

Whole exome analysis of primary immunodeficiency

E.S. Rahmani¹, H. Azarpara², M. Karimipour³, H. Rahimi³ 

¹ Department of Medical Biotechnology, College of Allied Medicine, Iran University of Medical Sciences, Tehran, Iran

² Department of Medicine, Iran University of Medical Sciences, Tehran, Iran

³ Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran

The human primary immunodeficiency diseases (PIDs) refer to a rare heterogeneous group of single-gene inherited disorders causing malfunctions in the immune system, and thus the affected patients have a predisposition to severe life-threatening infections. The heterogeneous nature of PIDs, which involves at list 300 different genes, makes diagnosis of the disease a complex issue. Although studies revealed that six million people have a kind of PID, but due to a complex diagnosis procedure many affected individuals have not gotten a correct diagnosis. However, thanks to advancing in the DNA sequencing method and availability of sophisticated sequencers molecular characterization of genetic disorders have been revolutionized. The whole exome sequencing (WES) method can help clinicians detect Mendelian disease and other complex genetic disorders. The presented study used WES to investigate two infants with symptoms of primary immunodeficiency including hemophagocytic lymphohistiocytosis (HLH) and severe combined immunodeficiency (SCID). It has been shown that the HLH patient had a mutation in the UNC13D gene (NM_199242.2:c.627delT), and the SCID patient had a mutation in the RAG1 gene (NM_000448.2:c.322C>G). It has been demonstrated that WES is a fast and cost-effective method facilitating genetic diagnosis in PID sufferers.

Key words: primary immunodeficiency; next-generation sequencing; UNC13D; RAG1.

Полноэкзомный анализ первичного иммунодефицита

Э.Ш. Рахмани¹, Х. Азарпара², М. Каримипур³,
Х. Рахими³ 

¹ Факультет медицинских биотехнологий, Школа комплементарной медицины, Иранский университет медицинских наук, Тегеран, Иран

² Факультет медицины, Иранский университет медицинских наук, Тегеран, Иран

³ Факультет молекулярной медицины, Исследовательский центр биотехнологий, Институт Пастера Ирана, Тегеран, Иран

Первичный иммунодефицит представляет собой гетерогенную группу редких наследственных мутаций единичного гена, вызывающих сбой в работе иммунной системы человека и проявляющихся в предрасположенности пациентов к тяжелым жизнеугрожающим инфекциям. Гетерогенная природа первичного иммунодефицита, при которой мутации могут быть подвержены по меньшей мере 300 различных генов, серьезно осложняет его диагностику. И хотя было подсчитано, что от этого заболевания могут страдать около шести миллионов человек, только немногие из них могут рассчитывать на постановку правильного диагноза. Однако развитие методов секвенирования ДНК и доступность высокотехнологичного оборудования позволили сделать значительный шаг вперед в области молекулярных исследований генетических заболеваний. Технология полноэкзомного анализа ДНК может оказать существенную помощь врачам при диагностировании менделевских предрасположенностей к микробактериальным инфекциям и других форм редких генетических заболеваний. В представленном исследовании мы использовали метод полноэкзомного анализа ДНК для обследования двух младенцев с симптомами первичного иммунодефицита, такими как гемофагоцитарный лимфогистиоцитоз (ГЛГ) и тяжелый комбинированный иммунодефицит (ТКИД). Полноэкзомный анализ выявил мутацию UNC13D гена (NM_199242.2:c.627delT) у пациента с ГЛГ и мутацию RAG1 гена (NM_000448.2:c.322C>G) – у пациента с ТКИД. Исследование показало, что полноэкзомный анализ – это быстрый и экономичный метод, помогающий поставить правильный диагноз пациентам с первичным иммунодефицитом.

Ключевые слова: первичный иммунодефицит; секвенирование нового поколения; UNC13D; RAG1.

HOW TO CITE THIS ARTICLE:

Rahmani E.S., Azarpara H., Karimipour M., Rahimi H. Whole exome analysis of primary immunodeficiency. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2018;22(5):620-626. DOI 10.18699/VJ18.403

Received 09.04.2018

Accepted for publication 28.06.2018

© AUTHORS, 2018

These authors contributed equally to this work.

Human primary immunodeficiencies (PIDs) include at least 300 genetically-defined single-gene inherited disorders causing malfunctions in the immune system (Bousfiha et al., 2015). Recent studies have revealed that six million people worldwide have a kind of PID, whereas only 27000–60000 individual have been diagnosed. Despite the fact that the prevalence of PIDs is highest among children, there are also many adult patients with PIDs (Bousfiha et al., 2013). Patients with PIDs are predisposed to severe infections such as the EBV, Neisseria, Papillomavirus, Streptococcus pneumoniae, weakly virulent mycobacteria, Herpes simplex virus, and Candida Albicans. Furthermore, the immune dysregulation and aberrant inflammatory responses (Abolhassani et al., 2014) such as Allergy, Angioedema, Hemophagocytosis, Autoinflammation, and Autoimmunity have also been diagnosed in such patients (Bousfiha et al., 2013). PIDs are classified into nine groups based on the clinical and laboratory parameters (Picard et al., 2015). The first group includes the most common combined T- and B-cell immunodeficiency (CID) and a severe PID form known as severe combined immunodeficiency (SCID) (Bonilla et al., 2005). SCID is characterized by profound defects of T-cell development and it affects some of the B and NK cells (Kwan, Puck, 2015). Hemophagocytic lymphohistiocytosis (HLH) is a heterogeneous group of disorders related to dysregulation of the immune system (PID class 4) that are classified into two groups, namely genetic (familial HLH (FHL)) and acquired forms based on the etiology. FHL includes five forms of loss-of-function mutation (FHL1-5) leading to defects in the cytotoxic granule secretion pathway in NK and CD8⁺ T-cells, and they consequently lead to failure in exocytosis of granules in immunologic synapsis and completely eradicate target cells, in situation of immune response (Sifers et al., 2016). HLH can also occur due to infections and autoinflammatory/autoimmune and malignant diseases which are known as the acquired forms (Janka, 2012).

Current procedures for PID diagnosis are very complex and involve using specialized immunologic tests, including lymphocyte proliferation and cytotoxic assay, evaluation of serum immunoglobulin level, flow cytometry, neutrophil function assays and complementary analysis (McCusker, Warrington, 2011). However, immunological evaluation is performed to assess a patient's immune status for primary PID diagnosis. In addition, phenotype – based PID diagnosis is often complex, expensive and not always successful. Genetic investigation of PIDs is also very complex, and covers more than 300 genes that may be involved. Allelic heterogeneity and locus heterogeneity also increase the complexity of genetic analysis (Moens et al., 2014; Stoddard et al., 2014).

Advances in next-generation sequencing (NGS), particularly in whole exome sequencing (WES) have revolutionized molecular diagnosis of Mendelian disorders (Gilissen et al., 2012), and thus the traditional methods can be replaced by interrogation of a large set of genes in the single test in a timely and cost-effective manner instead of gene-by-gene approaches (Stoddard et al., 2014; Vrijenhoek et al., 2015). Thirty four new gene defects in PIDs have been diagnosed using NGS technology, so far (Conley, Casanova, 2014). The most recent research by S. Tamura et al. (2015) described the identification of novel compound heterozygous mutation in the DNA ligase IV (LIG4) gene through WES.

The present study was aimed at using WES for molecular characterization of two families with PID-affected children to be confirmed by the Sanger sequencing. We found a mutation in UNC13D and RAG1 genes in HLH and SCID families respectively. Our results have proved WES to be a useful method for detecting pathogenic variants in PID sufferers.

Patients, materials, and methods

Patients. The patients selected for molecular characterization using the whole exome sequencing were two children from two Iranian families in consanguineous marriage hospitalized in Children's medical center (Tehran) and their PIDs were confirmed with specialized immunology methods. The first patient was a six-month female infant with clinical symptoms of HLH, and the second - a two-month male infant diagnosed with SCID. The HLH patient had prolonged fever, hepatosplenomegaly, an infection caused by Epstein-Barr virus (EBV) and increased levels of ferritin, while the SCID patient suffered of recurrent diarrhea, respiratory infections without any circulating T and B cells (T⁺ B⁻ SCID). For the children could participate in the study, the parents had signed an ethical consent form.

Blood sampling. Peripheral blood samples were taken from the children and their family members. The clinical information obtained from their medical records included date and year of diagnosis, disease class and severity, surgical history, medication and family history. Genomic DNA was extracted using salting-out protocol (Miller et al., 1988).

Whole exome sequencing and data analysis. One hundred ng/μl of high-quality genomic DNA was first used for whole exome enrichment in the IonAmpliSeq Exome RDY plates and Ion AmpliSeq HiFi Mix. After ligation of Proton adapters and quantification by qPCR, the final library was sequenced using the Ion Proton platform which produced raw FASTQ files at an average coverage depth of 50X.

The following general workflow was used for performing bioinformatics analysis of the FASTQ raw data to prioritize causative variants (Fig. 1).

The raw FASTQ files were processed using the NGS QC toolkit (Patel, Jain, 2012) to estimate the quality and states of sequence reads. Sequencing reads, which have some errors such as adaptor and primer contamination, low quality 5' and 3' end bases, short reads and those with quality scores (Phred score) below 20, were trimmed from the FASTQ files using the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). BWA-MEM algorithm (Caboche et al., 2014) was used to align reads against the reference genome sequence (GRCh37), and then the results were stored in the SAM (Sequence Alignment/Map) file format. BWA-MEM is fast and accurate alignment software for nucleotide sequences of about 70 bp–1 Mbp. Duplicate reads were marked by Picard tools (<https://broadinstitute.github.io/picard/>). Afterwards, the SNP and InDel calling was performed with the Genome Analysis Toolkit (GATK; v 3.6) (Van der Auwera et al., 2013). The functional variant annotation was performed using the following software and databases: ANNOVAR (Wang et al., 2010), KGGSeq (v1.0) (Li et al., 2012), the 1000-Genome Project (www.1000genomes.org/), NHLBI GO Exome Sequencing Project (ESP), Exome Variant Server (EVS) (<http://evs.gs.washington.edu/EVS/>), Exome Aggregation Consortium

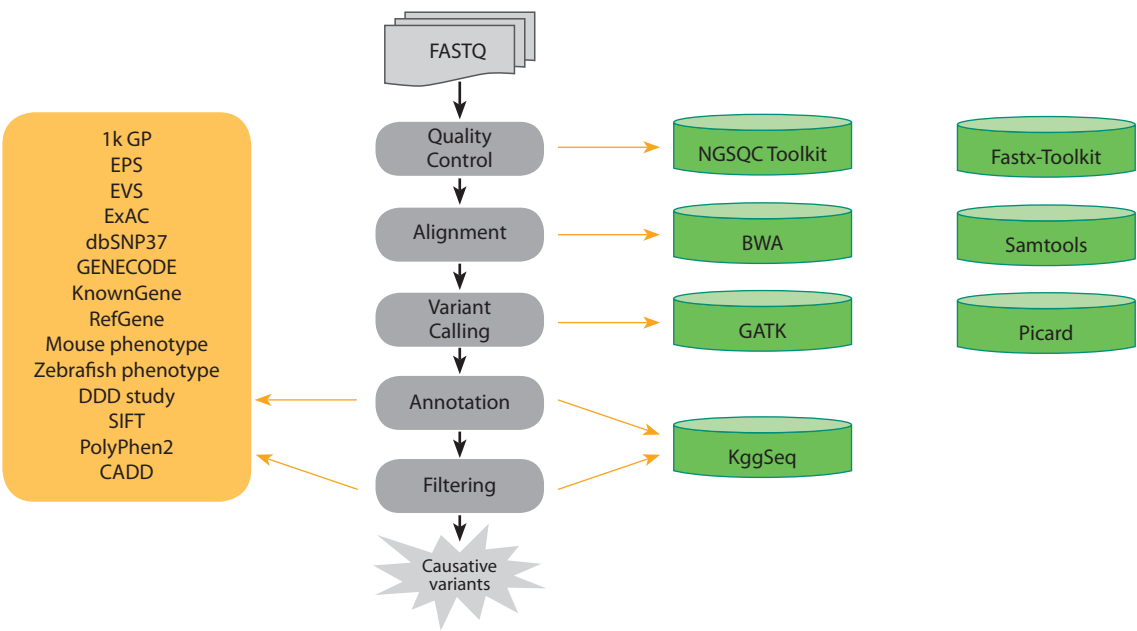


Fig. 1. WES data analysis flowchart.
The green circles denote software, while the orange box shows all the databases and algorithms provided by KGGSeq that were used for annotation and filtering.

Table 1. PCR primers

Sample	Chr	Pos	Ref	Alt	Primer-Forward	Tm, °C	Primer-Reverse	Tm, °C	Product length, bp
HLH	17	73836899	A	–	AGGTATGGAAGGGAAGGGATC	60.3	GCACCCAGCATCCAGTGTG	63.1	257
SCID	11	36595176	C/G	T	GGACTTGTTCATTGTTCTCAG	54.9	CGAGTCAACATCTGCCTTCAC	58.1	544

Chr – chromosome; Pos – position; Ref – reference sequence; Alt – alternative sequence.

(ExAC) (<http://exac.broadinstitute.org>; release 0.3), dbSNP, GENECODE (Harrow et al., 2012), knownGene, RefGene (UCSC), mouse phenotype (Eppig et al., 2015) and DDD study (Deciphering Developmental Disorders) (Firth et al., 2009). The KGGSeq filtering strategy was executed to filter out both common benign variants and recurrent artifact as well as to find causal variants. At the first level, variant quality control was checked using the KGGSeq software, which includes various filters such as genotype QC, variants QC and sample QC to filter out errors and low-quality variants. A cut-off of 20, 50 and 20 was done for variants' Phred quality score, mapping quality score and depth coverage respectively. SIFT (Ng, Henikoff, 2003), PolyPhen-2 (Adzhubei et al., 2010) and CADD (Kircher et al., 2014) tools were used for predicting and scoring the possible impact of amino acid substitution on the structure and function of human proteins. Prioritization was performed with the focus on a primary immunodeficiency panel, which was gathered from the NCBI gene database (<http://www.ncbi.nlm.nih.gov/gene/>) and recommended genes from National Immunology Society resulting in 400 genes and genomic regions. The causative variants were limited to following criteria: 1) a minor allele frequency (MAF) of less than 0.1 % in data from the 1000-Genome Project, EVS, and ExAC, 2) minimum CADD score of 20,

and 3) high or moderate effect determined by SNPEff (v 4.2) (Cingolani et al., 2012).

Sanger sequencing validation. The chromosomal regions containing the candidate causative variants in patients were amplified by polymerase chain reaction (Eppendorf, Germany) with Taq DNA polymerase (Amplicon, Germany) and designed primer for each of them at PCR conditions as summarized in Table 1. The sequencing diagrams were obtained using the Chromas software (v 2.6) (<http://technelysium.com.au>).

Results

Molecular diagnosis of PIDs by means of whole exome sequencing is a cost-effective and valuable approach for patients. In the present project, WES was used to study two families with HLH and SCID symptoms respectively. After performing the sequencing, the total reads were about 44 and 55 million for HHL and SCID patients respectively (Table 2). QC analysis of the FASTQ files showed that the average length of reads consisted of 190 nucleotides and 51 % of GC content, and more than 86 present of reads had quality scores equal or more than the cut-off (Q20) (see Table 2). After QC checking and trimming, the trimmed files were mapped to human reference genome version of GRCh37. Alignment statistics report using Samtools (Li et al., 2009) indicated that 95 %

of the reads were mapped to reference genome in the both files (see Table 2). 49426 variants were observed in the HLH patient, and 49629 variants – in the SCID patient after variant calling (see Table 2). After using the KGGSeq filters, the observed variations for the two patients were 48838 (HLH) and 49181 (SCID) respectively. Allele frequency trimming performed using 1000G, ESP6500 and ExAc indicated that 30569 and 29968 variants had greater than 1 % of MAF in the HLH and SCID patients respectively. Functional effect prediction of variants was performed using the SIFT, Polyphen-2 and CADD software resulting in 80 pathogenic and deleterious variants in the both patients. Top 10 variants with higher scores are presented in Table 3. Finally, two variants were considered as potential causative variants for the above-mentioned phenotype of HLH and SCID. Selected variants included a frameshift variant in Exon 8 of the UNC13D gene (NM_199242.2:c.627delT) and a premature stop codon variant in Exon 2 of the RAG1 gene (NM_000448.2:c.322C>G) in the HLH and SCID patient respectively (Table 4). Sanger sequencing confirmed the both variants (Fig. 2).

Table 2. Quality control and alignment statistics

Parameter	HLH	SCID
Total reads (Mb)	44.509	55.250
Total bases (Gb)	8.457	10.446
Average length of read (bp)	190	189
Q20	86.57 %	86.15 %
% GC	51.51 %	51.92 %
Mapped reads	44354714	55004005
Total aligned base reads	8307987731	10246246706
Percent reads on target	95.12 %	94.72 %
Percent base reads on target	94.22 %	93.94 %
Total variant count	49248	49456
SNV	46455	46670
INDEL	2793	2786

Q20 – percent of base number calls with quality value of 20 or higher; % GC – percentage of GC content; SNV – single nucleotide variant; INDEL – insertion and deletion; CNV – copy number variation.

Table 3. Top 10 variants with high scores for both patients

Chr	Pos	Allele	Gene	Type	Score	Causality
HLH						
19	57326850	T/C	PEG3	Missense	0.9967	Damaging
17	61601686	G/T	KCNH6	»	0.9518	»
16	2577847	C/G	AMDHD2	»	0.9176	»
12	121176083	G/A	ACADS	»	0.9158	»
2	189918622	G/A	COL5A2	»	0.8126	»
1	2938989	G/A	ACTRT2	»	0.7911	»
17	21319121	C/T	KCNJ12	»	0.7460	»
10	14862082	C/G	CDNF	»	0.7220	»
11	5758062	T/C	OR56B1	»	0.6986	»
14	20692643	T/C	OR11H6	»	0.6830	»
SCID						
17	61601686	G/T	KCNH6	»	0.9518	»
12	121176083	G/A	ACADS	»	0.9158	»
18	61387333	T/A	SERPINB11	»	0.9145	»
1	94466628	G/A	ABCA4	»	0.8452	»
10	54531235	C/T	MBL2	»	0.7919	»
16	84456014	C/T	ATP2C2	»	0.7698	»
7	158591753	G/A	ESYT2	»	0.7685	»
1	11856378	G/A	MTHFR	»	0.7582	»
14	24707479	G/A	GMPR2	»	0.7563	»
17	21319121	C/T	KCNJ12	»	0.7460	»

Table 4. Potential causative variants

Sample	Chromosome	Position	Damaging variant	Reference allele	Alternative allele	Type	rsID
HLH	17	73836899	Exon 8:c.627 del T	A	–	Homozygous	rs755619812
SCID	11	36595176	Exon 2: c.322 C>T	C/G	T	Homozygous	rs193922464

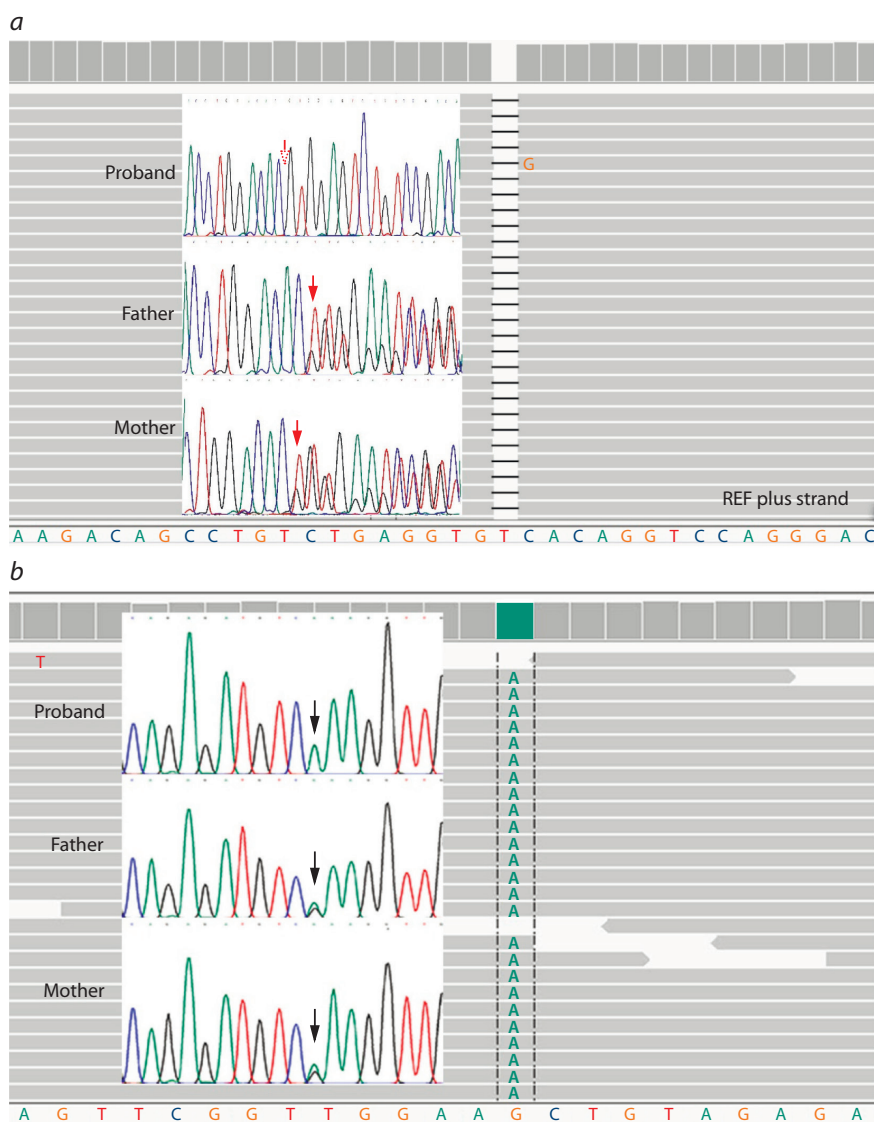


Fig. 2. Sanger sequencing confirmation of considered variants in HLH and SCID patients:

a – in the HLH patient, IGV screenshot shows a homozygous deletion in the UNC13D gene confirmed both in the proband and the parents; *b* – IGV screenshot from the SCID patient also shows a homozygous single nucleotide substitution in the RAG1 gene confirmed both in the proband and the parents.

Discussion

Molecular diagnosis of PIDs with at least 300 genes, allelic and locus heterogeneity is a challenging issue. Recent studies indicate that 121 gene defects have been identified in these diseases in addition to the list of genes involved in PID (179 genes) since 2011 (Al-Herz et al., 2011; Picard et al., 2015). Utilization of the Sanger sequencing as a direct sequencing method for finding mutations in genes has become a de facto issue in genetic diagnosis (Sanger et al., 1977). Despite the fact that the gene-by-gene approach is a conventional diagnostic test for monogenic diseases, such methods are too expensive and time-consuming in multigenic diseases such as the PIDs (Chou et al., 2012; Sikkema-Raddatz et al., 2013). In contrast, WES method has become a favorable test for genetic diagnosis of multigenic diseases and it sequences all coding regions using amplification and parallel sequencing in a single test (Sikkema-Raddatz et al., 2013; Stoddard et al., 2014). For example, E. Mukda et al. (2017) evaluated 25 HLH patients and diagnosed pathogenic mutations in PFR1, UNC13D, STXBP2, LYST and XIAP genes. G.S. Schulert et al. (2015) also conducted WES in 16 HLH patients and diagnosed the disease-causing mutations in PFR1 and LYST gene. CARD11 gene inactivation due to a premature stop codon was diagnosed in

a SCID patient by means of WES (Greil et al., 2013). Accordingly, the present study sought to obtain a molecular diagnosis of two patients with HLH and SCID by means of WES. A variant (rs755619812) in UNC13D gene and a variant (rs193922464) in RAG1 gene were detected in HLH and SCID patients respectively after exome sequencing and processing the raw data. In HLH case, the UNC13D gene (17q25.1) comprised 32 exons and 4.5-kb transcript with the highest expression in spleen, thymus, and peripheral blood leukocytes. In lymphocytes, the encoded protein (Munc13-4, 1090 amino acids, UniProt ID: Q70J99) played an important role in cytotoxic granule exocytosis. Munc13-4 is an essential protein for maturation, docking, and priming of cytotoxic granules in cytotoxic cells (CTLs and NK) (Feldmann et al., 2003). Pathogenic mutations in the UNC13D gene led to an ineffective protein create type 3 of familial HLH (FHL-3). We found a homozygous nucleotide deletion (c.627delT:p.V210Wfs*39) in exon 8 of UNC13D, and it caused a frameshift mutation (see Fig. 2, a). This damaging mutation led to the substitution of valine with tryptophan at position 210 resulting in a premature stop codon at 39 codons after this substitution which produced a truncated protein. This mutation results in defects in the killing ability of NK and CD8⁺ T-cells and uncontrolled hyperinflammation in HLH patients. In 2006, Stasdt et al. studied 63 HLH patients and described the c.627delT mutation in two infants in order to find the mutations spectra of the PFR1, UNC13D, STX11 and RAB27A genes (Stadt et al., 2006).

In the second patient with clinical symptoms of SCID, we found a homozygous single nucleotide substitution in the first position of Arginine codon 108 in Exon 2 of the RAG1 gene (c.322C>G: R108*) leading to a premature stop codon and consequently a truncated protein (see Fig. 2, b). The human RAG1 gene (11p12) consists of 2 exons and encodes a protein with 1043 amino acids (UniProt ID: P15918). Recombination-activating protein 1 (RAG1) is a catalytic component of the RAG complex as a multi-protein complex that mediates the DNA cleavage phase during V(D)J recombination. In the RAG complex (RAG1/2), RAG1 mediates the DNA-binding to the conserved recombination

signal sequences (RSS) and catalyzes DNA cleavage activities by introducing a double-strand break between RSS and the adjacency to each coding V, D, and J DNA segment (Melek, Gellert, 2000). This somatic recombination leads to the diversity of immunoglobulins and T-cell receptors (TCRs). The pathogenic mutations affecting the active core (amino acids 384-1009) of RAG1 produce the clinical symptoms of SCID disease such as absence of lymphocyte (B- and T-cells) circulation (Corneo et al., 2001) observed in our patient.

Conclusion

Whole exome sequencing has proved itself as a fast and cost-effective method for detection of causative rare mutations in PID patients.

Acknowledgements

We would like to thank the families who participated in this project as well as Pasteur Institute of Iran for supporting this project (grant No. 756).

Conflict of interest

The authors declare they have no conflict of interest.

References

- Abolhassani H., Wang N., Aghamohammadi A., Rezaei N., Lee Y.N., Frugoni F., Notarangelo L.D., Pan-Hammarström Q., Hammarström L. A hypomorphic recombination-activating gene 1 (RAG1) mutation resulting in a phenotype resembling common variable immunodeficiency. *J. Allergy Clin. Immunol.* 2014;134(6):1375-1380. DOI 10.1016/j.jaci.2014.04.042.
- Adzhubei I.A., Schmidt S., Peshkin L., Ramensky V.E., Gerasimova A., Bork P., Kondrashov A.S., Sunyaev S.R. A method and server for predicting damaging missense mutations. *Nat. Methods.* 2010;7(4):248-249. DOI 10.1038/nmeth0410-248.
- Al-Herz W., Bousfiha A., Casanova J.-L., Chapel H., Conley M.E., Cunningham-Rundles C., Etzioni A., Fischer A., Franco J.L., Geha R.S., Hammarström L., Nonoyama S., Notarangelo L.D., Ochs H.D., Puck J.M., Roifman C.M., Seger R., Tang M.L. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. *Front. Immunol.* 2011;2:54. DOI 10.1007/s10875-015-0201-1.
- Bonilla F.A., Bernstein I.L., Khan D.A., Ballas Z.K., Chinen J., Frank M.M., Kobrynski L.J., Levinson A.I., Mazer B., Nelson R.P., Orange J.S., Routes J.M., Shearer W.T., Sorensen R.U. Practice parameter for the diagnosis and management of primary immunodeficiency. *Ann. Allergy Asthma Immunol.* 2005;94(5):S1-S63. DOI 10.1016/S1081-1206(10)61142-8.
- Bousfiha A.A., Jeddane L., Ailal F., Benhsaien I., Mahlaoui N., Casanova J.L., Abel L. Primary immunodeficiency diseases worldwide: more common than generally thought. *J. Clin. Immunol.* 2013;33(1):1-7. DOI 10.1007/s10875-012-9751-7.
- Bousfiha A., Jeddane L., Al-Herz W., Ailal F., Casanova J.L., Chatila T., Conley M.E., Cunningham-Rundles C., Etzioni A., Franco J.L., Gaspar H.B., Holland S.M., Klein C., Nonoyama S., Ochs H.D., Oksenhendler E., Picard C., Puck J.M., Sullivan K.E., Tang M. The 2015 IUIS phenotypic classification for primary immunodeficiencies. *J. Clin. Immunol.* 2015;35(8):727-738. DOI 10.1007/s10875-015-0198-5.
- Caboche S., Audebert C., Lemoine Y., Hot D. Comparison of mapping algorithms used in high-throughput sequencing: application to Ion Torrent data. *BMC Genomics.* 2014;15(1):1. DOI 10.1186/1471-2164-15-264.
- Chou J., Ohsumi T.K., Geha R.S. Use of whole exome and genome sequencing in the identification of genetic causes of primary immunodeficiencies. *Curr. Opin. Allergy Clin. Immunol.* 2012;12(6):623-628. DOI 10.1097/ACI.0b013e3283588ca6.
- Cingolani P., Platts A., Wang L.L., Coon M., Nguyen T., Wang L., Land S.J., Lu X., Ruden D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly.* 2012;6(2):80-92. DOI 10.4161/fly.19695.
- Conley M.E., Casanova J.-L. Discovery of single-gene inborn errors of immunity by next generation sequencing. *Curr. Opin. Immunol.* 2014;30:17-23. DOI 10.1016/j.coi.2014.05.004.
- Corneo B., Moshous D., Gungör T., Wulffraat N., Philippot P., Le Deist F.L., Fischer A., de Villartay J.P. Identical mutations in RAG1 or RAG2 genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome. *Blood.* 2001;97(9):2772-2776. DOI 10.1182/blood.V97.9.2772.
- Eppig J.T., Blake J.A., Bult C.J., Kadin J.A., Richardson J.E. Mouse Genome Database Group. The Mouse Genome Database (MGD): Facilitating mouse as a model for human biology and disease. *Nucleic Acids Res.* 2015;43:D726-D736. DOI 10.1093/nar/gku967.
- Feldmann J., Callebaut I., Raposo G., Certain S., Bacq D., Dumont C., Lambert N., Ouachée-Chardin M., Chedeville G., Tamary H., Minard-Colin V. Munc13-4 is essential for cytolytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). *Cell.* 2003;115(4):461-473. DOI 10.1016/S0092-8674(03)00855-9.
- Firth H.V., Richards S.M., Bevan A.P., Clayton S., Corpes M., Rajan D., Van Voreen S., Moreau Y., Pettett R.M., Carter N.P. DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am. J. Hum. Genet.* 2009;84(4):524-533. DOI 10.1016/j.ajhg.2009.03.010.
- Gilissen C., Hoischen A., Brunner H.G., Veltman J.A. Disease gene identification strategies for exome sequencing. *Eur. J. Hum. Genet.* 2012;20(5):490-497. DOI 10.1038/ejhg.2011.258.
- Greil J., Rausch T., Giese T., Bandapalli O.R., Daniel V., Bekeredjian-Ding I., Stütz A.M., Drees C., Roth S., Ruland J., Korbel J.O., Kulzik A.E. Whole-exome sequencing links caspase recruitment domain 11 (CARD11) inactivation to severe combined immunodeficiency. *J. Allergy Clin. Immunol.* 2013;131(5):1376-1383. e1373. DOI 10.1016/j.jaci.2013.02.012.
- Harrow J., Frankish A., Gonzalez J.M., Tapanari E., Diekhans M., Kokocinski F., Aken B.L., Barrell D., Zadissa A., Searle S., Barnes I., Bignell A., Boychenko V., Hunt T., Kay M., Mukherjee G., Rajan J., Despacio-Reyes G., Saunders G., Steward C., Harte R., Lin M., Howald C., Tanzer A., Derrien T., Chrast J., Walters N., Balasubramanian S., Pei B., Tress M., Rodriguez J.M., Ezkurdia I., van Baren J., Brent M., Haussler D., Kellis M., Valencia A., Reymond A., Gerstein M., Guigó R., Hubbard T.J. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 2012;22(9):1760-1774. DOI 10.1101/gr.135350.111.
- Janka G. Familial and acquired hemophagocytic lymphohistiocytosis. *Annu. Rev. Med.* 2012;63:233-246. DOI 10.1146/annurev-med-041610-134208.
- Kircher M., Witten D.M., Jain P., O'Roak B.J., Cooper G.M., Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 2014;46(3):310. DOI 10.1038/ng.2892.
- Kwan A., Puck J.M. Newborn screening for severe combined immunodeficiency. *Curr. Pediatr. Rep.* 2015;3(1):34-42. DOI 10.1007/s40124-014-0068-2.
- Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079. DOI 10.1093/bioinformatics/btp352.
- Li M.-X., Gui H.-S., Kwan J.S., Bao S.-Y., Sham P.C. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. *Nucleic Acids Res.* 2012;gkr1257.

- McCusker C., Warrington R. Primary immunodeficiency. *Allergy Asthma Clin. Immunol.* 2011;7(1):1. DOI 10.1186/1710-1492-7-S1-S11.
- Melek M., Gellert M. RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. *Cell.* 2000;101(6):625-633. DOI 10.1016/S0092-8674(00)80874-0.
- Miller S.A., Dykes D.D., Polesky H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16:1215. DOI 10.1093/nar/16.3.1215.
- Moens L.N., Falk-Sörqvist E., Asplund A.C., Bernatowska E., Smith C.E., Nilsson M. Diagnostics of primary immunodeficiency diseases: a sequencing capture approach. *PloS One.* 2014;9(12):e114901. DOI 10.1371/journal.pone.0114901.
- Mukda E., Trachoo O., Pasomsub E., Tiyasirichokchai R., Iemwimangsa N., Sosothikul D., Chantaratita W., Pakakasama S. Exome sequencing for simultaneous mutation screening in children with hemophagocytic lymphohistiocytosis. *Int. J. Hematol.* 2017;1-9. DOI 10.1007/s12185-017-2223-3.
- Ng P.C., Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003;31(13):3812-3814. DOI 10.1093/nar/gkg509.
- Patel R.K., Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PloS One.* 2012;7(2):e30619. DOI 10.1371/journal.pone.0030619.
- Picard C., Al-Herz W., Bousfiha A., Casanova J.-L., Chatila T., Conley M.E., Cunningham-Rundles C., Etzioni A., Franco J.L., Gaspar H.B., Holland S.M., Klein C., Nonoyama S., Ochs H.D., Oksenhendler E., Picard C., Puck J.M., Sullivan K., Tang M.L. Primary immunodeficiency diseases: an update on the classification from the International Union of Immunological Societies Expert Committee for primary immunodeficiency 2015. *J. Clin. Immunol.* 2015;35(8):696-726. DOI 10.1007/s10875-015-0201-1.
- Sanger F., Nicklen S., Coulson A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci USA.* 1977;74(12):5463-5467. DOI 10.1073/pnas.74.12.5463.
- Schulert G.S., Zhang M., Fall N., Husami A., Kissell D., Hanosh A., Zhang K., Davis K., Jentzen J.M., Napolitano L., Siddiqui J., Smith L.B., Harms P.W., Grom A.A., Cron R.Q. Whole-exome sequencing reveals mutations in genes linked to hemophagocytic lymphohistiocytosis and macrophage activation syndrome in fatal cases of H1N1 influenza. *J. Infect. Dis.* 2015;213(7):1180-1188. DOI 10.1093/infdis/jiv550.
- Sifers T.M., Raje N., Dinakar C. Hemophagocytic lymphohistiocytosis: A concise review for the practicing physician. *Allergy Asthma Proc.* 2016;37(3):256-258. DOI 10.2500/aap.2016.37.3948.
- Sikkema-Raddatz B., Johansson L.F., Boer E.N., Almomani R., Boven L.G., van den Berg M.P., van Spaendonck-Zwarts K.Y., van Tintelen J.P., Sijmons R.H., Jongbloed J.D., Sinke R.J. Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum. Mutat.* 2013;34(7):1035-1042. DOI 10.1002/humu.22332.
- Stadt U.Z., Beutel K., Kolberg S., Schneppenheim R., Kabisch H., Janka G., Hennies H.C. 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. DOI 10.1002/humu.20274.
- Stoddard J.L., Niemela J.E., Fleisher T.A., Rosenzweig S.D. Targeted NGS: a cost-effective approach to molecular diagnosis of PIDs. *Front. Immunol.* 2014;5:531. DOI 10.3389/fimmu.2014.00531.
- Tamura S., Higuchi K., Tamaki M., Inoue C., Awazawa R., Mitsuki N., Nakazawa Y., Mishima H., Takahashi K., Kondo O., Imai K., Morio T., Ohara O., Ogi T., Furukawa F., Inoue M., Yoshiura K., Kanazawa N. Novel compound heterozygous DNA ligase IV mutations in an adolescent with a slowly-progressing radiosensitive-severe combined immunodeficiency. *Clin. Immunol.* 2015;160(2):255-260. DOI 10.1016/j.clim.2015.07.004.
- Van der Auwera G.A., Carneiro M.O., Hartl C., Poplin R., del Angel G., Levy-Moonshine A., Jordan T., Shakir K., Roazen D., Thibault J., Banks E., Garimella K.V., Altshuler D., Gabriel S., DePristo M.A. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics.* 2013;11:11.10.11-11.10.33. DOI 10.1002/0471250953.bi1110s43.
- Vrijenhoek T., Kraaijeveld K., Elferink M., De Ligt J., Kranendonk E., Santen G., Nijman I.J., Butler D., Claes G., Costessi A., Dorlijn W., van Eyndhoven W., Halley D.J., van den Hout M.C., van Hove S., Johansson L.F., Jongbloed J.D., Kamps R., Kockx C.E., de Koning B., Kriek M., Lekanne Dit Deprez R., Lunstroo H., Mannens M., Mook O.R., Nelen M., Ploem C., Rijnen M., Saris J.J., Sinke R., Sistermans E., van Slegtenhorst M., Sleutels F., van der Stoep N., van Tienhoven M., Vermaat M., Vogel M., Waisfisz Q., Marjan Weiss J., van den Wijngaard A., van Workum W., Ijntema H., van der Zwaag B., van IJcken W.F., den Dunnen J., Veltman J.A., Hennekam R., Cuppen E. Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects. *Eur. J. Hum. Genet.* 2015;23(9):1142-1150. DOI 10.1038/ejhg.2014.279.
- Wang K., Li M., Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164. DOI 10.1093/nar/gkq603.