 doi:10.20944/preprints202205.0277.v1

Article

Exploring potential causal genes for immunoglobulin A nephropathy: A summary data-based Mendelian randomization and FUMA analysis

Qian Zhang [1#], Kang Zhang [2#], Yining Zhu [3], Jingyun Yang [4,5,*], Minmin Zhang [1,*]

[1] Huashan Hospital, Fudan University, Shanghai, China; qianzhang@fudan.edu.cn (Q.Z.); zhang_minmin@fudan.edu.cn (M.Z.)
[2] Wuhu Hospital of Traditional Chinese Medicine, Anhui College of Traditional Chinese Medicine, Wuhu, Anhui, China; zkzkzk462@126.com
[3] School of Mathematical Sciences, Fudan University, Yangpu District, Shanghai, China; 18300120139@fudan.edu.cn
[4] Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA; jingyun_yang@rush.edu
[5] Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA; jingyun_yang@rush.edu
# Qian Zhang and Kang Zhang contributed equally to this paper and share first authorship.
* Correspondence: jingyun_yang@rush.edu; Tel.: 1-312-563-9035 (J.Y.); zhang_minmin@fudan.edu.cn; Tel.: 86-21-52888133 (M.Z.)

Abstract:

Background: Immunoglobulin A nephropathy (IgAN) is a complex autoimmune disease, and the exact pathogenesis remains to be elucidated.

Methods: We conducted summary data-based Mendelian randomization (SMR) analysis and performed functional mapping and annotation using FUMA to explore genetic loci that are potentially involved in the pathogenies of IgAN. Both analyses used summarized data of a recent genome-wide association study (GWAS) on IgANs, which included 477,784 Europeans (15,587 cases and 462,197 controls) and 175,359 East Asians (71 cases and 175,288 controls). We performed separate SMR analysis using CAGE and GTEx eQTL data.

Results: Using the CAGE eQTL data, our SMR analysis identified 32 probes tagging 25 unique genes that were pleiotropically/potentially causally associated with IgAN, with the top three probes being ILMN_2150787 (tagging HLA-C, $P_{SMR}=2.10\times10^{-18}$), ILMN_1682717 (tagging IER3, $P_{SMR}=1.07\times10^{-16}$) and ILMN_1661439 (tagging FLOT1, $P_{SMR}=1.16\times10^{-14}$). Using GTEx eQTL data, our SMR analysis identified 24 probes tagging 24 unique genes, with the top three probes being ENSG00000271581.1 (tagging XXbac-BPG248L24.12, $P_{SMR}=1.44\times10^{-10}$), ENSG00000186470.9 (tagging BTN3A2, $P_{SMR}=2.28\times10^{-10}$), and ENSG00000224389.4 (tagging C4B, $P_{SMR}=1.23\times10^{-9}$). FUMA analysis identified 3 independent, significant and lead SNPs, 2 genomic risk loci and 39 genes.

Conclusion: We identified many genetic variants/loci that are potentially involved in the pathogenesis of IgAN.

Keywords: immunoglobulin A nephropathy; expression quantitative trait loci; summary data-based Mendelian randomization; genome-wide association study; functional mapping

 doi:10.20944/preprints202205.0277.v1

1. Introduction

Immunoglobulin A nephropathy (IgAN), first described by Jean Berger in 1968[1], is one of the most common forms of glomerulonephritis (GN) in the world[2]. It is characterized by the deposition of IgA immune complexes (specifically the IgA1 subclass) in the glomerular mesangium, leading to frequent episodes of hematuria and/or proteinuria[3]. Approximately 20-40% of IgAN patients will progress to end-stage renal disease (ESRD) within 10-20 years of diagnosis[4,5], causing a critical public health burden.

IgAN is a complex autoimmune disease with contributions from multiple factors, such as preference of salty food [6] and a family history of chronic glomerulonephritis[7]. Previous studies also indicated the important role of genetics in the etiology of IgAN. The prevalence of IgAN varies considerably across ethnicities, being the highest in Asians, moderate in Caucasians and the lowest in the African population[8], implying that both environmental and genetic factors are likely to be involved in the pathogenesis of IgAN. Genome-wide association studies (GWAS) identified several independent risk alleles for IgAN in East Asians and Europeans, such as genetic loci in *CFH*, *TNFSF13*, *ST6GAL1* and *ACCS*[9-13]. Recent research also discovered two distinct genome-wide significant loci in *C1GALT1* and *C1GALT1C1* in association with defective O-glycosylation of serum immunoglobulin A1 (IgA1), the key pathogenic defect in IgAN[14]. However, the exact pathogenic mechanisms underlying the observed associations in general and the genetic associations in particular remains to be elucidated.

Mendelian randomization (MR), using genetic variants as a proxy of randomized clinical trial, is a valuable epidemiological technique to assess pleotropic/potentially causal effect of an exposure (e.g., gene expression) on the outcome[15]. By using MR, confounding and reverse causation, which are commonly encountered in traditional association studies, can be greatly minimized. MR has been successful in identifying gene expression or DNA methylation sites showing pleiotropic association with various phenotypes, such as systemic lupus erythematosus (SLE), educational attainment, and severity of COVID-19[16-18].

In this study, we adopted the summary data-based MR (SMR) approach integrating summarized cis-expression quantitative trait loci (cis- eQTL) data and GWAS data for IgAN to prioritize genes that are pleiotropically/potentially causally associated with IgAN. Moreover, we performed functional mapping and annotation to further explore genetic variants and genomic loci in the pathogenesis of IgAN.

2. Methods

GWAS data for IgAN

The GWAS summarized data for IgAN were provided by a recent genome-wide association meta-analysis of IgAN[19]. The results were based on meta-analyses of IgAN using data from three population-based projects: The BioBank Japan (BBJ)[20], the UK Biobank[21], and GWAS summary results from FinnGen (https://www.finngen.fi/), with the sample size being 175,359 (71 cases and 175,288 controls), 344,365 (15,418 cases and 328,947 controls), and 133,419 (169 cases and 133,250 controls) for the three projects, respectively. As a result, the meta-analysis included 477,784 Europeans (15,587 cases and 462,197 controls) and 175,359 East Asians (71 cases and 175,288 controls).

For BBJ, genotyping was done with the Illumina HumanOmniExpressExome BeadChip or a combination of the HumanOmniExpress and the HumanExome BeadChip, with subsequent imputation performed using 1000 Genome Project Phase 3 version 5 and Japanese whole-genome sequencing data. GWAS analysis was conducted using a generalized linear mixed model, adjusting for age, age^2, sex, age×sex, age^2×sex and the first 20 principal components. For UK Biobank, genotyping was done using the Applied Biosystems UK BiLEVE Axiom Array or the Applied Biosystems UK Biobank Axiom Array, with

 doi:10.20944/preprints202205.0277.v1

subsequent imputation performed using a combination of the Haplotype Reference Consortium, UK10K and 1000 Genomes Project Phase 3. For FinnGen, genotyping was done with Illumina and Affymetrix chip arrays, with subsequent imputation performed using the population specific SISu v3 imputation reference panel of 3,775 whole genomes (https://finngen.gitbook.io/documentation/methods/genotype-imputation). GWAS was conducted using a generalized linear mixed model, adjusting for age, sex, the first 10 principal components and genotyping batch. The GWAS summarized data can be downloaded at http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90018001-GCST90019000/GCST90018866/harmonised/.

eQTL data for SMR analysis

In the SMR analysis, cis-eQTL genetic variants were used as the instrumental variables (IVs) for gene expression. We performed separate SMR analysis using eQTL data from two sources. Specifically, we used the CAGE eQTL summarized data for whole blood, which included 2,765 participants[22], and the V7 release of the GTEx eQTL summarized data for whole blood, which included 338 participants[23]. The eQTL data can be downloaded at https://cnsgenomics.com/data/SMR/#eQTLsummarydata.

FUMA analysis

We conducted a FUMA analysis to functionally map and annotate the genetic associations to better understand the genetic mechanisms underlying IgAN. FUMA uses GWAS association results as the input and integrates information from multiple resources. It provides a friendly on-line platform for easy implementation of post-GWAS analysis, such as functional annotation and gene prioritization[24]. FUMA provides two major functions: SNP2GENE for annotating SNPs regarding their biological functions and SNP-to-genes mapping; and GENE2FUNC for annotating the mapped genes in biological contexts. In SNP2GENE, we performed both positional mapping and eQTL mapping using GTEx v8 of whole blood and kidney. We adopted the default settings otherwise for both SNP2GENE (e.g., maximum P-value of lead SNPs being 5×10^{-8} and r^2 threshold for independent significant SNPs being 0.6) and GENE2FUNC (e.g., using FDR to correct for multiple testing in the gene-set enrichment analysis).

SMR analysis

We conducted the SMR analysis as implemented in the software SMR, with cis-eQTL as the IV, gene expression as the exposure, and IgAN as the outcome. Detailed information regarding the SMR method can be found in a previous publication[25]. We conducted the heterogeneity in dependent instruments (HEIDI) test to evaluate the existence of linkage in the observed association. $P_{HEIDI}<0.05$ indicates pleiotropy (i.e., the observed association could be due to two distinct genetic variants in high linkage disequilibrium with each other). We adopted the default settings in SMR and used false discovery rate (FDR) to adjust for multiple testing.

Data cleaning and statistical/bioinformatical analysis was performed using R version 4.1.2 (https://www.r-project.org/), PLINK 1.9 (https://www.cog-genomics.org/plink/1.9/), SMR (https://cnsgenomics.com/software/smr/), and FUMA (https://fuma.ctglab.nl/).

 doi:10.20944/preprints202205.0277.v1

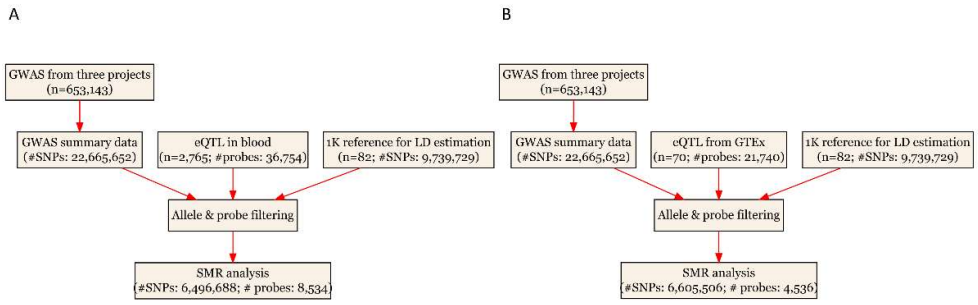

Figure 1: Flow chart for the SMR analysis.

A) SMR analysis using eQTL data from CAGE; and B) SMR analysis using eQTL data from GTEx

3. Results

Basic information of the summarized data

The GWAS summarized data included a total of 2,266,5652 SNPs. After checking of allele frequencies among the datasets and LD pruning, there were more than 6 million eligible SNPs in each SMR analysis. The CAGE eQTL has a much larger number of participants (2,765 vs. 70) and eligible probes (8,534 vs. 4,536) than that of the GTEx eQTL data. The detailed information was shown in **Table 1**.

Table 1. Basic information of the eQTL and GWAS data.

Data Source	Total # of participants	Number of eligible genetic variants or probes
eQTL data		
CAGE	2,765	8,534
GTEx	70	4,536
GWAS data for SMR analysis		
BBJ	71/175,288*	-
UKBB	15,418/328,947*	-
FinnGen	169/133,250*	-
Total	15,658 / 637,485*	CAGE: 6,496,688; GTEx: 6,605,506
GWAS data for FUMA analysis	15,658 /637,485*	22,665,652

*Data were represented as case/control.

GWAS: genome-wide association studies; QTL, quantitative trait loci; BBJ, BioBank Japan; UKBB, UK Biobank

Pleiotropic association with IgAN

 doi:10.20944/preprints202205.0277.v1

Using the CAGE eQTL data, our SMR analysis identified 32 probes tagging 25 unique genes that were pleiotropically/potentially causally associated with IgAN, with the top three probes being ILMN_2150787 (tagging *HLA-C*, $P_{SMR}=2.10\times10^{-18}$), ILMN_1682717 (tagging *IER3*, $P_{SMR}=1.07\times10^{-16}$) and ILMN_1661439 (tagging *FLOT1*, $P_{SMR}=1.16\times10^{-14}$; **Figure 2**, **Table 2**, **Table S1**). There were six genes, each of which was tagged by two probes, including *HLA-C*, *HLA-H*, *HLA-DRB6*, *HSPA7*, *HIST1H2BK* and *BTN3A2*. Using GTEx eQTL data, our SMR analysis identified 24 probes tagging 24 unique genes that were pleiotropically/potentially causally associated with IgAN, with the top three probes being ENSG00000271581.1 (tagging *XXbac-BPG248L24.12*, $P_{SMR}=1.44\times10^{-10}$; **Figure 3**), ENSG00000186470.9 (tagging *BTN3A2*, $P_{SMR}=2.28\times10^{-10}$; **Figure 4**), and ENSG00000224389.4 (tagging *C4B*, $P_{SMR}=1.23\times10^{-9}$; **Figure 3**, **Table 2**, **Table S1**).

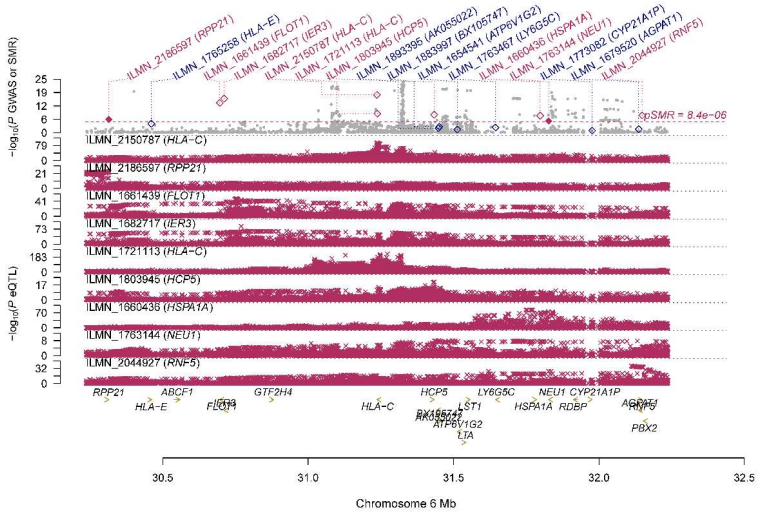

Figure 2: Pleiotropic association of *HLA-C* with IgAN.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 May 2022 doi:10.20944/preprints202205.0277.v1

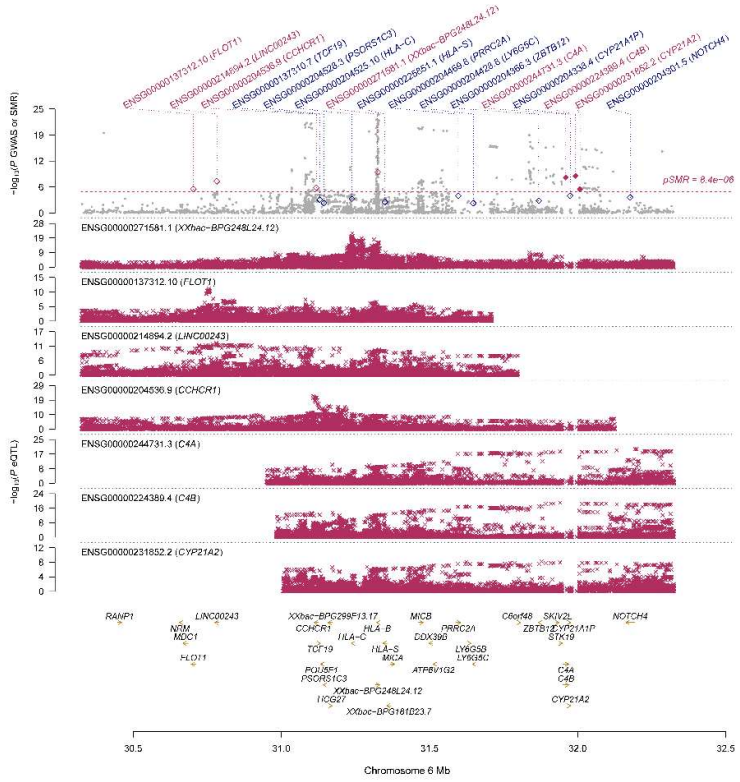

Figure 3: Pleiotropic association of *XXbac-BPG248L24.12*, *C4A* and *C4B* with IgAN.

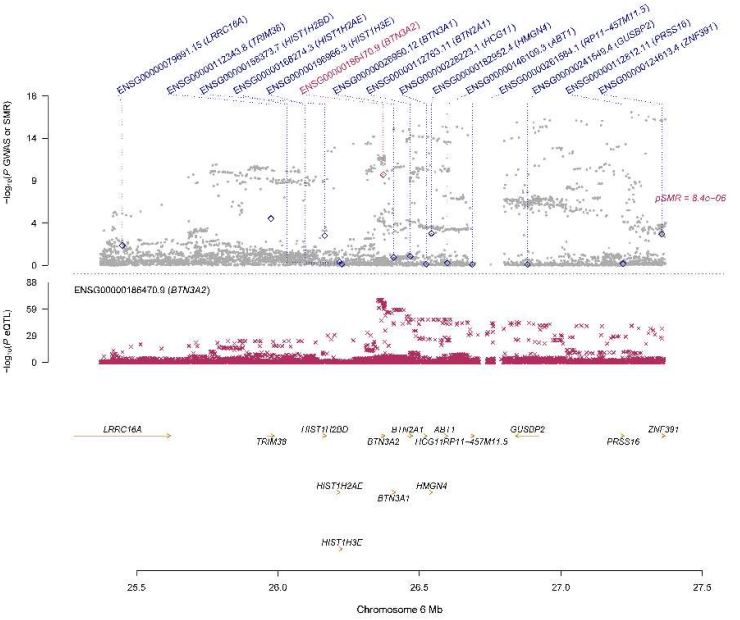

Figure 4: Pleiotropic association of *BTN3A2* with IgAN.

Table 2: The top ten hit probes identified in each SMR analysis.

eQTL data	Probe	Gene	CHR	Top SNP	P_{eQTL}	P_{GWAS}	Beta	SE	P_{SMR}	P_{HEIDI}	Q value

 doi:10.20944/preprints202205.0277.v1

CAGE sparse	ILMN_2150787	*HLA-C*	6	rs2523578	1.55×10^{-65}	2.62×10^{-24}	-0.2378	0.0272	2.10×10^{-18}	9.48×10^{-12}	5.73×10^{-13}
	ILMN_1682717	*IER3*	6	rs2233980	2.05×10^{-60}	5.47×10^{-22}	-0.2065	0.0249	1.07×10^{-16}	1.89×10^{-03}	7.30×10^{-12}
	ILMN_1661439	*FLOT1*	6	rs3130985	6.93×10^{-37}	2.55×10^{-22}	-0.2666	0.0345	1.16×10^{-14}	2.68×10^{-02}	3.53×10^{-10}
	ILMN_1700067	*BTN3A2*	6	rs9393710	9.06×10^{-237}	3.86×10^{-12}	0.0832	0.0123	1.24×10^{-11}	9.13×10^{-08}	2.11×10^{-07}
	ILMN_1820787	*BTN2A1*	6	rs3734544	3.34×10^{-189}	1.40×10^{-10}	-0.0877	0.0140	3.60×10^{-10}	1.40×10^{-08}	3.93×10^{-06}
	ILMN_1721113	*HLA-C*	6	rs9266075	2.99×10^{-183}	5.90×10^{-10}	-0.0867	0.0143	1.35×10^{-09}	1.12×10^{-03}	1.02×10^{-05}
	ILMN_1803945	*HCP5*	6	rs2596495	5.73×10^{-14}	1.58×10^{-21}	0.4716	0.0798	3.48×10^{-09}	3.20×10^{-02}	1.94×10^{-05}
	ILMN_2044927	*RNF5*	6	rs192471087	2.49×10^{-37}	1.28×10^{-10}	-0.1974	0.0343	8.59×10^{-09}	1.29×10^{-04}	3.00×10^{-05}
	ILMN_1804571	*ZKSCAN4*	6	rs13200462	2.61×10^{-19}	1.37×10^{-13}	-0.3193	0.0558	1.08×10^{-08}	1.88×10^{-02}	3.00×10^{-05}
	ILMN_1660436	*HSPA1A*	6	rs494620	1.40×10^{-74}	1.81×10^{-09}	0.1417	0.0248	1.10×10^{-08}	2.16×10^{-09}	3.00×10^{-05}
Whole blood	ENSG00000271581.1	*XXbac-BPG248L2 4.12*	6	rs9266244	1.52×10^{-16}	2.96×10^{-24}	0.2043	0.0319	1.44×10^{-10}	4.34×10^{-07}	6.22×10^{-06}
	ENSG00000186470.9	*BTN3A2*	6	rs68112369	9.15×10^{-70}	1.14×10^{-11}	0.1088	0.0172	2.28×10^{-10}	1.07×10^{-03}	6.22×10^{-06}
	ENSG00000224389.4	*C4B*	6	rs1270942	3.63×10^{-19}	1.44×10^{-16}	-0.1233	0.0203	1.23×10^{-09}	5.84×10^{-02}	1.48×10^{-05}
	ENSG00000244731.3	*C4A*	6	rs116667074	5.15×10^{-19}	1.67×10^{-15}	0.1382	0.0233	2.88×10^{-09}	1.08×10^{-01}	1.96×10^{-05}
	ENSG00000214894.2	*LINC00243*	6	rs3094222	8.03×10^{-12}	3.93×10^{-22}	-0.3086	0.0553	2.38×10^{-08}	5.63×10^{-03}	1.04×10^{-04}
	ENSG00000204644.5	*ZFP57*	6	rs2747431	2.03×10^{-85}	5.51×10^{-08}	-0.0566	0.0108	1.62×10^{-07}	7.31×10^{-05}	4.88×10^{-04}
	ENSG00000204536.9	*CCHCR1*	6	rs1265087	4.64×10^{-22}	1.01×10^{-08}	-0.2081	0.0423	8.60×10^{-07}	7.49×10^{-04}	1.91×10^{-03}
	ENSG00000137312.10	*FLOT1*	6	rs3094222	3.74×10^{-08}	3.93×10^{-22}	-1.3396	0.2802	1.74×10^{-06}	1.11×10^{-02}	2.49×10^{-03}
	ENSG00000231852.2	*CYP21A2*	6	rs7763805	2.24×10^{-10}	5.15×10^{-13}	-0.1665	0.0349	1.85×10^{-06}	3.40×10^{-01}	2.49×10^{-03}
	ENSG00000204622.6	*HLA-J*	6	rs1611336	6.54×10^{-20}	8.13×10^{-08}	0.1162	0.0251	3.78×10^{-06}	4.10×10^{-04}	4.10×10^{-03}

*The GWAS summarized data were provided by the study of Sakaue et al. and can be downloaded at

P_{eQTL} is the P-value of the top associated cis-eQTL in the eQTL analysis, and P_{GWAS} is the P-value for the top associated cis-eQTL in the GWAS analysis. Beta is the estimated effect size in SMR analysis, SE is the corresponding standard error, P_{SMR} is the P-value for SMR analysis and P_{HEIDI} is the P-value for the HEIDI test. Q value is the adjusetd p-value found using false discover rate

CAGE, Consortium for the Architecture of Gene Expression; CHR, chromosome; eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue Expression; HEIDI, heterogeneity in dependent instruments; SNP, single-nucleotide polymorphism; SMR, summary data-based Mendelian randomization; GWAS, genome-wide association studies; IgAN, immunoglobulin A nephropathy

Functional mapping and annotation

FUMA analysis identified 3 independent, significant and lead SNPs (rs2076030, rs469228, and rs1884937; **Table S2-S4**), and 2 genomic risk loci (**Figure 5**; **Table S5**). All the three SNPs are located on chromosome 6. In addition, FUMA identified 39 genes that are potentially involved in the pathogenesis of IgAN (**Table S6**). All the 39 genes are clustered in one genomic risk locus, with the other genomic locus containing no identified genes (**Figure 5** & **Table S6**). Of the 39 identified genes, four were also identified by SMR analysis using CAGE eQTL data, including *HIST1H2BK*, *ZSCAN16*, *ZKSCAN4* and *ZKSCAN3*; and one (*TRIM27*) was identified by SMR analysis using GTEx eQTL data. Expression of the prioritized genes in 54 tissues can be found in **Table S7** and **Figure S1**.

Gene-set enrichment analysis (GSEA) was undertaken to test the possible biological mechanisms of the 39 candidate genes implicated in IgAN (**Table S8**). A total of 310 gene

 doi:10.20944/preprints202205.0277.v1

sets with an adjusted P < 0.05 were identified. We found enrichment signals in intestinal immune network such as SLE (adjusted $P=9.00\times10^{-23}$) as revealed by a recent GWAS study[13], and chromatin-related pathways such as chromatin assembly (adjusted $P=5.58\times10^{-11}$) which are crucial regulators in cellular immunity[26].

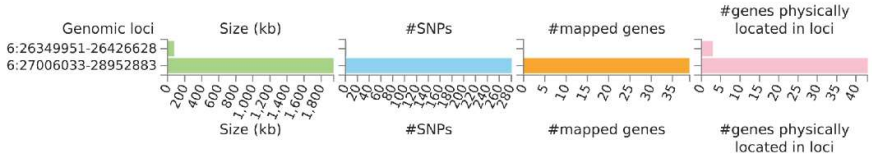

Figure 5: Genetic risk loci identified by FUMA analysis using GWAS data on IgAN.

4. Discussion

In this study, we conducted SMR and FUMA analysis to prioritize genetic loci associated with IgAN. We identified multiple genetic variants, genes, gene sets and two genomic loci that may be involved in the pathogenesis of IgAN. These findings provided helpful leads to a better understanding of the etiology of IgAN.

We found that multiple genes in the human leukocyte antigen (HLA) complex showed significantly pleiotropic association with IgAN using CAGE and/or GTEx eQTL data, such as *HLA-A*/*C*/*E*/*H*/*J*/*L*, *HLA-DQA1*/*A2*, and *HLA-DRB1*/*B6*. The HLA complex, known as major histocompatibility complex (MHC), plays important roles in enabling the immune system to recognize "self" versus "non-self" antigens. The first association between HLA and renal disease was reported more than 50 years ago. Since then, mounting evidence has demonstrated the importance of the HLA complex in IgAN[27]. Many genetic variants in the HLA complex have been found to be associated with the risk of IgAN[28-30]. Previous GWAS studies also identified a few genetic variants in MHC in association with IgAN in individuals of European and East-Asian ancestry[9,11-13]. A recent study found that the expression of *HLA-DQB1* and *HLA-DRB1* decreased on the peripheral blood lymphocytes (PBLs) in IgAN patients, compared with the controls, and that abnormal *HLA-DQB1* and *HLA-DRB1* expression may aggravate the progression of IgAN[31], suggesting the possible involvement of the abnormal expression of both genes in the pathogenesis of IgAN. Abnormal mRNA expression of some HLA genes has been observed in many autoimmune diseases such as lupus and was found to be related to DNA methylation[32,33]. These findings indicated that inflammation and DNA methylation might be two possible mechanisms underlying the HLA's involvement in IgAN. It should be noted that HEIDI test was significant for some of the HLA genes except *HLA-DRB6*, *HLA-DQA1*, *HLA-DRB1*, *HLA-DQA2* (**Tabel S1**), indicating the existence of pleiotropy. More studies are needed to elucidate the exact functions of HLA genes in the pathogenesis of IgAN.

In SMR analysis, we also found that two genes in the complement component C4 family, including *C4A* and *C4B*, showed significantly pleiotropic association with IgAN using GTEx eQTL data, with no significant pleiotropy (**Table 2**). Both genes are mapped in III region of the major histocompatibility complex (MHC) on chromosome 6p21.3[34]. The two genes, together with three other neighboring genes including *RP* (serine–threonine kinase), *CYP21* (steroid 21-hydroxylase) and *TNX* (tenascin-X), form a genetic unit called RCCX module (RP-C4A-CYP21-TNX or RP-C4B-CYP21-TNX) which determines gene copy number (GCN) variation[35]. GCN of the two genes was found to be associated with many autoimmune diseases. For example, a previous meta-analysis found that low *C4A* GCNs were associated with increased risk of SLE in Caucasian populations[36]. The expression of *C4A* and *C4B* was significantly upregulated in glomeruli of patients with IgAN[37]. The whole complement system can be activated by three pathways, including the classical pathway, the lectin binding pathway, and the alternative pathway, and *C4A* and *C4B* are likely involved in the classical pathway[38]. However, the exact roles of the

 doi:10.20944/preprints202205.0277.v1

two genes and the mechanisms underlying the pathogenesis of IgAN remain to be explored.

Our study has some limitations. The number of eligible probes and the sample size of the eQTL for the SMR analysis was limited, especially the GTEx eQTL. Moreover, we used FDR to correct for multiple testing. Together, we may miss some important genes that were not tagged in the eQTL data or filtered out by FDR. The HEIDI test indicated the existence of horizontal pleiotropy for some of the observed associations (**Table 2**). Our SMR analysis used eQTL data from the blood as eQTL data for kidney is not available in GTEx V7. It would be interesting to explore whether the findings still hold using GTEx V8 which has eQTL data for the cortex of kidney.

In summary, we performed SMR and FUMA analysis and identified many genetic variants/loci that are potentially involved in the pathogenesis of IgAN. More studies are needed to elucidate the exact mechanism of the identified genetic variants/loci in the etiology of IgAN.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1, Expression values of prioritized genes in 54 tissues; Table S1, The significant probes identified in SMR analysis; Table S2, Summary of SNPs and mapped genes for SNP2GENE of FUMA analysis.; Table S3, Independent significant SNPs (r^2<0.6) identified from IGA GWAS; Table S4, Lead SNPs identified from independent significant SNPs of IGA GWAS; Table S5, Genomic risk loci of interest from IGA GWAS; Table S6, Prioritized genes from IGA GWAS by functional mapping; Table S7, Expression values of prioritized genes in 54 tissues.

Author Contributions: Conceptualization, J.Y. and M.Z.; methodology, Q.Z. and K.Z.; software, Y.Z.; validation, Q.Z., J.Y. and M.Z.; formal analysis, Y.Z. and J.Y; investigation, Q.Z., K.Z. and M.Z; resources, J.Y.; data curation, Y.Z and J.Y.; writing—original draft preparation, Q.Z. and J.Y.; writing—review and editing, all authors; visualization, Y.Z. and J.Y; supervision, J.Y. and M.Z.; project administration, J.Y. and M.Z. All authors have read and agreed to the published version of the manuscript."

Funding: Dr. Jingyun Yang's research was supported by NIH/NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG033678, R01AG36042, U01AG61356, and 1RF1AG064312– 01.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The BioBank Japan (BBJ) [20], the UK Biobank[21], and GWAS summary results were from FinnGen (https://www.finngen.fi/). The GWAS summarized data can be downloaded at http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90018001-GCST90019000/GCST90018866/harmonised/. The eQTL data can be downloaded at https://cnsgenomics.com/data/SMR/#eQTLsummarydata.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Figure 1. Flow chart for the SMR analysis.
A) SMR analysis using eQTL data from CAGE; and B) SMR analysis using eQTL data from GTEx
eQTL, expression quantitative trait loci; GWAS, genome-wide association studies; LD, linkage disequilibrium; SMR, summary data-based Mendelian randomization; SNP, single nucleotide polymorphisms

Figure 2. Pleiotropic association of *HLA-C* with IgAN.
Top plot, grey dots represent the -log10(P values) for SNPs from the GWAS of IOP, with solid rhombuses indicating that the probes pass HEIDI test. Middle plot, eQTL results. Bottom plot, location of genes tagged by the probes.

 doi:10.20944/preprints202205.0277.v1

GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quan-titative trait loci; IgAN, immunoglobulin A nephropathy

Figure 3. Pleiotropic association of *XXbac-BPG248L24.12*, *C4A* and *C4B* with IgAN.
Top plot, grey dots represent the -log10(P values) for SNPs from the GWAS of IOP, with solid rhombuses indicating that the probes pass HEIDI test. Middle plot, eQTL results. Bottom plot, location of genes tagged by the probes.
GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quan-titative trait loci; IgAN, immunoglobulin A nephropathy

Figure 4. Pleiotropic association of *BTN3A2* with IgAN.
Top plot, grey dots represent the -log10(P values) for SNPs from the GWAS of IOP, with solid rhombuses indicating that the probes pass HEIDI test. Middle plot, eQTL results. Bottom plot, location of genes tagged by the probes.
GWAS, genome-wide association studies; SMR, summary data-based Mendelian randomization; HEIDI, heterogeneity in dependent instruments; eQTL, expression quan-titative trait loci; IgAN, immunoglobulin A nephropathy

Figure 5. Genetic risk loci identified by FUMA analysis using GWAS data on IgAN.
Genomic risk loci are displayed in the format of 'chromosome:start position-end position' on the Y axis. For each genomic locus, histograms from left to right depict the size, the number of candidate SNPs, the number of mapped genes (using positional mapping and eQTL mapping), and the number of genes known to be located within the genomic locus, respectively.
eQTL, expression quantitative trait loci; GWAS, genome-wide association studies; SNP, single nucleotide polymorphism; IgAN, immunoglobulin A nephropathy

References

1. Berger, J.; Hinglais, N. [Intercapillary deposits of IgA-IgG]. *J Urol Nephrol (Paris)* **1968**, *74*, 694-695.
2. D'Amico, G. The commonest glomerulonephritis in the world: IgA nephropathy. *Q J Med* **1987**, *64*, 709-727.
3. Roberts, I.S. Pathology of IgA nephropathy. *Nat Rev Nephrol* **2014**, *10*, 445-454, doi:10.1038/nrneph.2014.92.
4. Li, P.K.; Ho, K.K.; Szeto, C.C.; Yu, L.; Lai, F.M. Prognostic indicators of IgA nephropathy in the Chinese--clinical and pathological perspectives. *Nephrol Dial Transplant* **2002**, *17*, 64-69, doi:10.1093/ndt/17.1.64.
5. Le, W.; Liang, S.; Hu, Y.; Deng, K.; Bao, H.; Zeng, C.; Liu, Z. Long-term renal survival and related risk factors in patients with IgA nephropathy: results from a cohort of 1155 cases in a Chinese adult population. *Nephrol Dial Transplant* **2012**, *27*, 1479-1485, doi:10.1093/ndt/gfr527.
6. Wakai, K.; Kawamura, T.; Matsuo, S.; Hotta, N.; Ohno, Y. Risk factors for IgA nephropathy: a case-control study in Japan. *Am J Kidney Dis* **1999**, *33*, 738-745, doi:10.1016/s0272-6386(99)70228-3.
7. Wakai, K.; Nakai, S.; Matsuo, S.; Kawamura, T.; Hotta, N.; Maeda, K.; Ohno, Y.; Research Group on Ig, A.N. Risk factors for IgA nephropathy: a case-control study with incident cases in Japan. *Nephron* **2002**, *90*, 16-23, doi:10.1159/000046309.
8. Schena, F.P.; Nistor, I. Epidemiology of IgA Nephropathy: A Global Perspective. *Semin Nephrol* **2018**, *38*, 435-442, doi:10.1016/j.semnephrol.2018.05.013.
9. Feehally, J.; Farrall, M.; Boland, A.; Gale, D.P.; Gut, I.; Heath, S.; Kumar, A.; Peden, J.F.; Maxwell, P.H.; Morris, D.L.; et al. HLA has strongest association with IgA nephropathy in genome-wide analysis. *J Am Soc Nephrol* **2010**, *21*, 1791-1797, doi:10.1681/ASN.2010010076.

10. Gharavi, A.G.; Kiryluk, K.; Choi, M.; Li, Y.; Hou, P.; Xie, J.; Sanna-Cherchi, S.; Men, C.J.; Julian, B.A.; Wyatt, R.J.; et al. Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat Genet* **2011**, *43*, 321-327, doi:10.1038/ng.787.
11. Yu, X.Q.; Li, M.; Zhang, H.; Low, H.Q.; Wei, X.; Wang, J.Q.; Sun, L.D.; Sim, K.S.; Li, Y.; Foo, J.N.; et al. A genome-wide association study in Han Chinese identifies multiple susceptibility loci for IgA nephropathy. *Nat Genet* **2011**, *44*, 178-182, doi:10.1038/ng.1047.
12. Li, M.; Foo, J.N.; Wang, J.Q.; Low, H.Q.; Tang, X.Q.; Toh, K.Y.; Yin, P.R.; Khor, C.C.; Goh, Y.F.; Irwan, I.D.; et al. Identification of new susceptibility loci for IgA nephropathy in Han Chinese. *Nat Commun* **2015**, *6*, 7270, doi:10.1038/ncomms8270.
13. Kiryluk, K.; Li, Y.; Scolari, F.; Sanna-Cherchi, S.; Choi, M.; Verbitsky, M.; Fasel, D.; Lata, S.; Prakash, S.; Shapiro, S.; et al. Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. *Nat Genet* **2014**, *46*, 1187-1196, doi:10.1038/ng.3118.
14. Kiryluk, K.; Li, Y.; Moldoveanu, Z.; Suzuki, H.; Reily, C.; Hou, P.; Xie, J.; Mladkova, N.; Prakash, S.; Fischman, C.; et al. GWAS for serum galactose-deficient IgA1 implicates critical genes of the O-glycosylation pathway. *PLoS Genet* **2017**, *13*, e1006609, doi:10.1371/journal.pgen.1006609.
15. Davey Smith, G.; Hemani, G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* **2014**, *23*, R89-98, doi:10.1093/hmg/ddu328.
16. Mo, X.; Guo, Y.; Qian, Q.; Fu, M.; Lei, S.; Zhang, Y.; Zhang, H. Mendelian randomization analysis revealed potential causal factors for systemic lupus erythematosus. *Immunology* **2020**, *159*, 279-288, doi:10.1111/imm.13144.
17. Porcu, E.; Rueger, S.; Lepik, K.; e, Q.C.; Consortium, B.; Santoni, F.A.; Reymond, A.; Kutalik, Z. Mendelian randomization integrating GWAS and eQTL data reveals genetic determinants of complex and clinical traits. *Nat Commun* **2019**, *10*, 3300, doi:10.1038/s41467-019-10936-0.
18. Liu, D.; Yang, J.; Feng, B.; Lu, W.; Zhao, C.; Li, L. Mendelian randomization analysis identified genes pleiotropically associated with the risk and prognosis of COVID-19. *J Infect* **2021**, *82*, 126-132, doi:10.1016/j.jinf.2020.11.031.
19. Sakaue, S.; Kanai, M.; Tanigawa, Y.; Karjalainen, J.; Kurki, M.; Koshiba, S.; Narita, A.; Konuma, T.; Yamamoto, K.; Akiyama, M.; et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet* **2021**, *53*, 1415-1424, doi:10.1038/s41588-021-00931-x.
20. Nagai, A.; Hirata, M.; Kamatani, Y.; Muto, K.; Matsuda, K.; Kiyohara, Y.; Ninomiya, T.; Tamakoshi, A.; Yamagata, Z.; Mushiroda, T.; et al. Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol* **2017**, *27*, S2-S8, doi:10.1016/j.je.2016.12.005.
21. Bycroft, C.; Freeman, C.; Petkova, D.; Band, G.; Elliott, L.T.; Sharp, K.; Motyer, A.; Vukcevic, D.; Delaneau, O.; O'Connell, J.; et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **2018**, *562*, 203-209, doi:10.1038/s41586-018-0579-z.
22. Lloyd-Jones, L.R.; Holloway, A.; McRae, A.; Yang, J.; Small, K.; Zhao, J.; Zeng, B.; Bakshi, A.; Metspalu, A.; Dermitzakis, M.; et al. The Genetic Architecture of Gene Expression in Peripheral Blood. *Am J Hum Genet* **2017**, *100*, 371, doi:10.1016/j.ajhg.2017.01.026.
23. Consortium, G. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **2020**, *369*, 1318-1330, doi:10.1126/science.aaz1776.
24. Watanabe, K.; Taskesen, E.; van Bochoven, A.; Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **2017**, *8*, 1826, doi:10.1038/s41467-017-01261-5.

25. Zhu, Z.; Zhang, F.; Hu, H.; Bakshi, A.; Robinson, M.R.; Powell, J.E.; Montgomery, G.W.; Goddard, M.E.; Wray, N.R.; Visscher, P.M.; et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* **2016**, *48*, 481-487, doi:10.1038/ng.3538.
26. Mazzone, R.; Zwergel, C.; Artico, M.; Taurone, S.; Ralli, M.; Greco, A.; Mai, A. The emerging role of epigenetics in human autoimmune disorders. *Clin Epigenetics* **2019**, *11*, 34, doi:10.1186/s13148-019-0632-2.
27. Patel, R.; Mickey, M.R.; Terasaki, P.I. Leucocyte antigens and disease. I. Association of HL-A2 and chronic glomerulonephritis. *Br Med J* **1969**, *2*, 424-426, doi:10.1136/bmj.2.5654.424.
28. Jiyun, Y.; Guisen, L.; Li, Z.; Yi, S.; Jicheng, L.; Fang, L.; Xiaoqi, L.; Shi, M.; Cheng, J.; Ying, L.; et al. The genetic variants at the HLA-DRB1 gene are associated with primary IgA nephropathy in Han Chinese. *BMC Medical Genetics* **2012**, *13*, 33, doi:10.1186/1471-2350-13-33.
29. Akiyama, F.; Tanaka, T.; Yamada, R.; Ohnishi, Y.; Tsunoda, T.; Maeda, S.; Takei, T.; Obara, W.; Ito, K.; Honda, K.; et al. Single-nucleotide polymorphisms in the class II region of the major histocompatibility complex in Japanese patients with immunoglobulin A nephropathy. *J Hum Genet* **2002**, *47*, 532-538, doi:10.1007/s100380200080.
30. Cao, H.X.; Li, M.; Nie, J.; Wang, W.; Zhou, S.F.; Yu, X.Q. Human leukocyte antigen DRB1 alleles predict risk and disease progression of immunoglobulin A nephropathy in Han Chinese. *Am J Nephrol* **2008**, *28*, 684-691, doi:10.1159/000122111.
31. Zhan, X.; Deng, F.; Wang, A.Y.; Chen, Q.; Du, Y.; Wang, Q.; Zhong, X.; Zhang, P.; Wang, W.; Chen, S.; et al. HLA-DQB1 and HLA-DRB1 expression is associated with disease severity in IgAN. *Ann Palliat Med* **2021**, *10*, 9453-9466, doi:10.21037/apm-21-2065.
32. Miller, S.; Tsou, P.S.; Coit, P.; Gensterblum-Miller, E.; Renauer, P.; Rohraff, D.M.; Kilian, N.C.; Schonfeld, M.; Sawalha, A.H. Hypomethylation of STAT1 and HLA-DRB1 is associated with type-I interferon-dependent HLA-DRB1 expression in lupus CD8+ T cells. *Ann Rheum Dis* **2019**, *78*, 519-528, doi:10.1136/annrheumdis-2018-214323.
33. Kular, L.; Liu, Y.; Ruhrmann, S.; Zheleznyakova, G.; Marabita, F.; Gomez-Cabrero, D.; James, T.; Ewing, E.; Linden, M.; Gornikiewicz, B.; et al. DNA methylation as a mediator of HLA-DRB1*15:01 and a protective variant in multiple sclerosis. *Nat Commun* **2018**, *9*, 2397, doi:10.1038/s41467-018-04732-5.
34. Martinez, O.P.; Longman-Jacobsen, N.; Davies, R.; Chung, E.K.; Yang, Y.; Gaudieri, S.; Dawkins, R.L.; Yu, C.Y. Genetics of human complement component C4 and evolution the central MHC. *Front Biosci* **2001**, *6*, D904-913, doi:10.2741/martinez.
35. Yang, Z.; Mendoza, A.R.; Welch, T.R.; Zipf, W.B.; Yu, C.Y. Modular variations of the human major histocompatibility complex class III genes for serine/threonine kinase RP, complement component C4, steroid 21-hydroxylase CYP21, and tenascin TNX (the RCCX module). A mechanism for gene deletions and disease associations. *J Biol Chem* **1999**, *274*, 12147-12156, doi:10.1074/jbc.274.17.12147.
36. Li, N.; Zhang, J.; Liao, D.; Yang, L.; Wang, Y.; Hou, S. Association between C4, C4A, and C4B copy number variations and susceptibility to autoimmune diseases: a meta-analysis. *Sci Rep* **2017**, *7*, 42628, doi:10.1038/srep42628.
37. Park, S.; Yang, S.H.; Jeong, C.W.; Moon, K.C.; Kim, D.K.; Joo, K.W.; Kim, Y.S.; Lee, J.W.; Lee, H. RNA-Seq profiling of microdissected glomeruli identifies potential biomarkers for human IgA nephropathy. *Am J Physiol Renal Physiol* **2020**, *319*, F809-F821, doi:10.1152/ajprenal.00037.2020.
38. Selvaskandan, H.; Barratt, J.; Cheung, C.K. Immunological drivers of IgA nephropathy: Exploring the mucosa-kidney link. *Int J Immunogenet* **2022**, *49*, 8-21, doi:10.1111/iji.12561.