Search for the potential ‘second-hit’ mechanism underlying the onset of familial hemophagocytic lymphohistocytosis type 2 (FHL2) by whole-exome sequencing analysis

**LILI GAO^{1*}, XIAO DANG^{2,3*}, LIANG HUANG¹, LI ZHU¹, MINGYAN FANG³,
JIANGUO ZHANG³, XUN XU³, LIJUN ZHU¹, TONGJUAN LI¹, LEI ZHAO¹, JIA
WEI^{1&}, JIANFENG ZHOU^{1&}**

¹Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, *P. R. China*; ²BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, *P. R. China*;
³BGI-SHENZHEN, Shenzhen, *P. R. China*

& Correspondence

Jianfeng Zhou and Jia Wei, Department of Hematology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1095 Jie-Fang Avenue, Wuhan, 430030, *P. R. China*. Tel: +86-27-83662437; Fax: +86-27-83662680;
E-mail: jfzhou@tjh.tjmu.edu.cn; jiawei@tjh.tjmu.edu.cn

*These authors contributed equally to this work.

ABSTRACT

Familial hemophagocytic lymphohistocytosis type 2 (FHL2), caused by PRF1, is a genetic disorder of lymphocyte cytotoxicity that usually presents in the first two years of life and has a poor prognosis. Late onset of FHL2 has been sporadically reported, and the mechanism is largely unknown. A newly diagnosed FHL2 patient was detected to have compound mutations in both PRF1 alleles and positive EBV infection. Her brother carried the same mutations and EBV infection status but kept healthy. To search for the potential unknown mechanisms, we performed whole-exome sequencing analysis. The patient and her asymptomatic brother carried the same heterozygous missense (c.916G>A) and frameshift mutation (c.65delC) in PRF1. Germline mutations analysis demonstrated that only the proband was exclusively detected with a homozygous missense mutation (S1006L) in the PCDH18 gene, while others were found to have a heterozygous mutation (S1006L) of PCDH18. The calculated stability (free energy) changes showed that the mutation of PCDH18 mainly destabilized the protein structure. Furthermore, the mutation (S1006L) could lessen the PCDH18-induced inhibition of target cells activation and reduced the apoptosis of T lymphocytes. This study is the first to perform whole-exome sequencing analysis to search for the potential 'second-hit' mechanism that underlies the onset of FHL2. A novel type of compound heterozygous mutation has been found in PRF1. The detection of the homozygous germline mutation in PCDH18 strongly argues that the presence of a 'second' germline mutation besides the PRF1 gene might be potentially an important mechanism of triggering the onset of FHL2.

Keywords: Familial hemophagocytic lymphohistocytosis; Exome sequencing; PRF1; Epstein-Barr virus; PCDH18

Running heads: Whole-exome sequencing in FHL2

Abbreviations:

FHL = familial hemophagocytic lymphohistocytosis; PRF1= perforin 1; APC = antigen-presenting cells; CTLs = cytotoxic T cells; HLH = hemophagocytic lymphohistocytosis; PBS = phosphate balanced solution; SE = standard error; PBMCs = peripheral blood mononuclear cells; LDH = lactate dehydrogenase; EBV = Epstein-Barr virus; NGS = next generation sequencing; ISP = Ion sphereparticles; ATCC = American Type Culture Collection; DMEM = Dul-becco's modified Eagle's medium; rIL-2 = recombinant interleukin-2; 7-AAD = 7-Amino-Actinomycin; CT = computed tomography; FISH = Fluorescence in situ hybridization; EBER = EBV-encoded RNA; TCR = T cell receptor; H&E = Hematoxylin and Eosin; RT-PCR = Reverse transcription-PCR; ITK = IL-2-inducible T cell kinase; TIL = tumor-infiltrating T cells; AICD = activation-induced cell death.

INTRODUCTION

Familial hemophagocytic lymphohistocytosis (FHL) is a genetically heterogeneous clinical syndrome of abnormal immune regulation that is characterized by an uncontrolled hyperinflammatory reaction with high fever, hepatosplenomegaly, cytopenia and occasional central nervous system involvement¹⁻³. To date, four types of genetic defects have been identified in FHL, and almost all of the genes are involved in the regulation of cytotoxic lymphocytes for cytotoxic granule exocytosis or function. Among them, a mutation in the PRF1 gene was the first genetic defect to be identified as a causative factor of FHL, in 1999⁴. The PRF1 gene encodes perforin, a pore-forming protein that is essential for the introduction of cytotoxic effector molecules into the cytosol of target cells, which subsequently induces target cell death. PRF1 mutations have been estimated to account for 13% to 58% of FHL cases depending on the ethnic origin³. FHL cases with PRF1 mutations have been classified as FHL2, which forms the principal subtype of FHL.

The pathophysiology of FHL2 appears to involve the uncontrolled activation of lymphocytes and macrophages. An excessive hyperinflammatory reaction in most cases is triggered by infections. Under normal circumstances, the intensity and duration of an immune response to infection are precisely controlled by a number of negative-feedback mechanisms. Several feedback mechanisms have been proposed to be initiated and to stimulate the immune response, which is triggered by infections, including perforin- and Fas-dependent killing of antigen-presenting cells (APC) by cytotoxic T cells (CTLs), perforin-dependent killing of activated CD4 cells by NK

cells and the elimination of highly expanded CD8 T cells by cytotoxic CD4 T-regulatory cells. A number of studies have shown that the lack of perforin results in a failure to kill APC as well as activated T cells and leads to abnormal persistence of immune stimuli. Consequently, the profound hyperinflammation leads to multiple-organ dysfunction and the highly fatal clinical manifestation characteristics of FHL2.

Although most of the FHL2 patients typically present with FHL2 in early childhood, there are a few late-onset cases of teenage and adult FHL. FHL2 could develop beyond infancy and as late as the fifth decade of life⁵. Sometimes, FHL2 siblings with the same mutations within a family develop the disease in different ways, with various onset ages⁶. Furthermore, not all FHL2 patients who carry mutations in the perforin gene will develop FHL2 eventually⁷. The proposed explanation for the discrepancy in FHL2 onset is that various mutations in the PRF1 gene could lead to the loss of PRF1-dependent cytotoxicity to a varying extent. In addition, the FHL2 onset is also correlated with environmental or viral exposure factors. On the other hand, it is not known whether there are additional mechanisms that can synergize with the defective PRF1 function to enable the development of FHL2. To explore the potential unknown mechanisms of FHL2 would extend our insights into the genetic basis of FHL2 and allow a better understanding of the immune control during anti-infection and anti-cancer.

In the present study, a 25-year-old newly diagnosed FHL2 patient was detected with compound mutations in both PRF1 alleles and positive EBV infection.

Interestingly, her adult sibling carried the same PRF1 gene mutations and EBV infection status but did not develop FHL2. To search for the potential unknown mechanisms that are capable of synergizing with the defective PRF1 function, we have performed whole-exome sequencing analysis for the patient and four other unaffected family members. The acquired somatic mutations of the patient and germline mutations related to the family were analyzed. We proposed that mutation in the PCDH18 gene and an inhibitory signaling receptor of effector phase CD8⁺ T lymphocytes might involve the development of the FHL2 patient⁸. The functional analysis was also performed to elucidate the impact of this gene mutation on the pathogenesis of FHL2.

MATERIAL AND METHODS

Patient and Samples. The clinical data were obtained from a non-consanguineous Chinese parentage patient (without family history of HLH or cancer) who presented in Tongji Hospital and was diagnosed as familial hemophagocytic lymphohistocytosis according to Diagnostic Guidelines for hemophagocytic lymphohistocytosis (HLH) proposed by the Histiocyte Society^{9,10}. Peripheral blood and saliva samples were obtained from the patient and her four Chinese unaffected family members. Appropriate informed consent was obtained from all donors prior to specimen collection in accordance with the Declaration of Helsinki and under a research protocol approved by the ethics committees of Tongji Hospital.

Flow Cytometry. Whole blood samples were collected in sodium heparin-containing vials, and red cells were lysed for 5 minutes with 2 ml of FACSlyse (BD Immunocytometry Systems, San Jose, CA) and washed with Phosphate Balanced Solution (PBS). The resultant white cell pellet was resuspended with 100 μ l PBS and stained with the following antibodies: CD3-Percp, CD8-APC and CD56-FITC (BD Immunocytometry Systems, San Jose, CA) for 20 minutes at room temperature. After washing with PBS, the white cells were permeabilized with Cytofix/Cytoperm (BD Pharmingen, San Diego, CA) and stained with either phycoerythrin (PE)–conjugated anti-perforin or PE-conjugated mouse IgG2b (BD Pharmingen, San Diego, CA) for 30 minutes at room temperature. After washing, the cells were resuspended with 200 μ l PBS prior to analysis by flow cytometry. The samples were analyzed by flow cytometry using Cell Quest software (BD Biosciences, San Jose, USA). The following gates were used to distinguish the three populations of interest: NK cells were defined as being CD3⁻ and CD56⁺; CD56⁺ T cells were defined as being CD3⁺ and CD56⁺; and CD8⁺ T cells were defined as being CD3⁺CD8⁺CD56⁻. Moreover, all of the populations were restricted to a lymphocyte gate based on the forward *versus* side scatter. The perforin-positive region was set using an isotype-matched negative control, and the percent positive for each region was reported. Each data value represents the mean $\pm$ standard error (SE) of triplicates.

NK cell Activity Assay. Venous blood was obtained from a healthy volunteer, the patient and her family members for the NK cell activity assay. The assay was performed as previously described¹¹. Briefly, peripheral blood mononuclear cells

(PBMCs) were isolated by Ficoll-Paque density gradient (Pharmacia Biotech, Piscataway, NJ) centrifugation. After washing, the lymphocytes were resuspended in culture medium at a concentration of 1×10^7 cells/ml, and target cells (K562 cells) were prepared at a concentration of 1×10^5 cells/ml. Using a round-bottom 96-well culture plate, experimental, spontaneous and maximum wells were set up, and then, the plate was incubated at 37°C for 4 hours. Forty-five minutes prior to the supernatant harvest, 1% NP-40 was added to the target cell maximum lactate dehydrogenase (LDH) release control. Afterward, 50µl of the supernatant was transferred from each well of the assay plate to the corresponding well of a flat-bottom 96-well enzymatic assay plate, and 50µl of the substrate mix was added to each well of the plate. The plate was kept in a dark place and incubated at room temperature for 30 minutes. Then, 50µl of the stop solution was added to each well of the plate. The absorbance at 490nm was recorded and corrected by the computer. The percent of cytotoxicity for each test sample is calculated by the following formula, where cytotoxicity (%) = $(OD_{\text{experimental}} - OD_{\text{spontaneous}}) / (OD_{\text{maximum}} - OD_{\text{spontaneous}}) \times 100\%$. Each data value represents the mean $\pm$ SE of triplicates.

Quantitation of EBV-DNA. Whole blood obtained from the patient and her family members was centrifuged and separated into plasma and cell fractions. The cell fraction was separated into PBMC on Ficoll-Paque density gradients (Pharmacia Biotech, Piscataway, NJ). DNA was extracted from 2×10^6 PBMC using a QIAamp Blood Kit (Qiagen, Hilden, Germany) and resuspended in 50µl distilled water. A real-time quantitative PCR assay with a fluorogenic probe was performed using the

TaqMan PCR kit and a Model 7700 Sequence Detector (PE Applied Biosystems, Foster City, CA). The amount of EBV was calculated as the number of virus copies per microgram PBMC DNA. The EBV detection limit in this assay was approximately 10 copies/ μ g DNA for PBMC. Each data value represents the mean $\pm$ SE of triplicates.

Next Generation Sequencing (NGS). Genomic DNA was obtained from peripheral blood of the proband using a standard procedure and then was tested by NGS using a custom design for the HLH panel, based on an AmpliSeq strategy (Life Technologies, Carlsbad, CA, USA). The panel, which was created using the Ion AmpliSeq™ designer software, was designed to identify disease-causing mutations in the following genes: PRF1, UNC13D, STX11, STXBP2, XIAP, SH2D1A, Rab27a, AP3B1 and LYST. This design allowed analysis of all of the exons of these genes (padding: $\pm$ 50 bp). Libraries were constructed using the Ion AmpliSeq Library Kit v2.0 according to the manufacturer's instructions. Validation of enrichment and quantification of enriched target DNA were performed on the Qubit Fluorometer using the Qubit dsDNA HS Assay Kit (Invitrogen, NY, USA). Amplified libraries were submitted to emulsion PCR performed on the Ion OneTouch™2 system using the Ion PGM™ Template OT2 200 Kit (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. Ion sphereparticles (ISP) were enriched using the Ion OneTouch™ES Instrument. Resultant live ISPs were loaded and sequenced on an Ion 316 chip (Life Technologies, Carlsbad, CA, USA). Torrent Suite™ Software v4.0.2 (Life Technologies, Carlsbad, CA, USA) and Alamut® 2.3.6

(Interactive Biosoftware, Rouen, France) software were further used to perform primary to tertiary analysis. An initial analysis was performed in the NextGENE® software, and then, a.pjt file was created and opened in the NextGENE® Viewer. Variants identified as disease-related mutations that were reported publicly were further verified by conventional Sanger sequencing using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA).

Whole-exome Sequencing and Sanger Sequencing. Three micrograms of genomic DNA were randomly fragmented by Covaris, and the resulting fragments were ligated with adaptors (Illumina, CA, USA). After purification using Agencourt AMPure SPRI beads (Beckman Coulter Genomics, MA, USA), the adaptor-ligated templates were amplified by ligation-mediated PCR (LM-PCR) and hybridized to the SureSelect Biotinylated RNA Library (BAITSI Agilent, SC, USA) for enrichment. Captured LM-PCR products were subjected to Agilent 2100 Bioanalyzer to estimate the magnitude of enrichment. Each captured library was then loaded onto the HiSeq 2000 sequencing platform to ensure that each sample was covered to a depth of at least 72 ×. Raw image files were processed using the Illumina Pipeline version 1.7 for base calling with default parameters, and the sequences of each individual were generated as 90-bp paired-end reads. Sequencing reads were processed to remove adapter and low quality sequences before being aligned to the human reference genome (NCBI build 37.1, hg19) using SOAPaligner16. The CCDS database was used to calculate the depth and coverage for each exon, whereas four databases, CCDS, Refseq, Ensembl and Encode, were used to annotate the variation and print

only the most deleterious variants. On the basis of SOAP alignment results, SOAPsnp (V.1.03) software¹⁷ was used to assemble the consensus sequence and call the genotypes in the target regions. Insertions and deletions (indels) in the exome regions were identified using GATK¹⁸, which aligned to the reference genome by using BWA. Public data available from dbSNP¹³⁷, the 1000 Genomes Project, the HapMap project database and the YH database were used for excluding the polymorphism in general populations. The following cutoffs were used to identify high-priority candidate variants: homozygous in the proband, heterozygous in her family members or compound heterozygous mutations in a single gene; Quality Score ≥ 20 ; coverage depth ≥ 4 ; not overlapping a segmental duplication (USCS Genome Browser Assembly GRCh37), and minor allele frequency < 0.005 (dbSNP¹³⁷). Although FHL is a recessive disease, a dominant model of inheritance was also considered. Because the disease-causing mutation was not likely to be a polymorphism in the general population, dbSNP¹³⁷ (with a frequency greater than 0.5%), 1000 Genome variants databases, eight HapMap exomes, YH and in-house databases were used to filter out variants that had a high frequency. Called coding variants were categorized into nonsense, missense, read-through (stop codon to non-stop codon), splice-site mutations and coding-indel according to the variant types and impact on the sequence of amino acid products.

Sanger sequencing was performed to confirm the candidate variants that were found by whole-exome sequencing. PCR amplification was conducted using a GeneAmp PCR System 9700 (Applied Biosystems) under standard conditions. The

PCR primers were designed to amplify the entire exons and the intron-exon boundaries of the PRF1, C21orf2, FLG, HRNR, SSX4, PCDH18, ANKRD11, NOS1 and FAM4C genes. Detailed methodologies have either been described previously¹² or are available upon request. For further confirmation of the impact of these selected variants on the structure and function of related proteins, the variants were analyzed with SIFT¹³ and PolyPhen-2¹⁴.

Protein Stability Prediction. After homology modeling of human PCDH18, <PositionScan> command of FoldX version 3.0 beta 3 (<http://foldx.crg.es/>)¹⁵ was used to calculate the change in the PCDH18 protein stability upon the missense mutation (c.3017 C>T). Then, the <RepairPDB> command was used to minimize the model structure; if the mutation resulted in $\Delta\Delta G < 0$ kcal /mol, then we considered the protein to be stabilized, while if $\Delta\Delta G > 2$ kcal /mol, the protein would be considered to be destabilized¹⁶. The wild-type and mutant PCDH18 protein structure were built based on the SWISS-MODEL server by PyMOL (PyMOL Molecular Graphics System, Version1.3r1).

Generation of Wild-type and Mutated PCDH18 Plasmids. Human PCDH18 cDNA was synthesized and then amplified by PCDH18 primers and Flag primers, and subcloned into pLVX-CMV-CDGA-T2A-Puro vector *via* EcoR I and BamH I sites to generate the wild-type PCDH18 plasmid. The vector was used to generate the mutated PCDH18 plasmid with the use of the QuikChange Site-Directed Mutagenesis Kit (QIAGEN).

Histology and Immunohistochemistry. Paraffin-embedded tissue sections were

stained with Hematoxylin and Eosin (H&E) or dewaxed in environmental protection dewaxing liquid and subjected to immunohistochemical analysis as previously described¹⁷. Anti-CD3 (mouse monoclonal; Fuzhou Maixin Biotech. Co. Ltd, China), anti-CD68 (mouse monoclonal; Fuzhou Maixin Biotech. Co. Ltd, China), anti-PCDH18 (rabbit polyclonal; sigma, USA) and biotinylated immunoglobulin and conjugated with avidin-horseradish peroxidase (HRP) secondary antibody were used in the present study.

Cell Culture. All of the cell lines, including 293, HepG2, K562 and CTLL-2 cells, were obtained from American Type Culture Collection (ATCC, USA). The 293 cells and HepG2 cells were grown in Dul-becco's modified Eagle's medium (DMEM, Gibco, USA) supplemented with 10% bovine calf serum. K562 cells and CTLL-2 cells were cultured in RPMI1640 supplemented with 10% fetal calf serum, while CTLL-2 cells, which were dependent on recombinant interleukin-2 (rIL-2) for growth, were added with rIL-2 (500 IU/ml).

Transfection of 293 cells and HepG2 cells, RT-PCR and Western blot. The 293 cells and HepG2 cells were transiently transfected with wild-type or mutated PCDH18 expression plasmid, with plasmid control using X-tremeGENE HP DNA Transfection Reagent (Roche, Switzerland) following the standard procedure. At 48 hours after the transfection, 3×10^6 of 293 cells and HepG2 cells were harvested for mRNA and protein analysis. For the relative mRNA expression analysis, the total RNA was harvested from 1×10^6 293 cells or HepG2 cells using the RNeasy Mini Kit (QIAGEN); 1 μ g of total RNA was used to generate cDNA and subjected to Real-time PCR using

the universal 2× qPCR Master Mix. For the relative protein analysis, the total protein was isolated from 2×10^6 293 cells or HepG2 cells, and then, the protein was separated on 10% SDS-PAGE and then blotted onto a PVDF membrane. After blocking with 5% non-fat milk in TBST for 1 hour at room temperature, the membranes were incubated overnight at 4 °C with primary antibodies that were specific for Flag (mouse monoclonal; sigma, USA) and GAPDH (mouse monoclonal; sigma, USA). After the washing procedures, the membranes were incubated with horseradish peroxidase-conjugated secondary antibodies (Pierce Biotechnology, Rockford, IL, USA) of IgG for 1 hour at room temperature. After washing with TBST, the primary antibodies were visualized with the SuperSignal West Pico chemiluminescent substratedetection (Pierce Biotechnology, Rockford, IL, USA). Signals were detected by a chemiluminescence detection system (Bio-Rad, USA). The protein expression level was quantified by image-pro plus 6.0 software (Media Cybernetics, USA), and the value represents the mean $\pm$ SE of triplicates.

Transfection of CTLL-2 cells and function analysis. CTLL-2 cells (1×10^6 cells) were suspended in 100 μ l electroporation buffer, which was provided in Lonza® SF Cell Line 4D-Nucleofector® Kit (Lonza, Switzerland). Then, 4 μ g of plasmids were used to transfect CTLL-2 cells by the 4D-Nucleofector TM system (Lonza, Switzerland) with the CM-120 program. The post-transfected cells were cultured in RPMI1640 supplemented with 20% fetal calf serum with rIL-2, 24h later, CTLL-2 cells were separated into two groups, one was added with rIL-2, the other was not. After 24h, the cells were harvested, and the activation rate was analyzed by incubating

CTLL-2 cells with mouse monoclonal CD69 antibody (eBioscience, USA), which was an activation marker on the cell surface in the lymphocyte subsets; then, the frequency of apoptosis in the CTLL-2 cells was assessed using extracellular staining with recombinant Annexin V and 7-Amino-Actinomycin (7-AAD) (eBioscience, USA).

RESULTS

Clinical manifestation and diagnosis of the proband. The proband in the present study was a 25-year-old female who presented with fever of unknown origin for two months. In primary care, she received antibiotics, but the fever persisted. Physical examination showed enlargement of the right cervical and axillary lymph nodes. Enlargement of the liver and spleen can be easily palpated. At administration, the white blood cell count was $1.66 \times 10^9/\text{L}$ with $0.95 \times 10^9/\text{L}$ neutrophils and $0.57 \times 10^9/\text{L}$ lymphocytes. The hemoglobin level was 84 g/L. The platelet level was $32.0 \times 10^9/\text{L}$. The fibrinogen was 1.35 g/L (normal range: 2 g~4 g/L). The serum LDH was 1225 U/L. High serum ferritin (2000 ng/ml) and serum triglyceride (4.08mmol/L, normal range: 0.05~1.70 mmol/L) were detected. Hepatosplenomegaly could be confirmed by ultrasonography as well as abdomen computed tomography (CT) scan. EBV-DNA was 5.39×10^5 copies/ μg . The pulmonary CT scan showed mediastinal lymphadenopathy with bilateral pleural effusion. Histological examination of the resected cervical lymph nodes showed reactive hyperplasia of T lymphocytes with positive immune-reactive stain for CD2, CD3, CD5, CD7, CD43, CD4, CD8,

GrB and TIA-1 in the paracortex zone but negative for CD30, ALK, TdT, CD34 and MPO. Fluorescence *in situ* hybridization (FISH) showed negative stain for EBER. Hemophagocytosis can be observed in bone marrow smear without signs of lymphoma or other malignancies (Figure 1a). Both bone marrow and resected lymph node samples were detected with a positive monoclonal T cell receptor (TCR) rearrangement. However, no clonal population of cells could be detected by flow cytometry in either the bone marrow or lymph node samples. Furthermore, deficient activity of NK cells was observed in the proband and her brother, who is currently still being followed-up with no reported complaints, in comparison to those of their parents and the proband's son (Table 1).

A splenectomy and biopsy of liver had been performed later for diagnostic purposes and to rule out the possibility of malignant lymphoma. Tissue sections stained with HE showed a large number of histiocyte-engulfed mature blood cells in the splenic sinus (Figure 1b and c) with normal architecture of the spleen. CD68⁺ macrophagocytes could be easily observed in the red pulp of the spleen (Figure 1d, black arrows). The presence of periportal small lymphocyte infiltration that mimics chronic active hepatitis was readily observed (Figure 1e). Infiltrating cells expressed CD2, CD3 (Figure 1f), CD5, CD7, CD4, CD8, TIA-1 and GrB but were negative for CD56. FISH showed positive stain for EBER in a dispersed pattern. The analysis of bone marrow, spleen and lymph node samples showed normal karyotypes of the chromosomes.

According to the Diagnostic Guidelines for HLH proposed by the Histiocyte Society, the proband met seven of eight criteria and thereby was diagnosed as hemophagocytic lymphohistocytosis. She was then treated with the international HLH treatment protocol HLH-2004¹⁰.

NK activity, perforin protein expression and EBV test. Deficient activity of NK cells was observed in the proband and her brother in comparison to those of their parents (Table 1). Substantially reduced perforin expression on CD56⁺CD3⁺NK cells, CD8⁺CD3⁺CD56⁻T cells and CD3⁺CD56⁺T cells were determined in the proband as well as her brother in comparison to their parents and the proband's son (Figure 2, **P*<0.05). A high load of EBV detected in the PBMCs of the proband, whereas a lower load of EBV was found in the proband's family members by quantitative PCR (Table 1).

Mutation Analysis. Because flow cytometry analysis showed substantially reduced expression of perforin protein in the proband and her brother, the HLH gene mutation was then examined. The NGS and Sanger sequencing detected the mutation of the perforin gene in this family. The patient and her sibling brother carried the same heterozygous missense mutation (c.916G>A; p.Gly306Ser) in exon 3 of PRF1, and a heterozygous frameshift mutation (c.65delC; p.Pro22ArgfsX29) in exon 2 of PRF1. Apparently, the c.916G>A variant was inherited from their mother, whereas the c.65delC was inherited from their father, and thus, both alleles of PRF1 were mutated. The patient's son inherited a missense mutation in exon 3 of PRF1 (c.916G>A; p.Gly306Ser). Figure 3 depicts the pedigree for this three-generation late-onset FHL2.

The UNC13D, STX11, STXBP2, XIAP, SH2D1A, Rab27a, AP3B1 and LYST genes related to other types of FHL were not mutated in this family.

Alignment and coverage statistics. The phenomenon that the proband and her adult sibling carried the same perforin gene mutations as well as EBV positive status but exhibited different clinical outcomes in developing FHL2 prompts us to ask whether certain additional factors contributed to the onset of FHL2 in the patient. To search for the potential unknown mechanisms that are capable of synergizing with the defective PRF1 function, we have performed whole-exome sequencing analysis for the patient and four other unaffected family members. The total number of reads ranged from 62987322 to 82421666 per sample. Table 2 summarized the general information of the seven samples. Here, 86%~92% aligned to the reference human genome, and 82.57%~85.87% of the reads were covered at more than 20 times. The mean depth of the target region ranged from 64.45 to 84.61×. Each sample had a total of 94299-103482 single nucleotide substitutions. The number of indels was 9017 to 10172. An average total of 7,589 genetic variations, including non-synonymous mutations, splice-acceptor and donor site variations and insertions or deletions (indels), were identified per sample.

Search for possible somatic mutations of FHL2 patients. Theoretically, a possible additional factor besides defective PRF1 in the proband could be obtained in two ways, spontaneous somatic mutation or homozygous/compound germline mutation inherited from her heterozygous parents. We initially searched for the spontaneous somatic mutation, which could regulate the immune response and act as a

‘second-hit’ to help to start the FHL2 onset. Genetic variants were prioritized for further evaluation using the following filtering criterion: the variant was found exclusively in the peripheral blood sample of the affected individual but not in her saliva sample and the peripheral blood samples of her parents and brother. According to the filtering criterion, three heterozygous candidate mutations or indels within the FASTKD3, HOXA10 and SVEP1 genes were identified. Detailed information on these three genes is listed in Tables 3 and 4.

Search for potential germline mutated genes related to FHL2. Next, we searched the homozygous or compound germline mutation inherited from her heterozygous parents. Genetic variants were prioritized for further evaluation using the following filtering criterion: the mutation was detected with a priority of less than one and detected in all seven samples in either the recessive model or dominant model. Table 3 summarized the detailed information on the genetic variants that were related to all of the family members. Five germline gene mutations (PCDH18, CKD11B, MAGEC1, NOS1 and PPP1R14BP3) were detected in an autosomal recessive model, while six germline mutated genes (MLL3, PCDH18, ANKRD36, HNRNPC, PCMTD1 and MAGEC1) were detected in an autosomal dominant model. Notably, only the proband was exclusively detected with a homozygous missense mutation (S1006L) in the PCDH18 gene, while all of her family members were found to have a heterozygous mutation (S1006L) of the PCDH18 gene (Table 4, Figure 4A). Apparently, the homozygous missense mutation (S1006L) of the PCDH18 gene in the

proband was inherited from her heterozygous parents. The S1006L missense mutation of PCDH18 was predicted to be damaging, by Software Polyphen-2.

Functional analysis of the PCDH18 mutation. The missense mutation (c.3017C>T) resulted in a change in the protein amino acid sequence (p.S1006L) (Figure 4A). The total Gibb's free energy change that resulted from the mutation was greater than 2 kcal/mol ($\Delta\Delta G = 8.08$ kcal/mol), which demonstrated that the mutation destabilizes the protein structure. The wild-type, mutated (S1006L) PCDH18 or negative control plasmids were transfected into 293 and HepG2 cells. The S1006L mutant protein expression reduced significantly compared with WT controls (Figure 4B). While reverse transcription-PCR (RT-PCR) analysis demonstrated that there was no significant difference in the mRNA expression in the 293 cells or in the HepG2 cells (Figure 4B). Then the empty vector, wild-type and mutated plasmids were transfected into CTLL-2 cells which were dependent on rIL-2 for growth and activation, and then, rIL-2 deprivation suppressed their growth and activation. Our results showed that the wild-type PCDH18 exerted an inhibitory effect on activation and increased the apoptosis rates of CTLL-2 cells with or without rIL-2 compared with empty vector. The mutated type, however, could counteract these two effects (Figure 4C-D, $*P<0.05$). Immunohistochemistry data demonstrated that the expression level of PCDH18 protein was much lower in the proband than in her asymptomatic sibling and normal controls (Figure 4E.a-e).

DISCUSSION

To our knowledge, the present investigation is the first study to perform whole-exome sequencing analysis on an FHL2 family. Three reasons have prompted us to check the genetic makeup of the family from the global view of genomes rather than the well-known HLH gene approach. First, although seven genes (LYST, AP3, Rab27a, STX11, STXBP2, MUNC13-4 and PRF1) have been generally regarded as the causative factors of FHL and all of the genes are involved in the regulation of cytotoxic granule exocytosis or function, dozens of additional genes have also been identified to be causative factors of HLH²⁰. For example, a deficiency in IL-2-inducible T cell kinase (ITK) causes a fatal HLH response to EBV infection. Thus, while the current portrait of FHL pathophysiology is well accepted, our understanding of the molecular basis of FHL is still in its infancy and must be reinforced. Second, at the present time, clinical diagnosis of FHL relies on identification of disease-related mutations that are available on the publicly accessible database. However, the list of FHL-related mutations is far from complete and must be updated. Third, in the present study, the newly diagnosed FHL2 adult patient and her adult sibling carried the same perforin gene mutations and EBV infection status but differed in the onset of FHL2, which provided a good model for searching the potential unknown mechanisms that are capable of synergizing with the defective PRF1 function. Therefore, we hypothesized that the potential unknown mechanisms besides a defective PRF1 in the proband could be gained by spontaneous somatic mutations or homozygous/compound germline mutations inherited from her heterozygous parents. The whole-exome sequencing analysis was thus performed on an FHL2 patient and her three-generation family

members to test the hypothesis.

To date, more than 70 FHL-related perforin mutations have been described^{6,7,21-31}. The mutation sites are quite different among different ethnic groups. All of the individuals in African American patients carry the same 50delT mutation²⁷. Turkish people express the Trp374X perforin mutation with high frequency³¹, and Italian families express the A91V sequence variant. In Japan, the most common perforin gene mutation is the 1090-1091delCT (62.5%) and 207delC (37.5%)³⁰. Nevertheless, the genetic information on FHL2 in the Chinese population is rare based on the published literature. The proband in the present study carried a heterozygote missense mutation (c.916G>A) and a frameshift mutation (c.65delC) on another allele of the PRF1 gene. 65delC mutation is a novel FHL2-related PRF1 gene defect that was first reported in a fatal neonate case recently³² but was not previously reported in any adult cases. The c.916G>A mutation appears to be a previously unknown pathogenic mutation. The mutation had been described by other investigators despite that they had not provided clear evidence to support its pathogenic effect^{33,34}. The study of Nagafuji had described the c.916G>A mutation in an adult onset patient with residual perforin activity, which is similar to the findings revealed in the current paper³³. The c.916G>A could therefore be a hypomorphic mutation, that does not cause classic early onset FHLH, but together with a clear pathogenic PRF1 mutation could make the subject more sensitive to developing HLH. Our data emphasized that the combination of c.916G>A and c.65delC was a new type of compound heterozygous mutation, which was an FHL2-causing defect. A recent large-scale retrospective study established the

FHL-causing effect of the compound mutation in the PRF1 gene²⁰.

Siblings who had the same mutant phenotype within a family but showed different onset ages were previously reported⁶. The additional unknown genetic mechanisms might account for the deviations. The whole-exome sequencing analysis revealed that the proband acquired spontaneous somatic mutations in several genes, including FASTKD3, HOXA10 and SVEP1. FASTKD3 is an essential component of mitochondrial respiration, and it could modulate energy balance in cells that are exposed to adverse conditions by functionally coupling mitochondrial protein synthesis to respiration³⁵. HOXA10, which belongs to the homeobox genes family, is directly involved in the regulation of hematopoiesis and is implicated in the pathogenesis of acute myeloid leukemia³⁶. SVEP1 (sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1) is a gene that has an unknown function. While the relevance of these spontaneous somatic mutations to the onset of FHL2 is not straightforward and requires further investigation, the detection of somatic mutations on the late-onset FHL2 patient with a chronic EBV infection suggests that an FHL2 patient with positive EBV is to some extent genetically unstable and has the potential risk of transforming to hematological malignancies. Notably, the search for the homozygous or compound germline mutation revealed that the proband was exclusively detected with a homozygous missense mutation (S1006L) in the PCDH18 gene. Apparently, the potentially damaging homozygous missense mutation (S1006L) of the PCDH18 gene in the proband was inherited from her heterozygous parents. Surprisingly, although the biological function of PCDH18 remains largely unknown for

the time being and the mutation in PCDH18 we detected has a frequency in Japanese of 0.22% (HGVD database), the most recent study has shown PCDH18 to be a novel activation marker of CD8⁺ memory T cells and functions as an inhibitory signaling receptor to restrict the effector phase of the T cell activation. PCDH18 is transcribed upon *in vitro* or *in vivo* activation of all CD8⁺ central memory T cells. The transfection of PCDH18 in primary CD8⁺ effector cells induces the phenotype of nonlytic CD8⁺ tumor-infiltrating T cells (TIL) that showed defective proximal TCR signaling, cytokine secretion and cytolysis and enhanced the activation-induced cell death (AICD) of T cells⁸.

Given the critical role of CD8⁺ T cell expansion in the pathophysiology of HLH, PCDH18 is therefore proposed to be associated with the onset of the proband. Just as the functional analysis of the PCDH18 mutation (S1006L) demonstrated, the mutation destabilized the protein structure and reduced its protein level. Furthermore, the mutation could lessen the PCDH18-induced inhibition of target cells activation and reduce the apoptosis of CTLL-2 cells. These results gave us a hint that the PCDH18 missense mutation, accompanied with the PRF1 mutation, might attribute the ‘second-hit’ of FHL2 through lessening the PCDH18-induced inhibition of CTLs activation and reducing their apoptosis, which resulted in uncontrolled CTLs activation and defective apoptosis, consequently triggering the onset of FHL2.

The present study for the first time used whole-exome sequencing to search for the potential ‘second-hit’ mechanism that underlies the onset of FHL2. A new type of compound heterozygous mutation (c.916G>A and c.65delC) was identified as an

FHL2-causing defect. More importantly, this study has highlighted a novel mechanism of FHL2 besides the current classical mechanism that FHL2 is always caused by defects in the regulation of cytotoxic granule exocytosis or function. More potential susceptible genes should be further investigated to support the second-hit hypothesis of FHL in the future.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Science Fund for Distinguished Young Scholars of China (no.81025011), the National Science Foundation of China (NSFC 81300390), the Major Program of the National Science Fund (nos.81090414, 81230052), and the National High Technology Research and Development Program of China (863 Program, no.2012AA020801). All authors have read the journal's policy on disclosure of potential conflicts of interest and declare that they have no competing interests. All authors have read the journal's authorship agreement and the manuscript has been reviewed by and approved by all named authors.

REFERENCES

1. Cetica V, Pende D, Griffiths GM, Arico M. Molecular basis of familial hemophagocytic lymphohistiocytosis. *Haematologica*. 2010;95(4):538-541.
2. Filipovich AH. Hemophagocytic lymphohistiocytosis and other hemophagocytic disorders. *Immunol Allergy Clin North Am*. 2008;28(2):293-313, viii.
3. Janka GE. Familial and acquired hemophagocytic lymphohistiocytosis. *Annu Rev Med*. 2012;63:233-246.
4. Stepp SE, Dufourcq-Lagelouse R, Le Deist F, et al. Perforin gene defects in familial hemophagocytic lymphohistiocytosis. *Science*. 1999;286(5446):1957-1959.
5. Mancebo E, Allende LM, Guzman M, et al. Familial hemophagocytic lymphohistiocytosis in an adult patient homozygous for A91V in the perforin gene, with tuberculosis infection. *Haematologica*. 2006;91(9):1257-1260.
6. Ueda I, Kurokawa Y, Koike K, et al. Late-onset cases of familial hemophagocytic lymphohistiocytosis with missense perforin gene mutations. *Am J Hematol*. 2007;82(6):427-432.
7. Clementi R, Emmi L, Maccario R, et al. Adult onset and atypical presentation of hemophagocytic lymphohistiocytosis in siblings carrying PRF1 mutations. *Blood*. 2002;100(6):2266-2267.
8. Vazquez-Cintron EJ, Monu NR, Burns JC, et al. Protocadherin-18 is a novel differentiation marker and an inhibitory signaling receptor for CD8+ effector memory T cells. *PLoS One*. 2012;7(5):e36101.
9. Henter JI, Elinder G, Ost A. Diagnostic guidelines for hemophagocytic lymphohistiocytosis. The FHL Study Group of the Histiocyte Society. *Semin Oncol*. 1991;18(1):29-33.
10. Henter JI, Horne A, Arico M, et al. HLH-2004: Diagnostic and therapeutic guidelines for

hemophagocytic lymphohistiocytosis. *Pediatr Blood Cancer*. 2007;48(2):124-131.

11. Chan FK, Moriwaki K, De Rosa MJ. Detection of necrosis by release of lactate dehydrogenase activity. *Methods Mol Biol*. 2013;979:65-70.

12. Ihle MA, Fassunke J, Konig K, et al. Comparison of high resolution melting analysis, pyrosequencing, next generation sequencing and immunohistochemistry to conventional Sanger sequencing for the detection of p.V600E and non-p.V600E BRAF mutations. *BMC Cancer*. 2014;14:13.

13. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073-1081.

14. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.

15. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The FoldX web server: an online force field. *Nucleic Acids Res*. 2005;33(Web Server issue):W382-388.

16. An O, Gursoy A, Gurgey A, Keskin O. Structural and functional analysis of perforin mutations in association with clinical data of familial hemophagocytic lymphohistiocytosis type 2 (FHL2) patients. *Protein Sci*. 2013;22(6):823-839.

17. He YF, Zhang GM, Wang XH, et al. Blocking programmed death-1 ligand-PD-1 interactions by local gene therapy results in enhancement of antitumor effect of secondary lymphoid tissue chemokine. *J Immunol*. 2004;173(8):4919-4928.

18. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res*. 2001;11(5):863-874.

19. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic*

Acids Res. 2002;30(17):3894-3900.

20. Zhang K, Chandrakasan S, Chapman H, et al. Synergistic defects of different molecules in the cytotoxic pathway lead to clinical familial hemophagocytic lymphohistiocytosis. *Blood*. 2014.

21. Busiello R, Adriani M, Locatelli F, et al. Atypical features of familial hemophagocytic lymphohistiocytosis. *Blood*. 2004;103(12):4610-4612.

22. Clementi R, Locatelli F, Dupre L, et al. A proportion of patients with lymphoma may harbor mutations of the perforin gene. *Blood*. 2005;105(11):4424-4428.

23. Feldmann J, Le Deist F, Ouachee-Chardin M, et al. Functional consequences of perforin gene mutations in 22 patients with familial haemophagocytic lymphohistiocytosis. *Br J Haematol.* 2002;117(4):965-972.

24. Goransdotter Ericson K, Fadeel B, Nilsson-Ardnor S, et al. Spectrum of perforin gene mutations in familial hemophagocytic lymphohistiocytosis. *Am J Hum Genet.* 2001;68(3):590-597.

25. Kogawa K, Lee SM, Villanueva J, Marmer D, Sumegi J, Filipovich AH. Perforin expression in cytotoxic lymphocytes from patients with hemophagocytic lymphohistiocytosis and their family members. *Blood*. 2002;99(1):61-66.

26. Lee SM, Sumegi J, Villanueva J, et al. Patients of African ancestry with hemophagocytic lymphohistiocytosis share a common haplotype of PRF1 with a 50delT mutation. *J Pediatr.* 2006;149(1):134-137.

27. Molleran Lee S, Villanueva J, Sumegi J, et al. Characterisation of diverse PRF1 mutations leading to decreased natural killer cell activity in North American families with haemophagocytic lymphohistiocytosis. *J Med Genet.* 2004;41(2):137-144.

28. Risma KA, Frayer RW, Filipovich AH, Sumegi J. Aberrant maturation of mutant perforin

- underlies the clinical diversity of hemophagocytic lymphohistiocytosis. *J Clin Invest.* 2006;116(1):182-192.
29. Suga N, Takada H, Nomura A, et al. Perforin defects of primary haemophagocytic lymphohistiocytosis in Japan. *Br J Haematol.* 2002;116(2):346-349.
30. Ueda I, Morimoto A, Inaba T, et al. Characteristic perforin gene mutations of haemophagocytic lymphohistiocytosis patients in Japan. *Br J Haematol.* 2003;121(3):503-510.
31. Zur Stadt U, Beutel K, Kolberg S, et al. Mutation spectrum in children with primary hemophagocytic lymphohistiocytosis: molecular and functional analyses of PRF1, UNC13D, STX11, and RAB27A. *Hum Mutat.* 2006;27(1):62-68.
32. Kim JY, Shin JH, Sung SI, et al. A novel PRF1 gene mutation in a fatal neonate case with type 2 familial hemophagocytic lymphohistiocytosis. *Korean J Pediatr.* 2014;57(1):50-53.
33. Nagafuji K, Nonami A, Kumano T, et al. Perforin gene mutations in adult-onset hemophagocytic lymphohistiocytosis. *Haematologica.* 2007;92(7):978-981.
34. Wang Y, Wang Z, Chen H, Wang X. Adult onset of primary hemophagocytic syndrome in subjects carrying PRF1 mutations. *Ann Hematol.* 2012;91(9):1489-1490.
35. Simarro M, Gimenez-Cassina A, Kedersha N, et al. Fast kinase domain-containing protein 3 is a mitochondrial protein essential for cellular respiration. *Biochem Biophys Res Commun.* 2010;401(3):440-446.
36. Shah CA, Bei L, Wang H, Platanias LC, Eklund EA. HoxA10 protein regulates transcription of gene encoding fibroblast growth factor 2 (FGF2) in myeloid cells. *J Biol Chem.* 2012;287(22):18230-18248.

FIGURE LEGENDS

Figure 1. Morphologic findings in the proband with hemophagocytic

lymphohistocytosis. (A) Hemophagocytosis was readily detectable in BM aspirate smears (black arrow). BM smears were stained with Wright–Giemsa (original magnification, $\times 400$). (B) Tissue sections of spleen stained with Hematoxylin and Eosin (H&E) showed extensive infiltration of lymphocytes and histiocytes with normal architecture of spleen (original magnification, $\times 40$). (C) A large number of histiocytes (black arrows) engulfed mature blood cells in the splenic sinus (original magnification, $\times 600$). (D) Immunohistochemical staining of CD68 in spleen tissue (original magnification, $\times 600$) showed extensive infiltration of histiocytes (brown stain). (E) Tissue sections of resected liver stained with H&E showed extensive infiltration of lymphocytes (original magnification, $\times 40$). (F) Immunohistochemical staining of CD3 in resected liver showed extensive infiltration of CD3⁺ T cells (black arrows) in the portal area of the liver (original magnification, $\times 40$).

Figure 2. Expression of perforin protein in subpopulations of cytotoxic

lymphocytes. Cytotoxic lymphocytes (CD3⁻CD56⁺ NK cells, CD8⁺CD3⁺CD56⁻ T cells and CD3⁺ CD56⁺ T cells) were examined by flow cytometry for the expression of perforin protein in the proband and her family members. (A) Representative flow cytometry images of determination of perforin protein in the CD56⁺CD3⁻ NK cell population. CD56⁺CD3⁻ NK cells were first gated and subsequently analyzed for the perforin protein. The numbers refer to the cell percentages that are positive for perforin protein. (B) The expression of perforin protein in various subpopulations of cytotoxic

lymphocytes. Data represent the mean percentage $\pm$ SE of triplicates.* $P<0.05$ compared to the proband.

Figure 3. The pedigree of the three-generation family affected by the perforin gene. Filled-in symbols indicate the individuals with the c.916G>A and c.65delC mutations. The black arrow indicates the proband.

Figure 4. Functional analysis of mutated PCDH18. (A) PCDH18 protein domain structure. The PCDH18 gene contains four exons that encode several domains in the PCDH18 protein. All family members had missense mutations (c.3017C>T, p.Ser1006Leu) in the Coding Sequence (CDS) area. The wild-type and mutant PCDH18 protein structure were built based on the SWISS-MODEL server by PyMOL (PyMOL Molecular Graphics System, Version1.3r1). The red ball represents the unaffected Serine and the green ball represents (white arrow) the substitute Leucine. (B) The protein expression of PCDH18 reduced significantly after transfection with mutated PCDH18 compared with WT controls in both 293 and HepG2 cells (Figure 4B, upper panel and lower left panel, * $P<0.05$), while the mRNA expression was similar (Figure 4B, lower right panel, $P>0.05$). (C and D) The wild-type PCDH18 exerted an inhibitory effect on activation and increased the apoptosis rates of CTLL-2 cells with or without rIL-2 compared with empty vector. The mutated type, however, could counteract these two effects (Figure 4C-D, * $P<0.05$). (E) The expression level of PCDH18 protein was much lower in the proband's bone marrow (Figure 4E. a) and lymph node (Figure 4E.d) than that in the asymptomatic sibling's bone marrow (Figure 4E.b) and normal control's bone marrow (Figure 4E.c) and lymph node (Figure 4E.e).

Table 1. The genetic and clinical characteristics of the family

Famliy members	Proband	Sibling	Son	Father	Mother
Age (years)	25	23	1	48	47
NK activity (%)	1.18±0.32	4.27±0.58	ND	29.21±7.57	23.73±3.01
perforin (%) NK cells (CD56+CD3-)	0.44±0.53	0.77±0.42	56.4±4.36	54.68±6.68	69.11±3.39
CD8+T cells (CD8+CD3+CD56-)	0.23±0.12	0.06±0.05	15.16±2.34	0.32±0.22	7.66±2.25
CD56+T cells (CD56+CD3+)	0.43±0.36	0.04±0.03	17.09±3.14	8.74±2.67	45.19±3.78
PRF1 gene mutation	c.65delC;916G>A	c.65delC;916G>A	c.916G>A	c.65delC	c.916G>A
Neucleotide change	p. Pro22ArgfsX29; Gly306Ser	p.Pro22ArgfsX29; Gly306Ser	p.Gly306Ser	p.Pro22ArgfsX29	p.Gly306Ser
EBV-DNA Status (copies/μg)	(1.57±0.4)× 10 ⁵	(8.13±1.03) × 10 ²	(3.23±0.75) × 10 ³	(3.77±1.08) × 10 ³	(4.95±0.57) × 10 ²

ND: not done

Table 2. Alignment, coverage statistics and gene mutations by samples

	proband		brother		father	mother	son
	PB	S	PB	S	PB	PB	PB
%Coverage at 20X	84.69	83.25	84.38	85.87	83.17	84.54	82.57
Average Coverage	98.92	98.82	98.90	98.94	98.81	98.87	98.78
Mean depth of target region (X)	73.73	66.76	71.15	84.61	70.51	76.29	64.45
%PCR duplicates	5.01	5.72	5.84	9.16	5.48	5.78	5.48
Total number of SNPs	103482	102127	100973	94299	98994	102812	96350
Homozygous (SNP)	40019	39555	39130	36521	38454	39804	38448
Heterozygous (SNP)	63463	62572	61843	57778	60540	63008	57902
Missense	11592	11479	11571	11623	11346	11398	11267
Total number of Indels	10100	9927	9781	9017	9546	10172	9076
Homozygous (Indel)	4371	4302	4276	3954	4104	4339	4061
Heterozygous (Indel)	5729	5625	5505	5063	5442	5833	5015
Frameshift	235	234	243	247	235	231	208

Table 3. The general information of candidate variants

Gene Name	Chrom:Position	Mutaion Type	Prediction from Polyphen-2	Substitution	Gene description
The candidate somatic variants in FHL2 patient					
FASTKD3	chr5: 7868314	SpliceSite	-	-	FAST kinase domains 3
HOXA10	chr7:27213757	cds-Indel	-	-	-
SVEP1	chr9:113137745	SpliceSite	-	-	von Willebrand factor type A, EGF and pentraxin domain containing 1
The candidate variants related with the family under an autosomal recessive model					
PCDH18	chr4:138442574	missense	probably damaging	S1006L	protocadherin 18
CDK11B	chr1:1654067	missense	-	-	-
MAGEC1	chrX:140993945	missense	benign	F252S	melanoma antigen family C, 1
NOS1	chr12: 117691485	missense	benign	L869P	nitric oxide synthase 1 (neuronal)
PPP1R14BP3	chr4:140036422	missense	-	-	protein phosphatase 1, regulatory (inhibitor) subunit 14B pseudogene 3
The candidate variants related with the family under an autosomal dominant model					
MLL3	chr7: 151970859	missense	probably damaging	G315S	myeloid/lymphoid or mixed-lineage leukemia 3
PCDH18	chr4: 138442574	missense	probably damaging	S1006L	protocadherin 18
ANKRD36	chr2: 97830177	missense	-	G501V	-
HNRNPC	chr14: 21679406	missense	benign	S299R	heterogeneous nuclear ribonucleoprotein C (C1/C2)
PCMTD1	chr8: 52733144	missense	benign	V281T	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 1
MAGEC1	chrX: 140993945	missense	benign	F252S	melanoma antigen family C, 1

- means no prediction available; Indel , insertion and deletion

Table 4 Detailed information for the candidate variants related to the family

Gene Name	Inheritant model	Reference	Variant	Mutation type	Genotype (Quality Score)						
					Proband		Sibling		Father	Mother	Son
					PB	SA	PB	SA	PB	PB	PB
PCDH18	R/D	G	A	Non-synonymous: missense	1/1(85)	1/1(83)	0/1(99)	0/1(99)	0/1(99)	0/1(99)	0/1(99)
CDK11B	R	C	T	Non-synonymous: 5-UTR	1/1 (52)	1/1 (42)	1/1 (44)	1/1 (77)	1/1 (49)	1/1 (48)	1/1 (54)
MAGEC1	R	T	C	Non-synonymous: missense	0/1(23)	0/1(99)	1/1(99)	1/1(99)	1/1(99)	0/1(99)	1/1(99)
NOS1	R	A	G	Non-synonymous: missense	1/1(53)	1/1(90)	0/1(99)	0/1(99)	0/1(99)	0/1(99)	0/1(99)
PPP1R14BP3	R	T	C	Non-synonymous: NR_exon	0/1(99)	0/1(31)	1/1(35)	1/1(37)	1/1(26)	0/1(99)	0/1(57)
MLL3	D	C	T	Non-synonymous: missense	0/1(99)	0/1(99)	0/1(99)	0/1(92)	0/1(99)	0/1(99)	0/1(65)
ANKRD36	D	G	T	Non-synonymous: missense	0/1(99)	0/1(99)	0/1(91)	0/1(99)	0/1(99)	0/1(99)	0/1(99)
HNRNPC	D	G	C	Non-synonymous: missense	0/1(99)	0/1(99)	0/1(29)	0/1(99)	0/1(99)	0/1(70)	0/1(21)
PCMTD1	D	C	T	Non-synonymous: missense	0/1(99)	0/1(99)	0/1(99)	0/1(99)	0/1(99)	0/1(99)	0/1(99)

0 for reference allele and 1 for the variant allele; R, recessive; D, dominant; PB, peripheral blood; SA, saliva;

Figure 1

Figure 2

Figure 3

Figure 4