Ann Med Res 2021;28(5):1038-43

Annals of Medical Research

DOI: 10.5455/annalsmedres.2021.03.201

Original Article

Investigation of causative genetic defects in patients with primary immunodeficiency by next generation sequencing

Baran Erman[1], Tugba Arikoglu[2], Aylin Kont Ozhan[2], Ali Demirhan[2], Nazan Tokmeci[2], Cigdem Aydogmus[3], Caner Aytekin[4]

[1]Institute of Child Health, Can Sucak Research Laboratory for Translational Immunology, Hacettepe University, Ankara, Turkey
[2]Department of Pediatrics, Division of Pediatric Allergy and Immunology, Mersin University Faculty of Medicine, Mersin, Turkey
[3]Department of Pediatric Allergy and Immunology, Basaksehir Cam and Sakura City Hospital, Istanbul, Tukey
[4]Department of Pediatric Immunology, Dr. Sami Ulus Maternity and Childrens Health and Diseases Training and Research Hospital, Ankara, Turkey

Abstract

Aim: Inborn errors of immunity are rare diseases presented with a broad range of clinical symptoms. There are more than 450 causative genetic defects and the genetic diagnosis is very important for the patients. Use of next generation sequencing facilitated the molecular and genetic identification of these diseases for last 10 years. We aimed to search disease-causing defects in patients with primary immunodeficiencies (PIDs) by next generation sequencing.

Materials and Methods: The study included 12 PID patients without genetic diagnosis. We performed whole exome sequencing for the investigation of genetic defects and Sanger sequencing for variant validation.

Results: We found seven different disease-causing mutations in 6 patients with a diagnosis rate of 50%. There were three known pathogenic variants in *CYBA, SBDS* and *RAG2* genes. We identified two additional new causative variants in *NCF2* and *SBDS* genes and two novel mutations in *PGM3* and *SAMD9L* genes.

Conclusion: The result revealed that NGS-based methods especially whole exome sequencing can be used efficiently for genetic diagnosis of primary immunodeficiency diseases.

Keywords: Genetic diagnosis; next generation sequencing; primary immune deficiencies

INTRODUCTION

Inborn errors of immunity (IEI) or primary immunodeficiencies (PIDs) are a congenital group of disorders emerged by mutations in genes encoding immune system components. The most common clinical manifestations are susceptibility to infections, malignancies, autoimmunity and allergy (1). To date, more than 450 different genetic etiologies of PIDs have been identified (2). Most of them are monogenic disorders and inherited autosomal recessive. Primary immunodeficiencies are rare disorders but the prevalence of these diseases is relatively high in some countries like Turkey due to frequent consanguineous marriages (3). Therapeutic options for PIDs can be supportive treatment of clinical symptoms or targeted pharmacological approaches. However, for most PIDs, hematopoietic stem cell transplantation (HSCT) or gene therapies are the only definite treatment options (4).

Next generation sequencing (NGS) has been using effectively for last ten years. That is high-throughput sequencing technique which allows to determine the nucleotide sequence of whole genome, whole exome or a targeted nucleic acid region (5). Next generation based methods allow rapid genetic diagnosis for different congenital disorders like primary immunodeficiencies (6). Hence, there is a growing number of newly identified PIDs every year in the literature (2). Since there is a broad range of clinical presentations and life-threatening infections can emerge very early in life, rapid genetic diagnosis is very important and may be life-saving for immunodeficient patients.

In this study, we used whole exome sequencing to search genetic defects in patients with primary immunodeficiency.

MATERIALS and METHODS

Study population

Twelve genetically undiagnosed PID patients were included in the study (a). The patients were diagnosed as PID in different clinical immunology departments according to their clinical and laboratory findings in

Received: 01.03.2021 **Accepted:** 30.03.2021 **Available online:** 25.05.2021
Corresponding Author: Baran Erman, Institute of Child Health, Can Sucak Research Laboratory for Translational Immunology, Hacettepe University, Ankara, Turkey **E-mail:** baranerman@gmail.com

accordance with the fulfilling criteria of the European Society for Immunodeficiencies (ESID) (7).

The study was approved by the Ethics Committee of Hacettepe University (GO 20/407). Blood samples were obtained in accordance with the Declaration of Helsinki. Informed consent forms were provided from patients and parents.

Sample preparation
Genomic DNA was isolated from peripheral blood using the GenEx Blood, Cell, Tissue Kit (GeneAll Biotechnology). DNA concentrations were determined on Qubit 2.0 Fluorometer (Thermo).

DNA enrichment, whole exome sequencing and data analysis
Patient's DNA samples were enriched by The Nextera Exome Kit (Illumina). Then, 150 base pair paired-end-read sequencing was performed on Illumina NextSeq 550 system (Illumina). The raw data in BCL format was converted to FASTQ by BCL2FASTQ (Illumina) software. Aligment of the reads to human genome 19 (GRCh37/Hg19) and variant calling were performed by Dragen Joint Genotyping Pipeline v3.7.5 software (Illumina). Finally, variant annotation was done with wANNOWAR (Wang Genomics Lab.). Variants found in 1,000 Genome Project, ExAC, and GnomAD databases with a minor allele frequency of more than 0.01 were excluded, and the other variants were filtered for rare nonsense, missense, and indel changes within the exons and splice-site regions.

Variant validation
All possible disease-causing variants obtained from NGS data were validated by Sanger sequencing according to standard methods as previously described (8).

RESULTS

Patients
Among PID patients enrolled in the study, three were diagnosed as combined immunodeficiency (CID), two were diagnosed as chronic granulomatous disease (CGD) and two of them had severe congenital neutropenia. The other clinical diagnoses were one combined immunodeficiency (SCID), one hyper IgE syndrome (combined immunodeficiency with associated or syndromic features) and one ataxia pancytopenia syndrome. Two patients have been following as possible mendelian susceptibility to mycobacterial disease (MSMD) due to mycobacterium tuberculosis infection and intestinal granulomatous lesions in one and Salmonella infection with hypogammaglobulinemia in the other.

The age distribution was ranged between two months and 21 years. Gender of the five patients is female while seven of them are male. Nine patients were born to consanguineous parents (75%).

The clinical and laboratory findings of the patients are given in Table 1.

Whole exome sequencing
We sequenced all 12 samples in a single run and the entire study process including NGS library preparation, exome sequencing, data analyses and variant validation was completed in approximately three weeks. Sequence analysis viewer tool (SAV, Illumina) showed that the percentage of bases with a quality score of 30 or higher was 95.7%. In the same quality check, we found the number of clusters that passed Illumina's "Chastity filters" 96.4%.

Variant analyses
Exome sequencing revealed seven disease-causing variants in 6 patients. All causative variants were found in known PID genes (PGM3, SAMD9L, CYBA, NCF2, SBDS and RAG2). Of these 7 variants, three of them were known pathogenic mutations in CYBA, SBDS and RAG2 genes. Another two variants in NCF2 and SBDS genes are found in dbSNP database but they are considered as new likely disease-causing mutations. The other variants in the PGM3 and SAMD9L genes are novel mutations. The overall genetic diagnosis rate of the study is 50% (6/12).

We found a missense, homozygous, nonsynonymous PGM3 variant in patient 1. This patient has hyper IgE syndrome with severe atopy, viral infections and elevated serum IgE level. The missense variant (c.G214A:p.G72S) causes an amino acid change in codon 72. Another variant in the same cDNA position (c.C214A:p.G72C) has a reference SNP ID number (rs866671933). However, there is no information about pathogenic significance. The possible functional effects of the variant (c.G214A:p.G72S) were deleterious and damaging in SIFT and PolyPhen tools (SIFt score:0, PolyPhen score:1). In addition, Combined Annotation Dependent Depletion (CADD) score was 33. Glycine to serine change is located in a conserved region of catalytic serine domain of the protein. An identified missense mutation (p.Leu83Ser) in this domain from the study by Atfa Sassi et al. reduces catalytic activity and impaired protein function (9). Thus, c.G214A:p.G72S variant is considered a novel likely disease-causing mutation based on the respective patient's clinical phenotype.

Patient 2 has been following with hydrocephaly, cerebellar dysfunction, hypogammaglobulinemia and pancytopenia. We detected a missense single nucleotide variant (c.A2639C:p.H880P) leading to histidine-proline change in position 880 of SAMD9L protein. SAMD9L deficiency is inherited as autosomal dominant and causes ataxia pancytopenia syndrome in the patients. The detected variant in our patient have not been reported so far. However, Dong-Hui Chen et al. showed the other nucleotide change (c.2640C>A, p.His880Gln) affecting the same amino acid of the protein in a patient with ataxia-pancytopenia syndrome (10). Moreover, SIFT prediction is deleterious and the CADD score is 24.2. Consequently, this novel mutation explains our patient's immunodeficiency.

Table 1. Patient characteristics

Patient	Age	G	C	PID	WBC	ANC	ALC	CD3	CD4	CD8	CD19	NK	IgA	IgG	IgM
1	21 y	F	+	Hyper IgEsyndrome	3 (4.3-11.4)	**1.1**	**1.2** (1.7-5.7)	75 (58-82)	40 (26-48)	24 (16-32)	**6** (10-30)	10 (8-30)	101 (139-378)	1900 (913-1884)	**498** (88-322)
2	2 y	F	-	Ataxia pancytopenia	4.7 (4.4-12.9)	**0.06**	4.5 (2.2-8.1)	**90** (51-77)	54 (29-55)	**36** (15-33)	**0.2** (17-41)	8 (4-15)	**9** (26-296)	**229** (604-1941)	**11** (71-235)
3	2 y	F	+	CGD	**17.9** (4.4-12.9)	8.5	7.1 (2.2-8.1)	**40** (51-77)	32 (29-55)	**8** (15-33)	33 (17-41)	**16** (4-15)	36 (26-296)	1150 (604-1941)	81 (71-235)
4	11 y	M	+	CGD	10.4 (4.3-11.4)	8.1 (1.7-5.7)	1.7 (1.7-5.7)	63 (58-82)	27 (26-48)	**36** (16-32)	17 (10-30)	12 (8-30)	494 (67-433)	1790 (835-2094)	267 (47-484)
5	3 y	M	+	CID	**12.1** (4-10.4)	3.5	7.1 (2.6-10.4)	76 (54-76)	41 (31-54)	32 (12-28)	24 (15-39)	1 (3-17)	49 (26-296)	798 (463-1006)	40 (46-159)
6	2 m	F	+	CID	**16.5** (5.6-13.1)	5.2	9.2 (3.2-10.9)	51 (51-79)	43 (31-54)	10 (10-31)	**46** (14-44)	**2** (5-23)	29 (13-72)	**227** (294-1165)	**20** (33-154)
7	12 y	M	+	CID	7.6 (4.3-11.4)	5.8 (1.7-5.7)	**1.1** (1.7-5.7)	**52** (58-82)	**25** (26-48)	21 (16-32)	**3** (10-30)	**38** (8-30)	142 (67-433)	1240 (835-2094)	225 (47-484)
8	11 y	F	-	Congenital Neutrophenia	14.2 (4.3-11.4)	**0.3** (1.7-5.7)	**11.2** (1.7-5.7)	**49** (58-82)	43 (26-48)	**12** (16-32)	29 (10-30)	15 (8-30)	74 (67-433)	1830 (835-2094)	120 (47-484)
9	3 y	M	+	Congenital Neutrophenia	**3.1** (4-10.4)	**0.4**	2.2 (1.5-5.2)	78 (55-79)	44 (26-49)	31 (9-35)	17 (17-41)	**3** (4-15)	33 (26-296)	**452** (604-1941)	52 (52-297)
10	10 y	M	+	MSMD	11.1 (3.7-11.1)	8.2	1.5 (1.5-7.6)	60 (57-81)	37 (24-47)	18 (17-37)	27 (10-27)	8 (8-28)	126 (62-390)	1870 (842-1943)	344 (54-392)
11	4 m	M	-	MSMD	10.4 (5.6-13.1)	2.6	6.1 (4.4-12.9)	58 (51-79)	32 (31-54)	22 (10-31)	23 (14-44)	19 (5-23)	**5** (13-72)	**249** (294-1165)	**27** (33-154)
12	1 y	M	+	T- B- SCID	4.4 (4.4-12.9)	2.8	**0.4** (2.2-8.1)	61 (51-77)	40 (29-55)	20 (15-33)	**1** (17-41)	**35** (4-15)	**10** (30-107)	**426** (605-1430)	**24** (66-228)

G: Gender C: Consanguinity PID: Primary immunodeficiency WBC: White blood cell ANC: Absolut neutrophil count ALC: Absolut lymphocyte count NK: Natural killer y: year m: month F: Female M: Male CGD: Chronic granulomatous disease CID: Combined immunodeficiency MSMD: Mendelian susceptibility to mycobacterial disease SCID: Severe combined immunodeficiency
Number in the brackets indicate reference values Bold numbers indicate aberrant values

There were two CGD patients in our study (patient 3 and 4). Both were diagnosed after detection of impaired oxidative burst according to dihydro rhodamine tests. The first variant is a known mutation (c.G70A:p.G24R, homozygous) in CYBA gene (11). The second variant in the other CGD patient is a biallelic missense point mutation (c.G233A:p.G78E) affecting NCF2 protein. We found this variant in dbSNP database (rs137854519) but there is no clinical interpretation in Gnomad, TopMed or ClinVar databases. The prediction scores showed this variant as causative (SIFT: 0.001, PolyPhen: 1, CADD: 32). Thus, this is considered as a new disease-causing variant in a known CGD gene.

We detected two different heterozygous variants in SBDS gene in an eleven-year-old female patient (patient 8) who has been following with severe congenital neutropenia for last 10 years. The first variant (c.A184T:p.K62X) is found in ClinVar database as likely pathogenic/pathogenic and leads to a premature stop codon (12). The second one is a missense variant (c.T578C:p.L193P) and its effect is leucine to proline change in codon 193. This change is located in RNA recognition motif domain (RRM) of the protein (13). Mutations in SBDS gene are the main reason for Shwachman-Diamond syndrome which is resulted in congenital neutropenia (13). These compound heterozygous SBDS mutations are compatible with the patient's clinical phenotype.

The last genetic diagnosis of the study is a RAG2 deficiency in the patient (patient 12) with T- B- NK+ severe combined immunodeficiency. We found a known variant (c.104G>T: p.Gly35Val) which has been shown to cause T- B- NK+ SCID (14). Moreover, the patient has an affected twin brother suffering SCID. Then we showed the same mutation in the brother by Sanger sequencing.

All detected variants in the patients and segregations in the parents were validated by capillary sequencing and the details of the variants are given in Table 2.

Table 2. Disease-causing variants detected in the study

Patient	Gene	Variant	Aa change	Type	Hom/Het	Sanger Validation
1	PGM3	c.G214A	p.G72S	Missense	Hom	Yes
2	SAMD9L	c.A2639C	p.H880P	Missense	Het	Yes
3	CYBA	c.G70A	p.G24R	Missense	Hom	Yes
4	NCF2	c.G233A	p.G78E	Missense	Hom	Yes
8	SBDS	c.A184T	p.K62X	Nonsense	Het	Yes
8	SBDS	c.T578C	p.L193P	Missense	Het	Yes
12	RAG2	c.104G>T	p.Gly35Val	Missense	Hom	Yes

Aa: Amino acid Hom: Homozygous Het: Heterozygous

DISCUSSION

Primary immunodeficiency diseases have very heterogeneous clinical presentations. Thus, genetic and molecular characterization of the disease is crucial for the patients. The use of NGS methods facilitates this characterization by massive and rapid sequencing (15). Here, we performed whole exome sequencing in 12 PID patients and made genetic diagnosis in 6 of them.

Genetic diagnosis rate in PID patients by NGS based methods vary according to different parameters of the study such as consanguinity rate, disease population and the type of the NGS application. For example, in the study by Abolhassani et al. 76% the study population was coming from consanguineous marriage and the genetic diagnosis percentage was 78% (16). There are many studies reflecting the efficiency of NGS use in PID patients in the literature. Some of them included only a specific group of primary immunodeficiency like SCID, CID or antibody deficiencies (16-19). The study of Abolhassani et al. included only combined immunodeficiency patients and the results were highly satisfactory (16). However, Emily Edwards et al. reported that whole exome sequencing identifies pathogenic variants in only less than 20% of patients with predominantly antibody deficiency (PAD) (20). These studies included different number of patients ranged between 15 and 696 and causative mutations were found in between 15% and %78 of the patients (16-19). For the studies conducted with different PID groups, the diagnosis rate was between 15%- and 46% (21-24). This ratio was 50% in our study which was performed in a highly consanguineous population. Four of the six diagnosed patients (67%) were born to consanguineous parents.

Autosomal recessive inheritance in primary immunodeficiency is common in consanguineous communities like the particular regions of Turkey (3). Unsurprisingly all identified variants in our study were autosomal recessive inherited mutations except one (SAMD9L variant).

Genetic variants determined in the study were found in known PID genes. Of seven causative PID variants, four homozygous, one heterozygous and two compound heterozygous mutations were found in six patients. The variants in RAG2, CYBA genes and one of the SBDS variant

were identified in PID patients before (11,12,14). The NCF2 (rs137854519) and SBDS (rs1195681400) variants in the patients with CGD and Shwachman-Diamond syndrome had SNP identification numbers. However, pathogenicity of these variants has not been defined before and these are new disease-causing mutations based on the high prediction tool scores and the respective patients' clinical phenotypes. We also identified two novel mutations in PGM3 and SAMD9L genes leading to hyper IgE and Shwachman-Diamond syndrome according to predictive effects of the variants and compatibility with the clinical features.

We could not find any causative variants in half of the patients. The clinical diagnosis of these patients was three combined immunodeficiency, two possible mendelian susceptibility to mycobacterial disease (MSMD) and one congenital neutropenia. Although the quality of the study and the coverage of the targeted regions were really high, undetermined genetic defects of the patients could be located in uncovered or untargeted regions like promotor regions or introns. Another reason for the undetermined variants was copy number variations (CNVs). Due to the length of the CNV, targeted NGS panels and WES could not detect the variant because an exon generally does not cover the entire CNV (15). In this case, whole genome sequencing (WGS) may be applied because it is much more useful than whole exome sequencing for detecting CNVs (25). However, whole genome sequencing is more expensive than WES and the data analysis obtained from WGS is more complex and difficult (25). Thus, the use of WGS may cause a longer diagnosis period for the patients.

LIMITATIONS

As the limitations of our study, we did not perform a proper CNV analysis due to used methodology (whole exome sequencing). Another limitation of the study is may be the low number of the patients.

CONCLUSION

In conclusion, our results indicate that the percentage of the genetic diagnosis is encouraging for the further studies and whole exome sequencing is a useful method for a rapid and true diagnosis in primary immunodeficiency. Thus, we consider that NGS based methods can be applied efficiently in clinical immunology departments of our country.

Acknowledgements: We would like to thank "Can Sucak Candan Biseyler" Foundation (CSCBF) for their support and contributions during the study. CSCBF was founded in 2018 to honor Can Sucak who lost his life due to complications of primary immunodeficiency. CSCBF supports research in the field of primary immunodeficiency and promotes awareness.

Competing interests: The authors declare that they have no competing interest.

Financial Disclosure: This study was partly supported by the "Can Sucak Candan Biseyler" Foundation (CSCBF).

Ethical approval: Ethics Committee of Hacettepe University (GO 20/407).

REFERENCES

1. Parvaneh N, Casanova JL, Notarangelo LD, et al. Primary immunodeficiencies: a rapidly evolving story. J Allergy Clin Immunol 2013;131:314-23.
2. Tangye SG, Al-Herz W, Bousfiha A, et al. The Ever-Increasing Array of Novel Inborn Errors of Immunity: an Interim Update by the IUIS Committee. J Clin Immunol 2021;41:666-79.
3. Sanal O, Tezcan I. Thirty years of primary immunodeficiencies in Turkey. Ann N Y Acad Sci 2011;1238:15-23.
4. Notarangelo LD, Bacchetta R, Casanova JL, Su HC. Human inborn errors of immunity: An expanding universe. Sci Immunol 2020;5:1662.
5. Chou J, Ohsumi TK, Geha RS. Use of whole exome and genome sequencing in the identification of genetic causes of primary immunodeficiencies. Curr Opin Allergy Clin Immunol 2012;12:623-8.
6. Picard C, Fischer A. Contribution of high-throughput DNA sequencing to the study of primary immunodeficiencies. Eur J Immunol 2014;44:2854-61.
7. Seidel MG, Kindle G, Gathmann B, et al. The European Society for Immunodeficiencies (ESID) Registry Working Definitions for the Clinical Diagnosis of Inborn Errors of Immunity. J Allergy Clin Immunol Pract 2019;7:1763-70.
8. Halacli SO, Ayvaz DC, Sun-Tan C, et al. 8STK4 (MST1) deficiency in two siblings with autoimmune cytopenias: A novel mutation. Clin Immunol 2015;161:316-23.
9. Sassi A, Lazaroski S, Wu G, et al. Hypomorphic homozygous mutations in phosphoglucomutase 3 (PGM3) impair immunity and increase serum IgE levels. J Allergy Clin Immunol 2014;133:1410-9.
10. Chen DH, Below JE, Shimamura A, et al. Ataxia-Pancytopenia Syndrome Is Caused by Missense Mutations in SAMD9L. Am J Hum Genet 2016;98:1146-58.
11. Yamada M, Ariga T, Kawamura N, et al. Genetic studies of three Japanese patients with p22-phox-deficient chronic granulomatous disease: detection of a possible common mutant CYBA allele in Japan and a genotype-phenotype correlation in these patients. Br J Haematol 2000;108:511-7.
12. Pronicka E, Piekutowska-Abramczuk D, et al. New perspective in diagnostics of mitochondrial disorders: two years' experience with whole-exome sequencing at a national paediatric centre. J Transl Med 2016;14:174.
13. Liu Y, Liu F, Cao Y, et al. Shwachman-Diamond Syndrome Protein SBDS Maintains Human Telomeres by Regulating Telomerase Recruitment. Cell Rep 2018;22:1849-60.
14. Corneo B, Moshous D, Callebaut I, de Chasseval R, Fischer A, de Villartay JP. Three-dimensional clustering of human RAG2 gene mutations in severe combined immune deficiency. J Biol Chem 2000;275:12672-5.

15. Meyts I, Bosch B, Bolze A, et al. Exome and genome sequencing for inborn errors of immunity. J Allergy Clin Immunol 2016;138:957-69.
16. Abolhassani H, Chou J, Bainter W, et al. Clinical, immunologic, and genetic spectrum of 696 patients with combined immunodeficiency. J Allergy Clin Immunol 2018;141:1450-8.
17. Abolhassani H, Aghamohammadi A, Fang M, et al. Clinical implications of systematic phenotyping and exome sequencing in patients with primary antibody deficiency. Genet Med 2019;21:243-51.
18. Erman B, Bilic I, Hirschmugl T, et al. Investigation of Genetic Defects in Severe Combined Immunodeficiency Patients from Turkey by Targeted Sequencing. Scand J Immunol 2017;85:227-34.
19. Maffucci P, Filion CA, Boisson B, et al. Genetic Diagnosis Using Whole Exome Sequencing in Common Variable Immunodeficiency. Front Immunol 2016;7:220.
20. Edwards ESJ, Bosco JJ, Ojaimi S, O'Hehir RE, van Zelm MC. Beyond monogenetic rare variants: tackling the low rate of genetic diagnoses in predominantly antibody deficiency. Cell Mol Immunol 2020;18:588-603.
21. Al-Mousa H, Abouelhoda M, Monies DM, et al. Unbiased targeted next-generation sequencing molecular approach for primary immunodeficiency diseases. J Allergy Clin Immunol 2016;137:1780-7.
22. Gallo V, Dotta L, Giardino G, et al. Diagnostics of Primary Immunodeficiencies through Next-Generation Sequencing. Front Immunol 2016;7:466.
23. Moens LN, Falk-Sorqvist E, Asplund AC, Bernatowska E, Smith CI, Nilsson M. Diagnostics of primary immunodeficiency diseases: a sequencing capture approach. PLoS One 2014;9:114901.
24. Nijman IJ, van Montfrans JM, Hoogstraat M, et al. Targeted next-generation sequencing: a novel diagnostic tool for primary immunodeficiencies. J Allergy Clin Immunol 2014;133:529-34.
25. Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A 2015;112:5473-8.