Check for updates

Genetic characteristics of common variable immunodeficiency patients with autoimmunity

Zhihui Liu[1], Chenyang Lu[1], Pingying Qing[1], Ruijuan Cheng[1], Yujie Li[2], Xue Guo[2], Ye Chen[1], Zhiye Ying[3], Haopeng Yu[3,4]* and Yi Liu[1]*

[1]Department of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu, China, [2]Novogene Co. Ltd., Beijing, China, [3]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China, [4]Med-X Center for Informatics, Sichuan University, Chengdu, China

Background: The pathogenesis of common variable immunodeficiency disorder (CVID) is complex, especially when combined with autoimmunity. Genetic factors may be potential explanations for this complex situation, and whole genome sequencing (WGS) provide the basis for this potential.

Methods: Genetic information of patients with CVID with autoimmunity, together with their first-degree relatives, was collected through WGS. The association between genetic factors and clinical phenotypes was studied using genetic analysis strategies such as sporadic and pedigree.

Results: We collected 42 blood samples for WGS (16 CVID patients and 26 first-degree relatives of healthy controls). Through pedigree, sporadic screening strategies and low-frequency deleterious screening of rare diseases, we obtained 9,148 mutation sites, including 8,171 single-nucleotide variants (SNVs) and 977 Insertion-deletions (InDels). Finally, we obtained a total of 28 candidate genes (32 loci), of which the most common mutant was *LRBA*. The most common autoimmunity in the 16 patients was systematic lupus erythematosis. Through KEGG pathway enrichment, we identified the top ten signaling pathways, including "primary immunodeficiency", "JAK-STAT signaling pathway", and "T-cell receptor signaling pathway". We used PyMOL to predict and analyse the three-dimensional protein structures of the *NFKB1, RAG1, TIRAP, NCF2*, and *MYB* genes. In addition, we constructed a PPI network by combining candidate mutants with genes associated with CVID in the OMIM database via the STRING database.

Conclusion: The genetic background of CVID includes not only monogenic origins but also oligogenic effects. Our study showed that immunodeficiency and autoimmunity may overlap in genetic backgrounds.

Abbreviations: CVID, common variable immunodeficiency disorder; InDels, insertion-deletions; KEGG, kyoto encyclopedia of genes and genomes; OMIM, online mendelian inheritance in man; SNVs, single-nucleotide variants; STRING, search tool for recurring instances of neighbouring genes; WGS, whole genome sequencing.

1 Introduction

Common variable immunodeficiency disorder (CVID) is the most common primary symptomatic immunodeficiency disease in adults, with a prevalence of 1:50000 to 1:25000 (Sneller et al., 1993; Ameratunga, 2018). The clinical manifestations were highly heterogeneous, mainly showing recurrent infections and hypogammaglobulinemia. Some CVID patients also have autoimmune or inflammatory features, such as progressive interstitial lung disease, autoimmunity, gastrointestinal inflammatory disease, granuloma, lymphoproliferative disorders. The etiology of CVID is mostly unknown, and there is no unified definition. Expansion of the inborn errors of immunity landscape has increasingly revealed that monogenic, polygenic, and epigenetic factors are involved in the onset of CVID. In addition, genotypic and phenotypic overlap between CVID and other immune diseases is also increasing (Bogae et al., 2016; Peng et al., 2023). Therefore, some scholars have proposed that if pathogenic mutations are found (excluding *TNFRSF13B/TACI, TNFRSF13C/BAFFR*, etc.), these patients will be removed from the overall diagnosis of CVID and reclassified as having a CVID-like disease caused by a specific mutation (Ameratunga, 2018; Ameratunga et al., 2018). With the development of sequencing technologies, such as whole exome sequencing and whole genome sequencing (WGS), high-throughput sequencing technology has gradually become an important means for diagnosing and formulating treatment strategies for CVID-like patients (van Schouwenburg et al., 2015; Ameratunga et al., 2018). However, whole exome sequencing has identified pathogenic variants in less than 20% of patients with predominantly antibody deficiency. WGS can detect broader identifications of rare, underlying genetic variants, helping to further understand the contribution of common and rare variants to disease pathogenesis (Edwards et al., 2021). The early clinical manifestations of CVID often lack specificity, leading to delayed diagnosis and poor prognosis. Genetic testing can not only effectively assist diagnosis, but also provide basis for precision medicine (such as targeted immunotherapeutics). WGS is a good choice to solve this problem, which can effectively improve the diagnosis rate.

Current studies on CVID have mainly focused on Caucasians, whereas there are fewer studies on Asian populations (Gutierrez et al., 2018). According to the clinical manifestations, it can be divided into two types, one is mainly characterized by recurrent infections, while the other is often associated with noninfectious diseases, such as autoimmune or inflammatory disease, lymphoproliferative diseases and other complications (Chapel et al., 2008; Resnick et al., 2012). Studies have also shown that patients with the second type of CVID tend to have increased morbidity and poor prognosis (Chapel et al., 2008; Resnick et al., 2012). The cooccurrence of two contradictory situations,

immunodeficiency and autoimmunity, seems to be impenetrable. Shared genetic factors may be one reason, but this also adds to some extent to the complexity of CVID diagnosis (Xiao et al., 2014). Currently, the phenotypic spectrum of CVID syndromes or CVID-like disorders is expanding, and an increasing number of cases with autoimmune complications have been reported. Therefore, we reviewed the phenotypes of patients with CVID and autoimmunity enrolled in our hospital, and explored the clinical and genetic characteristics of these patients through WGS.

2 Materials and methods

2.1 Samples and collection

Patients were recruited into the study through the Department of Rheumatology and Immunology, West China Hospital, Sichuan University. All 16 patients and 26 first-degree healthy controls were Han Chinese. The patients met the ESID diagnostic criteria at the time of enrollment, and also fulfilled the classification criteria of coexisting autoimmunity or inflammation (Ameratunga et al., 2013). The cohort consisted of 4 unrelated patients and 12 core families with a clinical presentation of common variable immunodeficiency disorder (CVID). The baseline data in this study was collected retrospectively from our hospital database. The study protocol was approved by the Ethics Committee of West China Hospital, Sichuan University (ethics application reference number: 2020471). Consent was provided by all patients.

2.2 Sequencing and data analysis

WGS was performed on all subjects, including 16 probands and 26 first-degree relatives. Genomic DNA fragments with a size of approximately 350 bp were extracted from peripheral blood mononuclear cells. A DNA library was created by the Illumina paired-end protocol. Whole genome sequencing was performed using the Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA, United States) to produce 150-bp paired-end reads with a coverage of at least $4\% \times 99\%$ of the genome.

The sequenced data were subjected to basecall file conversion and demultiplexing by bcl2fastq software (Illumina) (Supplementary Figure S1). Low-quality reads in the resulting fastq data were eliminated using fastp software (Chen et al., 2018). Then, the reads were aligned to the reference human genome (hs37d5) by the Burrows-Wheeler Aligner (BWA), and duplicate reads were flagged by the sambamba tool (Li and Durbin, 2009; Tarasov et al., 2015).

SAMtools was used to call single-nucleotide variants (SNVs) and insertion deletions (InDels) to generate gVCF (Li et al., 2009). Filter thresholds for raw calls of SNVs and InDels: 1) Read depth >4; 2)

root-mean-square mapping quality of covering reads >30; 3) variant quality score >20.

The software used for annotation was ANNOVAR (June 8 2017). The annotation content includes the allele frequencies (AF), deleteriousness and conservation scores from public control datasets, enabling further screening and assessment of variants for possible pathogenicity (Wang et al., 2010).

2.3 Rare variant filtering

The rare variant screening was as follows: 1) Only SNVs occurring in exons or splice sites (10 bp splice junctions) were further analysed; 2) synonymous SNVs not related to amino acid changes predicted by dbscSNV or defined by RepeatMasker were also discarded small nonframeshifted (<10 bp) InDels in repetitive regions; 3) Screened for variants with AF less than 0.01 in 1000 genome data (1000 g-all), esp6500siv2-all (http://evs.gs.washington.edu/EVS), gnomAD data (gnomAD-all and gnomAD-EAS) (Auton et al., 2015); 4) Screened variants based on the scores of SIFT, Polyphen, MutationTaster and CADD software. Potentially deleterious variants were retained if more than half of the scores in these four pieces of software supported the deleteriousness of the variant (Karczewski et al., 2020).

We used Control-FREEC (v9.1) and LUMPY (v0.2.13) to detect copy number variable (CNV) and structural variable (SV) mutations, respectively (Boeva et al., 2012; Muona et al., 2015). Performing co-occurrence analysis on the detected CNVs and SVs, with the principle of screening for CNVs and SVs that are common to at least two or more patients and not found in normal individuals. If the reciprocal overlap of two mutations was more than 50%, it was considered the same CNV or SV.

2.4 Candidate gene analysis

We collected genes related to PID (including 13 known genes of CVID from the OMIM database) from the reported literature and disease database. PID-related genes were extracted after rare variant filtering analysis for candidate gene screening.

Given the characteristics of the pedigree, homozygous, compound heterozygous and de novo variants were considered to be candidate causal variations. Considering the complex inheritance of CVID, oligogenic inheritance patterns were also considered. For sporadic samples, the variants were reserved if only detected in the patient.

The American Academy of Medical genetics and Genomics (ACMG) was used to classify variants (pathogenicity, possible pathogenicity, uncertain significance (VUS), possible benign and benign) (Ri et al., 2015). At the same time, phenolyzer software was used to sort the selected candidate genes according to the disease correlation (Yang et al., 2015). We created protein-protein interactions (PPI) networks by String (version 11.0), focusing preferentially on proteins that interact with proteins encoded by known genes of CVID (Szklarczyk et al., 2019). KEGG was used to enrich candidate genes. In addition, the 3D structures of the mutant proteins from the RCSB protein database were displayed using PyMOL software (https://www.schrodinger.com/products/pymol),

and the hydrogen bond changes at the mutation sites of the wild-type and mutant proteins were compared (Goodsell et al., 2020). Sanger sequencing was performed to validate all candidate genes detected by WGS.

3 Results

3.1 Demographics and clinical parameters

We enrolled a total of 16 patients with CVID syndrome or CVID-like disorders in the Department of Rheumatology, West China Hospital, Sichuan University from June 2020 to May 2022, with their first-degree relatives as a control group. Finally, 4 sporadic samples, 12 probands and their 26 healthy family samples were collected (Figure 1A). The clinical characteristics of the patients are shown in Supplementary Table S1. Interestingly, in this study, 10/16 (62.5%) patients were complicated with connective tissue diseases, including systemic lupus erythematosus (SLE, 37.5%), Sjogren's syndrome (12.5%), undifferentiated connective tissue disease (6.25%), and eosinophilic granulomatosis with polyangiitis (6.25%). Others included inflammatory arthritis (12.5%), psoriatic arthritis (6.25%), immune thrombocytopenic purpura (6.25%), necrotizing lymphadenitis (6.25%) and pancytopenia (6.25%) (Figure 1B). Among them, one patient suffered from inflammatory arthritis and myelodysplastic syndrome.
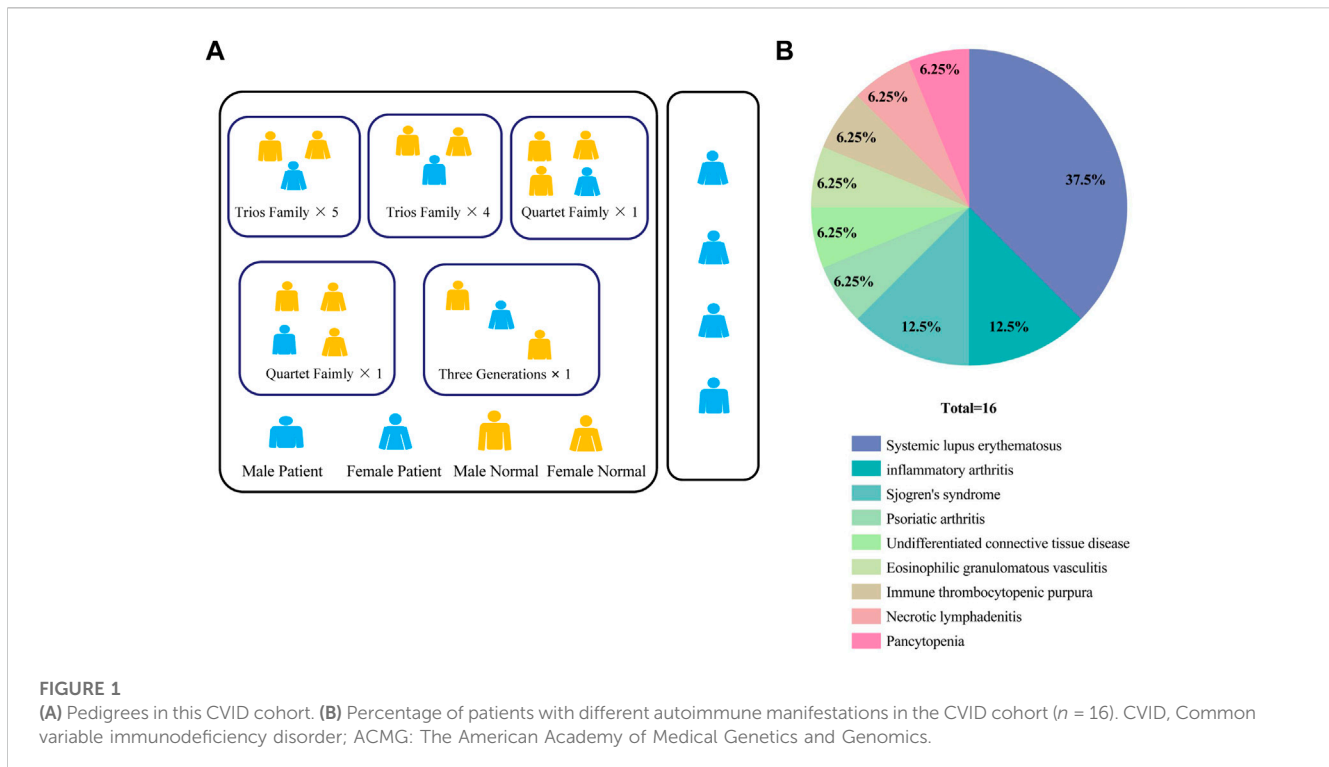
3.2 Genetic mutations associated with CVID found by whole genome sequencing

We first performed analysis by whole-genome sequencing of the patients and their first-degree healthy relatives and obtained a total of approximately 4,256 Gb of raw data. The average sequencing depth was 34 x, and the average mapping rate was 99.9%. In our cohort, 26 healthy control samples were selected from the patients' immediate family members. Therefore, we used two analysis strategies of pedigree and distribution to screen candidate genes.

3.2.1 SNV and InDel

Due to the potentially complex polygenicity and heterogeneity of CVID, we chose mild screening criteria (MAF less than 0.01) and detected a total of 9,578,118 SNVs and 2,758,295 InDels. Through low-frequency deleterious screens, such as frequency, functional region, and deleteriousness, a total of $9'148$ mutations were obtained, of which SNV and InDel were 8,171 and 977, respectively. The number of mutations carried by patients was 4,736 SNVs and 528 InDels (Supplementary Tables S2, S3). Meanwhile, to validate our data, we sorted out 551 PID and CVID-related genes (Supplementary Table S3) through the literature and databases (HGMD, OMIM, Phenolyzer), and then screened out the candidate genes related to CVID or PID in our patients.

A total of 9 CVID-related variants were found in our analysis (Supplementary Table S4), the most common of which was LPS-responsive beige-like ankyrin (LRBA), and 3 gene variants (p. G359D, p. I2744T and p. N2364S) were found (Figure 2A; Table 1). The three LRBA variants were all VUS according to the

FIGURE 1
(A) Pedigrees in this CVID cohort. **(B)** Percentage of patients with different autoimmune manifestations in the CVID cohort (*n* = 16). CVID, Common variable immunodeficiency disorder; ACMG: The American Academy of Medical Genetics and Genomics.

ACMG classification, which showed a very high level of evolutionary conservation (gerp ++ scores are 5.13, 5.8, and 5.82, respectively). We detected one of the *LRBA* variants (*p.G359D*) with extremely low population frequency as a homozygous, novel mutation that had not been previously reported. Evaluation of this variant by protein damage prediction software (SIFT, Polyphen2_HVAR, Polyphen2_HDIV, MutationTaster) revealed it to be deleterious.

Our analysis identified one patient who was heterozygous for the *p.R582G* variant in *RAG1*, the gene encoding the recombination activating 1 protein. A homozygous variant of CD40LG (p.H125P) was also found in one patient.In addition, four genes related to CVID were screened: *CTLA4 (p. P169A), NFKB1 (p. R533H), IL21R (c.867+10_867+28dup), and TNFRSF13B (p. R84T)*. Three of them were novel mutations, including *CTLA4 (p. P169A), NFKB1 (p. R533H)*, and *IL21R (c.867+10_867+28dup)*.

3.2.2 PID-associated genes

Given that the clinical phenotypes of CVID or CVID-like diseases often overlap with other primary immunodeficiency forms (PIDs), we hypothesized that PID-related genetic variants might exist in the patients in this study. In addition to CVID-related gene variants, we obtained a total of 23 PID-related variants. Our analysis found variants in the *PIK3R1 (c.336+1G>A)* and *PIK3CD (p. R408C)* genes. *PIK3R1 (c.336+1G>A)* is a *de novo*, heterozygous, splice mutation, which is classified as a pathogenic variant according to ACMG. In addition, we also confirmed that there were loss-of-function (LoF) variants in three other genes, *TAP1 (p. E46X), AIRE (c.652+1G>T), and GATA2 (p. R348X)*, respectively. Among them, *GATA2 (p. R348X)* is a stop-gain, heterozygous, pathogenic mutation. The clinical phenotype of the patient with this mutation is very similar to that of CVID. We have already

identified different transcript at this mutated site by whole exome sequencing of this patient (Liu et al., 2021).

MYB is an oncogene, but a heterozygous, *de novo* variant of this gene (*p. E132K*) emerged in our study. Two variants (*p. P30L, p. V233M*) were found in *IL10RA*, which have been reported to be closely associated with autoinflammatory diseases, such as inflammatory bowel disease (Kara et al., 2019). In addition, we identified heterozygous missense variants in 10 genes, including *CD8A, TNFSF1, PSTPIP1, IL17RA, TCIRG1, TIRAP, MYSM1, IGLL1, KMT2D, and STAT5B*, associated with T and B cells. Through our sorting of PID-related genes, we also found five possible variants of *TBK1, TIRAP, IL10RB, TFRC, and NCF2*. Finally, we screened a total of 28 candidate genes (32 loci in total). Meanwhile, 32 candidate loci were validated by Sanger sequencing. Details of all variants are shown in Table 1.

3.2.3 KEGG pathway enrichment analysis of candidate genes

CVID is clinically heterogeneous, with onset in different patients often caused by dysregulation of different signaling pathways. KEGG pathway enrichment was performed based on the candidate genes, and the enrichment results were statistically analysed by R software, with a priority given to immune-related and *p*-value significant pathways. KEGG analysis showed that the top 10 candidate gene-related pathways mainly included "primary immunodeficiency", "prolactin signaling pathway", "human cytomegalovirus infection", "JAK-STAT signaling pathway", "type B hepatitis", "T-cell receptor signaling pathway", "toll-like receptor signaling pathway", and so on (Supplementary Table S5, Figure 2B). We also performed disease ontology enrichment analysis through the STRING website based on 28 candidate genes. The results showed that the significantly enriched disease pathways included
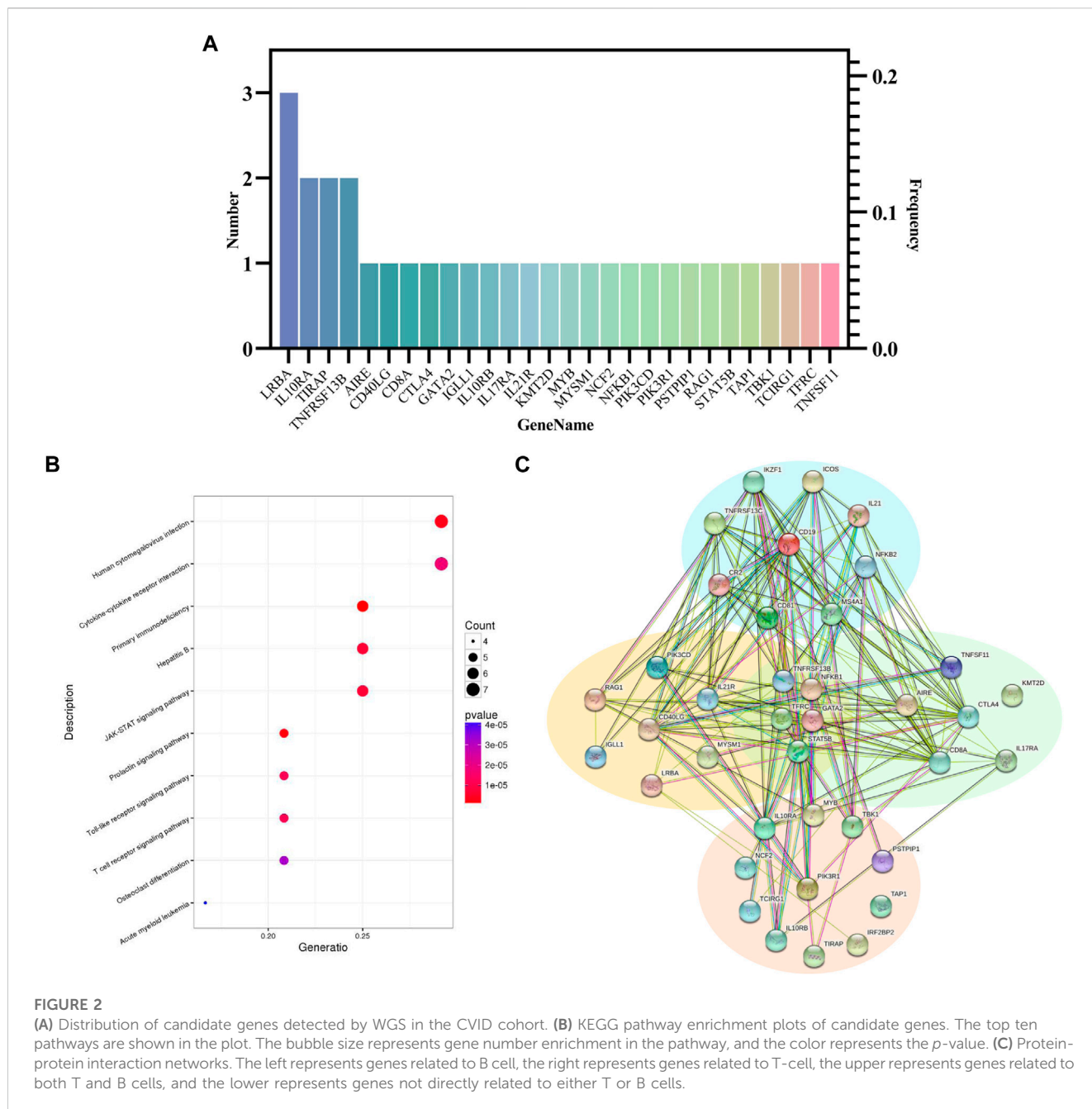
FIGURE 2
(A) Distribution of candidate genes detected by WGS in the CVID cohort. **(B)** KEGG pathway enrichment plots of candidate genes. The top ten pathways are shown in the plot. The bubble size represents gene number enrichment in the pathway, and the color represents the *p*-value. **(C)** Protein-protein interaction networks. The left represents genes related to B cell, the right represents genes related to T-cell, the upper represents genes related to both T and B cells, and the lower represents genes not directly related to either T or B cells.

common variable immunodeficiency and autoimmune diseases (Supplementary Table S6).

3.2.4 PPI analysis between CVID-related genes

Protein-protein Interaction networks (PPIs) are composed of proteins that interact with each other to participate in all aspects of life processes, such as biological signal transmission, gene expression regulation, energy and material metabolism, and cell cycle regulation. To further understand the protein-protein interaction, we constructed a visualized PPI network by combining 28 candidate genes and 13 genes associated with CVID in the OMIM database with the STRING database (Supplementary Table S7) As shown in Figures 2A, C total of 36 nodes were included in the PPI network, indicating that these proteins are highly connected with other proteins and interact functionally. In

addition, these proteins were divided into four clusters according to whether they were related to T and B lymphocytes. Coding genes for proteins are represented as nodes.

3.2.5 Protein 3D structure prediction

In addition, we performed 3D structure mapping of mutant and wild-type proteins using PyMOL software on candidate mutations for which protein 3D structure data were available in the PDB database, and further analysed whether the amino acid encoded by the mutation has changed hydrogen bonds. After excluding alternative splicing, stopgain, loss of function, incomplete protein structure analysis of missense mutations, and the absence of alterations in the H-bonds before and after mutation, we compared the wild-type and mutant structures of five candidate

TABLE 1 Candidate gene information.

Patient No	Normal name	Heredity	Gene	Function	Het/hom	AAChange	ACMG	Frenquency	Deleterious	gerp++
S1	—	sporadic	*RAG1*	missense	het	p.R582G	VUS	.|	D|B|B|D|12.19	2.13
S2#	—	sporadic	*LRBA*	missense	hom	p.G359D	VUS	.|	D|D|D|D|33	5.13
S3	—	sporadic	*CTLA4*	missense	het	p.P169A	VUS	0.000199681 | 0.00004471	D|B|P|D|12.47	<2
	—	sporadic	*TNFRSF13B*	missense	het	p.R84T	VUS	.|0.00002886	D|P|P|N|23	<2
S4	—	sporadic	*AIRE*	splicing	het	c.652+1G>T	VUS	0.000798722 | 0.00007992	1.0000,0.906 |-29.5134	3.61
F1	F1_Dad	oligogene	*TNFSF11*	missense	het	p.D317E	VUS	.|	D|P|D|D|	<2
	F1_Mom	oligogene	*PSTPIP1*	missense	het	p.F388L	VUS	0.00239617 | 0.00069299	T|B|B|D|11.9	<2
F2	F2_Mom	oligogene	*CD8A*	missense	het	p.R174Q	VUS	.|0.00001222	D|D|D|N|33	3.92
	—	*De novo*	*PIK3R1*	splicing	het	c.336+1G>A	Pathogenic	.|	1.0000,0.938 |-3.9537	5.1
	F2_Mom	oligogene	*TNFRSF13B*	missense	het	p.R84T	VUS	.|0.00002886	D|P|P|N|23	<2
F3	F3_Dad	oligogene	*LRBA*	missense	het	p.I2744T	VUS	0.000199681 | 0.00014432	D|P|P|D|24.5	5.8
	F3_Dad	oligogene	*PIK3CD*	missense	het	p.R408C	VUS	0.000599042 | 0.00018045	T|B|B|D|18.64	<2
	F3_Mom|F4_Mom	oligogene	*MYSM1*	missense	het	p.P760L	VUS	0.000599042 | 0.00053214	D|D|D|D|28.4	5.26
F4	F4_Dad	oligogene	*IL17RA*	missense	het	p.I763V	VUS	0.00279553 | 0.00054611	D|B|P|N|19.51	4.55
	F3_Mom|F4_Mom	oligogene	*MYSM1*	missense	het	p.P760L	VUS	0.000599042 | 0.00053214	D|D|D|D|28.4	5.26
F5	F5_Dad	oligogene	*TAP1*	stopgain	het	p.E46X	VUS	0.000199681 | 0.00010713	.|	2.01
	F5_Mom	oligogene	*IL21R*	splicing	het	c.867 + 10_867 + 28dup	VUS	.|0.00000409	.|	<2
F6	F6_Dad	oligogene	*IL10RA*	missense	het	p.P30L	VUS	.|0.00007214	D|B|P|N|12.45	3.08
	F6_Mom	oligogene	*TIRAP*	missense	het	p.D96N	VUS	0.00139776 | 0.00258373	D|D|D|D|29.4	5.41
F7	F7_Dad	oligogene	*TBK1*	missense	het	p.V421L	VUS	0.000199681 | 0.00001353	T|B|B|D|20.6	5.27
	F7_Son	oligogene	*IGLL1*	missense	het	p.T146M	VUS	0.00199681 | 0.00109133	D|P|P|N|	<2
F8	F8_Dad	oligogene	*LRBA*	missense	het	p.N2364S	VUS	.|0.00000407	T|B|B|D|23.7	5.82
	F8_Mom	oligogene	*IL10RA*	missense	het	p.V233M	VUS	0.0071885 | 0.00199517	D|P|D|N|19.48	3.27
	F8_Mom	oligogene	*IL10RB*	missense	het	p.A44V	VUS	0.00119808 | 0.00038956	D|P|D|N|22.8	3.47
F9	—	*De novo*	*MYB*	missense	het	p.E132K	VUS	.|	D|D|D|D|31	5.81
F10#	F10-Mom##	recessive	*CD40LG*	missense	hom	p.H125P	Likely Pathogenic	.|	D|P|D|D|25	5.5
F11	—	*De novo*	*GATA2*	stopgain	het	p.R348X	Pathogenic	.|	.|-0.2054	4.05
	F11_Mom	oligogene	*TFRC*	missense	het	p.D70A	VUS	.|0.00004117	T|B|B|D|10.37	<2
	F11_Mom	oligogene	*NFKB1*	missense	het	p.R533H	VUS		D|D|D|D|29	4.5

(Continued on following page)

TABLE 1 (*Continued*) Candidate gene information.

Patient No	Normal name	Heredity	Gene	Function	Het/hom	AAChange	ACMG	Frenquency	Deleterious	gerp++
								0.000199681 | 0.00053528		
	F11_Dad	oligogene	*KMT2D*	missense	het	p.T245S	VUS	.|0.00000455	T|D|D|D|14.18	5.05
	—	*De novo*	*STAT5B*	missense	het	p.N642H	VUS	.|0.00000406	T|D|D|D|24.3	5.33
F12	F12_Mom	oligogene	*NCF2*	missense	het	p.F116S	VUS	.|	D|D|D|D|32	5.56
	F12_Dad	oligogene	*TCIRG1*	missense	het	p.R60W	VUS	0.00159744 | 0.00023873	D|D|D|D|33	4.9
	F12_Mom	oligogene	*TIRAP*	missense	het	p.E132K	VUS	0.00159744 | 0.00030251	D|D|D|D|29.3	5.67

Note: "#" represents a homozygous mutation; "##" represents that the patient's mother carries a heterozygous mutation in CD40LG. "/" stands for none. AA Change represents amino acid change. ACMG: The American Academy of medical genetics and genomics. RS: reference snp. VUS: variants of uncertain significance. "." the black dots in column H (RS) represent that this locus is not included in dbSNP, database, "\|" (Frenquency) the black dots in column I represent that there is no mutation in the corresponding population database, indicating that the mutation frequency of this site is extremely low. "\|" represents 1000 g and GnomAD, database respectively. Deleterious: D: damaging/disease causing, P: possibly damaging, T: tolerated, N: polymorphism, B: benign.

genes, including *MYB* (AlphaFold), *NFKB1* (AlphaFold), *RAG1* (AlphaFold), *TIRAP* (3UB2), and *NCF2* (1HH8), to elucidate the structural changes in the protein. Since *MYB, NFKB1* and *RAG1* had no analytical results in PDB, we compared them based on the 3D structures predicted by the AlphaFold algorithm. Among them, the mutation occurred in exon 5 of the *MYB* gene (*p. E132K*), and the glutamic acid at position 132 is replaced by lysine, resulting in the disappearance of the hydrogen bond with arginine at position 131 and asparagine at position 136 (distance were 2.9 Å and 2.9 Å, respectively). This variant of *MYB* is haploinsufficient, and the patient carrying the mutant was a 30-year-old female with CVID and myelodysplastic syndrome, who eventually died of septic shock. Although this gene variant (*p. E132K*) has never been reported to be associated with CVID, ACMG classified it as a VUS, with a gerp++ score of 5.81, predicted to be deleterious. Therefore, we speculated that it may be a candidate gene for CVID, but further research is needed to verify its function in pathogenesis.

The mutation in exon 15 of the gene encoding *NFKB1* (*p. R533H*) results in the disappearance of the 2.9 Å long hydrogen bond with leucine at position 529, and the mutated histidine itself forms a 3.2 Å long hydrogen bond. The mutation located in exon 2 of the *RAG1* gene (*p. R582G*) replaced arginine at position 582 with glycine in the original protein structure, resulting in the disappearance of three hydrogen bonds (2.8 Å, 2.9 Å, 3.4 Å in distance) with aspartic acid at position 588 and glutamine at position 584, respectively. Mutations occurred in exon 4 of the *TIRAP* gene (*p. D96N*), asparagine replaces aspartic acid at position 96, resulting in the disappearance of a long 2.9 Å hydrogen bond with threonine at position 148. Asparagine itself formed an additional hydrogen bond, and the distance of the hydrogen bond with the serine at position 93 changed from 2.9 Å to 2.8 Å. A mutation in exon 3 of the *NCF2* gene (*p. F116S*) resulted in the replacement of phenylalanine at position 116 by serine, forming a 2.8 Å hydrogen bond with leucine at position 114 (Protein structures can be found in Figure 3).

3.3 CNV and SV

The results of CNV and SV are shown in Supplementary Tables S8-1, S8-2. Most of the CNVs and SVs shared by more than 2 patients and absent from normal individuals were located in intronic regions. However, no genes related to the clinical phenotype of the patients were found in CNV and SV that cover exons. Therefore, the analysis of pathogenic variants based on CNV and SV did not find the expected candidate variants.

4 Discussion

In the present study, WGS was applied to identify genetic mutations in 16 CVID or CVID-like patients with autoimmune features. There were more females than males in our study (10:6), which was lower than that found in a previous study from the rheumatology group of the USIDNET cohort (3.3:1) (Chapel et al., 2008). The results from a study in the United States that enrolled 870 CVID patients revealed that 5.9% had rheumatic disease, with inflammatory arthritis being the most common, and others including SLE, Sjogren's syndrome, Raynaud's syndrome, vasculitis, mixed connective tissue disease, and psoriasis (Gutierrez et al., 2018). However, in a recent systematic review, blood autoimmune diseases were the most common in CVID with autoimmunity (Rizvi et al., 2020). Contrary to their findings, the most common connective tissue disease in this study was SLE, which may be related to ethnic differences between the study populations.

In this study, candidate genes of patients with SLE included *TNFSF11, PSTPIP1, TAP1, IL21R, IL10RA, TIRAP, TBK1, IGLL1, LRBA, IL10RB, NCF2, and TCIRG1*, while *LRBA, IL10RA, TIRAP, TAP1, IL10RB and IL21R* have been reported to be associated with the clinical phenotype of SLE (Huang et al., 2004; Webb et al., 2009; Peng et al., 2013; Kara et al., 2019; Liphaus et al., 2020; Tayel et al., 2021). In addition, we found correlations between *TNFSF11, PSTPIP1, TBK1, IGLL1, LRBA, IL10RA, IL10RB* and 13 CVID-related genes recorded in the OMIM database by PPI. Therefore, we hypothesized that there is an overlap between immunodeficiency and autoimmune diseases (such as SLE) in genetic background. Unfortunately, research evidence in this area is scarce and needs further exploration.

LRBA deficiency has been reported to cause a CVID-like clinical phenotype (Costagliola et al., 2021). LRBA deficiency
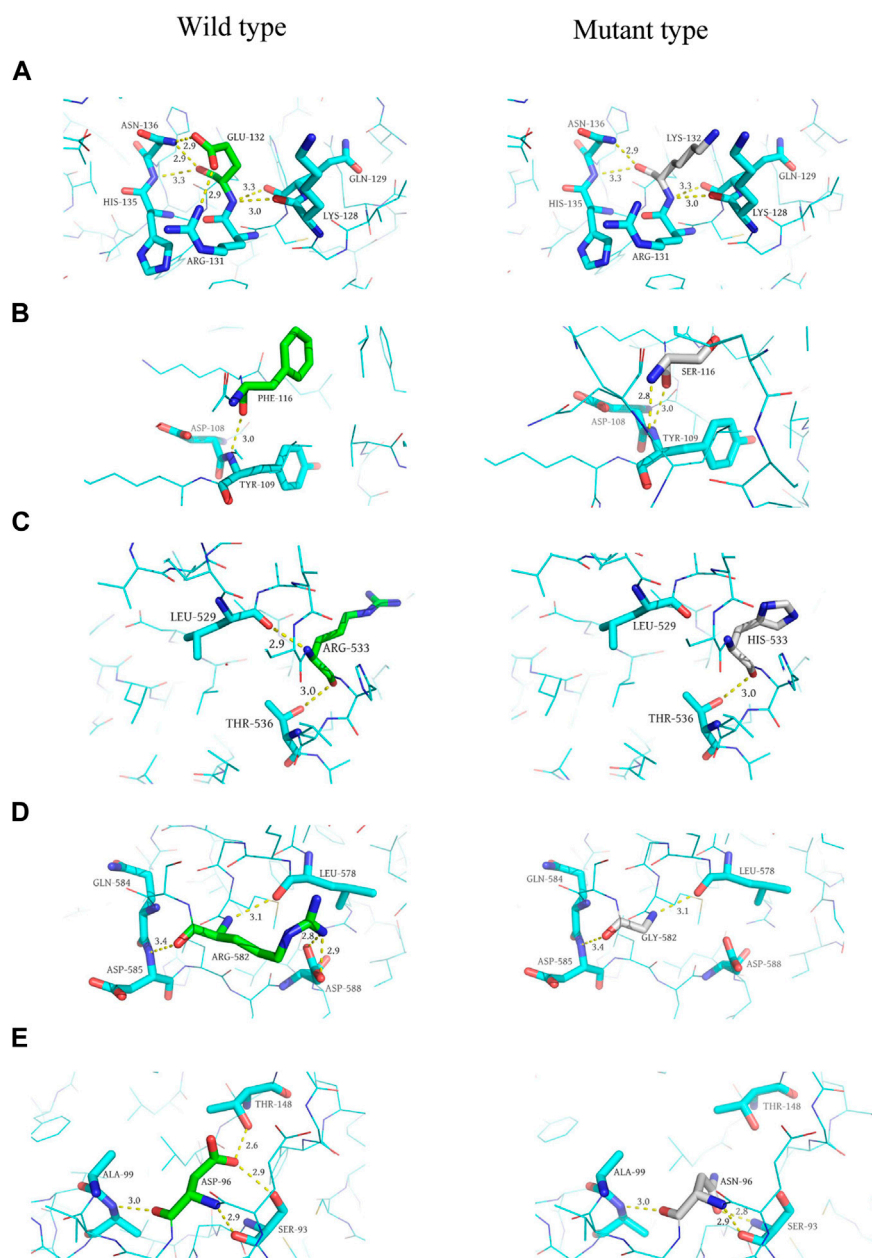
FIGURE 3
Predicted 3D structure of the protein. **(A, C, E, G, and I)** are the wild-type (left) 3D structures of MYB, NCF2, NFKB1, RAG1, and TIRAP, respectively. **(B, D, F, H, and J)** are the mutant (right) 3D structures of MYB, NCF2, NFKB1, RAG1, and TIRAP, respectively.

not only shows the characteristics of immunodeficiency, such as repeated infection, but is also often combined with autoimmune cytopenia, lymphoid hyperplasia, rheumatoid arthritis, inflammatory bowel disease and so on. A study in Iran showed that the most frequently mutated gene in CVID patients with an autoimmune phenotype was *LRBA*, which is consistent with our findings (Asgardoon et al., 2020). Three patients with *LRBA* variants in our study had psoriatic arthritis, EGPA, and SLE. The *LRBA* variant carried by patients with psoriatic arthritis is a homozygous, novel mutation (*p. G359D*) that has not been previously reported. LRBA is associated with protein kinase A and can participate in intracellular transport. Moreover, LRBA

plays a key role in promoting and maintaining the homeostasis of cytotoxic T lymphocyte antigen-4 (CTLA-4) in Tregs (foxp3+t regulatory cells) and activated T-cells (Burnett et al., 2017; Lo and Abdel-Motal, 2017). Many studies have considered variants in *NFKB1* as a monogenic cause of CVID, and considered heterozygous loss-of-function variants in *NFKB1* to be the most common known monogenic cause of CVID (Fliegauf et al., 2015; Kaustio et al., 2017; Tuijnenburg et al., 2018). However, in our study, there was a 30-year-old male patient with not only *NFKB1* missense mutations but also mutations in the *GATA2*, *TFRC*, *KMT2D*, and *STAT5B genes*. Therefore, the monogenic model cannot account for the complex clinical heterogeneity of CVID,

and more than 2 candidate gene loci each were present in at least 11/16 patients in our study. We also demonstrated the interaction between candidate genes by PPI network analysis, including the interaction between LRBA and CTLA-4 protein. (Figure 2C).

KEGG pathway enrichment analysis based on candidate genes showed that the highest ranking was primary immunodeficiency. Our analysis identified mutations in multiple genes involved in this pathway (AIRE, CD40LG, IGLL1, RAG1, TAP1, and TNFRSF13B). CD40L is mainly expressed on the surface of activated CD4+ T cells. It controls the proliferation and differentiation of B cells through interaction with CD40, leading to the conversion of immunoglobulin classes (Bhushan and Covey, 2001). CD40L mutations lead to CD40 ligand deficiency, which can lead to a rare X-linked recessive primary immunodeficiency disorder, hyper-IgM syndrome (HIGM), characterized by normal or elevated serum IgM levels and reduced or absent levels of other immunoglobulins (Lee et al., 2005). A homozygous missense mutation in CD40LG (p. H125P) was detected in a young male patient with a CVID phenotype in our cohort. Although this variant has not been previously reported, a different amino acid change at this site (c.374A>G:p. H125R) was found in a previous study (Katz et al., 1996). According to ACMG, the mutation is likely pathogenic. He suffered from recurrent pulmonary infections, severe diarrhea, and concomitant C. malfini infection. This patient had no family history of immunodeficiency, and his mother was a heterozygous carrier. Finally, according to the WGS analysis, the patient was diagnosed with X-linked hyper-IgM syndrome, and he died of septic shock. Typical HIGM has an early onset with recurrent infections and many complications. The onset of atypical mutations is usually late and can occur in young adulthood. This clinical phenotype is milder, has fewer complications, and is hard to distinguish from CVID (França et al., 2018).

Transcription factors encoded by AIRE play a critical role in the development of central immune tolerance, which promotes negative selection of T cells in the thymus, builds thymic microarchitecture, and induces specific subsets of regulatory T-cells (Bruserud et al., 2016). In our cohort, a middle-aged female patient with this gene mutation (LoF) had a clinical phenotype of CVID with UCTD. Studies have reported that mutations in this gene can lead to autoimmune injuries such as autoimmune polyglandular syndrome type 1 (APS-1) (Eldershaw et al., 2011; Bruserud et al., 2016). TAP plays an important role in the immune system to recognize harmful foreign invasion, while also having a tracking effect on tumor cells (Tabassum et al., 2021). Any defect in the TAP1 gene results in insufficient tumor tracking, and it may also be associated with some inflammatory or autoimmune diseases (Membrive Jiménez et al., 2021).

Mutations in the PIK3R1 and PIK3CD genes can lead to activated phosphoinositide 3-kinase δ syndrome (APDS), a primary immunodeficiency affecting both humoral and cellular immunity, with some phenotypic similarities to CVID (Bloomfield et al., 2021). In this study, we also reported a 19-year-old male with PIK3R1 gene mutation. He presented the CVID immunophenotype, accompanied by a significant decrease in CD19+CD5+ B cells (26/μL, normal range: 175–332/μL) and CD4+ T cells (272/μL, normal range: 471–1220/μL). He developed recurrent necrotizing lymphadenitis and died of sudden gastrointestinal bleeding at the age of 20. In pathway analysis, we found that this gene was involved in the JAK-STAT signaling pathway, T-cell receptor signaling pathway and Toll-like receptor

signaling pathway. In addition to PIK3R1, genes involved in T-cell receptor signaling include CD40LG, CTLA4, NFKB1, and PIK3CD, and genes involved in Toll-like receptor signaling include NFKB1, PIK3CD, TBK1, and TIRAP. We found that these genetic variants can be involved in multiple pathways simultaneously, suggesting that they may contribute to B-cell dysfunction by affecting T-B-cell interactions as well as innate immunity.

Previous studies have also reported that TIRAP (p. D96N, p. E132K) play a role in innate immunity (George et al., 1950; An et al., 2011). In addition, two patients carried these two missense mutations in the TIRAP gene. Interestingly, both patients exhibited the clinical phenotype of systemic lupus erythematosus. The 3D protein structure of this variant of TIRAP (p. D96N) was found to be altered, highly conserved (gerp++ score: 5.41), and predicted to be deleterious. The protein encoded by this TIRAP gene, is related to the TLR4 signaling pathway and plays an important role not only in innate immunity but also in autoimmunity (Summers et al., 2010; Liu et al., 2014). Studies have found that the TIRAP single nucleotide polymorphism is also a protective factor in systemic lupus erythematosus (Castiblanco et al., 2008).

The pathogenesis of CVID is related to defect in B-cell function and antibody production. In this study, 2/16 cases (12.5%) had the same TNFRSF13B mutation (p. R84T), which was reported in previous studies on CVID (de Valles-Ibáñez et al., 2018). The mutation of TNFRSF13B, which leads to defects in BAFF and April (proliferation inducing ligand) receptor (TACI), is one of the earliest mutations related to CVID (de Valles-Ibáñez et al., 2018). TACI deficiency may cause inefficient class-switching recombination of B cells, which could lead to decreased IgG or IgA levels, consistent with the immunological phenotype of the two patients in our cohort (Castigli et al., 2005; Mohammadi et al., 2009). In addition, TNFRSF13B/TACI mutation is also closely related to the development of SLE, but heterozygous mutations are also present in a subset of healthy individuals (Salzer et al., 2007). Therefore, scholars believe that TNFRSF13B mutations have no determinant role in CVID and can be considered modifier genes or risk factors for CVID (de Valles-Ibáñez et al., 2018).

The genetic background of CVID is very complex, and there is no definite and uniform genetic diagnosis model. To solve this query, WGS could detect a comprehensive range of variation sites, and can also detect splicing sites that are far away from the exonic region and affect variable splicing [e.g., IL21R (c.867+10_867+28dup)]. A study of Oxford University Hospital by WGS combined with RNA-sequencing in patients with CVID, the results of which identified the polygenicity of CVID and suggested a number of individual specific mutations (van Schouwenburg et al., 2015). In addition, there are mainly focused on the SNPs of CVID, the evidence is less about the genetic susceptibility of CNVs and SV. Keller M et al. analyzed the association between CNVs burden and CVID, and found that the increase in CNVs burden in CVID was static and intrinsic (Keller et al., 2014). And, Orange JS et al. found that CNVs had multiple novel susceptibility loci for CVID in their study, providing new insights into the genetic variation of CVID (Orange et al., 2011). In this study, we also attempted to analyze the association of CNVs and SV with genetic susceptibility to CVID, but the results were regrettable.

However, there are still some limitations. First, our current sample size is relatively small, and we need to expand the sample

size in the future to explore more reliable genetic features. Second, this study is a single-center study. Third, if a meta-analysis of GWAS in different regions can be performed, the conclusions can be made more convincing. This is also what we originally intended to do but did not achieve. Because there are few WGS studies on CVID, data from other studies were not available.

In conclusion, this study provided genetic variants in Chinese adults with autoimmune phenotypes with CVID or CVID-like syndrome. Monogenic and polygenic patterns may participate in the pathogenesis of CVID, and there may be a shared genetic background between immunodeficiency and autoimmunity.

Data availability statement

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Chen et al., 2021) in National Genomics Data Center (Xue et al., 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA005867) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human.

Ethics statement

The studies involving humans were approved by the Ethics Committee of West China Hospital, Sichuan University [ethics application reference number: 2020 (471)]. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZL was responsible for experiment design, sample collection, data analysis, and writing the manuscript. CL, PQ, RC, and YC were responsible for data collection, analysis and revising the manuscript. ZL and YjL are responsible for sample sequencing and data analysis. XG and ZY were responsible for suggestion during the experiment performance. HY and YiL were responsible for revising the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

Author YL and XG are employed by Novogene Co., Ltd., Beijing.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1209988/full#supplementary-material

References

Ameratunga, R., Woon, S., Gillis, D., Koopmans, W., and Steele, R. (2013). New diagnostic criteria for common variable immune deficiency (CVID), which may assist with decisions to treat with intravenous or subcutaneous immunoglobulin. *Clin. Exp. Immunol.* 174 (2), 203–211. doi:10.1111/cei.12178

Ameratunga, R., Lehnert, K., Woon, S., Gillis, D., Bryant, V., Slade, C., et al. (2018). Review: diagnosing common variable immunodeficiency disorder in the era of genome sequencing. *Clin. Rev. allergy & Immunol.* 54 (2), 261–268. doi:10.1007/s12016-017-8645-0

Ameratunga, R. (2018). Assessing disease severity in common variable immunodeficiency disorders (CVID) and CVID-like disorders. *Front. Immunol.* 9, 2130. doi:10.3389/fimmu.2018.02130

An, Y., Ohnishi, H., Matsui, E., Funato, M., Kato, Z., Teramoto, T., et al. (2011). Genetic variations in MyD88 adaptor-like are associated with atopic dermatitis. *Int. J. Mol. Med.* 27 (6), 795–801. doi:10.3892/ijmm.2011.645

Asgardoon, M., Azizi, G., Yazdani, R., Sohani, M., Pashangzadeh, S., Kalantari, A., et al. (2020). Monogenic primary immunodeficiency disorder associated with common variable immunodeficiency and autoimmunity. *Int. archives allergy Immunol.* 181 (9), 706–714. doi:10.1159/000508817

Auton, A., Brooks, L., Durbin, R., Garrison, E., Kang, H., Korbel, J., et al. (2015). A global reference for human genetic variation. *Nature* 526 (7571), 68–74. doi:10.1038/nature15393

Bhushan, A., and Covey, L. (2001). CD40:CD40L interactions in X-linked and non-X-linked hyper-IgM syndromes. *Immunol. Res.* 24 (3), 311–324. doi:10.1385/ir:24:3:311

Bloomfield, M., Klocperk, A., Zachova, R., Milota, T., Kanderova, V., and Sediva, A. (2021). Natural course of activated phosphoinositide 3-kinase delta syndrome in childhood and adolescence. *Front. Pediatr.* 9, 697706. doi:10.3389/fped.2021.697706

Boeva, V., Popova, T., Bleakley, K., Chiche, P., Cappo, J., Schleiermacher, G., et al. (2012). Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinforma. Oxf. Engl.* 28 (3), 423–425. doi:10.1093/bioinformatics/btr670

Bogaert, D., Dullaers, M., Lambrecht, B., Vermaelen, K., De Baere, E., and Haerynck, F. (2016). Genes associated with common variable immunodeficiency: one diagnosis to rule them all? *J. Med. Genet.* 53 (9), 575–590. doi:10.1136/jmedgenet-2015-103690

Bruserud, Ø., Oftedal, B., Wolff, A., and Husebye, E. (2016). AIRE-mutations and autoimmune disease. *Curr. Opin. Immunol.* 43, 8–15. doi:10.1016/j.coi.2016.07.003

Burnett, D., Parish, I., Masle-Farquhar, E., Brink, R., and Goodnow, C. (2017). Murine LRBA deficiency causes CTLA-4 deficiency in Tregs without progression to immune dysregulation. *Immunol. Cell. Biol.* 95 (9), 775–788. doi:10.1038/icb.2017.50

Castiblanco, J., Varela, D., Castaño-Rodríguez, N., Rojas-Villarraga, A., Hincapié, M., and Anaya, J. (2008). TIRAP (MAL) S180L polymorphism is a common protective factor against developing tuberculosis and systemic lupus erythematosus. *Infect. Genet. Evol.* 8 (5), 541–544. journal of molecular epidemiology and evolutionary genetics in infectious diseases. doi:10.1016/j.meegid.2008.03.001

Castigli, E., Wilson, S., Garibyan, L., Rachid, R., Bonilla, F., Schneider, L., et al. (2005). TACI is mutant in common variable immunodeficiency and IgA deficiency. *Nat. Genet.* 37 (8), 829–834. doi:10.1038/ng1601

Chapel, H., Lucas, M., Lee, M., Bjorkander, J., Webster, D., Grimbacher, B., et al. (2008). Common variable immunodeficiency disorders: division into distinct clinical phenotypes. *Blood* 112 (2), 277–286. doi:10.1182/blood-2007-11-124545

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinforma. Oxf. Engl.* 34 (17), i884–i890. doi:10.1093/bioinformatics/bty560

Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., et al. (2021). The genome sequence archive family: toward explosive data growth and diverse data types. *Genomics, Proteomics & Bioinformatics* 19 (4), 578–583. doi:10.1016/j.gpb.2021.08.001

Costagliola, G., Cappelli, S., and Consolini, R. (2021). Autoimmunity in primary immunodeficiency disorders: an updated review on pathogenic and clinical implications. *J. Clin. Med.* 10 (20), 4729. doi:10.3390/jcm10204729

de Valles-Ibáñez, G., Esteve-Solé, A., Piquer, M., González-Navarro, E., Hernandez-Rodriguez, J., Laayouni, H., et al. (2018). Evaluating the genetics of common variable immunodeficiency: monogenetic model and beyond. *Front. Immunol.* 9, 636. doi:10.3389/fimmu.2018.00636

Edwards, E., Bosco, J., Ojaimi, S., O'Hehir, R., and van Zelm, M. (2021). Beyond monogenetic rare variants: tackling the low rate of genetic diagnoses in predominantly antibody deficiency. *Cell. Mol. Immunol.* 18 (3), 588–603. doi:10.1038/s41423-020-00520-8

Eldershaw, S., Sansom, D., and Narendran, P. (2011). Expression and function of the autoimmune regulator (Aire) gene in non-thymic tissue. *Clin. Exp. Immunol.* 163 (3), 296–308. doi:10.1111/j.1365-2249.2010.04316.x

Fliegauf, M., Bryant, V., Frede, N., Slade, C., Woon, S., Lehnert, K., et al. (2015). Haploinsufficiency of the NF-κB1 subunit p50 in common variable immunodeficiency. *Am. J. Hum. Genet.* 97 (3), 389–403. doi:10.1016/j.ajhg.2015.07.008

França, T., Leite, L., Maximo, T., Lambert, C., Zurro, N., Forte, W., et al. (2018). A novel *de novo* mutation in the CD40 ligand gene in a patient with a mild X-linked hyper-IgM phenotype initially diagnosed as CVID: new aspects of old diseases. *Front. Pediatr.* 6, 130. doi:10.3389/fped.2018.00130

George, J., Kubarenko, A., Rautanen, A., Mills, T., Colak, E., Kempf, T., et al. (2010). MyD88 adaptor-like D96N is a naturally occurring loss-of-function variant of TIRAP. *J. Immunol.* 184 (6), 3025–3032. Baltimore, Md 1950. doi:10.4049/jimmunol.0901156

Goodsell, D., Zardecki, C., Di Costanzo, L., Duarte, J., Hudson, B., Persikova, I., et al. (2020). RCSB Protein Data Bank: enabling biomedical research and drug discovery. *Protein Sci.* 29 (1), 52–65. doi:10.1002/pro.3730

Gutierrez, M., Sullivan, K., Fuleihan, R., Bingham, C., and Bingham, C. O., 3rd (2018). Phenotypic characterization of patients with rheumatologic manifestations of common variable immunodeficiency. *Seminars arthritis rheumatism* 48 (2), 318–326. doi:10.1016/j.semarthrit.2018.02.013

Huang, C., Hang, L., Chen, C., Wu, J., and Tsai, F. (2004). Polymorphisms of TAP1 transporter genes in Chinese patients with systemic lupus erythematosus in Taiwan. *Rheumatol. Int.* 24 (3), 130–132. doi:10.1007/s00296-003-0338-8

Kara, S., Pirela-Morillo, G., Gilliam, C., and Wilson, G. (2019). Identification of novel susceptibility genes associated with seven autoimmune disorders using whole genome molecular interaction networks. *J. Autoimmun.* 97, 48–58. doi:10.1016/j.jaut.2018.10.002

Karczewski, K., Francioli, L., Tiao, G., Cummings, B., Alföldi, J., Wang, Q., et al. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* 581 (7809), 434–443. doi:10.1038/s41586-020-2308-7

Katz, F., Hinshelwood, S., Rutland, P., Jones, A., Kinnon, C., and Morgan, G. (1996). Mutation analysis in CD40 ligand deficiency leading to X-linked hypogammaglobulinemia with hyper IgM syndrome. *Hum. Mutat.* 8 (3), 223–228. doi:10.1002/(SICI)1098-1004(1996)8:3<223::AID-HUMU5>3.0.CO;2-A

Kaustio, M., Haapaniemi, E., Göös, H., Hautala, T., Park, G., Syrjänen, J., et al. (2017). Damaging heterozygous mutations in NFKB1 lead to diverse immunologic phenotypes. *J. allergy Clin. Immunol.* 140 (3), 782–796. doi:10.1016/j.jaci.2016.10.054

Keller, M., Glessner, J., Resnick, E., Perez, E., Chapel, H., Lucas, M., et al. (2014). Burden of copy number variation in common variable immunodeficiency. *Clin. Exp. Immunol.* 177 (1), 269–271. doi:10.1111/cei.12255

Lee, W., Torgerson, T., Schumacher, M., Yel, L., Zhu, Q., and Ochs, H. (2005). Molecular analysis of a large cohort of patients with the hyper immunoglobulin M (IgM) syndrome. *Blood* 105 (5), 1881–1890. doi:10.1182/blood-2003-12-4420

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma. Oxf. Engl.* 25 (14), 1754–1760. doi:10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinforma. Oxf. Engl.* 25 (16), 2078–2079. doi:10.1093/bioinformatics/btp352

Liphaus, B., Caramalho, I., Rangel-Santos, A., Silva, C., Demengeot, J., and Carneiro-Sampaio, M. (2020). LRBA deficiency: a new genetic cause of monogenic lupus. *Ann. rheumatic Dis.* 79 (3), 427–428. doi:10.1136/annrheumdis-2019-216410

Liu, Y., Yin, H., Zhao, M., and Lu, Q. (2014). TLR2 and TLR4 in autoimmune diseases: a comprehensive review. *Clin. Rev. allergy & Immunol.* 47 (2), 136–147. doi:10.1007/s12016-013-8402-y

Liu, Z., Qing, P., Zhao, Y., Liu, Y., and Marion, T. (2021). Combined mutation of the GATA2 gene and STAT5B gene in a patient with hypogammaglobulinemia and autoimmunity. *Tohoku J. Exp. Med.* 255 (2), 143–146. doi:10.1620/tjem.255.143

Lo, B., and Abdel-Motal, U. (2017). Lessons from CTLA-4 deficiency and checkpoint inhibition. *Curr. Opin. Immunol.* 49, 14–19. doi:10.1016/j.coi.2017.07.014

Membrive Jiménez, C., Pérez Ramírez, C., Sánchez Martín, A., Vieira Maroun, S., Arias Santiago, S., Ramírez Tortosa, M., et al. (2021). Influence of genetic polymorphisms on response to biologics in moderate-to-severe psoriasis. *J. personalized Med.* 11 (4), 293. doi:10.3390/jpm11040293

Mohammadi, J., Liu, C., Aghamohammadi, A., Bergbreiter, A., Du, L., Lu, J., et al. (2009). Novel mutations in TACI (TNFRSF13B) causing common variable immunodeficiency. *J. Clin. Immunol.* 29 (6), 777–785. doi:10.1007/s10875-009-9317-5

Muona, M., Berkovic, S., Dibbens, L., Oliver, K., Maljevic, S., Bayly, M., et al. (2015). A recurrent *de novo* mutation in KCNC1 causes progressive myoclonus epilepsy. *Nat. Genet.* 47 (1), 39–46. doi:10.1038/ng.3144

Orange, J., Glessner, J., Resnick, E., Sullivan, K., Lucas, M., Ferry, B., et al. (2011). Genome-wide association identifies diverse causes of common variable immunodeficiency. *J. allergy Clin. Immunol.* 127 (6), 1360–1367. doi:10.1016/j.jaci.2011.02.039

Peng, H., Liu, C., Zhou, M., Wen, P., Zhang, M., Qiu, L., et al. (2013). IL-10RB rs2834167 (A/G) polymorphism is associated with the susceptibility to systemic lupus erythematosus: evidence from a study in Chinese Han population. *Inflammation* 36 (6), 1218–1224. doi:10.1007/s10753-013-9658-3

Peng, X., Caballero-Oteyza, A., and Grimbacher, B. (2023). Common variable immunodeficiency: more pathways than roads to rome. *Annu. Rev. pathology* 18, 283–310. doi:10.1146/annurev-pathmechdis-031521-024229

Resnick, E., Moshier, E., Godbold, J., and Cunningham-Rundles, C. (2012). Morbidity and mortality in common variable immune deficiency over 4 decades. *Blood* 119 (7), 1650–1657. doi:10.1182/blood-2011-09-377945

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al., and ACMG Laboratory Quality Assurance Committee (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of medical genetics and genomics and the association for molecular pathology. *Genet. Med.* 17 (5), 405–424. official journal of the American College of Medical Genetics. doi:10.1038/gim.2015.30

Rizvi, F., Zainaldain, H., Rafiemanesh, H., Jamee, M., Hossein-Khannazer, N., Hamedifar, H., et al. (2020). Autoimmunity in common variable immunodeficiency: a systematic review and meta-analysis. *Expert Rev. Clin. Immunol.* 16 (12), 1227–1235. doi:10.1080/1744666x.2021.1850272

Salzer, U., Birmelin, J., Bacchelli, C., Witte, T., Buchegger-Podbielski, U., Buckridge, S., et al. (2007). Sequence analysis of TNFRSF13b, encoding TACI, in patients with systemic lupus erythematosus. *J. Clin. Immunol.* 27 (4), 372–377. doi:10.1007/s10875-007-9094-y

Sneller, M., Strober, W., Eisenstein, E., Jaffe, J., and Cunningham-Rundles, C. (1993). NIH conference. New insights into common variable immunodeficiency. *Ann. Intern. Med.* 118 (9), 720–730. doi:10.7326/0003-4819-118-9-199305010-00011

Summers, S., Hoi, A., Steinmetz, O., O'Sullivan, K., Ooi, J., Odobasic, D., et al. (2010). TLR9 and TLR4 are required for the development of autoimmunity and lupus nephritis in pristane nephropathy. *J. Autoimmun.* 35 (4), 291–298. doi:10.1016/j.jaut.2010.05.004

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids Res.* 47, D607–D613. doi:10.1093/nar/gky1131

Tabassum, A., Samdani, M., Dhali, T., Alam, R., Ahammad, F., Samad, A., et al. (2021). Transporter associated with antigen processing 1 (TAP1) expression and prognostic analysis in breast, lung, liver, and ovarian cancer. *J. Mol. Med. (Berlin, Ger.* 99 (9), 1293–1309. doi:10.1007/s00109-021-02088-w

Tarasov, A., Vilella, A., Cuppen, E., Nijman, I., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. *Bioinforma. Oxf. Engl.* 31 (12), 2032–2034. doi:10.1093/bioinformatics/btv098

Tayel, S., Muharram, N., Fotoh, D., Elbarbary, H., Abd-Elhafiz, H., El-Masry, E., et al. (2021). Prognostic impact of genetic variants of MECP2 and TIRAP on clinical

outcomes of systemic lupus erythematosus with and without nephritis. *Biomolecules* 11 (9), 1378. doi:10.3390/biom11091378

Tuijnenburg, P., Lango Allen, H., Burns, S., Greene, D., Jansen, M., Staples, E., et al. (2018). Loss-of-function nuclear factor κB subunit 1 (NFKB1) variants are the most common monogenic cause of common variable immunodeficiency in Europeans. *J. allergy Clin. Immunol.* 142 (4), 1285–1296. doi:10.1016/j.jaci.2018.01.039

van Schouwenburg, P., Davenport, E., Kienzler, A., Marwah, I., Wright, B., Lucas, M., et al. (2015). Application of whole genome and RNA sequencing to investigate the genomic landscape of common variable immunodeficiency disorders. *Clin. Immunol. Orl. Fla)* 160 (2), 301–314. doi:10.1016/j.clim.2015.05.020

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids Res.* 38 (16), e164. doi:10.1093/nar/gkq603

Webb, R., Merrill, J., Kelly, J., Sestak, A., Kaufman, K., Langefeld, C., et al. (2009). A polymorphism within IL21R confers risk for systemic lupus erythematosus. *Arthritis rheumatism* 60 (8), 2402–2407. doi:10.1002/art.24658

Xiao, X., Miao, Q., Chang, C., Gershwin, M., and Ma, X. (2014). Common variable immunodeficiency and autoimmunity–an inconvenient truth. *Autoimmun. Rev.* 13 (8), 858–864. doi:10.1016/j.autrev.2014.04.006

Xue, Y., Bao, Y., Zhang,, Z., Zhoa, W., Xiao, W., He, S., et al. (2014). Database resources of the national genomics data center, China national center for bioinformation in 2022. Nucleic Acids. *Res*. 50 (D1), D27–D38. doi:10.1093/nar/gkab951

Yang, H., Robinson, P., and Wang, K. (2015). Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. *Nat. methods* 12 (9), 841–843. doi:10.1038/nmeth.3484