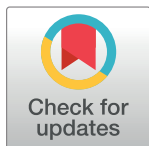


## RESEARCH ARTICLE

## Novel genetic variants of inborn errors of immunity

Farida Almarzooqi<sup>1</sup><sup>✉</sup>, Abdul-Kader Souid<sup>1</sup><sup>✉</sup>, Ranjit Vijayan<sup>2</sup><sup>✉</sup>, Suleiman Al-Hammadi<sup>1</sup>**1** Department of Pediatrics, College of Medicine and Health Sciences, UAE University, Al Ain, United Arab Emirates, **2** Department of Biology, College of Science, UAE University, Al Ain, United Arab Emirates These authors contributed equally to this work.<sup>✉</sup> Current address: Department of Pediatrics, College of Medicine and Health Sciences, UAE University, Abu Dhabi, United Arab Emirates<sup>✉</sup> Current address: Department of Biology, College of Science, UAE University, Abu Dhabi, United Arab Emirates<sup>✉</sup> These authors also contributed equally to this work.\* [famarzooqi@uaeu.ac.ae](mailto:famarzooqi@uaeu.ac.ae)

## Abstract

 OPEN ACCESS**Citation:** Almarzooqi F, Souid A-K, Vijayan R, Al-Hammadi S (2021) Novel genetic variants of inborn errors of immunity. PLoS ONE 16(1): e0245888. <https://doi.org/10.1371/journal.pone.0245888>**Editor:** Cinzia Ciccacci, Unicamillus, Saint Camillus International University of Health Sciences, ITALY**Received:** September 17, 2020**Accepted:** January 10, 2021**Published:** January 22, 2021**Copyright:** © 2021 Almarzooqi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.**Data Availability Statement:** All relevant data are within the paper and its [Supporting Information](#) files.**Funding:** This retrospective data collection study has no source of funding. The study has not received any funding from grants, organizations, or institutions (including our institution). The authors received no specific funding for this work. We have no conflicts of interest to disclose. This manuscript is not funded.**Competing interests:** NO authors have competing interests.

## Objectives

Inborn errors of immunity (IEI) are prevalent in tribal cultures due to frequent consanguineous marriages. Many of these disorders are autosomal recessive, resulting from founder mutations; hence they are amenable to prevention. The primary objective of this study was to evaluate the pathogenicity of novel variants of IEI found among Emiratis.

## Methods

This retrospective data collection study reports novel variants of IEI detected by diagnostic exome sequencing. Pathogenicity prediction was based on scoring tools, amino acid alignment, and Jensen–Shannon divergence values.

## Results

Twenty-one novel variants were identified; nine were frameshift, three nonsense, four intro-nic (one pathogenic), and five missense (two pathogenic). Fifteen variants were likely pathogenic, of which 13 were autosomal recessive and two uncertain inheritance. Their clinical spectra included combined immunodeficiency, antibody deficiency, immune dysregulation, defects in intrinsic/innate immunity, and bone marrow failure.

## Conclusion

The described novel pathogenic variants are core to a planned national screening program that aims toward IEI prevention. Future studies, however, are needed to confirm their natural history in individual patients and estimate their prevalence in the community.

## Introduction

There are over 400 genes implicated in inborn errors of immunity (IEI) [1]. Identifying their variants is critical for patient care, genetic prevention, and translational research [2]. For this purpose, diagnostic exome sequencing has proven helpful in preventing primary immunodeficiency (PID) in tribal populations where founder mutations and autosomal recessive (AR) disorders are exceptionally common [3].

A national premarital screening program was implemented in the United Arab Emirates (UAE) in 2005 (Personal Status Act No 28, Article 27). This endeavor has eliminated sickle cell disease and beta-thalassemia major from the community. This study examines indigenous (novel) pathogenic variants of IEI, and suggests including them in the existing premarital screening program.

## Methods

This retrospective data collection (Chart Review) study was approved through the ‘Tawam Human Research Ethics Committee’ (T-HREC); reference number: MF2058-2020-722 (AA/AJ/722; May 13, 2020), and informed consent to participate was waived. The study involved one center (Tawam Hospital, Al Ain City, Abu Dhabi, UAE). A review of 54 variants of IEI (detected by diagnostic exome sequencing between March 2016 and November 2019) was conducted between May 13, 2020 and July 10, 2020. Of these, 21 variants were novel and included in this report. Thirteen variants involved patients and eight were found on a screening test.

Variant information in public databases was consolidated using Ensembl Variant Effect Predictor (<https://www.ensembl.org/vep>). Gathered data included effect of variation, codon change, amino acid change, and variations from dbSNP, functional consequence, and pathogenicity score from algorithms, such as SIFT, PolyPhen, Condel, CADD, FATHMM, LRT, MetaLR, MetaSVM, MutationAssessor, MutationTaster, PROVEAN, REVEL, and VEST3 in dbNSFP (One-Stop Database of Functional Predictions and Annotations for Human Non-synonymous and Splice Site). The Mutation Significance Cutoff (MSC) Server was used to obtain gene-wise CADD MSC [4].

Multiple sequence alignment was used to evaluate conservation of amino acids at studied sites and compute the Jensen-Shannon divergence (JSD) values. Amino acid sequences of proteins from *Homo sapiens* (human), *Pan troglodytes* (chimpanzee), *Mus musculus* (house mouse), *Rattus norvegicus* (Norway rat), *Canis lupus familiaris* (dog), *Equus caballus* (horse), *Bos taurus* (bovine), *Xenopus tropicalis* (frog), *Gallus gallus* (chicken), and *Danio rerio* (zebrafish) were downloaded from NCBI RefSeq (S1 Table). Sequences were imported into Geneious 9.1.8 (<https://www.geneious.com>) using NCBI Accessions and multiple sequence alignment was performed using MUSCLE. Alignments were exported in FASTA format and used to obtain JSD values from “<https://compbio.cs.princeton.edu/conservation/score.html>” using the default options [5].

Homology model of the ATM (Ataxia telangiectasia mutated) protein with the variation (ATM:p.Val2823Gly) was generated using the SWISS-MODEL server (<https://swissmodel.expasy.org/>). The Protein Data Bank (PDB) structure 6K9L (4.27 Angstrom resolution cryo-EM structure of human dimeric ATM kinase) was used as the template. Protein structures were visualized using Visual Molecular Dynamics 1.9.3 [6].

## Statistics

Probability of having an autosomal recessive disorder(s) in the offspring was calculated on Excel, using binomial probability distribution (probability of having a recessive phenotype of 0.25 with variable number of genes).

## Results

[Table 1](#) lists the 21 novel variants of genes (19 single genes) related to IEI, sorted according to the International Union of Immunological Societies Phenotypical Classification [1]. Thirteen variants were found in symptomatic patients, while eight variants were found on a screening test in asymptomatic subjects. Fifteen variants are autosomal recessive (AR, 71%), three undetermined inheritance, one autosomal dominant (AD), and two X-linked. Nine variants are frameshift, three nonsense, four intronic (one pathogenic), and five missense (two pathogenic). [Table 2](#) shows sequence alignment of the proteins at positions of the missense variants from ten species.

### Two variants involve combined immunodeficiency

*DCLRE1C*:p.Cys182Tyr is predicted to be pathogenic based on the *in silico* prediction scores ([Table 1](#)) and phylogenetic analysis (Cys182 is conserved; JSD: 0.919, [Table 2](#)). *TAP2*:p.Arg252ThrfsTer11, which results in the truncation in the ABC transporter transmembrane domain making it non-functional, is likely to cause combined immunodeficiency with low CD8 [7]. A homozygous truncation at position 253 of TAP2, leading to a non-functional TAP2, has also been demonstrated to be associated with human lymphocyte antigen (HLA) class I deficiency [8].

### One variant involves ataxia-telangiectasia (AT)

*ATM*:p.Val2823Gly has conflicting predictions of pathogenicity. Val2823 is conserved, with a permissible change to isoleucine (JSD: 0.7041, [Table 2](#)). Val2823 ([Fig 1](#)) is located in the kinase domain of the alpha helical C terminus. Gly2823 is predicted to disturb the stability of the helical conformation, thus, affecting the structure and function of the protein [9,10]. This assessment is consistent with the SIFT score of 0.00 (damaging) and the CADD score of 28.9 ([Table 1](#)).

### Three variants involve combined immunodeficiency with syndromic features

*BCL11B*:p.Ser256Asn has benign predictions, except for CADD: 23.9. Ser256 is replaced by asparagine in chicken ([Table 2](#)), supporting the benign predictions. *DOCK8* variant p.Met1114TyrfsTer4 is likely pathogenic since the truncation resulting from the frameshift leads to the deletion of the DOCKER domain that is essential for the guanine nucleotide exchange factor activity of DOCK8. *ORAI1*:p.His165ProfsTer5, a truncation that results in the deletion of two transmembrane helices of the membrane protein, and *SP110*:p.Gln231Ter, a truncation in the transcription factor, are also predicted to be pathogenic. Of note, in this category, two microdeletions (2.55 Mb and 3.152 Mb) within 22q11.21 were found; they cause *TBX1* haploinsufficiency and hence are disease-causing (DiGeorge syndrome) [11].

### One variant involves antibody deficiency

Variant *BTK*:p.Gly27Asp is benign ([Table 1](#)). Gly27 is replaced by glutamate in chicken ([Table 2](#)).

### Two variants involve phagocyte defects

*CD8A*:c.49+2T>G (CADD: 28.1) and *CSF3R*:p.Asp339ThrfsTer94, in the third fibronectin type III domain in this protein is likely to cause it to misfold, and hence likely pathogenic.

Table 1. Novel variants of IEI.

Variants	Phenotypes	SIFT	PolyPhen	Condel	CADD (MSC)	FATHMM	LRT	MetalR	MetaSVM	Mut Assessor	Mut Taster	PROVEAN	REVEL	VEST3	Predicted Pathogenicity
<b>Severe combined immunodeficiency (CD3 T-cell count &lt; 300/μl)</b>															
NM_001033855.3(DCLRE1C):c.545G>A, p.Cys182Tyr DNA cross-link repair protein 1C; MIM#605988; missense; [IGt]>[At]	Omenn syndrome (MIM#603554) and severe combined immunodeficiency, Athabaskan type (MIM#602450); AR	0.00 D	0.993 D	0.895 D	25.3 (11.57)	0.797	0.843	0.875	0.889	0.694	0.588	0.899	0.868	0.886	Pathogenic
<b>Combined immunodeficiency with low CD8</b>															
NM_000544.3(TAP2):c.753dupA, p.Arg252HisTer11 Transporter, ATP-binding cassette, major histocompatibility complex; MIM#170261; frameshift; [-]>[A]	Bare lymphocyte syndrome, type I, due to TAP2 deficiency; MIM#604571; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
<b>Combined immunodeficiency with DNA repair defects</b>															
NM_000051.3(ATM):c.8468T>G, p.Val2823Gly; Ataxia-telangiectasia mutated gene; MIM#607585; missense; [gTt]>[gGt]	Ataxia-telangiectasia, AT; MIM#208900; AR	0.00 D	0.741 D	0.716 D	28.5 (6.38)	0.757	0.513	0.834	0.855	0.612	0.810	0.817	0.944	0.789	Pathogenic
<b>Combined immunodeficiency with syndromic features</b>															
NM_138576.3(BCL11B):c.767G>A, p.Ser256Asn BAF chromatin remodeling complex subunit BCL11B; MIM#606558; missense; [aGc]>[aAc]	Immunodeficiency 49 (MIM#617237) and intellectual developmental disorder with dysmorphic facies, speech delay, and T-cell abnormalities (MIM#618092); AD	0.34 T	0.084 B	0.029 N	23.9 (10.67)	0.453	0.255	0.201	0.088	0.045	0.411	0.212	0.080	0.068	Benign
NM_203447.3(DOCK8):c.3339dup, p.Met1114TyrfsTer4 Dedicator of cytokinesis 8; MIM#611432; frameshift; [-]>[T]	Hyper-IgE recurrent infection syndrome 2; MIM#243700; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
NM_032790.3(ORAI1):c.454C>A, p.His165ProfsTer2 ORAI calcium release-activated calcium modulator 1; MIM# 61027; frameshift; [ccc]>[cCc]	Combined immunodeficiency due to ORAI1 deficiency; MIM#612782; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
NM_080424.2(SPI10):c.691C>T, p.Gln231Ter Nuclear body protein SPI10; MIM#604457; nonsense; [Caa]>[Taa]	Hepatic venoocclusive disease with immunodeficiency; MIM#235550; AR	-	-	-	7.97 (12.46)	-	-	-	-	-	0.810	-	-	-	Pathogenic
<b>Antibody deficiency</b>															
NM_001287344.1(BTK):c.80G>A, p.Gly27Asp Bruton agammaglobulinemia tyrosine kinase; MIM#300300; missense; [gGt]>[gAt]	X-linked agammaglobulinemia; MIM#300755	-	0 B	-	12.49 (10.1)	-	-	-	-	-	0.320	-	-	-	Benign
<b>Defects in phagocytes</b>															
NM_001145873.1(CD8A):c.49+2T>G CD8 Antigen, alpha polypeptide; MIM#186910; splice-donor	CD8 deficiency, familial; MIM#608957; AR	-	-	-	28.1 (8.93)	-	-	-	-	-	-	-	-	-	Pathogenic

(Continued)

Table 1. (Continued)

Variants	Phenotypes	SIFT	PolyPhen	Condel	CADD (MSC)	FATHMM	LRT	MetalR	MetaSVM	Mut Assessor	Mut Taster	PROVEAN	REVEL	VEST3	Predicted Pathogenicity
NM_156039.3(CSF3R):c.1015delG, p.Asp339ThrfsTer94 Colony-stimulating factor 3 receptor, granulocyte; MIM#138971; frameshift; [Gac]>[ac]	Neutropenia, severe congenital, 7, autosomal recessive; MIM#617014; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
<b>Immune dysregulation</b>															
NM_006726.4(LRBA):c.974C>T; p.Ala325Val Lipopolysaccharide-responsive, beige-like anchor; MIM#606453; missense; [gCt]>[gTt]	Immunodeficiency, common variable, 8, with autoimmunity; MIM#606453; AR	0.62 T	0.145 B	0.017 N	26.7 (33)	0.62918	0.62929	0.52904	0.48446	0.04202	0.58761	0.27052	0.40608	0.84713	Benign
NM_006726.4(LRBA):c.8007G>A; p.Trp269Ter Lipopolysaccharide-responsive, beige-like anchor; MIM#606453; nonsense; [tgG]>[tGA]		-	-	-	32.0 (33)	-	0.843	-	-	-	0.810	-	-	-	Pathogenic
NM_006726.4(LRBA):c.534del, p.Asp179IlefsTer16 Lipopolysaccharide-responsive, beige-like anchor; MIM#606453; frameshift; [ggA]>[gg]		-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
<b>Defects in intrinsic and innate immunity</b>															
NM_001127593.1(FCGR3A):c.423dupT, p.Leu142PhefsTer13 Fc fragment of IgG, low affinity IIIa, receptor for; MIM#146740; frameshift; [-]>[T]	Immunodeficiency 20; MIM#615707; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
NM_005534.3(IFNGR2):c.123C>G, p.Tyr41Ter Interferon-gamma receptor 2; MIM#147569; nonsense; [taC]>[taG]	Immunodeficiency 28; MIM#614889; AR	-	-	-	35.0 (4.43)	-	0.315	-	-	-	0.810	-	-	-	Pathogenic
<b>Hemophagocytic lymphohistiocytosis (HLH) and Epstein-Barr virus (EBV) susceptibility</b>															
NM_032121.5(MAGT1):c.198+383G>C Magnesium transporter 1; MIM#300715; splice-donor	Immunodeficiency, X-linked recessive, with magnesium defect, EBV virus infection and neoplasia; MIM#300853	-	-	-	-	-	-	-	-	-	-	-	-	-	Benign
<b>Bone marrow failure</b>															
NM_007300.3(BRIP1):c.4358-2A>T BRCA1-interacting protein 1; MIM#605882; splice-acceptor	Fanconi anemia, complementation group J; MIM#609054; unknown inheritance	-	-	-	-	-	-	-	-	-	-	-	-	-	Uncertain
NM_033084.4(FANCD2):c.1275_1278+5delCTTAGTAAAGinsTTTAT, p.Tyr425_Leu426delins FANCD2 gene; MIM#613984; splice donor	Fanconi anemia, complementation group D2; MIM#227646; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Uncertain
NM_001113378.1(FANCD1):c.1289_1290delTT, p.Phe430fsTer FANCD1 gene; MIM#611360; frameshift; [tTT]>[t]	Fanconi anemia, complementation group I; MIM#609053; AR	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
<b>Others</b>															

(Continued)

Table 1. (Continued)

Variants	Phenotypes	SIFT	PolyPhen	Condel	CADD (MSC)	FATHMM	LRT	MetalR	MetaSVM	Mut Assessor	Mut Taster	PROVEAN	REVEL	VEST3	Predicted Pathogenicity
NM_002641.3(PIGA): c.1077_1078delinC, p. Leu359PhefsTer19 Phosphatidylinositol glycan anchor biosynthesis class a protein; MIM#311770; frameshift; [ttGGaa]>[ttaa]	Paroxysmal nocturnal hemoglobinuria, somatic; MIM#300818; unknown inheritance	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic
NM_018676.3(THSD1): c.1322_1329del, p. Arg441GlnfsTer66 Thrombospondin type 1 domain-containing protein 1; MIM#616821; frameshift; [cGGCCAGCC]>[c]	Aneurysm, intracranial berry, 12; MIM#618734; unknown inheritance	-	-	-	-	-	-	-	-	-	-	-	-	-	Pathogenic

SIFT (Range: 0–1; “Deleterious (D)” if score is less than 0.05; “Tolerated (T)” if score is greater than or equal to 0.05), PolyPhen-2 (Range: 0–1; “Probably Damaging (D)” if score is greater than 0.908; “Benign (B)” if score is less than or equal to 0.446), Mutation Assessor, Mutation Taster (integrates scores from Ensembl, UniProt, ClinVar, ExAC, 1000 Genomes Project, phyloP, and phastCons), LRT (likelihood ratio test, based on a probabilistic estimation of the phylogenetic relationship, with residue changes treated equally rather than weighting radical or conservative amino acid changes differently), PROVEAN (protein variation effect analyzer) and FATHMM (functional analysis through hidden Markov models) are pathogenicity prediction based on alignment of protein sequences and/or structures. Condel (consensus deleterious) is a consensus prediction based on SIFT and PolyPhen-2 scores. REVEL (rare exome variant ensemble learner), CADD (combined annotation-dependent depletion; scores range from 1 to 99, e.g., a score of 30 signifies a 0.1% top variant), MSC (Mutation Significance Cutoff at the lower boundary of the 95% confidence interval for the CADD score of the gene; the phenotypic impact is possibly damaging above this cutoff), MetalR (meta-analytic logistic regression) and MetaSVM (meta-analytic support vector machine) are ensemble-based prediction based on multiple deleteriousness, conservation and ensemble scoring methods derived from surrounding sequences, gene-model interpretations, evolutionary constraints, epigenetic extents, pathogenicity predictions and allele frequencies. B, benign; D, deleterious; T, tolerated; AR autosomal recessive; AD, autosomal dominant.

<https://doi.org/10.1371/journal.pone.0245888.t001>

**Table 2. Sequence alignment of protein sequences at positions of the missense variants with the Jensen-Shannon divergence (JSD) values.**

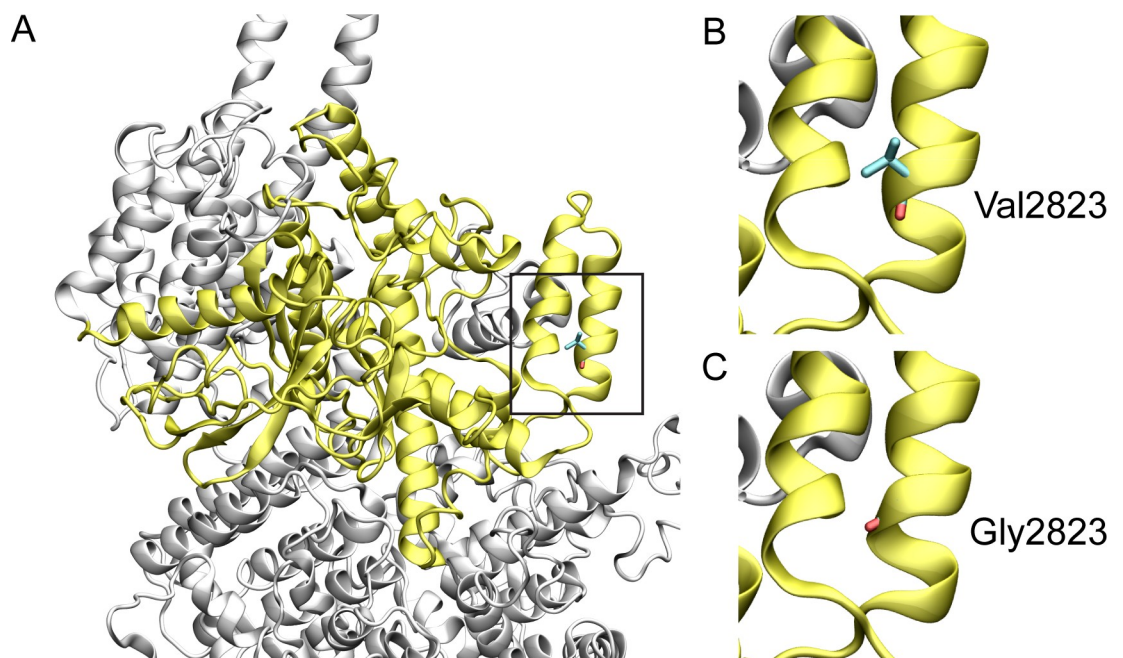
	<i>ATM</i> :p. Val2823Gly	<i>BCL11B</i> : pSer256Asn	<i>BTK</i> :p. Gly27Asp	<i>DCLRE1C</i> :p. Cys182Tyr	<i>LRBA</i> :p. Ala325Val
Human	Val	Ser	Gly	Cys	Ala
Chimpanzee	Val	Ser		Cys	Ala
Mouse	Ile	Ser		Cys	Ala
Rat	Ile	Ser		Cys	Ala
Dog	Ile	Ser		Cys	Ala
Horse	Ile	Ser	Ser	Cys	Ala
Bovine	Ile	Ser		Cys	Ala
Frog	Ile	Ser		Cys	Ala
Chicken	Val	Asn	Glu	Cys	Ala
Zebrafish	Val	Ser		Cys	Ala
<b>JSD value</b>	0.7041	0.7603		0.9189	0.7990

<https://doi.org/10.1371/journal.pone.0245888.t002>

*CSF3R* encodes the G-CSF (granulocyte colony-stimulating factor) receptor, essential for JAK/STAT-mediated granulopoiesis [12].

### Variants of LRBA involve immune dysregulation

*LRBA*:p.Ala325Val has benign scores, except for CADD: 26.7 and VEST3: 0.847 (Table 1). Ala325 is conserved (JSD: 0.799, Table 2). The *LRBA* variants p.Trp2669Ter (CADD: 32), affects the third WD repeat (short motifs often terminating at Trp-Asp) and likely impacts the folding of this region, and p.Asp179IlefsTer16, which results in the deletion of the BEACH domain and the WD repeats, are both likely to be pathogenic.



**Fig 1. Structural model of wild-type and variant ATM protein.** The ATM protein structure is shown in white cartoon representation and the kinase domain is shown in yellow. Amino acids are shown in stick representation. The black boxed region is enlarged in the subsequent two images. (A) Structure of the conserved kinase domain of ATM, (B) wild type Val2823, and (C) variant Gly2823.

<https://doi.org/10.1371/journal.pone.0245888.g001>

### Two variants involve intrinsic/innate immunity

*FCGR3A*:p.Leu142PhefsTer13, in the second immunoglobulin-like domain, and *IFNGR2*:p.Tyr41Ter (CADD: 35), in the first fibronectin type III domain, are both likely to affect the structure and function of these protein and hence likely pathogenic.

### One variant involve hemophagocytic lymphohistocytosis and Epstein-Barr virus susceptibility

*MAGT1*:c.198+383G>C is benign.

### Three variants involve bone marrow failure syndromes

*BRIP1*:c.4358-2A>T and *FANCD2*:c.1275\_1278+5delCTTAGTAAGinsTTTAT are intronic with uncertain clinical significance. *FANCI*:p.Phe430fsTer results in a truncation that deletes the DNA binding region of a protein associated DNA repair; it is therefore likely pathogenic.

*PIGA*:p.Leu359PhefsTer19, associated with the glycosyl transferases group 1 family, is likely pathogenic causing paroxysmal nocturnal hemoglobinuria. *THSD1*:p.Arg441GlnfsTer66, resulting from an eight-base deletion, is also likely pathogenic.

## Discussion

This study describes 21 novel variants of IEI. Their types include frameshift, nonsense, intronic, and missense. Our bioinformatics analysis shows the majority (15 of 21, or 71%) are pathogenic or likely pathogenic; of these, 13 are autosomal recessive. Thus, these mutations are amenable for genetic prevention via screening and counseling. Future studies, however, are needed to confirm their prevalence and natural history in individual patients.

The variants reported here are novel. Their prediction of pathogenicity is assessed by *in silico* computer tools described above. Thus, functional validation of these variants remains fundamental to attribute pathogenicity with certainty [13,14]. As previously shown for cancer predisposition, improved genomic analysis lowers the prevalence of variant of uncertain significance (VUS) [15]. The authors concluded that variant “interpretation by multiple experts in the context of personal and family histories maximizes actionable results and minimizes reports of VUS” [15]. Similar efforts are needed for novel variants of IEI.

Pathogenic variants of autosomal recessive disorders carry ‘fetal risk’ for homozygosity and compound heterozygosity, especially in tribal cultures with frequent consanguineous marriages. Fetal risk is, of course, higher in the presence of multiple pathogenic variants of a gene, especially within a tribe, such as the three variants of *LRBA* (Table 1). Another example is the six known pathogenic variants of ataxia telangiectasia (AT) in our community; one of these is novel and reported here. This entity is characterized by genomic instability due to defects in DNA repair [16]. Homozygous and double heterozygous affected individuals have immune derangements (e.g., thymic involution) and central nervous system degeneration. Heterozygous individuals may show mild clinical features, such as increased risk for cancer and hypersensitivity to radiation [17,18]. Thus, the AT variants carry an exceptional fetal risk for homozygosity and compound heterozygosity, especially in our tribal cultures that mainly practice consanguineous marriages. Such variants need to be built into a national screening program aiming at disease prevention [19].

Generation of homology models is often hampered by the unavailability of wild-type human or homologous protein structures. Hence, in this instance, the only variant that could be modelled was the *ATM*:p.Val2823Gly based on the 4.27 Å cryo-electron microscopy structure of ATM (PDB ID: 6K9L). Val283 is part of the helix designated as  $\alpha 4c$  in the



phosphatidylinositol 3-/4-kinase catalytic domain of the ATM serine-protein kinase [20]. This valine, or the physiochemically similar isoleucine, is conserved in all species considered here. While the precise role of the helix is not known, a change to a glycine, a notably achiral amino acid, at this position could potentially disrupt the stability of this helix.

Twenty-one IEI novel variants are categorized based on the IUIS Phenotypical Classification [1]. This allows simpler dissection of IEI molecular diagnosis in association with clinical and immunological profiles. It serves as a framework to determine IEI types and prevalence in the UAE. It will also enhance our immunological understanding of the variant to aid patient care with targeted therapy and prevention with family screening and counselling.

The novel IEI variants reported here include one variant of severe combined immunodeficiency (SCID); two variants of combined immunodeficiency (CID) with low CID and DNA repair defect; four variants of CID with syndromic features; one variant of antibody deficiency; two variants of defects in phagocytes; three variants of one gene related to immunodysregulation; two variants of defects in intrinsic and innate immunity; one variant of hemophagocytic lymphohistocytosis (HLH) and Epstein-Barr virus (EBV) susceptibility; three variants related to bone marrow failure; and two variants under 'others'. PIGA and THSD1 variants under 'Others' are not reported in IUIS. However, the PIGA variant causes paroxysmal nocturnal hemoglobinuria (PNH) and is associated with cellular immunodeficiency. It is hypothesized that PNH impacts lymphocyte count through decrease in lymphopoiesis and complement-mediated damage [21]. A case series in UAE shows the THSD1 variant causes intracranial berry aneurysm 12, nonimmune hydrops fetalis, congenital cardiac defects, hemangiomas and hypogammaglobulinemia [22].

One of the major limitations of diagnostic sequencing and genetic testing is that the effect of a variation on function cannot be inferred from sequence alone. Considering the volume of next generation data produced, a large proportion of variants currently remain classified as variants of uncertain significance (VUS) in bioinformatics databases. While phenotype data could address this problem partly, systematic *in vitro* functional and mutational assays are necessary to pinpoint the precise role of a variant on the structure and function of a protein and its pathways [14]. Often, a single assay may not be sufficient for this purpose elevating the difficulty in establishing the pathogenicity of VUS. Nonetheless, in this study, based on computational predictors, we have attempted to predict the pathogenicity of novel variants identified in our population. However, as indicated, functional assays would be necessary to ascertain the precise pathogenicity of these variants.

IEI are heterogeneous disorders and are often underdiagnosed. Knowing the causative genetic variant offers a better prediction of its course and opens up opportunities for participating in clinical trials related to the defect. Genetic screening, on the other hand, offers early identification of people at risk; thereby, allowing more prompt interventions [19]. Thus, genetic testing is required for all individuals with immunodeficiency. Our results will aid in setting the framework of a national screening program aiming at genetic prevention of IEI.

## Supporting information

**S1 Table. NCBI RefSeq accession of sequences used for protein sequence alignment.**  
(PDF)

## Acknowledgments

The contributions of patients and Tawam Hospital staff are greatly appreciated.

## Author Contributions

**Conceptualization:** Farida Almarzooqi, Abdul-Kader Souid, Suleiman Al-Hammadi.

**Data curation:** Farida Almarzooqi, Abdul-Kader Souid, Suleiman Al-Hammadi.

**Formal analysis:** Farida Almarzooqi, Abdul-Kader Souid, Ranjit Vijayan, Suleiman Al-Hammadi.

**Investigation:** Farida Almarzooqi, Abdul-Kader Souid.

**Methodology:** Farida Almarzooqi, Abdul-Kader Souid, Ranjit Vijayan.

**Software:** Farida Almarzooqi.

**Supervision:** Abdul-Kader Souid.

**Writing – original draft:** Farida Almarzooqi, Abdul-Kader Souid.

**Writing – review & editing:** Farida Almarzooqi, Abdul-Kader Souid, Ranjit Vijayan, Suleiman Al-Hammadi.

## References

1. Bousfiha A, Jeddane L, Picard C, Al-Herz W, Ailal F, Chatila T, et al. Human inborn errors of immunity: 2019 update of the IUIS phenotypical classification. *Journal of Clinical Immunology*. 2020 Feb 11:1–6. <https://doi.org/10.1007/s10875-020-00758-x> PMID: 32048120
2. Lo B, Zhang K, Lu W, Zheng L, Zhang Q, Kanellopoulou C, et al. Patients with LRBA deficiency show CTLA4 loss and immune dysregulation responsive to abatacept therapy. *Science*. 2015 Jul 24; 349(6246):436–40. <https://doi.org/10.1126/science.aaa1663> PMID: 26206937
3. Simon AJ, Golan AC, Lev A, Stauber T, Barel O, Somekh I, et al. Whole exome sequencing (WES) approach for diagnosing primary immunodeficiencies (PIDs) in a highly consanguineous community. *Clinical Immunology*. 2020 Mar 3:108376.
4. Itan Y, Shang L, Boisson B, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nature Methods* 2016; 13:109–110. <https://doi.org/10.1038/nmeth.3739> PMID: 26820543
5. Capra JA, Singh M. Predicting functionally important residues from sequence conservation. *Bioinformatics*. 2007 Aug 1; 23(15):1875–82. <https://doi.org/10.1093/bioinformatics/btm270> PMID: 17519246
6. Humphrey W, Dalke A, Schulten K. VMD—Visual Molecular Dynamics. *Journal of Molecular Graphics*. 1996; 14:33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) PMID: 8744570
7. Gadola SD, Moins-Teisserenc HT, Trowsdale J, Gross WL, Cerundolo V. TAP deficiency syndrome. *Clinical and Experimental Immunology*. 2000 Aug; 121(2):173. <https://doi.org/10.1046/j.1365-2249.2000.01264.x> PMID: 10931128
8. de la Salle H, Hanau D, Fricker D, Urlacher A, Kelly A, Salamero J, et al. Homozygous human TAP peptide transporter mutation in HLA class I deficiency. *Science*. 1994 Jul 8; 265(5169):237–41. <https://doi.org/10.1126/science.7517574> PMID: 7517574
9. Wang X, Chu H, Lv M, Zhang Z, Qiu S, Liu H, et al. Structure of the intact ATM/Tel1 kinase. *Nature Communications*. 2016 May 27; 7(1):1–8. <https://doi.org/10.1038/ncomms11655> PMID: 27229179
10. López-Llano J, Campos LA, Sancho J.  $\alpha$ -helix stabilization by alanine relative to glycine: Roles of polar and apolar solvent exposures and of backbone entropy. *Proteins: Structure, Function, and Bioinformatics*. 2006 Aug 15; 64(3):769–78. <https://doi.org/10.1002/prot.21041> PMID: 16755589
11. Zweier C, Sticht H, Aydin-Yaylagül I, Campbell CE, Rauch A. Human TBX1 missense mutations cause gain of function resulting in the same phenotype as 22q11.2 deletions. *The American Journal of Human Genetics*. 2007 Mar 1; 80(3):510–7. <https://doi.org/10.1086/511993> PMID: 17273972
12. Dong F, Qiu Y, Yi T, Touw IP, Lerner AC. The carboxyl terminus of the granulocyte colony-stimulating factor receptor, truncated in patients with severe congenital neutropenia/acute myeloid leukemia, is required for SH2-containing phosphatase-1 suppression of Stat activation. *The Journal of Immunology*. 2001 Dec 1; 167(11):6447–52. <https://doi.org/10.4049/jimmunol.167.11.6447> PMID: 11714811
13. Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD. Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. *Journal of Experimental Medicine*. 2014 Oct 20; 211(11):2137–49. <https://doi.org/10.1084/jem.20140520> PMID: 25311508

14. Bucciol G, Van Nieuwenhove E, Moens L, Itan Y, Meyts I. Whole exome sequencing in inborn errors of immunity: use the power but mind the limits. *Current Opinion in Allergy and Clinical Immunology*. 2017 Dec; 17(6):421–30. <https://doi.org/10.1097/ACI.0000000000000398> PMID: 28938278
15. Shirts BH, Casadei S, Jacobson AL, Lee MK, Gulsuner S, Bennett RL, et al. Improving performance of multigene panels for genomic analysis of cancer predisposition. *Genetics in Medicine*. 2016 Oct; 18(10):974–81. <https://doi.org/10.1038/gim.2015.212> PMID: 26845104
16. Leuzzi V, D'Agnano D, Menotta M, Caputi C, Chessa L, Magnani M. Ataxia-telangiectasia: A new remitting form with a peculiar transcriptome signature. *Neurology Genetics*. 2018; 4:e228. <https://doi.org/10.1212/NXG.0000000000000228> PMID: 29600275
17. Sandoval N, Platzer M, Rosenthal A, et al. Characterization of ATM gene mutations in 66 ataxia telangiectasia families. *Human Molecular Genetics*. 1999; 8:69–79. <https://doi.org/10.1093/hmg/8.1.69> PMID: 9887333
18. Prodosmo A, De Amicis A, Nisticò C, Gabriele M, Di Rocco G, Monteonofrio L, et al. p53 centrosomal localization diagnoses ataxia-telangiectasia homozygotes and heterozygotes. *Journal of Clinical Investigation*. 2013; 123:1335–1342. <https://doi.org/10.1172/JCI67289> PMID: 23454770
19. Al-Herz W, Al-Ahmad M, Al-Khabaz A, Husain A, Sadek A, Othman Y. The Kuwait National Primary Immunodeficiency Registry 2004–2018. *Front Immunol*. 2019; 10:1754. <https://doi.org/10.3389/fimmu.2019.01754> PMID: 31396239
20. Xiao J, Liu M, Qi Y, Chaban Y, Gao C, Pan B, et al. Structural insights into the activation of ATM kinase. *Cell Research* 2019 Jul 18; 29(8):683–685. <https://doi.org/10.1038/s41422-019-0205-0> PMID: 31320732
21. Fujimi A., Matsunaga T., Kogawa K., Ohnuma T., Takahira N., Abe T., et al. A patient with paroxysmal nocturnal haemoglobinuria in whom granulocyte colony-stimulating factor administration resulted in improvement of recurrent enterocolitis and its associated haemolytic attacks. *British Journal of Haematology* 2002; 119(3):858–862. <https://doi.org/10.1046/j.1365-2141.2002.03914.x> PMID: 12437672
22. Abdelrahman H. A., Al-Shamsi A., John A., Hertecant J., Lootah A., Ali B. R., et al. A recessive truncating variant in thrombospondin-1 domain containing protein 1 gene THSD1 is the underlying cause of nonimmune hydrops fetalis, congenital cardiac defects, and haemangiomas in four patients from a consanguineous family. *American Journal of Medical Genetics Part A* 2018; 176(9):1996–2003. <https://doi.org/10.1002/ajmg.a.40424> PMID: 30055085