A New Homozygous ABCB4 Mutation Identified in Two Chinese Siblings Based on Exome Sequencing

Hui Lin1, Rong Fu2, Xiumin Zhang1, Bihui Yao3, Jing Ye1, Jianguo Shi1, Wen Sui4 and Zengshan Li1*

1Department of Pathology, Xijing Hospital and Basic Medical College, Fourth Military Medical University, Xi’an 710032, Shaanxi, P.R. China
2Department of Pediatrics, Xijing Hospital, Fourth Military Medical University, Xi’an 710032, Shaanxi, P.R. China
3Department of Surgery, Baotou Central Hospital, Baotou 014040, Inner Mongolia, P.R. China
4Suzhou University Children’s Hospital, Suzhou 215003, Jiangsu, P.R. China

Abstract

Background: Many pediatric cholestatic liver diseases show unspecific symptoms, laboratory tests and histological features. A complete genetic background check is necessary to identify the genetic determinants as well as the diagnosis and differential diagnosis.

Method: The whole genome exome sequencing was performed in two Chinese sister siblings with unspecific cholestasis and the data was analyzed.

Results: A homozygous mutation, c.2176C-T transition (p.P726L) in exon 17 of ABCB4, was identified. Further study indicated that there are mutant alleles in the genome of their non consanguineous parents. The final diagnosis is progressive familial intrahepatic cholestasis (PFIC), type 3.

Conclusion: As a rare hereditary cholestatic liver disease, PFIC shares many similar features with other hereditary or acquired liver disease and requires a wide range of differential diagnosis.

Exome sequencing is a useful tool for mapping this kind of monogenic disease mutation and plays a very important role in genetic counseling and prenatal diagnosis.

Keywords: ABCB4; Mutation; Progressive familial intrahepatic cholestasis; Whole-exome sequencing

Abbreviations: PFIC- Progressive Familial Intrahepatic Cholestasis; MDR3 - Multidrug Resistance P-Glycoprotein 3; PC- Phosphatidylcholine; TM- Trans Membrane Segments

Introduction

Cholestasis is a common manifestation of hepatic disease in pediatric patients and a wide range of differential diagnosis, especially inherited diseases, need to be considered. Moreover, many of these entities have overlapped symptoms, histological findings as well as laboratory features. Under this kind of situations, genetic background assay is very important for a precise final diagnosis. There are different ways in which DNA sequencing can be undertaken, such as sanger sequencing, resequencing chip and denaturing high-performance liquid chromatography [1,2]. Anyway, the scope of these methods is limited sometimes, especially in detecting the unidentified genes. Whole-exome sequencing, as a new technique, is considered to be unbiased — it does not focus on specific genes or loci deemed of interest on the basis of assumptions about possible causal associations. Recent studies have shown that exome sequencing is a powerful, efficient and cost-effective strategy to identify the responsible genes for monogenic diseases such as Miller syndrome and familial combined hypolipidemia [3,4]. In this study we utilized exome sequencing in blood samples of two Chinese siblings who presented with cholestasis phenotypes to identify the underlying genetic defects responsible for the disease and confirm the clinical diagnosis. Finally a diagnosis of family progressive familial intrahepatic cholestasis was made and a new homozygous ABCB4 mutation was identified as well.

Progressive Familial Intrahepatic Cholestasis (PFIC) refers to a heterogeneous group of autosomal recessive liver disorders of childhood in which cholestasis of hepatocellular origin often presents in the neonatal period or first year of life and leads to death from liver failure at ages usually ranging from infancy to adolescence [5]. The estimated incidence of PFIC is about one per 50000-100000 births [6]. PFIC can be classified into three types with slightly different clinical, biochemical and histological features, associated with mutations in three distinct hepatobiliary transport proteins encoded by related ATP-Binding Cassette (ABC) genes, ATP8B1 (PFIC1), ABCB11 (PFIC2) and ABCB4 (PFIC3) [7]. It’s also possible that other unidentified genes involved in bile formation may be responsible for the PFIC phenotypes because mutations have not yet been identified in very rare PFIC patients [6]. Therefore molecular analysis is often recommended to confirm the diagnosis and typing of PFIC in affected patients. Histologically, PFIC exhibits extensive bile ductular proliferation, intranuclear cholestasis, portal inflammation, and fibrosis. All these findings are nonspecific and differential diagnosis should be considered in any neonate or infant with progressive symptomatic cholestasis.

Material and Methods

Patient samples

In this study, two Chinese siblings presented with pruritus, jaundice and hepatomegaly in 9 and 15 months old separately. They...
both have similar laboratory findings, including elevated alanine aminotransferase, aspartate aminotransferase, bile salt, alkaline phosphatase and gamma-glutamyl transpeptidase. The older sister once received mother-to-daughter living donor heterotopic auxiliary part liver transplantation at the age of 2. The postoperative liver biopsy showed extensive intracanalicular cholestasis, bile ductular proliferation, portal inflammation and fibrosis. All these findings are nonspecific and the final diagnosis is unclear. Six years later her younger sibling sister was born and presented with similar symptoms in 9 months old. Their parents, non-consanguineous and healthy, are both Chinese Hui ethnic minorities (Muslim ethnic group). For further diagnosis and classification, the blood samples of the siblings were collected for exome sequencing.

DNA preparation and exome capture sequencing

Genomic DNA was extracted from venous blood with a commercially available kit, Omega Blood DNA Midi Kit (D3494-01), using the manufacturer’s protocols. Qualified genomic DNA samples from each case were randomly fragmented by Covaris and enriched using Agilent’s Sure Select Biotinylated RNA Library (BAITS). The array covers 1.22% of human genomic regions corresponding to the NCBI consensus coding sequence database 2008. Each captured library was then loaded on HiSeq2000 platform, and sequencing was performed for each captured library independently to ensure each sample had at least 50-fold coverage. Raw image files were processed by Illumina Pipeline v1.6 for base-calling with default parameters and the coverage of target region was 98.3% 98.4%.

Bioinformatics analysis

For the basic analysis, low quality sequences in the raw Solexa reads were filtered out as follows: base quality is more than 20, depth is between 4 and 200, estimate copy number is equal or less than 2 and the distance between two SNPs must be longer than 4. Coverage and depth of each sample were also calculated.

For the advanced analysis, the sequencing reads were aligned to the Human reference genome (NCBI 36.3) using SOA Paligner (soap2.20) and the non-synonymous variants resulting in changes of proteins were extracted as candidate variants. Based on the hypothesis that the pathogenic mutation underlying the rare familial disease was not present in general population, these variants were then filtered against several public database, comprising dbSNP129, eight previously exome-sequenced Hap Map samples (Hap Map 8), the SNP release of the 1000 Genome Project (20100208 release) and Yan Huang (YH) 1 genome database, to remove the common variants. Then we choose the homozygous variants shared by both patients as the target variants. In addition, we used SIFT (Sorting Intolerant from Tolerant), available at http://sift.jcvi.org/, to predict that whether the target variants would affect protein functions. The only one non-synonymous and shared by two siblings is just the causative mutation in ABCB4 filtered out by another analysis strategy (Table 1).

Sanger sequencing

Sanger sequencing was used to screen all of the four family members. Exon 17 of ABCB4 gene, where the mutation located, was amplified using polymerase chain reaction (forward

Results

Exome sequencing and data analysis

For each of the two cases, we generated approximately 2.3 gigabases of effective sequence data allowing each targeted base to be read approximately 60 times. About 8000 synonymous variants were found in each case, of which approximately 7000 were missense mutations and approximately 40 were nonsense mutations (Table 1). We also detected more than 8000 insertions and deletions in each case and most of them were intragenic variants.

Given that non-synonymous variants were more likely to be the pathogenic mutations and unlikely to be present in general group, so we filtered them against several public database mentioned in methods and the number of the candidate variants was reduced to about 400. Considering that PFIC is an autosomal recessive inherited disorder and the same clinical phenotypes have been found in two siblings, so the target variants should be homozygous and shared by patients. There were only four variants meeting all the conditions. SIFT prediction showed that only one of the variants, the c.2176C-T transition (p.P726L) of ABCB4, occurred in a highly conserved position and was more likely to be functionally damaging (Table 1). The same mutation has been revealed compound heterozygous in the study of another group [8]. So we believed that this variant would be the very causation responsible for the two patients’ sufferings. Otherwise, we screened those related genes mentioned previously and most of the variants were unlikely to affect protein functions. The only one non-synonymous and shared by two siblings is just the causative mutation in ABCB4 filtered out by another analysis strategy (Table 1).

Sanger sequencing

Sanger sequencing was used to screen all of the four family members. Exon 17 of ABCB4 gene, where the mutation located, was amplified using polymerase chain reaction (forward

Table 1: Results of exome sequencing and details of data analysis.

<table>
<thead>
<tr>
<th>Case</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total effective yield(Mb)</td>
<td>3,801.77</td>
<td>3,935.91</td>
</tr>
<tr>
<td>Effective sequence on target(Mb)</td>
<td>2,290.43</td>
<td>2,296.60</td>
</tr>
<tr>
<td>Average sequencing depth on target</td>
<td>60.71</td>
<td>60.87</td>
</tr>
<tr>
<td>Coverage of target region</td>
<td>98.3%</td>
<td>98.4%</td>
</tr>
<tr>
<td>covered with at least 10X</td>
<td>89.1%</td>
<td>89.2%</td>
</tr>
<tr>
<td>covered with at least 20X</td>
<td>79.0%</td>
<td>79.2%</td>
</tr>
<tr>
<td>Total number of variants*</td>
<td>32,432</td>
<td>32,349</td>
</tr>
<tr>
<td>No. Non-synonymous variants</td>
<td>7432</td>
<td>7468</td>
</tr>
<tr>
<td>No. not in dbSNP</td>
<td>915</td>
<td>951</td>
</tr>
<tr>
<td>No. not in 1000 Genomes Project</td>
<td>572</td>
<td>579</td>
</tr>
<tr>
<td>No. not in HapMap 8</td>
<td>472</td>
<td>468</td>
</tr>
<tr>
<td>No. not in YH1 genome data</td>
<td>468</td>
<td>461</td>
</tr>
<tr>
<td>No. shared by 2 cases</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>No. homozygous</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SIFT prediction</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>No. Tolerated</td>
<td>No. Damaging</td>
<td></td>
</tr>
</tbody>
</table>

Coverage: the average number of times each base is represented in the sequence reads; dbSNP: the Single Nucleotide Polymorphism Database (build 129) of the National Center for Biotechnology Information; HapMap 8: eight previously exome-sequenced Hap Map samples; 1000 Genomes Project: the SNP release of the 1000 Genome Project (20100208 release); YH1 genome data; Yan Huang 1 genome database |

*Variants were filtered gradually according the criteria.

Table 1: Results of exome sequencing and details of data analysis.

Ethics approval

Xijing Hospital Human Research Ethics Committee

primer CGAACAACCCCATACTCGTATG, reverse primer GAGGTGGAAGAAGCAGCAGG)[9]. Sequence analysis showed that the mutation was homozygous in the patients and heterozygous in their parents (Figure 1). The variant has never been described in the Genbank database, so the sequence was submitted to Genbank (accession number HQ540315) and the variation was submitted to dbSNP (submitter snp number 263198237).

Discussion

In this study, we have identified a homozygous ABCB4 gene mutation using exome sequencing and finally confirmed the diagnosis as PFIC3. ABCB4 gene encodes the class III multidrug resistance P-glycoprotein 3 (MDR3) belonging to the family of ATP-binding cassette transporters. It is localized at the canalicular membrane of hepatocytes translocating Phosphatidyl Choline (PC) from the inner to the outer membrane leaflet [10]. MDR3 has 12 transmembrane segments (TM1–TM12) organized in two transmembrane domains. The mutation p.P726L is located in the TM7 which is conserved in mammals ABCB4. Mutation detected in our patients is categorized as damaging by SIFT (Score 0.01) and seems to induce a significant impairment in flippase activity. Site-directed mutagenetic analysis of P-glycoprotein has shown that mutations in transmembrane domains are important for substrate specificity [9]. Otherwise, many missense mutations may influence protein processing, causing the abnormal but potentially functional protein to be misfolded, trapped in the endoplasmic reticulum and subsequently degraded [11]. Defects of MDR3 may further induce an imbalance in the composition of primary bile, with lack of PC and a surplus of bile salts, the latter presumed to damage the biliary epithelium and cause progressive liver disease such as PFIC3 [12]. Previous study have shown that the phenotypic spectrum of PFIC3 ranges from neonatal cholestasis to cirrhosis in young adults [6]. The two siblings in our study reveal the similar clinical course characterized by progressive cholestasis and liver dysfunction.

Further sequencing identified that the mutation was inherited from the non-consanguineous parents. Notably, both parents are Hui Chinese and absent of symptoms. Lang et al. [13] have intended to establish genetic variability and haplotype structures of ABCB4 in healthy populations of different ethnic backgrounds comprising Caucasian, African-American, Japanese, and Korean origin. Their results showed that most ABCB4 variants were population-specific, and either their occurrence or their population frequency varied among different ethnic groups. The same mutation mentioned above was detected in a French patient and compound with another heterozygous mutation. Interestingly, Degiorgio et al. [9] have ever described a genotype in an Italian patient, showing three heterozygous mutations, one of which was located on the maternal allele and similar to our finding (c.2176C-A transition, p.P726T). It seems that mutations at this site may be more important than previous perspectives. Theses suggest that it would make sense to investigate the frequencies of the rare variants in specific groups, however there is still short of related epidemiologic data.

Another important task is to make genetic counseling and prenatal diagnosis to some extent. PFIC transmission is autosomal recessive and mutations on both alleles are characterized in most cases. This understanding has already allowed prenatal diagnosis. Camille et al. have shown that it is possible to perform reliable molecular prenatal diagnosis for PFIC [14]. However, objects of genetic counseling and antenatal diagnosis are usually limited in individuals with a family history or from close family. The presence of mutations in ABCB4 gene had never been described in siblings of non-consanguineous parents until Raj et al. reported it in two Sicilian siblings [15]. Therefore, it’s really difficult to discover individuals with genetic susceptibility using traditional methods in the situation just as we have encountered. For this reason, exome sequencing, enabling low-cost, rapid, broad sequencing, may be a potentially valuable tool in genetic counseling and prenatal diagnosis. Sequencing data reflect most of the genetic background including some possible genetic defects and has the potential to broaden our knowledge of the genetic basis of diseases. By combining with different methods or using different analysis strategies, exome sequencing is apt to be as promising a route to mapping genes for diseases comprising not only monogenic disease but also more complex diseases.[4,16,17] In our study, a total of 250 novel variants are both carried by the two siblings and some of the heterozygous variants, categorized as damaging by SIFT, may bring potential risk to next generations. Unfortunately, there are few related publications in this area.

![Figure 1: Pedigree and mutation analysis of ABCB4 in family. Squares indicate male family members, and circles female family members. This is a two-generation family with 4 members. The parents are nonconsanguineous and healthy with a heterozygous variant in ABCB4, while the two siblings are suffering PFIC3 with a homozygous variant. The variant is a missense mutation, resulting in a proline- to- leucine substitution.](image-url)
Conclusions

Briefly, we found a novel homozygous mutation in ABCB4 gene causing deleterious results. Further functional studies of this specific gene mutation are needed to gain a clear idea of the pathogenic mechanism. We also demonstrate that exome sequencing is a useful means for mapping monogenic disease mutation. In addition, we may investigate the genetic background of certain groups and predict risk of suffering diseases by applying it to genetic counseling and prenatal diagnosis. Mutation-specific therapy may be the most promising approach in the treatment of genetic diseases such as PFIC.

Authors’ Contributions

H. L. and R. F. contributed equally to this work as co-first author by providing data collection and data analysis. X. Z., B. Y., J. Y., J. S. provided valuable opinions about the manuscript. W. S. and Z. L. conceived the idea and wrote the manuscript as co-corresponding author.

Grant Sponsor

The National High Technology Research and Development Program of China (2007AA02Z470).

References