Received: 21 February 2022 | Revised: 24 March 2022 | Accepted: 2 April 2022

DOI: 10.1002/jcla.24438

RESEARCH ARTICLE

WILEY

Identification of potential pathogenic genes for severe aplastic anemia by whole-exome sequencing

Yang Zhang[1] | Yu Zhang[1] | Hongyu Ge[2] | Nianbin Li[1] | Chunyan Liu[1] | Ting Wang[1] | Rong Fu[1] | Zonghong Shao[1]

[1]Department of Hematology, Tianjin Medical University General Hospital, Tianjin, China

[2]Department of Infectious Disease, Tianjin Medical University General Hospital, Tianjin, China

Correspondence
Zonghong Shao, Rong Fu and Ting Wang, Department of Hematology, Tianjin Medical University General Hospital, 154 Anshan St, Heping District, Tianjin, 300052 China.
Emails: shaozonghong@tmu.edu.cn (Z.S.); furong8369@tmu.edu.cn (R.F.); 99675616@qq.com (T.W.)

Funding information
This study was supported by the National Natural Science Foundation of China [grant numbers 81600093, 81770118, 81770110, 81870101, 81970116, 81970115]; the Science and Technology research project of Tianjin Health Commission [grant number 16KG124]; the Science and Technology personnel training project of Tianjin Health Commission [grant number RC20007]; and the Scientific Research Project of Tianjin Municipal Education Commission [grant number 2020KJ149]

Abstract

Background: Severe aplastic anemia (SAA) is a syndrome of severe bone marrow failure due to hyperfunction of CD8+ T cells. While, the genetic background of SAA is still unknown. In this study, we tried to explore the possible genetic variants in CD8+ T cells of SAA patients.

Methods: We performed whole-exome sequencing (WES) in CD8+ T cells of 4 SAA patients and 7 normal controls. The mutations that existed in SAA but not in NCs were identified as candidate genes. Then, we compared them with genes in the enriched KEGG pathway of differently expressed genes (DEGs) from previous RNA-seq. After analyzing the types of mutations, we identified possible pathogenic genes and validated them by RT-PCR. Finally, we compared them with the autoimmune disease-related genes in DisGeNET database to select the most possible pathogenic genes.

Results: We found 95 candidate mutant genes in which, 4 possible pathogenic genes were identified: PRSS1, KCNJ18, PRSS2, and DGKK. RT-PCR results showed that compared with NCs, PRSS1 and KCNJ18 mRNA expression was significantly increased in SAA patients ($p < 0.05$), PRSS2 was also increased in SAA patients but without statistical difference, and DGKK gene could not be detected by RT-PCR in SAA patients. In addition, PRSS1 was associated with autoimmune diseases from the DisGeNET database.

Conclusion: The mutations of PRSS1, KCNJ18, PRSS2, and DGKK, especially PRSS1 in CD8+T cells, may be involved in the immune pathogenesis of SAA.

KEYWORDS
pathogenic gene, RNA-Seq, severe aplastic anemia, whole-exome sequencing

1 | INTRODUCTION

Severe aplastic anemia (SAA) is a syndrome of severe bone marrow failure characterized by pancytopenia.[1] Multiple etiologies of SAA have been proposed. Previous study found that the dysregulation of T cells, especially hyperfunction of activated CD8+ T cells (CTLs) played an important role in pathogenesis of SAA.[1,2] Our research found that activated CTLs in SAA patients damage hematopoietic cells through perforin, granzyme B pathway, Fas/FasL pathway, and TNF-related apoptosis-inducing ligand (TRAIL) pathway.[3,4] Vitro research suggested that CTLs in SAA patients have a killing effect on normal bone marrow hematopoietic stem/progenitor cells, granulocytes, and monocytes.[5] The number of regulatory cells (Tregs), CTLA-related antigen 4 (CTLA-4) expressed by Tregs, and the

expression of Foxp3 protein were decreased in SAA patients.[6] The reason that caused unbalance between CTL and Tregs was still not fully understood.

Along with the development of sequencing technology, several human leukocyte antigen (HLA) alleles have been reported to be associated with SAA by whole-exome sequencing (WES) such as HLA-A*33:03, HLA-A*68:01, HLA-B*14:02, and HLA-B*40:02 alleles.[7] Yoshizato et al.[8] investigated 156 SAA patients and found that these patients had multiple somatic mutations, such as BCOR/BCORL1, PIGA, DNMT3A, and ASXL1. Whether CD8+ T cell abnormalities are also caused by genetic mutations is still unknown. Although the human exon region which holds the protein coding information, only makes up 1%–2% of the total genome, it concentrates on pathogenic mutation sites for most diseases.[9] Therefore, we aim to explore abnormal mutations of CD8+T lymphocytes in SAA patients by using WES technology, search for new pathogenic genes, and identify the possible molecular mechanism of SAA.

2 | MATERIALS AND METHODS

2.1 | Patients

WES was performed in 4 SAA patients and 7 normal controls (NCs). And RNA sequencing (RNA-seq) was included in 11 SAA patients and 10 NCs. The clinical features of these 15 AA patients were shown in Table S1, who were diagnosed in the Hematology Department of Tianjin Medical University General Hospital from 2017 to 2019. The subsequent RT-PCR experiment enrolled 12 SAA patients and 13 NCs who were hospitalized in our department from September 2020 to February 2021. The study was approved by the Ethics Committee of Tianjin Medical University. Informed written consent was obtained from all patients in accordance with the Declaration of Helsinki.

2.2 | Isolation and purification of CD8+T lymphocytes

Peripheral blood mononuclear cells (PBMNCs) from SAA patients and NCs were sorted by density gradient centrifugation using Ficoll-Paque Plus solution (solarbio, P8610). CD8+ T cells were purified by using CD8/MACS separators (Miltenyi Biotec) and the extraction method was carried out as recommended by the manufacturer. After detection by the multiparameter flow cytometry (BD Biosciences) and analysis using the Cell Quest software program (Version 3.1, Becton Dickinson), the cell was used for the subsequent experiments.

2.3 | WES and selection of candidate genes

The DNA was extracted according to the operational instructions of the genome extraction kit (TIANamp Micro DNA Kit, TIANGEN). The qualified genomic DNA sample was randomly fragmented by Covaris technology and each resulting qualified captured library was then loaded on BGISEQ-500 sequencing platforms, and we performed high-throughput sequencing.

The bioinformatics analysis began with the sequencing data. First, all clean data of each sample was mapped to the human reference genome GRCh37/HG19. Burrows-Wheeler Aligner (BWA) software was used to do the alignment. We followed recommended Best Practices for variant analysis with the Genome Analysis Toolkit (GATK, https://www.broadinstitute.org/gatk/guide/best-practices). The mean sequencing depth on target regions was 138.03×. In addition, the strict data analysis quality control system (QC) in the whole pipeline was built to guarantee qualified sequencing data. The SnpEff tool (http://snpeff.sourceforge.net/SnpEff_manual.html) was applied to perform a series of annotations for variants. The final variants and annotation results were used in the downstream advanced analysis.

Remove variants with MAF ≥ 1% according to allele frequency from the 1000 Genome Project control database. 3′/5′ prime UTR variants, intron variants, and intergenic variants have been removed without further analysis. Finally, we selected variants that were present in SAA patients but absent in controls.

2.4 | GO and KEGG pathway enrichment analysis of candidate genes

In order to identify the biological pathways, we performed enrichment analysis using Metascape (http://metascape.org),[10] which is a convenient and reliable website that provides comprehensive gene list annotation and enrichment analysis resources for experimental biologists.

For each given gene list, pathway and process enrichment analysis has been carried out with the following ontology sources: GO Biological Processes (BP), GO Cellular Components (CC), GO Molecular Functions (MF), and KEGG Pathway. All genes in the genome have been used as the enrichment background. Terms with a p-value <0.05, a minimum count of 3, and an enrichment factor >1.5 (the enrichment factor is the ratio between the observed counts and the counts expected by chance) were collected and grouped into clusters based on their membership similarities.

2.5 | Analysis of KEGG enrichment pathway of differently expressed genes and the genes in the enrichment pathway

Our research group previously used RNA-seq to study the differently expressed genes (DEGs) of peripheral blood CD8+ T cells in 11 SAA patients and 10 NCs. We sorted CD8+T lymphocytes and extracted total RNA of each sample using TRIzol Reagent (Invitrogen Life Technologies). Then, mRNA sequencing was performed by Illumina HiSeq. Finally, differential expression analysis was performed by DESeq2 package.[11] R (version 4.0.2) package ggplot2 was used to

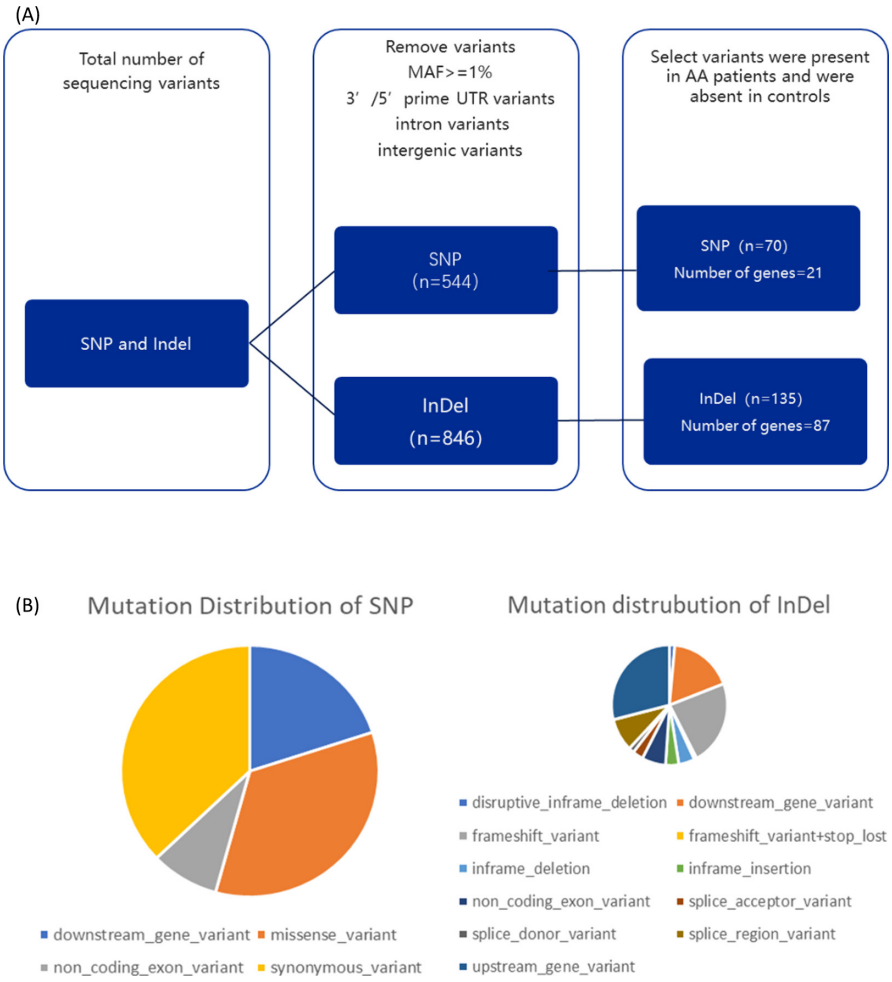

FIGURE 1 The discovery process of candidate mutant genes and distribution of mutations. (A) Schematic representation of the filter strategies used in our study. (B) Mutation distributions of SNP and InDel. Different colors indicate the frequency of different types of gene variants

paint a volcano plot. $p_{adj} < 0.05$ were set to select differential expressed genes. In this article, we used Metascape to analyze the KEGG enrichment pathway of the DEGs and the PathCards database to discover the genes in the enrichment pathway. Finally, observe whether the genes in the pathway overlap with the candidate genes of WES.

2.6 | RNA isolation, cDNA synthesis, and real-time quantitative reverse transcriptase PCR

Total RNA was extracted from the purified CD8+T cells using TRIzol reagent (Invitrogen Life Technologies). The extracted RNA was reverse transcribed using the reverse transcription kit (TaKaRa), according to the manufacturer's instructions. Gene expressions were analyzed by real-time quantitative PCR. β-Actin was used as the house-keeping gene for standardizing the expression of target mRNA. The primers used for this analysis are listed in Table S2. Reactions were performed using the Bio-Rad iQ5 Real-Time System and SYBR Premix Ex Taq (TaKaRa). The relative expression levels of target genes were calculated using the $2^{-\Delta\Delta C_t}$ method. All analyses were performed using the SPSS 20.0 software. The quantitative data were presented as mean ± SD (mean ± standard deviation). Differences in quantitative parameters between groups were determined using the t test. $p < 0.05$ was considered statistically significant.

2.7 | Analysis of genes associated with autoimmune diseases

DisGeNET is a discovery platform containing one of the largest publicly available collections of genes and variants associated with human diseases.[12-14] DisGeNET integrates data from expert curated repositories, GWAS catalogs, animal models, and scientific literature. Enter "Autoimmune Diseases," select "Summary of Gene-Disease Associations," and download genes related to autoimmune diseases.

3 | RESULTS

3.1 | WES and candidate genes analysis

According to the method mentioned above, we found 95 candidate mutant genes (SNP and Indel had 13 of the same genes) which were identified in cases but not in controls (Figure 1A). Mutation distributions of SNP and Indel were shown in Figure 1B.

3.2 | GO and KEGG pathway enrichment analysis of candidate genes

The GO-BP, GO-MF, and GO-CC analysis demonstrated that the candidate genes were concentrated in taste receptor activity, detection of chemical stimulus involved in sensory perception, collagen metabolic process, signaling adaptor activity, hemostasis, alcohol biosynthetic process, ribonucleoprotein complex biogenesis, meiotic nuclear division, myeloid leukocyte migration, positive regulation of innate immune response and so on (Figure 2A, Table S3). Additionally, the KEGG enrichment analysis indicated that the candidate genes were enriched in 3 pathways, which involved in taste transduction, olfactory transduction, and protein digestion and absorption pathways. (Figure 2B, Table S4). To further capture the relationships between the terms, we used Metascape to create a network plot. The network was visualized using Cytoscape5, where each node represents an enriched term and was colored first by its cluster ID (Figure 2C) and then by its p-value (Figure 2D). We also used Metascape for protein–protein interaction enrichment analysis. The resultant network contained the subset of proteins that form physical interactions with at least one other member in the list. Pathway and process enrichment analysis has been applied to each MCODE component independently, and the three best-scoring terms by *p*-value have been retained as the functional description of the corresponding components, shown in the tables underneath corresponding network plots within Figure 2E.

3.3 | Comparison of candidate genes with genes in the enriched KEGG pathway of DEGs

Our research group previously used RNA-seq to study the DEGs of peripheral blood CD8+ T cells and found 189 DEGs (36 up-regulated and 153 down-regulated genes) (Figure 3A). KEGG enrichment analysis indicated that the DEGs were enriched in 62 pathways ($p < 0.05$). To further show the relationship between terms, we used Metascape to create a network plot, which was colored by its cluster ID or p-value (Figure 3B,C). We also used Metascape for protein-protein interaction enrichment analysis (Figure 3D). Through the platform PathCards (https://pathcards.genecards.org/), we searched for the genes in the 62 KEGG pathways, and compared them with the candidate mutant genes ($n = 95$) of WES. PDE5A, AQP7, DGKK, AVPR2, KCNJ18, IQSEC1, PRSS1, PRSS2, FZD8, and KSR2 genes were found (Table S5). After analyzing the types of mutations, it was found that the mutations of PRSS1, KCNJ18, PRSS2, and DGKK could affect the amino acid structure of proteins (Table 1).

3.4 | PRSS1, KCNJ18, PRSS2, and DGKK mRNA expression levels in CD8+T cells of SAA patients

RT-PCR results showed that the mRNA expression level of PRSS1 was significantly increased in CD8+T cells of SAA patients compared

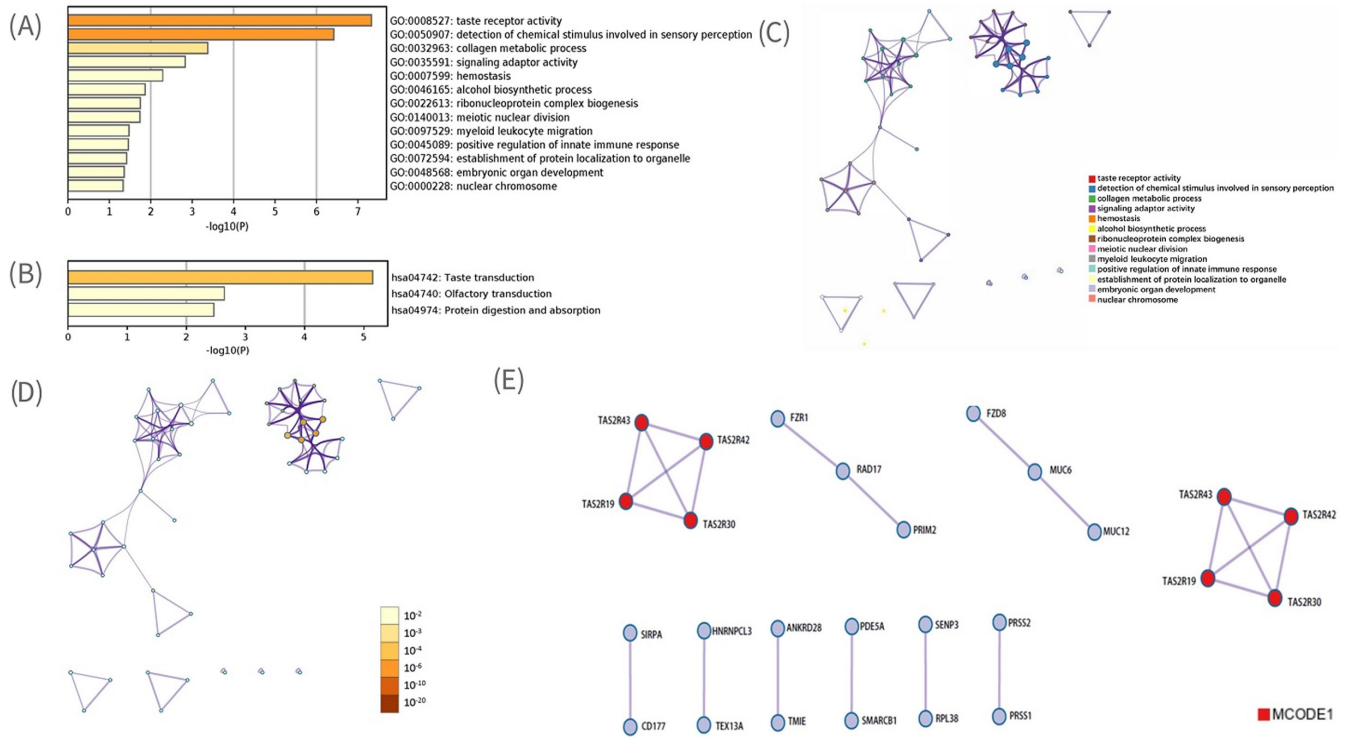

FIGURE 2 Enrichment and interaction analysis of candidate genes. (A) GO enrichment pathway analysis and (B) KEGG enrichment pathway analysis of candidate genes. Bar graph of enriched terms across input gene lists, colored by *p*-values. (C) Network plot of enriched terms for candidate genes. Colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (D) Network plot of enriched terms for candidate genes. Colored by *p*-value, where terms containing more genes tend to have a more significant *p*-value. (E) Protein-protein interaction network and MCODE components identified in the gene lists

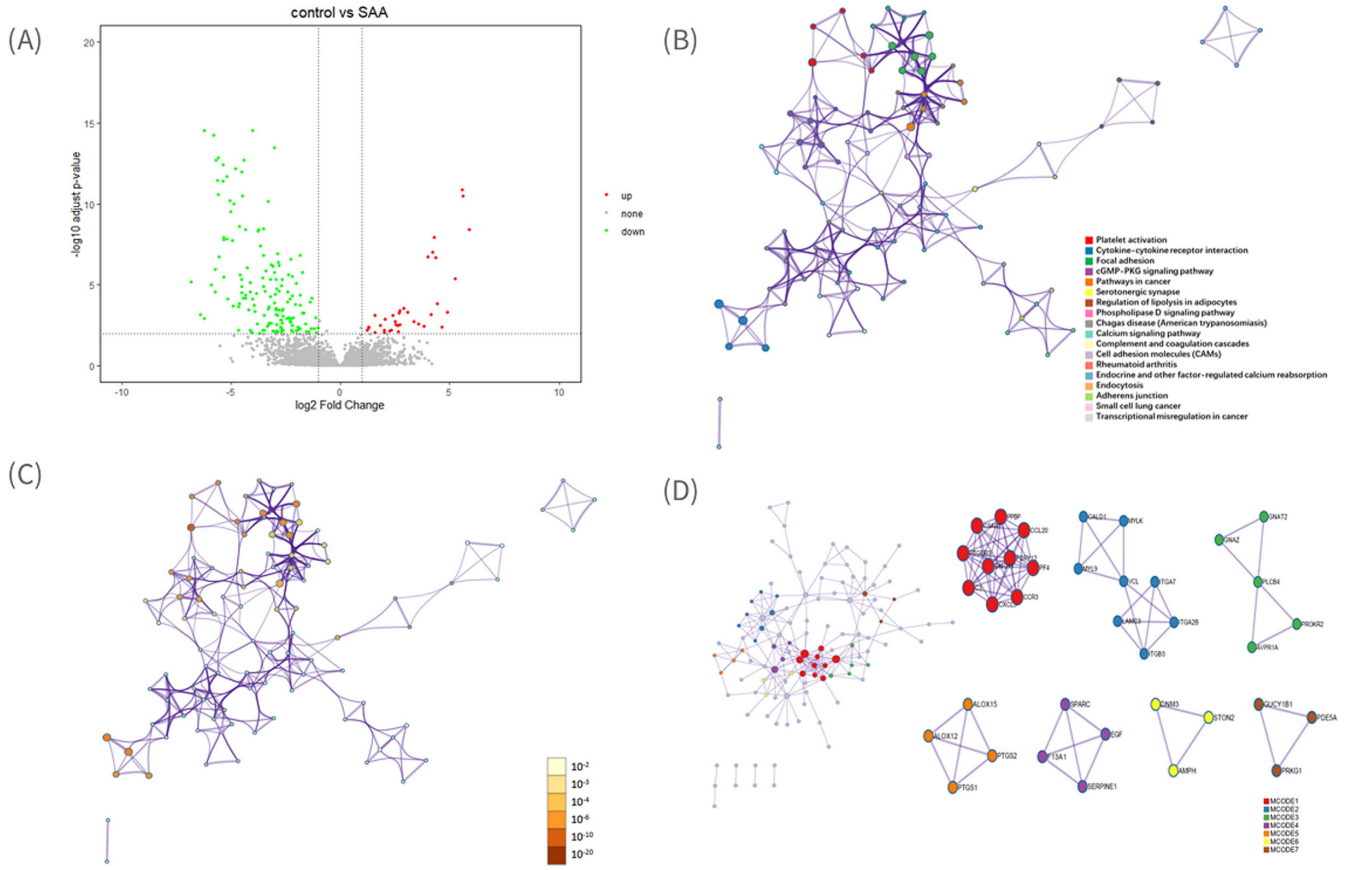

FIGURE 3 Enrichment and interaction analysis of DEGs between SAA patients and NC. (A) Volcano plot of gene expression profile data in SAA patients and NCs. (B) Network plot of enriched terms for DEGs. Colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (C) Network plot of enriched terms for DEGs. Colored by p-value, where terms containing more genes tend to have a more significant *p*-value. (D) Protein-protein interaction network and MCODE components identified in the gene lists

to NCs (5.518 ± 4.201 vs. 2.414 ± 1.808; $p < 0.05$; Figure 4A). The same for KCSNJ18, the mRNA levels of SAA patients and NCs were 2.873 ± 2.158 and 1.445 ± 1.110, respectively ($p < 0.05$; Figure 4B). For PRSS2, the mRNA expression level of PRSS2 in CD8+T cells of SAA patients was higher than that in NCs, but without statistical difference (2.113 ± 2.141 vs. 1.207 ± 0.8312; $p > 0.05$; Figure 4C). Moreover, we found that DGKK gene expression could not be detected in CD8+T cells of SAA patients, but could be detected in NC group by RT-PCR.

3.5 | Comparison of significant mutations with genes associated with autoimmune diseases

We downloaded 1758 autoimmune disease-related genes from the platform of DisGeNET and observed whether the genes overlap with the 4 possible pathogenic mutations (PRSS1, KCNJ18, PRSS2, and DGKK). PRSS1 gene was found as the overlapping gene. According to the results of WES, the mutation site of PRSS1 gene in SAA patients was c.508A>G, and amino acid changes corresponded to p. Lys170Glu (Figure 5).

4 | DISCUSSION

SAA was regarded as an autoimmune disease, and the initial research on its pathogenesis mainly focused on immunity. Young[15] found that the number of activated effector T lymphocytes (mainly CD8+T cells) in SAA patients increased and their function was hyperactive. Later, it was discovered that part of SAA patients had poor immunotherapy in clinical work. And with the development of next-generation sequencing technology, some researchers gradually discovered that certain gene mutations affected the condition and prognosis of SAA patients. WES technology utilizes its high-throughput characteristics to efficiently discover mutations in protein coding regions of all genes. Moreover, there were few studies on the WES of SAA patients, and even fewer studies focused on the major immune abnormal cells of SAA patients, such as CD8+T cells. Therefore, we targeted CD8+T cells of SAA patients by WES technology to look for possible pathogenic mutations, and laid a foundation for further large-scale experimental verification in SAA patients.

We performed GO and KEGG enrichment analysis on 95 candidate mutant genes which existed in CD8+ T cells of SAA patients but not in NCs. It was shown that they were enriched in the process of

TABLE 1 Analysis of gene mutation types, in which red marks represent significant mutations

Gene	Function	Transcript	Chr
PDE5A	Upstream gene variant	NM_001083.3	chr4
AQP7	Synonymous variant	NM_001170.1:p.Pro113Pro/c.339G>A NM_001170.1:p.Asn94Asn/c.282C>T	chr9
DGKK	Frameshift variant	NM_001013742.3:p. Gly1020_Leu1021fs/c.3060_3061insC	chrX
AVPR2	Upstream gene variant	NM_000054.4	chrX
KCNJ18	Synonymous variant	NM_001194958.2:p.Gly99Gly/c.297C>T	chr17
		NM_001194958.2:p.Ile105Ile/c.315C>T	
		NM_001194958.2:p.Asn199Asn/c.597C>T	
		NM_001194958.2:p.Ser266Ser/c.798G>A	
		NM_001194958.2:p.Asp291Asp/c.873C>T	
		NM_001194958.2:p.Asp397Asp/c.1191C>T	
		NM_001194958.2:p.Gly418Gly/c.1254C>T	
	Missense variant	NM_001194958.2:p.Gly145Ser/c.433G>A	
		NM_001194958.2:p.Thr290Met/c.869C>T	
		NM_001194958.2:p.Arg285Trp/c.853C>T	
	Inframe deletion	NM_001194958.2:p. Glu333del/c.997_999delGAG	
IQSEC1	Splice acceptor variant	NM_001134382.2:c.2977delG	chr3
PRSS1	Missense variant	NM_002769.4:p.Lys170Glu/c.508A>G	chr7
	Synonymous variant	NM_002769.4:p.Leu189Leu/c.567T>C	
PRSS2	Missense variant	NM_002770.2:p.Val3Ala/c.8T>C NM_002770.2:p.Pro4Ser/c.10C>T	chr7
	Upstream gene variant	NM_002770.2	
FZD8	Upstream gene variant	NM_031866.2	chr10
KSR2	Upstream gene variant	NM_173598.4	chr12

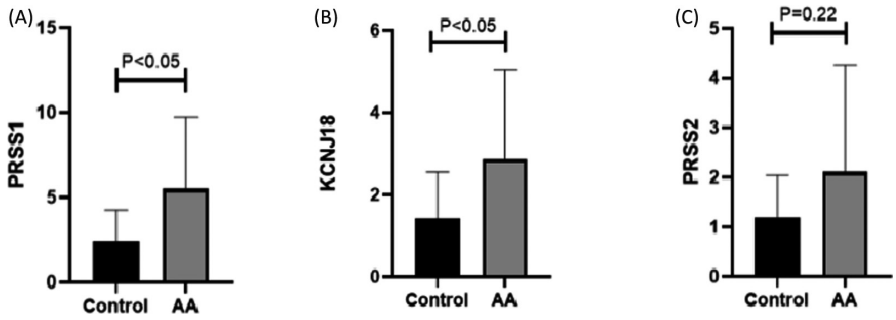

FIGURE 4 The mRNA expression levels of 4 candidate genes. (A) The relative mRNA levels of PRSS1was significantly increased in CD8+T cells of SAA patients compared to NCs ($p < 0.05$). (B) The relative mRNA levels of KCNJ18 was significantly increased in CD8+T cells of SAA patients compared to NCs ($p < 0.05$). (C) The relative mRNA levels of PRSS2 was increased in CD8+T cells of SAA patients compared to NCs but without statistic difference

taste and olfactory, positive regulation of innate immune response, myeloid leukocyte migration, hemostasis, protein digestion, and absorption pathways and so on. Previous studies have shown that patients with rheumatoid arthritis (RA) have similar pathway abnormalities. Ying Li et al.[16] used WES technology to study the pathogenic genes of RA patients and analyzed the enrichment pathways of the pathogenic genes. These pathogenic genes were also found to be enriched in the pathways such as olfactory transduction, protein digestion, and absorption, which were similar to the pathways found in our study, may be related to the incidence of RA. Seven genes involved in the olfactory transduction pathway were associated with the progression of RA. However, whether these pathways play a role in the pathogenesis of SAA patients and how they affect the pathogenesis of SAA needs further experimental verification.

Our research group previously discovered 189 DEGs in SAA patients through RNA-seq. In this paper, we performed KEGG enrichment pathway analysis on them. Furthermore, the genes in each pathway were further searched and compared with 95 candidate

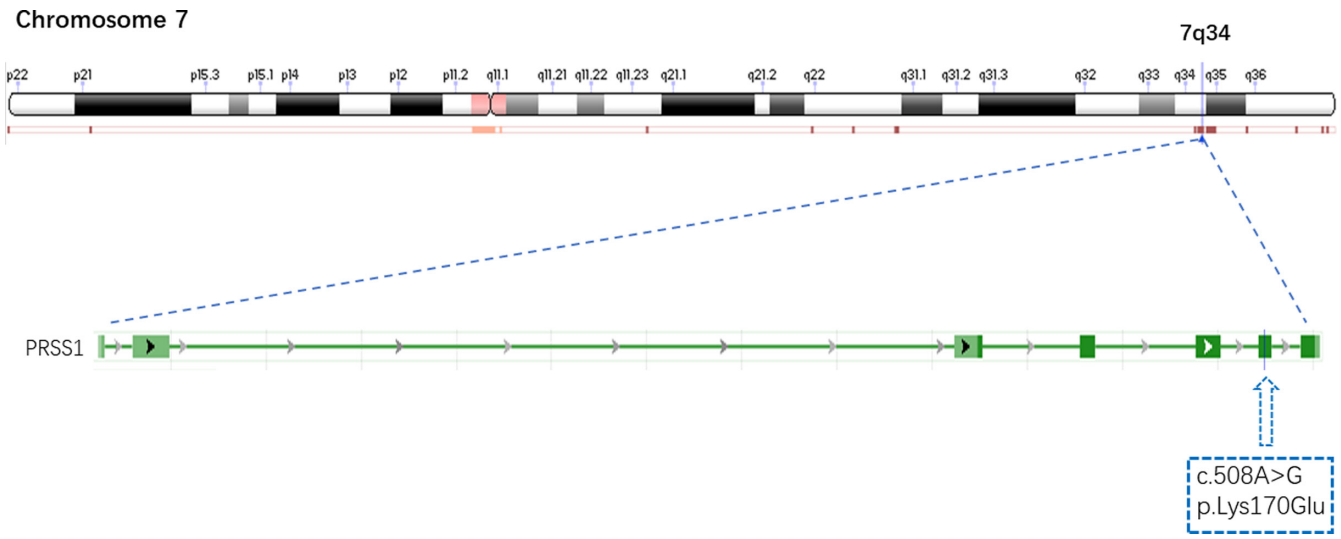

FIGURE 5 Schematic representation of the structure of PRSS1. cDNA mutation and amino acid change have been listed in the text box of blue

mutant genes, 10 shared genes including PDE5A, AQP7, DGKK, AVPR2, KCNJ18, IQSEC1, PRSS1, PRSS2, FZD8, and KSR2 were found. After analysis of the mutational types, 4 possible pathogenic mutations were found, including PRSS1, KCNJ18, PRSS2, and DGKK, because they could affect the amino acid structure of proteins. Further GO-BP analysis of these genes was performed. Although PRSS1 and PRSS2 were both part of the trypsinogen gene cluster located in the β locus of the T cell receptor, they encode different trypsinogens and participate in different biological processes. For example, PRSS1 is involved in the process of cobalamin metabolism (vitamin B12) and proteolysis, while PRSS2 participates in the antimicrobial humoral response and neutrophil activation involved in immune response. KCNJ18 encoded a member of the inwardly rectifying potassium channel family. DGKK is involved in the process of platelet activation and intracellular signal transduction.

Due to the limited sample size of this experiment, we would like to further prove the correlation between genes and AA through public databases. However, there is no specialized database for AA. Considering that AA is one of the typical autoimmune diseases, and autoimmune diseases often have a similar state of immune hyperfunction. Therefore, we downloaded the currently known genes that are clearly related to autoimmune diseases in the DisGeNET database and did an intersection analysis with the 4 pathogenic genes that had been screened. We found the common gene PRSS1.

As mentioned above, PRSS1 is a gene that encodes trypsinogen. This gene and several other trypsinogen genes are localized to the T cell receptor beta locus on chromosome 7. We firstly searched for the mutation rate of PRSS1 gene in normal people and found that PRSS1 mutation is very rare in the normal population. Sequencing of the PRSS1 gene in 200 French, 82 German, 420 Chinese, 28 Korean, and 150 Brazilians found the p.E79K variant in only one French and one Brazilian individual. Recently, 1000 German normal controls were sequenced, and the p.V123L mutation was only found in one subject.[17] Another study found no PRSS1 mutation in 220 Chinese Han individuals.[18] Then, we further searched the literature for the effect of PRSS1 gene mutation on autoimmune diseases. Autoimmune pancreatitis (AIP) is considered to be an immune-mediated inflammatory process targeting the pancreatic epithelial components. The serum of patients with AIP contained high titers of autoantibodies against trypsinogen PRSS1.[19] Multiple mutation types of PRSS1 gene were related to the occurrence of AIP, such as p. Leu81Met, PRSS1:ivs2+56_60delCCCAG, p.K92N, p.R116C, p.T137 M, and p.C139F.[20-22] Later, the effect of the mutation was further verified in a mouse model. The study[23] found that the levels of HSP70 and NRF2 in mice expressing PRSS1R122H were increased, and the level of chymotrypsin C (CTRC) was decreased. Pancreata from mice that express PRSS1R122H had more DNA damage, apoptosis, and collagen deposition and increased trypsin activity and infiltration by inflammatory cells than mice that express PRSS1 or control mice. We speculated that the SAA patients may have the mutation of this pathogenic gene, which affected the up-regulation or down-regulation of mRNA expression, thus further affecting the translation of proteins and eventually leading to the occurrence of diseases. Therefore, we performed an RT-PCR experiment to verify our hypothesis. Results showed that the expression levels of PRSS1-mRNA in CD8+T cells of SAA patients were significantly increased than normal controls ($p < 0.05$). We hypothesized that the mutations of PRSS1 may play an important role in the pathogenesis of SAA.

The mRNA expression of the other 3 genes was also examined in SAA patients. Results showed that the mRNA expression levels of KCNJ18 and PRSS2 in CD8+T cells of SAA patients were increased. However, DGKK could not be detected in CD8+T cells of SAA patients by RT-PCR, but could be detected in NC group. So, we suspected that the mutation resulted in the silencing of DGKK expression in CD8+T cells of SAA patients.

5 | CONCLUSIONS

In summary, our study is the first time for the analysis of CD8+ cells in SAA patients by WES technology. The mutations revealed in these genes (PRSS1, KCNJ18, PRSS2, and DGKK), especially PRSS1 may be involved in the pathogenesis of SAA.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

ETHICS STATEMENT

The study was approved by the Ethics Committee of Tianjin Medical University. Informed written consent was obtained from all patients in accordance with the Declaration of Helsinki.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Zonghong Shao https://orcid.org/0000-0003-4966-2956

REFERENCES

1. Shallis RM, Ahmad R, Zeidan AM. Aplastic anemia: etiology, molecular pathogenesis, and emerging concepts. *Eur J Haematol*. 2018;101:711-720.
2. Sun YY, Liu CY, Shao ZH. Advances of pathogenesis and genetic abnormalities in acquired aplastic anemia. *Zhonghua Nei Ke Za Zhi*. 2019;58:401-404.
3. Feng L, Fu R, Wang HQ, et al. Effect of CD8+ effector T cells on the hematopoiesis pathway damage in the patients with severe aplastic anemia. *Zhonghua Xue Ye Xue Za Zhi*. 2011;32:597-601.
4. Xing LM, Liu CY, Fu R, et al. CD8+HLA-DR+ T cells are increased in patients with severe aplastic anemia. *Mol Med Rep*. 2014;10:1252-1258.
5. Liu CY, Zheng MY, Fu R, et al. Target cells of cytotoxic T cells in severe aplastic anemia in vitro. *Zhonghua Yi Xue Za Zhi*. 2016;96:1728-1732.
6. Yan L, Fu R, Liu H, et al. Abnormal quantity and function of regulatory T cells in peripheral blood of patients with severe aplastic anemia. *Cell Immunol*. 2015;296:95-105.
7. Babushok DV, Duke JL, Xie HM, et al. Somatic HLA mutations expose the role of class I-mediated autoimmunity in aplastic anemia and its clonal complications. *Blood Adv*. 2017;1:1900-1910.
8. Yoshizato T, Dumitriu B, Hosokawa K, et al. Somatic mutations and clonal hematopoiesis in aplastic anemia. *N Engl J Med*. 2015;373:35-47.
9. Hansen MC, Haferlach T, Nyvold CG. A decade with whole exome sequencing in haematology. *Br J Haematol*. 2020;188:367-382.
10. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523-1623.
11. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
12. Piñero J, Queralt-Rosinach N, Bravo À, et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database*. 2015;2015:bav028.
13. Piñero J, Bravo À, Queralt-Rosinach N, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res*. 2017;45:D833-D839.
14. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res*. 2020;48:D845-D855.
15. Young NS, Bacigalupo A, Marsh JCW. Aplastic anemia: pathophysiology and treatment. *Biol Blood Marrow Transplant*. 2010;16:S119-S125.
16. Li Y, Lai-Han Leung E, Pan HD, et al. Identification of potential genetic causal variants for rheumatoid arthritis by whole-exome sequencing. *Oncotarget*. 2017;8:111119-111129.
17. Németh BC, Sahin-Tóth M. Human cationic trypsinogen (PRSS1) variants and chronic pancreatitis. *Am J Physiol Gastrointest Liver Physiol*. 2014;306(6):G466-G473.
18. Zeng K, Liu QC, Lin JH, et al. Novel mutations of PRSS1 gene in patients with pancreatic cancer among Han population. *Chin Med J (Engl)*. 2011;124(13):2065-2067.
19. Löhr JM, Faissner R, Koczan D, et al. Autoantibodies against the exocrine pancreas in autoimmune pancreatitis: gene and protein expression profiling and immunoassays identify pancreatic enzymes as a major target of the inflammatory process. *Am J Gastroenterol*. 2010;105:2060-2071.
20. Gao F, Li YM, Hong GL, et al. PRSS1_p.Leu81Met mutation results in autoimmune pancreatitis. *World J Gastroenterol*. 2013;19:3332-3338.
21. Chang MC, Jan IS, Liang PC, et al. Human cationic trypsinogen but not serine peptidase inhibitor, Kazal type 1 variants increase the risk of type 1 autoimmune pancreatitis. *J Gastroenterol Hepatol*. 2014;29:2038-2042.
22. Gao F, Li Y, Wang C, et al. Identification of a novel frame-shift mutation in PRSS1 gene in Han patients with autoimmune pancreatitis. *Curr Mol Med*. 2014;14:340-348.
23. Huang H, Swidnicka-Siergiejko AK, Daniluk J, et al. Transgenic expression of PRSS1 (R122H) sensitizes mice to pancreatitis. *Gastroenterology*. 2020;158:1072-82.e7.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Zhang Y, Zhang Y, Ge H, et al. Identification of potential pathogenic genes for severe aplastic anemia by whole-exome sequencing. *J Clin Lab Anal*. 2022;36:e24438. doi:10.1002/jcla.24438