Implementing targeted region capture sequencing for the clinical detection of Alagille syndrome: An efficient and cost-effective method

TIANHONG HUANG1,2*, GUILIN YANG3*, XIAO DANG1,2, FEIJIAN AO3, JIANKANG LI2, YIZHOU HE3, QIYUAN TANG3 and QING HE3

1BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, Guangdong 518083; 2BGI-Shenzhen, Shenzhen, Guangdong 518083; 3Liver Disease Division, The Third People’s Hospital of Shenzhen, Shenzhen, Guangdong 518112, P.R. China

Received February 17, 2016; Accepted February 21, 2017

DOI: 10.3892/mmr.2017.7429

Abstract. Alagille syndrome (AGS) is a highly variable, autosomal dominant disease that affects multiple structures including the liver, heart, eyes, bones and face. Targeted region capture sequencing focuses on a panel of known pathogenic genes and provides a rapid, cost-effective and accurate method for molecular diagnosis. In a Chinese family, this method was used on the proband and Sanger sequencing was applied to validate the candidate mutation. A de novo heterozygous mutation (c.3254_3255insT p.Leu1085PhefsX24) of the jagged 1 gene was identified as the potential disease-causing gene mutation. In conclusion, the present study suggested that targeted region capture sequencing is an efficient, reliable and accurate approach for the clinical diagnosis of AGS. Furthermore, these results expand on the understanding of the pathogenesis of AGS.

Introduction

Alagille syndrome (AGS), additionally known as arteriohepatic dysplasia, syndromic bile duct paucity and Alagille-Watson syndrome, is a highly variable autosomal dominant disease which affects the liver, heart and other parts of the body (1,2). AGS is one of the familial intrahepatic cholestatic syndromes associated with jagged 1 (JAG1), neurogenic locus notch homolog protein 2 (NOTCH2), uridine phosphorylase glucuronosyltransferase family 1 member A1 (UGT1A1), adenosine triphosphatase phospholipid transporting 8B1 (ATP8B1), adenosine triphosphate-binding cassette, sub-family B member 11 (ABCB11) and adenosine triphosphate-binding cassette, sub-family B member 4 (ABCB4) genes, and is one of the important causes of intrahepatic cholestasis in infancy (3-5). The incidence of AGS is ~1 in 70,000 newborns with neonatal jaundice (6), and the mortality rate is ~17% (7). AGS is the result of JAG1 mutations; however, the inherited pathogenic variants only account for 30-50% of all cases, with the remaining 50-70% of cases being de novo pathogenic variants (8-10). Traditionally, a diagnosis of AGS requires three of the five primary clinical features, or two features if the patient has a positive family history. The five primary clinical features are: Chronic cholestasis, cardiac disease, skeletal abnormalities, ocular abnormalities and a characteristic facial phenotype (11). However, following the demonstration that mutations in the JAG1 gene cause AGS (11,12), the diagnostic criteria has been modified; if an individual carries a harmful mutation in the JAG1 gene, AGS may be diagnosed even in the absence of clinical manifestations.

JAG1 and NOTCH2 genes are associated with AGS. Pathogenic mutations in JAG1 and NOTCH2 genes may impair the Notch signaling pathway, which is an evolutionarily conserved, intercellular signaling mechanism essential for healthy embryonic development (13-16). The majority of AGS cases (~90%) are caused by a detrimental mutation in the JAG1 gene (17). The JAG1 gene encodes a ligand of the notch receptor which is a key signaling molecule on the cell surface, and a component of the highly conserved Notch signaling pathway (14). Mutations located in almost all regions of the 26 exons of the JAG1 gene have been identified, and the majority are pathogenic variants. It has been reported that the pathogenic mutations of JAG1 include missense mutations (11%), nonsense and frame-shift mutations (69%), pathogenic splice site mutations (16%), and deletion of the entire JAG1 gene (4%) (8,9,11,12,18-26). A few reported cases (<1%) have been caused by a mutation in the NOTCH2 gene (17,27). The NOTCH2 gene encodes a member of the Notch family, and the Notch receptors (Notch1, 2, 3 and 4 in humans) share structural characteristics including an intracellular domain with...
seven ankyrin (ANK) repeats and an extracellular domain composed of multiple epidermal growth factor-like (EGF) repeats. To date, ten pathogenic variants associated with AGS in the NOTCH2 gene have been identified. Of these, seven are localized in the EGF-like repeats and ANK repeats domains in NOTCH2 protein (27). The remaining three pathogenic variants are splice site mutations, frame-shift variants and nonsense variants (28,29).

Materials and methods

Ethics statement. The present study was approved by the Institutional Review Board on Bioethics and Biosafety of The Third People's Hospital of Shenzhen (Shenzhen, China). Written informed consent was obtained from all participants. All of samples were obtained and analyzed according to appropriate ethical approvals.

Pedigree and subjects. Samples were obtained from two generations of a Chinese family. The proband (II:1; Fig. 1) was an 11-year-old boy diagnosed with chronic intrahepatic cholestasis resulting from AGS, who had a 9-year history of pruritus and a 7-year history of jaundice. He was the first full-term child of this family, and in 2005 presented with cutaneous pruritus without any clear cause, particularly on the lower limbs, accompanied with yellow urine. In 2007, the patient experienced loss of strength and appetite, jaundice, and occasional vomiting. In March 2015, liver function tests revealed that levels of total bilirubin (TBIL, 30.9 µmol/l), direct bilirubin (DBIL, 15.9 µmol/l), total bile acid (TBA, 49.6 µmol/l), alanine transaminase (ALT, 204 U/l), aspartate aminotransferase (AST, 128 U/l), γ-glutamyl transferase (γ-GGT, 517 U/l), total cholesterol (TC, 10.77 mmol/l) and free triiodothyronine (FT3, 3.10-6.80 pmol/l) were compared with healthy levels. Normal ranges are as follows: TBIL, 5.1-19.0 µmol/l; DBIL, 0.1-5.1 µmol/l; TBA 0-10 µmol/l; ALT, 0-40 U/l; AST, 0-40 U/l; γ-GGT, 7-32 U/l; TC, 3.35-6.45 mmol/l; FT3, 3.10-6.80 pmol/l. The right lobe of his liver was biopsied (length, 14 mm; diameter, 1 mm) and was observed to have minor lesions in the incomplete portal area (Fig. 2), suggesting that the proband had hepatobiliary hypoplasia. A B-scan ultrasound revealed that the liver, bladder, spleen and pancreas were healthy. The patient did not exhibit any other clear symptoms.

Preparation of DNA. Peripheral blood was obtained from the proband and his healthy parents. Genomic DNA was extracted from the blood using a blood DNA extraction kit (QIAamp DNA Blood Midi kit; Qiagen GmbH, Hilden, Germany) according to the manufacturer's protocol.

Targeted region capture sequencing. To construct the capture library, DNA was sheared randomly by sonication using an LE220 Focused-Ultrasonicator (Covaris, Inc.; Woburn, MA, USA); fragments of 200 to 250 bps were retained and subsequently purified by Ampure Beads (Beckman Coulter, Inc., Brea, CA, USA) according to the manufacturer's protocol. Following this, the two ends of the purified fragments were repaired, bound to A base and ligated with adapters. DNA fragments were subsequently amplified by ligation-mediated polymerase chain reaction (LM-PCR) using Platinum PfX DNA polymerase (Invitrogen; Thermo Fisher Scientific, Inc., Waltham, MA, USA) according to the manufacturer's protocol. Cycling conditions were as follows: An initial pre-denaturation step at 95°C for 5 min, followed by 30 cycles of denaturation at 95°C for 30 sec, annealing 58°C for 30 sec, extension at 72°C for 30 sec, and a final extension step at 72°C for 7 min. PCR products were purified and hybridized to a customized gene-trapping chip (Roche Applied Science, Madison, WI, USA) for enrichment. Non-hybridized fragments were washed out using a NimbleGen Wash and Elution kit (Roche NimbleGen, Inc., Madison, WI, USA) as per the manufacturer's protocol. A 2100 Bioanalyzer system (Agilent Technologies, Inc., Santa Clara, CA, USA) and an ABI StepOne Real-Time PCR system (Applied Biosystems; Thermo Fisher Scientific, Inc.) were used to estimate the magnitude of the enrichment, and the size and the concentration of the fragments in the two libraries. Subsequently, the library enriched for target regions was analyzed using a HiSeq 2500 Sequencing system (Illumina, Inc., San Diego, CA, USA) by paired-end sequencing to obtain raw reads, which were subsequently analyzed using Illumina Genome Analyzer Pipeline software version 1.3.4 (Illumina, Inc.).

Read mapping and variant analysis. The first step of the analysis was to assess the quality of the raw data and to remove contaminated or low quality reads. The reads were mapped to the reference hg19 (build 37.1) using Burrows Wheeler Aligner software version 0.7.12 (sourceforge.net/projects/bio-bwa/files) and estimated the effect capturing at the same time. Subsequently, Short Oligonucleotide Analysis Package single nucleotide polymorphism (SNP) software version 1.03 (soap.genomics.org.cn/soapsnp.html) and SAMtools software version 1.2 (samtools.sourceforge.net) were used to detect single nucleotide variants, and insertions and deletions (indels), respectively. The last step was to annotate and screen the suspected mutations. The coverage of targeted region and sequencing depth calculation were based on all mapped reads. All alterations were compared and filtered against exome data from the SNP database (dbSNP build 137; www.ncbi.nlm.nih.gov/snp), the 1000 human genome dataset (www.internationalgenome.org/), HapMap (www.ncbi.nlm.nih.gov/probe/docs/projhapmap/) and an in-house database of 100 Chinese healthy adults. Sorting Intolerant From Tolerant software version 1.03 (sift.jcvi.org) was used to predict if amino acid substitutions, insertions or deletions affected protein function (30).

Sanger sequencing for validation. The samples of the three individuals in this pedigree were used for validation. The primers were designed using Primer 3 software version 4.0.0 (primer3.sourceforge.net) and Sanger sequencing. PCR amplification and Sanger sequencing were conducted to validate the candidate mutation (human JAG1; locus, NM_000214; c.3254_3255insT) using standard protocols on an ABI 3730XL sequencer (Applied Biosystems). The forward primer for JAG1 was 5'-TTGGTGGTGTTGGTCCTCAGA-3' and the reverse primer was 5'-AGGGATAAAGGGCAGGAGAA-3' (product size, 244 bp). PCR amplification was performed using Ex Taq™ DNA polymerase (Takara Bio, Inc., Otsu, Japan) with an HUANG et al: TARGETED REGION CAPTURE SEQUENCING FOR THE CLINICAL DETECTION OF AGS 6877
Initial predenaturation step at 95˚C for 5 min, followed by 33 cycles of denaturation at 95˚C for 30 sec, annealing at 58˚C for 30 sec, extension at 72˚C for 30 sec and a final extension at 72˚C for 7 min. PCR products were pooled and subsequently purified with an AxyPrep DNA Gel Extraction kit (Axygen Biosciences, Inc., Union City, CA, USA) according to the manufacturer's protocol.

Results

Mutation screening. Targeted region capture sequencing of the proband in a Chinese family with chronic intrahepatic cholestasis linked with JAG1, NOTCH2, UGT1A1, ATP8B1, ABCB11 and ABCB4 genes was performed. The reads were aligned with the human genome reference. The average depth of the target region was 186.55-fold and the coverage of the target region was 99.67%. A total of 10,277 genetic variants were revealed, including 10,132 SNPs and 1435 indels. Of the variants, 758 missense, 3 nonsense, 1087 synonymous, 248 splice site, 7047 intron and 2188 3' untranslated region (UTR) mutations, 311 5'UTR and 8 frame-shift coding indels that were more likely to be pathogenic compared with the other variants. Priority was given to frame-shift, non-synonymous and splice site mutations. Using a preliminary screening process, ten candidate sites were revealed (Table I). Considering the frequency in the public and in-house databases, nine other candidate sites were excluded. A novel

<table>
<thead>
<tr>
<th>Gene</th>
<th>Transcript</th>
<th>Nucleotide change</th>
<th>AA change</th>
<th>Gene subregion</th>
<th>Location</th>
<th>Gene</th>
<th>Transcript</th>
<th>Nucleotide change</th>
<th>AA change</th>
<th>Gene subregion</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP8B1</td>
<td>NM_005603</td>
<td>c.3454G&gt;A</td>
<td>p.Ala1152Thr</td>
<td>EX27</td>
<td>Hom</td>
<td>Hom</td>
<td>chr18:55317676</td>
<td>rs222581</td>
<td>0.9991</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>ATP8B1</td>
<td>NM_005603</td>
<td>c.811A&gt;C</td>
<td>p.Arg271Arg</td>
<td>EX10</td>
<td>Hom</td>
<td>Hom</td>
<td>chr18:55364682</td>
<td>rs319443</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>ATP8B1</td>
<td>NM_005603</td>
<td>c.696T&gt;C</td>
<td>p.Asp232Asp</td>
<td>EX8</td>
<td>Hom</td>
<td>Hom</td>
<td>chr18:55368982</td>
<td>rs6668192</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>NOTCH2</td>
<td>NM_024408</td>
<td>c.7341T&gt;A</td>
<td>p.Gly2447Gly</td>
<td>EX34</td>
<td>Hom</td>
<td>Hom</td>
<td>chr1:120458004</td>
<td>rs6685892</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>JAG1</td>
<td>NM_000214</td>
<td>c.765C&gt;T</td>
<td>p.Tyr255Tyr</td>
<td>EX6</td>
<td>Hom</td>
<td>Hom</td>
<td>chr20:10633237</td>
<td>rs1131695</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>JAG1</td>
<td>NM_000214</td>
<td>c.588C&gt;T</td>
<td>p.Cys196Cys</td>
<td>EX4</td>
<td>Hom</td>
<td>Hom</td>
<td>chr20:10639222</td>
<td>rs1801138</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>ABCB11</td>
<td>NM_003742</td>
<td>c.3084A&gt;G</td>
<td>p.Ala1028Ala</td>
<td>EX10</td>
<td>Het</td>
<td>Het</td>
<td>chr2:169780916</td>
<td>rs497692</td>
<td>0.9908</td>
<td>Het</td>
<td></td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_003742</td>
<td>c.1331T&gt;C</td>
<td>p.Val444Ala</td>
<td>EX13</td>
<td>Hom</td>
<td>Hom</td>
<td>chr2:169830328</td>
<td>rs2287622</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
<tr>
<td>ABCB11</td>
<td>NM_003742</td>
<td>c.504C&gt;T</td>
<td>p.Asn168Asn</td>
<td>EX6</td>
<td>Het</td>
<td>Het</td>
<td>chr2:169830328</td>
<td>rs2287622</td>
<td>0.9908</td>
<td>Het</td>
<td></td>
</tr>
<tr>
<td>ABCB4</td>
<td>NM_003742</td>
<td>c.3245T&gt;C</td>
<td>p.Ala1082Asp</td>
<td>EX24</td>
<td>Hom</td>
<td>Hom</td>
<td>chr2:169830328</td>
<td>rs2287622</td>
<td>0.9908</td>
<td>Hom</td>
<td></td>
</tr>
</tbody>
</table>

Public database, 1,000 genomes database, in-house database; database of genome information from >200 healthy individuals; A: frequency from 0 to 0.01; B, frequency from 0.01-0.05; C, frequency from 0.05-1; AA, amino acid; Hom, homozygous; Het, heterozygous; ATP8B1, adenosine triphosphatase transporting 8B1; JAG1, jagged 1; ABCB11, adenosine triphosphate-binding cassette, sub-family B member 11; ABCB4, adenosine triphosphate-binding cassette, sub-family B member 4.

Figure 1. Family pedigree of the proband with Alagille syndrome. Squares indicate males, circles indicate females, the black symbol indicates an affected member, white symbols indicate unaffected members and the arrow indicates the proband.

Figure 2. Pathological section of the right lobe of liver of the proband. Minor lesions were observed in the incomplete portal area biopsied, as demonstrated by an arrow.
insertion mutation, c.3254_3255insT (p.Leu1085PhefsX24), was identified in exon 26 of the JAG1 gene in the proband. The mutation in the JAG1 gene was selected for further validation, as it is known to be the disease-causing gene.

Sanger sequencing: The mutation was subsequently sequenced in three members of the family via Sanger sequencing. The results of Sanger sequencing demonstrated that the mutation was present in the proband but absent in unaffected members (Fig. 3). The mutation results in a premature stop codon and a truncated protein, which affects the Notch signaling pathway, influencing multi-system development in the embryonic period.

Discussion

Originally, the proband was diagnosed with chronic intrahepatic cholestasis, which is a syndrome characterized by jaundice, pruritus and hepatomegaly. His serum levels of bilirubin, alkaline phosphatase, cholesterol and other indicators were greater compared with normal values (31). The primary function of the liver is the formation and secretion of bile, which is necessary for lipid digestion. However, cholestasis is a condition where bile does not form or does not flow from the liver to duodenum, and may be caused by numerous factors. In the process of bile formation, environmental factors, including viral hepatitis and intrahepatic cholestasis of pregnancy, may lead to hepatocellular cholestasis. However, the present study primarily focused on genetic factors.

To investigate the underlying mechanisms of chronic intrahepatic cholestasis, chip capture high-throughput sequencing was performed on the DNA of the proband and Sanger sequencing was conducted to validate the candidate mutation in this family. A novel frame-shift variant was identified in the JAG1 gene, which is the primary pathogenic gene of AGS. AGS, linked with the JAG1 and NOTCH2 genes, is associated with dysfunction of the liver, heart, skeleton and eyes, and a characteristic facial appearance. The underlying pathogenic mechanisms remain unclear; however, these may be attributed to haploinsufficiency (32-37). A dominant-negative effect of putative truncated proteins is an additional potential pathogenic factor (34). The majority of mutations in the JAG1 gene lead to a truncated protein, impairing its ability to attach to the cell membrane, and disrupting the Notch signaling pathway, which is involved in embryonic development. JAG1 protein is the ligand of the Notch receptor, and binding triggers a cascade of proteolytic cleavage that releases the Notch intracellular domain (NICD) from the plasma membrane. The NICD localizes to the nucleus, where it forms a complex with the DNA binding protein C-promoter binding factor 1, suppressor of hairless, lin-12 and glp-1 (CSL). Components of the activated complex are recruited to NICD-CSL, which may lead to the transcriptional activation of Notch target genes (15,16). The Notch signaling pathway has been reported to be linked to AGS, and is important in cell fate determination in Drosophila melanogaster and Caenorhabditis elegans (13,38,39). It is highly conserved and essential for proper embryonic development in all metazoan organisms. Furthermore, the JAG1 protein is a key signaling molecule on the cell surface. The present study demonstrated that an insert mutation causes a premature termination codon that encodes a truncated protein. The truncated protein has a dominant negative effect in in vitro cell culture systems and in vivo transgenic Drosophila melanogaster (35,40,41). The truncated protein lacks the transmembrane region necessary for the protein product to embed in the cell membrane and contribute to signaling. The conserved region of the JAG1 protein includes a signal peptide (aa:1-21; cDNA:414-477), a

The mutation identified in the present study, c.3254_3255insT (p.Leu1085PhefsX24), is located in the TM region.

In the conclusion, the present study identified a de novo heterozygous mutation c.3254_3255insT (p.Leu1085PhefsX24) in exon 26 of the JAG1 gene, which may be responsible for AGS in the proband. In addition, these data expanded the genotypic spectrum of JAG1 mutations associated with AGS, and provided evidence that gene detection may be used for clinical diagnosis. Additionally, these results suggested that targeted region capture sequencing is a reliable, cost-effective and accurate clinical molecular diagnosis method and may improve the accuracy of clinical diagnosis for AGS and associated disorders.

Acknowledgements

The authors would like to thank the family for participating in the present study, and the BGI-Shenzhen and The Third People’s Hospital of Shenzhen.

References