

Identification of Genetic Predisposition to Sjögren's Syndrome by Whole Exome Sequencing

Qiwei Guo^{a,b,c,d,1*}, Qiaowei Li^{e,1}, Huiqin Lu^{a,d,1}, Yingying Shi^{a,d}, Jintao Guo^d, Hao Wang^f, Qiuxiao Deng^d, Yihang Li^d, Yuan Liu^{a,c,d}, Guixiu Shi^{a,c,d*}, and Shiju Chen^{a,c,d*}

^a Department of Rheumatology and Clinical Immunology, The First Affiliated Hospital of Xiamen University, Xiamen 361001, China.

^b Department of Central Laboratory, Women and Children's Hospital, School of Medicine, Xiamen University, Xiamen 361003, China.

^c Xiamen Municipal Clinical Research Center for Immune Disease, Xiamen 361001, China.

^d School of Medicine, Xiamen University, Xiamen 361102, China.

^e School of Public Health, Xiamen University, Xiamen 361102, China.

^f School of Pharmaceutical sciences, Xiamen University, Xiamen 361102, China.

¹ These authors contributed equally.

* Corresponding author.

Email addresses: Qiwei Guo, guoqiwei@xmu.edu.cn; Guixiu Shi, gshi@xmu.edu.cn; Shiju Chen, shiju@xmu.edu.cn.

Abstract

A comprehensive understanding of the genetic predisposition associated with the initiation and development of Sjögren's syndrome (SjS) is imperative. This would not only enrich our knowledge of the pathogenesis underlying this autoimmune disease but also address the long-standing clinical challenges of more timely diagnosis and effective treatment to retain organ function and improve prognosis. In this study, we used whole exome sequencing analysis of 50 patients with SjS to investigate the predisposing variants, genes, and their associated biological functions. Hundreds of predisposing genes were identified, and numerous biological processes and pathways were highlighted; suggesting a heterogeneity of genetic predisposition to SjS. Female patients carrying a greater number of enriched variants tended to have higher levels of serum IgG and corresponding systemic involvement, demonstrating the pivotal role of genetic predisposition in the pathogenesis of SjS. Biological function analysis indicated that a subset of SjS and neuropathies may share a similar genetic predisposition. Our results showed that extracellular matrix-receptor interactions, macrophage-associated biological functions, and motor proteins may play important roles in the pathogenesis of SjS, and macrophage-associated biological functions may be associated with early onset SjS in female patients. Furthermore, the identification of highly enriched variants in the patient cohort provides the possibility of advancing the diagnosis of SjS. In conclusion, our study provides an extensive framework for analysis of the genetic predisposition to SjS which can facilitate further focused and in-depth investigation of the pathogenetic mechanisms of specific genes, biological processes, and pathways; thereby contributing to the pathophysiology, diagnosis, and therapeutics of SjS.

Keywords: genetic predisposition; pathogenesis, rare variant, Sjögren's syndrome, whole exome sequencing

1. Introduction

Sjögren's syndrome (SjS) is a systemic autoimmune disease that affects approximately 0.1–0.5% of the general population, and more than 90% of patients are women [1]. It is characterized by the disorganization and chronic inflammation of the exocrine glands, particularly the salivary and lacrimal glands, resulting in glandular dysfunction and sicca symptoms [2]. Heterogeneous extraglandular manifestations with multi-organ involvement, such as fatigue, synovitis, arthralgia, neuropathy, vasculitis, and interstitial lung disease, are also common in SjS [2]. Furthermore, approximately 5% of patients with SjS also develop non-Hodgkin lymphoma [3].

The pathogenesis of SjS is heterogeneous and remains to be fully elucidated. Generally, SjS initiates and develops when individuals with a genetic predisposition are exposed to environmental or hormonal factors [4]. In affected glands, both innate and adaptive immunity can be activated, causing inflammation and infiltration of immune cells; primarily T and B cells [5]. The pathogenic interplay between epithelial and immune cells can result in glandular tissue destruction, B-cell hyperactivity, formation of ectopic germinal centers, and local production of autoantibodies [6]. A comprehensive understanding of the genetic predisposition involved in the initiation and development of SjS is imperative as it not only enriches our knowledge of the pathogenesis underlying this disease but also addresses the long-standing clinical challenges of more timely diagnosis and effective treatment to retain organ function and improve prognosis [4].

In the past decade, several studies have generated initial insights into the genetic predisposition to SjS. Various predisposing human leukocyte antigen and non-human leukocyte antigen loci in different population backgrounds have been identified by genetic association analysis, a methodology frequently used for common diseases studies that focuses on common variants with allele frequencies ≥ 0.01 [7-11]. However, with the development of high-throughput sequencing, rare and deleterious variants with allele frequencies < 0.01 have been demonstrated to be more highly penetrant contributors to common diseases and have a strong impact on the predisposition to disease in individual patients [12, 13]. Johar et al. identified

several rare functional variants in multiple autoimmune syndromes and SjS using whole exome sequencing (WES); however, the sample size was limited (eight patients with SjS patients and other autoimmune diseases and four patients with SjS patients but no other autoimmune diseases), and only homozygous and compound heterozygous variants were analyzed [14]. Subsequently, Wang et al. used WES to analyze rare immune-related variants in 31 families (nine SjS, eleven systemic lupus erythematosus, six rheumatoid arthritis, and five mixed autoimmunity) and identified that the T cell receptor signaling pathway constitutes a shared genetic predisposition for these autoimmune diseases. However, non-immune-related genes were not analyzed [15]. Qi et al. demonstrated that the adenosine monophosphate-activated protein kinase signaling pathway and cell adhesion molecules are involved in the genetic predisposition to SjS, based on WES analysis of 12 patients with SjS (seven individuals from three pedigrees and five sporadic cases) [16]. However, our understanding of rare variant-based genetic predisposition to SjS is limited, and studies containing a larger cohort of sporadic cases are needed to delineate a more comprehensive profile.

In this study, we conducted WES analysis of 50 patients with SjS, most of whom were sporadic SjS cases, to investigate the predisposing variants, genes, and their associated biological functions; aiming to expand our knowledge of the pathogenesis of SjS and facilitate timely diagnosis and effective treatment of this autoimmune disease.

2. Materials and Methods

2.1. Patients and samples

Fifty patients who fulfilled the American College of Rheumatology/European League Against Rheumatism classification criteria for SjS were enrolled in the Department of Rheumatology and Clinical Immunology, First Affiliated Hospital of Xiamen University, China [6]. According to the available data, 40 patients were diagnosed with primary SjS, whereas 10 patients were diagnosed with associated SjS (concurrence of SjS with other autoimmune diseases). Currently, there is a lack of

sufficient evidence to distinguish primary from associated S_jS precisely, either from a clinical or biological perspective [1]; therefore, we analyzed these groups of patients non-distinctively. Except for two pairs of first-degree relatives, the patients were consanguineously unrelated. One milliliter of ethylenediaminetetraacetic acid-anticoagulated peripheral blood was collected from each patient for genomic DNA extraction. The patient information is listed in Supplemental Table 1.

2.2. Whole exome sequencing and variant output

Genomic DNA was extracted from 400 μ L of ethylenediaminetetraacetic acid-anticoagulated peripheral blood using the Super/HF16 Plus DNA Extraction System (MagCore, Xiamen, China) according to the manufacturer's instructions. Subsequently, the DNA samples were analyzed using a commercial WES service (Novogene, Beijing, China). Briefly, the exome sequences were enriched from 400 ng of genomic DNA using a liquid capture system (Agilent SureSelect Human All Exon V6; Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's protocol, and sequenced on the NovaSeq 6000 platform (Illumina, San Diego, CA, USA) to generate 150 bp paired-end reads. The mean exome coverage was above 100 \times , allowing examination of the selected region with sufficient depth to accurately match more than 99.5% of the targeted exome.

After quality control, high-quality reads were mapped to the human reference genome (hg38) using the Burrows Wheeler Aligner [17]. Subsequently, SAMtools was used for calling single nucleotide variants (SNVs) and insertions and deletions (indels) based on the following filter parameters: mapping quality \geq 30, quality by depth \geq 4, and Phred-scaled quality score of variant \geq 20 [18]. Called SNVs and indels were annotated with ANNOVAR to itemize bioinformatic content, such as the genomic location of the variant, amino acid changes, minor allele frequency (MAF), and *in silico* pathogenicity prediction [19].

2.3. Pipeline of variant filtering

Owing to the unsatisfactory accuracy of current WES for indel detection [20], only SNVs were analyzed in this study to reduce the impact of the innate error of WES and increase the specificity of candidate variant filtering. Genes in highly

homologous genomic regions that pose technical hurdles for variant analysis were also excluded [21].

A stepwise filtering pipeline was used to identify rare pathogenic variants (Fig. 1). Briefly, variants causing missense, frameshift, or nonsense mutations or splicing changes were retained, and rare variants were further retrieved according to MAF data in three multi-population databases (gnomAD, dbSNP, and 1000 Genomes) and three East Asian population databases (ChinaMAP, jMorp, and KRGDB) [22-26]. Finally, rare variants were filtered by *in silico* pathogenicity prediction using four algorithms, SIFT, PolyPhen-2, Mutation Taster, and CADD, to prioritize the possible pathogenic variants.

2.4. Variant validation by Sanger sequencing

Sanger sequencing is considered the gold standard for validation of candidate variants identified by WES. Therefore, DNA samples with target variants, and two control samples without target variants, were amplified with the corresponding sequencing primers (Supplemental Table 2) on a T100 Thermal Cycler (Bio-Rad, Hercules, CA, USA). The PCR conditions are described in detail in Supplementary Table 3. After amplification, the PCR products were sequenced, using Sanger sequencing, by a commercial sequencing service (Sangon, Shanghai, China).

2.5. Rarity validation by real-time PCR-based genotyping

Once the target variants were validated by Sanger sequencing, their rarity in the general population was confirmed using real-time PCR-based genotyping. Briefly, 100–300 genomic DNA samples from random female volunteers, equivalent to 200–600 alleles, were examined using real-time PCR in which hydrolysis probes were designed to identify target variants. Sanger sequencing-validated patient samples were used as positive and negative controls. The primers and hydrolysis probes are listed in Supplemental Tables 2 and 4, respectively.

2.6. Functional enrichment and pathway analysis

To determine the enriched biological processes and pathways associated with the candidate genes, The Database for Annotation, Visualization and Integrated Discovery (<http://david.abcc.ncifcrf.gov>) was used to perform Gene Ontology (GO) enrichment

and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses [27].

GraphPad Prism version 9 (GraphPad, Boston, MA, USA) was used to visualize the enriched GO terms and KEGG pathways.

2.7. Statistical analyses

Statistical analysis was performed using SPSS statistical software version 25.0 (IBM, Armonk, NY, USA). Pearson's (parametric) or Spearman's (non-parametric) correlation was performed to determine the correlations between the number of carried variants and different clinical indices.

2.8. Ethics statement

Signed informed consent was obtained from each participant, and all participants agreed to the use of their data for research purposes. This study was approved by the Research Ethics Committee of the First Affiliated Hospital of Xiamen University, China.

3. Results

3.1. Overview of candidate variants and genes

After filtering, 6,424 variants in 4,632 genes were identified. The allele count of a specific variant increased as the number of variants decreased (Fig. 2A). To confirm the accuracy of WES and our variant filtering pipeline, the 11 most enriched variants with allele counts of four and five were analyzed further using Sanger sequencing and real-time PCR-based genotyping. As shown in Supplemental Fig. 1, Sanger sequencing confirmed all 11 variants identified by WES. Examination of our sample cohort from the general population revealed that the allele frequencies of most variants (10 of 11) were lower than 0.01. However, the allele frequency of a *TRADD* variant was slightly higher than 0.01 (Fig. 2B), therefore, this was excluded from subsequent analyses. Our results suggest that the majority of candidate variants are rare in the general population, thus validating the variant filtering pipeline.

Most variants were heterozygous. When two or more heterozygous variants are detected in a single gene of a specific patient, both alleles of the gene can be affected in a compound heterozygous manner. Our results showed that 80 genes in 41 patients

could be affected by compound heterozygous variants (Supplemental Table 5). In addition to the heterozygous variants, eight variants were homozygous and located in eight genes from six patients (Supplemental Table 6).

Among the 4,632 candidate genes, 24 were identified as predisposing genes for SjS in previous studies, among which *RELN* was most frequently affected (Supplemental Fig. 2).

3.2. Functional analysis of enriched genes.

Considering that the presence of many variants with only one allele count could be a random rather than an enriched event in patients with SjS, this subset was excluded from further analysis to increase analytical specificity. Therefore, 617 variants with allele counts ≥ 2 (enriched variants) and 601 genes (enriched genes) were subsequently analyzed for their biological functions (Supplemental Tables 7).

Our results showed enrichment of several biological processes, including nervous system-associated biological processes, including nervous system development, single-organism behavior, cell projection organization, and neurological system process; cilium, including the motile and non-motile cilium, organization; immunological processes, including macrophage cytokine production, receptor-mediated endocytosis, and regulation of I-kappaB kinase/NF-kappaB signaling; process of female sex differentiation; and photoreceptor cell differentiation (Fig. 3A and Supplemental Table 8).

Several pathways were enriched for the 601 enriched genes, most of which were immune-related (Fig. 4A and Supplemental Table 9); underlining the contribution of an immune-related genetic predisposition to SjS. Among these enriched pathways, the extracellular matrix (ECM)-receptor interaction was the most enriched ($P=2.24E-06$).

3.3. Functional analysis of immune-related genes.

According to the immune-related genes list in the innateDB [28], 121 genes were immune-related among the 601 enriched genes (Supplemental Table 10). To further evaluate the genetic predisposition of immune factors in SjS, we performed functional analysis of the immune-related gene set.

Notably, cell adhesion was strongly enriched ($P=1.95E-18$); and the regulation of

signaling and associated cellular activities, such as cell activation, endocytosis, and cell migration, were also enriched. Furthermore, nervous system development was associated with this subset of genes. (Fig. 3B and Supplementary Table 11).

KEGG pathway analysis (Fig. 4B and Supplemental Table 12) showed that the ECM-receptor interaction pathway was strongly enriched ($P=1.13E-10$), which is consistent with the results from the full 601 enriched genes (Fig. 4A). The Fc gamma receptor (Fc γ R)-mediated phagocytosis and associated signaling pathways were also enriched.

3.4. Functional analysis of genes affected by homozygous variants and potentially compound heterozygous variants.

Compared to heterozygous variants, homozygous and compound heterozygous variants could result in impaired function of both alleles, and the affected genes would have a closer relationship with the pathogenesis of SjS. As shown in Supplemental Table 13, the main enriched biological processes were similar to those of the enriched genes (Supplemental Table 9), including regulation of homeostatic processes, especially calcium homeostasis; system processes, including neurological and muscle system processes; cell adhesion; and nervous system development, including glial and photoreceptor cell development.

KEGG pathway analysis revealed enrichment in the calcium signaling and pyrimidine metabolism pathways (Supplemental Table 14).

3.5. The presence of a greater number of enriched variants in female patients could promote autoimmunity and thus increase the risk of systemic involvement

In our patient cohort, 23 female patients did not receive any treatment at the time of enrolment (Supplemental Table 1), enabling us to investigate the effects of genetic predisposition on clinical indices. Statistical analyses were performed to evaluate the association between the number of enriched variants in an individual and various clinical indices. As shown in Table 2 and Fig. 5A and B, female patients who carried a greater number of enriched variants tended to have higher levels of IgG in their serum as well as higher European League Against Rheumatism Sjögren's syndrome disease activity index (ESSDAI) scores. Furthermore, higher serum IgG levels were

associated with higher ESSDAI scores (Fig. 5C). These results suggest that a higher number of predisposing variants in female patients could promote autoimmunity, thereby increasing the risk of systemic involvement.

3.6. Macrophage-associated processes are associated with early onset of disease in female patients.

Patients with an onset age of 35 or younger were defined as having early-onset SjS, whereas the others were defined as having regular- or late-onset SjS; according to the criteria from previous studies [29, 30]. In this context, 10 and 36 female patients were considered early onset and regular- and late-onset patients, respectively (Supplemental Table. 1). Eighty-one enriched genes were found to be specific to the early-onset subset, whereas 318 enriched genes were specific to the regular- and late-onset subset (Supplemental Fig. 3 and Supplemental Tables 15 and 16). Functional analysis showed that immune-associated processes, particularly macrophage-associated processes, were enriched in the early onset-specific gene set; whereas biological processes of nervous system development were enriched in the regular- and late-onset-specific gene sets (Fig. 3C and D, Supplemental Tables 17 and 18). Similarly, different enriched pathways were identified for the different age of onset subsets (Fig. 4C and D; Supplemental Tables 19 and 20).

4. Discussion

In this study, we investigated the genetic predisposition to SjS by evaluating enriched rare variants derived from WES data of 50 patients with SjS. By examining the allele frequencies of the 11 most enriched variants in our general population cohort (Fig. 2), we validated our filtering pipeline's ability to identify rare variants; which was a prerequisite for achieving the goal of this study. The enrichment of hundreds of rare variants in the SjS cohort, and the discovery that female patients who carry a greater number of enriched variants tend to manifest higher levels of serum IgG and corresponding disease activity, demonstrate the pivotal role of genetic predisposition in the initiation and development of this autoimmune disease.

Based on the functional analysis of enriched variant-associated gene sets, we

identified numerous associated biological processes and pathways (Fig. 3 and 4, Supplemental Tables 8–20); suggesting the heterogeneity of the genetic predisposition in SjS and providing an extensive framework for further focused and in-depth analyses to advance our knowledge of the pathogenesis of SjS.

A subset of SjS-associated predisposing genes was also involved in nervous system-associated biological processes, such as nervous system development and neurological system processes; suggesting an overlap in the genetic predisposition to SjS and neuropathies. Notably, this result was also identified in genes affected by homozygous and potentially compound heterozygous variants, which are considered closely related to the pathogenesis of SjS. An increasing number of studies have shown that neurological involvement in SjS has been previously underestimated or under-diagnosed [31-33]. More recently, Seeliger et al. demonstrated that nearly half of patients with SjS suffered from neurological involvement, including muscular, central, and peripheral nervous system, and this subset of patients have a different clinical phenotype from that of patients without neurological involvement [31]. These findings support our hypothesis that a subset of SjS and neuropathies share a similar genetic predisposition and suggest that a specific phenotype could be shaped by this underlying genetic predisposition. Seeliger et al. demonstrated that neurological involvement in SjS was associated with an older age of onset. This is consistent with our finding that genes involved in nervous system development were enriched in female patients with regular- and late-onset rather than early onset SjS (Fig. 3D and Supplemental Table 18), further suggesting the importance of genetic predisposition in phenotype shaping. Previously, neuropathies in SjS had been suggested to be secondary to the vasculopathy or glandular autoimmunity [34, 35]. However, along with the high incidence of neuropathies in SjS, the positive relationship between neuropathies and autoimmunity was challenged [33]. Furthermore, the therapeutic effects of immunosuppressants on SjS-associated neuropathies are limited in some cases [35]. These findings suggest that neuropathies, at least in part, result from a genetic predisposition rather than secondary injury to glandular autoimmunity [31, 33]. One of the putative etiologies underlying the hypothesis of shared genetic

predisposition is that specific predisposing genes could have multiple functions that simultaneously regulate the homeostasis of the nervous system, immune system, and salivary/lacrimal glands. Several lines of evidence associated with our identified enriched genes support this hypothesis. For example, the products of the four most enriched candidate genes, AQP1, ABL2, CHD3, and RELN, participate in inflammatory activation, cell adhesion and migration, wound healing, and organ regeneration [36-41]. In addition, they play important roles in the nervous system and are associated with central and periphery neuropathies [39, 42-46]. Motor proteins, which are the products of another subset of enriched genes, are essential for axonal transportation and function at immunological synapses [47, 48].

Macrophage-associated enriched variants can simultaneously affect the homeostasis of the nervous system and salivary/lacrimal glands by altering the function of tissue-resident macrophages [49, 50]. Another putative etiology underlying the hypothesis of shared genetic predisposition is that because the immune responses and gland morphogenesis, vascularization, and functionality can be modulated by the nervous system, its dysregulation disrupts homeostasis of the immune system or salivary/lacrimal glands [51-56].

Variants in genes encoding various ECM proteins, such as laminin, collagen, reelin, fibronectin, and thrombospondin, were enriched in our study; therefore, the ECM-receptor interaction was identified as an important player in the genetic predisposition to SjS (Fig. 3 and 4). The ECM provides various biomaterial scaffolds which cells can sense, adhere to, and remodel to control cell shape, motility, and signaling; and is essential for a broad array of biological processes such as development, hematopoiesis, immune response, wound healing, and organ regeneration [57]. In particular, several ECM-associated genes encode basement membrane (BM) proteins, such as LAMA2, LAMB1, LAMC3, FN1, COL6A3, COL9A2 [58], were identified in the present study. The BM is an ECM lining the mesenchymal-epithelial junction, which is physiologically pivotal for branching morphogenesis, cellular polarity of the salivary gland, and migration and differentiation of progenitor cells via interaction with associated receptors; thereby

maintaining glandular homeostasis [59, 60]. Disorganization of the BM owing to the pathogenic variants could not only result in the defective structure and function of salivary gland but also be associated with infection and mononuclear cell infiltration in SjS [61, 62] (Fig. 6). In the pathological state, ECM components can serve as damage-associated molecular patterns (DAMPs) that exacerbate and perpetuate inflammatory cascades and may provide a novel source of chronic B cell activation in SjS [63].

Macrophages are suggested to be an important immunological player in SjS, particularly in early onset, which manifests higher levels of autoimmunity and is associated with poor prognosis that should be monitored more strictly [29, 64]. Macrophages and their corresponding cytokines have been identified in the inflammatory glands of both patients with SjS and animal models and are positively correlated with the severity of glandular morbidity [50, 65, 66]. They are heterogeneous cell populations that are ubiquitous in all organs and serve as active participants in developmental processes and immune responses to maintain organic integrity and homeostasis [49, 50]. For example, macrophages remove cellular debris, apoptotic cells, senescent cells, invading pathogens, and immune complexes (ICs) by phagocytosis; thereby playing an important role in host defense, morphogenesis, and tissue homeostasis. In our study, Fc γ R-mediated phagocytosis and related signal pathways were associated with the genetic predisposition to SjS. Fc γ Rs are a group of key receptors involved in the engulfment of opsonized microorganisms and immune complexes [67]. Fc γ R signaling for phagocytosis involves the activation of serial signal cascades, including PI3K-Akt, phospholipase D, NF- κ B, JNK, and MAPK signaling pathway, to reorganize the cytoskeleton and induce an inflammatory response [68]. One putative pathogenic association between dysregulated Fc γ R-mediated phagocytosis and SjS is that inefficient clearance of invading pathogens in glands could activate epithelial and innate immune cells by other receptors, such as toll-like receptors; thereby inducing tissue injury and generating inflammatory milieu for autoimmunity [69] (Fig. 6). Another putative pathogenic association between dysregulated Fc γ R-mediated phagocytosis and SjS suggests the

inefficient clearance of ICs. Physically, ICs are efficiently eliminated by phagocytes to prevent their accumulation. However, in the pathological state, such as defective phagocytosis resulting from loss-of-function variants, ICs could be deposited on tissues and initiate several immune cascades; thus causing glandular and extra-glandular injuries [70] (Fig. 6). Several studies have indicated the presence of ICs, which can be formed by autoantigens and autoantibodies, in the salivary glands, blood vessels, glomeruli, and serum of patients with SjS [5, 71, 72]. Furthermore, endogenous defects in phagocytosis mediated by other receptors are associated with autoimmune diseases [73]. In addition to dysregulated phagocytosis, impaired phagolysosome function can also result in the accumulation of residual DAMPs or pathogen-associated molecular patterns and trigger local inflammation through pattern recognition receptors. In this study, a PLD3 p.T430I variant was found in a early-onset female patient that drastically impaired its enzymolytic activity for auto-DNA in the phagolysosome of macrophages (unpublished data). This could contribute to the initiation of SjS when triggered by a local overload of auto-DNA since PLD3 has been recently recognized as an important nuclease that limits auto-inflammation by eliminating DNA and RNA in the endolysosome [74]. In addition to their phagocytic ability, macrophages also participate in glandular morphogenesis, wound healing, regeneration, and disinfection by releasing biochemical factors such as trophic factors, coagulation factors, matrix metalloproteases, cytokines, and chemokines [75-79]. Therefore, loss-of-function variants could impair the above-mentioned process, thus interrupting glandular homeostasis. In contrast, macrophages with gain-of-function variants can produce dysregulated inflammatory cytokines, such as interleukin-1 and tumor necrosis factor, to facilitate autoimmunity and cause glandular and peripheral nerve injuries [80, 81]. Notably, several studies have shown that predisposing variants in Alzheimer's Disease are associated with genes expressed in macrophages [82, 83], which could partly explain why an increased risk of Alzheimer's Disease is found in patients with SjS and further supports the hypothesis of shared genetic predisposition between SjS and neuropathies [84]. However, the specific contribution of macrophages to early onset

SjS requires more comprehensive and in-depth investigation.

Motor proteins, including myosin, dynactin, kinesin, and axonemal dynein, were enriched in the patients with SjS. These proteins function within the cytoskeleton and are essential for various biological processes. Myosin is involved in the cytoskeleton reorganization, which is important for the development of the salivary gland, migration of progenitor and immune cells, cell adhesion, formation of immune synapse, axonal transport, and phagocytosis [48, 85-87]. The dynein/dynactin complex and kinesin transport cytoplasmic cargo, such as mitochondria, viral particles, synaptic vesicles, lysosomes, signaling endosomes, autophagosomes, and microtubules, in a retrograde and anterograde manner, respectively [88, 89]. Disruption of cargo transportation by defects in these motor proteins could result in a range of neurodevelopmental and neurodegenerative diseases, dysfunction of lacrimal acinar cells, disturbed innate and acquired immune response to viral infection, as well as enhancing the activity of NF- κ B; a well-known transcriptional factor involved in the pathogenesis of SjS [47, 90-93]. Furthermore, the disruption of dynein and dynactin can impair autophagy and cause defects in the salivary glands, muscles, and nervous system in a motor-independent manner [94]. Overall, the involvement of motor proteins in the genetic predisposition to SjS provides persuasive evidence for the hypothesis of a shared genetic predisposition between SjS and neuropathies. Interestingly, several genes encoding axonemal dyneins were identified in this study (Supplemental Table 7). Axonemal dyneins are foundational proteins of cilia that drive cilium beating, which is essential for various biological processes such as cerebrospinal fluid circulation, mucociliary clearance defense, and spermatozoa motion [95-97]. Therefore, primary ciliary dyskinesia caused by pathogenic variants of the axonemal dynein genes can manifest in various systems [98, 99]. However, this is the first time that axonemal dyneins have been associated with SjS. One putative cause is that ciliary dyskinesia impairs the mucociliary clearance defense, which increases the risk of respiratory infection that is associated with both seropositive and seronegative SjS [96, 98, 100]. Furthermore, axonemal dyneins may also have cilium-independent functions. For example, DNAH5 is required for sound sensation

in *Drosophila* larvae and is associated with myeloma [101, 102]. However, the role of axonemal dyneins in the pathogenesis of SjS requires further investigation.

Numerous predisposing genes, biological processes, and pathways were also identified in our study, such as female sex differentiation, cell adhesion, and PI3K-Akt/mTOR signaling, which deserve further investigation for their pathogenic roles in SjS. Genes affected by homozygous and potentially compound heterozygous variants, those frequently identified in different genetic predisposition studies (Supplemental Fig. 2), and their associated biological processes and pathways would be of considerable interest in future studies. For example, *RELN* was frequently affected in our patient cohort and has been identified as a predisposing gene for SjS in previous genome-wide association studies or WES-based analyses [9, 103].

The identification of genetic predisposition, especially highly enriched variants, provides the possibility to advance the diagnosis of SjS. For example, nine variants were identified in 24 of 50 patients with SjS compared to only two of the 100 individuals of the healthy volunteer cohort in our study, suggesting high sensitivity and specificity of these variants for SjS identification (Supplemental Table 21). Larger sample sizes and samples from contrasting diseases, such as systemic lupus erythematosus and rheumatoid arthritis, are required to further evaluate the diagnostic sensitivities and specificities of these candidate variants. However, the diagnostic utility of testing genetic predisposition, whether independently or combined with other phenotype-based measurements [104], would be a promising topic; not only because genetic testing is more stable than serological testing and less invasive than labial salivary gland biopsy but also because it could facilitate the identification of patients at an early stage, which is critical for improving management strategies [105].

The management of SjS is challenging. To date, the vast majority of drugs have failed to demonstrate clinical efficacy based on the outcomes of corresponding randomized placebo-controlled trials; therefore, no drugs have been officially licensed for the treatment of SjS [106]. However, when stratifying patients based on symptoms, the effects of drugs can be positive for specific subgroups of patients [107], and the heterogeneity of SjS pathogenesis is attributable to these findings and highlights the

importance of personalized treatment [106]. Our study provides a foundation for developing more personalized therapeutics based on the genetic predisposition that reflects the underlying pathogenesis and for determining the optimal treatment options for subgroups of patients, or even at the individual level for those affected by specific genetic defects. For example, glandular dysfunction and systemic inflammation of patients with *AQP1* variants could be resolved via aquaporin gene therapy [108]; the salivary glands of patients with ECM deficiency could be regenerated by the injection of decellularized ECM hydrogel [109]; and the glandular dysfunction and autoimmunity of patients with *TSC2* variants could be alleviated by seletalisib, a phosphatidylinositol 3-kinase inhibitor [110].

This study was based on an East Asian population, and the genetic predisposition to SjS could differ markedly according to ancestry [9]. Therefore, future studies involving different populations are needed to enable a more comprehensive understanding of the genetic predisposition to SjS. Furthermore, with the development of sequencing technologies and data analysis methodologies, a more accurate analysis of SNVs, indels, copy number variants, and epigenetic changes in whole genomes may further advance our knowledge of genetic predisposition to SjS.

In conclusion, our study provides an extensive framework for understanding the genetic predisposition to SjS based on WES and rare variant analysis and facilitates further focused and in-depth investigation of the pathogenetic mechanisms of specific genes, biological processes, and pathways; thus contributing to the pathophysiology, diagnostics, and therapeutics of SjS.

Acknowledgement

We thank Drs Zhibin Li, Jinxiu Xuan, Yan He, Hongyan Qian, Xinwei Zhang, Yan Li, and Minjie Zhang of the First Affiliated Hospital of Xiamen University for their kind supports to this work.

Funding

This work was supported by the National Natural Science Foundation of China

[grant numbers 82171779] and the Xiamen Municipal Bureau of Science and Technology [grant number 2022XMSLCYX01].

Declarations of interest

None

References

- [1] M. Beydon, S. McCoy, Y. Nguyen, T. Sumida, X. Mariette, R. Seror, Epidemiology of Sjogren syndrome, *Nat Rev Rheumatol* 20 (2024) 158-169.
- [2] X. Mariette, L. A. Criswell, Primary Sjogren's Syndrome, *N Engl J Med* 378 (2018) 931-9.
- [3] E. Theander, G. Henriksson, O. Ljungberg, T. Mandl, R. Manthorpe, L. T. Jacobsson, Lymphoma and other malignancies in primary Sjogren's syndrome: a cohort study on cancer incidence and lymphoma predictors, *Ann Rheum Dis* 65 (2006) 796-803.
- [4] G. E. Thorlacius, A. Bjork, M. Wahren-Herlenius, Genetics and epigenetics of primary Sjogren syndrome: implications for future therapies, *Nat Rev Rheumatol* 19 (2023) 288-306.
- [5] P. Brito-Zeron, C. Baldini, H. Bootsma, S. J. Bowman, R. Jonsson, X. Mariette, K. Sivils, E. Theander, A. Tzioufas, M. Ramos-Casals, Sjogren syndrome, *Nat Rev Dis Primers* 2 (2016) 16047.
- [6] G. M. Verstappen, S. Pringle, H. Bootsma, F. G. M. Kroese, Epithelial-immune cell interplay in primary Sjogren syndrome salivary gland pathogenesis, *Nat Rev Rheumatol* 17 (2021) 333-48.
- [7] Y. Li, K. Zhang, H. Chen, F. Sun, J. Xu, Z. Wu, P. Li, L. Zhang, Y. Du, H. Luan et al., A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjogren's syndrome at 7q11.23, *Nat Genet* 45 (2013) 1361-5.
- [8] I. W. Song, H. C. Chen, Y. F. Lin, J. H. Yang, C. C. Chang, C. T. Chou, M. M. Lee, Y. C. Chou, C. H. Chen, Y. T. Chen et al., Identification of susceptibility gene associated with female primary Sjogren's syndrome in Han Chinese by genome-wide association study, *Hum Genet* 135 (2016) 1287-94.
- [9] K. E. Taylor, Q. Wong, D. M. Levine, C. McHugh, C. Laurie, K. Doheny, M. Y. Lam, A. N. Baer, S. Challacombe, H. Lanfranchi et al., Genome-Wide Association Analysis Reveals Genetic Heterogeneity of Sjogren's Syndrome According to Ancestry, *Arthritis Rheumatol* 69 (2017) 1294-305.
- [10] C. J. Lessard, H. Li, I. Adrianto, J. A. Ice, A. Rasmussen, K. M. Grundahl, J. A. Kelly, M. G. Dozmorov, C. Miceli-Richard, S. Bowman et al., Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjogren's syndrome, *Nat Genet* 45 (2013) 1284-92.
- [11] B. Khatri, K. L. Tessneer, A. Rasmussen, F. Aghakhanian, T. R. Reksten, A. Adler, I. Alevizos, J. M. Anaya, L. A. Aqrabi, E. Baecklund et al., Genome-wide association study identifies Sjogren's risk loci with functional implications in immune and glandular cells, *Nat Commun* 13 (2022) 4287.
- [12] E. T. Cirulli, D. B. Goldstein, Uncovering the roles of rare variants in common disease

- through whole-genome sequencing, *Nat Rev Genet* 11 (2010) 415-25.
- [13] L. Bomba, K. Walter, N. Soranzo, The impact of rare and low-frequency genetic variants in common disease, *Genome Biol* 18 (2017) 77.
- [14] A. S. Johar, C. Mastronardi, A. Rojas-Villarraga, H. R. Patel, A. Chuah, K. Peng, A. Higgins, P. Milburn, S. Palmer, M. F. Silva-Lara et al., Novel and rare functional genomic variants in multiple autoimmune syndrome and Sjogren's syndrome, *J Transl Med* 13 (2015) 173.
- [15] Y. Wang, S. Chen, J. Chen, X. Xie, S. Gao, C. Zhang, S. Zhou, J. Wang, R. Mai, Q. Lin et al., Germline genetic patterns underlying familial rheumatoid arthritis, systemic lupus erythematosus and primary Sjogren's syndrome highlight T cell-initiated autoimmunity, *Ann Rheum Dis* 79 (2020) 268-75.
- [16] X. Qi, X. Q. Wang, L. Jin, L. X. Gao, H. F. Guo, Uncovering potential single nucleotide polymorphisms, copy number variations and related signaling pathways in primary Sjogren's syndrome, *Bioengineered* 12 (2021) 9313-31.
- [17] J. M. Abuin, J. C. Pichel, T. F. Pena, J. Amigo, BigBWA: approaching the Burrows-Wheeler aligner to Big Data technologies, *Bioinformatics* 31 (2015) 4003-5.
- [18] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools, *Bioinformatics* 25 (2009) 2078-9.
- [19] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, *Nucleic Acids Res* 38 (2010) e164.
- [20] S. W. Kong, I. H. Lee, X. Liu, J. N. Hirschhorn, K. D. Mandl, Measuring coverage and accuracy of whole-exome sequencing in clinical context, *Genet Med* 20 (2018) 1617-26.
- [21] D. Mandelker, R. J. Schmidt, A. Ankala, K. McDonald Gibson, M. Bowser, H. Sharma, E. Duffy, M. Hegde, A. Santani, M. Lebo et al., Navigating highly homologous genes in a molecular diagnostic setting: a resource for clinical next-generation sequencing, *Genet Med* 18 (2016) 1282-9.
- [22] A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, J. O. Korbel, J. L. Marchini, S. McCarthy, G. A. McVean, G. R. Abecasis, A global reference for human genetic variation, *Nature* 526 (2015) 68-74.
- [23] S. Gudmundsson, M. Singer-Berk, N. A. Watts, W. Phu, J. K. Goodrich, M. Solomonson, H. L. Rehm, D. G. MacArthur, A. O'Donnell-Luria, Variant interpretation using population databases: Lessons from gnomAD, *Hum Mutat* 43 (2022) 1012-30.
- [24] Y. Cao, L. Li, M. Xu, Z. Feng, X. Sun, J. Lu, Y. Xu, P. Du, T. Wang, R. Hu et al., The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals, *Cell Res* 30 (2020) 717-31.
- [25] K. S. Jung, K. W. Hong, H. Y. Jo, J. Choi, H. J. Ban, S. B. Cho, M. Chung, KRGDB: the large-scale variant database of 1722 Koreans based on whole genome sequencing, *Database (Oxford)* 2020 (2020).
- [26] S. Tadaka, E. Hishinuma, S. Komaki, I. N. Motoike, J. Kawashima, D. Saigusa, J. Inoue, J. Takayama, Y. Okamura, Y. Aoki et al., jMorp updates in 2020: large enhancement of multi-omics data resources on the general Japanese population, *Nucleic Acids Res* 49 (2021) D536-D44.
- [27] G. Dennis, Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, R. A. Lempicki, DAVID: Database for Annotation, Visualization, and Integrated Discovery, *Genome Biol* 4

- (2003) P3.
- [28] K. Breuer, A. K. Foroushani, M. R. Laird, C. Chen, A. Sribnaia, R. Lo, G. L. Winsor, R. E. Hancock, F. S. Brinkman, D. J. Lynn, InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation, *Nucleic Acids Res* 41 (2013) D1228-33.
- [29] M. E. Yayla, Z. Karaman, M. Torgutalp, A. B. Kelesoglu Dincer, E. G. Aydemir Guloksuz, S. Sezer, D. Sahin Eroglu, M. L. Yuksel, T. M. Turgay, G. Kinikli et al., Early onset primary Sjogren syndrome, clinical and laboratory characteristics, *Clin Rheumatol* 39 (2020) 2689-96.
- [30] G. Fulvio, G. La Rocca, L. G. Chatzis, F. Ferro, I. C. Navarro Garcia, G. Cafaro, A. V. Goules, E. Bartoloni, C. Baldini, Impact of gender and age at onset on Sjogren's syndrome presentation and outcome: state of the art, *Clin Exp Rheumatol* 41 (2023) 2547-54.
- [31] T. Seeliger, E. Kramer, F. F. Konen, N. Zehrfeld, S. Beider, N. K. Prenzler, V. Godecke, T. Witte, T. Skripuletz, D. Ernst, Sjogren's syndrome with and without neurological involvement, *J Neurol* 270 (2023) 2987-96.
- [32] J. Zheng, J. Zhang, Y. Jin, Y. Wang, L. Xu, H. Zheng, H. Jiang, C. Yuan, Characteristics of primary Sjogren's syndrome-associated peripheral nervous system lesions, *J Neurol* 270 (2023) 5527-35.
- [33] M. Jaskolska, M. Chylinska, A. Masiak, M. Sieminski, M. Zietkiewicz, Z. Czuszynska, Z. Smolenska, Z. Zdrojewski, Neuro-Sjogren: Uncommon or underestimated problem?, *Brain Behav* 10 (2020) e01665.
- [34] T. Gono, Y. Kawaguchi, Y. Katsumata, K. Takagi, A. Tochimoto, S. Baba, Y. Okamoto, Y. Ota, H. Yamanaka, Clinical manifestations of neurological involvement in primary Sjogren's syndrome, *Clin Rheumatol* 30 (2011) 485-90.
- [35] S. J. Rogers, C. S. Williams, G. C. Roman, Myelopathy in Sjogren's syndrome: role of nonsteroidal immunosuppressants, *Drugs* 64 (2004) 123-32.
- [36] S. Saadoun, M. C. Papadopoulos, M. Hara-Chikuma, A. S. Verkman, Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption, *Nature* 434 (2005) 786-92.
- [37] C. Delporte, M. Soyfoo, Aquaporins: Unexpected actors in autoimmune diseases, *Autoimmun Rev* 21 (2022) 103131.
- [38] C. D'Agostino, O. A. Elkashty, C. Chivasso, J. Perret, S. D. Tran, C. Delporte, Insight into Salivary Gland Aquaporins, *Cells* 9 (2020).
- [39] A. Alexander, J. Herz, L. Calvier, Reelin through the years: From brain development to inflammation, *Cell Rep* 42 (2023) 112669.
- [40] J. Amado-Azevedo, A. D. van Stalborch, E. T. Valent, K. Nawaz, J. van Bezu, E. C. Eringa, F. P. M. Hoevenaars, I. M. De Cuyper, P. L. Hordijk, V. W. M. van Hinsbergh et al., Depletion of Arg/Abl2 improves endothelial cell adhesion and prevents vascular leak during inflammation, *Angiogenesis* 24 (2021) 677-93.
- [41] Q. Chen, H. Li, Y. Liu, M. Zhao, Epigenetic Regulation of Immune and Inflammatory Responses in Rheumatoid Arthritis, *Front Immunol* 13 (2022) 881191.
- [42] M. Pechlivanidou, K. Xenou, D. Tzanetakos, E. Koutsos, C. Stergiou, E. Andreadou, K. Voumvourakis, S. Giannopoulos, C. Kilidireas, E. Tuzun et al., Potential Role of Antibodies against Aquaporin-1 in Patients with Central Nervous System Demyelination, *Int J Mol Sci* 24 (2023).
- [43] Y. Hua, X. Ying, Y. Qian, H. Liu, Y. Lan, A. Xie, X. Zhu, Physiological and pathological

- impact of AQP1 knockout in mice, *Biosci Rep* 39 (2019).
- [44] M. C. Papadopoulos, A. S. Verkman, Aquaporin water channels in the nervous system, *Nat Rev Neurosci* 14 (2013) 265-77.
- [45] F. A. Jacobsen, A. N. Scherer, J. Mouritsen, H. Bragadottir, B. Thomas Backstrom, S. Sardar, D. Holmberg, A. J. Koleske, A. Andersson, A Role for the Non-Receptor Tyrosine Kinase Abl2/Arg in Experimental Neuroinflammation, *J Neuroimmune Pharmacol* 13 (2018) 265-76.
- [46] J. V. Goodman, A. Bonni, Regulation of neuronal connectivity in the mammalian brain by chromatin remodeling, *Curr Opin Neurobiol* 59 (2019) 59-68.
- [47] P. Guedes-Dias, E. L. F. Holzbaur, Axonal transport: Driving synaptic function, *Science* 366 (2019).
- [48] N. B. Martin-Cofreces, F. Sanchez-Madrid, Sailing to and Docking at the Immune Synapse: Role of Tubulin Dynamics and Molecular Motors, *Front Immunol* 9 (2018) 1174.
- [49] N. Cox, M. Pokrovskii, R. Vicario, F. Geissmann, Origins, Biology, and Diseases of Tissue Macrophages, *Annu Rev Immunol* 39 (2021) 313-44.
- [50] D. Zhou, N. A. McNamara, Macrophages: important players in primary Sjogren's syndrome?, *Expert Rev Clin Immunol* 10 (2014) 513-20.
- [51] G. B. Proctor, The physiology of salivary secretion, *Periodontol* 2000 70 (2016) 11-25.
- [52] D. A. Dartt, Neural regulation of lacrimal gland secretory processes: relevance in dry eye diseases, *Prog Retin Eye Res* 28 (2009) 155-77.
- [53] E. M. Sternberg, Neuroendocrine regulation of autoimmune/inflammatory disease, *J Endocrinol* 169 (2001) 429-35.
- [54] K. Jin, T. Imada, R. Hisamura, M. Ito, H. Toriumi, K. F. Tanaka, S. Nakamura, K. Tsubota, Identification of Lacrimal Gland Postganglionic Innervation and Its Regulation of Tear Secretion, *Am J Pathol* 190 (2020) 1068-79.
- [55] V. N. Patel, M. P. Hoffman, Salivary gland development: a template for regeneration, *Semin Cell Dev Biol* 25-26 (2014) 52-60.
- [56] G. B. Proctor, G. H. Carpenter, Regulation of salivary gland function by autonomic nerves, *Auton Neurosci* 133 (2007) 3-18.
- [57] F. Wirth, A. Lubosch, S. Hamelmann, I. A. Nakchbandi, Fibronectin and Its Receptors in Hematopoiesis, *Cells* 9 (2020).
- [58] R. Jayadev, M. Morais, J. M. Ellingford, S. Srinivasan, R. W. Naylor, C. Lawless, A. S. Li, J. F. Ingham, E. Hastie, Q. Chi et al., A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations, *Sci Adv* 8 (2022) eabn2265.
- [59] M. Hajiabbas, C. D'Agostino, J. Siminska-Stanny, S. D. Tran, A. Shavandi, C. Delporte, Bioengineering in salivary gland regeneration, *J Biomed Sci* 29 (2022) 35.
- [60] Y. T. Konttinen, P. Porola, L. Konttinen, M. Laine, P. Poduval, Immunohistopathology of Sjogren's syndrome, *Autoimmun Rev* 6 (2006) 16-20.
- [61] C. Molina, C. Alliende, S. Aguilera, Y. J. Kwon, L. Leyton, B. Martinez, C. Leyton, P. Perez, M. J. Gonzalez, Basal lamina disorganisation of the acini and ducts of labial salivary glands from patients with Sjogren's syndrome: association with mononuclear cell infiltration, *Ann Rheum Dis* 65 (2006) 178-83.
- [62] H. Tomlin, A. M. Piccinini, A complex interplay between the extracellular matrix and the innate immune response to microbial pathogens, *Immunology* 155 (2018) 186-201.
- [63] J. Kiripolsky, E. M. Kasperek, C. Zhu, Q. Z. Li, J. Wang, G. Yu, J. M. Kramer,

- Immune-Intrinsic Myd88 Directs the Production of Antibodies With Specificity for Extracellular Matrix Components in Primary Sjogren's Syndrome, *Front Immunol* 12 (2021) 692216.
- [64] C. Anquetil, E. Hachulla, F. Machuron, X. Mariette, V. Le Guern, O. Vittecoq, E. Dernis, C. Larroche, J. J. Dubost, A. Perdriger et al., Is early-onset primary Sjogren's syndrome a worse prognosis form of the disease?, *Rheumatology (Oxford)* 58 (2019) 1163-7.
- [65] M. N. Manoussakis, S. Boiu, P. Korkolopoulou, E. K. Kapsogeorgou, N. Kavantzas, P. Ziakas, E. Patsouris, H. M. Moutsopoulos, Rates of infiltration by macrophages and dendritic cells and expression of interleukin-18 and interleukin-12 in the chronic inflammatory lesions of Sjogren's syndrome: correlation with certain features of immune hyperactivity and factors associated with high risk of lymphoma development, *Arthritis Rheum* 56 (2007) 3977-88.
- [66] M. I. Christodoulou, E. K. Kapsogeorgou, H. M. Moutsopoulos, Characteristics of the minor salivary gland infiltrates in Sjogren's syndrome, *J Autoimmun* 34 (2010) 400-7.
- [67] S. Ben Mkaddem, M. Benhamou, R. C. Monteiro, Understanding Fc Receptor Involvement in Inflammatory Diseases: From Mechanisms to New Therapeutic Tools, *Front Immunol* 10 (2019) 811.
- [68] E. Uribe-Querol, C. Rosales, Phagocytosis: Our Current Understanding of a Universal Biological Process, *Front Immunol* 11 (2020) 1066.
- [69] M. Z. Tay, K. Wiehe, J. Pollara, Antibody-Dependent Cellular Phagocytosis in Antiviral Immune Responses, *Front Immunol* 10 (2019) 332.
- [70] E. Chalayer, B. Gramont, F. Zekre, R. Goguyer-Deschaumes, L. Waeckel, L. Grange, S. Paul, A. W. Chung, M. Killian, Fc receptors gone wrong: A comprehensive review of their roles in autoimmune and inflammatory diseases, *Autoimmun Rev* 21 (2022) 103016.
- [71] T. J. Lawley, H. M. Moutsopoulos, S. I. Katz, A. N. Theofilopoulos, T. M. Chused, M. M. Frank, Demonstration of circulating immune complexes in Sjogren's syndrome, *J Immunol* 123 (1979) 1382-7.
- [72] K. Yamane, H. Nakamura, M. Hamasaki, Y. Minei, N. Aibara, T. Shimizu, A. Kawakami, M. Nakashima, N. Kuroda, K. Ohyama, Immune complexome analysis reveals the presence of immune complexes and identifies disease-specific immune complex antigens in saliva samples from patients with Sjogren's syndrome, *Clin Exp Immunol* 204 (2021) 212-20.
- [73] S. W. Tas, P. Quartier, M. Botto, L. Fossati-Jimack, Macrophages from patients with SLE and rheumatoid arthritis have defective adhesion in vitro, while only SLE macrophages have impaired uptake of apoptotic cells, *Ann Rheum Dis* 65 (2006) 216-21.
- [74] A. L. Gavin, D. Huang, T. R. Blane, T. C. Thinnis, Y. Murakami, R. Fukui, K. Miyake, D. Nemazee, Cleavage of DNA and RNA by PLD3 and PLD4 limits autoinflammatory triggering by multiple sensors, *Nat Commun* 12 (2021) 5874.
- [75] G. A. Sathi, M. Farahat, E. S. Hara, H. Taketa, H. Nagatsuka, T. Kuboki, T. Matsumoto, MCSF orchestrates branching morphogenesis in developing submandibular gland tissue, *J Cell Sci* 130 (2017) 1559-69.
- [76] A. Nimmerjahn, F. Kirchhoff, F. Helmchen, Resting microglial cells are highly dynamic surveillants of brain parenchyma in vivo, *Science* 308 (2005) 1314-8.
- [77] D. Zanolini, S. Merlin, M. Feola, G. Ranaldo, A. Amoroso, G. Gaidano, M. Zaffaroni, A. Ferrero, S. Brunelleschi, G. Valente et al., Extrahepatic sources of factor VIII potentially contribute to the coagulation cascade correcting the bleeding phenotype of mice with

- hemophilia A, *Haematologica* 100 (2015) 881-92.
- [78] G. A. Finlay, L. R. O'Driscoll, K. J. Russell, E. M. D'Arcy, J. B. Masterson, M. X. FitzGerald, C. M. O'Connor, Matrix metalloproteinase expression and production by alveolar macrophages in emphysema, *Am J Respir Crit Care Med* 156 (1997) 240-7.
- [79] B. A. Shook, R. R. Wasko, G. C. Rivera-Gonzalez, E. Salazar-Gatzimas, F. Lopez-Giraldez, B. C. Dash, A. R. Munoz-Rojas, K. D. Aultman, R. K. Zwick, V. Lei et al., Myofibroblast proliferation and heterogeneity are supported by macrophages during skin repair, *Science* 362 (2018).
- [80] Y. T. Chen, S. Lazarev, A. F. Bahrami, L. B. Noble, F. Y. Chen, D. Zhou, M. Gallup, M. Yadav, N. A. McNamara, Interleukin-1 receptor mediates the interplay between CD4+ T cells and ocular resident cells to promote keratinizing squamous metaplasia in Sjogren's syndrome, *Lab Invest* 92 (2012) 556-70.
- [81] Y. Wang, L. Guo, X. Yin, E. C. McCarthy, M. I. Cheng, A. T. Hoang, H. C. Chen, A. Y. Patel, D. Allard Trout, E. Xu et al., Pathogenic TNF-alpha drives peripheral nerve inflammation in an Aire-deficient model of autoimmunity, *Proc Natl Acad Sci U S A* 119 (2022).
- [82] A. McQuade, M. Blurton-Jones, Microglia in Alzheimer's Disease: Exploring How Genetics and Phenotype Influence Risk, *J Mol Biol* 431 (2019) 1805-17.
- [83] A. Nott, I. R. Holtman, N. G. Coufal, J. C. M. Schlachetzki, M. Yu, R. Hu, C. Z. Han, M. Pena, J. Xiao, Y. Wu et al., Brain cell type-specific enhancer-promoter interactome maps and disease-risk association, *Science* 366 (2019) 1134-9.
- [84] P. C. Liliang, C. L. Liang, K. Lu, S. N. Yang, M. T. Hsieh, Y. C. Tai, K. W. Wang, Population-based study suggests an increased risk of Alzheimer's disease in Sjogren's syndrome, *Clin Rheumatol* 37 (2018) 935-41.
- [85] J. J. Nirschl, A. E. Ghiretti, E. L. F. Holzbaur, The impact of cytoskeletal organization on the local regulation of neuronal transport, *Nat Rev Neurosci* 18 (2017) 585-97.
- [86] T. P. Le, S. Chung, Regulation of apical constriction via microtubule- and Rab11-dependent apical transport during tissue invagination, *Mol Biol Cell* 32 (2021) 1033-47.
- [87] P. P. Ostrowski, S. Grinstein, S. A. Freeman, Diffusion Barriers, Mechanical Forces, and the Biophysics of Phagocytosis, *Dev Cell* 38 (2016) 135-46.
- [88] N. Hirokawa, Y. Noda, Y. Tanaka, S. Niwa, Kinesin superfamily motor proteins and intracellular transport, *Nat Rev Mol Cell Biol* 10 (2009) 682-96.
- [89] J. T. Canty, R. Tan, E. Kusakci, J. Fernandes, A. Yildiz, Structure and Mechanics of Dynein Motors, *Annu Rev Biophys* 50 (2021) 549-74.
- [90] S. Tati, L. Alisaraie, Recruitment of dynein and kinesin to viral particles, *FASEB J* 36 (2022) e22311.
- [91] L. Qian, J. Xie, C. M. Rose, E. Sou, H. Zeng, S. F. Hamm-Alvarez, A. K. Mircheff, Altered traffic to the lysosome in an ex vivo lacrimal acinar cell model for chronic muscarinic receptor stimulation, *Exp Eye Res* 79 (2004) 665-75.
- [92] S. H. Berth, T. E. Lloyd, Disruption of axonal transport in neurodegeneration, *J Clin Invest* 133 (2023).
- [93] R. L. Lamason, A. Kupfer, J. L. Pomerantz, The dynamic distribution of CARD11 at the immunological synapse is regulated by the inhibitory kinesin GAKIN, *Mol Cell* 40 (2010) 798-809.
- [94] Y. Batlevi, D. N. Martin, U. B. Pandey, C. R. Simon, C. M. Powers, J. P. Taylor, E. H.

- Baehrecke, Dynein light chain 1 is required for autophagy, protein clearance, and cell death in *Drosophila*, *Proc Natl Acad Sci U S A* 107 (2010) 742-7.
- [95] R. Faubel, C. Westendorf, E. Bodenschatz, G. Eichele, Cilia-based flow network in the brain ventricles, *Science* 353 (2016) 176-8.
- [96] X. M. Bustamante-Marin, L. E. Ostrowski, Cilia and Mucociliary Clearance, *Cold Spring Harb Perspect Biol* 9 (2017).
- [97] A. Sironen, A. Shoemark, M. Patel, M. R. Loebinger, H. M. Mitchison, Sperm defects in primary ciliary dyskinesia and related causes of male infertility, *Cell Mol Life Sci* 77 (2020) 2029-48.
- [98] M. Legendre, L. E. Zaragosi, H. M. Mitchison, Motile cilia and airway disease, *Semin Cell Dev Biol* 110 (2021) 19-33.
- [99] C. Wyart, M. Carbo-Tano, Y. Cantaut-Belarif, A. Orts-Del'Immagine, U. L. Bohm, Cerebrospinal fluid-contacting neurons: multimodal cells with diverse roles in the CNS, *Nat Rev Neurosci* 24 (2023) 540-56.
- [100] A. Bjork, J. Mofors, M. Wahren-Herlenius, Environmental factors in the pathogenesis of primary Sjogren's syndrome, *J Intern Med* 287 (2020) 475-92.
- [101] B. Li, S. Li, Z. Yan, Axonemal Dynein DNAH5 is Required for Sound Sensation in *Drosophila* Larvae, *Neurosci Bull* 37 (2021) 523-34.
- [102] B. A. Walker, C. P. Wardell, L. Melchor, S. Hulkki, N. E. Potter, D. C. Johnson, K. Fenwick, I. Kozarewa, D. Gonzalez, C. J. Lord et al., Intracloonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma, *Blood* 120 (2012) 1077-86.
- [103] Q. Shen, K. Lee, S. K. Han, H. J. Ahn, S. Kim, J. H. Lee, Variants at potential loci associated with Sjogren's syndrome in Koreans: A genetic association study, *Clin Immunol* 207 (2019) 79-86.
- [104] C. H. Shiboski, S. C. Shiboski, R. Seror, L. A. Criswell, M. Labetoulle, T. M. Lietman, A. Rasmussen, H. Scofield, C. Vitali, S. J. Bowman et al., 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjogren's syndrome: A consensus and data-driven methodology involving three international patient cohorts, *Ann Rheum Dis* 76 (2017) 9-16.
- [105] B. Wang, S. Chen, Q. Zheng, Y. Li, X. Zhang, J. Xuan, Y. Liu, G. Shi, Early diagnosis and treatment for Sjogren's syndrome: current challenges, redefined disease stages and future prospects, *J Autoimmun* 117 (2021) 102590.
- [106] S. Arends, G. M. Verstappen, L. de Wolff, S. Pringle, F. G. M. Kroese, A. Vissink, H. Bootsma, Why do drug treatments fail in Sjogren's disease? Considerations for treatment, trial design and interpretation of clinical efficacy, *Expert Rev Clin Immunol* 19 (2023) 1187-94.
- [107] J. R. Tarn, N. Howard-Tripp, D. W. Lendrem, X. Mariette, A. Saraux, V. Devauchelle-Pensec, R. Seror, A. J. Skelton, K. James, P. McMeekin et al., Symptom-based stratification of patients with primary Sjogren's syndrome: multi-dimensional characterisation of international observational cohorts and reanalyses of randomised clinical trials, *Lancet Rheumatol* 1 (2019) E85-E94.
- [108] Z. Lai, H. Yin, J. Cabrera-Perez, M. C. Guimaro, S. Afione, D. G. Michael, P. Glenton, A. Patel, W. D. Swaim, C. Zheng et al., Aquaporin gene therapy corrects Sjogren's syndrome phenotype in mice, *Proc Natl Acad Sci U S A* 113 (2016) 5694-9.

- [109] T. Wang, Q. Huang, Z. Rao, F. Liu, X. Su, X. Zhai, J. Ma, Y. Liang, D. Quan, G. Liao et al., Injectable decellularized extracellular matrix hydrogel promotes salivary gland regeneration via endogenous stem cell recruitment and suppression of fibrogenesis, *Acta Biomater* 169 (2023) 256-72.
- [110] S. Nayar, J. Campos, C. G. Smith, V. Iannizzotto, D. H. Gardner, S. Colafrancesco, E. Pippi, F. Kollert, K. J. Hunter, C. Brewer et al., Phosphatidylinositol 3-kinase delta pathway: a novel therapeutic target for Sjogren's syndrome, *Ann Rheum Dis* 78 (2019) 249-60.

Figure 1. Filtering pipeline for pathogenic rare variants. *Due to the provisional close of KRGDB, the variants with only one allele count were not examined for their minor allele frequencies in this database.

Figure 2. Validation of the variant filtering pipeline. (A) Numbers of alleles with different allele counts. (B) The candidate variants were rare in healthy volunteer cohort.

Figure 3. Top 20 enriched biological processes for different gene sets. (A)

Enriched biological processes for the genes enriched in patients with SjS, (B)

immune-related genes among the enriched genes, (C) early-onset-specific genes, and

(D) regular- and late-onset-specific genes.

Figure 4. Significantly enriched KEGG pathways ($P<0.05$) for different gene sets.

(A) Enriched pathways for the genes enriched in patients with SjS, (B)

immune-related genes among enriched genes, (C) early-onset-specific genes, and (D)

regular- and late-onset-specific genes.

Figure 5. Carrying a greater number of enriched variants could promote autoimmunity and thus increase the risk of systemic involvement in female patients. (A) Associations between the number of variants in an individual and serum IgG levels. (B) Associations between the number of variants in an individual and ESSDAI scores. (C) Associations between serum IgG levels and ESSDAI scores. IgG, immunoglobulin G; ESSDAI, European League Against Rheumatism Sjögren’s syndrome disease activity index.

Figure 6. Putative pathogenesis of Sjögren's Syndrome associated with genetic predisposition. ECM, extracellular matrix; ICs, immune complexes, TLR, toll-like receptor; DAMPs, damage-associated molecular patterns; PAMPs, pathogen-associated molecular patterns.

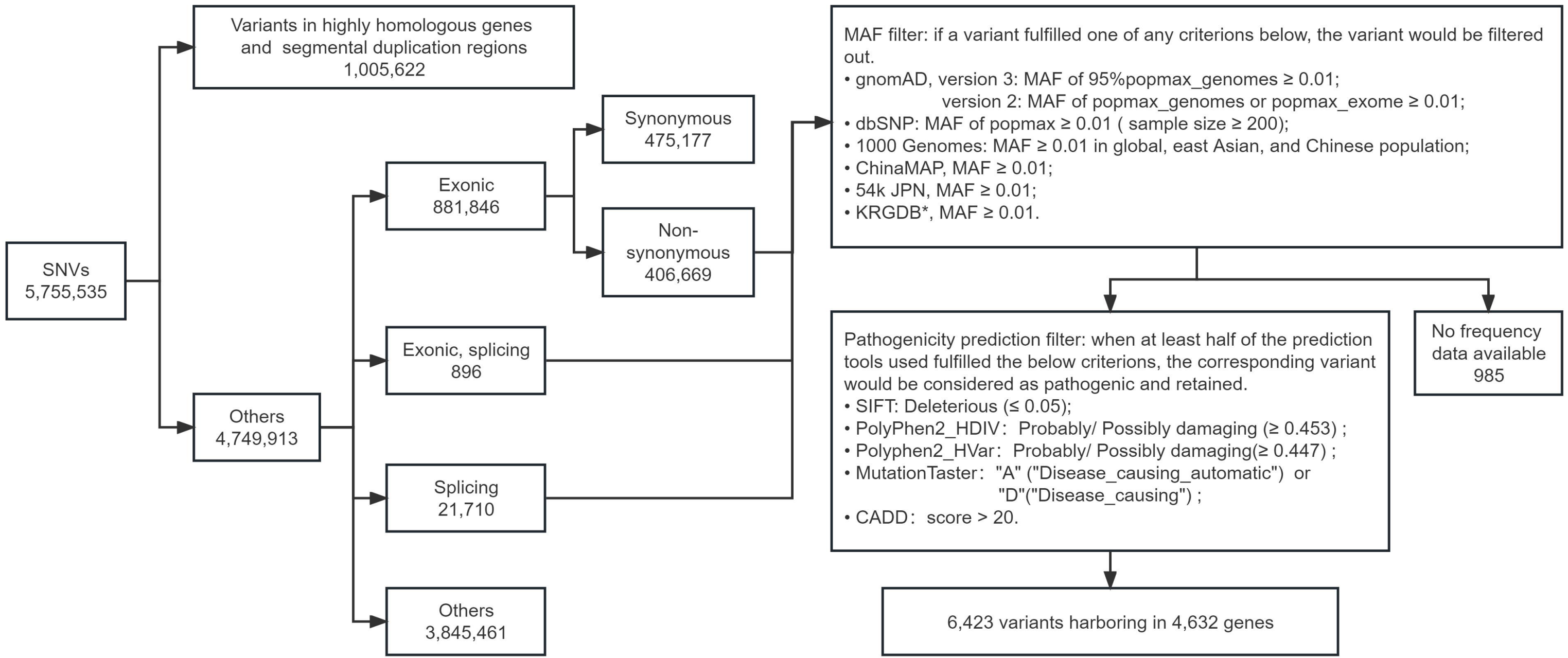
Table 1. Associations between the number of enriched variants in an individual and various clinical indices in female patients.

Clinical indexes	Case number	^a r/rho	<i>P</i> value
Onset age	46	-0.029	^b 0.848
Focus score	22	0.176	^b 0.434
Schirmer's test, left	21	0.029	^b 0.902
Schirmer's test, right	21	0.244	^b 0.287
Schirmer's test, mean	21	0.186	^b 0.419
ANA	20	0.002	^b 0.992
ACA (IgG)	23	0.141	^b 0.520
Anti-β2-GP1 Ab (IgG)	23	0.039	^b 0.860
Anti-SSA Ab	22	0.113	^b 0.618
Anti-Ro-52 Ab	22	0.268	^b 0.229
Anti-SSB Ab	22	0.027	^b 0.905
Anti-TPO Ab	17	0.255	^b 0.324
IgG	23	0.537	^c 0.008
C3	23	-0.032	^c 0.886
C4	23	-0.149	^c 0.498
RF	17	0.179	^b 0.491
ESSDAI	23	0.503	^b 0.014

^ar, Pearson's correlation coefficient; rho, Spearman's rank correlation coefficient.

^b*P* value was calculated with Spearman's rank correlation analysis.

^c*P* value was calculated with Pearson's correlation analysis.



SNVs
5,755,535

Variants in highly homologous genes
and segmental duplication regions
1,005,622

Others
4,749,913

Exonic
881,846

Exonic, splicing
896

Splicing
21,710

Others
3,845,461

Synonymous
475,177

Non-synonymous
406,669

MAF filter: if a variant fulfilled one of any criteria below, the variant would be filtered out.

- gnomAD, version 3: MAF of 95%popmax_genomes ≥ 0.01 ;
version 2: MAF of popmax_genomes or popmax_exome ≥ 0.01 ;
- dbSNP: MAF of popmax ≥ 0.01 (sample size ≥ 200);
- 1000 Genomes: MAF ≥ 0.01 in global, east Asian, and Chinese population;
- ChinaMAP, MAF ≥ 0.01 ;
- 54k JPN, MAF ≥ 0.01 ;
- KRGDB*, MAF ≥ 0.01 .

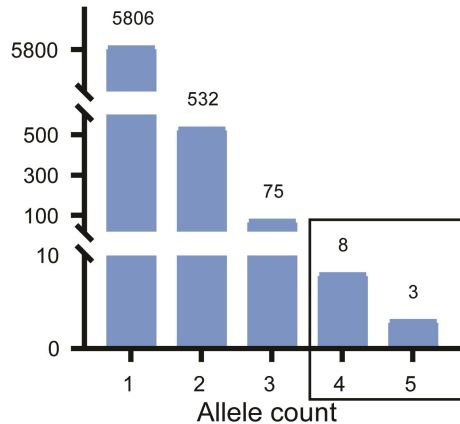
Pathogenicity prediction filter: when at least half of the prediction tools used fulfilled the below criteria, the corresponding variant would be considered as pathogenic and retained.

- SIFT: Deleterious (≤ 0.05);
- PolyPhen2_HDIV: Probably/ Possibly damaging (≥ 0.453) ;
- Polyphen2_HVar: Probably/ Possibly damaging(≥ 0.447) ;
- MutationTaster: "A" ("Disease_causing_automatic") or "D"("Disease_causing") ;
- CADD: score > 20 .

No frequency data available
985

6,423 variants harboring in 4,632 genes

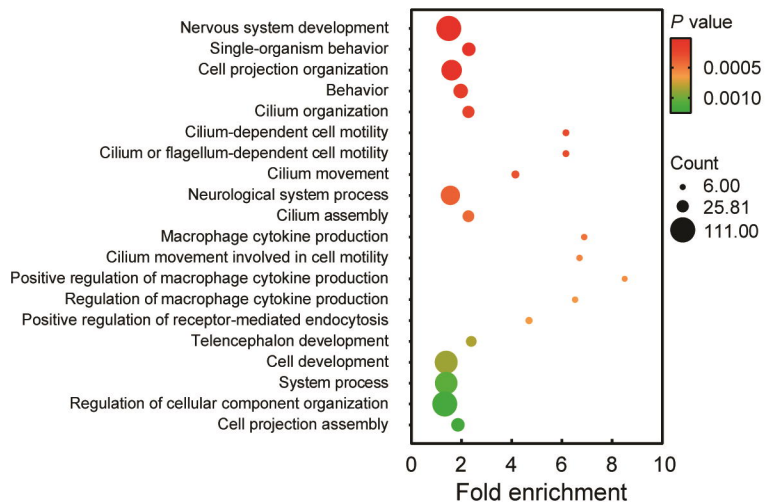
A



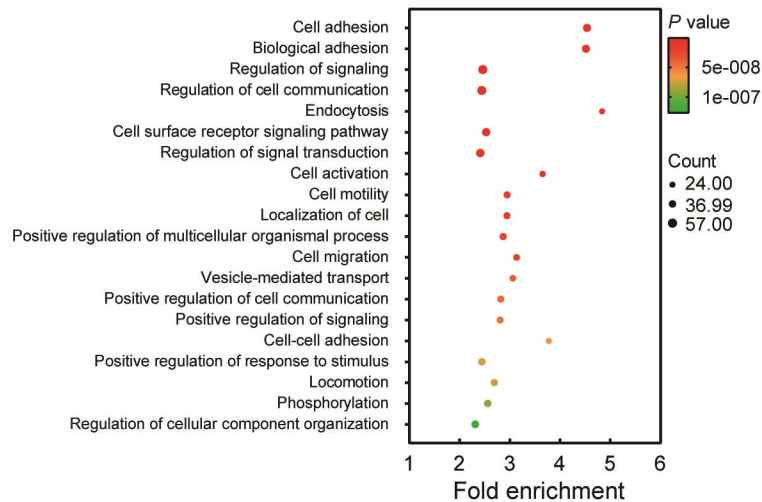
B

Allele count in pSS cohort	rs number	Gene name	General population cohort		
			Allele count	Total allele	Frequency
5	rs570527576	<i>AQP1</i>	0	200	0
5	rs190732812	<i>RRP12</i>	0	200	0
5	rs754990814	<i>TRADD</i>	5	300	0.016
4	rs193133475	<i>PPP4R4</i>	0	200	0
4	rs188038052	<i>PRDM2</i>	0	200	0
4	rs55655202	<i>ABL2</i>	1	200	0.005
4	rs181753558	<i>OR4K14</i>	0	200	0
4	rs534101761	<i>ALG5</i>	1	200	0.005
4	rs201048464	<i>CHD3</i>	4	600	0.007
4	rs114003896	<i>RELN</i>	0	200	0
4	rs192063486	<i>SH3D19</i>	0	200	0

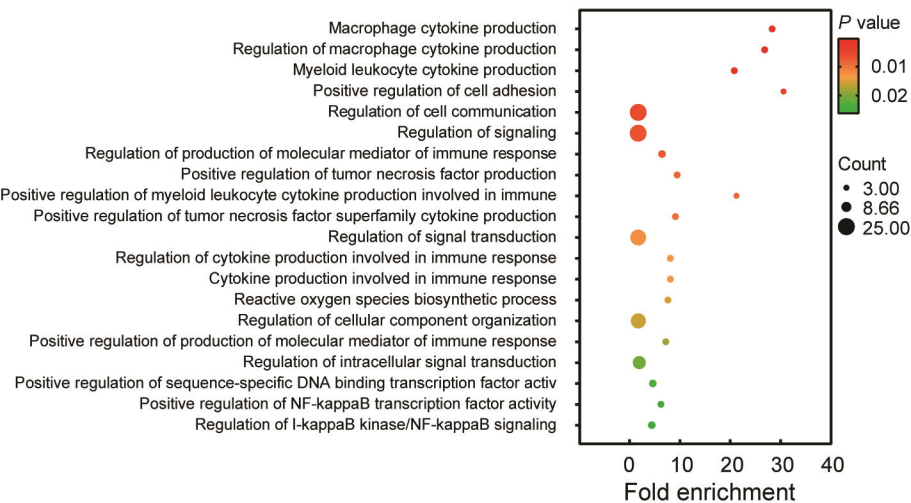
A Enriched genes



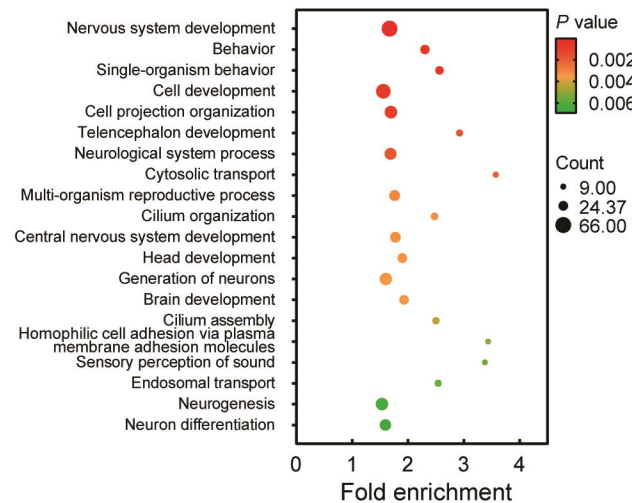
B Immune-related genes



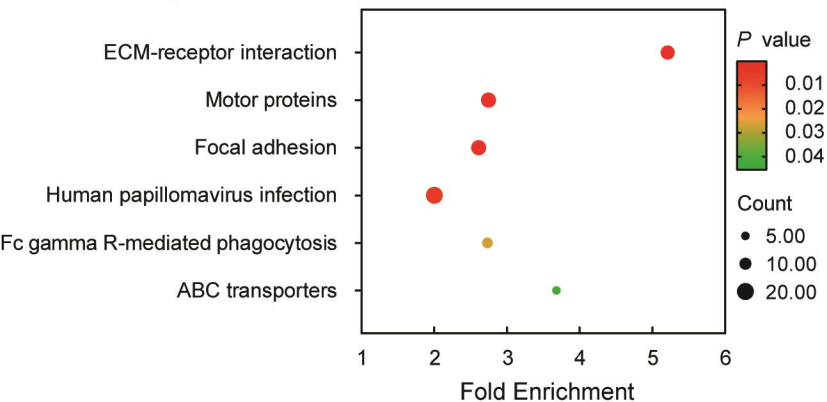
C Early-onset-specific genes



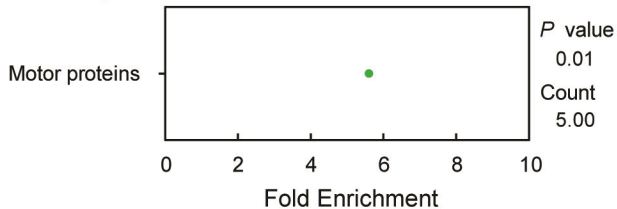
D Regular- and late-onset-specific genes



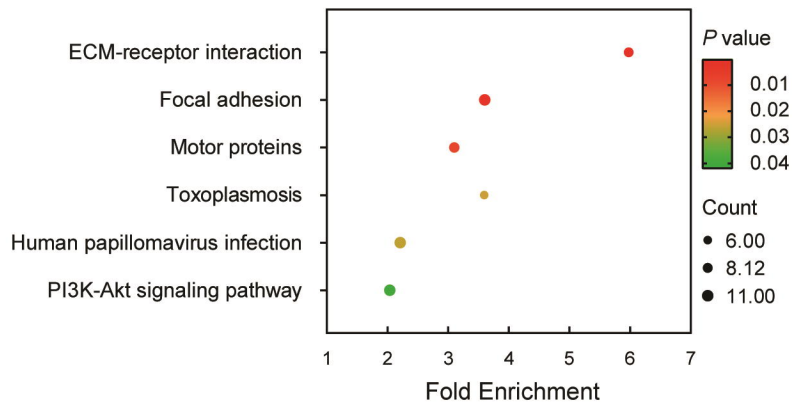
A Enriched genes



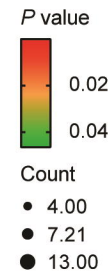
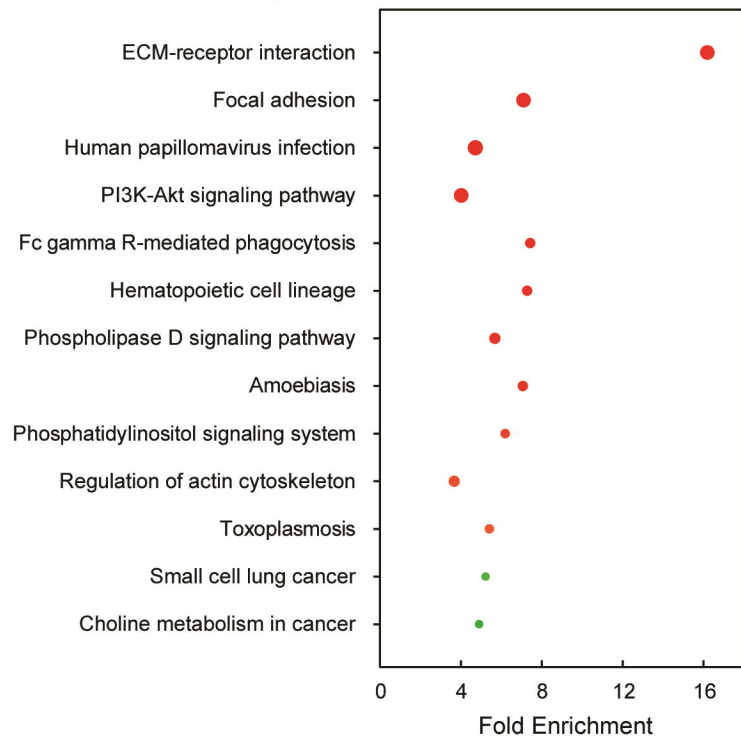
C Early-onset-specific genes



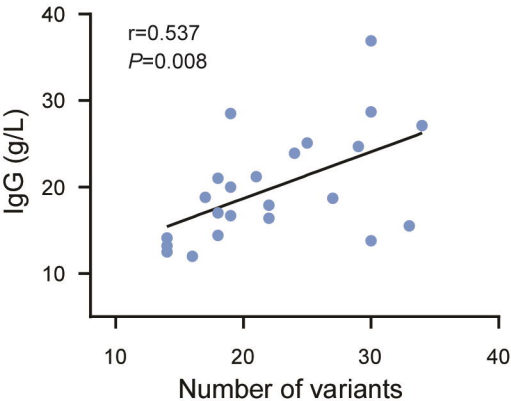
D Regular- and late-onset-specific genes



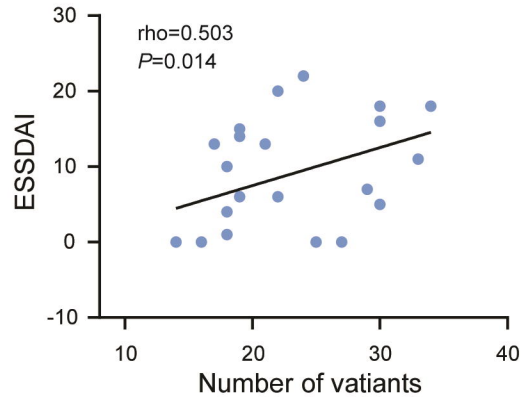
B Immune-related genes



A



B



C

