MINISTRY OF EDUCATION AND TRAINING

VIETNAM ACADEMY OF SCIENCE AND TECHNOLOGY

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY

Nguyen Van Tung

STUDY ON MUTATIONS AND POLYMORPHISMS OF SEVERAL GENES IN PEDIATRIC PATIENTS WITH BILIARY ATRESIA

SUMMARY OF DISSERTATION ON APPLIED BIOLOGY

Major: Biotechnology

Mã số: 9420201

Ha Noi - 2024

The dissertation is completed at: Graduate University of Science and Technology, Vietnam Academy Science and Technology

Supervisors:

1. Supervisor 1: Prof.Dr. Nguyen Huy Hoang – Institute of genome research – Vietnam Academy of Science and Technology
2. Supervisor 2: Dr. Nguyen Thi Kim Lien – Institute of genome research – Vietnam Academy of Science and Technology

Referee 1: Prof.Dr. Tran Huy Thinh

Referee 2: Assoc.Prof.Dr. Vu Chi Dung

Referee 3: Prof.Dr. Nguyen Duy Bac

The dissertation is examined by Examination Board of Graduate University of Science and Technology, Vietnam Academy of Science and Technology at…….. (time, date……)

The dissertation can be found at:

1. Graduate University of Science and Technology Library
2. National Library of Vietnam

INTRODUCTION

Biliary atresia is a rare disease which characterized by progressive inflammation, fibrosis of both the extrahepatic and intrahepatic bile ducts, resulting in biliary obstruction. Defects in genes involved in bile duct formation and the immune system are a factor in biliary atresia. Therefore, identifying genetic variations associated with the disease would be a marker that could aid in early diagnosis and improve surgical outcomes.

In Viet Nam, genetic study of biliary atresia related genes is lacking. So far, only clinical studies of BA patients who were treated in Hanoi and Ho Chi Minh City. Those studies only focused on evaluation of the effectiveness of surgery and postoperative treatment in biliary atresia patients. This study, "Study on mutations and polymorphisms of several genes in pediatric patients with biliary atresia," was performed to determine genetic changes in genes related to congenital biliary atresia. Changes in genes related to the disease, including mutations and single nucleotide polymorphisms, will be verified in a large number of patients to find the relationship between genotype and phenotype.

Research objective

- Identification of genetic variants associated with congenital biliary atresia by whole-exome sequencing.

- Study on the association between single nucleotide polymorphisms and the risk of biliary atresia in Vietnamese people

Research content

- Collection of blood samples from patients with biliary atresia and control.

- Whole exome sequencing of several patients with typical biliary atresia.

- Analysis of data to identify genetic variants associated with biliary atresia.

- Study on the association of single nucleotide polymorphisms with the risk of congenital biliary atresia in a large number of patients and controls.

New contributions of the thesis

1. This study has identified a novel double heterozygous novel mutation c.412G>A (p.Gly138Arg) in the *FAH* gene and c.2225A>G (p.Tyr742Cys) in the *ERCC4* gene of two siblings. Beside, we report a homozygous frameshift mutation c.50_51insG (p.Gly17Glyfs77*) in the *KRT18* gene, which involves biliary atresia, and a compound heterozygous mutation c.314C>A (p.Ser105∗), c.2975C>T (p.Pro992Leu) in the *ATP7B* gene in a patient with biliary atresia combined with Wilson disease.
2. The association between single nucleotide polymorphisms rs2287622 in the *ABCB11* gene, rs927344 in the *ABCC2* gene, and rs1815930 in the *MYO5B* gene with the risk of biliary atresia has been evaluated.

CHAPTER 1. LITERATURE OVERVIEW

1.1. Biliary atresia

Biliary atresia is characterized by abnormalitíe in the duct system, leading to blockage of the bile ducts, resulting cholestasis. The current diagnosis bases on clinical and subclinical such as bilirubin, liver biopsy,… and surgery for definitive diagnosis. The best time of operation is before 8 weeks-old. Without diagnosis and surgery, 50-80% of patients will die due to cirrhosis at 1 year of age. Therefore, finding a diagnosis method that helps doctor and the patient's family to decide an early operation is necessary. Genetic defects of the morphogenesis of the biliary system and immune systems are factors causing biliary atresia, therefore, identify genetic

alterations in genes associated with biliary atresia may be a marker that may support diagnosis and improve surgical outcomes

The disease is treated by surgery in the newborn. The late surgery will lead to secondary cirrhosis resulting in death of 1 year-old. Even though when surgery is successful, inflammation of bile ducts leads to cirrhosis and 70-80% of patients need liver transplants. The incidence in Europe and North America ranges from 1/14,000-1/20,000 live births. The incidence in Asian and Western Pacific countries is high with an incidence of 1:9,000 in Japan and 1:3,000 in Taiwan.

1.1.1. Clinical and classification of biliary atresia

The main clinical symptoms of congenital biliary atresia are jaundice, pale stools, and enlarged liver. Jaundice occurs from the postnatal period, following the physiological jaundice. However, in some children, jaundice and pale stools appear later. It is found that in more than half of children, yellow or light-yellow meconium, dark yellow urine and enlarged liver are due to cholestasis. Therefore, the enlarged liver increases, according to the age of the child. It is noteworthy that most children with congenital biliary atresia have normal weight and normal body development in the early months, even until surgery. However, there are still a few children with malnutrition, anemia or developmental delay. Symptoms included decreased levels of prothrombin due to poor vitamin K absorption, intracranial hemorrhage, skin hemorrhagic, hyperbilirubinemia predominantly bilirubin, elevated levels of bile acid concentrations, lipo- Protein X (Lp-X), and gamma-glutamyl transpeptidase:

According to the site of atresia in the extrahepatic biliary systems, biliary atresia may also be divided into 3 separate:

- Type I: atresia in the common bile duct (5%).

- Type II: atresia in the hepatic duct (2%).
- Type III: atresia in the most proximal part of the bile ducts (up to 90%).

Based on clinical biliary atresia can be divided into:

- Congenital biliary atresia without any other abnormalities (about 70-85%). In this group, children are with progressive jaundice in the first 2 months of life; in combination with pale colored infant stools.

- Biliary atresia splenic malformation (BASM) accounts for about 10 to 15% of cases. The malformations include splenic abnormalities, reversal of visceral, abnormal rotation, interrupted inferior vena cava, and cardiac anomalies. Postoperative prognosis is not good compared to no malformations.

1.1.2. Etiology of biliary atresia

Up to now, understanding of the etiology/pathophysiology of liver lesions and biliary system in congenital biliary atresia has not been clarified. Theories of the etiology have been suggested such as genetic disorders, autoimmune, immune abnormalities, toxicity, viral infections…..

1.2. The role of genetic factors in the pathogenesis of biliary atresia

Most investigators working in this area believe that biliary atresia is not a disease with a single etiology but a combination of different phenotypes that share certain clinical features, such as obliteration of the biliary tree early in life.

It should be noted that each genes do not imply that they exert only one role in the pathogenesis of biliary atresia, as the majority of genes contribute to many biological processes. However, biliary atresia is not an isolate disease but a common pathology of various pathogenetic pathways. As based on the pathological features of BA and gene's function, potential genetic variants of BA can be classified into 4 pathways: hepatobiliary development, fibrosis, inflammation, ciliopathy.

1.3. Research situation of biliary atresia

1.3.1. Research on the biliary atresia in the world

Congenital biliary atresia is known to be a highly heterogeneous hereditary disease that is difficult to diagnose and treat. Not being properly diagnosed and treated early may result in unsuccessful treatment or death in infants. Therefore, many studies have been conducted in the world to determine the exact genetic cause of the disease to guide treatment and genetic counseling for the family of patients. In recent years, new generation sequencing (Whole-genome sequencing and Whole-exome sequencing) has been used in studies in biliary atresia patients to identify new genetic alterations and new candidate genes related to biliary atresia disease.

In the world, scientists have found a relationship between variation/ polymorphism in multiple genes associated with biliary atresia. Whole-exome sequencing has been considered as an effective tool for the detection of novel disease-causing genes Tens of thousands of genetic variants can be identified in each coding region in many complex diseases such as cardiovascular, neurological, metabolic ... Therefore, we aim to conduct WES in biliary atresia patients to find out the genetic cause of the disease.

1.3.2. Research situation in Vietnam

In Vietnam, the prevalence of biliary atresia disease is estimated at 1: 2400 live births. Each year, approximately 70 pediatric BA patients are treated at Vietnam National Children's Hospital. However, genetic study of BA related genes is lacking. So far, only clinical studies of biliary atresia patients who were treated in Hanoi and Ho Chi Minh City. These publications only focused evaluation of the effectiveness of surgery and postoperative treatment in biliary atresia patients. Therefore, this research will become the first study in Vietnam in genetics in Vietnamese biliary

atresia patients. The results of the study will contribute to the better understanding of genetic aspects of biliary atresia in Vietnamese patients.

Chapter 2: MATERIALS AND METHODS

2.1. Study subjects

Two hundred and sixty-six of Vietnamese patient with biliary atresia A control cohort of 250 healthy people was formed.

2.2. Instruments and equipment

All instruments and equipment used in this study were provided by Institute of Genome Research, Vietnam Academy of Science and Technology

2.3. Methodologies

2.3.1. Total DNA extraction

Blood from probands were collected into EDTA tubes and stored at -20^{o}C. Genomic DNA was extracted from subject's blood using QIAamp DNA Mini Kit (QIAGEN, Germany) according to manufacturer's protocol.

2.3.2. Whole exome sequencing

2.3.2.1. Preparation and quantification of DNA library

DNA library was prepared using SureSelect XT Library Prep Kit ILM (Agilent) following the manufacturer's instruction.

2.3.2.2. Whole exome sequencing using Illumina NextSeq 500

Whole-exome sequencing was processed using kit NexSeq 500TM High Output kit (150 cycles) Illumina following the manufacturer's instruction.

2.3.3. Variant calling and annotation

After sequenced by Illumina platform, raw data was assessed and subjected to quality control using FastQC. The paired-end reads were aligned to the reference human genome (GRChr37/hg19) using BWA 0.7.10. Picard tools was used to processed post-alignment data. Genome Analysis Toolkit

v3.4 was used for variant calling.The The effects of variants on genes such as amino acid changes were predicted using SnpEff v4.1. *In-silico* analyses to confirm the effect of the mutations on the structure and function of the proteins was performed.

2.3.4. Screening of potential pathogenic variants/polymorphisms in the genes associated with biliary atresia

The potential pathogenic variants/polymorphisms in the genes associated with biliary atresia were filtered . The pathogenic variants with MAF<1% were *in silico analyzed* to confirm the effect of the mutations on the structure and function of the proteins. The variants with MAF>1% (single nucleotide polymorphism) were evaluated in a large number of patients and controls.

2.3.5. Genotyping of rs2287622 using RFLP

A fragment of *ABCB11* gene which contained SNP rs 2287622 was amplified using specific primer. The primer designed using Primer 3 software was shown on Method section. PCR product of 333 bp was obtained from all of 200 BA samples and 150 healthy controls. The PCR products were electrophored with agarose 1%.

2.3.6. Genotyping of SNP using ARMS-PCR

In this study, genotyping of SNP rs927344 in *ABCC2* gene and rs1815930 in *MYO5B* gene were performed using ARMS-PCR.

2.3.7. Sanger sequencing

PCR products were purified using Multiscreen PCR 96 Filter Plate. The purified PCR products were then sequenced using ABI Prism BigDye Terminator Cycle Sequencing Kit Version 3.1on an ABI genetic analyzer 3500.

2.3.8. Statistical analysis

Chi-square test was used to test whether allele distribution of the SNP follows Hardy-Weinberg Equilibrium (HWE). The odds ratios and 95% confidence intervals were calculated based on the formula described by Szumilas in 2010 (Szumilas, 2010). Chi-squared test was used to test three models (additive, dominant, recessive) for associations of the SNP with risk of biliary atresia.

Chapter 3: RESULTS

3.1. Samples collection and DNA extraction

Genomic DNA was extracted from blood of a total of 266 patients with Biliary atresia and 250 healthy controls. The DNA were electrophored with agarose 1%. Results were shown in supplementary 1 and 2. DNA after extracted will be stored at -20°C.

3.2. Whole exome sequencing

3.2.1. Preparation and quantification of DNA library

DNA library of 05 patients were prepared using SureSelect XT Library Prep Kit ILM (Agilent) following the manufacturer's instruction. The libraries had high quality and concentration; most of the DNA fragments had sizes ranging from 250 to 500 bp.

3.2.2. Whole exome sequencing

The data generated from the new generation sequencer has a high capacity, good quality, the ratio of nucleotides with a quality score greater than 20. Statistics of sequential quality are shown in Table 3.4

Table 3.4. Information of WES sequence of 05 patients with biliary atresia

	BA01	**BA02**	**BA03**	**BA04**	**BA05**
Total read base	9.933 Mb	9.247 Mb	9.693 Mb	10.973 Mb	10.797 Mb

	BA01	BA02	BA03	BA04	BA05
Total reads	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	
Average Read Length (bp)	149,24	149,15	149,43	149,22	148,77
Average Throughput Depth of Target regions (X)	164,3	152,9	160,3	181,5	157,6

3.3. Bioinformatics analysis and variants screening

3.3.1. Alignment data with reference genome

Whole exome sequencing data of 05 patients wwith biliary atresia were aligned to the reference human genome (GRChr37/hg19) using BWA 0.7.10. It can be seen that the results of sequencing data alignment of the patient with the reference genome are good (successful alignment rate of 99.8-99.9%), the rate of aligned data segments located in the exone region is high (70-79%).

3.3.2. Variant calling and annotation

After alignment, data is processed using Picard software. Variation is identified by GATK. *In-silico* analyses to confirm the effect of the mutations

on the structure and function of the proteins was performed using SIFT. The result was shown in table 3.6

Table 3.6: Results of analysis and variation prediction for the entire coding region of patients with biliary atresia

Parameter metric	BA01	Mẫu BA02	BA03	BA04	BA05
SNP	**23.295**	**23.743**	**23.413**	**23.508**	**23.774**
Synonymous Variant	11.870	12.033	11.958	12.012	12.036
Missense Variant	11.271	11.559	11.317	11.368	11.579
Stop Gained	114	112	99	89	119
Stop Lost	40	39	39	39	40
Indel	**724**	**711**	**738**	**776**	**712**
Frameshift Variant	325	318	348	350	317
Inframe Insertion	182	174	175	196	187
Inframe Deletion	217	219	215	230	208

3.4. Variants screening

After filtering, these variants located in the exone regions of genes associated with biliary atresia were retained. The list of genes used for screening of mutant variants is shown in Appendix 2. These variants can be divided into two categories: (i) mutation—variants with an allele frequency less than 1%; (ii) single nucleotide polymorphisms—variants with an allele frequency >1%.

3.5. In silico analysis

Potential mutations were considered as pathogenic if all four tools including SIFT, PolyPhen2, Mutation Taster, Provean assessed them as damage. As a result, patients BA01 and BA02 carried a double heterozygous c.412G>A (p.Gly138Arg) in the *FAH* gene and c.2225A>G (p.Tyr742Cys) in the *ERCC4* gene; patients BA03 and BA04 carried a heterozygous variant c.368A>C in *CYP27A1* gene; patient BA05 had a homozygous frameshift mutation c.50_51insG (p.Gly17Glyfs77 ∗) in *KRT18* gene.

3.5.1. Patients BA01 and BA02

Patients BA01 and BA02 are two siblings. They carried compound heterozygous mutations (double heterozygotes) in the FAH and ERCC4 genes. Both patients inherited the c.412G>A (p.Gly138Arg) mutation in the FAH gene from their father and the c.2225A>G (p.Tyr742Cys) mutation in the ERCC4 gene from their mother (Figure 3.3).

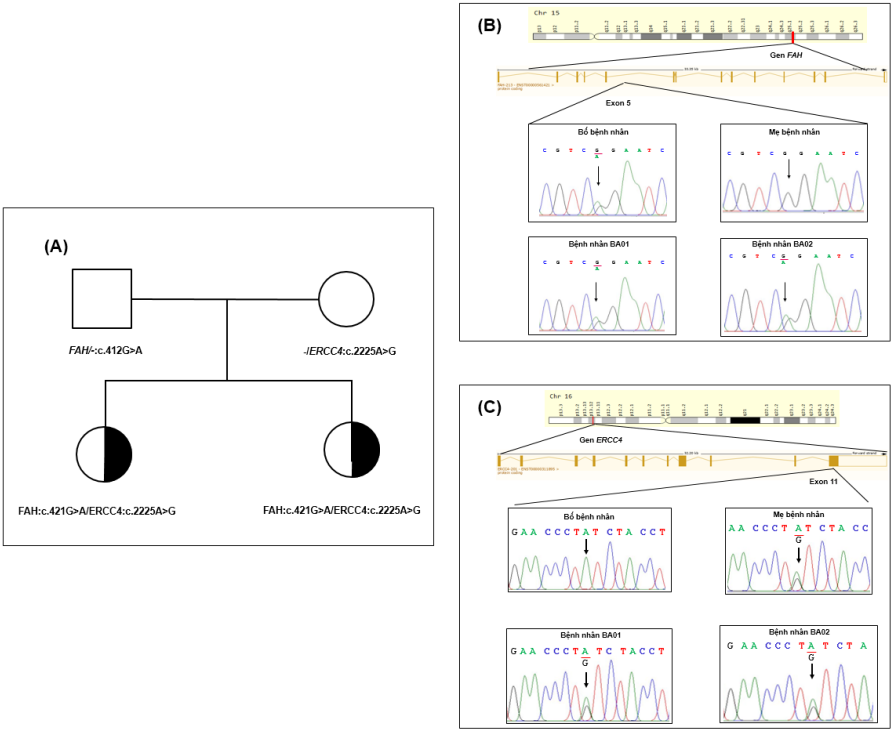

Figure 3.3: Double heterozygous mutation in patients BA01and BA02 (A) Genealogy diagram (B) Sanger sequencing of mutation c.412G>A in FAH gene; (C) Sanger sequencing of mutation c.2225A>G in ERCC4 gene

The comparation of amino acid sequences of two mutation in different species including*H.sapien*, *P.troglodytes*, *M.mulatta*, *F.catus* and *M.musculus* showed that Gly138 residue in protein coding by *FAH* and Tyr742 residue in protein coding by *ERCC4* are highly conservative region (Figure 3.4).

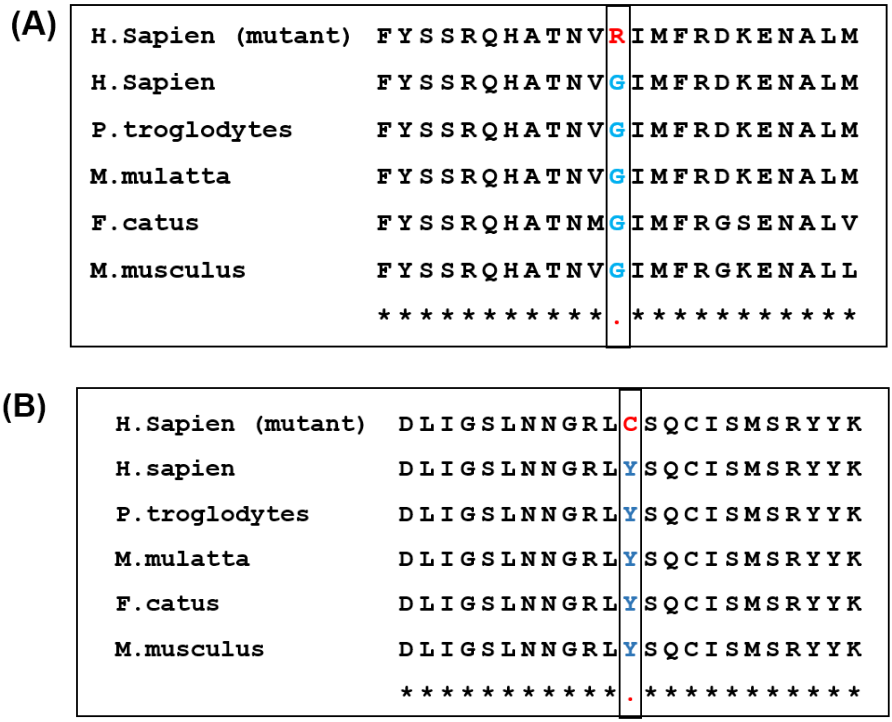

Figure 3.4: Conservation of the amino acid across different species (A) Mutation p.Gly138Arg in FAH protein; (B) Mutation p.Tyr742Cys in ERCC4 protein

3.5.2. Patients BA03 and BA04

The analysis result shown that two patients BA03 and BA4 carried a heterozygous mutation c.368A>C in *CYP27A1* gene. However, Sanger sequencing did not detect this change. Therefore, we could not determine the causative agent of the two diseases BA03 and BA04.

3.5.3. Patient BA05

The analysis results showed that patient BA05 carried a homozygous mutation c.50_51insG (p.Gly17Glyfs77 ∗) in the *KRT18* gene (Figure 3.7). The insertion of a Guanine (G) nucleotide at position 51 resulted in a

frameshift, causing the protein to be truncated 77 amino acids after the mutation site.

Patient BA05 was also diagnosed with Wilson's disease, so we expanded the scope of mutation screening on genes related to Wilson's disease. The results found that the patient carried two heterozyous mutations c.314C>A (p.Ser105∗) and c.2975C>T (p.Pro992Leu) in the *ATP7B* gene. These two mutations were recorded as pathogenic in the ClinVar database and reported in the dbSNP database with the code numbers rs753236073 and rs201038679, respectively.

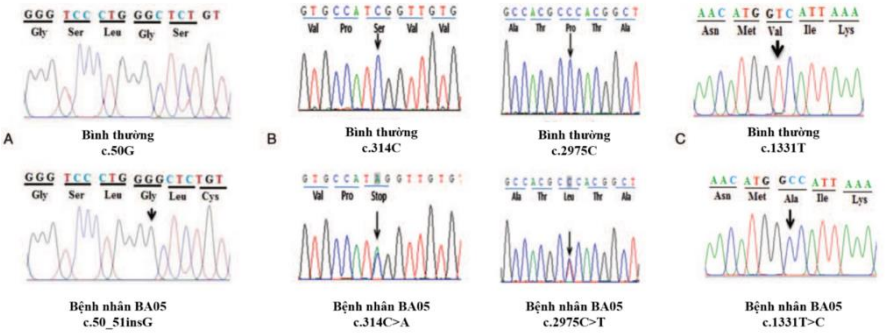

Figure 3.7: Mutation in patient BA05

(A) Homozygous mutation c.50_51insG in KRT18 gene; (B) Compound heterozygous c.314C>A and c.2975C>T in ATP7B gene; (C) Homozygous mutation c.1331T>C in ABCB11 gene

In addition, patient BA05 carried the homozygous variant c.1331T>C (p.Val444Ala) in the *ABCB11* gene. This is a single nucleotide polymorphism recorded in the dbSNP database under the code rs2287622. The influence of this SNP on the susceptibility to congenital biliary atresia will be evaluated in this study.

3.6. The association between single nucleotide polymorphisms with the risk of biliary atresia

3.6.1. The association between SNP rs2287622 with the risk of biliary atresia

PCR products were digested using restriction enzyme *HaeIII*. Base on recognition site of *HaeIII*, the homozygote genotype TT was not digested, the homozygote genotype CC was digested to produce bands of 206 bp and 127 bp, the heterozygote genotype TC was digested to produce bands of 333bp, 206 bp and 127 bp. The genotypes of rs2287622 were shown on table 3.11

Table 3.11: Genotype and alleles of rs2287622 polymorphism of ABCB11 gene in the BA patients and controls

	Allele			Allele frequency (%)		HWE p-value	HWE
	TT	TC	CC	T	C		
Patients	56	130	80	45,5	54,5	0,81	+
Controls	89	115	46	58,6	41,4	0.138	+
Total	145	245	126	51,84	48,16	0,06	+

The genotype distribution of SNP rs2287622 in controls group were in accordance with HWE (p>0.05) (Table 3.11). The frequency of allele C in the patient group (54.5%) was much higher than the frequency of allele C in the control group (33%).

Three models (additive, dominant, recessive) were tested for association of the variant p.Val444Ala (c.1331T>C, rs2287622) with susceptibility to the disease. The odds ratio (OR) and 95% confidence interval (CI) were calculated to detect the risk ratio between patients and controls (Table 3.12)

Table 3.12: Association of variant c.1331T>C (rs2287622) in ABCB11 gene with biliary atresia

Genotype	Patients	Controls	OR	95% CI	Giá trị P
Additive					
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
Dominant					
TT	56 (21,05%)	89 (35,60%)	1		
TC + CC	210 (78,95%)	161 (64,40%)	13,13	8,44-20,39	0,0001
Recesssive					
TT + TC	186 (69,92%)	204 (81,6%)	1		
CC	80 (30,08%)	46 (18,40%)	1,91	1,26-2,88	0,002
Allele					
T	242 (45,49%)	293 (58,60%)	1		
C	290 (54,51%)	207 (41,40%)	1,69	1,33-2,17	0,0001

OR: Odds ratio; 95%CI: 95% confidence interval of odds ratio

A significant difference of genotypes was obtained in all of three models ($p<0.01$). The frequency of CC genotype differed significantly between the two groups and was associated with an increased risk of biliary atresia (OR = 5,14, 95% CI: 2,82 – 9,39, $p<0,01$). The genotype TC also increases the risk of biliary atresia, but less than the CC genotype. (OR = 2,88, 95% CI: 1,81 - 4,59, $p<0,01$).

In the recessive model, the frequency of the homozygote genotype CC in biliary atresia patients was significantly higher than that in TT+TC group (OR = 2,79, 95% CI: 1,63 – 4,79, p<0,01). The frequency of C allele in biliary atresia group was significantly higher than that in control group (OR = 2,47, 95% CI: 1,84 – 3,32, p<0,01) implying a correlation between this risk allele and disease susceptibility.

3.6.2. The association between SNP rs927344 with the risk of biliary atresia

In this study, genotype of rs927344 in *ABCC2* gene were performed using ARMS-PCR. Chi-square test was used to test whether allele distribution of the SNP follows Hardy-Weinberg Equilibrium (HWE) (Table 3.13).

Table 3.13: Genotype and allele frequency ABCC2 rs927344

	Genotype			Allele frequency		HWE p value	HWE
	AA	AT	TT	A	T		
Patients	5	40	221	0,09	0,91	0.451	+
Controls	3	39	208	0,12	0,88	0,587	+
Total	8	79	429	0,10	0,90	0,056	+

Chi-squared test was used to test three models (additive, dominant, recessive) for associations of the SNP with risk of biliary atresia (Table 3.14). The result showed that no significant difference of allele frequencies between biliary patients and controls was detected for rs927344.

Table 3.14: Association of variant ABCC2 rs927344 with biliary atresia

Genotype	Patients	Control	OR	95% CI	Giá trị P
Additive					

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

The result showed that there was no association of *rs927344* with the risk of biliary atresia in Vietnamese population.

3.6.3. The association between SNP rs1815930 with the risk of biliary atresia

Statistics of rs1815930 was showed in table 3.15.

Table 3.15: Genotype and allele frequency MYO5B rs1815930

	Genotype			**Allele frequency**		**HWE p value**	**HWE**
	AA	**AG**	**GG**	**A**	**G**		
Patients	2	39	225	0,08	0,92	0.829	+

Controls	4	33	213	0,08	0,92	0,051	+
Total	6	72	438	0,08	0,92	0,128	+

Chi-squared test was used to test three models (additive, dominant, recessive) for associations of the SNP with risk of biliary atresia (Table 3.16).

Table 3.16: Association of variant MYO5B rs1815930 with biliary atresia

Kiểu gen	Nhóm bệnh	Nhóm đối chứng	OR	95% CI	Giá trị P
Additive					
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
Dominant					
AA	2 (0,75%)	4 (1,60%)	1		
AG + GG	264 (99,25%)	246 (98,40%)	2,15	0,39 - 11,82	0,38
Recesssive					
AA + AG	41 (15,41%)	37 (14,80%)	1		
GG	225 (84,59%)	213 (85.20%)	0,95	0,59 - 1,54	0,85
Allele					
A	43 (8,08%)	41 (8,20%)	1		
G	489 (91,92%)	259 (91,80%)	1,02	0,65 - 1,59	0,95

The result showed that p value higher than 0.05 in all of three model. Therefore there was no association of rs1815930 with the risk of biliary atrsia in Vietnamese population

Chapter 4: DISCUSSION

4.1. The role of genetic mutation in the biliary atresia

For the 05 patients were selected for WES in this study, genes that may be related to bile duct developmental disorders such as *CFTR, JAG1, ZIC3, CFC1, INV, MIF, VEGF, IFN-γ* and genes closely related to congenital biliary atresia such as *GPC1, ICAM1, ITGB2, NOTCH1, NOTCH2, NOTCH3, ZIC3, FOXA2, PKD1L1, ADD3, XPNPEP1* were either not identified mutation or carried mutations that were identified as benign on the ClinVar database.

Patients BA01 and BA02 carried a double heterozygous mutation c.412G>A (p.Gly138Arg) in the FAH gene and c.2225A>G (p.Tyr742Cys) in the ERCC4 gene. The FAH gene is located on chromosome 15, contains 14 exons and encodes 420 amino acids. Analysis using the STRING v.11 protein interaction network database showed that the protein encoded by the FAH gene strongly interacted with proteins encoded by the GSTZ1 (score: 0.986), HPD (score: 0.92) and FAHD1 (0.980) genes (Figure 4.1). These are genes that play an important role in the tyrosine metabolic pathway. Improper degradation of tyrosine leads to abnormal accumulation of tyrosine and its metabolites in the liver and kidneys, potentially leading to hepatobiliary syndromes such as cirrhosis or hepatocellular carcinoma.

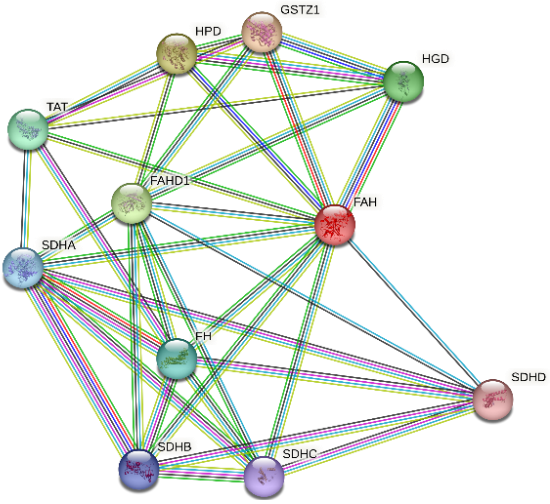

Figure 4.1: Interactions between FAH and other proteins

The HGMD Human Gene Mutation Database recorded 69 mutations in the ERCC4 gene, mainly missense mutations, nonsense mutations, and mRNA splice site mutations. Mutations in the ERCC4 gene in patients with congenital biliary atresia were also discovered by Sangkhathat and co-authors when using whole-genome sequencing to search for gene mutations in 20 patients diagnosed with biliary atresia in Thailand. Although this result is controversial, this study has opened a new approach to finding many other genes that may be involved in congenital biliary atresia.

Patient BA05 carries the c.50_51insG (p.Gly17Glyfs77∗) mutation in the *KRT18* gene. Keratins are a large family of intermediate filament proteins that are important components of the cytoskeleton. In mammals, keratin protein synthesis is a prominent feature of epithelial cells, including hepatocytes and cholangiocytes. The biliary system consists of a network of ducts in the liver that transport bile secreted by hepatocytes to large extrahepatic ducts adjacent to the gallbladder and intestine. In 2020, Babu and co-authors studied gene expression levels and beta-amyloid

accumulation around the bile ducts in patients with congenital biliary atresia. The results of the study showed that the keratin protein family, including *KRT8, KRT18,* and *KRT19* had increased expression in patients with congenital biliary atresia. This suggests an influence of the KRT18 gene on the pathogenesis of congenital biliary atresia.

4.2. The association between single nucleotide polymorphisms with the risk of biliary atresia

Many studies on the influence of rs2287622 nucleotide polymorphism in the ABCB11 gene on the risk of hepatobiliary syndromes and intrahepatic cholestasis have been conducted on many different groups of subjects. In a study on Han Chinese patients in Hunan with hepatitis C, Lei and co-authors found that the CC and TC genotypes of the rs2287622 single nucleotide polymorphism were associated with a higher risk of hepatitis C than the TT genotype. In a study on rs2287622 polymorphism in a group of Egyptian hepatitis C patients, Besheer and co-authors found an association between the CC genotype and the C allele with high bile acid concentrations in these patients. Accordingly, the homozygous genotype CC and allele C were higher in the group of patients with cirrhosis than in the group without cirrhosis, and the genotype CC and allele C were also more common in the group of patients with advanced fibrosis than in patients with early fibrosis. Furthermore, the observed bile acid concentrations in patients with chronic hepatitis with the CC genotype were much higher than those with the TC or TT genotypes. The survey of the single nucleotide polymorphism rs2287622 in the ABCB11 gene in 266 patients and 250 controls showed that the CC and TC genotypes were significantly different between patients with congenital biliary atresia and healthy individuals ($P < 0.01$), and the C allele was associated with an increased risk of the disease (OR = 2.47; 95%CI: 1.84–3.32; $P < 0.01$). The TC + CC genotype frequency

of the p.Val444Ala variant (c.1331T>C, rs2287622) in the ABCB11 gene was significantly different in Vietnamese patients with congenital biliary atresia. This variant may be closely related to cholestasis and lead to severe biliary tract damage causing biliary obstruction in patients.

In addition to studies on mutations, many scientists are also interested in evaluating the relationship between single nucleotide polymorphisms and disease risk. *ABCC2* and *MYO5B* are genes that have been studied by a group of Thai scientists and found to have variants related to congenital biliary atresia. Variants in the *ABCC2* and *MYO5B* genes have the common feature of leading to cholestasis and increased bilirubin. The *ABCC2* gene encodes the protein MRP2 (multidrug resistance protein 2). Variants in the *ABCC2* gene disrupt MRP2 function, leading to increased bilirubin in the blood. At the same time, the risk of cholestasis during pregnancy is associated with single nucleotide polymorphisms of MRP2. Variants in the *ABCC2* gene are thought to be associated with cholestasis in patients with Dubin-Johnson syndrome. In this study, we found no association between the single nucleotide polymorphism rs927344 in the *ABCC2* gene and the risk of congenital biliary atresia in Vietnamese patients.

Variations in the *MYO5B* gene have been identified in some patients with progressive familial intrahepatic cholestasis (PFIC). In addition, variations in the *MYO5B* gene have also been associated with biliary tract anomalies leading to clinical symptoms of congenital biliary atresia such as jaundice and intrahepatic cholestasis. *MYO5B* mutations have been reported to affect hepatobiliary function, increase serum bile acid levels, and lead to cholestasis in patients with congenital biliary atresia. In this study, we did not find an association between the single nucleotide polymorphism rs1815930 in the *MYO5B* gene and the risk of congenital biliary atresia in Vietnamese patients.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

1. This study has identified a novel double heterozygous novel mutation c.412G>A (p.Gly138Arg) in the *FAH* gene and c.2225A>G (p.Tyr742Cys) in the *ERCC4* gene of two siblings..

2. This study has identified a homozygous frameshift mutation c.50_51insG (p.Gly17Glyfs77*) in the *KRT18* gene, which involves biliary atresia, and a compound heterozygous mutation c.314C>A (p.Ser105∗), c.2975C>T (p.Pro992Leu) in the *ATP7B* gene in a patient with biliary atresia combined with Wilson disease.

3. The association between single nucleotide polymorphisms rs2287622 in the *ABCB11* gene, rs927344 in the *ABCC2* gene, and rs1815930 in the *MYO5B* gene with the risk of biliary atresia has been evaluated. Trong đó:

- SNP rs2287622 in *ABCB11* gene were determined related to biliary atresia in the Vietnamese population. TC and CC genotypes were significantly different between biliary atresia patients and healthy people, and the C allele was associated with an increased risk of biliary atresia.

- There was no significant difference in allele frequencies between biliary atresia patients and controls detected for rs927344 in the ABCC2 gene and rs1815930 in the MYO5B gene.

Recommendations

Although this study has detected some mutations with potential to cause disease, further research is needed on the effects of these mutations on protein function and clarifying the pathogenesis of the biliary atresia mechanism. And other single nucleotide polymorphisms can be studied to determine the risk of biliary atresia in Vietnamese people.

LIST OF THE PUBLICATIONS RELATED TO THE DISSERTATION

1. **Nguyen Van Tung**, Nguyen Thi Kim Lien, Nguyen Thi Phuong Mai, Nguyen Pham Anh Hoa, Nguyen Huy Hoang. Study on the association *ABCC2* rs927344 and *MYO5B* rs1815930 with congenital biliary atresia in Vietnamese population. *National Conference on Biotechnology 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