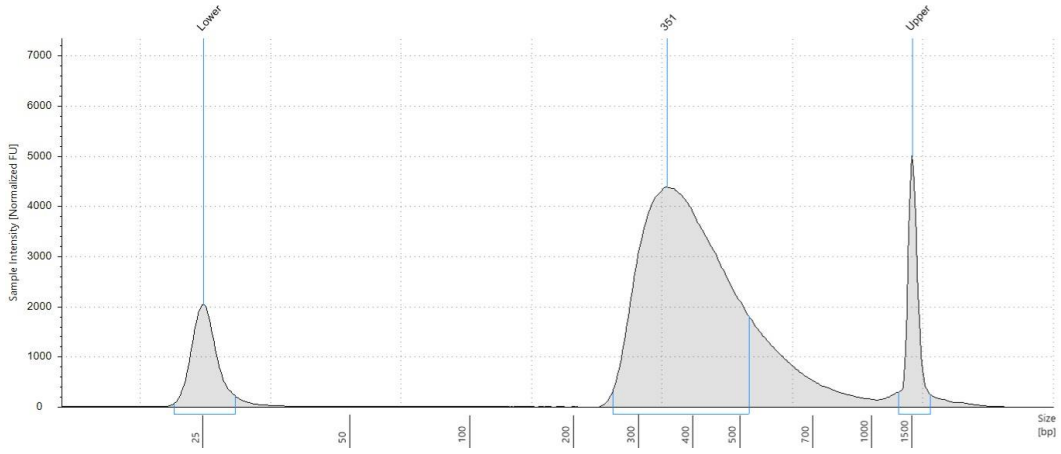
Chuẩn bị thư viện các mảnh DNA có gắn adapter là bước đầu tiên trong quy trình giải trình tự thế hệ mới và cũng là bước quan trọng xác định sự thành công của thí nghiệm, đánh giá độ tin cậy của kết quả. Thư viện DNA cần có nồng độ cao và phân bố kích thước của các mảnh DNA phải nằm trong khoảng

giới hạn phù hợp với hệ thống máy giải trình tự. Nồng độ của mảnh DNA trng nghiên cứu này được thể hiện ở Bảng 3.3.

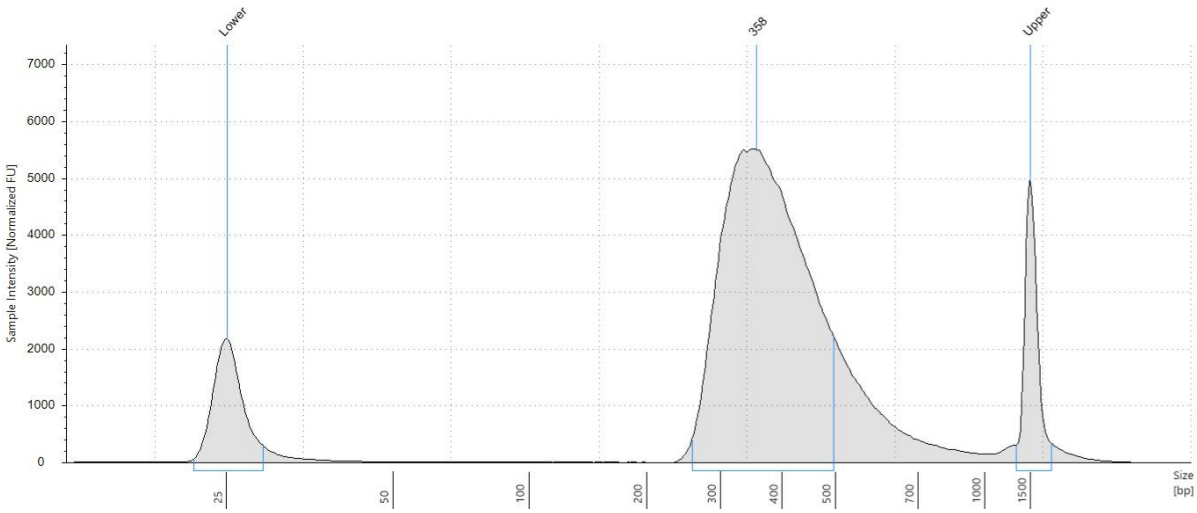
Bảng 3.3: Kích thước và nồng độ thư viện DNA

Mẫu	Kích thước (bp)	Nồng độ (ng/µl)
BA01	351	109,67
BA02	358	117,90
BA03	351	116,58
BA04	351	100,77
BA05	362	122,27

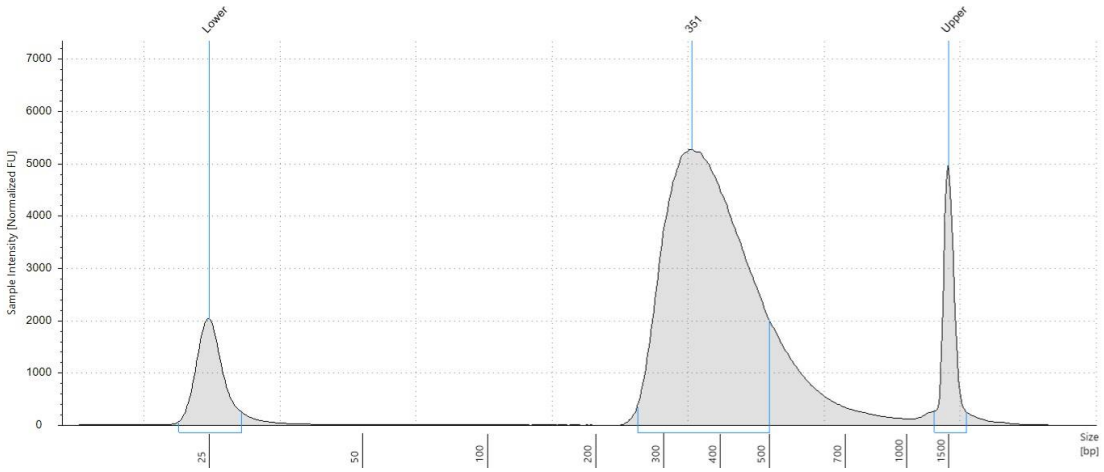
Để kiểm tra kích thước của mảnh PCR khuếch đại, chúng tôi kiểm tra sự phân bố kích thước mẫu bằng cách chạy mẫu trên máy Agilent Technologies 2100 Bioanalyzer sử dụng chip DNA 1000. Kết quả được thể hiện ở Hình 3.1



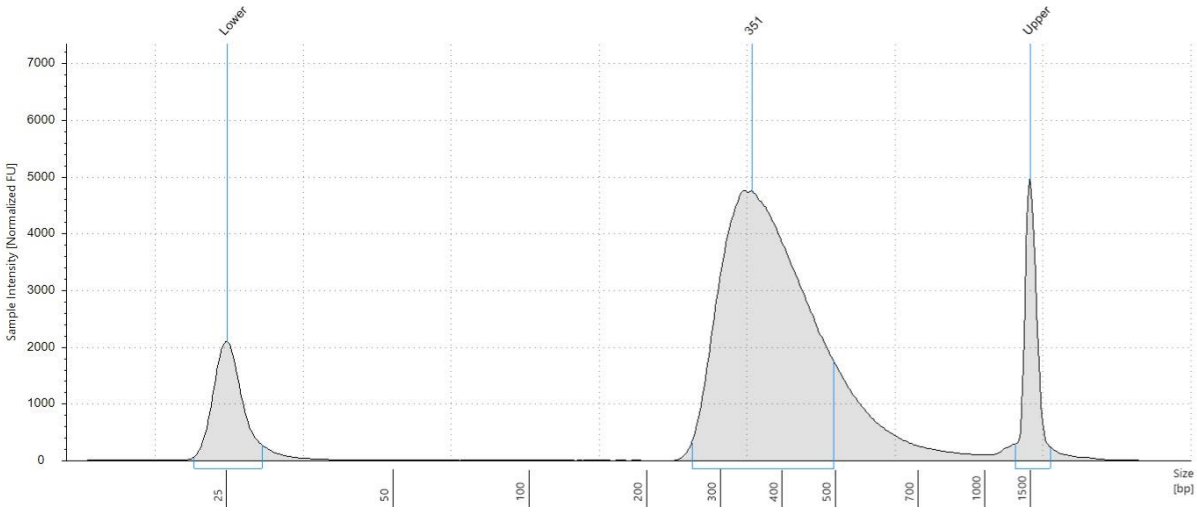
BA01



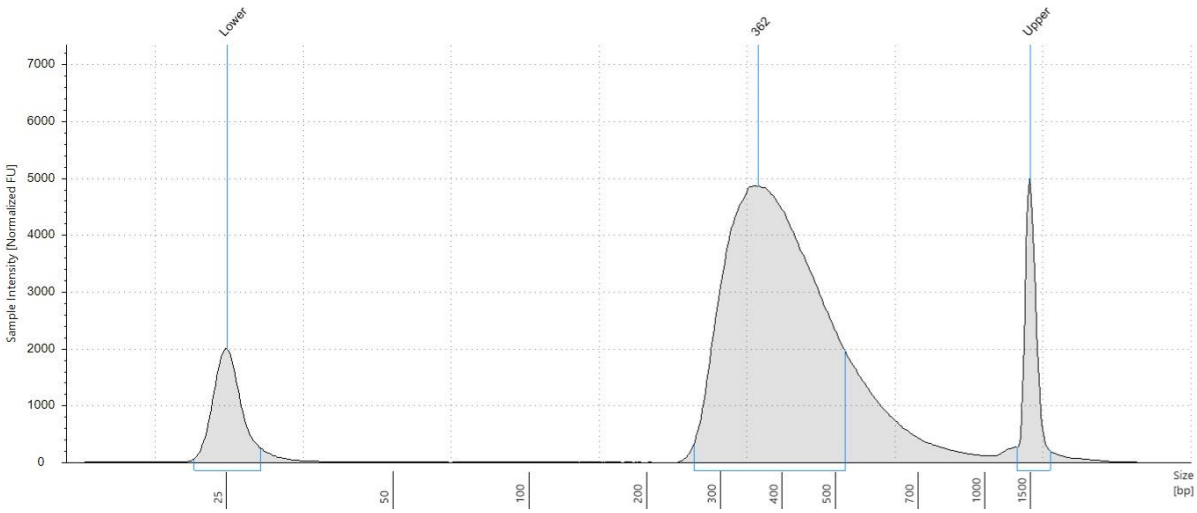
BA02



BA03



BA04



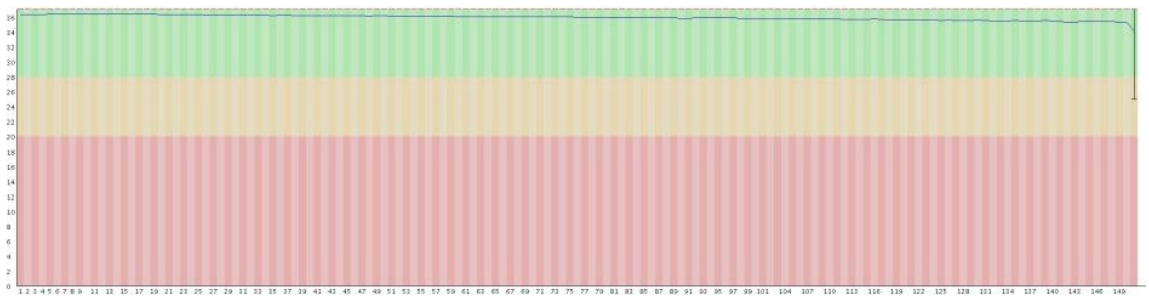
BA05

Hình 3.1: Kết quả kiểm tra chất lượng thư viện DNA

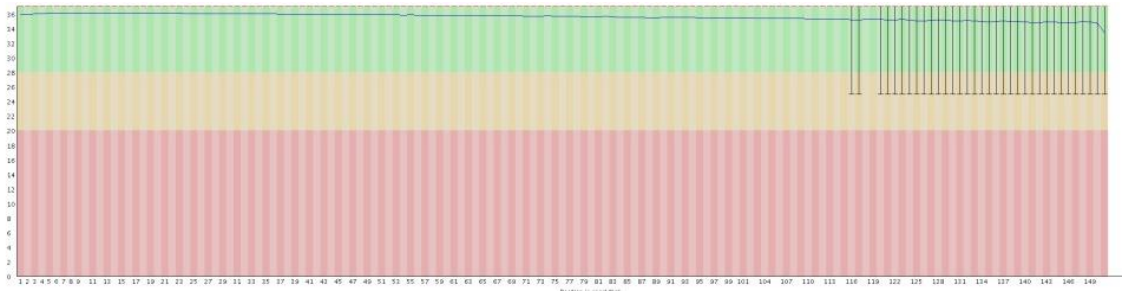
Kết quả kiểm tra chất lượng thư viện DNA được thể hiện trên Hình 3.1, tất cả các thư viện đều có chất lượng tốt, nồng độ cao, đạt yêu cầu cho bước nghiên cứu tiếp theo với phần lớn đoạn DNA sau khi được phần mảnh có kích thước nằm trong khoảng từ 250 – 500 bp.

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

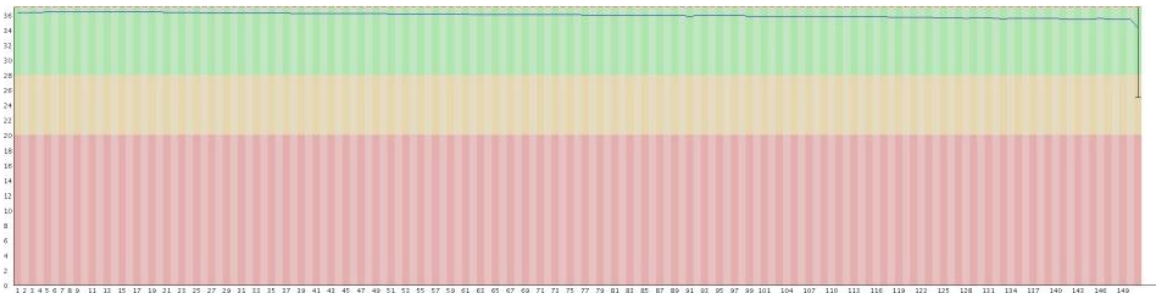
Trình tự toàn bộ vùng gen mã hóa của 05 bệnh nhân được xác định bằng thiết bị Illimina NextSeq 500. Chất lượng dữ liệu được đánh giá bằng phần mềm FastQC. Chỉ số thống kê được tóm tắt dưới dạng đồ thị và bảng biểu. Biểu đồ chất lượng từng nucleotide dữ liệu được phân chia thành 3 mức, có màu sắc khác nhau. Trong đó, chất lượng lớn hơn 28 (Q>28 – vùng màu xanh) được đánh giá là rất tốt, lớn hơn 20 (Q>20 – vùng màu vàng) là tốt, nhỏ hơn 20 (Q<20 – vùng màu hồng) là những dữ liệu có khả năng bị lỗi cao hơn mức chấp nhận được và có thể xem xét để cắt bỏ. Kết quả được thể hiện ở Hình 3.2.



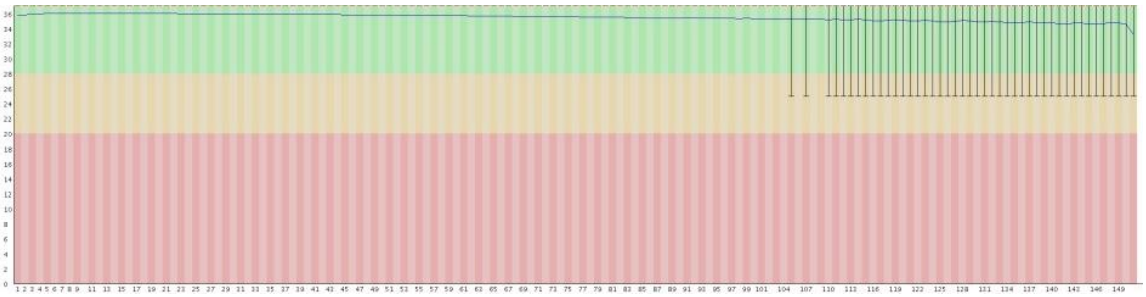
Mẫu BA01 – Read 1



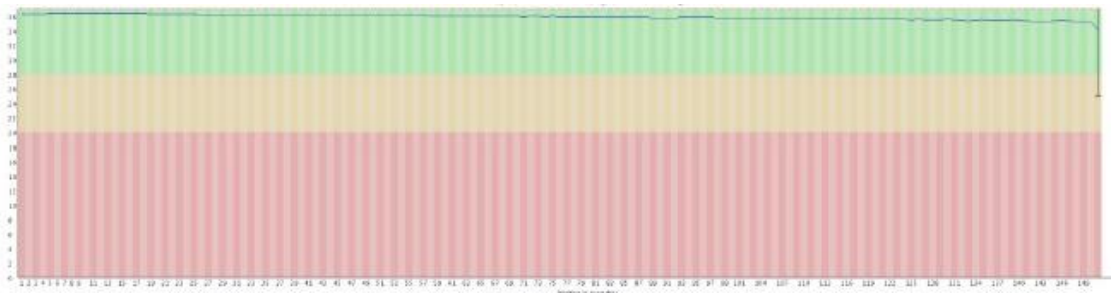
Mẫu BA01 – Read 2



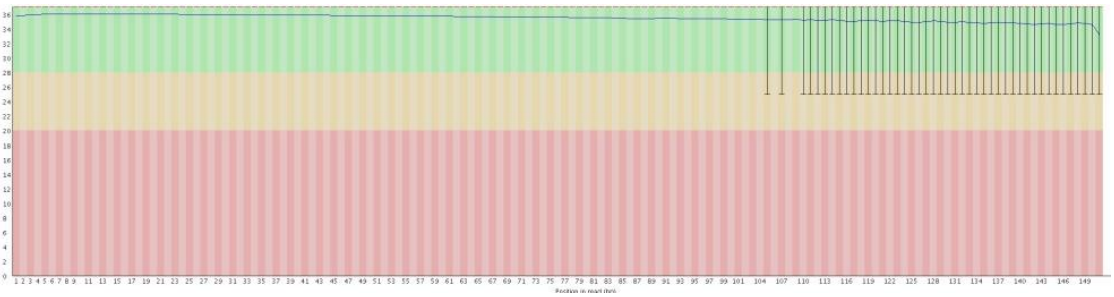
Mẫu BA02 – Read 1



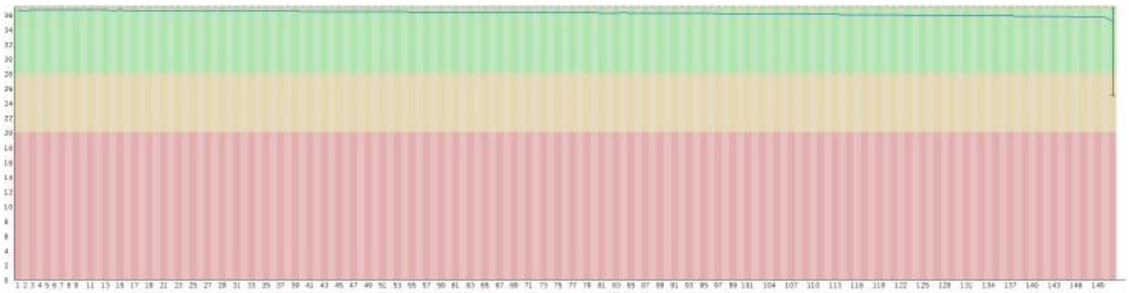
Mẫu BA02 – Read 2



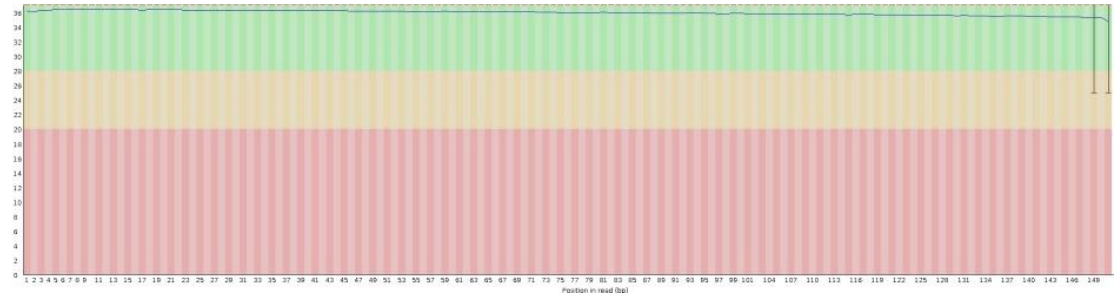
Mẫu BA03 – Read 1



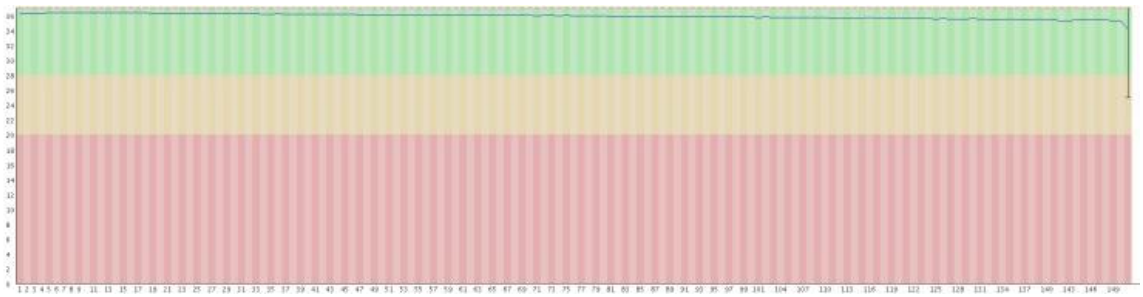
Mẫu BA03 – Read 2



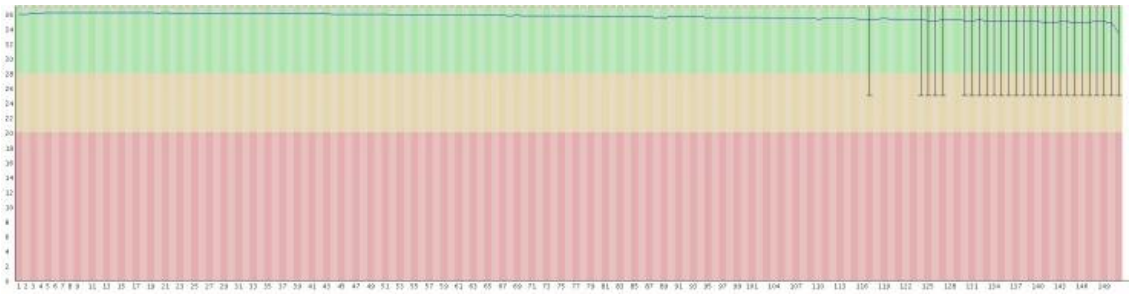
Mẫu BA04 – Read 1



Mẫu BA04 – Read 4



Mẫu BA05 – Read 1



Mẫu BA05 – Read 2

Hình 3.2: Chất lượng giải trình tự 05 bệnh nhân mắc teo đường mật bẩm sinh

Dữ liệu được tạo ra từ máy giải trình tự thế hệ mới có dung lượng cao, độ dài các đoạn đọc đồng đều (~150 bp mỗi đoạn), độ sâu bao phủ cao (>150X mỗi mẫu). Dữ liệu giải trình tự toàn bộ vùng gen mã hóa có 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
%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	95,1
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. Kết quả sắp hàng dữ liệu được thể hiện ở Bảng 3.5.

Bảng 3.5: Kết quả sắp hàng các mẫu giải trình tự với hệ gen tham chiếu

	BA01	**BA02**	**BA03**	**BA04**	**BA05**

Số đoạn trình tự sắp hàng thành công	67.207.673 (99,8%)	60.415.842 (99,8%)	58.264.414 (99,8%)	74.737.214 (99,9%)	70.500.265 (99,8%)
Số đoạn trình tự được sắp hàng vào vùng mã hóa	48.606.456 (79,0%)	42.739.604 (77,6%)	41.779.482 (78,2%)	44.832.640 (70,0%)	49.173.678 (76,5%)
Tỉ lệ độ sâu bao phủ >20X (%)	91,2	90,2	90,4	92,5	93,9
Tỉ lệ với độ sâu bao phủ >30X (%)	82,9	80,6	81,6	94,5	88,6
Tỉ lệ độ sâu bao phủ >50X (%)	63,2	59,1	61,5	65,3	72,7
Độ sâu bao phủ trung bình trong vùng exone (X)	88,0	80,5	85,5	94,6	85,2

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 exon cao (70-79%). Độ sâu bao phủ trung bình trong vùng exome sau khi sắp hàng của tất cả các mẫu bệnh nhân đều đạt >80X. trong đó, tỉ lệ dữ liệu có mức độ bao phủ trên 50X đều chiếm tỉ lệ lớn (59,1 – 72,7 %).

3.3.2. Xác định và chú giải biến thể

Nhiều công cụ khác nhau được phát triển để phát hiện biến thể từ dữ liệu NGS. Các nghiên cứu cho thấy mỗi công cụ phân tích có ưu thế đối với một loại dữ liệu giải trình tự nhất định. Quy trình xử lý số liệu sử dụng công cụ Picard/GATK được sử dụng để phát hiện biến thể trong nghiên cứu này. Phần mềm Picard để xử lý dữ liệu sau sắp hàng nhằm hiệu chỉnh chất lượng của từng nucleotide, sắp xếp, loại bỏ những đoạn lặp…. Bước phân tích này cũng nhằm mục đích loại bỏ những đoạn trình tự không được sắp hàng hoặc

những được sắp hàng có độ tin cậy thấp qua đó giảm tỉ lệ dương tính giả. Phần mềm GATK được sử dụng để xác định biến thể có trong dữ liệu hệ gen của mỗi bệnh nhân. Ngoài ra, việc biến đổi trình tự nucleotide ảnh hưởng thế nào đến amino axit 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 nucleotide	**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

Số lượng biến thể gen xuất hiện ở mỗi bệnh nhân là rất lớn, bao gồm chủ yếu là các thường biến lành tính, do đó cần có quá trình sàng lọc để tìm ra biến thể có tiềm năng là nguyên nhân gây bệnh ở mỗi bệnh nhân. Như đã trình bày ở trên, teo đường mật bẩm sinh là căn bệnh nhân có cơ sở di truyền rất phức tạp, tính không đồng nhất về gen cao. Teo đường mật có liên quan đến một số hội chứng bẩm sinh, bất thường nhiễm sắc thể và đa hình gen ở một số quần

thể cụ thể. Do đó, trong nghiên cứu này danh sách các gen được sử dụng để sàng lọc biến thể được mở rộng so với nhiều nghiên cứu khác, không chỉ bao gồm những gen đã được chỉ ra là nhạy cảm với teo đường mật bẩm sinh mà còn co cả những gen liên quan đến các hội chứng gan mật khác (133 gen trong nghiên cứu này so với 25 gen trong nghiên cứu của Trần Trung Kiên và đồng tác giả [106]). Danh sách những gen được sử dụng để sàng lọc đột biến được trình bày ở phụ lục 1. Sau quá trình sàng lọc, những biến thể nằm trong vùng exon của các gen liên quan đến bệnh teo đường mật bẩm sinh được giữ lại. 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 xuất hiện trong cơ sở dữ liệu 1000 Genome phase 3). Kết quả sàng lọc đột biến của 05 bệnh nhân được thể hiện ở Bảng 3.7.

Bảng 3.7: Kết quả sàng lọc đột biến xảy ra trong những gen liên quan đến bệnh teo đường mật bẩm sinh

Bệnh nhân	Tổng số đột biến trên gen liên quan	Số đột biến xảy ra trong exon
BA01	786	14
BA02	781	12
BA03	721	12
BA04	745	08
BA05	817	10

Số lượng biến thể xảy ra trong số 133 gen liên quan đến bệnh teo đường mật bẩm sinh là rất lớn (số lượng đột biến ở bệnh nhân BA01 là 786, ở bệnh nhân BA02 là 781, ở bệnh nhân BA03 là 721, ở bệnh nhân BA04 là 745, ở bệnh nhân BA05 là 817 đột biến). Tuy nhiên, phần lớn các đột biến này nằm trong vùng intron của gen, có rất ít đột biến nằm trong vùng exon (số lượng đột biến trong vùng exon 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). Danh sách cụ thể các đột biến xảy ra trong vùng exon của những gen liên quan đến bệnh teo đường mật bẩm sinh tìm thấy trong mỗi bệnh nhân được trình bày ở Bảng 3.8.

Bảng 3.8: Đột biến xảy ra trong gen liên quan đến bệnh teo đường mật bẩm sinh của các bệnh nhân

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p
Mẫu BA01					
1	chr1	Sai nghĩa	*NOTCH2*	c.57C>G	p.Cys19Trp
2	chr1	Dịch khung	*NOTCH2*	c.17_18delCC	p.Pro6fs
3	chr2	Sai nghĩa	*ABCG8*	c.1160C>T	p.Pro387Leu
4	chr2	Sai nghĩa	*CPS1*	c.1048A>G	p.Thr350Ala
5	chr5	Sai nghĩa	*TTC37*	c.1427C>G	p.Thr476Arg
6	chr5	Sai nghĩa	*TTC37*	c.829G>A	p.Gly277Ser
7	chr6	Sai nghĩa	*PKHD1*	c.1675C>T	p.Arg559Trp
8	chr12	Sai nghĩa	*KRT18*	c.300C>G	p.Ser100Arg
9	chr12	Sai nghĩa	*KRT18*	c.308C>A	p.Thr103Asn
10	chr12	Sai nghĩa	*KRT18*	c.316C>T	p.Arg106Trp
11	chr15	Sai nghĩa	*FAH*	c.412G>A	p.Gly138Arg
12	chr16	Sai nghĩa	*ERCC4*	c.2225A>G	p.Tyr742Cys
13	chr18	Sai nghĩa	*NPC1*	c.110A>G	p.Asp37Gly
14	chrX	Sai nghĩa	*AMER1*	c.85G>A	p.Ala29Thr

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p
Mẫu BA02					
1	chr1	Sai nghĩa	*NOTCH2*	c.57C>G	p.Cys19Trp
2	chr1	Dịch khung	*NOTCH2*	c.17_18delCC	p.Pro6fs
3	chr2	Sai nghĩa	*MPV17*	c.404A>G	p.Tyr135Cys
4	chr2	Sai nghĩa	*CPS1*	c.1048A>G	p.Thr350Ala
5	chr9	Sai nghĩa	*IARS*	c.337A>T	p.Ile113Phe
6	chr12	Sai nghĩa	*KRT18*	c.300C>G	p.Ser100Arg
7	chr12	Sai nghĩa	*KRT18*	c.308C>A	p.Thr103Asn
8	chr12	Sai nghĩa	*KRT18*	c.316C>T	p.Arg106Trp
9	chr15	Sai nghĩa	*FAH*	c.412G>A	p.Gly138Arg
10	chr16	Sai nghĩa	*ERCC4*	c.241G>T	p.Val81Phe
11	chr18	Sai nghĩa	*NPC1*	c.110A>G	p.Asp37Gly
12	chrX	Sai nghĩa	*AMER1*	c.85G>A	p.Ala29Thr
Mẫu BA03					
1	chr1	Sai nghĩa	*PEX14*	c.208T>G	p.Ser70Ala
2	chr1	Sai nghĩa	*NOTCH2*	c.57C>G	p.Cys19Trp
3	chr1	Dịch khung	*NOTCH2*	c.17_18delCC	p.Pro6fs
4	chr1	Sai nghĩa	*EPHX1*	c.128G>C	p.Arg43Thr
5	chr2	Sai nghĩa	*CPS1*	c.1048A>G	p.Thr350Ala

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p
6	chr2	Sai nghĩa	*CYP27A1*	c.368A>C	p.Lys123Thr
7	chr2	Sai nghĩa	*GPC1*	c.1084G>A	p.Gly362Ser
8	chr9	Sai nghĩa	*BAAT*	c.685G>A	p.Val229Ile
9	chr12	Sai nghĩa	*KRT18*	c.300C>G	p.Ser100Arg
10	chr12	Sai nghĩa	*KRT18*	c.308C>A	p.Thr103Asn
11	chr12	Sai nghĩa	*KRT18*	c.316C>T	p.Arg106Trp
12	chr12	Sai nghĩa	*KRT18*	c.361C>T	p.Pro121Ser
13	chr22	Sai nghĩa	*TRMU*	c.75G>T	p.Arg25Ser
Mẫu BA04					
1	chr1	Sai nghĩa	*NOTCH2*	c.57C>G	p.Cys19Trp
2	chr1	Dịch khung	*NOTCH2*	c.17_18delCC	p.Pro6fs
3	chr2	Sai nghĩa	*NBAS*	c.5984G>A	p.Arg1995Gln
4	chr2	Sai nghĩa	*CPS1*	c.1048A>G	p.Thr350Ala
5	chr2	Sai nghĩa	*CYP27A1*	c.368A>C	p.Lys123Thr
6	chr2	Sai nghĩa	*GPC1*	c.1084G>A	p.Gly362Ser
7	chr6	Sai nghĩa	*PKHD1*	c.10252G>A	p.Ala3418Thr
8	chr22	Sai nghĩa	*TRMU*	c.75G>T	p.Arg25Ser
Mẫu BA05					
1	chr1	Sai nghĩa	*NOTCH2*	c.137A>G	p.Asn46Ser

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p
2	chr1	Sai nghĩa	*NOTCH2*	c.112G>A	p.Glu38Lys
3	chr1	Sai nghĩa	*NOTCH2*	c.57C>G	p.Cys19Trp
4	chr2	Sai nghĩa	*NBAS*	c.3212A>T	p.Glu1071Val
5	chr7	Sai nghĩa	*CFTR*	c.1251C>A	p.Asn417Lys
6	chr12	Sai nghĩa	*KRT18*	c.50_51insG	p.Gly17Glyfs77∗
7	chr16	Sai nghĩa	*PMM2*	c.110A>T	p.Gln37Leu
8	chr18	Sai nghĩa	*MYO5B*	c.2292C>G	p.Phe764Leu
9	chr22	Sai nghĩa	*TRMU*	c.75G>T	p.Arg25Ser

Ngoài những nghiên cứu về đột biến, đánh giá mối liên quan giữa các đa hình nucleotide đơn (SNP) với nguy cơ mắc bệnh cũng được nhiều nhà khoa học quan tâm. Đây là những biến thể được phát hiện với tần số lớn hơn 1% trong quần thể. SNP thường xuất hiện sau mỗi 1000 nucleotide và được sử dụng làm dấu hiệu di truyền trong quần thể để nghiên cứu khuynh hướng đối với một số tính trạng, bao gồm cả sự mẫn cảm với bệnh. Danh sách đa hình nucleotide đơn của các bệnh nhân teo đường mật bẩm sinh được thể hiện ở phụ lục 2.

3.5. Đánh giá ảnh hưởng của các đột biến

Từ những đột biến xảy ra trong các gen liên quan đến bệnh teo đường mật bẩm sinh ở bảng 3.8 những đột biến đã được báo cáo là lành tính trên cơ sở dữ liệu ClinVar bị loại bỏ. Những đột biến còn lại được đánh giá bằng các phần mềm SIFT, PolyPhen2, Mutation Taster, Provean. Kết quả đánh giá ảnh hưởng của đột biến bằng các công cụ tin sinh được thể hiện ở Bảng 3.9.

Bảng 3.9: Kết quả đánh giá ảnh hưởng của các điểm đột biến bằng công cụ tin sinh

Gen	HGVS.p	HGVS.c	SIFT (Score)	PolyPhen2 (Score)	Mutation Taster (Score)	Provean (Score)
Mẫu BA01						
NOTCH2	p.Cys19Trp	c.57C>G	T (0,185)	B (0,002)	B (0,997)	N (-0,767)
ABCG8	p.Pro387Leu	c.1160C>T	T (0,229)	D (0,487)	B (0,999)	N (-1,03)
TTC37	p.Thr476Arg	c.1427C>G	D (0,009)	B (0,213)	D (0,906)	N (-1,25)
TTC37	p.Gly277Ser	c.829G>A	D (0,013)	D (0,961)	D (0,999)	N (-1,23)
PKHD1	p.Arg559Trp	c.1675C>T	T (0,221)	B (0,017)	B (0,999)	N (-0,5)
KRT18	p.Ser100Arg	c.300C>G	D (0,024)	B (0,054)	D (0,999)	N (0,21)
KRT18	p.Thr103Asn	c.308C>A	T (0,285)	B (0,007)	B (0,999)	N (-1,68)
KRT18	p.Arg106Trp	c.316C>T	T (0,070)	B (0,001)	B (0,999)	N (-1,79)
FAH	**p.Gly138Arg**	**c.412G>A**	**D (0,001)**	**D (1,000)**	**D (0,999)**	**D (-7,38)**
ERCC4	**p.Tyr742Cys**	**c.2225A>G**	**D (0,000)**	**D (1,000)**	**D (0,999)**	**D (-8,23)**
NPC1	p.Asp37Gly	c.110A>G	T (0,455)	B (0,001)	D (0,999)	N (-1,39)
AMER1	p.Ala29Thr	c.85G>A	D (0,004)	B (0,001)	B (0,834)	N (-0,38)
Mẫu BA02						
NOTCH2	p.Cys19Trp	c.57C>G	T (0,185)	B (0,002)	B (0,997)	N (-0,767)
MPV17	p.Tyr135Cys	c.404A>G	T (0,101)	D (0,996)	D (0,999)	D (-7,24)
IARS	p.Ile113Phe	c.337A>T	D (0,001)	D (0,975)	D (0,999)	N (-1,51)
KRT18	p.Ser100Arg	c.300C>G	D (0,024)	B (0,054)	D (0,999)	N (0,21)
KRT18	p.Thr103Asn	c.308C>A	T (0,285)	B (0,007)	B (0,999)	N (-1,68)

Gen	HGVS.p	HGVS.c	SIFT (Score)	PolyPhen2 (Score)	Mutation Taster (Score)	Provean (Score)
KRT18	p.Arg106Trp	c.316C>T	T (0,070)	B (0,001)	B (0,999)	N (-1,79)
FAH	**p.Gly138Arg**	**c.412G>A**	**D (0,001)**	**D (1,000)**	**D (0,999)**	**D (-7,38)**
ERCC4	**p.Val81Phe**	**c.241G>T**	**D (0,002)**	**D (0,965)**	**D (0,999)**	**D (-2,78)**
NPC1	p.Asp37Gly	c.110A>G	T (0,455)	B (0,001)	D (0,999)	N (-1,39)
AMER1	p.Ala29Thr	c.85G>A	D (0,004)	B (0,001)	B (0,834)	N (-0,38)
Mẫu BA03						
PEX14	p.Ser70Ala	c.208T>G	T (0,456)	D (0,950)	D (0,999)	N (-1,02)
NOTCH2	p.Cys19Trp	c.57C>G	T (0,185)	B (0,002)	B (0,997)	N (-0,767)
EPHX1	p.Arg43Thr	c.128G>C	T (0,262)	B (0,002)	B (0,997)	N (-1,40)
CYP27A1	**p.Lys123Thr**	**c.368A>C**	**D (0,003)**	**D (0,774)**	**D (0,999)**	**D (-3,71)**
GPC1	p.Gly362Ser	c.1084G>A	D (0,025)	B (0,002)	D (0,994)	D (-2,98)
BAAT	p.Val229Ile	c.685G>A	T (1,00)	B (0,015)	B (0,958)	N (0,56)
KRT18	p.Ser100Arg	c.300C>G	D (0,024)	B (0,054)	D (0,999)	N (0,21)
KRT18	p.Thr103Asn	c.308C>A	T (0,285)	B (0,007)	B (0,999)	N (-1,68)
KRT18	p.Arg106Trp	c.316C>T	T (0,070)	B (0,001)	B (0,999)	N (-1,79)
KRT18	p.Pro121Ser	c.361C>T	D (0,012)	D (0,593)	D (0,999)	D (-5,40)
Mẫu BA04						
NOTCH2	p.Cys19Trp	c.57C>G	T (0,185)	B (0,002)	B (0,997)	N (-0,767)
NBAS	p.Arg1995Gln	c.5984G>A	T (0,112)	D (1,00)	D (0,999)	N(-2,28)
CYP27A1	**p.Lys123Thr**	**c.368A>C**	**D (0,003)**	**D (0,774)**	**D (0,999)**	**D (-3,71)**
GPC1	p.Gly362Ser	c.1084G>A	D (0,025)	B (0,002)	D (0,994)	D (-2,98)

Gen	HGVS.p	HGVS.c	SIFT (Score)	PolyPhen2 (Score)	Mutation Taster (Score)	Provean (Score)
PKHD1	p.Ala3418Thr	c.10252G>A	T (0,170)	D (0,659)	B (0,997)	N (-0,49)
Mẫu BA05						
NOTCH2	p.Asn46Ser	c.137A>G	T (0,129)	D (0,713)	D (0,998)	N (-1,39)
NOTCH2	p.Glu38Lys	c.112G>A	T (0,058)	B (0,002)	D (0,545)	N (-0,38)
NOTCH2	p.Cys19Trp	c.57C>G	T (0,185)	B (0,002)	B (0,997)	N (-0,767)
NBAS	c.3212A>T	c.3212A>T	T (0,959)	D (1,000)	D (0,998)	D (-3,45)
CFTR	c.1251C>A	c.1251C>A	T (0,510)	B (0,000)	B (0,997)	N (0,33)
KRT18	**p.Gly17Glyfs77***	**c.50_51insG**	**-**	**-**	**D (1,000)**	**-**
PMM2	c.110A>T	c.110A>T	D (0,028)	B (0,007)	D (0,939)	D (-3,13)
MYO5B	c.2292C>G	c.2292C>G	T (1,000)	B (0,000)	D (0,999)	N (0,82)

SIFT: T – Tolerated, D – Damaging; PolyPhen2: B – Benign, D – Porobably Damaging; Mutation Taster: B – Benign, D – Disease causing; Provean: N – Neutral, D – Deleterious.

Phần mềm SIFT có giá trị điểm SIFT Score nằm trong khoảng từ 0 đến 1, giá trị này càng nhỏ thì đột biến càng có khả năng gây hại. SIFT đánh giá một đột biến là lành tính – T (Tolerated) nếu điểm SIFT Score > 0,05, gây bệnh – D (Damaging) nếu SIFT Score < 0,05. Phần mềm PolyPhen 2 có giá trị PolyPhen Score nằm trong khoảng từ 0 đến 1, giá trị này càng lớn thì đột biến càng có khả năng gây hại. Polyphen2 đánh giá đột biến là lành tính – B (Benign) nếu đột biến có PolyPhen Score < 0,452, có khả năng gây bệnh – P (Possibly Damaging) nếu PolyPhen Score nằm trong khoảng từ 0,453 đến 0,956, gây bệnh – D (porobably damaging) nếu PolyPhen Score lớn hơn 0,956. Phần mềm Provean có giá trị cutoff là -2,5. Nếu điểm đột biến có điểm Score > - 2,5 thì

được đánh giá là lành tính – N (Neutral), nếu đột biến có điểm Score < -2,5 thì được phần mềm đánh giá là gây bệnh – D (Deleterious).

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. Danh sách các đột biến được thể hiện ở Bảng 3.10.

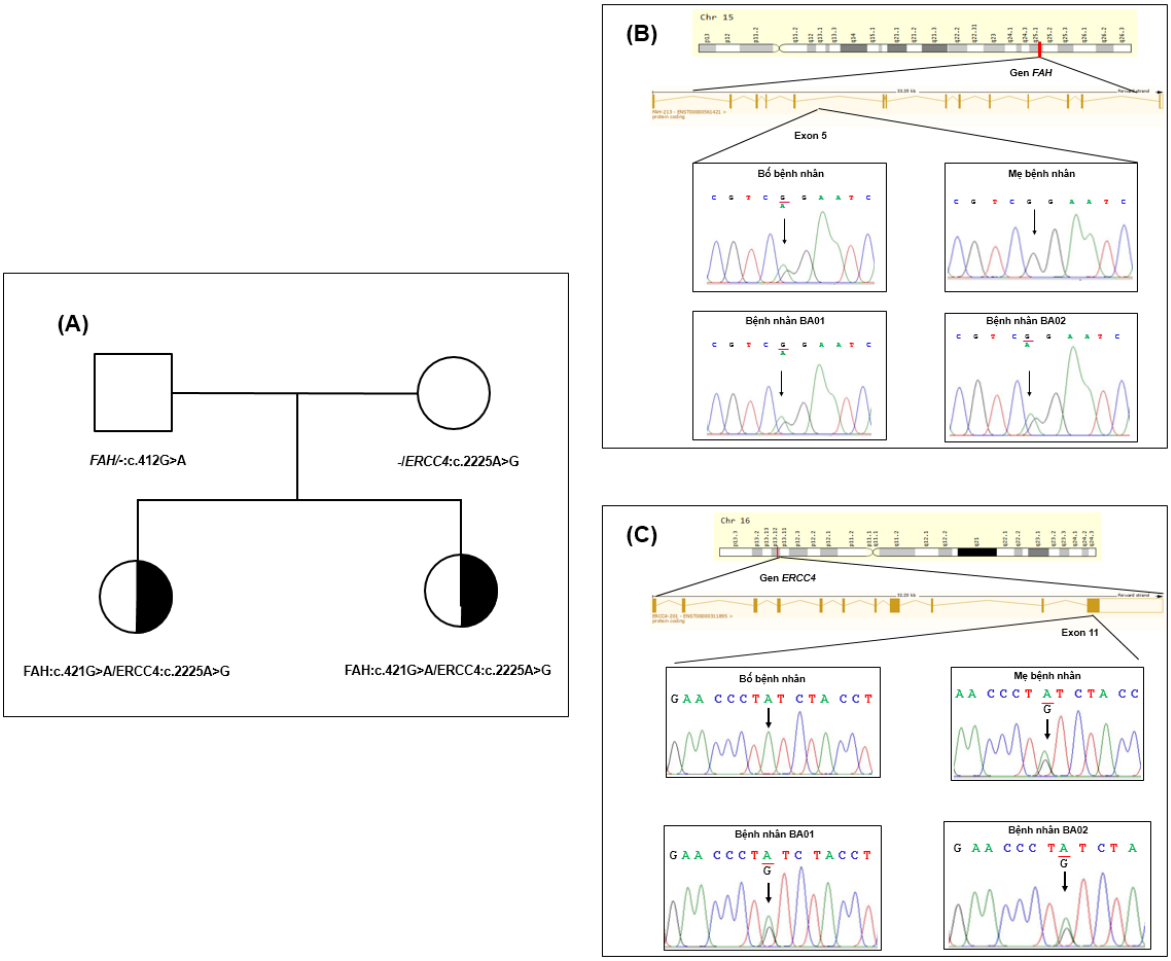
Bảng 3.10: Các đột biến có tiềm năng gây bệnh

Gen	NST	HGVS.p	HGVS.c	Loại đột biến	Trạng thái
Mẫu BA01 và BA02					
FAH	15	p.Gly138Arg	c.412G>A	Sai nghĩa	Dị hợp tử
ERCC4	12	p.Tyr742Cys	c.2225A>G	Sai nghĩa	Dị hợp tử
Mẫu BA03 và BA04					
CYP27A1	2	p.Lys123Thr	c.368A>C	Sai nghĩa	Dị hợp tử
Mẫu BA05					
KRT18	12	p.Gly17Glyfs77∗	c.50_51insG	Dịch khung	Đồng hợp tử

3.5.1. Bệnh nhân BA01 và BA02

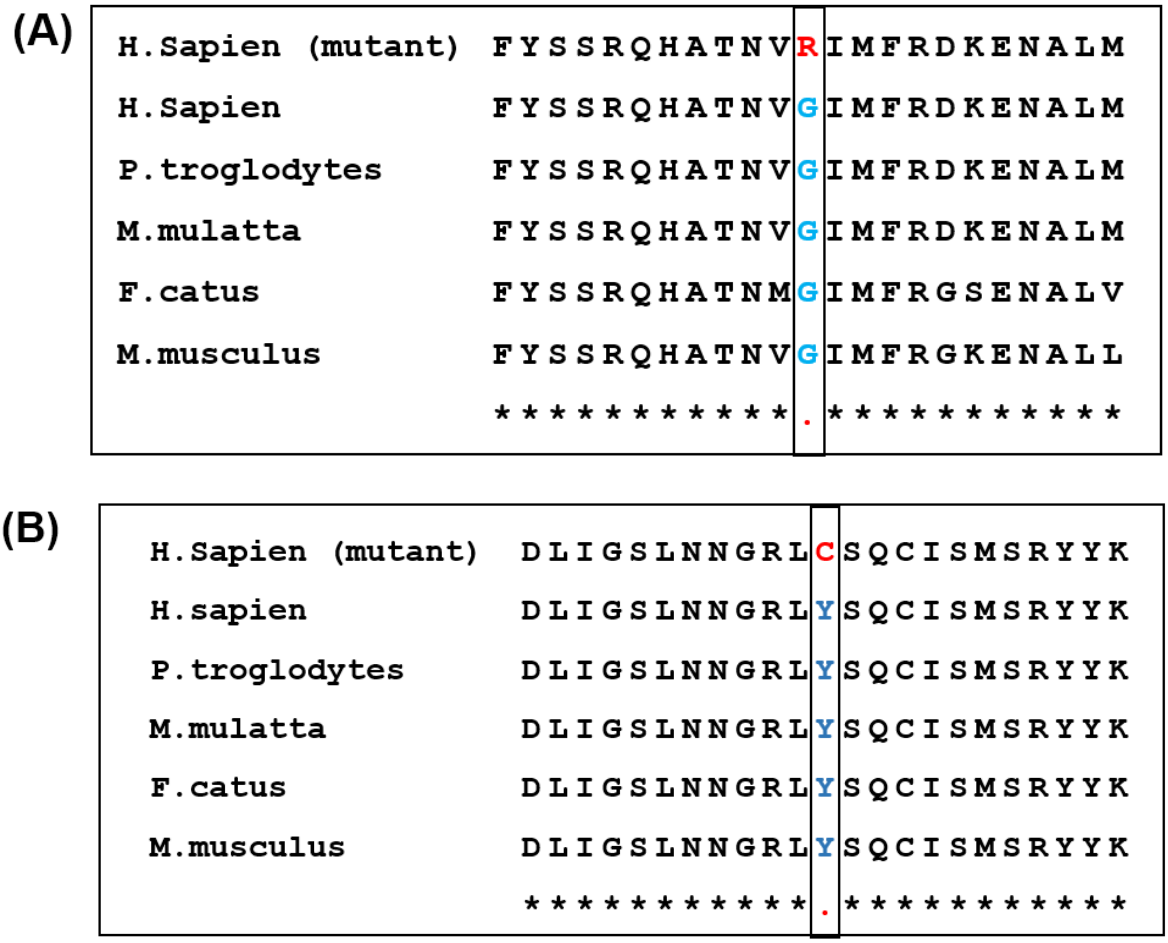
Hai bệnh nhân BA01 và BA02 là chị em ruột trong một gia đình. Hai bệnh nhân này mang đột biến dị hợp tử kép (double heterozygous) trên hai gen *FAH* và *ERCC4*. Trong đó, cả hai bệnh nhân di truyền đột biến c.412G>A

(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 và đột biến c.2225A>G trên gen ERCC4; (B) Kết quả giải trình tự Sanger đột biến c.412G>A trên gen FAH; (C) Kết quả giải trình tự Sanger đột biến c.2225A>G trên gen ERCC4

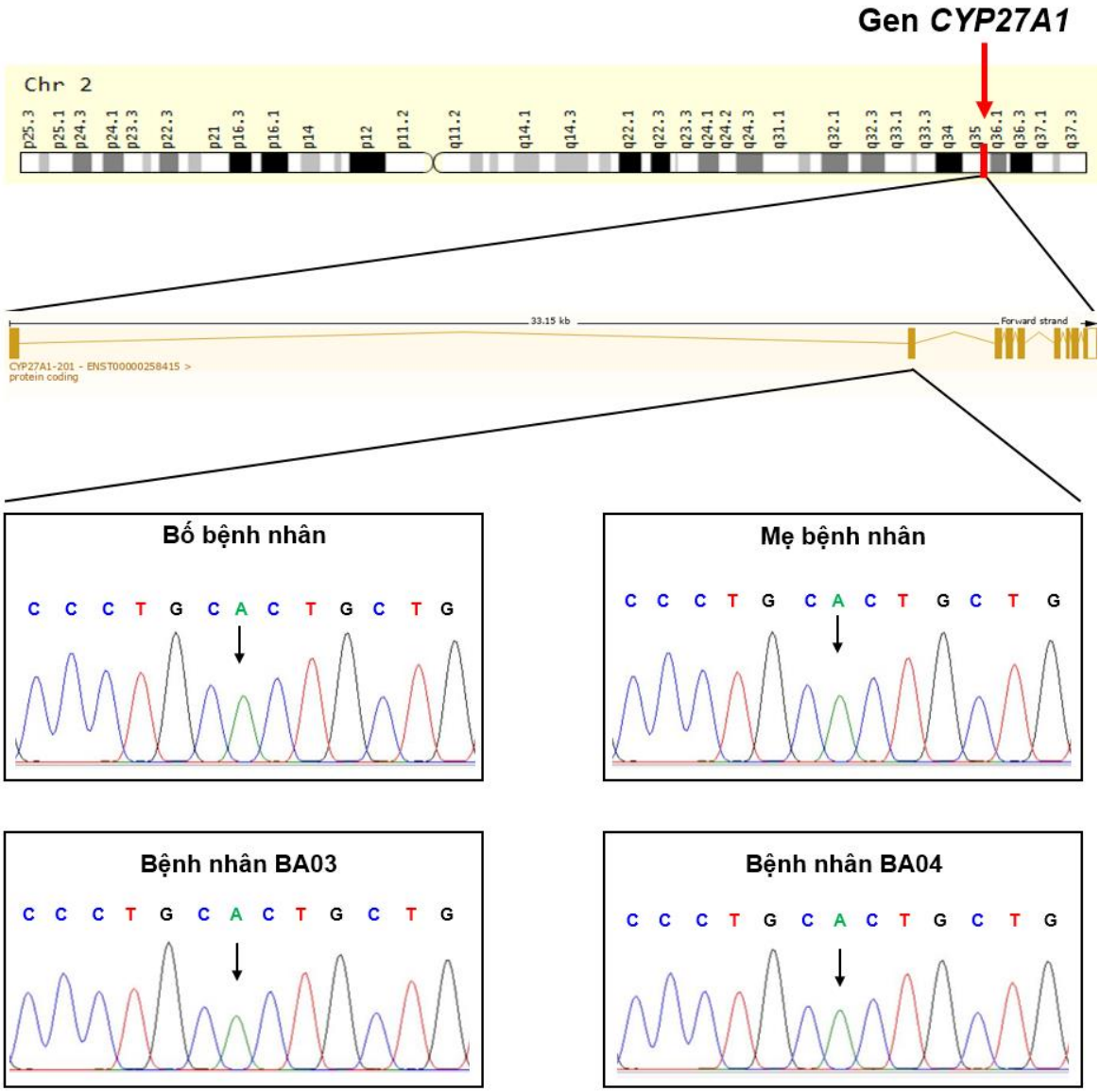
Ngoài ra, kết quả so sánh trình tự axit amin xung quanh vị trí đột biến giữa các loài bao gồm thể đột biến, loài người (*H.sapien*), tinh tinh (*P.troglodytes*), khỉ vàng (*M.mulatta*), mèo (*F.catus*) và chuột nhắt (*M.musculus*) cho thấy Gly138 trên protein mã hóa bởi gen *FAH* và Tyr742 trên protein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

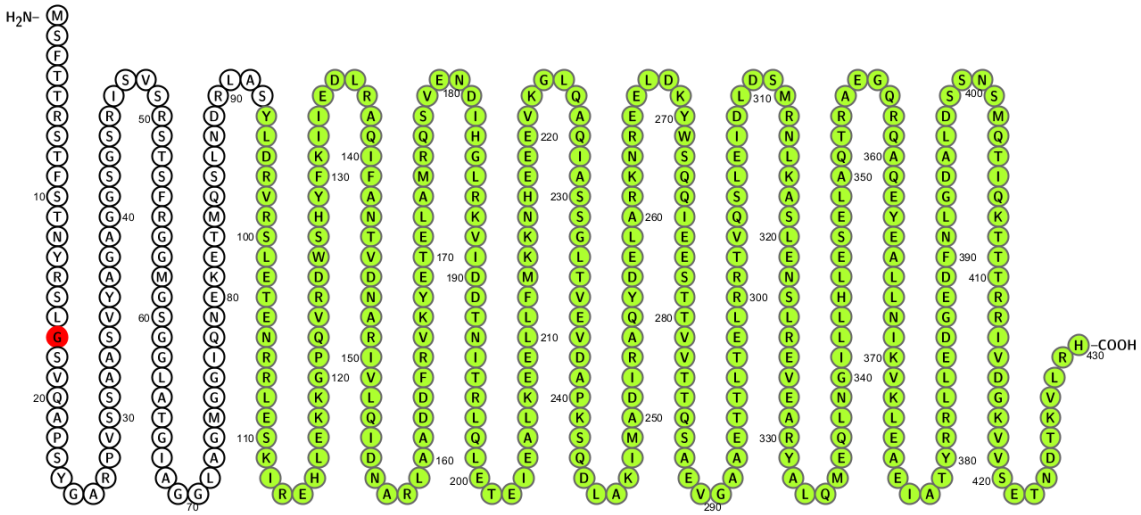
Kết quả phân tích dữ liệu WES cho thấy hai bệnh nhân BA03 và BA04 mang đột biến c.368A>C dạng dị hợp tử trên exon 2 của gen *CYP27A1*, đột biến làm thay đổi trình tự amino axit ở vị trí 123 từ Lysine thành Threonine. Tuy nhiên, kết quả giải trình tự bằng Sanger cho thấy cả hai bệnh nhân đều có kiểu gen bình thường, không mang đột biến này (Hình 3.5). Do đó, chúng tôi không xác định được đột biến gây bệnh trên hai bệnh nhân BA03 và BA04.



Hình 3.5: Kết quả giải trình tự Sanger cho thấy bệnh nhân BA03 và BA04 không mang đột biến c.368A>C dạng dị hợp tử trên gen CYP27A1

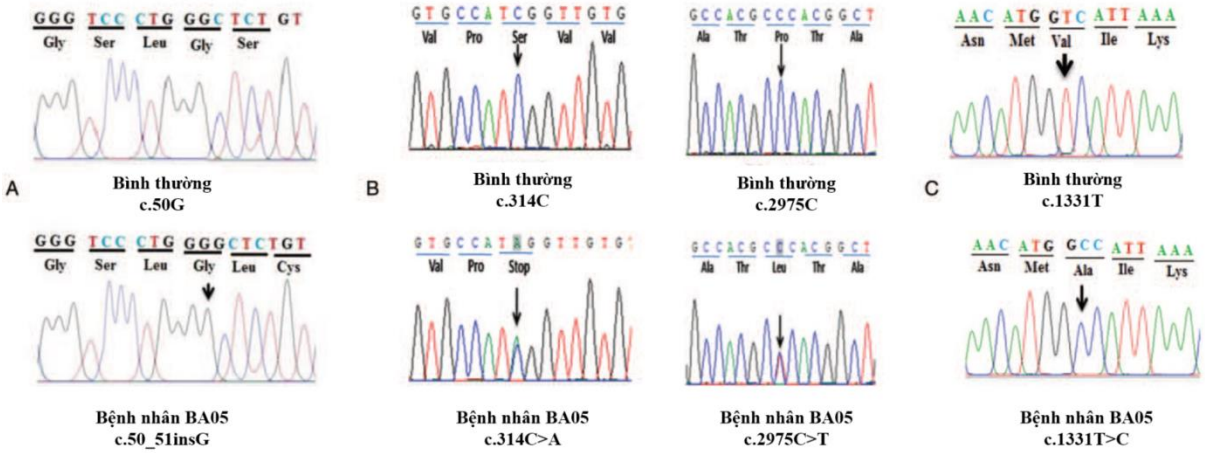
3.5.3. Bệnh nhân BA05

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 3.7). Phần mềm Mutation Taster dự đoán đây là đột biến gây bệnh. Đây là đột biến mới, chưa được báo cáo trên các cơ sở dữ liệu ClinVar và dbSNP. 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 (Hình 3.6).



Hình 3.6: Đột biến c.50_51insG xảy ra ở amino axit thứ 17 (màu đỏ) làm protein bị cắt ngắn sau vị trí đột biến 77 amino axit (phần màu xanh)

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 trên những gen liên quan đến bệnh Wilson. 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 nucleotide đơn rs2287622 trên gen *ABCB11* được xác định bằng kỹ thuật RFLP. Đoạn gen *ABCB11* có chứa đa hình nucleotide đơn có độ dài 333 bp được khuếch đại bằng phản ứng PCR với cặp mồi đặc hiệu. Dựa vào số lượng và kích thước băng điện di DNA sau khi xử lý với enzyme *Hae*III kiểu gen của mẫu nghiên cứu sẽ được xác định là TT, TC hay CC. Kết quả điện di của mẫu có kiểu gen TT sẽ có 1 băng với kích thước 333bp, mẫu có kiểu gen CC sẽ có 2 băng với kích thước 206 bp và 127 bp, mẫu có kiểu gen TC sẽ có cả 3 băng với kích thước 333 bp, 206 bp và 127 bp. Thống kê kiểu gen và tần suất alen của đa hình nucleotide đơn rs2287622 trên gen *ABCB11* 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	+

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

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		

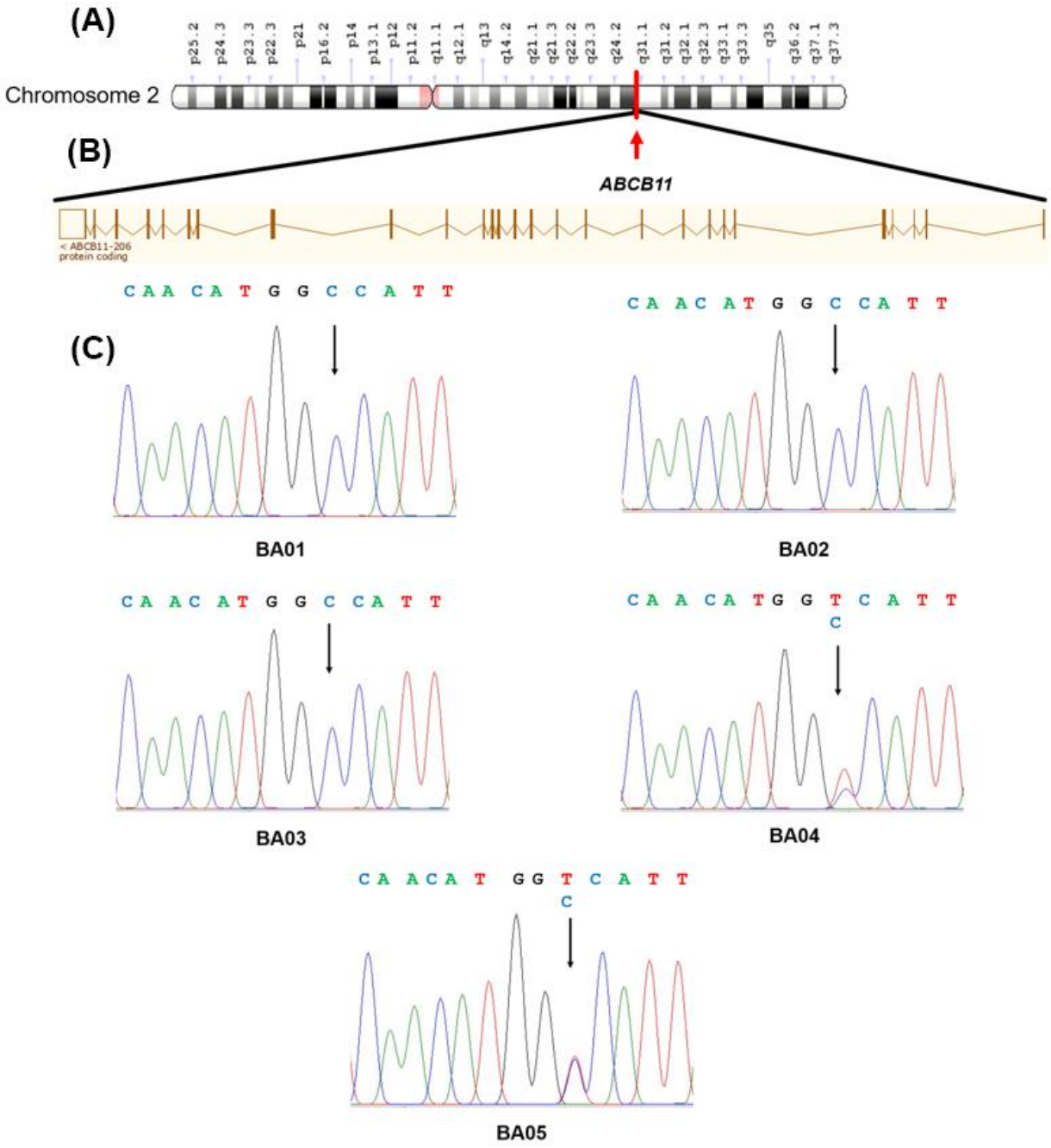
Kiểu gen	Nhóm bệnh	Nhóm đối chứng	OR	95% CI	Giá trị P
CC	80 (30,08%)	46 (18,40%)	1,91	1,26-2,88	0,002
Alen					
T	242 (45,49%)	293 (58,60%)	1		
C	290 (54,51%)	207 (41,40%)	1,69	1,33-2,17	0,0001

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

Có sự khác biệt đáng kể giữa tần số kiểu gen và alen C ở nhóm bệnh và nhóm đối chứng ở cả ba mô hình nghiên cứu. Sự khác nhau này có liên quan đến việc làm tăng nguy cơ mắc bệnh teo đường mật so với alen T. Ở mô hình cộng gộp, kiểu gen CC làm tăng nguy cơ mắc bệnh so với kiểu gen TT (OR = 2,76, 95%CI: 1,69-4,53, p<0,01). Sự xuất hiện kiểu gen TC trong mô hình này cũng làm tăng nguy cơ mắc bệnh teo đường mật bẩm sinh, nhưng ít hơn so với kiểu gen CC (OR = 1,79, 95% CI: 1,18-2,73, 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 = 13,13, 95% CI: 8,44-20,39, 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 = 1,91, 95% CI: 11,26-2,88, p<0,01). Xét tổng thể, sự xuất hiện của alen C làm tăng nguy cơ mắc bệnh teo đường mật bẩm (OR = 1,69, 95% CI: 1,33-2,17, p<0,01) so với alen T.

.



Hình 3.8: Phân tích di truyền đa hình nucleotide đơn rs2287622 ở 05 bệnh nhân teo đường mật bẩm sinh được lựa chọn để giải trình tự WES

(A) Gen ABCB11C nằm ở vị trí q31.1 trên nhánh dài của NST 2. (B) Sơ đồ vùng mã hóa và vùng không mã hóa của gen ABCB11. (C) Kết quả giải trình tự Sanger cho thấy đa hình xuất hiện ở cả 05 bệnh nhân

Kết quả giải trình tự Sanger cho thấy đa hình nucleotide đơn rs2287622 trên gen *ABCB11* xuất hiện ở cả 5 bệnh nhân teo đường mật bẩm sinh được

chọn để giải trình tự toàn bộ vùng gen mã hóa, trong đó biến thể này xuất hiện dạng đồng hợp tử ở các bệnh nhân BA01, BA02 và BA03, dạng dị hợp tử ở bệnh nhân BA04 và BA05 (Hình 3.8).

3.6.2. Ảnh hưởng của đa hình rs927344

Trong nghiên cứu này, kiểu gen của đa hình nucleotide đơn rs927344 trên gen *ABCC2* được xác định bằng phương pháp ARMS-PCR. Kiểu gen của các mẫu tham gia nghiên cứu được ghi nhận dựa vào số băng và kích thước sản phẩm PCR quan sát được khi chạy điện di. Theo thiết kế thí nghiệm, hình ảnh điện di trên gel agarose của những mẫu có kiểu gen AA sẽ xuất hiện 2 băng với kích thước 401 bp và 282 bp, mẫu có kiểu gen TT sẽ có 2 băng với kích thước 401 bp và 157 bp, mẫu có kiểu gen dị hợp tử AT sẽ có 3 băng với kích thước 401 bp, 282 bp và 157 bp. Thống kê kiểu gen và tần suất alen của đa hình rs927344 ở tất cả các mẫu được thể hiện trong Bảng 3.13. Kiểm định Chi bình phương cho thấy sự phân bố kiểu gen của đa hình này được xác định tuân theo định luật cân bằng di truyền HardyWeinberg với giá trị p>0,05, từ đó khẳng định tính ngẫu nhiên và đại diện của quần thể tham gia trong nghiên cứu.

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

	Kiểu alen			**Tần suất alen**		**HWE giá trị p**	**HWE**
	AA	**AT**	**TT**	**A**	**T**		
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 nguy cơ mắc bệnh teo đường mật bẩm sinh được đánh giá bằng phép kiểm nghiệm Chi-square trên cả ba mô hình: cộng gộp, trội và mô hình lặn (Bảng 3.14).

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 của phép thống kê Chi bình phương cho thấy ở cả ba mô hình cộng gộp, trội và mô hình lặn giá trị p của kiểm nghiệm 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 trên gen *ABCC2* 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

Trong nghiên cứu này, kiểu gen của SNP rs1815930 trên gen *MYO5B* được xác định thông quan kỹ thuật ARMS-PCR như đã trình bày ở phần phương pháp. Kiểu gen của các mẫu tham gia nghiên cứu được ghi nhận dựa vào số băng và kích thước sản phẩm PCR quan sát được khi chạy điện di. Hình ảnh điện di trên gel agarose của những mẫu có kiểu gen AA sẽ xuất hiện 2 băng với kích thước 349 bp và 246 bp, mẫu có kiểu gen GG sẽ có 2 băng với kích thước 349 bp và 154 bp, mẫu có kiểu gen dị hợp tử AG sẽ có 3 băng với kích thước 349 bp, 246 bp và 154 bp. Thống kê kiểu gen và tần suất alen của đa hình rs1815930 ở tất cả các mẫu được thể hiện trong Bảng 3.15. 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 Hardy Weinberg với giá trị p > 0,05, từ đó khẳng định tính ngẫu nhiên và đại diện của quần thể người Việt tham gia trong nghiên cứu.

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 liên quan giữa của đa hình nucleotide đơn rs1815930 trên gen *MYO5B* và khả năng mắc bệnh teo đường mật bẩm sinh được đánh giá thông qua phép kiểm nghiệm Chi-square trên cả ba mô hình: cộng gộp, trội và mô

hình lặn. Kết quả phân tích, đánh giá trên các mô hình này được trình bày ở 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 về sự liên quan giữa đã hình rs1815930 trên gen *MYO5B* với nguy cơ mắc bệnh teo đường mật bẩm sinh cho thấy ở cả ba mô hình cộng gộp, mô hình trội và mô hình lặn giá trị p của phép kiểm định Chi-square đều lớn hơn 0,05. Do đó, sự khác biệt về tỉ lệ kiển gen, kiểu alen trong phép kiểm định này không có ý nghĩa thống kê. Như vậy, tỉ lệ kiểu gen và kiểu alen ở đa hình rs1815930 trên gen *MYO5B* không có mối 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

Dịch mật là dịch đẳng trương với huyết tương, có thành phần chủ yếu là nước, các chất điện giải, chất hữu cơ như muối mật, phospholipid, bilirubin cholesterol, axit amin, steroid, enzyme, porphyrin, vitamin và kim loại nặng. Trong mật cũng có thể chứa thuốc ngoại sinh, xenobiotic và chất độc môi trường. Mỗi ngày, gan của người khỏe mạnh tạo ra khoảng 500 – 600ml mật. Sự hình thành mật bắt đầu từ ống gan (khoảng 75% sản lượng mật hàng ngày) sau đó là những thay đổi dọc theo ống mật (chiếm khoảng 25% sản lượng mật). Sự hấp thu của các chất khác nhau qua màng đáy được theo sau bởi các phản ứng liên hợp có sự tham gia của enzyme UGT1A1. Ống mật hình thành quá trình lọc thẩm thấu của nước và chất điện giải để đáp ứng với gradient thẩm thấu được tạo ra bởi hệ thống vận chuyển tích cực muối mật và các anion hữu cơ khác. Ở giai đoạn sau, sự bài tiết xảy ra và các chất bao gồm cả bilirubin được tiết vào mật – quá trình này có sự tham gia của protein MRP2 cùng một số chất vận chuyển khác tại màng tế bào gan. Sự bài tiết mật phụ thuộc vào: (1) hệ thống vận chuyển xuyên màng hoạt động tạo ra độ dốc thẩm thấu; (2) bộ máy tế bào cho phép các túi dịch chuyển động và sự co bóp của ống mật; (3) các điểm tiếp xúc của tế bào bịt kín các tiểu quản và duy trì sự phân cực của tế bào; (4) cơ chế truyền tín hiệu điều hòa và điều phối quá trình hình thành mật. Việc vận chuyển muối mật qua gan được thực hiện bởi một hệ thống bơm bài tiết muối mật (BSEP). Sự bài tiết mật từ tế bào gan bị suy giảm hoặc dòng chảy mật bị tắc nghẽn đều dẫn đến tình trạng ứ mật. Nguyên nhân phổ biến nhất gây ứ mật ở trẻ sơ sinh và trẻ nhỏ là teo đường mật (BA), thiếu hụt các ống mật ngoài gan với nguyên nhân không rõ ràng [126]. Theo đó, người ta ước tính rằng 25–50% trường hợp ứ mật xảy ra do các biến thể gen có thể xác định [127]. Những biến thể gen này liên quan đến nhiều loại gen có ảnh hưởng trực tiếp hoặc gián tiếp đến quá trình tổng hợp, vận chuyển và dòng chảy của mật. Đối với một số tính trạng lặn trên nhiễm sắc thể thường liên quan đến bệnh teo đường mật bẩm sinh và tình trạng ứ mật, thông thường bệnh nhân phải có kiểu gen đồng hợp tử (mang 2 alen lặn) mới bị ảnh hưởng để phát triển bệnh. Tuy nhiên, gần đây người ta đã đề xuất rằng ngay cả những thay đổi dị hợp tử ở một số gen nhất định cũng có thể dẫn đến tình trạng ứ mật, đặc biệt là khi có thêm

tác động của các yếu tố bên ngoài như thuốc, hormone hoặc các chất trung gian gây viêm [128]. Ví dụ như các biến thể gây bệnh dạng dị hợp tử trên gen *ATP8B1, ABCB11* và *ABCB4* [129,130]. Vai trò của các biến thể dị hợp tử trong nguyên nhân gây ra tình trạng ứ mật rất phức tạp do không có sự chắc chắn trong việc dự đoán khả năng gây bệnh. Do đó, mối tương quan giữa các kết quả lâm sàng, sinh hóa và di truyền là cần thiết để chỉ ra nguyên nhân gây bệnh. Theo đó, một biến thể dị hợp tử vẫn có thể gây ra bệnh theo ba cơ chế. Đầu tiên một biến thể có thể dẫn đến việc tạo ra một gen bị biến đổi với chức năng phân tử mới hoặc có thể làm thay đổi kiểu biểu hiện gen. Thứ hai là một biến thể dị hợp tử có thể khiến cho kiểu gen bị thay đổi cản trở chức năng của kiểu gen ban đầu. Thứ ba, một biến thể dị hợp tử dạng mất chức năng có thể dẫn đến hiện tượng đơn bội, trong đó lượng kiểu gen bình thường được tạo ra bởi alen bình thường không đủ để duy trì kiểu hình bình thường [131].

Trong bệnh teo đường mật bẩm sinh, sự phá hủy các ống dẫn mật bên trong và bên ngoài, kết quả của quá trình viêm không rõ nguyên nhân, dẫn đến xơ hóa, tắc mật tiến triển và xơ gan. Sự phá hủy và tiến triển nhanh chóng của xơ hóa đường mật, có thể do ứ mật không ngừng và giữ lại các thành phần của mật bao gồm cả axit mật [132]. Tình trạng ứ mật có thể ảnh hưởng đến hiệu quả của phẫu thuật Kansai và sự phục hồi chức năng đường mật sau này. Ở trẻ sơ sinh và trẻ nhỏ, tình trạng ứ mật có thể xảy ra do nhiều tình trạng bệnh lý có biểu hiện tương tự hoặc chồng chéo. Điều này có thể khiến việc chẩn đoán dựa trên các đặc điểm lâm sàng, sinh hóa, X quang và mô học trở nên khó khăn. Mối quan hệ giữa teo đường mật bẩm sinh và các bệnh liên quan đến trình trạng ứ mật khác đã được tìm thấy [133,134]. Bên cạnh đó, yếu tố di truyền còn tham gia vào quá trình thích ứng và đáp ứng với các bệnh lý đường mật và ứ mật để cung cấp các mục tiêu điều trị và sàng lọc di truyền [132].

Do chi phí giải trình tự DNA giảm nhanh chóng, nhiều bệnh và hội chứng rối loạn di truyền hiếm gặp hoặc chưa được chẩn đoán trước đây hiện có thể được phát hiện. Trong 10 năm qua, công nghệ giải trình tự gen thế hệ mới đã tạo ra một cuộc cách mạng trong nghiên cứu di truyền học ở con người với mức độ chính xác cao, hiệu quả chi phí và khả năng thông lượng cao. NGS đang dần trở thành một tiêu chuẩn trong thực hành chẩn đoán thông thường [125]. Trong đó, giải trình tự toàn bộ vùng mã hóa được coi là công cụ đáng tin cậy để hỗ

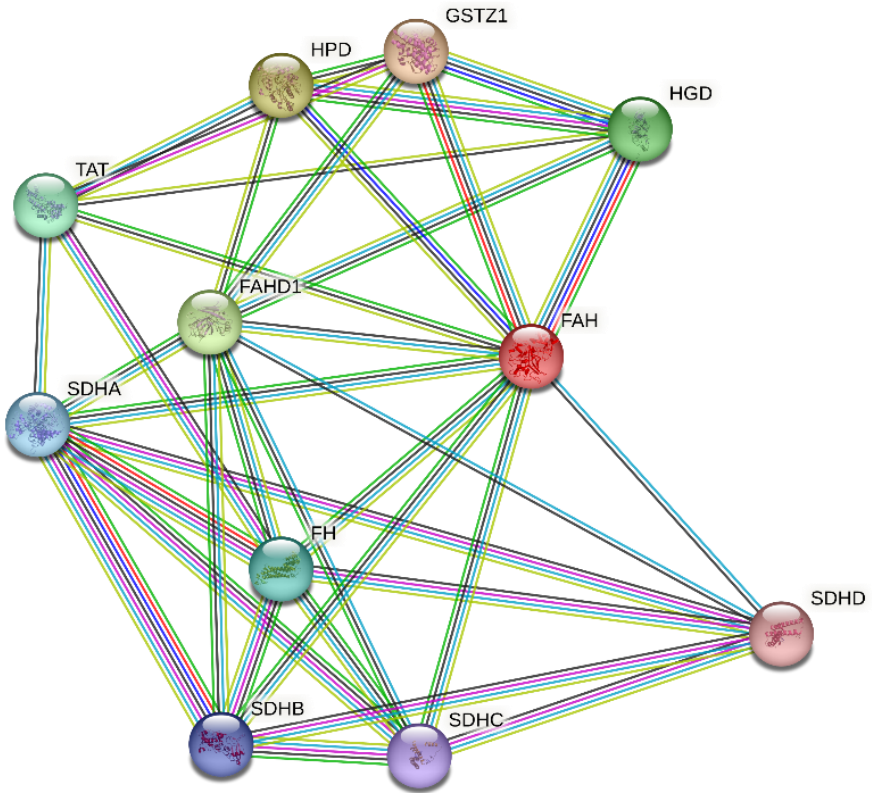
trợ chẩn đoán đối với những trường hợp có triệu chứng lâm sàng không rõ ràng. Đối với các bệnh về gan mật, NGS có vai trò ngày càng quan trọng, giúp chẩn đoán các loại rối loạn ứ mật ở trẻ sơ sinh, xác định nguyên nhân di truyền từ đó phân biệt những bệnh/hội chứng có kiểu hình lâm sàng tương tự nhau. Với đặc điểm là tốc độ giải trình tự nhanh và thông lượng cao, giải trình tự gen thế hệ mới cho phép phát hiện ra các gen và biến thể gen mới liên quan đến bệnh teo đường mật bẩm sinh và các bệnh gan mật khác có nhiều triệu chứng lâm sàng giống nhau.

Giải trình tự toàn bộ vùng mã hóa đã chỉ ra nhiều gen có tiềm năng liên quan đến bệnh teo đường mật bẩm sinh như gen mã hóa protein vận chuyển ABC (ATP-binding cassette transporters) thuộc siêu họ ABC [135] hoặc những gen liên quan đến con đường tín hiệu Notch như *JAG1* [136] và *NOTCH2* [16]. Rajagopalan và đồng tác giả đã thực hiện phân tích trình tự vùng mã hóa ở 101 bệnh nhân BA từ đó tìm thấy các đột biến *de novo* trên gen *REV1* nằm trên nhiễm sắc thể 2 và trên gen *STIP1* trên nhiễm sắc thể 11 [137]. Gen *REV1* mã hóa một protein chịu trách nhiệm sửa chữa DNA trong khi đó *STIP1* là một là một chaperone hỗ trợ việc chuyển các protein từ HSP70 sang HSP90. Cũng sử dụng phương pháp giải trình tự toàn bộ vùng mã hóa, Kubilay và các đồng tác giả đã xác định được nhiều đột biến ở hai gen *MAP2K3* và *PRIM2* [138] trên bệnh nhân teo đường mật bẩm sinh. Mặc dù vậy, không có nghiên cứu nào trước đó chứng minh sự liên quan giữa *MAP2K3* và *PRIM2* với bệnh teo đường mật bẩm sinh. Gen *PRIM2* tham gia vào quá trình tổng hợp các đoạn RNA sơ khai được sử dụng để tổng hợp các đoạn Okazaki trong quá trình tổng hợp DNA. Theo cơ sở dữ liệu Trực tuyến Mendelian Inheritance in Man (OMIM), gen này không có mối liên hệ nào với các bệnh lý đã được xác định trước đó. Gen còn lại, *MAP2K3*, mã hóa một enzym protein kinase có tính chất phân bào, và gen này cũng chưa từng được liên kết trước đây với 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* (fumarylacetoacetate hydrolase) nằm trên nhiễm sắc thể 15, chứa 14 exon và mã hóa cho 420 axit amin. Gen mã hóa enzyme fumarylacetoacetase là enzyme cuối cùng trong chuỗi 5 enzyme phân hủy axit amin tyrosine. Enzyme này có nhiều trong gan và thận, và một lượng nhỏ hơn được tìm thấy trong nhiều mô khắp cơ thể. Phân tích bằng cơ sở dữ liệu mạng tương tác protein STRING v.11 [139,140] cho thấy protein mã hóa bởi gen *FAH* tương tác mạnh với các protein mã hóa bởi các gen *GSTZ1* (score: 0,997), *HPD* (score: 0,993) và *FAHD1* (0,974), *HGD* (score: 0,945), *TAT* (score: 0.935) (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

Cho tới nay, có hơn 100 đột biến trên gen *FAH* được ghi nhận trên cơ sở dữ liệu đột biến gen ở người HGMD [141], trong đó chủ yếu là đột biến ở dạng sai nghĩa và đột biến điểm cắt nối mRNA (63 đột biến sai nghĩa, 25 đột biến điểm cắt nối mRNA). Các đột biến trên gen *FAH* thường được biết đến là nguyên nhân gây bệnh Tyrosinemia type I với những kiểu hình đặc trưng như vàng da, ứ mật, gan và lá lách to bất thường. Năm 2020, nhóm nghiên cứu Kubilay và các đồng tác giả thực hiện giải trình tự toàn bộ vùng mã hóa nhằm xác định các gen đột biến phổ biến trên 10 bệnh nhân teo đường mật bẩm sinh từ 55 đến 70 ngày tuổi đã được phẫu thuật Kasai [138]. Kết quả cho thấy có nhiều đột biến chung xuất hiện đồng thời ở 10 bệnh nhân, trong đó có 1 biến thể nằm trên gen *FAH*. Mặc dù không đưa ra được cơ chế rõ ràng ảnh hưởng các đột biến gen đến tình trạng bệnh, nhóm nghiên cứu này cho rằng bệnh có thể là bệnh lý do sự tương tác của các đột biến trong một gen hoặc nhiều gen. Vì 10 bệnh nhân tham gia nghiên cứu không có bất thường hoặc bệnh lý nào khác nên nhóm nghiên cứu giả thiết rằng teo đường mật bẩm sinh có thể xảy ra khi tất cả các đột biến đồng thời xuất hiện.

Gen *ERCC4* mã hóa protein sửa chữa DNA endonuclease XPR, còn được gọi là FANCQ. Endonuclease XPF mã hóa bởi *ERCC4* cùng với protein điều hòa của nó được mã hóa bởi *ERCC1* tạo thành phức hợp enzyme ERCC1 – XPR tham gia vào quá trình sửa chữa và tái tổ hợp DNA, đóng vai trò thiết yếu trong quá trình sửa chữa NER (nucleotide excision repair) phổ biến [142,143]. Cơ sở dữ liệu đột biến gen ở người HGMD ghi nhận 121 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. Các đột biến trên gen *ERCC4* được biết đến là nguyên nhân gây ra bệnh khô da sắc tố, một số loại ung thư, hội chứng Cockayne và bệnh thiếu máu Fanconi. 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. Nhóm tác giả này hướng đến 19 gen có liên quan đến tình trạng teo và ứ mật ở trẻ sơ sinh, trong đó tập trung vào những biến thể có tần số alen nhỏ hơn 0,01 ở cơ sở dữ liệu dbSNP150. Trong số 20 bệnh nhân được nghiên cứu, có 13 biến thể hiếm gặp đã được phát hiện ở 9 gen: 4 đột biến trên gen *JAG1* (hội chứng Alagille),

2 đột biến trên gen *MYO5B* (ứ mật trong gan tiến triển trong gia đình [PFIC] loại 6) và một đột biến trên gen *ABCB11* (PFIC loại 2), *ABCC2* (Hội chứng Dubin-Johnson), *ERCC4* (thiếu máu Fanconi), *KCNH1* (hội chứng Zimmermann-Laband), *MLL2* (hội chứng Kabuki), *RFX6* (hội chứng Mitchell-Riley) và *UG1A*1 (hội chứng Crigler-Najjar). Đột biến c.241G > T (p.Val81Phe) trên gen *ERCC4* có mã số rs55761944 phát hiện bởi Sangkhathat là đột biến duy nhất liên quan đến bệnh teo đường mật bẩm sinh ngoài gan được ghi nhận trên cở sở dữ liệu HGMD. 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 [144]. Trong nghiên cứu này, hai bệnh nhân BA01 và BA02 mang đột biến c.2225A>G (p.Tyr742Cys) trên gen *ERCC4*. Phần mềm CADD đánh giá c.2225A>G (p.Tyr742Cys) trên gen *ERCC4* mức điểm Phred score là 26,6 cho thấy đây là đột biến nằm trong nhóm 1% có hại nhất đối với chức năng protein.

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. Trong quá trình phát triển, các tế bào biểu mô đường mật trải qua quá trình tăng sinh và tái cấu trúc rộng rãi, đồng thời có thể trải qua những thay đổi về cấu trúc hoặc chức năng hơn nữa trong quá trình bệnh lý. Cơ sở dữ liệu HGMD ghi nhận 19 biến thể trên gen *KRT18* chủ yếu xảy ra ở dạng sai nghĩa. Các biến thể trên gen này được biết là có liên quan đến các bệnh về gan, tổn thương gan do thuốc. Nghiên cứu của Benjamin cùng các đồng tác giả trên mô hình cá ngựa vằn cho thấy *KRT18* được biểu hiện mạnh ở hệ thống đường mật đang phát triển [145]. 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 [146]. 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

Gen *ABCB11* nằm trên nhánh dài của nhiễm sắc thể số 2, tại vị trí q31.1. Gen có độ dài 110.54 kb, gồm 28 exon. Gen mã hóa cho protein BSEP có chức năng bơm bài xuất muối mật, vận chuyển muối mật ra khỏi gan. Muối mật là một thành phần của mật, được sử dụng để tiêu hóa chất béo. Muối mật được sản xuất bởi các tế bào gan và sau đó được vận chuyển ra khỏi tế bào để tạo ra mật. Sự giải phóng muối mật từ các tế bào gan là rất quan trọng đối với việc tiết mật bình thường. Cùng với các protein đa kháng thuốc như MDR1 (gen *ABCB1*), MDR3 (gen *ABCB4*) và MRP2 (gen *ABCC2*), bơm bài xuất muối mật (BSEP, gen *ABCB11*) là những chất vận chuyển dòng mật ra ngoài quan trọng nhất. Các biến thể di truyền di truyền xảy ra trên các chất vận chuyển họ ABC này có thể điều chỉnh biểu hiện của chất vận chuyển và có thể liên quan đến nhiều bệnh gan ứ mật di truyền và mắc phải. Ứ mât, tăng bilirubin máu xảy ra khi có sự rối loạn cân bằng nội môi của bilirubin và axit mật do các protein tham gia vào quá trình tổng hợp, hấp thu, giải độc và vận chuyển axit mật và bilirubin. Siêu họ của các chất vận chuyển cassette gắn ATP ở gan (ABC) đóng một vai trò quan trọng trong việc xử lý axit mật, bilirubin và nhiều loại thuốc. Những khiếm khuyết trong tổng hợp hoặc chức năng của BSEP dẫn đến giảm bài tiết muối mật và tích tụ muối mật. Tầm quan trọng của BSEP được thể hiện qua tình trạng ứ mật di truyền gây ra bởi đột biến gen ở *ABCB11* từ mức độ ứ mật dai dẳng và nghiêm trọng cần ghép gan (như đã thấy ở ứ mật trọng gan tiến triển có tính chất gia đình) đến các dạng ứ mật nhẹ hơn (như tình trạng ứ mật trong gan tái phát lành tính - benign recurrent intrahepatic cholestasis (BRIC)). Cùng với gen *ABCB4*, các biến thể di truyền ở gen *ABCB11* được cho là có liên quan đến khoảng 20% bệnh nhân bị ứ mật nặng trong gan khi mang thai, khi các hormone thai kỳ thúc đẩy sự phát triển của tình trạng ứ mật, điển hình là trong tháng thứ hai của thai kỳ khi nồng độ estrogen cao nhất [84,85]. Các biến thể trên gen này gây ảnh hưởng đến hoạt động của protein, làm ức chế quá trình vẫn chuyển muối mật ra khỏi gan dẫn đến tình trạng ứ mật trong gan. Một số biến thể trong gen *ABCB11* được biết là có liên quan đến teo đường mật bẩm sinh ngoài gan. Sangkhathat và đồng tác giả phát hiện đột biến rs372939910 (c.2135A > G, p.Leu712Ser) trên bệnh nhân teo đường mật bẩm sinh người

Thái Lan [45]. Tuy nhiên, cho đến nay, chưa có nghiên cứu nào được thực hiện để đánh giá vai trò của biến thể rs2287622 đối với tính mẫn cảm với bệnh teo đường mật bẩm sinh.

Khảo sát đa hình nucleotide đơn rs2287622 trong gen *ABCB11* trên 266 bệnh nhân và 250 đối chứng trong nghiên cứu này 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 biqệt có ý nghĩa ở bệnh nhân Việt Nam mắc bệnh teo đường mật bẩm sinh. Cụ thể, người có mang alen C có nguy cơ mắc bệnh nhân cao hơn người mang alen T. 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. Mối liên quan tương tự giữa bệnh nhân có kiểu gen CC và nguy cơ ứ mật và ứ mật trong gan khi mang thai cũng đã được báo cáo [157], theo đó người mang alen C có nguy cơ mắc ứ mật trọng gan khi mang thai cao hơn người mang alen T (OR: 1,7, 95% CI: 1,4 - 2,1, p<0.001), ngoài ra, người có kiểu gen đồng hợp tử CC có nhiều khả năng mắc bệnh hơn người có kiểu gen TT (OR 2,8, 95% CI: 1,7 – 4,4, p<0.0001)). Tuy nhiên, một nghiên cứu tương tự ở người Nhật Bản không tìm thấy mối liên quan giữa tình trạng ứ mật với biến thể này [158]. Cho đến nay, cơ chế của sự liên quan giữa đa hình nucleotide đơn rs2287622 với tính nhạy cảm của tình trạng ứ mật chưa được giải thích rõ ràng. Tuy nhiên, đây là biến thể rất phổ biến, với tần số alen là 53% ở người Mỹ gốc Phi, 57% ở người Mỹ gốc Âu, và 74% hoặc cao hơn ở người Trung Quốc đại lục và các nhóm người châu Á khác. Dựa trên các công cụ *silico* được sử dụng để dự đoán tác động chức năng của SNP, đa hình nucleotide rs2287622 trên gen *ABCB11* được dự đoán là lành tính.

Mối liên quan giữa đa hình nucleotide đơn rs2287622 trên gen *ABCB11* với nguy cơ ứ mật, tăng bilirubin ở bệnh nhân tổn thương gan do thuốc được nghiên cứu [148,159]. Tuy nhiên nhóm tác giả không tìm được mối liên quan nào giữa đa hình nucleotide này với nguy cơ ứ mật, tăng bilirubin ở nhóm đối tượng này. 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. Đặc biệt, mối liên quan này chỉ tìm thấy ở những người trên 40 tuổi, những người trong độ tuổi từ 18 – 40 không phát hiện được mối liên quan này [160]. 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.

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* ở người nằm trên nhiễm sắc thể 10, tại vị trí 10q23 - q24 và chứa 32 exon. Hơn 200 biến thể trình tự đã được xác định ở exon, intron và ở vùng 5’ - và 3′- flanking của gen *ABCC2*. Nhiều biến thể trong số này là đa hình đơn nucleotide, không dẫn đến thay đổi axit amin và không gây hậu quả về mặt chức năng. Một số biến thể của gen này có thể ảnh hưởng đến hoạt động của nó, bao gồm c.-1549A>G (rs1885301), c.-24C>T (rs717620), c.1249G>A (rs2273697), c.3972C>T (rs3740066) và c.4544G>A (rs8187710) [147].

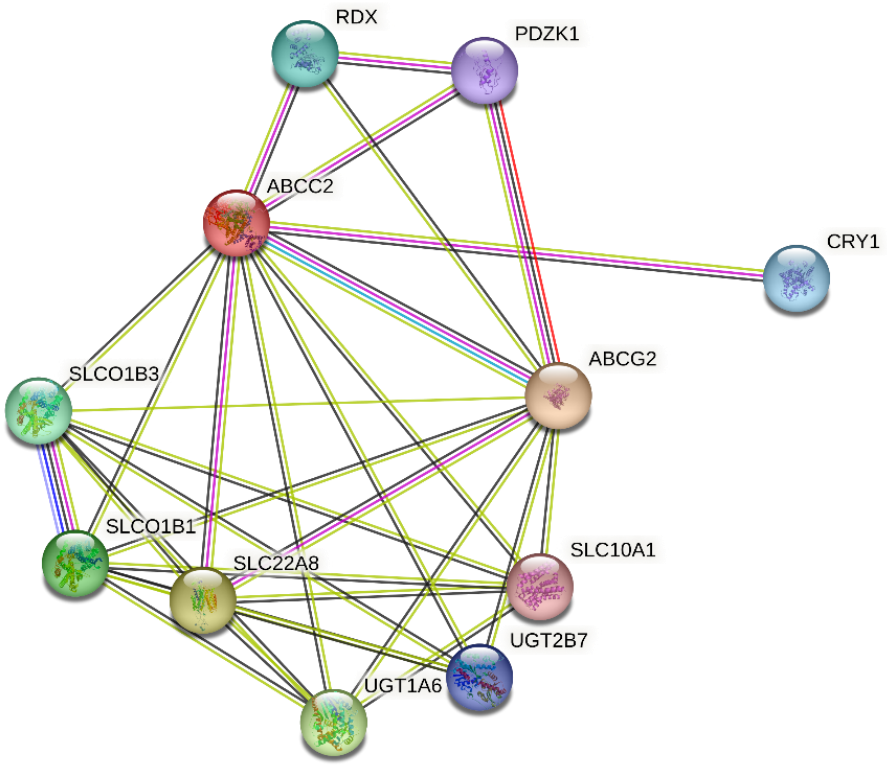
Các biến thể trên gen *ABCC2* làm rối loạn chức năng MRP2 dẫn đến làm tăng bilirubin trong máu. Có nghiên cứu đã chứng minh rằng biến thể *ABCC2* rs717620 T làm giảm hoạt động của chất MRP2 khoảng 40% [148], alen T của đa hình nucleotide đơn này liên quan đến nguy cơ tăng bilirubin trong máu ở những bệnh nhân tổn thương gan do thuốc (OR: 3,83, 95% CI: 1,73 – 8,48, p = 0.001). Đồng thời, nguy cơ ứ mật trong thai kỳ có liên quan đến đa hình nucleotide đơn của *MRP2* [149]. Biến thể trên gen *ABCC2* được cho là liên quan đến tình trạng ứ mật ở bệnh nhân mắc hội chứng Dubin-Johnson [150].

Các biến thể gen ở bệnh nhân mắc hội chứng Dubin-Johnson có liên quan đến sự vắng mặt của chức năng protein MRP2 ở tế bào gan bao gồm đột biến tại vị trí cắt nối mRNA, đột biến sai nghĩa dẫn đến thay thế axit amin, đột biến chèn/xóa và các đột biến vô nghĩa dẫn đến xuất hiện các stop codon sớm [86]. Các biến thể tại vùng cắt nối mRNA có thể dẫn đến sự trưởng thành kém của protein, khả năng chứa protein MRP2 ở màng ống tủy bị suy giảm hoặc protein mã hóa bởi gen *ABCC2* không có chức năng. Các đa hình nucleotide đơn xảy ra trên vùng promoter/enhancer hoặc trong protein bám DNA có thể dẫn đến mức độ biểu hiện MRP2 khác nhau.Việc chức năng của MRP2 bị ảnh hưởng cũng tác động đến dược động học của nhiều loại thuốc, chất độc nội sinh dẫn đến những thay đổi trong quá trình hấp thu, phân phối và thanh thải của các cơ quan, làm tăng khả năng gây độc và tương tác thuốc. Các biến thể gen trong MRP2 có liên quan đến sự thay đổi độ nhạy cảm với một số loại thuốc bao gồm thuốc chống ung thư, thuốc chống động kinh và kháng sinh [151]. Tuy nhiên, các nghiên cứu nhằm xác định mối tương quan giữa biểu hiện MRP2 và tác động lâm sàng của nó chưa đạt được thành công và do đó hậu quả chức năng của các biến thể di truyền này vẫn chưa được biết rõ.

Các biến thể trên gen *ABCC2* được ghi nhận trên cơ sở dữ liệu HGMD chủ yếu liên quan đến hội chứng Dubin – Jonson (80 biến thể), hội chứng Dubin – Jonson kết hợp với ứ mật trong gan khi mang thai (03 biến thể), các hội chứng ứ mật trong gan, vàng da từng đợt…Chỉ có 01 biến thể liên quan đến bệnh teo đường mật bẩm sinh được ghi nhận trên cơ sở dữ liệu này. Laochareonsuk và đồng tác giả phát hiện đột biến c.2601_2602insTA (p.Pro868TyrfsTer21) ở trạng thái dị hợp tử khi nghiên cứu trên các bệnh nhân teo đường mật bẩm sinh người Thái Lan bằng công nghệ giải trình tự gen thế hệ mới [152]. Nghiên cứu của Sangkhathat và đồng tác giả (2018) trên các bệnh nhân Thái Lan đã phát hiện đột biến rs370049107 (c.3379G > A, p.Val1127Ile) ở 2 trên tổng số 20 bệnh nhân. Những bệnh nhân có biến thể trên gen *ABCC2* cũng đồng thời có biến thể trên gen *JAG1*, do đó Sangkhathat giả thiết rằng bệnh lý ứ mật ở bệnh nhân teo đường mật bẩm sinh là kết quả tác động tổng hợp của biến thể trên cả hai gen *ABCC2* và *JAG1*.

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. Cơ sở dữ liệu mạng tương tác protein - protien STRING dự đoán protein mã hóa bởi gen *ABCC2* tương tác nhiều nhất với protein mã hóa bởi các gen *ABCG2* (score: 0,979), *SLC22A8* (score: 0,913), *UGT1A6* (score: 0,844), *SLCO1B1* (score: 0,842) và *SLCO1B3* (score: 0,817) (Hình 4.2). Trong đó *ABCG2* mã hóa protein liên kết màng thuộc siêu họ vận chuyển cassette liên kết ATP (ABC), đóng vai trò quan trọng trong việc vận chuyển các phân tử khác nhau qua màng ngoài và màng tế bào. Tuy nhiên, những tương tác giữa MRP2 được mã hóa bở *ABCC2* với các protein khác được STRING dự đoán khá mờ nhạt, thể hiện qua việc có ít đường gạch nối giữa protein này với protein khác, chủ yếu là tương tác ở dạng đồng biểu hiện – co-expression và). Như vậy, những thay đổi nhỏ trên protein này có thể ít ảnh hưởng đến sự hoạt động của các protein khác.



Hình 4.2: Tương tác giữa protein mã hóa bởi ABCC2 với các protein khác

Các biến đổi trên gen *MYO5B* được tìm thấy trên bệnh nhân ứ mật trong gan tiến triển gia đình (Progressive familial intrahepatic cholestasis-PFIC) [153]. Ngoài ra, các biến đổi 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 [59]. 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 [154]. Nghiên cứu trên một bệnh nhân nữ mắc PFIC với biểu hiện vàng da và ứ mật dai dẳng, Khanal và đồng tác giả đã xác định được hai biến thể dị hợp tử phức: 1 biến thể xóa ở exon 19 dẫn đến dịch khung và cắt ngắn sớm 5 axit amin và một biến thể sai nghĩa ở exon 20 được coi là nguyên nhân gây bệnh ở bệnh nhân này [155]. Gen *MYO5B* ảnh hưởng đến sự hình thành và phát triển của các vi lông mao. Bằng cách so sánh với người khỏe mạnh và bệnh nhân mắc bệnh gan ứ mật khác, Chu và các đồng tác giả nhận thấy rằng lông mao nguyên phát ở bệnh nhân teo đường mật bẩm sinh ngắn hơn, ít đa dạng hơn và có sự phát triển bất thường trong gan của bệnh nhân [156].

Cho đến nay, cơ sở dữ liệu đột biến gen trên người HGMD ghi nhận 88 biến thể trên gen *MYO5B,* chủ yếu là các biến thể liên quan đến bệnh vi nhung mao thể vùi (61 biến thể), 21 biến thể được ghi nhận là liên quan đến tình trạng ứ mật và chỉ có 2 biến thể được ghi nhận là liên quan đến bệnh teo đường mật bẩm sinh ngoài gan.

Bảng 4.1: Các biến thể trong gen MYO5B được phát hiện ở bệnh nhân teo đường mật bẩm sinh

SNP ID	HGVS.c	HGVS.p	MAF	Tài liệu tham khảo
rs756813138	c.4922G>A (HET)	p.Arg1641His	0.00	[152]
Không có	c.4407G>A (HET)	p.Met1469Ile	Không có	[152]
rs757769301	c.3254G > A (HET)	p.Arg1085Gln	<0.0001	[152]
rs117920737	c.197A > C (HET)	p.Asp66Ala	0.001	[45,152]

MAF: Tần số alen (được lấy theo cơ sở dư liệu dbSNP)

HET: dị hợp tử

Bằng cách sử dụng công nghệ NGS, Laochareonsuk và đồng tác giả phát hiện 3 đột biến trên gen *MYO5B* ở các bệnh nhân mắc bệnh teo đường mật bẩm sinh [152]. Sangkhathat và đồng tác giả phát hiện hai đột biến với tần số alen thấp (MAF<0,001) là rs757769301 (c.3254G > A, p.Arg1085Gln) và rs117920737 (c.197A > C, p.Asp66Ala) trên hai bệnh nhân mắc teo đường mật bẩm sinh (Bảng 4.1). Các bệnh nhân mắc teo đường mật bẩm sinh được phát hiện bởi Sangkhathat và đồng tác giả đều mắc đồng thời 2 bệnh là PFIC và teo đường mật bẩm sinh. Do đó nhóm tác giả giả thiết rằng teo đường mật bẩm sinh có thể là biểu hiện bệnh lý chung của nhiều hội chứng liên quan đến tình trạng ứ mật, tăng bilirubin ở trẻ em. 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

Nghiên cứu này đã thu thập mẫu máu của 266 bệnh nhân mắc teo đường mật bẩm sinh và 250 mẫu đối chứng. Bằng cách giải trình tự toàn bộ vùng gen mã hóa của 05 bệnh nhân điển hình (04 bệnh nhân thuộc hai gia đình, 01 bệnh nhân mắc đồng thời hai bệnh là teo đường mật bẩm sinh và Wilson), nghiên cứu này đã phát hiện được các đột biến có tiềm năng gây bệnh teo đường mật bẩm sinh, cụ thể như sau:

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 có khả năng liên quan đến bệnh teo đường mật bẩm sinh.

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, 02 đột biến c.314C>A (p.Ser105∗), c.2975C>T (p.Pro992Leu) trên gen *ATP7B* đã được báo cáo là có liên quan đến bệnh nhân Wilson.

3. Đã phân tích 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, alen C có liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh, 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* không có mối liên quan đến nguy cơ mắc bệnh teo đường mật bẩm sinh ở người Việt Nam.

- 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

Teo đường mật bẩm sinh là căn bệnh đa yếu tố với nhiều nguyên nhân gây bệnh khác nhau. Cơ sở di truyền của bệnh phức tạp với sự tham gia của nhiều gen, nhiều con đường gây bệnh. Hơn nữa, tác động của yếu tố di truyền đến cơ chế bệnh sinh của bệnh còn nhiều điểm chưa được làm rõ. Mặc dù nghiên cứu này đã phát hiện được một số điểm đột biến có tiềm năng gây bệnh tuy nhiên cần 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 CÔNG TRÌNH CÔNG BỐ 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

TÀI LIỆU THAM KHẢO

1. Nomden M., van Wessel D.B.E., Ioannou S., Verkade H.J., de Kleine R. H., Alizadeh B Z, Bruggink J.L.M., Hulscher J.B.F., 2021, A Higher Incidence of Isolated Biliary Atresia in Rural Areas: Results From an Epidemiological Study in The Netherlands, *J Pediatr Gastroenterol Nutr*, 72, pp.202–2099.
2. Hsiao C.H., Chang M.H., Chen H.L., Lee H.C., Wu T.C., Lin C.C., Yang Y.J., Chen A.C., Tiao M.M., Lau B.H., Chu C.H., Lai M.W., Taiwan Infant Stool Color Card Study Group., 2008, Universal screening for biliary atresia using an infant stool color card in Taiwan, *Hepatology*, 47, pp. 1233–1240.
3. Livesey E., Cortina B.M., Sharif K., Alizai N., McClean P., Kelly D., Hadzic N., Davenport M., 2009, Epidemiology of biliary atresia in England and Wales (1999-2006), *Arch Dis Child Fetal Neonatal Ed*, 94, pp. F451-455.
4. Hopkins P.C., Yazigi N., Nylund C.M., 2017, Incidence of Biliary Atresia and Timing of Hepatoportoenterostomy in the United States, *J Pediatr*, 187, pp. 253–257.
5. Chardot C., 2006, Biliary atresia, *Orphanet Journal of Rare Diseases*, 1, 28.
6. Zhan J., Feng J., Chen Y., Liu J., Wang B., 2017, Incidence of biliary atresia associated congenital malformations: A retrospective multicenter study in China, *Asian J Surg*, 40, pp. 429–433.
7. Girard M., Jannot A.S., Besnard M., Jacquemin E., Henrion-Caude A., 2012, Biliary atresia: does ethnicity matter? *J Hepatol*, 57, pp700–1; author reply 702.
8. Muraji T., Tanaka H., Ieiri S., 2018, Ethnic variation in the incidence of biliary atresia correlates with the frequency of the most prevalent haplotype in its population, *Hum Immunol*, 79, pp. 668–671.
9. Xu X., Zhan J., 2020, Biliary atresia in twins: a systematic review and meta-analysis, *Pediatr Surg Int*, 36, pp. 953–958.

10. Nio M., Ohi R., Miyano T., Saeki M., Shiraki K., Tanaka K.., and Japanese Biliary Atresia Registry 2003 Five- and 10-year survival rates after surgery for biliary atresia: a report from the Japanese Biliary Atresia Registry *J Pediatr Surg* 38 997–1000

11. Bezerra J A., Wells R G., Mack C L., Karpen S J., Hoofnagle J., Doo E and Sokol R J 2018 BILIARY ATRESIA: Clinical and Research Challenges for the 21st Century *Hepatology* 68 1163–73

12. Mack C L and Sokol R J 2005 Unraveling the Pathogenesis and Etiology of Biliary Atresia *Pediatr Res* 57 87–94

13. Kilgore A and Mack C L 2017 Update on investigations pertaining to the pathogenesis of biliary atresia *Pediatr Surg Int* 33 1233–41

14. Malik A., Thanekar U., Mourya R and Shivakumar P 2020 Recent developments in etiology and disease modeling of biliary atresia: a narrative review *Digestive Medicine Research* 3

15. Scottoni F and Davenport M 2020 Biliary atresia: Potential for a new decade *Semin Pediatr Surg* 29 150940

16. Shaul E., Kogan-Liberman D., Schuckalo S., Jan D., Ewart M., Nguyen T., Martinez M and Ovchinsky N 2019 Novel mutations in NOTCH2 gene in infants with neonatal cholestasis *Pediatr Rep* 11 8206

17. Koh H., Park G-S., Shin S-M., Park C E., Kim S., Han S J., Pham H Q., Shin J-H and Lee D-W 2018 Mitochondrial Mutations in Cholestatic Liver Disease with Biliary Atresia *Sci Rep* 8 905

18. Fabre A., Roman C and Roquelaure B 2017 Somatic mutation., a cause of biliary atresia: A hypothesis *Med Hypotheses* 102 91–3

19. Song W., Sun L-Y., Zhu Z-J., Wei L., Qu W., Zeng Z-G., Liu Y., Zhang H-M and Guo W 2021 Association of Gut Microbiota and Metabolites With Disease Progression in Children With Biliary Atresia *Front Immunol* 12 698900

20. Moore S W., Zabiegaj-Zwick C and Nel E 2012 Problems related to CMV infection and biliary atresia *S Afr Med J* 102 890–2

21. Lakshminarayanan B and Davenport M 2016 Biliary atresia: A comprehensive review *J Autoimmun* 73 1–9

22. Tarr P I., Haas J E and Christie D L 1996 Biliary atresia., cytomegalovirus., and age at referral *Pediatrics* 97 828–31

23. Oliveira N L G de., Kanawaty F R., Costa S C B and Hessel G 2002 Infection by cytomegalovirus in patients with neonatal cholestasis *Arq Gastroenterol* 39 132–6

24. Tyler K L., Sokol R J., Oberhaus S M., Le M., Karrer F M., Narkewicz M R., Tyson R W., Murphy J R., Low R and Brown W R 1998 Detection of reovirus RNA in hepatobiliary tissues from patients with extrahepatic biliary atresia and choledochal cysts *Hepatology* 27 1475–82

25. Rauschenfels S., Krassmann M., Al-Masri A N., Verhagen W., Leonhardt J., Kuebler J F and Petersen C 2009 Incidence of hepatotropic viruses in biliary atresia *Eur J Pediatr* 168 469–76

26. Riepenhoff-Talty M., Gouvea V., Evans M J., Svensson L., Hoffenberg E., Sokol R J., Uhnoo I., Greenberg S J., Schäkel K., Zhaori G., Fitzgerald J., Chong S., el-Yousef M., Nemeth A., Brown M., Piccoli D., Hyams J., Ruffin D and Rossi T 1996 Detection of group C rotavirus in infants with extrahepatic biliary atresia *J Infect Dis* 174 8–15

27. Bobo L., Ojeh C., Chiu D., Machado A., Colombani P and Schwarz K 1997 Lack of Evidence for Rotavirus by Polymerase Chain Reaction/Enzyme Immunoassay of Hepatobiliary Samples from Children with Biliary Atresia *Pediatr Res* 41 229–34

28. Brown W R., Sokol R J., Levin M J., Silverman A., Tamaru T., Lilly J R., Hall R J and Cheney M 1988 Lack of correlation between infection with reovirus 3 and extrahepatic biliary atresia or neonatal hepatitis *J Pediatr* 113 670–6

29. Shah I and Bhatnagar S 2012 Biliary atresia and cytomegalovirus and response to valganciclovir *Indian Pediatr* 49 484–6

30. Broomé U., Nemeth A., Hultcrantz R and Scheynius A 1997 Different expression of HLA-DR and ICAM-1 in livers from patients with biliary atresia and Byler's disease *J Hepatol* 26 857–62

31. Kobayashi H., Horikoshi K., Long L., Yamataka A., Lane G J and Miyano T 2001 Serum concentration of adhesion molecules in postoperative biliary atresia patients: relationship to disease activity and cirrhosis *J Pediatr Surg* 36 1297–301

32. Sokol R J and Mack C 2001 Etiopathogenesis of biliary atresia *Semin Liver Dis* 21 517–24

33. Bezerra J A., Tiao G., Ryckman F C., Alonso M., Sabla G E., Shneider B., Sokol R J and Aronow B J 2002 Genetic induction of proinflammatory immunity in children with biliary atresia *Lancet* 360 1653–9

34. Ahmed A F., Ohtani H., Nio M., Funaki N., Shimaoka S., Nagura H and Ohi R 2001 CD8+ T cells infiltrating into bile ducts in biliary atresia do not appear to function as cytotoxic T cells: a clinicopathological analysis *J Pathol* 193 383–9

35. Mack C L., Tucker R M., Sokol R J., Karrer F M., Kotzin B L., Whitington P F and Miller S D 2004 Biliary atresia is associated with CD4+ Th1 cell-mediated portal tract inflammation *Pediatr Res* 56 79–87

36. Ohya T., Fujimoto T., Shimomura H and Miyano T 1995 Degeneration of intrahepatic bile duct with lymphocyte infiltration into biliary epithelial cells in biliary atresia *J Pediatr Surg* 30 515–8

37. Mack C L 2007 The pathogenesis of biliary atresia: evidence for a virus-induced autoimmune disease *Semin Liver Dis* 27 233–42

38. Silveira T R., Salzano F M., Donaldson P T., Mieli-Vergani G., Howard E R and Mowat A P 1993 Association between HLA and extrahepatic biliary atresia *J Pediatr Gastroenterol Nutr* 16 114–7

39. Yuasa T., Tsuji H., Kimura S., Niwa N., Yurugi K., Egawa H., Tanaka K., Maruya E., Saji H-O., Asano H and Maekawa T 2005 Human leukocyte antigens in Japanese patients with biliary atresia: retrospective

analysis of patients who underwent living donor liver transplantation *Hum Immunol* 66 295–300

40. Hernandez H M., Kovarik P., Whitington P F and Alonso E M 2001 Autoimmune hepatitis as a late complication of liver transplantation *J Pediatr Gastroenterol Nutr* 32 131–6

41. Coffinier C., Gresh L., Fiette L., Tronche F., Schütz G., Babinet C., Pontoglio M., Yaniv M and Barra J 2002 Bile system morphogenesis defects and liver dysfunction upon targeted deletion of HNF1beta *Development* 129 1829–38

42. Clotman F., Lannoy V J., Reber M., Cereghini S., Cassiman D., Jacquemin P., Roskams T., Rousseau G G and Lemaigre F P 2002 The onecut transcription factor HNF6 is required for normal development of the biliary tract *Development* 129 1819–28

43. Balistreri W F., Bezerra J A., Jansen P., Karpen S J., Shneider B L and Suchy F J 2005 Intrahepatic cholestasis: summary of an American Association for the Study of Liver Diseases single-topic conference *Hepatology* 42 222–35

44. Zhang D-Y., Sabla G., Shivakumar P., Tiao G., Sokol R J., Mack C., Shneider B L., Aronow B and Bezerra J A 2004 Coordinate expression of regulatory genes differentiates embryonic and perinatal forms of biliary atresia *Hepatology* 39 954–62

45. Sangkhathat S., Laochareonsuk W., Maneechay W., Kayasut K and Chiengkriwate P 2018 Variants Associated with Infantile Cholestatic Syndromes Detected in Extrahepatic Biliary Atresia by Whole Exome Studies: A 20-Case Series from Thailand *J Pediatr Genet* 7 67–73

46. Cheng G., Tang C S-M., Wong E H-M., Cheng W W-C., So M-T., Miao X., Zhang R., Cui L., Liu X., Ngan E S-W., Lui V C-H., Chung P H-Y., Chan I H-Y., Liu J., Zhong W., Xia H., Yu J., Qiu X., Wu X-Z., Wang B., Dong X., Tou J., Huang L., Yi B., Ren H., Chan E K-W., Ye K., O'Reilly P F., Wong K K-Y., Sham P-C., Cherny S S., Tam P K-H and

Garcia-Barceló M-M 2013 Common genetic variants regulating ADD3 gene expression alter biliary atresia risk *J Hepatol* 59 1285–91

47. Cui S., Leyva-Vega M., Tsai E A., EauClaire S F., Glessner J T., Hakonarson H., Devoto M., Haber B A., Spinner N B and Matthews R P 2013 Evidence from human and zebrafish that GPC1 is a biliary atresia susceptibility gene *Gastroenterology* 144 1107-1115.e3

48. Tsai E A., Grochowski C M., Loomes K M., Bessho K., Hakonarson H., Bezerra J A., Russo P A., Haber B A., Spinner N B and Devoto M 2014 Replication of a GWAS signal in a Caucasian population implicates ADD3 in susceptibility to biliary atresia *Hum Genet* 133 235–43

49. Zeng S., Sun P., Chen Z., Mao J., Wang J., Wang B and Liu L 2014 Association between Single Nucleotide Polymorphisms in the ADD3 Gene and Susceptibility to Biliary Atresia *PLOS ONE* 9 e107977

50. Ningappa M., So J., Glessner J., Ashokkumar C., Ranganathan S., Min J., Higgs B W., Sun Q., Haberman K., Schmitt L., Vilarinho S., Mistry P K., Vockley G., Dhawan A., Gittes G K., Hakonarson H., Jaffe R., Subramaniam S., Shin D and Sindhi R., 2015, The Role of ARF6 in Biliary Atresia, *PLoS One*, 10, pp. e0138381.

51. Schön P., Tsuchiya K., Lenoir D., Mochizuki T., Guichard C., Takai S., Maiti A K., Nihei H., Weil J., Yokoyama T., Bouvagnet P., 2002, Identification genomic organization., chromosomal mapping and mutation analysis of the human INV gene the ortholog of a murine gene implicated in left-right axis development and biliary atresia, *Hum Genet*, 110, pp.157–165.

52. Yuasa T., Tsuji H., Kimura S., Niwa N., Yurugi K., Egawa H., Tanaka K., Maruya E., Saji H.O., Asano H., Maekawa T., 2005, Human leukocyte antigens in Japanese patients with biliary atresia: retrospective analysis of patients who underwent living donor liver transplantation, *Hum Immunol*, 66, pp. 295–300.

53. Donaldson P T., Clare M., Constantini P K., Hadzic N., Mieli-Vergani G., Howard E and Kelley D 2002 HLA and cytokine gene polymorphisms in biliary atresia, *Liver,* 22, pp.213–219.

54. Tsai E A., Grochowski C.M., Falsey A.M., Rajagopalan R., Wendel D., Devoto M., Krantz I.D., Loomes K.M., Spinner N.B., 2015, Heterozygous Deletion of FOXA2 Segregates with Disease in a Family with Heterotaxy, Panhypopituitarism, and Biliary Atresia, *Hum Mutat,* 36, pp. 631–637.

55. Asai A., Miethke A., Bezerra J.A., 2015, Pathogenesis of biliary atresia: defining biology to understand clinical phenotypes, *Nat Rev Gastroenterol Hepatol,* 12, pp. 342–352.

56. Wu L.N., Zhu Z.J., Sun L.Y., 2022, Genetic Factors and Their Role in the Pathogenesis of Biliary Atresia, *Front Pediatr,* 10, pp. 912154.

57. Luo Y., Brigham D., Bednarek J., Torres R., Wang D., Ahmad S., Mack C.L., 2021, Unique cholangiocyte-targeted IgM autoantibodies correlate with poor outcome in biliary atresia, *Hepatology*, 73, pp. 1855–1867.

58. Holterman A., Nguyen H P A., Nadler E., Vu G H., Mohan P., Vu M., Trinh T T., Bui H T T., Nguyen B T., Quynh A T and Pham H D 2021 Granulocyte-colony stimulating factor GCSF mobilizes hematopoietic stem cells in Kasai patients with biliary atresia in a phase 1 study and improves short term outcome *J Pediatr Surg* 56 1179–85

59. Vij M and Rela M 2020 Biliary atresia: pathology., etiology and pathogenesis *Future Science OA* 6 FSO466

60. Omenetti A and Diehl A M 2011 HEDGEHOG SIGNALING IN CHOLANGIOCYTES *Curr Opin Gastroenterol* 27 268–75

61. Frassetto R., Parolini F., Marceddu S., Satta G., Papacciuoli V., Pinna M A., Mela A., Secchi G., Galleri G., Manetti R., Bercich L., Villanacci V., Dessanti A., Antonucci R., Tanda F., Alberti D., Schwarz K B and Clemente M G 2018 Intrahepatic bile duct primary cilia in biliary atresia *Hepatol Res* 48 664–74

62. Jung H Y., Jing J., Lee K B and Jang J-J 2015 Sonic hedgehog (SHH) and glioblastoma-2 (Gli-2) expressions are associated with poor jaundice-free survival in biliary atresia *J Pediatr Surg* 50 371–6

63. Garcia-Barceló M-M., Yeung M-Y., Miao X-P., Tang C S-M., Cheng G., Chen G., So M-T., Ngan E S-W., Lui V C-H., Chen Y., Liu X-L., Hui K-J W S., Li L., Guo W-H., Sun X-B., Tou J-F., Chan K-W., Wu X-Z., Song Y-Q., Chan D., Cheung K., Chung P H-Y., Wong K K-Y., Sham P-C., Cherny S S and Tam P K-H 2010 Genome-wide association study identifies a susceptibility locus for biliary atresia on 10q24.2 *Hum Mol Genet* 19 2917–25

64. Bai M-R., Niu W-B., Zhou Y., Gong Y-M., Lu Y-J., Yu X-X., Wei Z-L., Wu W., Song H-L., Yu W-W., Gu B-L., Cai W and Chu X 2020 Association of common variation in ADD3 and GPC1 with biliary atresia susceptibility *Aging (Albany NY)* 12 7163–82

65. Cui S., Leyva-Vega M., Tsai E A., Eauclaire S F., Glessner J T., Hakonarson H., Devoto M., Haber B A., Spinner N B and Matthews R P 2013 Evidence From Human and Zebrafish That GPC1 Is a Biliary Atresia Susceptibility Gene *Gastroenterology* 144 1107-1115.e3

66. Leyva-Vega M., Gerfen J., Thiel B.D., Jurkiewicz D., Rand E.B., Pawlowska J., Kaminska D., Russo P., Gai X., Krantz I.D., Kamath B.M., Hakonarson H., Haber B.A., Spinner N.B., 2010, Genomic alterations in biliary atresia suggest region of potential disease susceptibility in 2q37.3, *Am J Med Genet A*, 152A, pp. 886–895.

67. Ke J., Zeng S., Mao J., Wang J., Lou J., Li J., Chen X., Liu C., Huang L-M., Wang B., Liu L., 2016, Common genetic variants of GPC1 gene reduce risk of biliary atresia in a Chinese population, *J Pediatr Surg*, 51, pp. 1661–1664.

68. Ningappa M., Min J., Higgs B.W., Ashokkumar C., Ranganathan S., Sindhi R., 2015, Genome-wide association studies in biliary atresia, *Wiley Interdiscip Rev Syst Biol Med*, 7, pp. 267–273.

69. Shan Y., Shen N., Han L., Chen Q., Zhang J., Long X., Xia Q., 2016, MicroRNA-499 Rs3746444 polymorphism and biliary atresia, *Dig Liver Dis*, 48, pp. 423–428.

70. Gawish E., El-Monem E A., El-Abd M., Sobhy G.A., Ghanem H., 2020, MicroRNA-499 rs3746444 polymorphism in Egyptian children with biliary atresia, *Clin Exp Hepatol*, 6, 263–269.

71. Zhao R., Song Z., Dong R., Li H., Shen C., Zheng S, 2013, Polymorphism of ITGB2 gene 3'-UTR+145C/A is associated with biliary atresia, *Digestion,* 88, pp. 65–71.

72. Arikan C., Berdeli A., Kilic M., Tumgor G., Yagci R.V., Aydogdu S., 2008, Polymorphisms of the ICAM-1 gene are associated with biliary atresia, *Dig Dis Sci,* 53, pp. 2000–2004.

73. Arikan C., Berdeli A., Ozgenc F., Tumgor G., Yagci R V and Aydogdu S 2006 Positive association of macrophage migration inhibitory factor gene-173G/C polymorphism with biliary atresia *J Pediatr Gastroenterol Nutr* 42 77–82

74. Sadek K H., Ezzat S., Abdel-Aziz S A., Alaraby H., Mosbeh A and Abdel-Rahman M H 2017 Macrophage Migration Inhibitory Factor (MIF) Gene Promotor Polymorphism Is Associated with Increased Fibrosis in Biliary Atresia Patients., but Not with Disease Susceptibility *Ann Hum Genet* 81 177–83

75. Lee H-C., Chang T-Y., Yeung C-Y., Chan W-T., Jiang C-B., Chen W-F., Chan H-W., Liu H-F., Lin M and Lee Y-J 2010 Genetic variation in the vascular endothelial growth factor gene is associated with biliary atresia *J Clin Gastroenterol* 44 135–9

76. Liu B., Wei J., Li M., Jiang J., Zhang H., Yang L., Wu H and Zhou Q 2017 Association of common genetic variants in VEGFA with biliary atresia susceptibility in Northwestern Han Chinese *Gene* 628 87–92

77. Liu F., Zeng J., Zhu D., Xu X., Lan M., Wang M., Zhao J., Xia H., Zhang Y and Zhang R 2020 PDGFA gene rs9690350 polymorphism increases biliary atresia risk in Chinese children *Biosci Rep* 40 BSR20200068

78. Chen Y., Gilbert M A., Grochowski C M., McEldrew D., Llewellyn J., Waisbourd-Zinman O., Hakonarson H., Bailey-Wilson J E., Russo P., Wells R G., Loomes K M., Spinner N B and Devoto M 2018 A genome-wide association study identifies a susceptibility locus for biliary atresia on 2p16.1 within the gene EFEMP1 *PLoS Genet* 14 e1007532

79. Lin Z., Xie X., Lin H., Fu M., Su L., Tong Y., Chen H., Wang H., Zhao J., Xia H., Zhang Y and Zhang R 2018 Epistatic Association of CD14 and NOTCH2 Genetic Polymorphisms with Biliary Atresia in a Southern Chinese Population *Mol Ther Nucleic Acids* 13 590–5

80. Udomsinprasert W., Tencomnao T., Honsawek S., Anomasiri W., Vejchapipat P., Chongsrisawat V and Poovorawan Y 2012 +276 G/T single nucleotide polymorphism of the adiponectin gene is associated with the susceptibility to biliary atresia *World J Pediatr* 8 328–34

81. Liang J., Wen Z., Zhao J., Liang Q., Liu T., Xia H., Zhang Y and Zhang R 2018 Association of IL18 genetic polymorphisms with increased risk of Biliary atresia susceptibility in Southern Chinese children *Gene* 677 228–31

82. Berauer J-P., Mezina A I., Okou D T., Sabo A., Muzny D M., Gibbs R A., Hegde M R., Chopra P., Cutler D J., Perlmutter D H., Bull L N., Thompson R J., Loomes K M., Spinner N B., Rajagopalan R., Guthery S L., Moore B., Yandell M., Harpavat S., Magee J C., Kamath B M., Molleston J P., Bezerra J A., Murray K F., Alonso E M., Rosenthal P., Squires R H., Wang K S., Finegold M J., Russo P., Sherker A H., Sokol R J and Karpen S J 2019 Identification of PKD1L1 Gene Variants in Children with the Biliary Atresia Splenic Malformation Syndrome *Hepatology* 70 899–910

83. Lam W-Y., Tang C S-M., So M-T., Yue H., Hsu J S., Chung P H-Y., Nicholls J M., Yeung F., Lee C-W D., Ngo D N., Nguyen P A H., Mitchison H M., Jenkins D., O'Callaghan C., Garcia-Barceló M-M., Lee S-L., Sham P-C., Lui V C-H and Tam P K-H 2021 Identification of a wide spectrum of ciliary gene mutations in nonsyndromic biliary atresia

patients implicates ciliary dysfunction as a novel disease mechanism *EBioMedicine* 71 103530

84. Zöllner J., Finer S., Linton K J., Team G and H R., Heel D A van., Williamson C and Dixon P H 2023 Rare variant contribution to cholestatic liver disease in a South Asian population in the United Kingdom *Scientific Reports* 13

85. Hu G., He P., Liu Z., Chen Q., Zheng B and Zhang Q 2014 Diagnosis of ABCB11 gene mutations in children with intrahepatic cholestasis using high resolution melting analysis and direct sequencing *Mol Med Rep* 10 1264–74

86. Morais M B and Machado M V 2022 Benign inheritable disorders of bilirubin metabolism manifested by conjugated hyperbilirubinemia–A narrative review *United European Gastroenterol J* 10 745–53

87. Anon 2020 UniProt: the universal protein knowledgebase in 2021 *Nucleic Acids Res* 49 D480–9

88. Tang V., Cofer Z C., Cui S., Sapp V., Loomes K M and Matthews R P 2016 Loss of a Candidate Biliary Atresia Susceptibility Gene., add3a., Causes Biliary Developmental Defects in Zebrafish *J Pediatr Gastroenterol Nutr* 63 524–30

89. Bangs F and Anderson K V 2017 Primary Cilia and Mammalian Hedgehog Signaling *Cold Spring Harb Perspect Biol* 9 a028175

90. Pan J and Ho M 2021 Role of glypican-1 in regulating multiple cellular signaling pathways *Am J Physiol Cell Physiol* 321 C846–58

91. Grisaru S., Cano-Gauci D., Tee J., Filmus J and Rosenblum N D 2001 Glypican-3 modulates BMP- and FGF-mediated effects during renal branching morphogenesis *Dev Biol* 231 31–46

92. Yan D and Lin X 2007 Drosophila glypican Dally-like acts in FGF-receiving cells to modulate FGF signaling during tracheal morphogenesis *Dev Biol* 312 203–16

93. Jassal B., Matthews L., Viteri G., Gong C., Lorente P., Fabregat A., Sidiropoulos K., Cook J., Gillespie M., Haw R., Loney F., May B., Milacic M., Rothfels K., Sevilla C., Shamovsky V., Shorser S., Varusai T., Weiser J., Wu G., Stein L., Hermjakob H and D'Eustachio P 2020 The reactome pathway knowledgebase *Nucleic Acids Res* 48 D498–503

94. Yang L., Dong C., Yang J., Yang L., Chang N., Qi C and Li L 2019 MicroRNA-26b-5p Inhibits Mouse Liver Fibrogenesis and Angiogenesis by Targeting PDGF Receptor-Beta *Mol Ther Nucleic Acids* 16 206–17

95. So J., Ningappa M., Glessner J., Min J., Ashokkumar C., Ranganathan S., Higgs B W., Li D., Sun Q., Schmitt L., Biery A C., Dobrowolski S., Trautz C., Fuhrman L., Schwartz M C., Klena N T., Fusco J., Prasadan K., Adenuga M., Mohamed N., Yan Q., Chen W., Horne W., Dhawan A., Sharif K., Kelly D., Squires R H., Gittes G K., Hakonarson H., Morell V., Lo C., Subramaniam S., Shin D and Sindhi R 2020 Biliary-Atresia-Associated Mannosidase-1-Alpha-2 Gene Regulates Biliary and Ciliary Morphogenesis and Laterality *Front Physiol* 11 538701

96. Tsunoda T., Kakinuma S., Miyoshi M., Kamiya A., Kaneko S., Sato A., Tsuchiya J., Nitta S., Kawai-Kitahata F., Murakawa M., Itsui Y., Nakagawa M., Azuma S., Sogo T., Komatsu H., Mukouchi R., Inui A., Fujisawa T., Nakauchi H., Asahina Y and Watanabe M 2019 Loss of fibrocystin promotes interleukin-8-dependent proliferation and CTGF production of biliary epithelium *J Hepatol* 71 143–52

97. Kohsaka T., Yuan Z-R., Guo S-X., Tagawa M., Nakamura A., Nakano M., Kawasasaki H., Inomata Y., Tanaka K and Miyauchi J 2002 The significance of human jagged 1 mutations detected in severe cases of extrahepatic biliary atresia *Hepatology* 36 904–12

98. Wang J-Z., Fu X., Fang Z., Liu H., Zong F-Y., Zhu H., Yu Y-F., Zhang X-Y., Wang S-F., Huang Y and Hui J 2021 QKI-5 regulates the alternative splicing of cytoskeletal gene ADD3 in lung cancer *J Mol Cell Biol* 13 347–60

99. Li T-F., Ke X-Y., Zhang Y-R and Zhan J-H 2023 The correlation between rs2501577 gene polymorphism and biliary atresia: a systematic review and meta-analysis *Pediatr Surg Int* 39 206

100. Lee H-C., Chang T-Y., Yeung C-Y., Chan W-T., Jiang C-B., Chen W-F., Chan H-W., Liu H-F., Lin M and Lee Y-J 2010 Association of Interferon-Gamma Gene Polymorphisms in Taiwanese Children with Biliary Atresia *J Clin Immunol* 30 68–73

101. Kaewkiattiyot S., Honsawek S., Vejchapipat P., Chongsrisawat V and Poovorawan Y 2011 Association of X-prolyl aminopeptidase 1 rs17095355 polymorphism with biliary atresia in Thai children *Hepatology Research* 41 1249–52

102. Wildhaber B E 2012 Biliary Atresia: 50 Years after the First Kasai *ISRN Surgery* 2012 e132089

103. Nguyen H P A., Ren J., Butler M., Li H., Qazi S., Sadiq K., Dao H T and Holterman A 2022 Study protocol of Phase 2 open-label multicenter randomized controlled trial for granulocyte-colony stimulating factor (GCSF) in post-Kasai Type 3 biliary atresia *Pediatr Surg Int* 38 1019–30

104. Liem N T., Pham H D., Dung L A., Son T N and Vu H M 2012 Early and intermediate outcomes of laparoscopic surgery for choledochal cysts with 400 patients *J Laparoendosc Adv Surg Tech A* 22 599–603

105. Liu M B., Huong T B., Hoang X., Doan L., Trinh S., Anh Nguyen H P., Thanh Le H and Holterman A-X 2017 Biliary atresia in Vietnam: Management and the burden of disease *Surgery* 161 533–7

106. Tran K T., Le V S., Dao L T M., Nguyen H K., Mai A K., Nguyen H T., Ngo M D., Tran Q A and Nguyen L T 2021 Novel findings from family-based exome sequencing for children with biliary atresia *Sci Rep* 11 21815

107. Girard M and Panasyuk G 2019 Genetics in biliary atresia *Curr Opin Gastroenterol* 35 73–81

108. Besheer T., Arafa M., El-Maksoud M A., Elalfy H., Hasson A., Zalata K., Elkashef W., Elshahawy H., Raafat D., Elemshaty W., Elsayed E.,

Zaghloul H., Razek A A and El-Bendary M 2018 Diagnosis of cirrhosis in patients with chronic hepatitis C genotype 4: Role of ABCB11 genotype polymorphism and plasma bile acid levels *Turk J Gastroenterol* 29 299–307

109. Marine R., Polson S W., Ravel J., Hatfull G., Russell D., Sullivan M., Syed F., Dumas M and Wommack K E 2011 Evaluation of a Transposase Protocol for Rapid Generation of Shotgun High-Throughput Sequencing Libraries from Nanogram Quantities of DNA *Applied and Environmental Microbiology* 77 8071–9

110. Cock P J A., Fields C J., Goto N., Heuer M L and Rice P M 2010 The Sanger FASTQ file format for sequences with quality scores., and the Solexa/Illumina FASTQ variants *Nucleic Acids Res.* 38 1767–71

111. Li H and Durbin R 2009 Fast and accurate short read alignment with Burrows-Wheeler transform *Bioinformatics* 25 1754–60

112. Zhang J., Lin H., Balaji P and Feng W 2013 Optimizing Burrows-Wheeler Transform-Based Sequence Alignment on Multicore Architectures *2013 13th IEEE/ACM International Symposium on Cluster., Cloud., and Grid Computing* pp 377–84

113. McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M and DePristo M A 2010 The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data *Genome Res* 20 1297–303

114. Yi M., Zhao Y., Jia L., He M., Kebebew E and Stephens R M 2014 Performance comparison of SNP detection tools with illumina exome sequencing data—an assessment using both family pedigree information and sample-matched SNP array data *Nucleic Acids Res* 42 e101–e101

115. Liu X., Han S., Wang Z., Gelernter J and Yang B-Z 2013 Variant Callers for Next-Generation Sequencing Data: A Comparison Study *PLOS ONE* 8 e75619

116. Wang K., Li M and Hakonarson H 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data *Nucleic Acids Research* 38 e164

117. Cingolani P., Platts A., Wang L L., Coon M., Nguyen T., Wang L., Land S J., Lu X and Ruden D M 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms., SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3 *Fly (Austin)* 6 80–92

118. Ng P C and Henikoff S 2003 SIFT: Predicting amino acid changes that affect protein function *Nucleic Acids Res* 31 3812–4

119. Adzhubei I., Jordan D M and Sunyaev S R 2013 Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2 *Curr Protoc Hum Genet* 0 7 Unit7.20

120. Choi Y and Chan A P 2015 PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels *Bioinformatics* 31 2745–7

121. Schwarz J M., Cooper D N., Schuelke M and Seelow D 2014 MutationTaster2: mutation prediction for the deep-sequencing age *Nature Methods* 11 361–2

122. Chen Q., Lu P., Jones A V., Cross N C P., Silver R T and Wang Y L 2007 Amplification Refractory Mutation System., a Highly Sensitive and Simple Polymerase Chain Reaction Assay., for the Detection of JAK2 V617F Mutation in Chronic Myeloproliferative Disorders *J Mol Diagn* 9 272–6

123. Ye S., Dhillon S., Ke X., Collins A R and Day I N M 2001 An efficient procedure for genotyping single nucleotide polymorphisms *Nucleic Acids Res* 29 e88

124. Collins A and Ke X 2012 Primer1: Primer Design Web Service for Tetra-Primer ARMS-PCR *The Open Bioinformatics Journal* 6

125. Nicastro E and D'Antiga L 2018 Next generation sequencing in pediatric hepatology and liver transplantation *Liver Transplantation* 24 282–93

126. Feldman A G and Sokol R J 2019 Neonatal cholestasis: emerging molecular diagnostics and potential novel therapeutics *Nat Rev Gastroenterol Hepatol* 16 346–60

127. Goldberg A and Mack C L 2020 Inherited Cholestatic Diseases in the Era of Personalized Medicine *Clin Liver Dis (Hoboken)* 15 105–9

128. Jeyaraj R., Bounford K M., Ruth N., Lloyd C., MacDonald F., Hendriksz C J., Baumann U., Gissen P and Kelly D 2021 The Genetics of Inherited Cholestatic Disorders in Neonates and Infants: Evolving Challenges *Genes (Basel)* 12 1837

129. Liu L-Y., Wang X-H., Lu Y., Zhu Q-R and Wang J-S 2013 Association of variants of ABCB11 with transient neonatal cholestasis *Pediatr Int* 55 138–44

130. Jacquemin E., Malan V., Rio M., Davit-Spraul A., Cohen J., Landrieu P and Bernard O 2010 Heterozygous FIC1 deficiency: a new genetic predisposition to transient neonatal cholestasis *J Pediatr Gastroenterol Nutr* 50 447–9

131. Hijikata A., Tsuji T., Shionyu M and Shirai T 2017 Decoding disease-causing mechanisms of missense mutations from supramolecular structures *Sci Rep* 7 8541

132. Mezina A and Karpen S J 2015 Genetic contributors and modifiers of biliary atresia *Dig Dis* 33 408–14

133. Verkade H J., Bezerra J A., Davenport M., Schreiber R A., Mieli-Vergani G., Hulscher J B., Sokol R J., Kelly D A., Ure B., Whitington P F., Samyn M and Petersen C 2016 Biliary atresia and other cholestatic childhood diseases: Advances and future challenges *J Hepatol* 65 631–42

134. Zhang B-P., Huang Z-H and Dong C 2019 Biliary atresia combined with progressive familial intrahepatic cholestasis type 3: A case report and review of the literature *Medicine (Baltimore)* 98 e15593

135. Mezina A., Gandhi K., Sabo A., Muzny D., Gibbs R., Hegde M and Karpen S J 2014 845 Whole Exome Sequencing Identifies ABCB4 Gene

Variants As Modifiers of Biliary Atresia Outcomes *Gastroenterology* 146 S-928

136. Fischetto R., Palmieri V V., Tripaldi M E., Gaeta A., Michelucci A., Delvecchio M., Francavilla R and Giordano P 2019 Alagille Syndrome: A Novel Mutation in JAG1 Gene *Front Pediatr* 7 199

137. Rajagopalan R., Tsai E A., Grochowski C M., Kelly S M., Loomes K M., Spinner N B and Devoto M 2020 Exome Sequencing in Individuals with Isolated Biliary Atresia *Sci Rep* 10 2709

138. Gürünlüoğlu K., Koç A., Durmuş K., Gözükara Bağ H., Ceran C., Gürünlüoğlu S., Yıldız T., Gül M and Demircan M 2020 Whole exome sequencing analysis for mutations in isolated type III biliary atresia patients *Clin Exp Hepatol* 6 347–53

139. Szklarczyk D., Gable A L., Nastou K C., Lyon D., Kirsch R., Pyysalo S., Doncheva N T., Legeay M., Fang T., Bork P., Jensen L J and von Mering C 2021 The STRING database in 2021: customizable protein-protein networks., and functional characterization of user-uploaded gene/measurement sets *Nucleic Acids Res* 49 D605–12

140. Szklarczyk D., Gable A L., Lyon D., Junge A., Wyder S., Huerta-Cepas J., Simonovic M., Doncheva N T., Morris J H., Bork P., Jensen L J and Mering C von 2019 STRING v11: protein-protein association networks with increased coverage., supporting functional discovery in genome-wide experimental datasets *Nucleic Acids Res* 47 D607–13

141. Stenson P D., Mort M., Ball E V., Evans K., Hayden M., Heywood S., Hussain M., Phillips A D and Cooper D N 2017 The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research., genetic diagnosis and next-generation sequencing studies *Hum Genet* 136 665–77

142. Faridounnia M., Folkers G E and Boelens R 2018 Function and Interactions of ERCC1-XPF in DNA Damage Response *Molecules* 23 3205

143. Mori T., Yousefzadeh M J., Faridounnia M., Chong J X., Hisama F M., Hudgins L., Mercado G., Wade E A., Barghouthy A S., Lee L., Martin G M., Nickerson D A., Bamshad M J., Niedernhofer L J and Oshima J 2018 ERCC4 Variants Identified in a Cohort of Patients with Segmental Progeroid Syndromes *Hum Mutat* 39 255–65

144. Sergi C M and Gilmour S 2022 Biliary Atresia: A Complex Hepatobiliary Disease with Variable Gene Involvement., Diagnostic Procedures., and Prognosis *Diagnostics* 12 330

145. Wilkins B J., Gong W and Pack M 2014 A novel keratin18 promoter that drives reporter gene expression in the intrahepatic and extrahepatic biliary system allows isolation of cell-type specific transcripts from zebrafish liver *Gene Expr Patterns* 14 62–8

146. Babu R O., Lui V C H., Chen Y., Yiu R S W., Ye Y., Niu B., Wu Z., Zhang R., Yu M O N., Chung P H Y., Wong K K Y., Xia H., Zhang M Q., Wang B., Lendahl U and Tam P K H 2020 Beta-amyloid deposition around hepatic bile ducts is a novel pathobiological and diagnostic feature of biliary atresia *Journal of Hepatology* 73 1391–403

147. Yang K., Battista C., Woodhead J., Stahl S., Mettetal J., Watkins P., Siler S and Howell B 2017 Systems pharmacology modeling of drug-induced hyperbilirubinemia: Differentiating hepatotoxicity and inhibition of enzymes/transporters *Clinical Pharmacology & Therapeutics* 101 501–9

148. Huang Y-S., Chang T-E., Perng C-L and Huang Y-H 2021 The association of transporter ABCC2 (MRP2) genetic variation and drug-induced hyperbilirubinemia *Journal of the Chinese Medical Association* 84 129

149. Sookoian S., Castaño G., Burgueño A., Gianotti T F and Pirola C J 2008 Association of the multidrug-resistance-associated protein gene (ABCC2) variants with intrahepatic cholestasis of pregnancy *J. Hepatol.* 48 125–32

150. Kim K Y., Kim T H., Seong M-W., Park S S., Moon J S and Ko J S 2020 Mutation spectrum and biochemical features in infants with neonatal Dubin-Johnson syndrome *BMC Pediatr* 20

151. van der Schoor L W E., Verkade H J., Kuipers F and Jonker J W 2015 New insights in the biology of ABC transporters ABCC2 and ABCC3: impact on drug disposition *Expert Opin Drug Metab Toxicol* 11 273–93

152. Laochareonsuk W., Surachat K., Chiengkriwate P and Sangkhathat S 2022 A novel pathogenesis concept of biliary atresia approached by combined molecular strategies *PLoS One* 17 e0277334

153. Gonzales E., Taylor S A., Davit-Spraul A., Thébaut A., Thomassin N., Guettier C., Whitington P F and Jacquemin E 2017 MYO5B mutations cause cholestasis with normal serum gamma-glutamyl transferase activity in children without microvillous inclusion disease *Hepatology* 65 164–73

154. Aldrian D., Vogel G.F., Frey T.K., Ayyıldız Civan H., Aksu A.Ü., Avitzur Y., Ramos B.E., Çakır M., Demir A.M., Deppisch C., Duba H.C., Düker G., Gerner P., Hertecant J., Hornová J., Kathemann S., Koeglmeier J., Koutroumpa A., Lanzersdorfer R., Lev-Tzion R., Lima R., Mansour S., Meissl M., Melek J., Miqdady M., Montoya J.H., Posovszky C., Rachman Y., Siahanidou T., Tabbers M., Uhlig H.H., Ünal S., Wirth S., Ruemmele F.M., Hess M.W., Huber L.A., Müller T., Sturm E., Janecke A.R., 2021, Congenital Diarrhea and Cholestatic Liver Disease: Phenotypic Spectrum Associated with MYO5B Mutations, *J Clin Med*, 10, pp. 481.

155. Khanal M., Jha A.K., Sharma A.K., 2022, Compound Heterozygous MYO5B Mutation, a Cause of Infantile Cholestasis: A Case Report, *JNMA J Nepal Med Assoc*, 60, pp. 821–824.

156. Chu A.S., Russo P.A., Wells R.G., 2012, Cholangiocyte cilia are abnormal in syndromic and non-syndromic biliary atresia, *Mod Pathol*, 25, pp. 751–757.

157. Dixon P.H., van Mil.S.W.C., Chambers J., Strautnieks S., Thompson R.J., Lammert F., Kubitz R., Keitel V., Glantz A., Mattsson L.A., Marschall

H.U., Molokhia M., Moore G.E., Linton K.J., Williamson C., 2009, Contribution of variant alleles of ABCB11 to susceptibility to intrahepatic cholestasis of pregnancy, *Gut,* 58, pp. 537–544.

158. Kagawa T., Hirose S., Arase Y., Oka A., Anzai K., Tsuruya K., Shiraishi K., Orii R., Ieda S., Nakazawa T., Tomita K., Hokari R., Miura S., Ebinuma H., Saito H., Kitamura T., Horie Y., Okuse C., Wasada M., Inoko H., Tohkin M., Saito Y., Maekawa K., Takikawa H., Mine T., 2015, No contribution of the ABCB11 p.444A polymorphism in Japanese patients with drug-induced cholestasis, *Drug Metab Dispos,* 43, pp. 691–697.

159. Ali I., Khalid S., Stieger B., Brouwer K.L.R., 2019, Effect of a Common Genetic Variant (p.V444A) in the Bile Salt Export Pump on the Inhibition of Bile Acid Transport by Cholestatic Medications, *Mol Pharm*, 16, pp. 1406–1411

160. Lei J.H., Yang X., Xiao X.Q., Chen Z., Peng F., 2017, A preliminary investigation on single nucleotide polymorphism rs2287622 of bile salt export pump gene in patients with chronic hepatitis C virus infection in Hunan, China, *BMC Gastroenterol,* 17, pp. 42.

PHỤ LỤC

Phụ lục 1: Danh sách 133 gen liên quan đến bệnh teo đường mật bẩm sinh được sử dụng để sàng lọc đột biến

NST	Gen
NST 1	*ABCD3; ACADM; AGL; CPT2; EPHX1; GALE; GBA; HMGCL; KCNH1; NOTCH2; PEX14; PEX19; PGM1; PKLR*
NST 2	*ABCB11; ABCG5; ABCG8; BCS1L; CCDC115; CFC1; CPS1; CYP27A1; DGUOK; EIF2AK3; GPC1; HADHA; MPV17; NBAS; PEX13; STAT4; SUCLG1; UGT1A1; VIL1*
NST 3	*CLDN1; GBE1; GFM1; NGLY1; SLC25A20; STT3B; APM1*
NST 4	*CC2D2A; ETFDH; PKD2; SLC39A8*
NST 5	*AMACR; CD14; LARS; TTC37*
NST 6	*DCDC2; FARS2; PEX3; PEX6; PEX7; PKHD1; RFX6; VEGF*
NST 7	*ABCB4; AKR1D1; ASL; CFTR; DLD; PDGFA; PEX1; SBDS*
NST 8	*CYP7A1; CYP7B1; PEX2; TMEM67*
NST 9	*AGPAT2; ALAD; ALDOB; ASS1; BAAT; GALT; IARS; INVS; TJP2*
NST 10	*ABCC2; ADD3; ADK; C10orf2; LIPA; MRPS16; XPNPEP1*
NST 11	*BSCL2; CPT1A; PEX16; SMPD1*

b

NST 12	*GYS2; KRT18; KRT8; MLL2; NR1H4; PEX5; TSFM*
NST 13	*ATP7B; COG6; SLC10A2*
NST 14	*ARF6; NPC2; SERPINA1; VIPAS39*
NST 15	*ETFA; FAH; MPI; POLG; VPS33B*
NST 16	*CIRH1A; COG7; ERCC4; HSD3B7; PMM2; RPGRIP1L*
NST 17	*ACADVL; COG1; PEX12; PTRF; SCO1*
NST 18	*ATP8B1; MYO5B; NPC1*
NST 19	*ETFB; ICAM1*
NST 20	*JAG1*
NST 21	*ITGB2*
NST 22	*MIF; PEX26; TRMU; TYMP*
NST X	*AMER1; OCRL; OTC; ZIC3*

Phụ lục 2: Danh sách các đa hình nucleotide trên 05 bệnh nhân

Bảng 1: Đa hình nucleotide đơn bệnh nhân BA01

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
1	chr1	Sai nghĩa	*CPT2*	c.1055T>G	p.Phe352Cys	0.046526
2	chr1	Sai nghĩa	*CPT2*	c.1102G>A	p.Val368Ile	0.413139
3	chr1	Sai nghĩa	*PGM1*	c.1312T>C	p.Tyr438His	0.225839
4	chr1	Sai nghĩa	*EPHX1*	c.337T>C	p.Tyr113His	0.313299
5	chr2	Sai nghĩa	*NBAS*	c.3026G>C	p.Cys1009Ser	0.036741
6	chr2	Sai nghĩa	*NBAS*	c.2845G>C	p.Val949Leu	0.026158
7	chr2	Sai nghĩa	*NBAS*	c.130C>G	p.Gln44Glu	0.01278
8	chr2	Sai nghĩa	*ABCG8*	c.161A>G	p.Tyr54Cys	0.434704
9	chr2	Sai nghĩa	*ABCG8*	c.1895T>C	p.Val632Ala	0.922923
10	chr2	Sai nghĩa	*EIF2AK3*	c.2110G>T	p.Ala704Ser	0.720647
11	chr2	Sai nghĩa	*EIF2AK3*	c.497A>G	p.Gln166Arg	0.648962
12	chr2	Sai nghĩa	*EIF2AK3*	c.407C>G	p.Ser136Cys	0.280551
13	chr2	Sai nghĩa	*ABCB11*	c.1331T>C	p.Val444Ala	0.588658
14	chr3	Sai nghĩa	*GBE1*	c.1000A>G	p.Ile334Val	0.983626
15	chr3	Sai nghĩa	*GBE1*	c.568A>G	p.Arg190Gly	0.321486
16	chr4	Sai nghĩa	*CC2D2A*	c.1127A>C	p.Glu376Ala	0.117612
17	chr4	Sai nghĩa	*ETFDH*	c.92C>T	p.Thr31Ile	0.672724
18	chr5	Sai nghĩa	*AMACR*	c.829G>A	p.Glu277Lys	0.696086
19	chr5	Sai nghĩa	*AMACR*	c.717G>T	p.Gln239His	0.114417
20	chr5	Sai nghĩa	*AMACR*	c.602T>C	p.Leu201Ser	0.714856
21	chr6	Sai nghĩa	*FARS2*	c.839A>G	p.Asn280Ser	0.235224
22	chr6	Sai nghĩa	*DCDC2*	c.661A>G	p.Ser221Gly	0.516973
23	chr6	Sai nghĩa	*PKHD1*	c.12143A>G	p.Gln4048Arg	0.565695
24	chr6	Sai nghĩa	*PKHD1*	c.11696A>G	p.Gln3899Arg	0.553315
25	chr6	Sai nghĩa	*PKHD1*	c.5608T>G	p.Leu1870Val	0.955871
26	chr6	Sai nghĩa	*PKHD1*	c.3785C>T	p.Ala1262Val	0.354633
27	chr6	Sai nghĩa	*PKHD1*	c.2278C>T	p.Arg760Cys	0.355831
28	chr7	Sai nghĩa	*CFTR*	c.1408G>A	p.Val470Met	0.582069
29	chr8	Sai nghĩa	*PEX2*	c.550T>C	p.Cys184Arg	0.996206
30	chr8	Sai nghĩa	*TMEM67*	c.1810A>G	p.Ile604Val	0.692891
31	chr9	Sai nghĩa	*TJP2*	c.1539C>A	p.Asp513Glu	0.772165
32	chr9	Sai nghĩa	*BAAT*	c.59G>A	p.Arg20Gln	0.562899
33	chr10	Sai nghĩa	*LIPA*	c.46A>C	p.Thr16Pro	0.286142
34	chr10	Sai nghĩa	*ABCC2*	c.116A>T	p.Tyr39Phe	0.994609
35	chr11	Sai nghĩa	*SMPD1*	c.107T>C	p.Val36Ala	0.561302
36	chr11	Sai nghĩa	*PEX16*	c.346G>A	p.Val116Ile	1
37	chr13	Sai nghĩa	*COG6*	c.28G>A	p.Ala10Thr	0.484425
38	chr13	Sai nghĩa	*COG6*	c.94T>A	p.Cys32Ser	0.484225
39	chr13	Sai nghĩa	*ATP7B*	c.1216T>G	p.Ser406Ala	0.376198

d

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
40	chr13	Sai nghĩa	*SLC10A2*	c.511T>G	p.Ser171Ala	0.840256
41	chr14	Sai nghĩa	*SERPINA1*	c.1200A>C	p.Glu400Asp	0.280351
42	chr14	Sai nghĩa	*SERPINA1*	c.374G>A	p.Arg125His	0.160543
43	chr15	Sai nghĩa	*VPS33B*	c.1540G>A	p.Gly514Ser	0.767772
44	chr16	Sai nghĩa	*HSD3B7*	c.748A>G	p.Thr250Ala	0.493411
45	chr16	Sai nghĩa	*RPGRIP1L*	c.3790G>A	p.Asp1264Asn	0.099042
46	chr16	Sai nghĩa	*RPGRIP1L*	c.3073G>A	p.Gly1025Ser	0.166134
47	chr17	Sai nghĩa	*COG1*	c.1175A>G	p.Asn392Ser	0.520966
48	chr18	Sai nghĩa	*NPC1*	c.2572A>G	p.Ile858Val	0.485823
49	chr18	Sai nghĩa	*NPC1*	c.1926G>C	p.Met642Ile	0.831669
50	chr18	Sai nghĩa	*MYO5B*	c.3962G>A	p.Gly1321Glu	0.144369
51	chr18	Sai nghĩa	*MYO5B*	c.376A>G	p.Thr126Ala	0.973442
52	chr18	Sai nghĩa	*ATP8B1*	c.3454G>A	p.Ala1152Thr	0.9998
53	chr19	Sai nghĩa	*ICAM1*	c.167A>T	p.Lys56Met	0.084465
54	chr19	Sai nghĩa	*ICAM1*	c.1405A>G	p.Lys469Glu	0.358826
55	chr19	Sai nghĩa	*ETFB*	c.734C>T	p.Thr245Met	0.419728
56	chr21	Sai nghĩa	*ITGB2*	c.1062A>T	p.Gln354His	1

e

Bảng 2: Đa hình nucleotide đơn bệnh nhân BA02

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
1	chr1	Sai nghĩa	*CPT2*	c.1102G>A	p.Val368Ile	0.4131
2	chr1	Sai nghĩa	*CPT2*	c.1939A>G	p.Met647Val	0.1004
3	chr1	Sai nghĩa	*PGM1*	c.1312T>C	p.Tyr438His	0.2258
4	chr1	Sai nghĩa	*AGL*	c.3199C>T	p.Pro1067Ser	0.1106
5	chr1	Sai nghĩa	*EPHX1*	c.337T>C	p.Tyr113His	0.3133
6	chr2	Sai nghĩa	*NBAS*	c.3026G>C	p.Cys1009Ser	0.0367
7	chr2	Sai nghĩa	*NBAS*	c.2845G>C	p.Val949Leu	0.0262
8	chr2	Sai nghĩa	*NBAS*	c.130C>G	p.Gln44Glu	0.0128
9	chr2	Sai nghĩa	*ABCG8*	c.161A>G	p.Tyr54Cys	0.4347
10	chr2	Sai nghĩa	*ABCG8*	c.1895T>C	p.Val632Ala	0.9229
11	chr2	Sai nghĩa	*EIF2AK3*	c.2110G>T	p.Ala704Ser	0.7206
12	chr2	Sai nghĩa	*EIF2AK3*	c.497A>G	p.Gln166Arg	0.649
13	chr2	Sai nghĩa	*EIF2AK3*	c.407C>G	p.Ser136Cys	0.2806
14	chr2	Sai nghĩa	*ABCB11*	c.1331T>C	p.Val444Ala	0.5887
15	chr3	Sai nghĩa	*GBE1*	c.1000A>G	p.Ile334Val	0.9836
16	chr3	Sai nghĩa	*GBE1*	c.568A>G	p.Arg190Gly	0.3215
17	chr4	Sai nghĩa	*CC2D2A*	c.1127A>C	p.Glu376Ala	0.1176
18	chr4	Sai nghĩa	*PKD2*	c.568G>A	p.Ala190Thr	0.1028
19	chr4	Sai nghĩa	*ETFDH*	c.92C>T	p.Thr31Ile	0.6727
20	chr5	Sai nghĩa	*AMACR*	c.829G>A	p.Glu277Lys	0.6961
21	chr5	Sai nghĩa	*AMACR*	c.602T>C	p.Leu201Ser	0.7149
22	chr5	Sai nghĩa	*AMACR*	c.524G>A	p.Gly175Asp	0.2526
23	chr5	Sai nghĩa	*AMACR*	c.25G>A	p.Val9Met	0.2869
24	chr6	Sai nghĩa	*FARS2*	c.839A>G	p.Asn280Ser	0.2352
25	chr6	Sai nghĩa	*DCDC2*	c.661A>G	p.Ser221Gly	0.517
26	chr6	Sai nghĩa	*PKHD1*	c.12143A>G	p.Gln4048Arg	0.5657
27	chr6	Sai nghĩa	*PKHD1*	c.11696A>G	p.Gln3899Arg	0.5533
28	chr6	Sai nghĩa	*PKHD1*	c.5608T>G	p.Leu1870Val	0.9559
29	chr6	Sai nghĩa	*PKHD1*	c.3785C>T	p.Ala1262Val	0.3546
30	chr6	Sai nghĩa	*PKHD1*	c.2489A>G	p.Asn830Ser	0.0635
31	chr6	Sai nghĩa	*PKHD1*	c.2278C>T	p.Arg760Cys	0.3558
32	chr7	Sai nghĩa	*CFTR*	c.1408G>A	p.Val470Met	0.5821
33	chr8	Sai nghĩa	*PEX2*	c.550T>C	p.Cys184Arg	0.9962
34	chr8	Sai nghĩa	*TMEM67*	c.1810A>G	p.Ile604Val	0.6929
35	chr9	Sai nghĩa	*TJP2*	c.1539C>A	p.Asp513Glu	0.7722

f

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
36	chr9	Sai nghĩa	*BAAT*	c.59G>A	p.Arg20Gln	0.5629
37	chr10	Sai nghĩa	*LIPA*	c.46A>C	p.Thr16Pro	0.2861
38	chr10	Sai nghĩa	*ABCC2*	c.116A>T	p.Tyr39Phe	0.9946
39	chr10	Sai nghĩa	*ABCC2*	c.1249G>A	p.Val417Ile	0.1865
40	chr11	Sai nghĩa	*SMPD1*	c.107T>C	p.Val36Ala	0.5613
41	chr11	Sai nghĩa	*PEX16*	c.346G>A	p.Val116Ile	1
42	chr12	Sai nghĩa	*GYS2*	c.1087A>G	p.Met363Val	0.7406
43	chr13	Sai nghĩa	*COG6*	c.28G>A	p.Ala10Thr	0.4844
44	chr13	Sai nghĩa	*COG6*	c.94T>A	p.Cys32Ser	0.4842
45	chr13	Sai nghĩa	*ATP7B*	c.3419T>C	p.Val1140Ala	0.5405
46	chr13	Sai nghĩa	*ATP7B*	c.2855G>A	p.Arg952Lys	0.531
47	chr13	Sai nghĩa	*ATP7B*	c.2495A>G	p.Lys832Arg	0.5024
48	chr13	Sai nghĩa	*ATP7B*	c.1366G>C	p.Val456Leu	0.377
49	chr13	Sai nghĩa	*ATP7B*	c.1216T>G	p.Ser406Ala	0.3762
50	chr13	Sai nghĩa	*SLC10A2*	c.511T>G	p.Ser171Ala	0.8403
51	chr15	Sai nghĩa	*VPS33B*	c.1540G>A	p.Gly514Ser	0.7678
52	chr16	Sai nghĩa	*HSD3B7*	c.748A>G	p.Thr250Ala	0.4934
53	chr16	Sai nghĩa	*RPGRIP1L*	c.3790G>A	p.Asp1264Asn	0.099
54	chr16	Sai nghĩa	*RPGRIP1L*	c.3073G>A	p.Gly1025Ser	0.1661
55	chr18	Sai nghĩa	*NPC1*	c.3797G>A	p.Arg1266Gln	0.1502
56	chr18	Sai nghĩa	*NPC1*	c.2572A>G	p.Ile858Val	0.4858
57	chr18	Sai nghĩa	*NPC1*	c.1926G>C	p.Met642Ile	0.8317
58	chr18	Sai nghĩa	*MYO5B*	c.3962G>A	p.Gly1321Glu	0.1444
59	chr18	Sai nghĩa	*MYO5B*	c.376A>G	p.Thr126Ala	0.9734
60	chr18	Sai nghĩa	*ATP8B1*	c.3454G>A	p.Ala1152Thr	0.9998
61	chr19	Sai nghĩa	*ETFB*	c.734C>T	p.Thr245Met	0.4197
62	chr21	Sai nghĩa	*ITGB2*	c.1062A>T	p.Gln354His	1

g

Bảng 3: Đa hình nucleotide đơn bệnh nhân BA03

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
1	chr1	Sai nghĩa	*CPT2*	c.1102G>A	p.Val368Ile	0.4131
2	chr1	Sai nghĩa	*CPT2*	c.1939A>G	p.Met647Val	0.1004
3	chr1	Sai nghĩa	*PGM1*	c.715C>T	p.Arg239Cys	0.2414
4	chr1	Sai nghĩa	*AGL*	c.1160G>A	p.Arg387Gln	0.0797
5	chr1	Sai nghĩa	*EPHX1*	c.128G>C	p.Arg43Thr	0.0102
6	chr1	Sai nghĩa	*EPHX1*	c.337T>C	p.Tyr113His	0.3133
7	chr2	Sai nghĩa	*NBAS*	c.3011G>A	p.Arg1004Gln	0.0677
8	chr2	Sai nghĩa	*ABCG5*	c.1810C>G	p.Gln604Glu	0.24
9	chr2	Sai nghĩa	*ABCG8*	c.161A>G	p.Tyr54Cys	0.4347
10	chr2	Sai nghĩa	*ABCG8*	c.1895T>C	p.Val632Ala	0.9229
11	chr2	Sai nghĩa	*EIF2AK3*	c.2110G>T	p.Ala704Ser	0.7206
12	chr2	Sai nghĩa	*EIF2AK3*	c.407C>G	p.Ser136Cys	0.2806
13	chr2	Sai nghĩa	*ABCB11*	c.1331T>C	p.Val444Ala	0.5887
14	chr2	Sai nghĩa	*UGT1A1*	c.211G>A	p.Gly71Arg	0.0343
15	chr2	Sai nghĩa	*GPC1*	c.1498A>G	p.Ser500Gly	0.6643
16	chr3	Sai nghĩa	*GBE1*	c.1000A>G	p.Ile334Val	0.9836
17	chr3	Sai nghĩa	*GBE1*	c.568A>G	p.Arg190Gly	0.3215
18	chr4	Sai nghĩa	*PKD2*	c.568G>A	p.Ala190Thr	0.1028
19	chr4	Sai nghĩa	*ETFDH*	c.92C>T	p.Thr31Ile	0.6727
20	chr5	Sai nghĩa	*AMACR*	c.829G>A	p.Glu277Lys	0.6961
21	chr5	Sai nghĩa	*AMACR*	c.717G>T	p.Gln239His	0.1144
22	chr5	Sai nghĩa	*AMACR*	c.602T>C	p.Leu201Ser	0.7149
23	chr5	Sai nghĩa	*AMACR*	c.524G>A	p.Gly175Asp	0.2526
24	chr5	Sai nghĩa	*AMACR*	c.25G>A	p.Val9Met	0.2869
25	chr6	Sai nghĩa	*FARS2*	c.839A>G	p.Asn280Ser	0.2352
26	chr6	Sai nghĩa	*DCDC2*	c.661A>G	p.Ser221Gly	0.517
27	chr6	Sai nghĩa	*PEX6*	c.2816C>A	p.Pro939Gln	0.3315
28	chr6	Sai nghĩa	*PEX6*	c.2426C>T	p.Ala809Val	0.0751
29	chr6	Sai nghĩa	*PKHD1*	c.12143A>G	p.Gln4048Arg	0.5657
30	chr6	Sai nghĩa	*PKHD1*	c.11696A>G	p.Gln3899Arg	0.5533
31	chr6	Sai nghĩa	*PKHD1*	c.5608T>G	p.Leu1870Val	0.9559
32	chr7	Sai nghĩa	*CFTR*	c.1408G>A	p.Val470Met	0.5821
33	chr8	Sai nghĩa	*PEX2*	c.550T>C	p.Cys184Arg	0.9962
34	chr8	Sai nghĩa	*TMEM67*	c.1810A>G	p.Ile604Val	0.6929
35	chr9	Sai nghĩa	*TJP2*	c.1539C>A	p.Asp513Glu	0.7722

h

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
36	chr9	Sai nghĩa	*BAAT*	c.59G>A	p.Arg20Gln	0.5629
37	chr10	Sai nghĩa	*LIPA*	c.46A>C	p.Thr16Pro	0.2861
38	chr10	Sai nghĩa	*ABCC2*	c.116A>T	p.Tyr39Phe	0.9946
39	chr11	Sai nghĩa	*SMPD1*	c.107T>C	p.Val36Ala	0.5613
40	chr11	Sai nghĩa	*PEX16*	c.346G>A	p.Val116Ile	1
41	chr12	Sai nghĩa	*GYS2*	c.1087A>G	p.Met363Val	0.7406
42	chr13	Sai nghĩa	*COG6*	c.28G>A	p.Ala10Thr	0.4844
43	chr13	Sai nghĩa	*COG6*	c.94T>A	p.Cys32Ser	0.4842
44	chr13	Sai nghĩa	*SLC10A2*	c.511T>G	p.Ser171Ala	0.8403
45	chr15	Sai nghĩa	*VPS33B*	c.1540G>A	p.Gly514Ser	0.7678
46	chr16	Sai nghĩa	*HSD3B7*	c.748A>G	p.Thr250Ala	0.4934
47	chr17	Sai nghĩa	*COG1*	c.1175A>G	p.Asn392Ser	0.521
48	chr18	Sai nghĩa	*NPC1*	c.1926G>C	p.Met642Ile	0.8317
49	chr18	Sai nghĩa	*MYO5B*	c.376A>G	p.Thr126Ala	0.9734
50	chr18	Sai nghĩa	*ATP8B1*	c.3454G>A	p.Ala1152Thr	0.9998
51	chr21	Sai nghĩa	*ITGB2*	c.1062A>T	p.Gln354His	1
52	chr22	Sai nghĩa	*TYMP*	c.1427C>T	p.Ser476Leu	0.1428

Bảng 4: Đa hình nucleotide đơn bệnh nhân BA04

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
1	chr1	Sai nghĩa	*CPT2*	c.1102G>A	p.Val368Ile	0.413139
2	chr1	Sai nghĩa	*CPT2*	c.1939A>G	p.Met647Val	0.100439
3	chr1	Sai nghĩa	*PGM1*	c.715C>T	p.Arg239Cys	0.241414
4	chr1	Sai nghĩa	*EPHX1*	c.416A>G	p.His139Arg	0.215455
5	chr2	Sai nghĩa	*NBAS*	c.3011G>A	p.Arg1004Gln	0.0676917
6	chr2	Sai nghĩa	*NBAS*	c.1964A>G	p.Lys655Arg	0.360623
7	chr2	Sai nghĩa	*NBAS*	c.727A>G	p.Ile243Val	0.389177
8	chr2	Sai nghĩa	*ABCG8*	c.161A>G	p.Tyr54Cys	0.434704
9	chr2	Sai nghĩa	*ABCG8*	c.1895T>C	p.Val632Ala	0.922923
10	chr2	Sai nghĩa	*EIF2AK3*	c.2110G>T	p.Ala704Ser	0.720647
11	chr2	Sai nghĩa	*EIF2AK3*	c.407C>G	p.Ser136Cys	0.280551
12	chr2	Sai nghĩa	*ABCB11*	c.1331T>C	p.Val444Ala	0.588658
13	chr2	Sai nghĩa	*UGT1A1*	c.211G>A	p.Gly71Arg	0.034345
14	chr2	Sai nghĩa	*GPC1*	c.1498A>G	p.Ser500Gly	0.664337
15	chr3	Sai nghĩa	*GBE1*	c.1000A>G	p.Ile334Val	0.983626
16	chr3	Sai nghĩa	*GBE1*	c.568A>G	p.Arg190Gly	0.321486
17	chr4	Sai nghĩa	*PKD2*	c.568G>A	p.Ala190Thr	0.102835
18	chr4	Sai nghĩa	*ETFDH*	c.92C>T	p.Thr31Ile	0.672724
19	chr5	Sai nghĩa	*AMACR*	c.829G>A	p.Glu277Lys	0.696086
20	chr5	Sai nghĩa	*AMACR*	c.717G>T	p.Gln239His	0.114417
21	chr5	Sai nghĩa	*AMACR*	c.602T>C	p.Leu201Ser	0.714856
22	chr6	Sai nghĩa	*FARS2*	c.839A>G	p.Asn280Ser	0.235224
23	chr6	Sai nghĩa	*DCDC2*	c.661A>G	p.Ser221Gly	0.516973
24	chr6	Sai nghĩa	*PEX6*	c.2426C>T	p.Ala809Val	0.0750799
25	chr6	Sai nghĩa	*PKHD1*	c.12143A>G	p.Gln4048Arg	0.565695
26	chr6	Sai nghĩa	*PKHD1*	c.11696A>G	p.Gln3899Arg	0.553315
27	chr6	Sai nghĩa	*PKHD1*	c.5608T>G	p.Leu1870Val	0.955871
28	chr7	Sai nghĩa	*CFTR*	c.1408G>A	p.Val470Met	0.582069
29	chr8	Sai nghĩa	*PEX2*	c.550T>C	p.Cys184Arg	0.996206
30	chr8	Sai nghĩa	*TMEM67*	c.1810A>G	p.Ile604Val	0.692891
31	chr9	Sai nghĩa	*TJP2*	c.1539C>A	p.Asp513Glu	0.772165
32	chr9	Sai nghĩa	*BAAT*	c.59G>A	p.Arg20Gln	0.562899
33	chr10	Sai nghĩa	*LIPA*	c.46A>C	p.Thr16Pro	0.286142
34	chr10	Sai nghĩa	*ABCC2*	c.116A>T	p.Tyr39Phe	0.994609
35	chr11	Sai nghĩa	*SMPD1*	c.107T>C	p.Val36Ala	0.561302
36	chr11	Sai nghĩa	*PEX16*	c.346G>A	p.Val116Ile	1
37	chr12	Sai nghĩa	*GYS2*	c.1087A>G	p.Met363Val	0.740615
38	chr13	Sai nghĩa	*COG6*	c.28G>A	p.Ala10Thr	0.484425
39	chr13	Sai nghĩa	*COG6*	c.94T>A	p.Cys32Ser	0.484225
40	chr13	Sai nghĩa	*ATP7B*	c.3419T>C	p.Val1140Ala	0.540535
41	chr13	Sai nghĩa	*ATP7B*	c.2855G>A	p.Arg952Lys	0.53095
42	chr13	Sai nghĩa	*ATP7B*	c.2495A>G	p.Lys832Arg	0.502396
43	chr14	Sai nghĩa	*SERPINA1*	c.1200A>C	p.Glu400Asp	0.280351
44	chr14	Sai nghĩa	*SERPINA1*	c.374G>A	p.Arg125His	0.160543

j

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
45	chr15	Sai nghĩa	*VPS33B*	c.1540G>A	p.Gly514Ser	0.767772
46	chr16	Sai nghĩa	*HSD3B7*	c.748A>G	p.Thr250Ala	0.493411
47	chr16	Sai nghĩa	*RPGRIP1L*	c.3790G>A	p.Asp1264Asn	0.0990415
48	chr16	Sai nghĩa	*RPGRIP1L*	c.3073G>A	p.Gly1025Ser	0.166134
49	chr18	Sai nghĩa	*NPC1*	c.1926G>C	p.Met642Ile	0.831669
50	chr18	Sai nghĩa	*MYO5B*	c.3962G>A	p.Gly1321Glu	0.144369
51	chr18	Sai nghĩa	*MYO5B*	c.376A>G	p.Thr126Ala	0.973442
52	chr18	Sai nghĩa	*ATP8B1*	c.3454G>A	p.Ala1152Thr	0.9998
53	chr21	Sai nghĩa	*ITGB2*	c.1062A>T	p.Gln354His	1

k

Bảng 5: Đa hình nucleotide đơn bệnh nhân BA05

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
1	chr1	Sai nghĩa	*CPT2*	c.1102G>A	p.Val368Ile	0.413139
2	chr1	Sai nghĩa	*PGM1*	c.715C>T	p.Arg239Cys	0.241414
3	chr1	Sai nghĩa	*PGM1*	c.1312T>C	p.Tyr438His	0.225839
4	chr1	Sai nghĩa	*AGL*	c.686A>G	p.Gln229Arg	0.0483227
5	chr1	Sai nghĩa	*EPHX1*	c.337T>C	p.Tyr113His	0.313299
6	chr2	Sai nghĩa	*NBAS*	c.3011G>A	p.Arg1004Gln	0.0676917
7	chr2	Sai nghĩa	*ABCG8*	c.161A>G	p.Tyr54Cys	0.434704
8	chr2	Sai nghĩa	*ABCG8*	c.1895T>C	p.Val632Ala	0.922923
9	chr2	Sai nghĩa	*EIF2AK3*	c.2110G>T	p.Ala704Ser	0.720647
10	chr2	Sai nghĩa	*EIF2AK3*	c.497A>G	p.Gln166Arg	0.648962
11	chr2	Sai nghĩa	*ABCB11*	c.1331T>C	p.Val444Ala	0.588658
12	chr2	Sai nghĩa	*GPC1*	c.1498A>G	p.Ser500Gly	0.664337
13	chr3	Sai nghĩa	*GBE1*	c.1000A>G	p.Ile334Val	0.983626
14	chr3	Sai nghĩa	*GBE1*	c.568A>G	p.Arg190Gly	0.321486
15	chr3	Sai nghĩa	*GFM1*	c.643G>A	p.Val215Ile	0.583267
16	chr4	Sai nghĩa	*CC2D2A*	c.1127A>C	p.Glu376Ala	0.117612
17	chr4	Sai nghĩa	*ETFDH*	c.92C>T	p.Thr31Ile	0.672724
18	chr5	Sai nghĩa	*AMACR*	c.829G>A	p.Glu277Lys	0.696086
19	chr5	Sai nghĩa	*AMACR*	c.602T>C	p.Leu201Ser	0.714856
20	chr5	Sai nghĩa	*AMACR*	c.524G>A	p.Gly175Asp	0.252596
21	chr5	Sai nghĩa	*AMACR*	c.25G>A	p.Val9Met	0.286941
22	chr6	Sai nghĩa	*FARS2*	c.839A>G	p.Asn280Ser	0.235224
23	chr6	Sai nghĩa	*DCDC2*	c.661A>G	p.Ser221Gly	0.516973
24	chr6	Sai nghĩa	*PEX6*	c.2426C>T	p.Ala809Val	0.0750799
25	chr6	Sai nghĩa	*PKHD1*	c.12143A>G	p.Gln4048Arg	0.565695
26	chr6	Sai nghĩa	*PKHD1*	c.11696A>G	p.Gln3899Arg	0.553315
27	chr6	Sai nghĩa	*PKHD1*	c.5608T>G	p.Leu1870Val	0.955871
28	chr6	Sai nghĩa	*PKHD1*	c.3785C>T	p.Ala1262Val	0.354633
29	chr7	Sai nghĩa	*CFTR*	c.1408G>A	p.Val470Met	0.582069
30	chr8	Sai nghĩa	*PEX2*	c.550T>C	p.Cys184Arg	0.996206
31	chr8	Sai nghĩa	*TMEM67*	c.1810A>G	p.Ile604Val	0.692891
32	chr9	Sai nghĩa	*TJP2*	c.1539C>A	p.Asp513Glu	0.772165
33	chr9	Sai nghĩa	*BAAT*	c.59G>A	p.Arg20Gln	0.562899
34	chr10	Sai nghĩa	*ABCC2*	c.116A>T	p.Tyr39Phe	0.994609
35	chr11	Sai nghĩa	*SMPD1*	c.107T>C	p.Val36Ala	0.561302
36	chr11	Sai nghĩa	*PEX16*	c.346G>A	p.Val116Ile	1
37	chr11	Sai nghĩa	*BSCL2*	c.803A>G	p.Lys268Arg	0.134984
38	chr12	Sai nghĩa	*GYS2*	c.1087A>G	p.Met363Val	0.740615
39	chr13	Sai nghĩa	*COG6*	c.28G>A	p.Ala10Thr	0.484425
40	chr13	Sai nghĩa	*COG6*	c.94T>A	p.Cys32Ser	0.484225
41	chr13	Sai nghĩa	*COG6*	c.1340T>C	p.Met447Thr	0.0473243
42	chr13	Sai nghĩa	*ATP7B*	c.3419T>C	p.Val1140Ala	0.540535
43	chr13	Sai nghĩa	*ATP7B*	c.2855G>A	p.Arg952Lys	0.53095
44	chr13	Sai nghĩa	*ATP7B*	c.2495A>G	p.Lys832Arg	0.502396

l

STT	NST	Loại đột biến	Tên gen	HGVS.c	HGVS.p	1000Gp3_AF
45	chr13	Sai nghĩa	*ATP7B*	c.1366G>C	p.Val456Leu	0.376997
46	chr13	Sai nghĩa	*ATP7B*	c.1216T>G	p.Ser406Ala	0.376198
47	chr13	Sai nghĩa	*SLC10A2*	c.511T>G	p.Ser171Ala	0.840256
48	chr15	Sai nghĩa	*VPS33B*	c.1540G>A	p.Gly514Ser	0.767772
49	chr16	Sai nghĩa	*COG7*	c.1814C>T	p.Thr605Met	0.0880591
50	chr16	Sai nghĩa	*HSD3B7*	c.748A>G	p.Thr250Ala	0.493411
51	chr16	Sai nghĩa	*RPGRIP1L*	c.2231G>A	p.Arg744Gln	0.0621006
52	chr17	Sai nghĩa	*COG1*	c.1175A>G	p.Asn392Ser	0.520966
53	chr18	Sai nghĩa	*NPC1*	c.2572A>G	p.Ile858Val	0.485823
54	chr18	Sai nghĩa	*NPC1*	c.1926G>C	p.Met642Ile	0.831669
55	chr18	Sai nghĩa	*NPC1*	c.644A>G	p.His215Arg	0.21865
56	chr18	Sai nghĩa	*MYO5B*	c.2753G>A	p.Arg918His	0.135383
57	chr18	Sai nghĩa	*MYO5B*	c.376A>G	p.Thr126Ala	0.973442
58	chr18	Sai nghĩa	*ATP8B1*	c.3454G>A	p.Ala1152Thr	0.9998
59	chr19	Sai nghĩa	*ETFB*	c.734C>T	p.Thr245Met	0.419728
60	chr21	Sai nghĩa	*ITGB2*	c.1062A>T	p.Gln354His	1
61	chr22	Sai nghĩa	*TYMP*	c.1427C>T	p.Ser476Leu	0.142772

m

Phụ lục 3: Danh sách bệnh nhân và đối chứng

Bảng 1: Danh sách bệnh nhân tham gia nghiên cứu

STT	Labcode	MSBV	Họ và tên
1	BA181257	170460741	Phạm Bảo A
2	BA181258	110038735	Nguyễn Tiến Q
3	BA181238	180578351	Nguyễn Phúc K
4	BA181260	180182063	Nguyễn Khả Minh K
5	BA181261	160560511	Đỗ Khánh C
6	BA181262	170369178	Nguyễn Việt V
7	BA181263	179041746	Vũ Thị Kim L
8	BA181264	179463789	Lê Phương C
9	BA181265	170147613	Nguyễn Anh T
10	BA181266	186357835	Phạm Hoài
11	BA181267	179463997	Nguyễn Anh M
12	BA181268	130076252	Ngô Mỹ A
13	BA181269	120428010	Hoàng Minh C
14	BA181270	140128688	Trần Đức An
15	BA181271	170086887	Hoàng Thảo N
16	BA181272	150124693	Nguyễn Vũ Hà V
17	BA181273	180429201	Nguyễn Đỗ Danh T
18	BA181274	140212933	Cao Thị H
19	BA181275	160354514	Nguyễn Lâm Nhật A
20	BA181276	124646646	Lê Vân A
21	BA181277	160016423	Ngô Huy H
22	BA181278	160315164	Lương Gia H
23	BA181279	170542401	Nguyễn Minh K
24	BA181280	170330015	Lê Hùng A
25	BA181281	170396148	Lê Hà M
26	BA181282	110248371	Hà Văn Trung K
27	BA181283	180497737	Cao Bảo N
28	BA181284	180360507	Trịnh Hoài A
29	BA181285	170150090	Dương Bảo C
30	BA181286	150360954	Vũ Trần Bảo N
31	BA181287	180398743	Ngô Hồng N
32	BA181288	140044819	Hồ Thùy D
33	BA181289	160046591	Phạm Ngọc H
34	BA181290	150123064	Đinh Phương L
35	BA181291	120299809	Nguyễn Hải A
36	BA181292	140417057	Lê Gia H

n

STT	Labcode	MSBV	Họ và tên
37	BA181293	150293899	Khuất Thị Hoài A
38	BA181294	120336577	Lê Văn Đ
39	BA181295	160322921	Nguyễn Đắc Nhân T
40	BA190101	180104139	Trần Minh C
41	BA190102	180148874	Nguyễn Thùy L
42	BA190103	180451419	Lê Quang N
43	BA190104	180137113	Trần Trung K
44	BA190105	168542173	Nguyễn Ngọc Bảo N
45	BA190106	130356858	Ngô Quốc Đ
46	BA190107	160070280	Hà Thanh V
47	BA190108	140110382	Đỗ Hà N
48	BA190109	150381710	Đào Nguyễn Minh C
49	BA190110	188451212	Nguyễn Ngọc Khánh H
50	BA190111	170147613	Nguyễn Anh T
51	BA190112	170576708	Phạm Ngọc Khánh N
52	BA190113	150442271	Vũ Hải P
53	BA190114	140170258	Trần Khánh L
54	BA190115	170098510	Hà Như Q
55	BA190116	180532925	Nguyễn Khương D
56	BA190117	170408681	Khương Huyền T
57	BA190118	150355311	Trần Bảo A
58	BA190119	160500639	Ngô Minh K
59	BA190120	188510524	Lê Văn Minh K
60	BA190121	185569744	Trần Thị D
61	BA190122	170048054	Trần Thị Tiến A
62	BA190123	180017255	Nguyễn Quốc T
63	BA190124	120088518	Đỗ Huy H
64	BA190125	180580339	Phạm Thảo Ng
65	BA190126	180051847	Võ Văn Gia P
66	BA190127	160106337	Dương Đình Đức N
67	BA190128	140232868	Phạm Thế A
68	BA190129	120427680	Trần Anh T
69	BA190130	186258354	Hạ Gia H
70	BA190131	160026871	Phạm Đức T
71	BA190132	140422206	Nguyễn Thị Nguyệt N
72	BA190133	130011380	Nguyễn Ngọc A
73	BA190134	180398743	Ngô Hồng N
74	BA190135	160059308	Trần Diệu A
75	BA190136	100033417	Nguyễn Thị Ánh N

o

STT	Labcode	MSBV	Họ và tên
76	BA190137	139403132	Hồ Thanh P
77	BA190138	170228016	Nguyễn Bảo N
78	BA190139	190137132	Trương Ngọc B
79	BA190140	170015879	Nguyễn Thị Hà A
80	BA190141	165318624	Trương Thùy A
81	BA190142	170487688	Trần Trung N
82	BA190143	170533720	Nguyễn Tùng A
83	BA190144	170330015	Lê Hùng A
84	BA190145	150332150	Vũ Hoàng D
85	BA190146	180419103	Lê Ngọc A
86	BA190147	170029698	Phạm Anh T
87	BA190148	140029865	Phùng Bảo C
88	BA190149	140192994	Ngô Minh Đ
89	BA190150	110033009	Nguyễn Nguyên V
90	BA190151	170098934	Nguyễn Thạch Tùng L
91	BA190152	150348129	Nguyễn Nhật D
92	BA190153	170014111	Nguyễn Hoàng M
93	BA190154	186254354	Phạm Hải Đ
94	BA190155	170150090	Dương Thảo C
95	BA190156	170584025	Lưu Xuân Đ
96	BA190157	170543761	Nguyễn Viết Đăng Q
97	BA190158	170383509	Vũ Nguyễn Ngọc D
98	BA190159	180543632	Nguyễn Thiên A
99	BA190160	178798747	Trần Phương Anh
100	BA190161	180537889	Nguyễn Đăng Bảo P
101	BA190162	180533212	Hà Khánh L
102	BA190163	195453465	Nguyễn Quỳnh Bảo N
103	BA190164	170042295	Đặng Thị Ngọc A
104	BA190165	170525341	Nguyễn Thị Thanh T
105	BA190166	180370660	Phạm Nguyên Đại D
106	BA190167	179463789	Nguyễn Đình Thành L
107	BA190168	110351717	Phạm Minh Q
108	BA190169	150403546	Nguyễn Hoàng N
109	BA190170	140194568	Nguyễn Quang T
110	BA190171	170416324	Nguyễn Lê Diệu V
111	BA190201	180452946	Nguyễn Hữu Đ
112	BA190202	170575536	Phạm Phương Thúy V
113	BA190203	180310456	Nguyễn Thị Minh C
114	BA190204	180275903	Đặng Minh K

p

STT	Labcode	MSBV	Họ và tên
115	BA190205	170516225	Nguyễn Thị Th
116	BA190206	180600990	Đỗ Xuân B
117	BA190207	180504077	Nguyễn Thị Hạ V
118	BA190208	140044819	Hồ Thùy D
119	BA190209	130416915	Nguyễn Bích N
120	BA190210	180568676	Lê Ngọc Thùy D
121	BA190211	180471333	Vũ Anh T
122	BA190212	180630611	Nguyễn Nhật L
123	BA190213	190121437	Trần Minh C
124	BA190214	160085930	Nguyễn Ngọc A
125	BA190215	110381468	Đào Nguyên K
126	BA190216	180598430	Vũ Ngọc Bảo A
127	BA190217	170091398	Đặng Ngọc H
128	BA190218	169804165	Ngô Bảo C
129	BA190219	179148664	Bùi Thị Hồng D
130	BA190220	160518830	Nguyễn Minh K
131	BA190221	140110382	Đỗ Hà N
132	BA190222	140192994	Ngô Minh Đ
133	BA190223	170533720	Nguyễn Tùng A
134	BA190224	160122208	Đỗ Thị Hồng N
135	BA190225	100523451	Nguyễn Ngọc Gia H
136	BA190226	180310465	Nguyễn Thị Minh C
137	BA190301	170346955	Trần Thị Diễm M
138	BA190302	187612746	Phạm Thị Kim N
139	BA190303	154777773	Dương Mạnh H
140	BA190304	180484152	Nguyễn Minh K
141	BA190305	180504077	Nguyễn Thị Hạ V
142	BA190306	150058648	Trần Khánh L
143	BA190307	180517764	Nguyễn Anh T
144	BA190308	150392976	Trần Bảo A
145	BA190309	180108444	Nguyễn Khả H
146	BA190310	160029155	Nguyễn Minh K
147	BA190311	120144297	Hoàng Quốc H
148	BA190312	160372337	Đỗ Yến T
149	BA190313	160214206	Bùi Ngọc Thảo N
150	BA190314	190019963	Lương Diễm K
151	BA190315	180369329	Nguyễn Thúy N
152	BA181214	170024866	Ngô Ngọc H
153	BA181215	180502293	Bùi Yến M

q

STT	Labcode	MSBV	Họ và tên
154	BA181216	180471216	Đỗ Mạnh T
155	BA181217	150341124	Trần Ngọc Khánh L
156	BA181218	180325486	Ngô Ngọc Q
157	BA181219	179880633	Trần Quốc B
158	BA181220	110925151	Nguyễn Hoàng Bảo Y
159	BA181221	110183665	Hà Lê Thùy D
160	BA181222	140113461	Vũ Nguyễn Hoài A
161	BA181223	160932431	Nguyễn Hữu Quang A
162	BA181224	170476103	Đinh Khánh N
163	BA181225	170360190	Lưu Huy H
164	BA181226	182530401	Trần Lê An N
165	BA181227	130543534	Nguyễn Hoàng B
166	BA181228	170539266	Xa Văn T
167	BA181229	170383509	Vũ Nguyễn Ngọc D
168	BA181230	170175336	Mai Hà A
169	BA190320	110079519	Lê Văn T
170	BA190321	180591597	Phạm Tuấn H
171	BA190322	120088728	Nguyễn Quang M
172	BA190323	170197964	Đỗ Tuấn Đ
173	BA190324	180641085	Bùi Trung K
174	BA190325	190102286	Phùng Đức D
175	BA190326	180066404	Nguyễn Ngọc N
176	BA190327	116545132	Lê Tứ Thảo C
177	BA190328	160071909	Nguyễn Bảo B
178	BA190329	190136665	Đinh Minh N
179	BA190330	160275222	Lê Ngọc Bảo A
180	BA190331	180478485	Nguyễn Hoàng Thanh N
181	BA190332	150381710	Đào Nguyễn Minh C
182	BA190333	160991879	Trần Hữu Đ
183	BA190334	190126564	Nguyễn Quốc H
184	BA190335	170097633	Ngô Gia H
185	BA190336	180551707	Trần Tùng Diệp C
186	BA190337	170383509	Vũ Nguyễn Ngọc D
187	BA190338	190051547	Nguyễn Trung H
188	BA190339	160078557	Ngô Bảo N
189	BA190340	190026980	Nguyễn Như Q
190	BA190341	180256897	Đỗ Khánh D
191	BA190342	100894015	Dương Tiến Đ
192	BA190343	170048054	Trần Tiền A

r

STT	Labcode	MSBV	Họ và tên
193	BA190402	170024737	Lê Hữu Hồng P
194	BA190403	130001045	Nguyễn Trọng T
195	BA190404	180003873	Nguyễn Thị Ngọc H
196	BA190405	160244486	Nguyễn Tuyết P
197	BA190406	160315164	Lường Gia H
198	BA190407	120352720	Hoàng Thu H
199	BA190408	180562972	Đặng Hải Đ
200	BA190409	190107991	Vũ Thanh H
201	BA190410	110326139	Lê Lan H
202	BA190411	170379867	Trương Nhã O
203	BA190412	190202764	Nguyễn Thành T
204	BA190413	160549261	Nguyễn Huy Đ
205	BA190414	160052971	Lê Bảo T
206	BA190415	160393446	Nguyễn Hải Đ
207	BA190416	190123523	Nguyễn Kiên G
208	BA190417	190126156	Nguyễn Hoàng Hải Y
209	BA190418	169874452	Tạ Ngọc Hương T
210	BA190419	170189823	Nguyễn Xuân P
211	BA190501	180022727	Nguyễn An H
212	BA190502	196413216	Phạm Quang Gia M
213	BA190503	160518830	Nguyễn Minh Kh
214	BA190504	160029155	Nguyễn Minh K
215	BA190505	150377354	Nguyễn Thị Tố N
216	BA190506	180373984	Nguyễn Văn Tùng K
217	BA190507	187788333	Trần Công Việt T
218	BA190508	180462483	Vũ Bảo N
219	BA190509	160075443	Trần Vương Ánh V
220	BA190510	130000012	Lê Hương G
221	BA190511	130055755	Nguyễn Thọ H
222	BA190512	170319270	Nguyễn Đức Minh T
223	BA190513	120254512	Đỗ Nguyễn Nhật P
224	BA190514	140170258	Trần Khánh L
225	BA190515	150113335	Nguyễn Hữu K
226	BA190516	180110316	Trần Bảo A
227	BA190517	180370660	Phạm Nguyễn Đại D
228	BA190518	160400007	Nguyễn Thị Huyền T
229	BA190519	180603870	Vũ Nguyễn Tuấn P
230	BA190520	170203490	Đào Đức T
231	BA190521	190036087	Trịnh Ngọc Hoài A

s

STT	Labcode	MSBV	Họ và tên
232	BA190522	190050643	Liễu Cát Bảo T
233	BA190523	160175078	Nguyễn Thùy C
234	BA190524	140060357	Đào Phương L
235	BA190525	187245749	Nguyễn Cẩm L
236	BA190526	198542665	Vũ Minh K
237	BA190527	190200362	Ngô Hoài Vĩnh A
238	BA190528	180256676	Trần Vũ Hà V
239	BA190529	160046473	Vũ Thùy T
240	BA190530	170569303	Hoàng Thị N
241	BA190531	130041532	Lê Bảo N
242	BA190532	100223795	Nguyễn Khánh N
243	BA181231	130371426	Phạm Thị Quỳnh C
244	BA181232	140048525	Trần Thị Hải Y
245	BA181233	165005006	Vũ Gia K
246	BA181234	180134439	Trần Tiến Đ
247	BA181235	170320377	Lê Thái S
248	BA181236	180498003	Bàng Gia H
249	BA181237	170558730	Phạm Khánh B
250	BA181239	180452946	Nguyễn Hữu Đ
251	BA181240	180630611	Nguyễn Nhật L
252	BA181241	180434791	Lương Khánh L
253	BA181242	170042295	Đặng Thị Ngọc A
254	BA181243	130149075	Nguyễn Huy Q
255	BA181244	180619268	Vi Bảo Y
256	BA181245	150058648	Trần Khánh L
257	BA181246	170463371	Đặng Gia H
258	BA181247	170584025	Lưu Xuân Đ
259	BA181248	160024193	Nguyễn Huy H
260	BA181249	180609006	Nguyễn Thị Bảo A
261	BA181250	170018227	Nguyễn Gia H
262	BA181251	160396869	Vũ Hoàng M
263	BA181252	170181317	Chu Thị Thanh N
264	BA181254	160071909	Nguyễn Bảo B
265	BA181255	160456781	Nguyễn Trọng D
266	BA181256	180340052	Nguyễn Minh K

t

Bảng 2: Danh sách mẫu đối chứng

STT	Labcode	Họ và tên	Năm sinh
1	DC1	Nguyễn Đăng N	2016
2	DC2	Phạm Hoài A	2018
3	DC3	Hoàng Thảo N	2017
4	DC4	Lê Hà M	2017
5	DC5	Nguyễn Hải A	2012
6	DC6	Nguyễn Thị Bảo A	2018
7	DC7	Nguyễn Gia H	2017
8	DC8	Nguyễn Khả Minh K	2018
9	DC9	Trần Thị Hải Y	2014
10	DC10	Trần Tiến Đ	2018
11	DC11	Phạm Bảo A	2017
12	DC12	Nguyễn Bảo B	2015
13	DC13	Nguyễn Hoàng B	2013
14	DC14	Nguyễn Hoàng Bảo Y	2011
15	DC15	Nguyễn Vũ Hà V	2015
16	DC16	Phạm Hải Đ	2018
17	DC17	Bùi Yến M	2018
18	DC18	Nguyễn Đỗ Danh T	2018
19	DC19	Lương Gia H	2016
20	DC20	Lê Quang N	2018
21	DC21	Nguyễn Trọng D	2016
22	DC22	Hồ Thanh P	2017
23	DC23	Trần Quốc B	2017
24	DC24	Đỗ Hà N	2013
25	DC25	Vũ Nguyễn Ngọc D	2017
26	DC26	Dương Thảo C	2017
27	DC27	Vũ Ngọc Bảo A	2018
28	DC28	Trương Thùy A	2017
29	DC29	Vi Bảo Y	2018
30	DC30	Ngô Ngọc Q	2018
31	DC31	Lê Hùng A	2016
32	DC32	Nguyễn Hải A	2012
33	DC33	Nguyễn Nhật L	2017
34	DC34	Trần Bảo N	2018
35	DC35	Vũ Hoàng M	2016
36	DC36	Phạm Bảo P	2016
37	DC37	Đỗ Khánh C	2016
38	DC38	Trần Ngọc Khánh L	2015
39	DC39	Nguyễn Huy Q	2013

u

STT	Labcode	Họ và tên	Năm sinh
40	DC40	Phạm Thị Quỳnh C	2013
41	DC41	Phùng Bảo C	2014
42	DC42	Hà Văn Trung K	2011
43	DC43	Nguyễn Hữu Đ	2018
44	DC44	Đặng Gia H	2017
45	DC45	Nguyễn Lâm Nhật A	2011
46	DC46	Nguyễn Thị Bảo A	2018
47	DC47	Nguyễn Hoàng Bảo Y	2011
48	DC48	Trịnh Hoài A	2018
49	DC49	Vi Bảo Y	2018
50	DC50	Phạm Phương Thúy V	2017
51	DC51	Nguyễn Phúc K	2018
52	DC52	Ngô Ngọc H	2016
53	DC53	Nguyễn Thiên A	2018
54	DC54	Nguyễn Ngọc Bảo N	2018
55	DC55	Nguyễn Thị Minh C	2017
56	DC56	Đặng Minh K	2018
57	DC57	Lương Khánh L	2010
58	DC58	Nguyễn Thị T	2017
59	DC59	Đỗ Xuân B	2018
60	DC60	Nguyễn Thị Hạ V	2018
61	DC61	Nguyễn Thùy L	2017
62	DC62	Đặng Gia H	2017
63	DC63	Nguyễn Ngọc Khánh H	2017
64	DC64	Nguyễn Gia H	2017
65	DC65	Mai Hà A	2017
66	DC66	Hồ Thùy D	2013
67	DC67	Nguyễn Bích N	2013
68	DC68	Lê Ngọc Thùy D	2018
69	DC69	Vũ Anh T	2018
70	DC70	Nguyễn Thị Ánh N	2010
71	DC71	Nguyễn Huy H	2015
72	DC72	Nguyễn Nhật L	2018
73	DC73	Trần Minh C	2006
74	DC74	Nguyễn Ngọc A	2016
75	DC75	Đào Nguyên K	2011
76	DC76	Trần Khánh L	2014
77	DC77	Đặng Thị Ngọc A	2016
78	DC78	Hà Khánh L	2018
79	DC79	Lê Quang N	2018
80	DC80	Vũ Ngọc Bảo A	2018

STT	Labcode	Họ và tên	Năm sinh
81	DC81	Nguyễn Thị Bảo A	2018
82	DC82	Đặng Ngọc H	2016
83	DC83	Nguyễn Lê Diệu V	2017
84	DC84	Ngô Hồng N	2018
85	DC85	Nguyễn Thị Nguyệt N	2014
86	DC86	Ngô Bảo C	2017
87	DC87	Bùi Thị Hồng D	2017
88	DC88	Nguyễn Minh K	2016
89	DC89	Đỗ Hà N	2013
90	DC90	Ngô Minh Đ	2014
91	DC91	Nguyễn Tùng A	2017
92	DC92	Đỗ Thị Hồng N	2016
93	DC93	Lê Gia H	2014
94	DC94	Trần Thị Hải Y	2014
95	DC95	Phạm Hoài A	2018
96	DC96	Hồ Thanh P	2017
97	DC97	Nguyễn Ngọc Gia H	2013
98	DC98	Nguyễn Anh T	2017
99	DC99	Nguyễn Thị Minh C	2017
100	DC100	Trần Thị Diễm M	2017
101	DC101	Phạm Thị Kim N	2018
102	DC102	Dương Mạnh H	2015
103	DC103	Nguyễn Minh K	2018
104	DC104	Nguyễn Thị Hạ V	2018
105	DC105	Trần Khánh L	2014
106	DC106	Nguyễn Anh T	2018
107	DC107	Trần Bảo A	2015
108	DC108	Nguyễn Khả H	2018
109	DC109	Nguyễn Minh K	2016
110	DC110	Hoàng Quốc H	2012
111	DC111	Đỗ Yến T	2016
112	DC112	Bùi Ngọc Thảo N	2018
113	DC113	Lương Diễm K	2018
114	DC114	Nguyễn Thúy N	2018
115	DC115	Đỗ Thị Hồng N	2016
116	DC116	Vũ Gia K	2016
117	DC117	Cao Bảo N	2018
118	DC118	Nguyễn Huy Q	2013
119	DC119	Nguyễn Khả Minh K	2018
120	DC120	Lê Phương C	2017
121	DC121	Nguyễn Anh M	2017

STT	Labcode	Họ và tên	Năm sinh
122	DC122	Đỗ Minh A	2018
123	DC123	Nguyễn Phạm H	2016
124	DC124	Tô Tiến P	2012
125	DC125	Nguyễn Ngọc P	2017
126	DC126	Nguyễn Mai T	2018
127	DC127	Phan Đỗ V	2016
128	DC128	Nguyễn Trà M	2016
129	DC129	Nguyễn gia L	2016
130	DC130	Hoàng Thị H	2015
131	DC131	Nguyễn Bảo T	2013
132	DC132	Nguyễn Thị T	2013
133	DC133	Hàn Nam P	2014
134	DC134	Lê Minh A	2011
135	DC135	Nguyễn Công T	2014
136	DC136	Cù Bảo N	2016
137	DC137	Phan Minh L	2018
138	DC138	Nguyễn Xuân N	2018
139	DC139	Nguyễn Hương B	2018
140	DC140	Tạ Minh A	2018
141	DC141	Trần Nguyễn M	2016
142	DC142	Lê Quang V	2017
143	DC143	Lê Ngọc L	2018
144	DC144	Bùi Hà L	2014
145	DC145	Nghiêm Thị H	2017
146	DC146	Nguyễn Ngọc Y	2017
147	DC147	Nguyễn Tuyết C	2016
148	DC148	Lê Thị A	2013
149	DC149	Trần Đức H	2014
150	DC150	Hà Phương U	2017
151	DC151	Nông Triệu Bảo D	2018
152	DC152	Hoàng Tuấn K	2011
153	DC153	Lê Hữu K	2013
154	DC154	Kha Hoàng N	2016
155	DC155	Trần Thiên A	2018
156	DC156	Nguyễn Văn M	2017
157	DC157	Nguyễn Minh Q	2015
158	DC158	Nguyễn Thùy L	2011
159	DC159	Vũ Diễm X	2010
160	DC160	Sòng Minh D	2014
161	DC161	Lê Ngọc H	2012
162	DC162	Nguyễn Ngọc C	2012

STT	Labcode	Họ và tên	Năm sinh
163	DC163	Dương Phúc K	2013
164	DC164	Trịnh Minh Q	2015
165	DC165	Trần Trung T	2018
166	DC166	Phạm Kim N	2016
167	DC167	Nguyễn Thiện L	2017
168	DC168	Đặng Nguyễn Hạ L	2014
169	DC169	Trần Anh K	2014
170	DC170	Phạm Trần Đăng K	2016
171	DC171	Nguyễn Phú H	2013
172	DC172	Nguyễn Hải N	2012
173	DC173	Chu Nam A	2012
174	DC174	Tạ Lâm T	2017
175	DC175	Sòng Anh T	2015
176	DC176	Lê Đăng Minh T	2011
177	DC177	Nguyễn Khánh N	2016
178	DC178	Trần Thị Mai P	2016
179	DC179	Nguyễn Hữu P	2017
180	DC180	Phạm Ngọc Gia H	2015
181	DC181	Chu Bá N	2015
182	DC182	Nguyễn Thị Phương A	2013
183	DC183	Trương Long V	2014
184	DC184	Nguyễn Minh T	2017
185	DC185	Nguyễn Tuyết V	2011
186	DC186	Nguyễn Thùy L	2015
187	DC187	Chu Mạnh C	2013
188	DC188	Nguyễn Trần An N	2014
189	DC189	Trần Gia Cát T	2014
190	DC190	Nguyễn Thanh T	2018
191	DC191	Phạm Tường A	2018
192	DC192	Phạm Tuấn B	2013
193	DC193	Ứng Minh K	2011
194	DC194	Nguyễn Gia K	2015
195	DC195	Bùi Văn H	2017
196	DC196	Diệp Thảo A	2010
197	DC197	Bùi Thị Gia H	2010
198	DC198	Ninh Đăng K	2018
199	DC199	Nguyễn Viết M	2012
200	DC200	Hoàng Quang M	2010
201	DC201	Hà Trần Khánh N	2010
202	DC202	Bùi Kế N	2010
203	DC203	Phạm Minh D	2015

y

STT	Labcode	Họ và tên	Năm sinh
204	DC204	Bùi Yến N	2012
205	DC205	Phạm Vân A	2017
206	DC206	Quách Ngọc T	2011
207	DC207	Phạm Quang P	2014
208	DC208	Lò Thị Phương T	2011
209	DC209	Hà Minh V	2018
210	DC210	Đinh Nhật A	2017
211	DC211	Nguyễn Hoài T	2011
212	DC212	Nguyễn Ngọc Bảo T	2011
213	DC213	Trần Đăng K	2012
214	DC214	Bùi Đức Tuệ L	2011
215	DC215	Quản Hà M	2012
216	DC216	Nguyễn Kiên C	2012
217	DC217	Tô Anh D	2016
218	DC218	Nguyễn Hải N	2012
219	DC219	Hoàng Tuấn M	2010
220	DC220	La Hữu T	2016
221	DC221	Phạm Chí T	2016
222	DC222	Nguyễn Viết T	2017
223	DC223	Nguyễn Thảo M	2017
224	DC224	Vũ Thị Minh T	2017
225	DC225	Hoàng Khánh A	2014
226	DC226	Nguyễn Ngọc Minh C	2011
227	DC227	Phạm Thị Anh T	2016
228	DC228	Đỗ Bảo H	2016
229	DC229	Lê Bảo A	2014
230	DC230	Ví Việt P	2018
231	DC231	Lê Thùy L	2013
232	DC232	Đỗ Nhật V	2011
233	DC233	Lữ Duy M	2017
234	DC234	Nguyễn Hương Q	2016
235	DC235	Nguyễn Trung D	2014
236	DC236	Phạm Nhật M	2013
237	DC237	Đào Minh D	2011
238	DC238	Nguyễn Phương T	2015
239	DC239	Phan Minh K	2016
240	DC240	Nguyễn Khánh D	2012
241	DC241	Hoàng Minh H	2015
242	DC242	Nguyễn Thùy D	2017
243	DC243	Nguyễn Thúy A	2011
244	DC244	Nguyễn Phạm Hoàng Y	2017

z

STT	Labcode	Họ và tên	Năm sinh
245	DC245	Đỗ Thị Mai P	2017
246	DC246	Phạm Trung P	2015
247	DC247	Nguyễn Vũ Hải B	2016
248	DC248	Trần Trí K	2012
249	DC249	Phạm Dương Kiều A	2013
250	DC250	Bùi Quốc D	2017