Target genes regulated by *CLEC16A* intronic region associated with common variable immunodeficiency

Xubo Huang, MD, MS,[a]* Jinxia Huang, MD, MS,[a]*‡ Xiumei Li, MS,[b]* Jingxian Fan, MS,[b,g] Desheng Zhou, MD, MS,[a] Hui-Qi Qu, MD, PhD,[c] Joseph T. Glessner, PhD,[c,k,l] Dandan Ji, MS,[b] Qi Jia, PhD,[d] Zhiyong Ding, PhD,[e] Nan Wang, PhD,[e] Wei Wei, MD,[f] Xing Lyu, MD,[f] Mulin Jun Li, PhD,[g] Zhe Liu, PhD,[h] Wei Liu, MD,[i,j] Yongjie Wei, PhD,[a] Hakon Hakonarson, MD, PhD,[c,k,l] Qianghua Xia, PhD,[b,g] and Jin Li, PhD[b,f] *Guangzhou, Tianjin, Dalian, Jinan, and China; and Philadelphia, Pa*

Background: *CLEC16A* intron 19 has been identified as a candidate locus for common variable immunodeficiency (CVID).

Objectives: This study sought to elucidate the molecular mechanism by which variants at the *CLEC16A* intronic locus may contribute to the pathogenesis of CVID.

Methods: The investigators performed fine-mapping of the *CLEC16A* locus in a CVID cohort, then deleted the candidate functional SNP in T-cell lines by the CRISPR-Cas9 technique and conducted RNA-sequencing to identify target gene(s). The interactions between the *CLEC16A* locus and its target genes were identified using circular chromosome conformation capture. The transcription factor complexes mediating the chromatin interactions were determined by proteomic approach. The molecular pathways regulated by the *CLEC16A* locus were examined by RNA-sequencing and reverse phase protein array.

Results: This study showed that the *CLEC16A* locus is an enhancer regulating expression of multiple target genes including a distant gene *ATF7IP2* through chromatin interactions. Distinct transcription factor complexes mediate the chromatin interactions in an allele-specific manner. Disruption of the *CLEC16A* locus affects the AKT signaling pathway, as well as the molecular response of $CD4^+$ T cells to immune stimulation.

Conclusions: Through multiomics and targeted experimental approaches, this study elucidated the underlying target genes and signaling pathways involved in the genetic association of *CLEC16A* with CVID, and highlighted plausible molecular targets for developing novel therapeutics. (J Allergy Clin Immunol 2024;■■■:■■■-■■■.)

Key words: *AKT, autoimmunity, chromatin interaction,* CLEC16A, *common variable immunodeficiency disorder, primary immunodeficiency*

Coexistence of symptoms and phenotypic overlap are frequently observed between autoimmune diseases and primary immunodeficiency disorders (PIDDs).[1,2] Among patients with PIDD, there is a high prevalence of autoimmune diseases.[2] For example, hypothyroidism, alopecia areata, vitiligo, type 1 diabetes, rheumatoid arthritis (RA), and systemic lupus erythematosus (SLE) frequently occur to patients with common variable immunodeficiency (CVID), which makes CVID a human model disease for comorbid autoimmunity and PIDDs.[3,4] The molecular mechanism underlying these "paradoxical" immune actions is of particular interest and importance for the development of clinical therapeutics.

The intronic region of *CLEC16A* has been identified of having associations with both PIDDs[5] and autoimmune disorders, such as type 1 diabetes (for which *CLEC16A* was first reported by its alias *KIAA0350* by our group and others),[6-8] multiple sclerosis,[9] SLE,[10] and RA.[11] *CLEC16A* has been suggested as a possible link between autoimmunity and immunodeficiency according to our previous study.[5] However, molecular mechanisms of this pleotropic locus in the immune system have been poorly understood. For the intron 19 of *CLEC16A*, the promoter of *DEXI* has been shown to have long-range chromosomal interaction with this region in EBV-transformed B cells and monocytes.[12] However, its pleiotropy in autoimmune disorders implies that the intron 19 of *CLEC16A* might function as an enhancer regulating different genes in a variety of cell types.

To dissect the role of *CLEC16A* in the pathogenesis of CVID, we fine-mapped this locus in a CVID cohort, which has been densely genotyped by the Immunochip (Illumina, San Diego, Calif), and identified rs11642009 as the candidate functional SNP at this locus. Using gene-editing, genomic, epigenomic,

From [a]the Affiliated Cancer Hospital and Institute of Guangzhou Medical University; [b]the Department of Cell Biology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Tianjin Key Laboratory of Medical Epigenetics, Tianjin Institute of Immunology, School of Basic Medical Sciences, Tianjin Medical University; the Departments of [g]Bioinformatics, and [h]Immunology, School of Basic Medical Sciences, Tianjin Medical University; [f]the Department of Rheumatology and Immunology, Tianjin Medical University General Hospital; [i]the Tianjin Children's Hospital (Tianjin University Children's Hospital); [j]the Tianjin Key Laboratory of Birth Defects for Prevention and Treatment; [c]the Center for Applied Genomics and [k]the Division of Human Genetics, The Children's Hospital of Philadelphia; [l]the Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia; [d]the International School of Information Science Engineering, Dalian University of Technology; and [e]the Mills Institute for Personalized Cancer Care, Fynn Biotechnologies Ltd, Jinan.

*These authors contributed equally to the study.

‡Present address: Foresea Life Insurance Guangzhou General Hospital, Guangzhou, China.

Received for publication August 17, 2022; revised December 25, 2023; accepted for publication December 29, 2023.

Corresponding author: Jin Li, PhD, Department of Cell Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China. E-mail: jli01@tmu.edu.cn. Or: Qianghua Xia, PhD, Department of Bioinformatics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China. E-mail: qhxia@tmu.edu.cn.

0091-6749/$36.00

https://doi.org/10.1016/j.jaci.2023.12.023

Abbreviations used
CHOP: Children's Hospital of Philadelphia
CVID: Common variable immune deficiency
DEG: Differentially expressed gene
FC: Fold change
GWAS: Genome-wide association study
KO: Knockout
MS: Mass spectrometry
PIDD: Primary immunodeficiency disorder
PMA: Phorbol-12-myristate-13-acetate
RA: Rheumatoid arthritis
RNA-seq: RNA-sequencing
SILAC: Stable isotope labeling with amino acids in cell culture

and proteomic approaches, we demonstrated the regulation of the expression of multiple target genes through cell type–specific chromatin interactions. Distinct transcription factor complexes mediate the chromatin interactions in an allele-specific manner. Perturbation of *CLEC16A* locus affects immune and metabolic pathways and molecular response of $CD4^+$ T cells to phorbol-12-myristate-13-acetate (PMA) stimulation. These findings offer potential opportunities for developing novel therapeutics.

METHODS

Detailed methods are presented in the Methods section in this article's Online Repository (available at www.jacionline.org). Briefly, the imputed Immunochip data of patients with CVID (n = 299) and controls (n = 900) were selected from the database of Center for Applied Genomics, the Children's Hospital of Philadelphia (CHOP) and were used to compute the linkage disequilibrium between SNPs at the *CLEC16A* intron 19 region for fine-mapping. Ethical approval was obtained from CHOP Institutional Review Board and carried out in accordance with the nationally approved guidelines. CRISPR/Cas9 genome editing was conducted on T-cell lines (Table E1 in this article's Online Repository at www.jacionline.org) followed by RNA-sequencing, quantitative real-time PCR (Table E2 in this article's Online Repository at www.jacionline.org) and reverse phase protein array analysis. Circular chromosome conformation capture and data analysis were performed. Stable isotope labeling with amino acids in cell culture (SILAC) experiment and oligonucleotide pull-down were conducted followed by mass spectrometry (MS). Small interfering RNA and Western blotting were also performed.

RESULTS

SNP rs11642009 is associated with several autoimmune disorders and is a candidate functional variant for CVID

To dissect the contribution of the pleotropic locus *CLEC16A* intron 19 to CVID, we conducted an extensive bioinformatic and experimental study by integrating genomic, epigenomic, transcriptomic, and proteomic data (Fig 1).

Our previous genome-wide association study (GWAS) on 778 CVID cases and 10,999 controls reported that *CLEC16A* intron 19 locus is strongly associated with CVID based on the Immunochip data.[5] We performed fine-mapping and found an imputed SNP

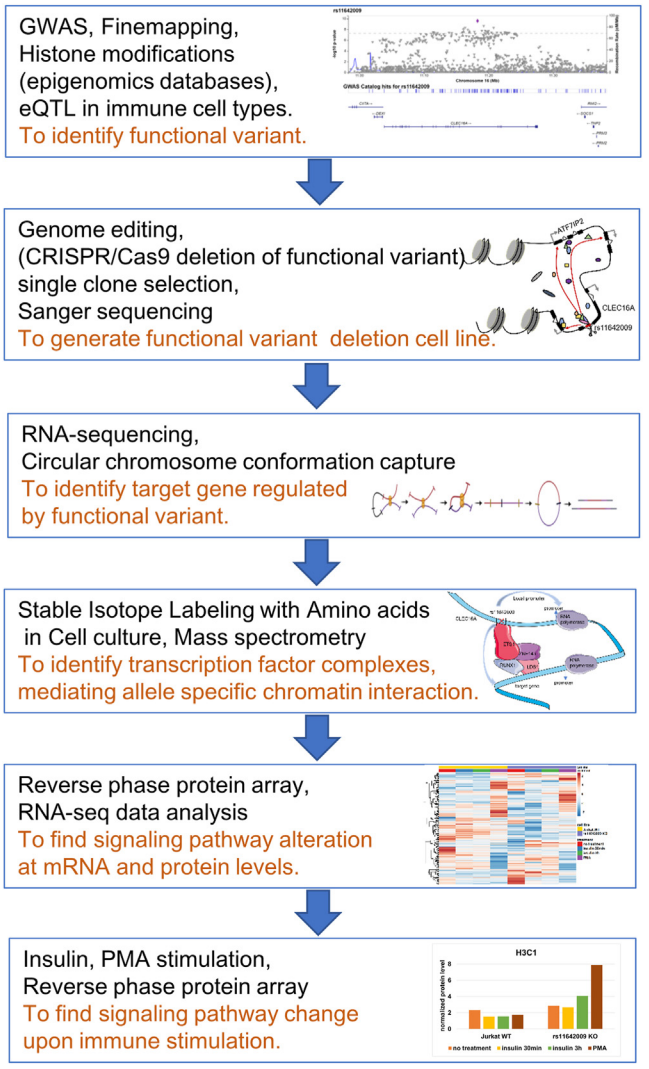

FIG 1. The design of the study. The experiments and the aim of each step are shown. *eQTL*, expression quantitative trait loci.

rs11642009 (associated with CVID at $P = 2.46E{-}10$)[5] is predicted to be in a 95% credible set with posterior probability higher than the other SNPs based on multiple fine-mapping tools (Fig 2, *A*). The SNP overlaps with enhancer-specific modifications in multiple immune cell types, including T cells (Fig 2, *B*, and Fig E1 in this article's Online Repository at www.jacionline.org). It is correlated with the expression level of *CLEC16A*, *ATF7IP2*, *DEXI*, *RMI2*, and *SNX29* at nominally significant level in $CD4^+$T cells and other genes in other cell types in expression quantitative trait loci studies (Table E3 in this article's Online Repository at www.jacionline.org). The SNP is associated with several immunological traits and immune diseases, including lymphocyte counts ($\beta = -0.0138$, $P = 2.81E{-}10$),[13] SLE ($\beta = 0.19$, $P = 1.65E{-}12$),[14] and atopy ($\beta = -0.0612$, $P = 5.69E{-}20$)[15] (Fig 2, *C*, and Table E4 in this article's Online Repository at www.jacionline.org). The *CLEC16A* intron 19 region was also highlighted of significant association with PIDD indexed by SNP rs2286974 in another study.[16] However, rs2286974 lacks major enhancer/promoter histone modifications in T cells or B cells, in contrast to the various epigenomic histone marks

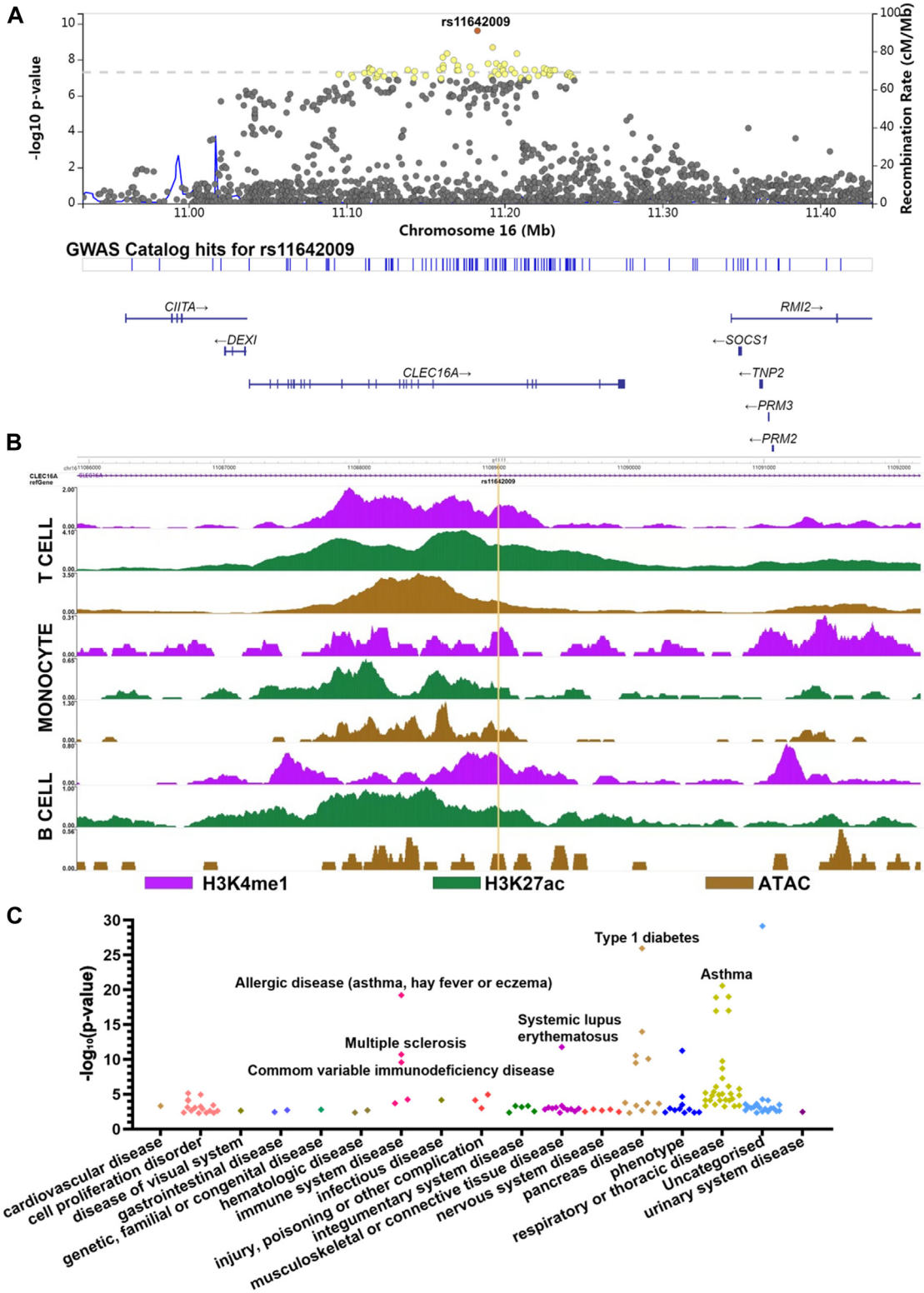

FIG 2. Genomic and epigenomic analyses of the *CLEC16A* locus. **A,** The regional association plot of rs11642009 associated with CVID. The SNP rs11642009 is at the intron 19 of *CLEC16A,* and it is most strongly associated with CVID at this locus. The color of the dots represents the 95% credible set membership (*brown*: most contribution; *yellow*: some contribution; *gray*: no contribution). The plot was generated via LocusZoom web portal. **B,** ChIP-seq (Chromatin immunoprecipitation followed by sequencing) signal of H3K4me1, H3K27Ac, and ATAC-Seq (assay for transposase-accessible chromatin with sequencing) signal from Roadmap database in the vicinity of SNP rs11642009 in T cells, B cells, and monocytes. **C,** Immunological traits and immune disorders associated with the SNP rs11642009.

overlapped with SNP rs11642009 (Fig E1). Therefore, the genetic and epigenomic data suggest that rs11642009 is a functional variant for CVID at this locus.

The genomic region harboring rs11642009 regulates the expression of multiple genes

With rs11642009 as a potential enhancer in T-cell lineage and the important roles of $CD4^+$ T cells in CVID and autoimmunity,[2,17,18] we used Jurkat cells (a $CD4^+$ T-cell line) for molecular experiments to delineate the regulatory roles of rs11642009. We generated a 443-bp deletion of the genomic region harboring rs11642009 in the Jurkat cells (Fig 3, *A* and *B*), followed by RNA-sequencing (RNA-seq) analysis. On deletion of rs11642009, 1291 genes showed significantly differential expression (with a *Q* value cutoff of ≤0.05 and a fold-change [FC] ≥2), among which 787 genes were downregulated and the remaining 504 were upregulated (Fig 3, *C*, and Table E5 in this article's Online Repository at www.jacionline.org). GWAS functional variants and their target genes are likely to reside in the same topologically associating domain within which most enhancer-promoter cis-regulatory contacts happen. Among the 15 genes located in the flanking region of rs11642009 (within 1 Mb distance), the genes *CLEC16A*, *SOCS1,* and *CIITA*, which have been implicated in the pathogenesis of autoimmune diseases,[19] showed decreased expression on deletion of the rs11642009 region, though not reaching $Q < 0.05$. Interestingly, a distal gene *ATF7IP2* was significantly reduced (FC = 2.0; Q = 4.61E−20). Additional 6 genes also showed expression change of FC > 1.2, $Q < 0.05$, including *GSPT1, ZC3H7A, LITAF, SNX29, RMI2,* and *TVP23A* (Table E6 in this article's Online Repository at www.jacionline.org). We tested 3 genes with the most significant statistics ($Q < 1.0E-10$) by quantitative PCR and confirmed their expression change (Fig 3, *D*). To further confirm the transcriptional regulatory effect of rs11642009, we similarly deleted a 482-bp genomic region containing rs11642009 in another human T-cell line CCRF-CEM, and subsequently conducted an RNA-seq experiment. The expression level of 12 genes among 15 genes located in the flanking region of rs11642009 showed the same direction of change as in Jurkat cells, especially *ATF7IP2,* which demonstrated consistent profound reduction on the deletion of rs11642009 region (Fig E2 and Table E7 in this article's Online Repository at www.jacionline.org). These genes in the proximity to rs11642009 are likely to be directly impacted by the deletion of rs11642009, rather than as a consequence of indirect downstream signal transduction.

SNP rs11642009 interacts with the regulatory region of multiple genes

Because SNP rs11642009 is located in the intron of *CLEC16A* but distant to other genes, especially *ATF7IP2*, we speculated that rs11642009 may function as an enhancer affecting gene expression level via chromatin loops. We then carried out a circular chromosome conformation capture sequencing (4C-seq) experiment in Jurkat cells to find the physical contacts between the genomic region harboring rs11642009 and the regulatory region of target genes. We found 44 peaks mapped to 19 nearest genes (Fig 4, *A*). As expected, we detected strong interactions between rs11642009 and the promoter regions of *CLEC16A* and *SOCS1*, which are also enriched for H3K4me3 and H3K27ac based on the ENCODE (Encyclopedia of DNA Elements) database (Fig 4, *A* and *B*). We observed a distal interaction between the rs11642009 genomic region and the upstream region of *ATF7IP2*. H3K27me3 peaks were detected in this region (Fig 4, *A* and *B*), which represents a dynamic and important epigenetic modification during $CD4^+$ T-cell activation.[20] This chromatin interaction persisted under insulin or PMA stimulation (Fig E3 in this article's Online Repository at www.jacionline.org). We also observed interaction between the rs11642009 genomic region and the upstream region of *ATF7IP2* in other $CD4^+$ T-cell lines (CCRF-CEM and Molt4 cells) (Fig E4 in this article's Online Repository at www.jacionline.org).

To examine the cell type specificity of the chromatin interactions, we carried out 4C-seq experiments in cell lines K562 and LCL, which are human erythroleukemic cells and immortalized B-lymphoblastoid cells, respectively. There are 29 4C-seq peaks mapped to 11 genes in K562 cells and 32 peaks mapped to 14 genes in LCL cells (Fig 4, *A*). We found that 4C-seq peaks shared by Jurkat cells, and these 2 cell lines are located at the promoter region of 6 genes (Fig 4, *A, C,* and *D*, and Table E8 in this article's Online Repository at www.jacionline.org). In addition, a number of peaks were uniquely found in 1 cell line (Fig 4, *C*, Table E8). These data suggest the shared and cell type–specific chromatin interactions between the *CLEC16A* intronic region and the regulatory regions of other genes.

We further detected a stronger interaction between the G allele of rs11642009 and *ATF7IP2* than the T allele of rs11642009 in Jurkat cells through 4C experiment, suggesting the allele-specific effect (Fig E5 in this article's Online Repository at www.jacionline.org).

Distinct transcription factor complexes mediate the allele specific chromatin looping

We next sought to explore the transcription factor complex mediating the chromatin interactions between rs11642009 and the target genes. We performed SILAC experiment followed by MS, to identify the proteins that bind to the 61-bp oligonucleotides containing the 2 alleles of rs11642009, respectively (Fig 5, *A*). Jurkat cells were grown in SILAC medium containing regular lysine and arginine (light medium) or ^{13}C6-lysine and ^{13}C6-arginine (heavy medium), so that the cells grown in these 2 conditions were labeled with light and heavy isotopes, respectively. The nuclear extracts from the "light" cells were used to pull down the G allele and those from the "heavy" cells were used to pull down the T allele in this "2-way" oligo pull-down experiments, and then were 1:1 mixed for SDS-PAGE and MS. A reverse pull-down experiment was performed with the SILAC-labeled extracts switched to assess the reliability and reproducibility of the experimental approach. The proteins pulled down were quantified by MS, and multiple proteins showed differential binding to the 2 alleles of rs11642009 (FC of peptide abundance in the pull-down complex was consistently >1.5 between the alleles). The differentially bound proteins include transcription factors RUNX1, YY1, ZNF143, and KLF16, which were predicted to bind the genomic sequence where SNP rs11642009 is located (Fig 5, *B*). Interestingly, RUNX1 is involved in regulation of DNA-methylation at hematopoietic gene promoters.[21] $CD4^+$ T cells with Runx1 knockout (KO) tend to have spontaneous activation, leading to the development of autoimmune disease in mice.[22,23] Higher level of RUNX1

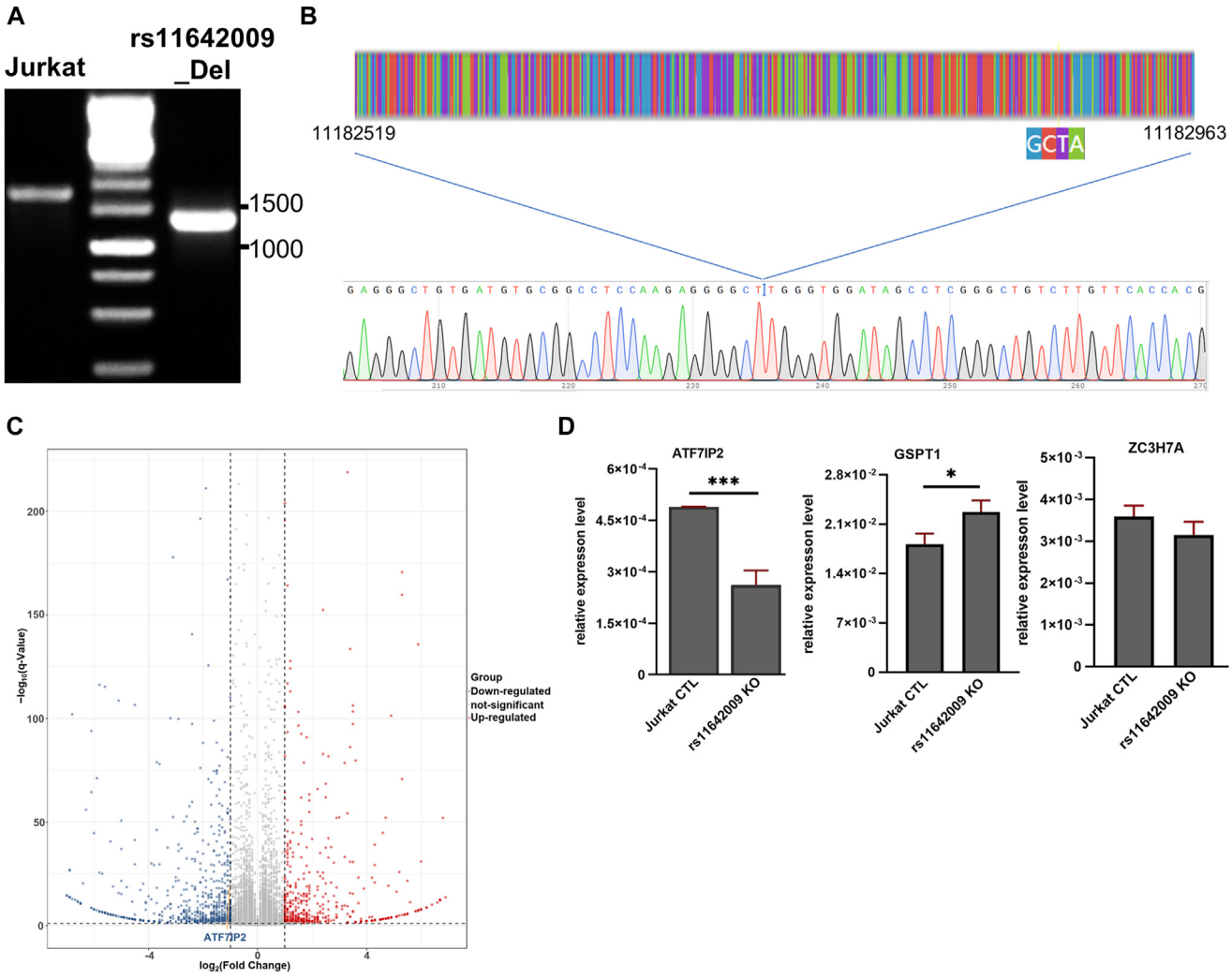

FIG 3. Generation of Jurkat cell line with the deletion of SNP rs11642009 genomic region. **A,** PCR confirming the 443-bp deletion at rs11642009 genomic region. PCR primers are located at the flanking region of rs11642009. The PCR product for the Jurkat cells is 1748 bp (lane 1) and that for the cells with the deletion is 1305 bp (lane 3). Molecular marker is shown in lane 2. **B,** Sanger sequencing confirmed the deletion of rs11642009 genomic region. The deleted sequence is represented by colors: *blue* for G, *red* for C, *purple* for T, *green* for A. **C,** The volcano plot showing RNA-seq result comparing rs11642009 deletion cell line versus parental Jurkat cell line. The significantly upregulated genes are shown in *red* and the significantly downregulated genes are shown in *blue*. **D,** Quantitative PCR validation of the expression level change of genes in the vicinity of rs11642009. $*P < .05$, $***P < .001$.

binding was detected for the G-allele of rs11642009, which showed a stronger chromatin interaction with *ATF7IP2* upstream region. Differential analysis demonstrated that the abundance of peptides from 38 proteins was higher for G allele than T allele, while the peptide abundance for 25 proteins was higher for T allele (Table E9 in this article's Online Repository at www.jacionline.org). Many of these proteins are known to interact with each other (Fig 5, *C*). As expected, these proteins are involved in pathways related to transcriptional regulation (Fig 5, *D*, and Table E10 in this article's Online Repository at www.jacionline.org). The results suggest transcription factor complexes may bind to the genomic sequence of rs11642009 in an allele-specific manner to mediate its interaction with target genes.

SNP rs11642009 deletion induces potent immune signaling alteration

In RNA-seq analysis, the deletion of rs11642009 induced the altered expression of 1291 genes. Pathway analysis demonstrated the overrepresentation of primary immunodeficiency pathway that included 7 differentially expressed genes DEGs: *BTK, CD4, CD79A, CD8B, IGLL1, RAG1,* and *RAG2* (Fig 6, *A*). BTK and CD79A function in the B-cell receptor signaling pathway;[24] CD4 and CD8B play important roles in T-cell receptor signaling;[25] and RAG1 and RAG2 initiate V(D)J (variable [diversity] joining gene segment) recombination, which is critical for B-cell and T-cell function.[26] The genes carrying deleterious mutations for familial forms of CVID also fall into these 3 important functional categories.[27] The discovery of these DEGs suggests

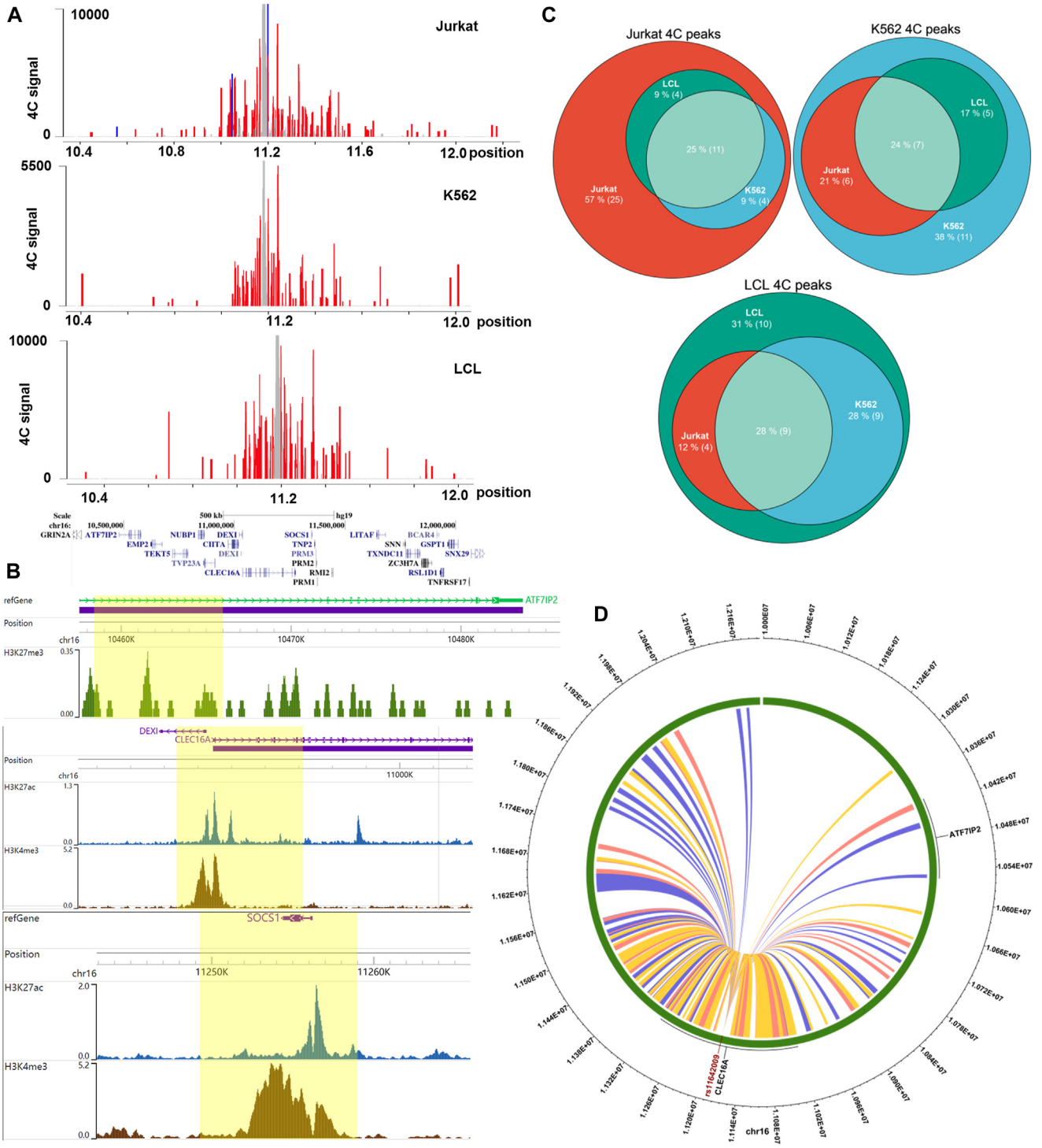

FIG 4. The 4C-seq analyses showing the chromatin interactions between rs11642009 and other genomic regions. **A,** The 4C-seq signals in Jurkat, K562 and LCL cells. The viewpoint is shown in *gray*; the significant 4C-seq peaks are shown in *red*; the most significant peak in or close to genes *ATF7IP2*, *CLEC16A*, and *SOCS1* in Jurkat cells are shown in *blue*. **B,** The 4C-seq signals in or close to genes *ATF7IP2*, *CLEC16A*, and *SOCS1* overlap with the different histone modifications from Roadmap database. **C,** The shared and unique 4C-seq signals in Jurkat, K562, and LCL cells. **D,** Circos plots of cis-chromatin interactions between rs11642009 and other genomic regions on chr6 shown in our study. The viewpoint rs11642009 is indicated in *red*. The 4C-seq interactions from the viewpoint to other genomic regions on chromosome 16 are indicated with a line colored according to the cell type: *blue*: Jurkat cells; *red*: K562 cells; *yellow*: LCL cells.

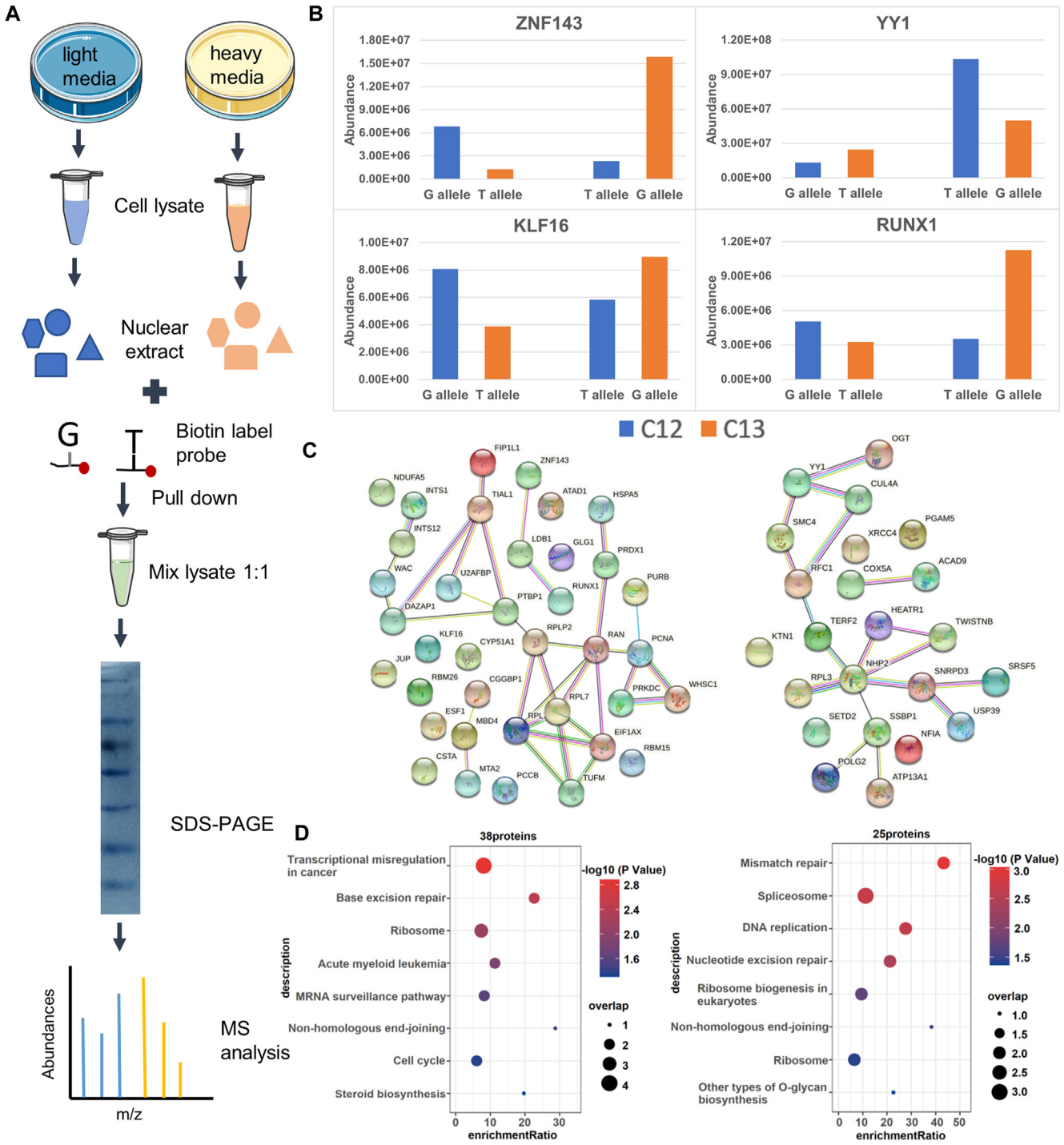

FIG 5. The experiment of SILAC showing the transcription factor complexes bound to the 2 alleles of rs11642009. **A,** The workflow of SILAC experiment. **B,** The peptide abundance of 4 transcription factors is significantly different between the 2 alleles of rs11642009. **C,** The protein-protein interactions between the 38 proteins showing higher peptide abundance in the protein complex pulled down by the G allele or the 25 proteins showing higher peptide abundance in the protein complex pulled down by the T allele (FC > 1.5) in the SILAC experiment. **D,** The overrepresented pathways among the 38 proteins or the 25 proteins for the G allele or the T allele, respectively.

the validity of our study. Furthermore, the pathway hematopoietic cell lineage was highly significantly overrepresented (Table E11 in this article's Online Repository at www.jacionline.org). Interestingly, genes in the mitogen-activated protein kinase signaling pathway and PI3K/AKT pathway showed significant change on rs11642009 deletion (Fig 6, *A*). In gene-set enrichment analysis, pathway antigen processing and presentation showed highly significant negative enrichment (Fig 6, *B*). Deletion of the rs11642009 genomic region in CCRF-CEM cells similarly resulted in alterations of the PI3K/AKT pathway and various

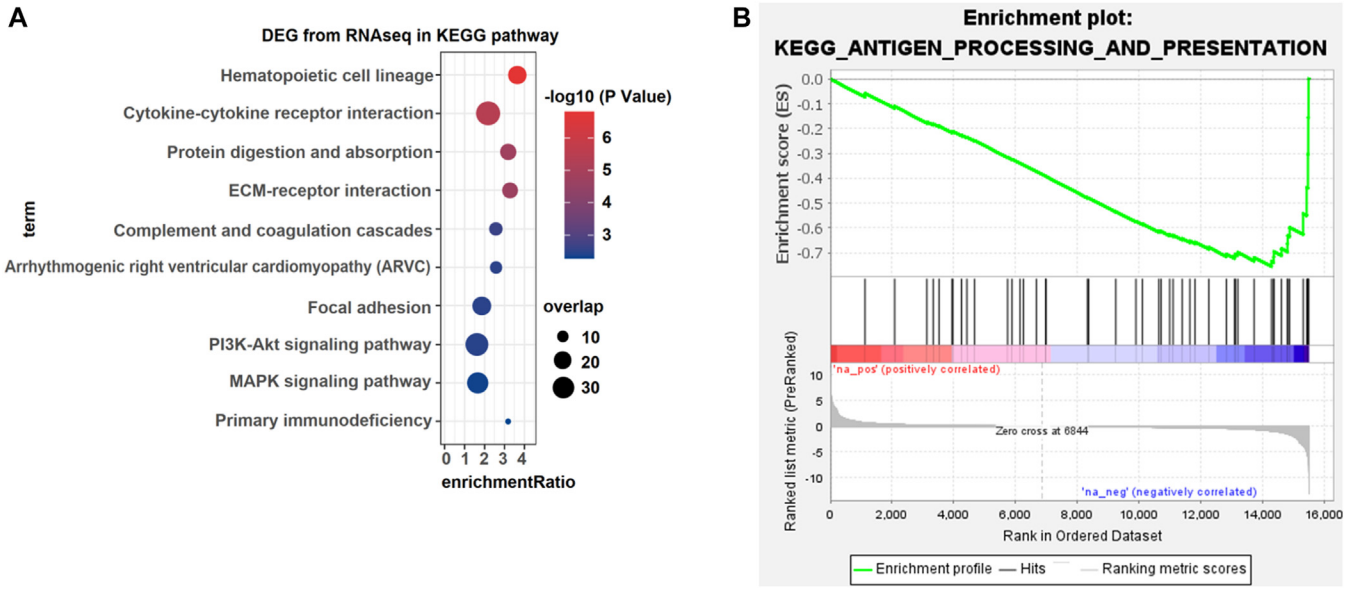

FIG 6. Pathway analysis of DEGs from RNA-seq. **A,** The result of pathway overrepresentation analysis on the DEGs in RNA-seq experiment. **B,** The result of pathway enrichment analysis on the RNA-seq result. *ECM,* extracellular matrix; *KEGG,* Kyoto Encyclopedia of Genes and Genomes; *MAPK,* mitogen-activated protein kinase.

immune signaling pathways (Fig E2, and Table E12 in this article's Online Repository at www.jacionline.org). These results suggest the deletion of rs11642009 leads to potent immune signaling alterations, which are likely to be transduced by the genes within its vicinity.

The transduction of immune signaling usually occurs via a series of translational modifications. To examine the cellular change at the protein level and decipher the alteration of immune signaling pathways, we conducted a reverse phase protein array experiment to evaluate 306 proteins in Jurkat cells and rs11642009-deletion cells (Fig 7, *A*). We observed the decreased level of 30 proteins or phosphoproteins and increased level of 11 proteins (FC >1.5) (Fig 7, *B*, and Table E13 in this article's Online Repository at www.jacionline.org), including genes *YAP1*, *NR2F2*, *GAB2,* and *VCL,* which have been involved in immune regulation.[26,28-30] The phosphorylation level of AKT was prominently enhanced (Fig 7, *C*, and Fig E6 in this article's Online Repository at www.jacionline.org).

ATF7IP2 is likely to be one of the genes mediating the effect of rs11642009 region deletion on the AKT signaling. Copy number loss at the chr16:9624870-10467886 (hg18) covering genes *ATF7IP2* and *GRIN2A* was found among CVID cases but not in 3031 controls[31] (Fig E7 in this article's Online Repository at www.jacionline.org). *ATF7IP2* is abundantly expressed in $CD4^+$ T cells and other immune cell types[32-35] (Figs E8 and E9 in this article's Online Repository at www.jacionline.org), but the expression level of *GRIN2A* in immune tissues and cells is very low. Therefore, *ATF7IP2* is likely to be the key gene in this region contributing to the CVID pathogenesis. ATF7IP2 encodes an epigenetic mediator interacting with MBD1, SETDB1, and SP1 and that is involved in forming transcriptionally silent heterochromatin domains.[35] In our study, knocking down *ATF7IP2* expression in Jurkat and CCRF-CEM cells by small interfering RNA led to the increased phosphorylation level of AKT (Fig E10 in this article's Online Repository at www.jacionline.org).

rs11642009 deletion induces profound changes in response to stimulation

To investigate how signaling pathways involved in T-cell activation were affected by rs11642009 deletion, we treated Jurkat cells and rs11642009-deletion cells with immune stimulants. PMA activates protein kinase C and multiple intracellular signaling pathways, leading to T-cell activation. With the deletion of rs11642009, Jurkat cells exhibited distinct molecular changes from its parental cell line on PMA treatment (Fig 7, *A*). Though it has been shown that insulin receptor signaling controls T-cell proliferation and cytokine production through regulation of cell metabolism,[36] the molecular changes on insulin stimulation is much weaker than those under PMA treatment (Fig 7, *A*). Twenty-two proteins showed distinct response to insulin or PMA treatment in rs11642009-deletion cells compared to Jurkat cells, as defined by the criteria that a protein in one cell line showed significant change (FC ≥2 between treatment group and no treatment) but not in the other cell line. These proteins interact with each other and impact on immune systems and metabolism (Fig 7, *D-F*, and Table E14 in this article's Online Repository at www.jacionline.org). We observed an increase in the levels of AKT–phosphorylated S473 and AKT–phosphorylated T308 levels following PMA treatment in Jurkat cells, consistent with AKT activation in response to immune stimulation. However, the phosphorylation levels at these 2 residues were reduced in rs11642009-deletion cells on stimulation (Fig 7, *F*). We further validated this observation by Western blotting (Fig E6). These results suggest that rs11642009 deletion has an impact on the molecular response of Jurkat cells to PMA stimulation.

DISCUSSION

Understanding the molecular mechanism behind the coexistence of autoimmune diseases and PIDD is critical for the

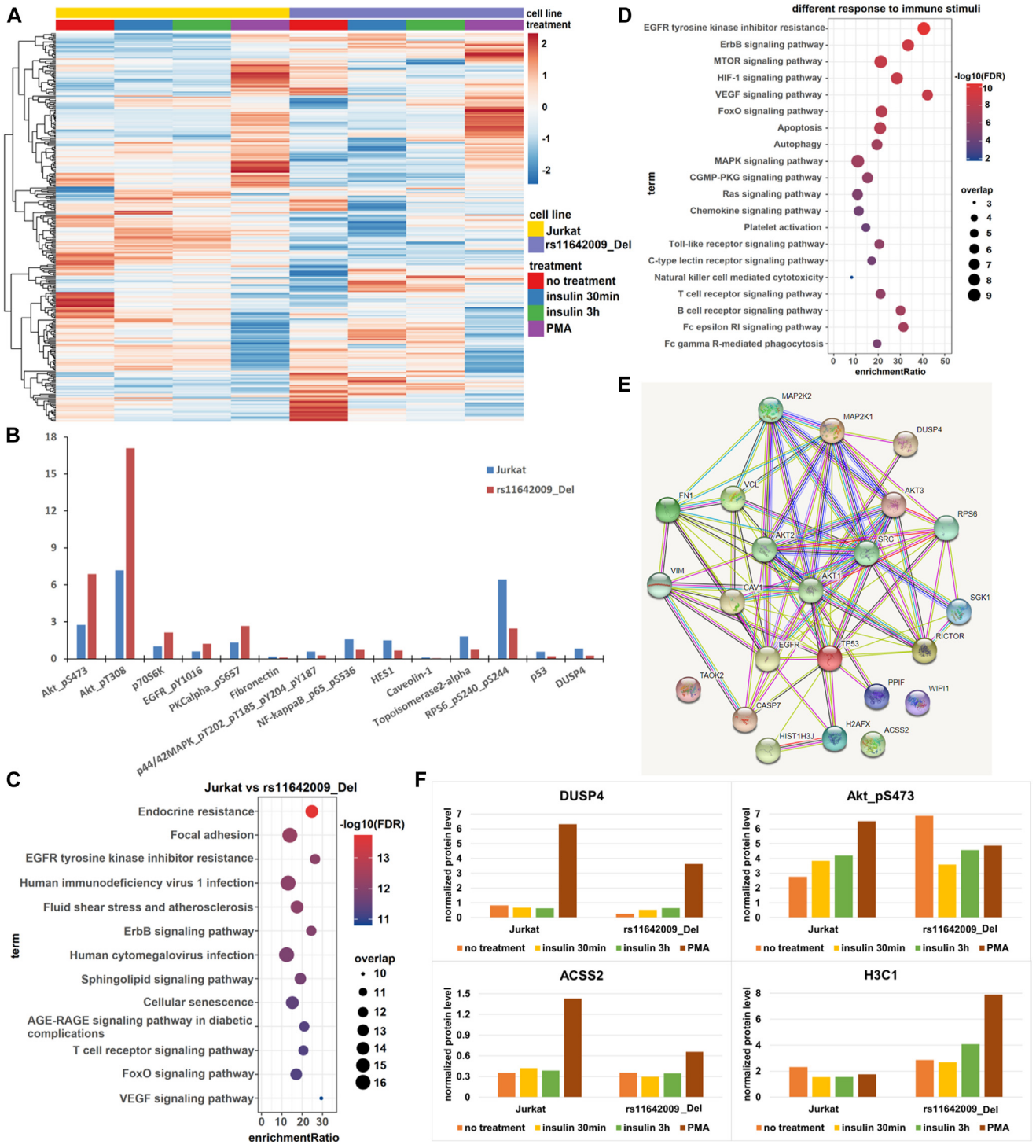

FIG 7. The experiment of reverse phase protein array examining alterations at the protein level on rs11642009 deletion. **A,** The heatmap showing the level of 306 proteins in 2 cell lines (Jurkat cells and Jurkat cells with rs11642009 deletion) under 4 conditions: no treatment, insulin treatment for 30 minutes, insulin treatment for 3 hours, PMA overnight. **B,** The 14 proteins with FC >2 in Jurkat cells and Jurkat cells with rs11642009 deleted. **C,** The overrepresented pathways among the proteins with FC ≥1.5 in the 2 cell lines. **D,** The overrepresented pathways among the 22 proteins showing different changes in Jurkat cells and Jurkat cells with rs11642009 deletion under insulin/PMA treatment. **E,** Interactions among the 22 proteins. **F,** Examples of proteins that showed different changes in Jurkat cells and Jurkat cells with rs11642009 deletion under insulin/PMA treatment. *AGE*, advanced glycation end product; *RAGE*, receptor for AGE.

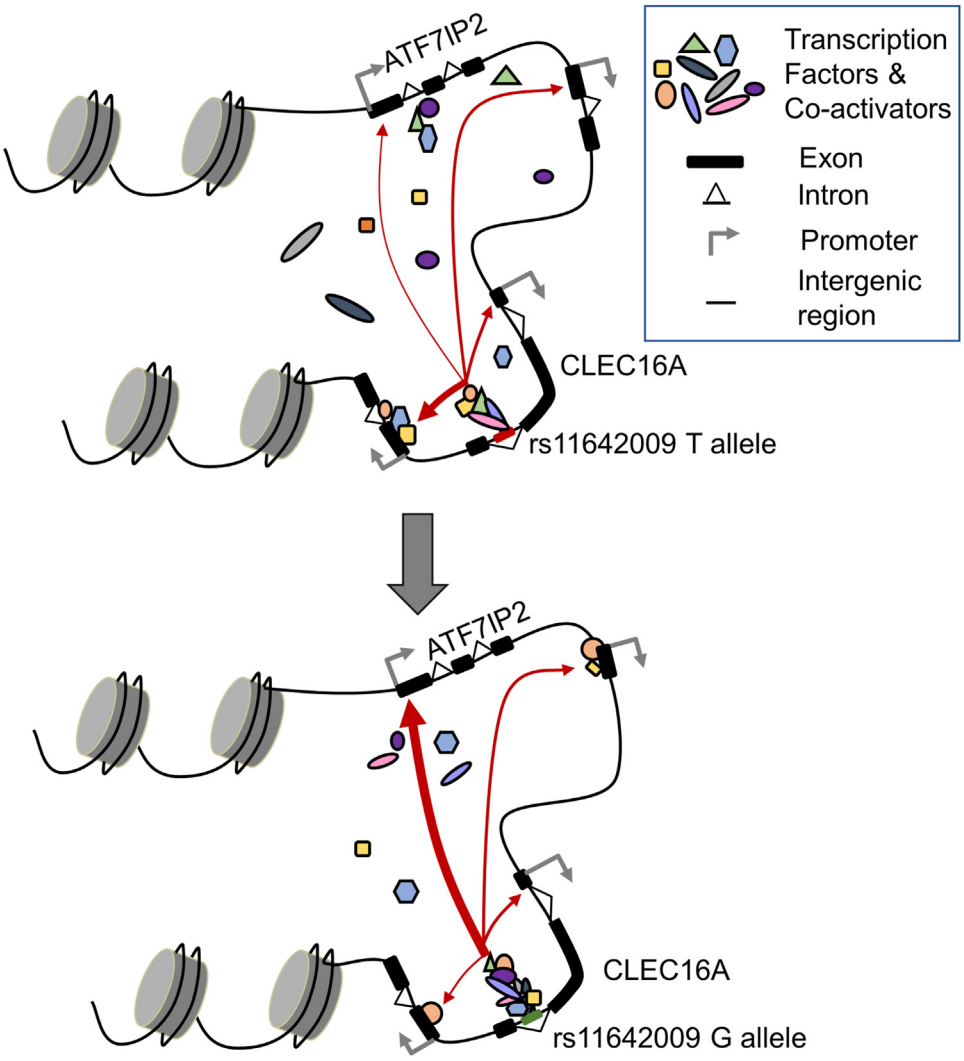

FIG 8. The model showing the enhancer role of rs11642009 in $CD4^+$ T cells. The *CELC16A* intronic SNP rs11642009 functions as an enhancer regulating the transcription of multiple genes. The cell type–dependent chromatin interactions between rs11642009 and the other genomic regions are mediated by different transcription factor complexes in an allele-specific manner, which may result in different interaction strength.

development of effective therapeutics.[2] *CLEC16A* intron 19 has been identified of association with numerous immunological traits and autoimmune diseases.[5] This locus is also associated with atopic diseases.[15] However, the effect size in this context is very small ($\beta = -0.0612$) when compared to its association with CVID ($\beta = -0.402$). The highly significant *P* value ($P = 5.69E-20$) for its association with atopy is driven by the extremely large sample size (n [cases] = 180,129; n [controls] = 360,838). Therefore, the strong association of *CLEC16A* intron 19 with CVID is unlikely to be confounded by contribution from atopy and it might offer a potential key to explaining the comorbidity of autoimmunity and PIDD.

The identified *CLEC16A* intronic variants are usually common variants with minor allele frequency >5%, while their associated autoimmune diseases and PIDD, such as type 1 diabetes, multiple sclerosis, and CVID, are of much lower prevalence. The contribution of common variants to such uncommon or even rare diseases has been reported in many studies.[37-40] Common variants play important roles in the pathoetiology of rare diseases in several aspects. Common variants confer disease susceptibility and predisposition, though not as the detrimental disease-causing mutations. Common variants can have disease-modifier effects on the penetrance or severity of the diseases, even for Mendelian disease.[41] Another aspect would be the less well-studied interactions between common variants and rare variants and common variants and environmental factors. Therefore, complex human diseases, including both common diseases and rare diseases, are shaped by the interplay among common SNPs, rare variants, and environmental factors.[42,43] Given the well-recognized contribution of common variants in different types of complex human diseases, the molecular mechanisms underlying the genetic findings remained elusive to our knowledge.

Through multiomics approaches, we elucidated a molecular mechanism underlying the genetic susceptibility of CVID and its comorbid autoimmunity mediated by *CLEC16A* variants. One of the plausible reasons for the difference in disease susceptibility

between the 2 alleles of the locus could be the diverse chromatin interactions mediated by differential binding of transcription factors to the 2 alleles. *ATF7IP2*, an interesting epigenetic mediator in this locus, is highlighted for potential gene-environmental interactions (Fig 8).

The association of *CLEC16A* intronic region with immunodeficiency and autoimmune diseases is related to its role as a cell type–dependent enhancer. In addition to B-cell defects, T-cell abnormalities frequently occur in patients with immunodeficiency and autoimmune comorbidity.[44-47] In our study, we showed the enhancer role of rs11642009 in $CD4^+$ T cells, which adds to our previous finding that *CLEC16A* KO in mice results in B-cell defects.[5] Others have shown that *CLEC16A* intronic SNPs are associated with higher expression of *CLEC16A* in $CD4^+$ T cells and T-cell activation stimulates *CLEC16A* expression, but *CLEC16A* knockdown did not affect T-cell activation.[48] Our results showed that multiple genes in the proximity of rs11642009 exhibited much more profound alteration than *CLEC16A* in $CD4^+$ T cells with the deletion of *CLEC16A* intron deletion, suggesting that these genes are the major target genes of this locus and the contribution of the *CLEC16A* intronic SNPs to immune disorders may be mediated by different genes in diverse cell types. Gene expression of *CIITA*, *DEXI,* and *SOCS1* did not show significant change could be due to redundant regulation in the other genomic regions.[49]

ATF7IP2, one of the target genes regulated by *CLEC16A* intronic variants, encodes an epigenetic mediator involved in forming transcriptionally silent heterochromatin domains. The KO mice for *ATF7IP2*'s interacting proteins—MBD1, SETDB1, or SP1—each showed defects in the immune system. The outstanding phenotypes include increased mean corpuscular hemoglobin (MBD1 KO), increased $CD4^+$, α/β T-cell number (SETDB1 KO), and decreased erythroid progenitor (SP1 heterozygous null).[50] *ATF7IP*, the ortholog of *ATF7IP2*, represses IL2 production by the deposition of H3K9me3 in the Il2-Il21 intergenic region. T cells with *ATF7IP* deletion had abnormally increased production of IL-2 on T-cell receptor stimulation and deficient T_H17 differentiation.[51] Our study showed the prominent reduction of *ATF7IP2* expression on rs11642009 deletion, as well as the chromatin interactions between rs11642009 and proximal regions of *ATF7IP2* in Jurkat cells and CCRF-CEM cells. Given the important functions of its ortholog *ATF7IP* and its interactors in T cells, the involvement of *ATF7IP2* in the immune system represents an intriguing epigenetic mechanism mediated by the *CLEC16A* locus.

Dysregulation of the balance between subtypes of $CD4^+$ T cells including T_H1, T_H2, T_H17, and regulatory T cells is a major factor driving pathogenesis of autoimmune diseases.[52] AKT signaling plays a crucial role in the development, differentiation, activation, and function of $CD4^+$ T cells. The regulation of AKT signaling by ATF7IP2 interactors have been demonstrated in cancer cells, but have been less investigated in T cells.

It has been shown that *CLEC16A* KO mice exhibited increased AKT phosphorylation levels in whole splenocyte lysates, which was effectively reversed by PI3K inhibitor Wortmannin, suggesting the induced activation of Akt signaling in *CLEC16A* KO mice.[53] In our study, the mRNA level of *CLEC16A* did not show remarkable change with rs11642009 deletion. In contrast, we observed simultaneously the prominent elevation of AKT phosphorylation level, suggesting that differential expression of other target gene(s) of rs11642009 also contributes to the altered PI3K/AKT signaling in $CD4^+$ T cells. It is in line with the association of this locus with autoimmunity in which abnormality of $CD4^+$ T-cell activation occurs. On antigen stimulation, for example, by PMA under experimental conditions, AKT phosphorylation was reduced with rs11642009 deletion, suggesting that rs11642009-deletion cells altered their ability to further respond to stimulation signal. AKT signaling has been an important therapeutic target for inflammation, autoimmunity, and cancer.[54,55] The findings in our study may direct the development of novel therapeutics targeting at AKT signaling for PIDD and autoimmune diseases.

There are limitations in our study, due primarily to technical constraints. First, our approach involved deleting a 443-bp genomic region rather than a single nucleotide. In this region, SNP rs11642009 is the only common variant, with all other 149 variants having a minor allele frequency <0.001.[56] None of these variants has been reported in ClinVar,[57] identified as expression quantitative trait loci variant,[58] recognized as functional rare genetic variation,[59] or predicted to affect splicing.[60] However, the possibility that this region contains additional, yet unrecognized, functional variants cannot be ruled out. Second, the use of a homozygous deletion approach limits our ability to pinpoint allelic or haploinsufficiency effects, an area that warrants further investigation. Third, we employed $CD4^+$ T-cell lines. In the future, the inclusion of primary cells would enhance the study's relevance and applicability. The genetic diversity inherent in different individuals necessitates the comparison of multiple subjects, including those without the variant, to obtain comprehensive insights. In addition, it would be highly advantageous for future studies to employ single-nucleotide editing techniques on rs11642009 within primary $CD4^+$ T cells. Such an approach would allow for a more nuanced and accurate investigation of the SNP's effects, facilitated by an isogenic background and conditions that more closely mimic physiological realities, thereby providing a clearer understanding of the specific impacts of rs11642009.

In summary, through multiomics approaches, we elucidated the transcriptional regulatory role of *CLEC16A* intron 19 locus by which it may contribute to the pathogenesis of CVID. We demonstrated that the rs11642009 functions as an enhancer, regulating the expression of multiple target genes, while the latter (especially AKT signaling) represents plausible targets for developing novel therapeutics for autoimmunity and primary immunodeficiency.

Data sharing statement

The CVID GWAS summary statistics can be downloaded from https://github.com/CAG-CNV/CVID_GWAS_summary_statistics_immunochip. The other high-throughput sequencing data are deposited at the Genome Sequence Archive for Human at China National Center for Bioinfomation (https://ngdc.cncb.ac.cn) with accession number HRA005829.

DISCLOSURE STATEMENT

This project was funded by National Natural Science Foundation of China (81771769), the Natural Science Foundation of Guangdong Province (2021A1515012392), Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-040A), Tianjin Outstanding Health Professional Selection and

Training Program (TJSJMYXYC-D2-031), and institute development funds to the Center for Applied Genomics at CHOP.

Disclosure of potential conflict of interest: Z. Ding and N. Wang are employees of Fynn Biotechnologies Ltd. The rest of the authors declare that they have no relevant conflicts of interests.

We thank Prof Yong Cui at the Department of Dermatology, China–Japan Friendship Hospital, for providing us the summary statistics of rs11642009 in the meta-analysis of genomic studies on SLE in the East Asian population. We thank the Core Facility of Research Center of Basic Medical Sciences and the High-Performance Computing Platform at Tianjin Medical University, National Supercomputer Center of Tianjin for technical support.

Clinical implications: The *CLEC16A* intronic variants may contribute to the pathological development of CVID and its autoimmune comorbidity by affecting its target genes and the AKT signaling, which represent plausible targets for developing novel therapeutics.

REFERENCES

1. Goyal R, Bulua AC, Nikolov NP, Schwartzberg PL, Siegel RM. Rheumatologic and autoimmune manifestations of primary immunodeficiency disorders. Curr Opin Rheumatol 2009;21:78-84.
2. Schmidt RE, Grimbacher B, Witte T. Autoimmunity and primary immunodeficiency: two sides of the same coin? Nat Rev Rheumatol 2017;14:7-18.
3. Agarwal S, Cunningham-Rundles C. Autoimmunity in common variable immunodeficiency. Curr Allergy Asthma Rep 2009;9:347-52.
4. Azar A, Aldaoud N, Hardenbergh D, Krimins R, Son J, Shiroky J, et al. Systemic lupus erythematosus and common variable immunodeficiency. J Clin Rheumatol 2022;28:e245-8.
5. Li J, Jorgensen SF, Maggadottir SM, Bakay M, Warnatz K, Glessner J, et al. Association of CLEC16A with human common variable immunodeficiency disorder and role in murine B cells. Nat Commun 2015;6:6804.
6. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41:703-7.
7. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.
8. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature 2007;448:591-4.
9. De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet 2009;41:776-82.
10. Bentham J, Morris DL, Graham DSC, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nat Genet 2015;47:1457-64.
11. Skinningsrud B, Lie BA, Husebye ES, Kvien TK, Førre Ø, Flatø B, et al. A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-cyclic citrullinated peptide antibody negative rheumatoid arthritis. Ann Rheum Dis 2010;69:1471-4.
12. Davison LJ, Wallace C, Cooper JD, Cope NF, Wilson NK, Smyth DJ, et al. Long-range DNA looping and gene expression analyses identify DEXI as an autoimmune disease candidate gene. Hum Mol Genet 2012;21:322-33.
13. Chen MH, Raffield LM, Mousas A, Sakaue S, Huffman JE, Moscati A, et al. Trans-ethnic and ancestry-specific blood-cell genetics in 746,667 individuals from 5 global populations. Cell 2020;182:1198-213e.14.
14. Yin X, Kim K, Suetsugu H, Bang SY, Wen L, Koido M, et al. Meta-analysis of 208370 East Asians identifies 113 susceptibility loci for systemic lupus erythematosus. Ann Rheum Dis 2021;80:632-40.
15. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet 2017;49:1752-7.
16. Thaventhiran JED, Allen HL, Burren OS, Rae W, Greene D, Staples E, et al. Whole-genome sequencing of a sporadic primary immunodeficiency cohort. Nature 2020;583:90-5.
17. Carbone J, Sarmiento E, Micheloud D, Rodriguez-Molina J, Fernandez-Cruz E. Elevated levels of activated CD4 T cells in common variable immunodeficiency: association with clinical findings. Allergol Immunopathol (Madr) 2006;34:131-5.
18. Azizi G, Rezaei N, Kiaee F, Tavakolinia N, Yazdani R, Mirshafiey A, et al. T-cell abnormalities in common variable immunodeficiency. J Investig Allergol Clin Immunol 2016;26:233-43.
19. Schuster C, Gerold KD, Schober K, Probst L, Boerner K, Kim MJ, et al. The autoimmunity-associated gene CLEC16A modulates thymic epithelial cell autophagy and alters T cell selection. Immunity 2015;42:942-52.
20. LaMere SA, Thompson RC, Meng X, Komori HK, Mark A, Salomon DR. H3K27 methylation dynamics during CD4 T cell activation: regulation of JAK/STAT and IL12RB2 expression by JMJD3. J Immunol 2017;199:3158-75.
21. Suzuki T, Shimizu Y, Furuhata E, Maeda S, Kishima M, Nishimura H, et al. RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. Blood Adv 2017;1:1699-711.
22. Naoe Y, Setoguchi R, Akiyama K, Muroi S, Kuroda M, Hatam F, et al. Repression of interleukin-4 in T helper type 1 cells by Runx/Cbf beta binding to the Il4 silencer. J Exp Med 2007;204:1749-55.
23. Korinfskaya S, Parameswaran S, Weirauch MT, Barski A. Runx transcription factors in T cells—what is beyond thymic development? Front Immunol 2021;12:701924.
24. Nemazee D. Mechanisms of central tolerance for B cells. Nat Rev Immunol 2017;17:281-94.
25. Nomura A, Taniuchi I. The role of CD8 downregulation during thymocyte differentiation. Trends Immunol 2020;41:972-81.
26. Gan T, Wang Y, Liu Y, Schatz DG, Hu J. RAG2 abolishes RAG1 aggregation to facilitate V(D)J recombination. Cell Rep 2021;37:109824.
27. Maggadottir SM, Li J, Glessner JT, Li YR, Wei Z, Chang X, et al. Rare variants at 16p11.2 are associated with common variable immunodeficiency. J Allergy Clin Immunol 2015;135:1569-77.
28. Olson WJ, Jakic B, Labi V, Schoeler K, Kind M, Klepsch V, et al. Orphan nuclear receptor NR2F6 suppresses T follicular helper cell accumulation through regulation of IL-21. Cell Rep 2019;28:2878-91.e5.
29. Wang Z, Vaughan TY, Zhu W, Chen Y, Fu G, Medrzycki M, et al. Gab2 and Gab3 redundantly suppress colitis by modulating macrophage and CD8(+) T-cell activation. Front Immunol 2019;10:486.
30. Saez de Guinoa J, Barrio L, Carrasco YR. Vinculin arrests motile B cells by stabilizing integrin clustering at the immune synapse. J Immunol 2013;191:2742-51.
31. Orange JS, Glessner JT, Resnick E, Sullivan KE, Lucas M, Ferry B, et al. Genome-wide association identifies diverse causes of common variable immunodeficiency. J Allergy Clin Immunol 2011;127:1360-79.e6.
32. Uhlen M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics: tissue-based map of the human proteome. Science 2015;347:1260419.
33. Wu CL, Jin XF, Tsueng G, Afrasiabi C, Su AI. BioGPS: building your own mash-up of gene annotations and expression profiles. Nucleic Acids Res 2016;44:D313-6.
34. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of genetic polymorphisms on human immune cell gene expression. Cell 2018;175:1701-15.e16.
35. Ichimura T, Watanabe S, Sakamoto Y, Aoto T, Fujita N, Nakao M. Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins. J Biol Chem 2005;280:13928-35.
36. Tsai S, Clemente-Casares X, Zhou AC, Lei H, Ahn JJ, Chan YT, et al. Insulin receptor-mediated stimulation boosts T cell immunity during inflammation and infection. Cell Metab 2018;28:922-34.e4.
37. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet 2013;45:664-9.
38. Li YR, Li J, Zhao SD, Bradfield JP, Mentch FD, Maggadottir SM, et al. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. Nat Med 2015;21:1018-27.
39. Barc J, Tadros R, Glinge C, Chiang DY, Jouni M, Simonet F, et al. Genome-wide association analyses identify new Brugada syndrome risk loci and highlight a new mechanism of sodium channel regulation in disease susceptibility. Nat Genet 2022;54:232-9.
40. Bezzina CR, Barc J, Mizusawa Y, Remme CA, Gourraud JB, Simonet F, et al. Common variants at SCN5A-SCN10A and HEY2 are associated with Brugada syndrome, a rare disease with high risk of sudden cardiac death. Nat Genet 2013;45:1404-9.
41. Blair DR, Hoffmann TJ, Shieh JT. Common genetic variation associated with Mendelian disease severity revealed through cryptic phenotype analysis. Nat Commun 2022;13:3675.
42. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. Cell 2017;169:1177-786.
43. Mathieson I. The omnigenic model and polygenic prediction of complex traits. Am J Hum Genet 2021;108:1558-63.
44. Bateman EA, Ayers L, Sadler R, Lucas M, Roberts C, Woods A, et al. T cell phenotypes in patients with common variable immunodeficiency disorders:

associations with clinical phenotypes in comparison with other groups with recurrent infections. Clin Exp Immunol 2012;170:202-11.
45. Comte D, Karampetsou MP, Kis-Toth K, Yoshida N, Bradley SJ, Kyttaris VC, et al. Brief report: CD4+ T cells from patients with systemic lupus erythematosus respond poorly to exogenous interleukin-2. Arthritis Rheumatol 2017;69:808-13.
46. Suarez-Fueyo A, Bradley SJ, Tsokos GC. T cells in systemic lupus erythematosus. Curr Opin Immunol 2016;43:32-8.
47. Wong GK, Huissoon AP. T-cell abnormalities in common variable immunodeficiency: the hidden defect. J Clin Pathol 2016;69:672-6.
48. Eriksson AM, Leikfoss IS, Abrahamsen G, Sundvold V, Isom MM, Keshari PK, et al. Exploring the role of the multiple sclerosis susceptibility gene CLEC16A in T cells. Scand J Immunol 2021;94:e13050.
49. Zuvich RL, Bush WS, McCauley JL, Beecham AH, De Jager PL, International Multiple Sclerosis Genetics Consortium, et al. Interrogating the complex role of chromosome 16p13. 13 in multiple sclerosis susceptibility: independent genetic signals in the CIITA–CLEC16A–SOCS1 gene complex. Hum Mol Genet 2011;20:3517-24.
50. Kruger I, Vollmer M, Simmons DG, Elsasser HP, Philipsen S, Suske G. Sp1/Sp3 compound heterozygous mice are not viable: impaired erythropoiesis and severe placental defects. Dev Dyn 2007;236:2235-44.
51. Sin JH, Zuckerman C, Cortez JT, Eckalbar WL, Erle DJ, Anderson MS, et al. The epigenetic regulator ATF7ip inhibits Il2 expression, regulating Th17 responses. J Exp Med 2019;216:2024-37.
52. Jager A, Kuchroo VK. Effector and regulatory T-cell subsets in autoimmunity and tissue inflammation. Scand J Immunol 2010;72:173-84.
53. Pandey R, Bakay M, Hain HS, Strenkowski B, Elsaqa BZB, Roizen JD, et al. CLEC16A regulates splenocyte and NK cell function in part through MEK signaling. PLoS One 2018;13:e0203952.
54. He Y, Sun MM, Zhang GG, Yang J, Chen KS, Xu WW, et al. Targeting PI3K/Akt signal transduction for cancer therapy. Signal Transduct Target Ther 2021;6:425.
55. Zarrin AA, Bao K, Lupardus P, Vucic D. Kinase inhibition in autoimmunity and inflammation. Nat Rev Drug Discov 2021;20:39-63.
56. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434-43.
57. Landrum MJ, Chitipiralla S, Brown GR, Chen C, Gu B, Hart J, et al. ClinVar: improvements to accessing data. Nucleic Acids Res 2020;48:D835-44.
58. Li J, Kong N, Han B, Sul JH. Rare variants regulate expression of nearby individual genes in multiple tissues. PLoS Genet 2021;17:e1009596.
59. Ferraro NM, Strober BJ, Einson J, Abell NS, Aguet F, Barbeira AN, et al. Transcriptomic signatures across human tissues identify functional rare genetic variation. Science 2020;369:eaaz5900.
60. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell 2019;176:535-48.e24.