博士論文

Genome-wide association study and HLA fine-mapping for childhood steroid-sensitive nephrotic syndrome in the Japanese population

(日本人小児ステロイド感受性ネフローゼ症候群に関するゲノムワイド関連解析と HLA マッピング)

賈　暁媛
Jia Xiaoyuan

【論文タイトル】

Genome-wide association study and HLA fine-mapping for childhood steroid-sensitive nephrotic syndrome in the Japanese population (日本人小児ステロイド感受性ネフローゼ症候群に関するゲノムワイド関連解析と HLA マッピング)

【所属】

人類遺伝学

【指導教員】

徳永　勝士

【申請者名】

賈　暁媛 (Jia Xiaoyuan)

Publication

Strong Association of *HLA-DR/DQ* Locus with Childhood Steroid-Sensitive Nephrotic Syndrome in the Japanese Population. *Journal of the American Society of Nephrology*. In press.

Conferences

Poster: “Genome-wide association study identifies *HLA-DR/DQ* region for childhood nephrotic syndrome in Japanese”. The American Society of Human Genetics 2017 Annual Meeting. USA, Orlando, October 17-21, 2017.

Oral presentation: “Genome-wide association study reveals strong association of *HLA-DR/DQ* region in childhood nephrotic syndrome in the Japanese population”. The 26th Annual Meeting of the Japanese Society for Histocompatibility and Immunogenetics (JSHI, 2017). Japan, Hiroshima, October 27-29, 2017.

Oral presentation: “Genome-wide association study reveals strong association of *HLA-DR/DQ* region in childhood nephrotic syndrome in the Japanese population”. The 62nd Annual Meeting of the Japan Society of Human Genetics (JSHG, 2017). Japan, Kobe, November 15-18, 2017.

Table of Contents

Abstract 5
Abbreviations 6
General Introduction 8

Chapter 1: Genome-wide association study for childhood steroid-sensitive nephrotic syndrome in the Japanese population 12
Introduction 13
Materials and Methods 14
Results 19
Discussion 23

Chapter 2: Fine-mapping of HLA association 26
Introduction 27
Materials and Methods 29
Results 32
Discussion 37

Summary and Conclusions 43
Supplementary 45
Acknowledgements 47
References 48
Tables and figures 60

Abstract

Nephrotic syndrome (NS) is the most common cause of chronic glomerular disease in children. In Japan, the estimated incidence of idiopathic NS (INS) is 6.49 cases/100,000 children annually. Between 80 and 90% of patients achieve remission with steroid therapy (steroid-sensitive NS [SSNS]), whereas 10-20% exhibit resistance to the therapy (steroid-resistant NS [SRNS]).

I conducted a genome-wide association study (GWAS) to identify the genetic factors in susceptibility to childhood SSNS in the Japanese population. In the discovery stage, 224 patients with childhood SSNS and 419 adult healthy controls were genotyped using Affymetrix 'Japonica Array', then whole-genome imputation was performed. The most significant association was detected in the *HLA-DR/DQ* region and top SNPs were replicated in an independent Japanese sample set including 216 cases and 719 healthy controls by DigiTag2 assay [rs4642516, combined $P = 7.84\times10^{-23}$, odds ratio (OR)=0.33 under allelic model, combined $P = 4.87\times10^{-23}$, OR=0.29 under dominant model; rs3134996, combined $P = 1.72\times10^{-25}$, OR=0.29 under allelic model, combined $P = 9.57\times10^{-21}$, OR=0.28 under dominant model]. High-accuracy HLA-imputation and direct HLA-typing were conducted to fine-map the *HLA* association. *HLA-DRB1*08:02* ($Pc = 1.82\times10^{-9}$, OR=2.62), and *HLA-DQB1*06:04* ($Pc = 2.09\times10^{-12}$, OR=0.10) were considered as primary *HLA* alleles associated with Japanese childhood SSNS. *HLA-DRB1*08:02-DQB1*03:02* ($Pc = 7.01\times10^{-11}$, OR=3.60) was identified as the most significant susceptibility haplotype and showed more significant and stronger association than single *HLA* allele. *HLA-DRB1*13:02-DQB1*06:04* ($Pc = 4.18\times10^{-12}$, OR=0.10) was identified as the most significant protective haplotype. The findings indicate that HLA fine-mapping after SNP-based analysis is essential for a better understanding of disease mechanism.

Abbreviations

bp	base pair
CEU	Northern and Western Europeans from Utah
CHR/Chr	Chromosome
CHB	Han Chinese in Beijing
DNA	Deoxyribonucleic Acid
eQTL	Expression Quantitative Trait Loci
ESRD	End-Stage Renal Disease
FRNS	Frequently-Relapse Nephrotic Syndrome
FSGS	Focal Segmental Glomerulosclerosis
GBM	Glomerular Basement Membrane
GC	Genomic Control
GWAS	Genome-wide Association Study
HLA	Human Leukocyte Antigen
HWE	Hardy-Weinberg Equilibrium
IBD	Identity-By-Descent
IL	Interleukin
INDELs	Insertions or Deletions
INS	Idiopathic Nephrotic Syndrome
JPT	Japanese in Tokyo
kb	kilobase
LD	Linkage Disequilibrium
LOD	Logarithm of the Odds

MAF	Minor Allele Frequency
Mb	Megabases
MCD	Minimal Change Disease
MHC	Major Histocompatibility Complex
OR	Odds Ratio
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
QC	Quality Control
Q-Q	Quantile-quantile
RR	Relative Risk
SDNS	Steroid-Dependent Nephrotic Syndrome
SNP	Single Nucleotide Polymorphism
SRNS	Steroid-Resistant Nephrotic Syndrome
SSNS	Steroid-Sensitive Nephrotic Syndrome
TNF	Tumor Necrosis Factor
VEGF	Vascular Endothelial Growth Factor
YRI	Yoruba in Nigeria

General Introduction

Nephrotic syndrome (NS) is a kidney disorder characterized by heavy proteinuria, hypoalbuminemia, edema and dyslipidemia. NS can be classified as primary NS (also called idiopathic NS) which only affects the kidneys, or secondary NS caused by systemic disorders. Idiopathic nephrotic syndrome (INS) is the most common glomerular disease in children, accounting for about 90% of children with NS. Geographic or ethnic differences have been reported about the incidence of childhood INS. The annual incidence of NS in children has been estimated to be 2.0-2.7 per 100,000 in USA with a cumulative prevalence of 16 cases per 100,000 children[1, 2]. South Asian children had a six-time higher incidence than Europeans, whereas East/Southeast Asians had a similar rate as European children[3]. In Japan, the estimated incidence of INS is 6.49 cases/100,000 children annually[4]. The main pathological types of INS in pediatric patients are minimal change disease (MCD, 85%), focal segmental glomerulosclerosis (FSGS, 10%) and mesangial proliferation (5%)[5].

Corticosteroids are used as the first-line treatment to induce the remission of INS and are effective in 80-90% of pediatric patients. 80-90% of patients with childhood INS achieved complete remission in 4 weeks after steroid treatment (60mg/m^2 oral prednisolone per day), which are defined as steroid-sensitive NS (SSNS). While 10-20% of pediatric cases exhibit steroid-resistant NS (SRNS), defined by persistent proteinuria after 60mg/m^2 oral prednisolone per day for 4 weeks. Up to 90% of steroid responders will relapse, with approximately half becoming frequently-relapsing NS (FRNS) or steroid-dependent NS (SDNS)[6] (Table 1). Alternative immunosuppressive agents including alkylating agents, cyclosporine, mycophenolate mofetil (MMF) and rituximab are used in patients suffering significant side effects of steroid therapy, or the ones who do not respond to steroid treatment. Generally speaking, childhood INS

patients, especially the ones with SSNS, have a favorable long-term outcome. However, the prognosis of patients with SRNS is poor, with 50% developing end-stage renal disease (ESRD) within 15 years[7]. The long-term course of disease and the adverse effects caused by steroid and/or other immunosuppressive agents often have detrimental effects on the life quality of patients with INS and their families[8].

The underlying pathogenesis and genetic background of INS is not fully understood. Previous studies have highlighted the crucial role of podocytes in the development of INS[9, 10]. Podocytes are terminally differentiated epithelial cells that locate at the outermost layer of the glomerular basement membrane (GBM). Podocytes form the final barrier to urinary protein loss by the formation and maintenance of podcyte foot processes and the interposed slit diaphragms, which explain why podocyte injury is typically associated with marked proteinuria[11]. The podocyte dysfunction or injury can be caused by genetic abnormalities of podocyte-related genes, circulating factors, infections, drugs and toxins[12, 13]. As the most common pathological type of INS in pediatric patients, both MCD and FSGS have been suggested for the category of podocytopathies[14]. Under electron microscopy, podocyte foot process effacement can be observed in MCD and FSGS cases, indicating the podocyte injury[15].

Immune system plays an important role in the development and maintenance of INS. Some INS patients are reported to have infection history or allergen exposure before disease onset or the relapse of NS[16]. About 30% of patients with SSNS have been shown to have allergic symptoms[16, 17]. The favorable response of INS to immunosuppressive treatment has been considered as additional proof of the involvement of immune system. The dysfunctions of T-cells and B-cells contribute to the predisposition of INS[18]. Many researches focused on the relationship between circulating factors and INS, and cytokines are considered as the most likely

pathogenic factors. Patients with MCD relapse were found to have increased levels of various cytokines in serum or urine, including interleukin (IL)-2, soluble IL-2 receptor, interferon-γ, IL-4, IL-12, IL-18, tumor necrosis factor (TNF)-α and vascular endothelial growth factor (VEGF)[11].

A large number of genetic studies have been conducted to elucidate possible genetic factors influencing the susceptibility to INS, including linkage analyses and candidate gene studies. Most of the previous genetic studies focused on families with SRNS (FSGS) in order to discover mutations in podocyte-related genes. In the past 20 years, over 45 recessive or dominant genes encoding proteins in podocytes have been identified in patients with SRNS of hereditary NS, explaining 57-100% of familial and infant-onset NS, and 10-20% of sporadic cases[7, 19, 20]. Genetic alterations in podocyte-related genes were recognized in 29.5% of patients with SRNS[21]. In contrast, SSNS is thought to be an immune-mediated disease[22]. Since SSNS in two generations is rare, the knowledge of specific genetic mutation for familial SSNS is still limited[23, 24]. The strong association between human leukocyte antigen (*HLA*), especially *HLA-DR/-DQ* genes, and the susceptibility to INS has been repeatedly reported in different populations[25-42].

As a powerful approach to investigate disease-associated genetic factors, genome-wide association studies (GWASs) have also been utilized in NS field. Missense variants in *HLA-DQA1* were reported as susceptibility factors associated childhood SSNS in South Asian population and Europeans[43]. Common variant in *GPC5* was identified as a genetic factor contributing to acquired NS in Japanese adults[44]. However, there has been no published GWAS working on childhood-onset INS in the Japanese population yet.

As a heterogenetic disease, the associated genetic factors and the pathogenesis of INS have not been completely elucidated yet. Since the histopathologic spectrum of INS in children differs from that in adults, and there may be ethnic differences in susceptible genes, a GWAS examining childhood INS in the Japanese population is needed.

In this study, I focused on pediatric patients with SSNS and performed a two-stage GWAS in the Japanese population, aiming to identify the genetic factors related to the susceptibility of childhood SSNS. The identification and further fine-mapping of disease-associated genes can provide a better understanding of disease mechanism.

Chapter 1:

Genome-wide association study for childhood steroid-sensitive nephrotic syndrome in the Japanese population

Introduction

Genome-wide association study (GWAS) is a powerful approach to analyze the whole genome using genome-wide SNP array based on linkage disequilibrium (LD) block structures. As there's no prior hypothesis, GWAS can provide an unbiased genetic association with a given phenotype. Since the advent of GWAS in 2005, GWASs have been strikingly successful in identifying common genetic variants associated with complex diseases, contributing substantially to our knowledge of disease mechanisms.

To date, two GWASs for NS patients have been reported. A GWAS including both primary and secondary adult-onset NS patients was conducted in the Japanese population[44]. *GPC5* (rs16946160) was reported as a significant genetic factor associated with severe proteinuria in Japanese patients with adult-onset NS irrespective of the etiology. Gbadegesin RA et al reported a genome-wide association study using exome array[42]. Four exome-wide significant variants in or around *HLA-DQA1* and -*DQB1* were identified in South Asian children with SSNS. Two missense variants in *HLA-DQA1* [C34Y (rs1129740) and F41S (rs1071630)] were further replicated in children of European ancestry.

As a heterogenetic disease, the histopathologic spectrum of INS in children markedly differs from that in adult-onset cases, and there may be ethnic differences in susceptible genes, a GWAS examining childhood INS in the Japanese population is necessary for understanding the disease mechanism.

Here, for the purpose of identifying loci conferring susceptibility to childhood-onset INS, especially for SSNS, I performed a two-stage GWAS for childhood SSNS in the Japanese population.

Materials and Methods

Samples and clinical data

In the discovery stage of GWAS, 224 Japanese patients with childhood-onset INS based on International Study of Kidney Disease in Children (ISKDC) criteria [45] (onset age < 18 years, urine protein to creatinine ratio ≥ 2.0 and serum albumin ≤ 2.5g/dl) and responded to steroid treatment (achieved complete remission within 4 weeks after starting 60 mg/m^2 oral prednisolone per day, considered as SSNS) were included. Patients with the history of steroid resistance during follow-up were not included in this study. The related definitions of NS are shown in Table 1. Another 216 cases were recruited for replication purpose. In total, 440 patients were included, with male to female ratio 2.6:1 and median age at onset 4.0 years (Table 2). There was no onset-age bias between patients in the two sample sets (P=0.12, Figure 1A and 1B). Renal biopsy was performed in 248 of the 440 patients (56%) [Minimal change disease (n=227, 91.5%), Focal segmental glomerulosclerosis (n=14, 5.7%) and Mesangial proliferative glomerulonephritis (n=7, 2.8%)]. DNA samples and clinical information regarding subjects were collected from 39 hospitals across Japan (see Supplementary). Clinical data for the patients were collected using a simple questionnaire. Genomic DNA was extracted from peripheral blood following a standard protocol.

Considering healthy adults have passed childhood period without the disease onset, healthy Japanese adults were selected as controls for this study. In total, 1,138 Japanese healthy adults (age > 18 years) without current diseases or disease history were recruited from Tokyo area in Japan. In the discovery stage, a total of 419 healthy adults were referred by the Department of Human Genetics, Graduate School of Medicine, the University of Tokyo. For the

replication phase, 719 healthy adult controls were recruited from the Pharma SNP Consortium (Tokyo, Japan).

This study was approved by the Research Ethics Committees of Graduate School of Medicine, the University of Tokyo, Graduate School of Medicine, Kobe University, as well as all of the universities and hospitals that participated in this collaborative study. Written informed consent was obtained from all participants (or the parents or guardians of patients) for participation in this study.

Genome-wide SNP genotyping in the discovery GWAS

In the discovery stage of GWAS, 224 patients with childhood SSNS and 419 adult healthy controls were genotyped using Affymetrix 'Japonica Array' (Toshiba, Japan)[46], which was specially designed for the Japanese population based on the whole-genome sequencing data of 1,070 healthy Japanese individuals.

Two hundred nanograms of genomic DNAs were amplified, fragmented and labeled according to the manufacturer's instruction with Nimbus automated system (Hamilton, Reno, NV, USA) controlled by Hamilton Run Control-Axiom (v1.1.0 med, Affymetrix) and Gene Titan Multi-channel instrument operated by AGCC Gene Titan Instrument Control (version 4.1.0.1567, Affymetrix)[46].

Genotype calling and data quality control

The genotype calling for discovery samples was conducted using Genotyping Console software v4.2.0.26 (Affymetrix). Quality control (QC) steps were applied on the genotyped data using PLINK 1.9[47]. I removed SNPs with genotype call rate < 97%, SNPs with minor allele

frequency (MAF) < 5% and SNPs with Hardy-Weinberg equilibrium (HWE) test *p*-value < 0.0001 in healthy controls. Identify-by-descent (IBD) test was done to check the cryptic relatedness among samples and sample duplicates. A Pi-Hat (proportion identity by descent) threshold of > 0.1875 was used to exclude first and second degree relatives. To check the population stratification between cases and controls that may potentially confound the association results, I conducted principal component analysis (PCA) using GCTA (version 1.26.0)[48] for samples in the discovery stage and HapMap Phase III data. HapMap Phase III reference data included 113 individuals of Utah of Northern and Western European ancestry (CEU), 113 individuals of Yoruban descent (YRI), 84 individuals of Han Chinese (CHB) and 86 individuals of Japanese (JPT).

Whole-genome imputation using 2KJPN panel

Before whole-genome imputation, the clustering plots of SNPs were classified by the Ps classification function in the SNPolisher package (version 1.5.2; Affymetrix). SNPs assigned as 'recommended' by the Ps classification were retained. SNPs with call rate < 99%, HWE *p*-value < 0.0001 or MAF < 0.5% were excluded. Prephasing was conducted first with the SNPs which passed the filtering step with SHAPEIT (v.2.r644); the options were –burn 10, –prune 10, and –main 25. Genotype imputation was performed on the phased genotypes with IMPUTE2 (version 2.3.1) using a phased reference panel of 2,049 healthy Japanese individuals (2KJPN panel). For IMPUTE2, the applied options were –Ne 2000, –k hap 1000, –k 120, –burnin 15, and –iter 50.

After whole-genome imputation, QC was conducted using the following thresholds: individual call rate ≥97%, SNP/short insertion or deletion (INDEL) call rate ≥97%, MAF ≥5% and Hardy-Weinberg equilibrium (HWE) test *p*-value ≥0.0001 in healthy controls.

Statistical analyses and selecting candidate loci of interest

Association analyses of SNPs and short INDELs were performed using PLINK 1.9[47]. Since Cochran-Armitage trend test is commonly used in GWAS analysis with the advantage of not relying on the assumption of HWE, SNP-based association analyses were done by Cochran-Armitage trend test. Genomic inflation factor λ was calculated by dividing the observed by the expected median test statistics. The quantitle-quantitile (Q-Q) plots were obtained using the expected distribution of association test statistics under null hypothesis against the observed *P*-values across all tested SNPs/short INDELs. *P*-values for SNPs/short INDELs associations were further corrected by genomic control (GC).

After whole-genome imputation, SNPs/short INDELs were selected from the candidate loci with suggestive significant associations (GC-corrected $P < 1 \times 10^{-5}$) under Cochran-Armitage trend test. All cluster plots of the SNPs with a GC-corrected $P<1\times10^{-5}$ were checked by visual inspection. For replicable SNPs, the associations under allelic model, dominant model and recessive model were further checked using Fisher's exact test.

Manhattan plots and Q-Q plots were generated using R package 'qqman'. Regional plots were prepared using Locuszoom (http://locuszoom.org)[49]. The power of GWAS was calculated using R package 'CATS'[50]. NCBI database (https://www.ncbi.nlm.nih.gov), GTEx Portal database (https://gtexportal.org/home/), HaploReg v4.1[51, 52] (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php) and the NephQTL browser (http://nephqtl.org)[53] were used for variants interpretation. Proxy SNPs were selected by SNAP (SNP Annotation and Proxy Search)[54] using 1000 Genomes Pilot 1 (CHB-JPT panel).

SNP genotyping for validation and replication by DigiTag 2 assay

Genotyping of candidate variants was performed using DigiTag2 assay based on the protocol[55, 56]. Ten nanograms of genomic DNAs were amplified by multiplex PCR. Then multiplexed oligonucleotide ligation assay was performed, and the labeling reaction was achieved with two 5' query probes and one common probe prepared for a single SNP site. The 5' query probes had a sequence complementary to that of the 5' sequence flanking the target SNP and each of the probes had an allele-specific sequence. Then hybridization reaction was performed using a DNA microarray (NGK Insulators) with separated areas containing oligonucleotide probes for the tested SNPs. The genotype calls were determined using SNPStar software (version 0.0.0.8, Olympus).

As the validation study, genotyping was performed in 224 cases in discovery sample set. The genotypes determined by genome-wide SNP array and DigiTag2 assay were compared to assess the typing accuracy.

As the replication study, genotyping of candidate variants was done in an independent Japanese sample set including 216 cases and 719 healthy controls. Combined analysis was done for the replicable SNPs using typed data in discovery GWAS and data in replication stage.

Results

Genotyping and QC procedure in the discovery GWAS

224 patients with childhood SSNS and 419 adult healthy controls with 645,708 autosomal SNPs were genotyped using Affymetrix 'Japonica Array'. All genotyped samples passed the recommended sample quality control metric for the AXIOM arrays (dish QC > 0.82). One control sample with an overall call rate < 97% was excluded. Then the remaining 642 samples were recalled. After the second round of genotype calling, all the samples had an overall call rate ≥ 97%.

SNP QC was conducted using the following thresholds: SNP call rate ≥ 97% (21,199 SNPs were excluded), MAF ≥ 5% (127,207 SNPs were excluded), and HWE *p*-value ≥ 0.0001 in healthy controls (1,407 SNPs were excluded).

No subject was excluded in IBD test. In PCA, when plotted samples with 3 global populations (CEU, YRI and JPT+CHB), no outlier was detected (Figure 2A). Further PCA was conducted using my dataset together with JPT individuals and CHB individuals, 6 healthy controls were located outside of the main Japanese cluster and identified as outliers (Figure 2B). After the removal of the 6 outliers, there was no obvious population stratification between the cases and controls in discovery stage (Figure 2C).

Association analyses in the discovery GWAS

Association analyses of typed SNPs

After stringent QC and filtering steps, 224 cases and 412 healthy controls with 495,895 autosomal SNPs were retained for association analysis. The inflation factor, λ, was 1.045 for all tested SNPs (Figure 3). Considering the 4% of inflation was not negligible, I adjusted the *P*-

values using genomic control (GC). Figure 4 presents the genome-wide view of the SNP association data (GC-corrected *P*) using Cochran-Armitage trend test. The *HLA-DR/DQ* region on chromosome 6 exhibited the most significant association with childhood SSNS in the initial GWAS (rs4642516, 21kb 5' of *HLA-DQB1*, GC-corrected $P=5.44\times10^{-10}$)(Figure 5). No other locus outside the *HLA* region reached genome-wide significance (GC-corrected $P<5\times10^{-8}$).

Association analyses after whole-genome imputation

After whole-genome imputation, there were 32,509,374 autosomal SNPs and short INDELs with info score > 0.5. QC was conducted using the following thresholds: SNP/short INDEL call rate ≥ 97% (980,814 SNPs/short INDELs were excluded), MAF ≥ 5% (27,422,486 SNPs/short INDELs were excluded), and HWE test *p*-value ≥ 0.0001 in healthy controls (531 SNPs/short INDELs were excluded). After QC, 224 cases and 412 controls with 4,105,543 autosomal SNPs and short INDELs were included in association analyses. The λ value was 1.023 for all tested variations (Figure 6A) and decreased to 1.019 when SNPs/short INDELs in *HLA* region (Hg19: chr6: 29,691,116–33,054,976) were excluded (Figure 6B). GC correction was performed to adjust the *P*-value of each variant (GC-corrected *P*).

Manhattan plot of variants after whole-genome imputation was shown in Figure 7. Besides the genome-wide significant association detected in *HLA-DR/DQ* region, two loci on chromosome 5 achieved the suggestive significance (GC-corrected $P<1\times10^{-5}$) using Cochran-Armitage trend test.

For further validation and replication, in total 10 SNPs/short INDELs were selected based on the association analysis after whole-genome imputation. In *HLA* region, 6 SNPs/short INDELs in high LD ($r^2>0.8$) at the top hit were picked up. Another 4 SNPs from 2 candidate loci

outside *HLA* region (1 genotyped SNP and 1 imputed SNP with minimum *P*-value in each candidate locus) were selected (Table 3).

Next I performed conditional analysis in *HLA* region, to clarity if there is secondary disease-associated gene (locus) in *HLA* region. 18,747 variants in *HLA* region were extracted and conditional analysis was conducted using logistic regression by PLINK 1.9. After conditioning on the top typed SNP in *HLA-DR/DQ* region (rs4642516), there was no secondary signal in *HLA* region after Bonferroni correction (significance threshold conditional $P< 0.05/18,747=2.67 \times 10^{-6}$) (Figure 8).

Validation of candidate variants

Genotyping for candidate SNPs/short INDELs was performed by DigiTag2 assay. Genotyping failed for four SNPs/short INDELs in *HLA-DR/DQ* region (rs79193269, rs9275103, rs9282090 and rs9275101) and one SNP outside *HLA* region (rs4865828). 5 candidate SNPs were genotyped successfully, with mean concordance rate as 97% (90-100%).

Replication and combined analysis of candidate variants

In the replication stage, the associations of rs4642516 and rs3134996 in *HLA-DR/DQ* region were confirmed in an independent sample set using Cochran-Armitage trend test (rs4642516, $P=6.69 \times 10^{-13}$; rs3134996, $P=9.44 \times 10^{-10}$). Other candidate loci did not achieve significance after Bonferroni correction (significance threshold *P*-value < 0.05/5=0.01)(Table 3).

To further evaluate the appropriate genetic model for the two replicable variants, I performed the association analyses under dominant, recessive and allelic model using Fisher's exact test, separately (Table 4). rs4642516 and rs3134996 showed highly significant associations

under allelic and dominant models [rs4642516, $P = 4.51 \times 10^{-15}$, OR=0.30 (0.22-0.42) under allelic model, $P = 3.43 \times 10^{-13}$, OR= 0.28 (0.20-0.41) under dominant model; rs3134996, $P = 1.26 \times 10^{-15}$, OR=0.30 (0.22-0.41) under allelic model, $P = 6.24 \times 10^{-11}$, OR=0.29 (0.19-0.43) under dominant model].

Combined analysis was done using typed data in discovery GWAS and replication stage. rs4642516 [minor allele G, combined *P*-value=7.84×10^{-23}, OR=0.33 (0.26-0.41) under allelic model; combined *P*-value=4.87×10^{-23}, OR=0.29 (0.22-0.38) under dominant model; combined *P*-value=2.40×10^{-7}, OR=0.17 (0.07-0.39) under recessive model] and rs3134996 [minor allele A, combined *P*-value=1.72×10^{-25}, OR=0.29 (0.23-0.37) under allelic model; combined *P*-value=9.57×10^{-21}, OR=0.28 (0.21-0.38) under dominant model; combined *P*-value=1.31×10^{-10}, OR=0.23 (0.14-0.39) under recessive model] showed significant associations with the disease (Table 5).

Discussion

To identify genetic variants influencing childhood SSNS, I performed a two-stage GWAS using Japanese-specific SNP array 'Japonica Array' and followed by whole-genome imputation in the Japanese population. This study is the first GWAS for childhood SSNS patients using genome-wide SNP array based on LD block structures. The most significant association was detected in *HLA-DR/DQ* region, and the top SNPs were further replicated in an independent Japanese sample set [rs4642516 (21kb 5'of *HLA-DQB1*, allelic *P*-combined=7.84×10^{-23}, OR=0.33; dominant *P*-combined=4.87×10^{-23}, OR=0.29) and rs3134996 (705bp 5' of *HLA-DQB1*, allelic *P*-combined=1.72×10^{-25}, OR=0.29; dominant *P*-combined=9.57×10^{-21}, OR=0.28)]. In the current experiences, majority of the disease-associated SNPs mapped by GWASs were located in intergenic (43%) or intronic (45%) regions[57]. These non-coding SNPs may play an important role in the regulatory network[58]. Here, the replicated SNPs, rs4642516 and rs3134996, were previously reported as expression quantitative trait loci (eQTL) correlated with the expression of many genes including *HLA-DQA2/-DQB1/-DQB2/-DRB1* in various tissues, as well as in glomerulus and/or tubulointerstitium tissues (GTEx database and NephQTL browser).

The strong association between *HLA* genes, especially *HLA-DR/-DQ* genes, and the susceptibility to INS has been reported in different populations, including the Japanese population[25-42]. However, most of the previous genetic studies were conducted using candidate gene approaches with limited numbers of samples. For the first time, the present study confirmed the predominant role of *HLA-DR/DQ* region in the pathogenesis of Japanese childhood SSNS by a hypothesis-free method.

Previously, an exome array association study identified 4 exome-wide significant variants in or around *HLA-DQA1* and -*DQB1* in South Asian children with SSNS. Two missense variants in *HLA-DQA1* (C34Y [rs1129740, allele A], allelic $P=1.187\times10^{-6}$, OR=2.11; F41S [rs1071630, allele G], allelic $P=1.187\times10^{-6}$, OR=2.11) were further replicated in children of European ancestry[42]. As genome-wide approaches without prior hypothesis, both the current GWAS and the exome-wide association study identified the most significantly associated variants in *HLA-DR/DQ* region, providing solid evidence that the *HLA-DR/DQ* region plays the predominant role in the susceptibility to childhood SSNS. Although the two reported SNPs were not included in Japanese-specific genotyping array or whole-genome imputation reference, another two SNPs that are in LD with the missense variants achieved suggestive significant associations in discovery stage of present GWAS (rs9272346, $r^2=0.76$, D'=1 with rs1129740, allele A, GC-corrected $P= 9.54\times10^{-8}$, OR=1.97 under allelic model; rs1063355, $r^2=0.65$, D'=1 with rs1071630, allele G, GC-corrected $P=1.60\times10^{-7}$, OR=1.93 under allelic model), indicating that the susceptibility loci within the *HLA* region detected by SNP-based analyses may be shared across different populations.

Another published GWAS was conducted in Japanese adult-onset NS patients including both primary and secondary NS, aiming to identify the genetic variants associated with severe proteinuria in NS patients[44]. A SNP rs16946160 (allele A, $P = 6.0\times10^{-11}$, OR=2.54 under recessive model) in *GPC5* was reported as a significant genetic factor contributing to the common pathway of massive proteinuria in acquired NS, whereas the *HLA* region did not show significant association. Although both studies were performed in the same ethnic group, the reported SNP was not replicated in Japanese children with SSNS in the current GWAS (rs16946160, P=0.60 under recessive model in the discovery stage). One possible explanation is

the different disease spectra in the two Japanese GWASs. In the present study, I focused on pediatric patients with primary NS and who responded to steroid treatment only, in order to decrease the potential confounding factors brought by disease heterogeneity. About half of patients in this study were undertaken renal biopsy (248 in 440 patients), since renal biopsy is not the clinical routine for pediatric patients with INS. As the most common pathologic type in childhood SSNS, MCD accounted for 91.5% of the 248 patients. However, addressing the question of underlying genetic susceptibility factors shared by NS patients with different causes, the GWAS of Japanese adult-onset NS included both primary and secondary patients with various pathological types. MCD only accounted for 18% of cases in the GWAS of adult NS. Therefore, the different disease backgrounds in patients led to the different disease-associated genes identified in the two Japanese GWASs.

In non-*HLA* regions, although two loci with suggestive significant associations were observed on chromosome 5 in the discovery stage, the significant associations were not confirmed in the replication sample set. The power of this two-stage GWAS exceeded 80% to detect common alleles (MAF >= 5%) with genotypic relative risk (RR) above 2.95, at a significant *P*-value threshold of 5×10^{-8} (Figure 9). The current sample size put limits on the ability to detect associated variants with modest associations. Although this study did not reveal any other disease-associated genetic factor outside of *HLA* region, it does point out that *HLA-DR/DQ* genes are the strongest genetic factors associated with Japanese childhood SSNS, and other potentially associated common variants in non-*HLA* regions could have moderate disease associations. GWAS with larger sample size is necessary for detecting more candidate loci associated with Japanese childhood SSNS (especially for non-*HLA* genes).

Chapter 2:

Fine-mapping of *HLA* association

Introduction

The human leukocyte antigen (*HLA*) region is a gene complex located on chromosome 6p21, encoding the major histocompatibility complex (MHC) proteins in humans. The classical *HLA* region spans 3.6 megabases (Mb) and encompasses more than 200 genes, playing important roles in the regulation of immune system as well as some fundamental cellular processes[59]. This region can be subdivided into classical class I, classical class III and classical class II region. *HLA* region is known as the most polymorphic genetic system in humans, which means there are many different alleles in a population. *HLA* region is characterized by strong LD extending across *HLA* genes and non-*HLA* genes, making it challenging to ascertain the primary disease-risk *HLA* gene(s).

The strong association between *HLA* genes, especially *HLA-DR/-DQ* genes and the susceptibility to INS has been reported in different populations, including Japanese patients[25-42]. However, prior works on the associations between *HLA* genes and NS have the following limitations: First, most of the studies had relatively small sample sizes (mostly < 100 patients). Second, limited *HLA* alleles were tested or only low-resolution *HLA* types were determined. Third, there were very few *HLA* allele/haplotype analyses focusing on INS in the Japanese population.

Using SNP-based association analyses, the most significant association with Japanese childhood SSNS was detected in *HLA-DR/DQ* region and the associations of top SNPs (rs4642516, 21kb 5'of *HLA-DQB1*; rs3134996, 705bp 5' of *HLA-DQB1*) were further confirmed in an independent Japanese sample set. However, since SNPs in *HLA-DR* region do not conform to the Hardy-Weinberg equilibrium because of the marked copy number polymorphism of *HLA-*

DRB genes, commercially available SNP genotyping arrays are unable to analyze SNPs in *HLA-DR* region. Therefore, even the *HLA-DQ* region may appear to be the most significantly associated genetic factor governing disease susceptibility in SNP-based GWASs, *HLA-DR* region, which is in strong LD with *HLA-DQ* region, must also be considered. *HLA* fine-mapping is necessary to clarify the disease-associated *HLA* genes and alleles[60].

Since 4-digit *HLA* alleles correspond to specific *HLA* molecules, high-resolution HLA imputation and typing at 4-digit level were conducted in this study, aiming to dissect the *HLA* association detected by SNP-based analysis and provide a better understanding of disease mechanism at molecular and functional levels.

Materials and Methods

HLA imputation

Four-digit HLA imputation was done using typed SNP data by R package 'HIBAG'[61]. After SNP quality control (SNP call rate ≥97%, MAF ≥5%, and HWE test *p*-value ≥0.0001 in healthy controls), SNP data for each individual was extracted from an extended MHC region, ranging from 25,759,242 to 33,534,827 base pair (bp) based on the hg19 position. Our department in-house Japanese imputation reference was used for the imputation of *HLA* genotype[62]. After HLA-imputation, post QC was conducted using a call-threshold (CT) > 0.4[62].

In the discovery stage, six classical *HLA* genes in both class I and class II regions (*HLA-A*, *-C*, *-B*, *-DRB1*, *-DQB1*, and *-DPB1*) were imputed in 224 cases and 412 healthy controls. *HLA* haplotypes were determined using Arlequin algorithm (Arlequin version 3.5.2.2)[63].

In the replication stage, HLA-imputation was performed for *HLA-DRB1* and *-DQB1* using the same approach in 269 healthy controls whose *HLA* genotypes were not determined by direct genotyping. SNP genotyping was performed in these 269 controls using Affymetrix Axiom Genome-Wide ASI 1 Array previously[64]. The SNPs passed QC using the same criteria (SNP call rate ≥ 97%, MAF ≥ 5%, and HWE test *p*-value ≥ 0.0001 in healthy controls) were extracted for HLA imputation.

HLA genotyping

Genotyping of *HLA-DRB1* and *-DQB1* was performed using the PCR-SSO (sequence-specific oligonucleotide probing) method on Luminex platform with the WAKFlow *HLA* typing

kit (Wakunaga, Hiroshima, Japan). Briefly, target DNA was amplified by PCR with biotinylated primers specifically designed for each *HLA* locus. The PCR product was then denatured and hybridized to complementary oligonucleotide probes immobilized on fluorescently coded microsphere beads. At the same time, biotinylated PCR product was labeled with phycoerythrin-conjugated streptavidin and finally examined using the Luminex system. Genotype determination and data analysis were performed automatically using the WAKFlow typing software.

To evaluate the accuracy of HLA-imputation, direct *HLA-DRB1/-DQB1* typing was conducted in 224 cases in the discovery set. 409 controls in the discovery set were genotyped previously using the same method. The typed data was used as validation.

The *HLA-DRB1/-DQB1* genotype data of 450 controls genotyped using the same method was obtained from published studies[65, 66] and used here.

Genotyping of single-tag SNP rs3129888

Genotyping was performed for the reported single-tag SNP rs3129888 by DigiTag2 assay following the experiment protocol. Genotyping was performed in 224 cases in discovery stage to compare the concordance with the genotype determined by 'Japonica Array' typing. Then genotyping was done in the replication sample set including 216 cases and 719 controls.

Statistical methods and software

SNP analyses were done by PLINK 1.9[47]. HLA-imputation was performed by R package 'HIBAG'[61]. *HLA* haplotypes were determined using the Arlequin algorithm (Arlequin version 3.5.2.2)[63]. *HLA* allele and haplotype frequencies were compared between case and

control groups. *P*-values were calculated using the Pearson's chi-square test or Fisher's exact test in presence vs. absence of each *HLA* allele/*HLA* haplotype. *HLA* alleles/haplotypes with frequencies < 0.5% in cases or controls were excluded from the association analyses. Then, Bonferroni correction was performed by the standard method in which *P*-values were corrected for the number of alleles/haplotypes tested in each analysis [shown as *P*-corrected (*Pc*)]. An association was considered to be significant when the *P*-value was <0.05 after correction for multiple comparisons. For the evaluation of disease variance explained in this study, logistic regression and the calculation of Nagelkerke's pseudo-R^2 were done by R programming and R package 'rcompanion'.

Results

HLA-imputation and analyses in the discovery stage

HLA-imputation in discovery sample set

In the discovery stage, six classical *HLA* genes (*HLA-A, -C, -B, -DRB1, -DQB1,* and *-DPB1*) were imputed in 224 cases and 412 healthy controls. A total of 197 cases and 411 controls passed post-imputation QC (CT>0.4) and were included in subsequent analyses.

Association analyses of HLA alleles

Significantly associated *HLA* alleles were identified in all six *HLA* genes after multiple corrections (Table 6A to 6F and Table 8). *HLA-A*02:06* (Pc=8.84×10^{-3}, OR=1.91), *HLA-DRB1*08:02* (Pc=2.56×10^{-4}, OR=2.79), and *HLA-DQB1*03:02* (Pc=4.14×10^{-4}, OR=2.08) were highly associated with the susceptibility to Japanese childhood SSNS. *HLA-A*33:03* (Pc=2.43×10^{-3}, OR=0.30), *HLA-B*44:03* (Pc=1.25×10^{-4}, OR=0.13), *HLA-C*14:03* (Pc=5.25×10^{-5}, OR=0.13), *HLA-DRB1*13:02* (Pc=7.47×10^{-4}, OR=0.18), and *HLA-DQB1*06:04* (Pc=3.40×10^{-5}, OR=0.07) showed significantly protective effects.

Association analyses of HLA haplotypes

HLA haplotypes were determined using imputed alleles by Arlequin algorithm and significantly associated haplotypes were detected after multiple corrections (Table 7A to 7D and Table 8). *HLA-DRB1*08:02-DQB1*03:02* (Pc=4.34×10^{-6}, OR=4.33) and *HLA-DRB1*13:02-DQB1*06:04* (Pc=6.17×10^{-5}, OR=0.07) demonstrated the most significant positive and negative association, respectively. *HLA-DRB1*08:02-DQB1*03:02* even showed a stronger and more robust association than *HLA-DRB1*08:02* or *HLA-DQB1*03:02* alone. The two significant

HLA-DRB1-DQB1 haplotypes (*HLA-DRB1*08:02-DQB1*03:02* and *HLA-DRB1*13:02-DQB1*06:04*) showed more significant associations than their related *HLA-DRB1-DQB1-DPB1* haplotypes (*HLA-DRB1*08:02-DQB1*03:02-DPB1*05:01*, $Pc=3.91\times10^{-2}$, OR=2.88; *HLA-DRB1*13:02-DQB1*06:04-DPB1*04:01*, $Pc=9.52\times10^{-3}$, OR=0.11).

Validation of *HLA-DRB1/-DQB1*

HLA genotyping was performed for *HLA-DRB1/-DQB1* in 224 cases in the discovery stage successfully. The concordance rate between the imputed (after post-imputation QC using CT>0.4) and genotyped *HLA-DRB1/-DQB1* genotypes was 99.2%, indicating the HLA-imputation using Japanese-specific reference achieved a high accuracy.

Replication of *HLA-DRB1/-DQB1*

Replication of HLA-DRB1/-DQB1

In the replication stage, HLA genotyping for *HLA-DRB1* and *HLA-DQB1* was done in 213 of the 216 cases successfully. HLA-DRB1 genotyping failed in 3 cases. The genotyped data of 450 controls in the replication sample set were obtained previously using the same approach [65, 66]. High-accuracy HLA-imputation was performed for *HLA-DRB1* and *-DQB1* in 269 healthy controls whose *HLA* genotypes were not determined by direct genotyping. 260 of the 269 controls passed the post-imputation QC (CT>0.4) and were included in the replication data set.

Almost all the significant *HLA-DRB1* and *-DQB1* alleles and *HLA-DRB1-DQB1* haplotypes identified in the discovery stage were replicated except for *HLA-DQB1*03:02* (Table 9A to 9C and Table 10). Although *HLA-DQB1*03:02* did not show significance in the replication stage (*P*-value=3.07×10^{-1} before multiple correction), *HLA-DRB1*08:02-*

*DQB1*03:02* was still the most significant susceptibility haplotype in the same sample set (Pc=1.25×10^{-4}, OR=3.15) (Table 10).

Combined analyses and identification of primary HLA alleles

Combined analyses were performed using the *HLA-DRB1/-DQB1* genotype data in the validation and replication steps (Table 9A to 9C and Table 10). *HLA-DRB1*08:02* (Pc=1.82×10^{-9}, OR=2.62), *HLA-DRB1*13:02* (Pc=7.31×10^{-11}, OR= 0.16), and *HLA-DQB1*06:*04 (Pc=2.09×10^{-12}, OR=0.10) were identified as the most significantly associated *HLA* alleles. *HLA-DRB1*08:02-DQB1*03:02* (Pc=7.01×10^{-11}, OR=3.60) and *HLA-DRB1*13:02-DQB1*06:04* (Pc=4.18×10^{-12}, OR=0.10) were highly associated with childhood SSNS.

Considering the tight LD between *HLA-DRB1* and *-DQB1* gene, I performed the reciprocal conditional analyses using logistic regression in combined dataset, aiming to elucidate which allele in the *HLA-DRB1-DQB1* haplotype has independent effect on disease. When *HLA-DQB1*03:02* was conditioned on, the significance of *HLA-DRB1*08:02* decreased from P=2.88×10^{-10} to conditional P=3.13×10^{-7}, but still show robustly independent effect. However, when I conditioned on *HLA-DRB1*08:02*, the significance of *HLA-DQB1*03:02* disappeared (P=5.62×10^{-5} to conditional P=0.10). Similarly, *HLA-DQB1*06:04* showed some independent effect when *HLA-DRB1*13:02* was conditioned on (P=2.63×10^{-9} to conditional P=7.22×10^{-5}), while *HLA-DRB1*13:02* didn't show any significance when *HLA-DQB1*06:04* was conditioned on (P=2.32×10^{-9} to conditional P=0.37). Both *HLA-DRB1*08:02* (P=2.88×10^{-10} to conditional P=1.14×10^{-8} when *HLA-DQB1*06:04* was conditioned on) and *HLA*DQB1*06:04* (P=2.63×10^{-9} to conditional P=8.47×10^{-9} when *HLA-DRB1*08:02* was conditioned on) still showed strikingly significant associations in reciprocal conditional analyses (Table 11).

Single-tag SNP for capturing *HLA-DRB1*08:02* and *HLA-DRB1*08:02-DQB1*03:02*

rs3129888 (G) was reported as a tag SNP for *HLA-DRB1*08:02* and *HLA-DRB1-DQB1* haplotypes carrying *HLA-DRB1*08:02* (sensitivity was 92.3% and specificity was 98.9%) in the Japanese population[67, 68].

Genotyping of rs3129888 using DigiTag2 assay was performed in 224 cases in the discovery sample set as validation (concordance rate=100%) and the replication sample set including 216 cases and 719 healthy controls successfully.

To check the sensitivity and specificity of the tag SNP, I conducted the analyses using the combined data sets of *HLA-DRB1* alleles and *HLA-DRB1-DQB1* haplotypes including 437 cases and 1,119 controls, separately. In *HLA-DRB1* allele dataset, 177 of 184 *HLA-DRB1*08:02* had the G allele of rs3129888 (sensitivity 96.2%), and 2,865 of 2,928 other *HLA-DRB1* sub-alleles had the A allele of rs3129888 (specificity 97.8%) (Table 12A). For *HLA-DRB1*08:02-DQB1*03:02*, 110 of 112 *HLA-DRB1*08:02-DQB1*03:02* had the G allele of rs3129888 (sensitivity 98.2%), and 2,870 of 3,000 other *HLA-DRB1-DQB1* haplotypes had the A allele of rs3129888 (specificity 95.7%) (Table 12B).

Variance explained by *HLA* association

Logistic regression and Nagelkerke's pseudo-R^2 were used to measure the proportion of variance explained by genetic factor(s). To use informative and independent genetic marker, I included the top SNP in the *HLA* region (rs4642516), which was detected in the discovery stage and replicated in an independent sample set. The results of the conditional analysis in the *HLA* region demonstrated that there was no other SNP showing an independent effect in the *HLA*

region. Logistic regression was performed in the replication dataset including 213 cases and 710 controls. Case-control status was used as outcome while the genotype of rs4642516 was used as covariate in the logistic regression. 9.7% of variance (Nagelkerke's pseudo-R^2) in Japanese childhood SSNS can be explained by rs4642516. Considering the complicated LD structure within *HLA* region, there may be strong intercorrelations between the index SNP in GWAS and the two independent classical *HLA* alleles (rs4642516 and *HLA-DRB1*08:02*, r^2=0.02, D'=1; rs4642516 and *HLA-DQB1*06:04*, r^2=0.21, D'=1 in combined dataset). I also conducted the logistic regression using the two independent classical *HLA* alleles, *HLA-DRB1*08:02* and *HLA-DQB1*06:04*, in the replication dataset. *HLA-DRB1*08:02* and *HLA-DQB1*06:04* explained 8.1% of the variance in childhood SSNS in the Japanese population (Nagelkerke's pseudo-R^2).

Discussion

In the first chapter, the strong association between *HLA-DR/DQ* region and Japanese childhood SSNS was identified by SNP-based analyses. *HLA-DR/-DQ* genes belong to *HLA* class II genes, encoding the HLA-DR, HLA-DQ molecules, respectively. HLA-DR/-DQ molecules play crucial roles in the process of antigen presentation in the immune reaction.

However, in the post-GWAS era, interpretation of *HLA*-association is still challenging. Based on the design of GWAS typing arrays, only tag SNPs which can capture most common SNPs of the genome through LD were considered and genotyped, the SNP-disease association detected by GWASs dose not provide the causality link directly. The associated SNP may only tag the causal variants, and less likely is the causal variant itself. In *HLA*-associated diseases, the situation could be more complicated because of the high gene density, extreme polymorphism, strong and complicated LD structures within the *HLA* region as well as the lack of SNPs in *HLA-DR* region in commercial available arrays[60, 69]. Since current knowledge for the immunogenetic background of INS is still limited, clarifying the disease-associated allele(s)/haplotype(s) by HLA fine-mapping is important for improving our understanding of disease mechanism from the molecular and functional aspects, and may have huge impact on the individualized treatment. In this chapter, high-accuracy HLA-imputation and direct HLA-typing were conducted and *HLA* alleles/haplotypes associated with Japanese childhood SSNS were elucidated.

In the previous Japanese study, *HLA-DRw8* was reported as a risk factor associated with Japanese adult NS patients with MCD[40]. Consistently, in the present study, *HLA-DRB1*08:02* exhibited a significant positive association with Japanese childhood SSNS patients and was further determined as a primary susceptibility factor by reciprocal conditional analyses. The

increased allele frequency of *HLA-DQB1*03:02* was reported in Japanese pediatric patients with SSNS[27], and the same tendency was also observed in the present study (14.6% in SSNS cases vs. 9.5% in healthy controls). Since *HLA-DRB1*08:02* was concluded as the primary genetic factor, the increased allele frequency of *HLA-DQB1*03:02* is likely to be caused by the tight LD with *HLA-DRB1*08:02*.

HLA haplotype associated with INS was rarely reported in previous studies. In this study, *HLA-DRB1*13:02-DQB1*06:04* and *HLA-DRB1*08:02-DQB1*03:02* were identified as the most significant protective and susceptibility haplotype, respectively. In *HLA* allele analyses, the strong and complicated LD structures within *HLA* region make it challenging to clarify the primary genetic factor. Here, the results given by reciprocal conditional analyses supported that *HLA-DQB1*06:04* and *HLA-DRB1*08:02* was the primary genetic factor in the related haplotype, respectively. More importantly, *HLA-DRB1*08:02-DQB1*03:02* showed a stronger and more significant association than *HLA-DRB1*08:02* or *HLA-DQB1*03:02* alone. There could be two possible hypotheses: First, this associated *HLA* haplotype is a 'marker' for the susceptibility gene/locus by virtue of statistically significant LD, tagging the true susceptibility variants better than single allele. Second, the susceptibility haplotype *HLA-DRB1-DQB1* itself may play a more important role than single *HLA-DRB1/-DQB1* allele in the pathogenesis of Japanese childhood SSNS[69]. The HLA molecules encoded by these two alleles might work together or have some functional interaction in the disease mechanism.

For autoimmune diseases with a main association signal coming from *HLA* class II region, such as type 1 diabetes, celiac disease, multiple sclerosis, systemic lupus erythematosus, inflammatory bowel disease and narcolepsy, the reported variance explained by *HLA* alleles varies from 2-58%[70, 71]. In the current study, 9.7% (Nagelkerke's pseudo-R^2) of the disease

variance in Japanese childhood SSNS was explained by rs4642516 (the top SNP in *HLA* region), while *HLA-DRB1*08:02* and *HLA-DQB1*06:04* explained 8.1% of the disease variance. The less disease variance explained by the two *HLA* alleles compared to that explained by the top GWAS hit (rs4642516) indicated that, there could be more independent *HLA* alleles contributing to the disease pathogenesis.

Interestingly, consistent with previous researches in the Japanese population, single-tag SNP rs3129888 works with high sensitivity and specificity for capturing *HLA-DRB1*08:02* and *HLA-DRB1*08:02-DQB1*03:02* in this study. Single-tag SNP rs3129888 could be used as an cost-effective biomarker to capture the susceptibility genetic factors in clinical practice.

Not surprisingly, the significantly associated *HLA* alleles detected in this Japanese study are different from previous studies in other populations. The increased frequency of *HLA-DR7* antigen in INS has been reported in several populations[25, 26, 29, 30, 32, 34, 38, 39], but it was not replicated in the previous Japanese study[40] and the current study. Based on the reported frequencies in the Allele Frequency Net Database (http://www.allelefrequencies.net)[72], *HLA-DR7* is common in European, African, Chinese, and Saudi Arabian populations, with an allele frequency of 10.3-18%[73-77]. By contrast, the frequency of *HLA-DR7* is only 0.3-0.8% in the Japanese population[78, 79]. On the contrary, as the significant susceptibility allele in the present study, *HLA-DRB1*08:02* is common in Japanese (4.2%)[80], while rare in European and Chinese populations (0-0.9%)[81] and relatively rare in South Koreans (2.4%)[82]. The susceptibility haplotype *HLA-DRB1*08:02-DQB1*03:02* seems to be extremely rare in European, since there is no haplotype frequency reported. These observations indicated that the disease-associated *HLA* alleles may vary depending on geographic and ethnic origins. Given the diversity seen in *HLA* region, analyzing the disease associated *HLA* allele/haplotype in certain

population/area at high-resolution level is necessary to get a better understanding of the disease mechanism.

The current study not only provided the solid evidence for the predominant role of *HLA-DR/DQ* region in the pathogenesis of Japanese childhood SSNS, as well as clarified the disease associated *HLA* alleles and haplotypes, but also gave a particularly vivid example to emphasize the importance of HLA fine-mapping after SNP-based analysis.

An exome array association study of childhood SSNS published in 2015 identified four variants in or around *HLA-DQA1* and *-DQB1* with exome-wide significance in South Asian children with SSNS. Two missense variants in *HLA-DQA1* (C34Y [rs1129740], allelic P=1.187×10^{-6}, OR=2.11; F41S [rs1071630], allelic P=1.187×10^{-6}, OR=2.11) were further replicated in children of European ancestry[42]. Although the two reported SNPs were not included in the current study, another two SNPs that are in LD with the missense variants achieved suggestive significant associations in discovery GWAS (rs9272346, r^2=0.76, D'=1 with rs1129740, GC-corrected P= 9.54×10^{-8}, OR=1.97 under allelic model; rs1063355, r^2=0.65, D'=1 with rs1071630, GC-corrected P=1.60×10^{-7}, OR=1.93 under allelic model), indicating that the susceptibility loci within the *HLA* region detected by SNP-based analyses may be shared across different populations. Notably, refinement of the *HLA-DQA1* association was recently done by HLA-imputation using exome array data in the same South Asian samples[43]. *HLA-DRB1*07:01* (P=4.6×10^{-6}, OR=2.510), *HLA-DQA1*02:01* (P=5.5×10^{-6}, OR=2.492), *HLA-DQB1*02:01* (P=1.0×10^{-6}, OR=2.282) and *HLA-DQA1*01:01* (P=3.3×10^{-4}, OR=0.465) were significantly associated with childhood SSNS. However *HLA-DRB1*07:01* and *HLA-DQB1*02:01* were not replicated in the current study, because they are at a low allele frequency in the Japanese population [*HLA-DRB1*07:01* (0.375%), *HLA-DQB1*02:01* (0.134%),

http://hla.or.jp]. Considering the LD patterns and allele frequencies in *HLA* region are highly differentiated across populations and areas, it is essential to clarify the disease-associated and the primary *HLA* alleles/haplotypes though HLA fine-mapping after SNP-based association analysis. High-resolution HLA-typing and imputation techniques at 4-digit level can provide more information for investigating presumable molecular mechanism of disease pathogenesis as compared with SNP-based results. HLA-imputation using dense SNP data with population-specific reference provides a fast, accurate and cost-effective approach to examine *HLA*-disease associations in large-scale studies, largely improving our immunogenetic knowledge of diseases.

Nowadays, interpretation of *HLA*-disease associations in term of the functions of variants has been problematic because of several reasons[69]. First, because of the existence of extended LD within *HLA* region, ascertaining the primary *HLA* gene(s) is difficult. Moreover, *HLA*-associated diseases can be the result of the combination of different *HLA* molecules encoded by various loci rather than the results of one *HLA* variant only[83]. Second, almost all of the *HLA*-associated diseases are multifactorial polygenic diseases. Disease-associated *HLA* allele(s) may work together with other genetic variants as well as environmental factors in the disease mechanisms. However, our knowledge of associated genetic variants for a given disease is still limited, and it also challenging to evaluate or control the possible environmental factors in current genetic analyses. For example, although infection history, allergen exposure, exposures to lead or heavy metals are reported as associated environmental factors contributing to the onset, relapse and course of NS, the previous studies are mostly cross-sectional studies or prospective studies with limited follow-up, the roles of these exposures in NS have not been explicitly identified[84]. It is still difficult to interpret the disease mechanism considering the potential pathways activated by different pathogens and allergens. Third, in most *HLA*-associated diseases

(for example MCD, the most common INS), the existence of autoantigens or the disease-relevant antigens are largely unknown, which prevents the three-dimensional analysis of *HLA*-peptide interactions.

Summary and Conclusions

This is the first GWAS of childhood SSNS using genome-wide SNP array based on LD block structures. The most significant association was detected in *HLA-DR/DQ* region in the Japanese population and the top SNPs were replicated in an independent Japanese sample set.

The *HLA* association with Japanese childhood SSNS was further dissected by HLA fine-mapping. *HLA-DRB1*08:02-DQB1*03:02* and *HLA-DRB1*13:02-DQB1*06:04* were identified as the most significant susceptibility and protective haplotype to the disease, respectively. *HLA-DRB1*08:02* and *HLA-DQB1*06:04* were further identified as the primary factor in each related haplotype. Moreover, *HLA-DRB1*08:02-DQB1*03:02* showed a stronger and more significant association to the disease susceptibility than single allele alone, providing the possible hypothesis that the haplotype itself may play a more important role than single *HLA-DRB1/-DQB1* allele in the disease pathogenesis, or this associated *HLA* haplotype tags the true susceptibility variants better than single allele. Interestingly, rs3129888 works as single-tag SNP for capturing *HLA-DRB1*08:02* and *HLA-DRB1*08:02-DQB1*03:02* with high sensitivity and specificity. rs3129888 could be used as an cost-effective biomarker to capture the susceptibility genetic factors in clinical practice.

The present study demonstrated the predominant role of *HLA-DR/DQ* locus in the pathogenesis of Japanese patients with childhood SSNS. In the current study, rs4642516 (the top SNP in *HLA* region) explained 9.7% (Nagelkerke's pseudo-R^2) of the disease variance in the Japanese childhood SSNS. However, this study did not reveal any other association in non-*HLA* regions. As a two-stage GWAS, the current sample size puts limits on the ability of detecting genetic variants with modest associations. Larger-scale GWAS in the Japanese population is still

needed in the future, with the purpose to identify more disease-associated loci (especially in non-*HLA* regions). Common genetic factors with moderate associations, and disease-associated rare variants (MAF<5% in the population) may be detected in GWAS with larger sample size. It is also meaningful to conduct replication in other populations using the top SNPs detected in *HLA* region in this GWAS. Functional study focused on associated *HLA* allele/haplotype could improve our knowledge of the role of *HLA* genes in the pathogenesis of childhood SSNS. It would be helpful to compare the structures and functions of disease-associated *HLA* alleles/haplotypes detected in this study, with the ones reported in other populations. Clarifying the common mechanism shared by different population-specific disease-associated *HLA*-alleles/haplotypes is important for the understanding of disease pathogenesis.

Supplementary

Hospitals and universities contributing to the sample collections for the current study:

Hokkaido University Hospital, Sapporo.

Obihiro Kyokai Hospital, Obihiro.

Nikko Memorial Hospital, Muroran.

Kushiro Red Cross Hospital, Kushiro.

Sapporo Kosei Hospital, Sapporo.

Obihiro Kosei Hospital, Obihiro.

Oji General Hospital, Tomakomai.

Hakodate Goryoukaku Hospital, Hakodate.

Hirosaki University Hospital, Hirosaki.

Tohoku University Graduate School of Medicine, Sendai.

Tokyo Medical and Dental University, Tokyo.

Musashino Red Cross Hospital, Musashino.

National Center for Child Health and Development, Tokyo.

Tokyo Women's Medical University, Tokyo.

Tokyo Kita Medical Center, Tokyo.

Faculty of Medicine, University of Yamanashi, Yamanashi.

Fujiyoshida Manucipal Hospital, Yamanashi.

Dokkyo Medical University School of Medicine, Tochigi.

Tokyo Metropolitan Children’s Medical Center, Tokyo.

National Hospital Organization Kofu National Hospital, Kofu.

Kanazawa Medical Center, Kanazawa.

Japanese Red Cross Fukui Hospital, Fukui.

Kobe University Graduate School of Medicine, Kobe.

Hyogo Prefectural Kobe Children's Hospital, Kobe.

Hyogo College of Medicine, Nishinomiya.

Takatsuki General Hospital, Takatsuki.

Kakogawa Central City Hospital, Kakogawa.

Osaka City General Hospital, Osaka.

Osaka Medical College, Osaka.

Shiga University of Medical Science, Shiga.

Wakayama Medical University, Wakayama.

Kagawa Prefecture Central Hospital, Takamatsu.

Uwajima City Hospital, Uwajima.

Fukuoka Children’s Hospital, Fukuoka.

Japanese Red Cross Fukuoka Hospital, Fukuoka.

Faculty of Medicine, Saga University, Saga.

Saga-ken Medical Centre Koseikan, Saga.

Faculty of Life Sciences, Kumamoto University, Kumamoto.

Graduate School of Medicine, University of the Ryukyus, Okinawa.

Acknowledgements

My deepest ad most sincere gratitude to Professor Katsushi Tokunaga, for the continuous support of my study and related research, for his patience, motivation, immense knowledge and insightful supervision over the 3 years. I am extremely thankful for the opportunity he provided for me to pursue my PhD degree at the University of Tokyo. I also thank all the members of the Department of Human Genetics for their kind support and hearty discussions that allowed me to deepen my understanding of genetics and finish this project.

I would like to extend my gratitude to Dr. Yuki Hitomi for his continuous support and warm encouragement in my daily study and in this project. I would like to extend my gratitude to Dr. Akihiko Mabuchi, Dr. Yosuke Kawai, Dr. Seik-Soon Khor, Dr. Yosuke Omae, Dr. Taku Miyagawa, Dr. Hiromi Sawai and Dr. Khun Zaw Latt for their help and advice regarding experiments and analysis skills.

Additionally, I would like to express my special appreciation and thanks to Professor Kazumoto Iijima and his team in Kobe University, as well as all the collaborators involved in this study, for their help and support regarding the sample collection and the whole project, without which this study would not have been possible. I would like to extend my thanks to all the patients and their families, as well as all the healthy volunteers providing samples for this project.

Last but not the least, I would like to express my special thanks and utmost gratitude to my parents and my husband for always supporting me, and being there for me.

References

1. Eddy AA, Symons JM: Nephrotic syndrome in childhood. *Lancet,* 362**:** 629-639, 2003.

2. Vats AN: Genetics of idiopathic nephrotic syndrome. *Indian J Pediatr,* 72**:** 777-783, 2005.

3. Banh TH, Hussain-Shamsy N, Patel V, Vasilevska-Ristovska J, Borges K, Sibbald C, Lipszyc D, Brooke J, Geary D, Langlois V, Reddon M, Pearl R, Levin L, Piekut M, Licht CP, Radhakrishnan S, Aitken-Menezes K, Harvey E, Hebert D, Piscione TD, Parekh RS: Ethnic Differences in Incidence and Outcomes of Childhood Nephrotic Syndrome. *Clin J Am Soc Nephrol,* 11**:** 1760-1768, 2016.

4. Kikunaga K, Ishikura K, Terano C, Sato M, Komaki F, Hamasaki Y, Sasaki S, Iijima K, Yoshikawa N, Nakanishi K, Nakazato H, Matsuyama T, Ando T, Ito S, Honda M, Japanese Pediatric Survey Holding Information of NssotJSGoRDiC: High incidence of idiopathic nephrotic syndrome in East Asian children: a nationwide survey in Japan (JP-SHINE study). *Clin Exp Nephrol,* 21**:** 651-657, 2017.

5. Kayes MI: Recent Update in The Management of Childhood Nephrotic Syndrome. *Journal of Bangladesh College of Physicians and Surgeons,* 34**:** 26-32, 2016.

6. Metz DK, Kausman JY: Childhood nephrotic syndrome in the 21st century: What's new? *J Paediatr Child Health,* 2014.

7. Ha TS: Genetics of hereditary nephrotic syndrome: a clinical review. *Korean J Pediatr,* 60**:** 55-63, 2017.

8. Niaudet P: Long-term outcome of children with steroid-sensitive idiopathic nephrotic syndrome. *Clin J Am Soc Nephrol,* 4**:** 1547-1548, 2009.

9. Pereira Wde F, Brito-Melo GE, Guimaraes FT, Carvalho TG, Mateo EC, Simoes e Silva AC: The role of the immune system in idiopathic nephrotic syndrome: a review of clinical and experimental studies. *Inflamm Res,* 63**:** 1-12, 2014.

10. van den Berg JG, Weening JJ: Role of the immune system in the pathogenesis of idiopathic nephrotic syndrome. *Clin Sci (Lond),* 107**:** 125-136, 2004.

11. Kaneko K, Tsuji S, Kimata T, Kitao T, Yamanouchi S, Kato S: Pathogenesis of childhood idiopathic nephrotic syndrome: a paradigm shift from T-cells to podocytes. *World J Pediatr,* 11**:** 21-28, 2015.

12. Nagata M: Podocyte injury and its consequences. *Kidney Int,* 89**:** 1221-1230, 2016.

13. Radhakrishnan J, Perazella MA: Drug-induced glomerular disease: attention required! *Clin J Am Soc Nephrol,* 10**:** 1287-1290, 2015.

14. Wiggins RC: The spectrum of podocytopathies: a unifying view of glomerular diseases. *Kidney Int,* 71**:** 1205-1214, 2007.

15. Downie ML, Gallibois C, Parekh RS, Noone DG: Nephrotic syndrome in infants and children: pathophysiology and management. *Paediatr Int Child Health,* 37**:** 248-258, 2017.

16. Abdel-Hafez M, Shimada M, Lee PY, Johnson RJ, Garin EH: Idiopathic nephrotic syndrome and atopy: is there a common link? *Am J Kidney Dis,* 54**:** 945-953, 2009.

17. Meadow SR, Sarsfield JK, Scott DG, Rajah SM: Steroid-responsive nephrotic syndrome and allergy: immunological studies. *Arch Dis Child,* 56**:** 517-524, 1981.

18. Frank C, Herrmann M, Fernandez S, Dirnecker D, Boswald M, Kolowos W, Ruder H, Haas JP: Dominant T cells in idiopathic nephrotic syndrome of childhood. *Kidney Int,* 57**:** 510-517, 2000.

19. McCarthy HJ, Bierzynska A, Wherlock M, Ognjanovic M, Kerecuk L, Hegde S, Feather S, Gilbert RD, Krischock L, Jones C, Sinha MD, Webb NJ, Christian M, Williams MM, Marks S, Koziell A, Welsh GI, Saleem MA, Group RtUSS: Simultaneous sequencing of 24 genes associated with steroid-resistant nephrotic syndrome. *Clin J Am Soc Nephrol,* 8**:** 637-648, 2013.

20. Rood IM, Deegens JK, Wetzels JF: Genetic causes of focal segmental glomerulosclerosis: implications for clinical practice. *Nephrol Dial Transplant,* 27**:** 882-890, 2012.

21. Sadowski CE, Lovric S, Ashraf S, Pabst WL, Gee HY, Kohl S, Engelmann S, Vega-Warner V, Fang H, Halbritter J, Somers MJ, Tan W, Shril S, Fessi I, Lifton RP, Bockenhauer D, El-Desoky S, Kari JA, Zenker M, Kemper MJ, Mueller D, Fathy HM, Soliman NA, Group SS, Hildebrandt F: A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. *J Am Soc Nephrol,* 26**:** 1279-1289, 2015.

22. Bierzynska A, Saleem M: Recent advances in understanding and treating nephrotic syndrome. *F1000Res,* 6**:** 121, 2017.

23. Landau D, Oved T, Geiger D, Abizov L, Shalev H, Parvari R: Familial steroid-sensitive nephrotic syndrome in Southern Israel: clinical and genetic observations. *Pediatr Nephrol,* 22**:** 661-669, 2007.

24. Motoyama O, Sugawara H, Hatano M, Fujisawa T, Iitaka K: Steroid-sensitive nephrotic syndrome in two families. *Clin Exp Nephrol,* 13**:** 170-173, 2009.

25. Ruder H, Scharer K, Opelz G, Lenhard V, Waldherr R, Muller-Wiefel DE, Wingen AM, Dippell J: Human leucocyte antigens in idiopathic nephrotic syndrome in children. *Pediatr Nephrol,* 4**:** 478-481, 1990.

26. Konrad M, Mytilineos J, Bouissou F, Scherer S, Gulli MP, Meissner I, Cambon-Thomsen A, Opelz G, Scharer K: HLA class II associations with idiopathic nephrotic syndrome in children. *Tissue Antigens,* 43**:** 275-280, 1994.

27. Abe KK, Michinaga I, Hiratsuka T, Ogahara S, Naito S, Arakawa K, Tsuru N, Tokieda K: Association of DQB1*0302 alloantigens in Japanese pediatric patients with steroid-sensitive nephrotic syndrome. *Nephron,* 70**:** 28-34, 1995.

28. Adhikari M, Coovadia HM, Hammond MG: Associations between HLA antigens and nephrotic syndrome in African and Indian children in South Africa. *Nephron,* 41**:** 289-292, 1985.

29. Al-Eisa AA, Haider MZ, Srivasta BS: HLA-DRB1 alleles in Kuwaiti children with idiopathic nephrotic syndrome. *Pediatr Nephrol,* 15**:** 79-81, 2000.

30. Alfiler CA, Roy LP, Doran T, Sheldon A, Bashir H: HLA-DRw7 and steroid-responsive nephrotic syndrome of childhood. *Clin Nephrol,* 14**:** 71-74, 1980.

31. Bakr AM, El-Chenawy F: HLA-DQB1 and DRB1 alleles in Egyptian children with steroid-sensitive nephrotic syndrome. *Pediatr Nephrol,* 12**:** 234-237, 1998.

32. Bouissou F, Meissner I, Konrad M, Sommer E, Mytilineos J, Ohayon E, Sierp G, Barthe B, Opelz G, Cambon-Thomsen A, et al.: Clinical implications from studies of HLA antigens in idiopathic nephrotic syndrome in children. *Clin Nephrol,* 44**:** 279-283, 1995.

33. Cheung W, Ren EC, Chan SH, Gong WK, Yap HK: Increased HLA- A*11 in Chinese children with steroid-responsive nephrotic syndrome. *Pediatr Nephrol,* 17**:** 212-216, 2002.

34. de Mouzon-Cambon A, Bouissou F, Dutau G, Barthe P, Parra MT, Sevin A, Ohayon E: HLA-DR7 in children with idiopathic nephrotic syndrome. Correlation with atopy. *Tissue Antigens,* 17**:** 518-524, 1981.

35. Huang YY, Lin FJ, Fu LS, Lan JL: HLA-DR, -DQB typing of steroid-sensitive idiopathic nephrotic syndrome children in Taiwan. *Nephron Clin Pract,* 112**:** c57-64, 2009.

36. Tumer N, Ekim M, Yalcinkaya F, Cakar N, Ensari C: HLA-system and the frequency of relapses in childhood minimal change nephrotic syndrome in Turkish children. *Acta Paediatr Jpn,* 37**:** 419-421, 1995.

37. Gulati S, Tripathi P, Patil SJ, Sharma RK, Agarwal S: Is typing for HLA class II alleles beneficial in Indian children with idiopathic nephrotic syndrome? *Pediatr Nephrol,* 22**:** 528-532, 2007.

38. Clark AG, Vaughan RW, Stephens HA, Chantler C, Williams DG, Welsh KI: Genes encoding the beta-chains of HLA-DR7 and HLA-DQw2 define major susceptibility determinants for idiopathic nephrotic syndrome. *Clin Sci (Lond),* 78**:** 391-397, 1990.

39. Zhou GP, Guo YQ, Ji YH, Zhang GL: Major histocompatibility complex class II antigens in steroid-sensitive nephrotic syndrome in Chinese children. *Pediatr Nephrol,* 8**:** 140-141, 1994.

40. Kobayashi Y, Chen XM, Hiki Y, Fujii K, Kashiwagi N: Association of HLA-DRw8 and DQw3 with minimal change nephrotic syndrome in Japanese adults. *Kidney Int,* 28**:** 193-197, 1985.

41. Kobayashi T, Ogawa A, Takahashi K, Uchiyama M: HLA-DQB1 allele associates with idiopathic nephrotic syndrome in Japanese children. *Acta Paediatr Jpn,* 37**:** 293-296, 1995.

42. Gbadegesin RA, Adeyemo A, Webb NJ, Greenbaum LA, Abeyagunawardena A, Thalgahagoda S, Kale A, Gipson D, Srivastava T, Lin JJ, Chand D, Hunley TE, Brophy PD, Bagga A, Sinha A, Rheault MN, Ghali J, Nicholls K, Abraham E, Janjua HS, Omoloja A, Barletta GM, Cai Y, Milford DD, O'Brien C, Awan A, Belostotsky V, Smoyer WE, Homstad A, Hall G, Wu G, Nagaraj S, Wigfall D, Foreman J, Winn MP, Mid-West Pediatric Nephrology C:

HLA-DQA1 and PLCG2 Are Candidate Risk Loci for Childhood-Onset Steroid-Sensitive Nephrotic Syndrome. *J Am Soc Nephrol,* 26**:** 1701-1710, 2015.

43. Adeyemo A, Esezobor C, Solarin A, Abeyagunawardena A, Kari JA, El Desoky S, Greenbaum LA, Kamel M, Kallash M, Silva C, Young A, Hunley TE, de Jesus-Gonzalez N, Srivastava T, Gbadegesin R: HLA-DQA1 and APOL1 as Risk Loci for Childhood-Onset Steroid-Sensitive and Steroid-Resistant Nephrotic Syndrome. *Am J Kidney Dis,* 71**:** 399-406, 2018.

44. Okamoto K, Tokunaga K, Doi K, Fujita T, Suzuki H, Katoh T, Watanabe T, Nishida N, Mabuchi A, Takahashi A, Kubo M, Maeda S, Nakamura Y, Noiri E: Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nat Genet,* 43**:** 459-463, 2011.

45. Primary nephrotic syndrome in children: clinical significance of histopathologic variants of minimal change and of diffuse mesangial hypercellularity. A Report of the International Study of Kidney Disease in Children. *Kidney Int,* 20**:** 765-771, 1981.

46. Kawai Y, Mimori T, Kojima K, Nariai N, Danjoh I, Saito R, Yasuda J, Yamamoto M, Nagasaki M: Japonica array: improved genotype imputation by designing a population-specific SNP array with 1070 Japanese individuals. *J Hum Genet,* 60**:** 581-587, 2015.

47. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**:** 559-575, 2007.

48. Yang J, Lee SH, Goddard ME, Visscher PM: GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet,* 88**:** 76-82, 2011.

49. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ: LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics,* 26**:** 2336-2337, 2010.

50. Skol AD, Scott LJ, Abecasis GR, Boehnke M: Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet,* 38**:** 209-213, 2006.

51. Ward LD, Kellis M: HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res,* 40**:** D930-934, 2012.

52. Ward LD, Kellis M: HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res,* 44**:** D877-881, 2016.

53. Christopher E. Gillies RP, Rajasree Menon, Edgar Otto, Kalyn Yasutake, Viji Nair, Paul Hoover, David Lieb, Shuqiang Li, Sean Eddy, Damian Fermin, Nephrotic Syndrome Study Network (NEPTUNE), Nir Hacohen, Krzysztof Kiryluk, William Wen, Matthias Kretzler, Matthew G. Sampson: An eQTL landscape of kidney tissue in human nephrotic syndrome.

54. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI: SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics,* 24**:** 2938-2939, 2008.

55. Nishida N, Tanabe T, Takasu M, Suyama A, Tokunaga K: Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal Biochem,* 364**:** 78-85, 2007.

56. Nishida N, Mawatari Y, Sageshima M, Tokunaga K: Highly parallel and short-acting amplification with locus-specific primers to detect single nucleotide polymorphisms by the DigiTag2 assay. *PLoS One,* 7**:** e29967, 2012.

57. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A,* 106**:** 9362-9367, 2009.

58. Zhang F, Lupski JR: Non-coding genetic variants in human disease. *Hum Mol Genet,* 24**:** R102-110, 2015.

59. Shiina T, Inoko H, Kulski JK: An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens,* 64**:** 631-649, 2004.

60. Tokunaga K: Lessons from Genome-Wide Search for Disease-Related Genes with Special Reference to HLA-Disease Associations. *Genes (Basel),* 5**:** 84-96, 2014.

61. Zheng X, Shen J, Cox C, Wakefield JC, Ehm MG, Nelson MR, Weir BS: HIBAG--HLA genotype imputation with attribute bagging. *Pharmacogenomics J,* 14**:** 192-200, 2014.

62. Khor SS, Yang W, Kawashima M, Kamitsuji S, Zheng X, Nishida N, Sawai H, Toyoda H, Miyagawa T, Honda M, Kamatani N, Tokunaga K: High-accuracy imputation for HLA class I and II genes based on high-resolution SNP data of population-specific references. *Pharmacogenomics J,* 15**:** 530-537, 2015.

63. Excoffier L, Lischer HE: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour,* 10**:** 564-567, 2010.

64. Kawashima M, Hitomi Y, Aiba Y, Nishida N, Kojima K, Kawai Y, Nakamura H, Tanaka A, Zeniya M, Hashimoto E, Ohira H, Yamamoto K, Abe M, Nakao K, Yamagiwa S, Kaneko S, Honda M, Umemura T, Ichida T, Seike M, Sakisaka S, Harada M, Yokosuka O, Ueno Y, Senju

M, Kanda T, Shibata H, Himoto T, Murata K, Miyake Y, Ebinuma H, Taniai M, Joshita S, Nikami T, Ota H, Kouno H, Kouno H, Nakamuta M, Fukushima N, Kohjima M, Komatsu T, Komeda T, Ohara Y, Muro T, Yamashita T, Yoshizawa K, Nakamura Y, Shimada M, Hirashima N, Sugi K, Ario K, Takesaki E, Naganuma A, Mano H, Yamashita H, Matsushita K, Yamauchi K, Makita F, Nishimura H, Furuta K, Takahashi N, Kikuchi M, Masaki N, Tanaka T, Tamura S, Mori A, Yagi S, Shirabe K, Komori A, Migita K, Ito M, Nagaoka S, Abiru S, Yatsuhashi H, Yasunami M, Shimoda S, Harada K, Egawa H, Maehara Y, Uemoto S, Kokudo N, Takikawa H, Ishibashi H, Chayama K, Mizokami M, Nagasaki M, Tokunaga K, Nakamura M: Genome-wide association studies identify PRKCB as a novel genetic susceptibility locus for primary biliary cholangitis in the Japanese population. *Hum Mol Genet,* 26**:** 650-659, 2017.

65. Thiri M, Honda K, Kashiwase K, Mabuchi A, Suzuki H, Watanabe K, Nakayama M, Watanabe T, Doi K, Tokunaga K, Noiri E: High-density Association Mapping and Interaction Analysis of PLA2R1 and HLA Regions with Idiopathic Membranous Nephropathy in Japanese. *Sci Rep,* 6**:** 38189, 2016.

66. Kawashima M, Ohashi J, Nishida N, Tokunaga K: Evolutionary analysis of classical HLA class I and II genes suggests that recent positive selection acted on DPB1*04:01 in Japanese population. *PLoS One,* 7**:** e46806, 2012.

67. de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, Morrison J, Richardson A, Walsh EC, Gao X, Galver L, Hart J, Hafler DA, Pericak-Vance M, Todd JA, Daly MJ, Trowsdale J, Wijmenga C, Vyse TJ, Beck S, Murray SS, Carrington M, Gregory S, Deloukas P, Rioux JD: A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet,* 38**:** 1166-1172, 2006.

68. Nakanishi K, Shima Y: Capture of type 1 diabetes-susceptible HLA DR-DQ haplotypes in Japanese subjects using a tag single nucleotide polymorphism. *Diabetes Care,* 33**:** 162-164, 2010.

69. Trowsdale J, Knight JC: Major histocompatibility complex genomics and human disease. *Annu Rev Genomics Hum Genet,* 14**:** 301-323, 2013.

70. Matzaraki V, Kumar V, Wijmenga C, Zhernakova A: The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome Biol,* 18**:** 76, 2017.

71. Yamasaki M, Miyagawa T, Toyoda H, Khor SS, Liu X, Kuwabara H, Kano Y, Shimada T, Sugiyama T, Nishida H, Sugaya N, Tochigi M, Otowa T, Okazaki Y, Kaiya H, Kawamura Y, Miyashita A, Kuwano R, Kasai K, Tanii H, Sasaki T, Honda Y, Honda M, Tokunaga K: Evaluation of polygenic risks for narcolepsy and essential hypersomnia. *J Hum Genet,* 61**:** 873-878, 2016.

72. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, Teles e Silva AL, Ghattaoraya GS, Alfirevic A, Jones AR, Middleton D: Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. *Nucleic Acids Res,* 43**:** D784-788, 2015.

73. Fort M, de Stefano GF, Cambon-Thomsen A, Giraldo-Alvarez P, Dugoujon JM, Ohayon E, Scano G, Abbal M: HLA class II allele and haplotype frequencies in Ethiopian Amhara and Oromo populations. *Tissue Antigens,* 51**:** 327-336, 1998.

74. Ruiz TM, da Costa SM, Ribas F, Luz PR, Lima SS, da Graca Bicalho M: Human leukocyte antigen allelic groups and haplotypes in a brazilian sample of volunteer donors for bone marrow transplant in Curitiba, Parana, Brazil. *Transplant Proc,* 37**:** 2293-2296, 2005.

75. Miao KR, Pan QQ, Tang RC, Zhou XP, Fan S, Wang XY, Zhao X, Xue M, Zhou XY, Wang CY: The polymorphism and haplotype analysis of HLA-A, -B and -DRB1 genes of population in Jiangsu province of China. *Int J Immunogenet,* 34**:** 419-424, 2007.

76. Hajeer AH, Sawidan FA, Bohlega S, Saleh S, Sutton P, Shubaili A, Tahan AA, Al Jumah M: HLA class I and class II polymorphisms in Saudi patients with myasthenia gravis. *Int J Immunogenet,* 36**:** 169-172, 2009.

77. Leffell MS, Cherikh WS, Land G, Zachary AA: Improved definition of human leukocyte antigen frequencies among minorities and applicability to estimates of transplant compatibility. *Transplantation,* 83**:** 964-972, 2007.

78. Hashimoto M, Kinoshita T, Yamasaki M, Tanaka H, Imanishi T, Ihara H, Ichikawa Y, Fukunishi T: Gene frequencies and haplotypic associations within the HLA region in 916 unrelated Japanese individuals. *Tissue Antigens,* 44**:** 166-173, 1994.

79. Ikeda N, Kojima H, Nishikawa M, Hayashi K, Futagami T, Tsujino T, Kusunoki Y, Fujii N, Suegami S, Miyazaki Y, Middleton D, Tanaka H, Saji H: Determination of HLA-A, -C, -B, -DRB1 allele and haplotype frequency in Japanese population based on family study. *Tissue Antigens,* 85**:** 252-259, 2015.

80. Nakajima F, Nakamura J, Yokota T: Analysis of HLA haplotypes in Japanese, using high resolution allele typing. *MHC,* 8**:** 1-32, 2001.

81. Qin Qin P, Su F, Xiao Yan W, Xing Z, Meng P, Chengya W, Jie S: Distribution of human leucocyte antigen-A, -B and -DR alleles and haplotypes at high resolution in the population from Jiangsu province of China. *Int J Immunogenet,* 38**:** 475-481, 2011.

82. Huh JY, Yi DY, Eo SH, Cho H, Park MH, Kang MS: HLA-A, -B and -DRB1 polymorphism in Koreans defined by sequence-based typing of 4128 cord blood units. *Int J Immunogenet,* 40**:** 515-523, 2013.

83. Caillat-Zucman S: Molecular mechanisms of HLA association with autoimmune diseases. *Tissue Antigens,* 73**:** 1-8, 2009.

84. Hussain N, Zello JA, Vasilevska-Ristovska J, Banh TM, Patel VP, Patel P, Battiston CD, Hebert D, Licht CP, Piscione TD, Parekh RS: The rationale and design of Insight into Nephrotic Syndrome: Investigating Genes, Health and Therapeutics (INSIGHT): a prospective cohort study of childhood nephrotic syndrome. *BMC Nephrol,* 14**:** 25, 2013.

Tables and figures

Table 1. Definitions of nephrotic syndrome.

Nephrotic syndrome	Urine protein to creatinine ratio ≥2.0 and serum albumin ≤2.5 g/dl
Complete remission	Negative protein on urine dipstick test or urine protein to creatinine ratio 0.2 for 3 consecutive days
Relapse	Protein ≥3+ on urine dipstick test for 3 consecutive days
Steroid-sensitive nephrotic syndrome (SSNS)	Complete remission within 4 weeks after starting 60 mg/m² oral prednisolone per day
Steroid-resistant nephrotic syndrome (SRNS)	Persistent proteinuria after 60 mg/m² oral prednisolone per day for 4 weeks
Steroid-dependent nephrotic syndrome (SDNS)	Two relapses of nephrotic syndrome during the reduction of steroid treatment or within 2 weeks of discontinuation of steroid treatment
Frequently-relapsing nephrotic syndrome (FRNS)	≥2 relapses of nephrotic syndrome within 6 months after initial remission, or ≥4 relapses within any 12-month period

Table 2. Clinical information of patients in the discovery stage and the replication stage.

Characteristic	Discovery stage	Replication stage	Total
Number of patients	224	216	440
Sex ratio (Male: Female)	2.3:1 (156:68)	3.1:1 (163:53)	2.6:1
Onset age (Median, Range) (yr)	3.8 (0.4-16.3)	4.3 (1.3-22.0)	4.0 (0.4-22.0)
Follow-up (Median, Range) (yr)	5.5 (0-48.2)	6.3 (0-30.3)	5.8 (0-48.2)

Table 3. Candidate SNPs/INDELs identified in the discovery stage using Cochran-Armitage trend test and the replication of candidate SNPs.

CHR	SNP/INDEL	BP	A1	A2	Annotation	Discovery stage (224 cases vs. 412 controls)			Replication (216 cases vs. 719 controls)		
						Freq in cases	Freq in controls	GC-corrected *P*-value	Freq in cases	Freq in controls	*P*-value
5	rs4242036	53733306	A	G	Genotyped	0.11	0.04	4.46E-06	0.06	0.05	0.49
5	rs4865828	53733241	A	G	Imputed	0.11	0.04	4.46E-06	/	/	/
5	rs6864175	84134415	T	A	Imputed	0.29	0.18	6.67E-06	0.24	0.21	0.26
5	rs1875277	84127057	C	T	Genotyped	0.29	0.19	1.50E-05	0.25	0.22	0.33
6	rs79193269	32636661	C	CCTT	Imputed	0.08	0.23	1.54E-10	/	/	/
6	rs9275103	32649386	T	C	Imputed	0.11	0.27	2.46E-10	/	/	/
6	rs9282090	32629047	A	AC	Imputed	0.13	0.29	4.60E-10	/	/	/
6	rs9275101	32649355	G	A	Imputed	0.11	0.26	4.69E-10	/	/	/
6	rs4642516	32657543	G	T	Genotyped	0.11	0.26	4.69E-10	0.10	0.28	6.69E-13
6	rs3134996	32636866	A	T	Genotyped	0.09	0.23	7.06E-10	0.12	0.32	9.44E-10

CHR: Chromosome; A1: Minor allele (test allele); A2: Major allele (reference allele).
Freq: Minor allele frequency.
GC-corrected *P*-value: *P*-values were calculated using Cochran-Armitage trend test, corrected by genomic control in the discovery GWAS.
P-value: *P*-values were calculated using Cochran-Armitage trend test.
/: Genotyping of 5 candidate SNPs/INDELs was failed.
Bonferroni correction: The *P*-value required for significance in the replication stage was 0.05/5=0.01.

Table 4. Replication of two candidate SNPs in *HLA-DR/DQ* region under different genetic models.

CHR	SNP	A1	A2	Genetic model	Cases (N=216)	Controls (N=719)	*P*-Fisher	OR (95% CI)
6	rs4642516	G	T	Allelic	45/387	398/1036	4.51E-15	0.30 (0.22-0.42)
				Dominant	43/173	336/381	3.43E-13	0.28 (0.20-0.41)
				Recessive	2/214	62/655	9.51E-06	0.10 (0.01-0.41)
6	rs3134996	A	T	Allelic	49/345	403/853	1.26E-15	0.30 (0.22-0.41)
				Dominant	34/163	265/363	6.24E-11	0.29 (0.19-0.43)
				Recessive	15/182	138/490	1.75E-06	0.29 (0.17-0.51)

CHR: Chromosome; A1: Minor allele (test allele); A2: Major allele (reference allele); *P*-Fisher: *P*-value was calculated using Fisher's exact test.
OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.

Table 5. Combined analysis of SNPs in *HLA-DR/DQ* region.

CHR	SNP	A1/A2	Genetic model	Discovery GWAS (224 cases vs. 412 controls)				Replication (216 cases vs. 719 controls)				Combined analysis (440 cases vs. 1131 controls)			
				Cases	Controls	*P*-Fisher	OR (95% CI)	Cases	Controls	*P*-Fisher	OR (95% CI)	Cases	Controls	*P*-Fisher	OR (95% CI)
6	rs4642516	G/T	Allelic	0.11[a]	0.26[b]	1.06E-10	0.36 (0.26-0.50)	0.10[a]	0.28[b]	4.51E-15	0.30 (0.22-0.42)	0.11[a]	0.27[b]	7.84E-23	0.33 (0.26-0.41)
			Dominant	47/177	194/218	5.29E-11	0.30 (0.21-0.43)	43/173	336/381	3.43E-13	0.28 (0.20-0.41)	90/350	530/599	4.87E-23	0.29 (0.22-0.38)
			Recessive	4/220	24/388	2.41E-02	0.29 (0.10-0.86)	2/214	62/655	9.51E-06	0.10 (0.02-0.41)	6/434	86/1043	2.40E-07	0.17 (0.07-0.39)
6	rs3134996	A/T	Allelic	0.09[a]	0.23[b]	4.85E-11	0.32 (0.22-0.46)	0.12[a]	0.32[b]	1.26E-15	0.30 (0.22-0.41)	0.10[a]	0.28[b]	1.72E-25	0.29 (0.23-0.37)
			Dominant	37/187	168/244	1.86E-10	0.29 (0.19-0.43)	34/163	265/363	6.24E-11	0.29 (0.19-0.43)	71/350	433/607	9.57E-21	0.28 (0.21-0.38)
			Recessive	2/222	21/391	6.35E-03	0.17 (0.04-0.72)	15/182	138/490	1.75E-06	0.29 (0.17-0.51)	17/404	159/881	1.31E-10	0.23 (0.14-0.39)

CHR: Chromosome; A1: Minor allele (test allele); A2: Major allele (reference allele). *P*-Fisher: *P*-value calculated using Fisher's exact test; OR Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%. [a]: Minor allele frequency in cases; [b]: Minor allele frequency in controls.

Table 6A. *HLA* allele association analysis of *HLA-A* with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-A alleles[a]	Allele frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*A*02:01*	42	10.7	84	10.2	1.05 (0.71-1.55)	NS	-
*A*02:06*	53	13.5	62	7.5	1.91 (1.29-2.81)	9.82E-04	8.84E-03
*A*02:07*	16	4.1	24	2.9	1.41 (0.74-268)	NS	-
*A*11:01*	42	10.7	80	9.7	1.11 (0.75-1.64)	NS	-
*A*24:02*	141	35.8	326	39.7	0.85 (0.66-1.09)	NS	-
*A*26:01*	35	8.9	69	8.4	1.06 (0.70-1.63)	NS	-
*A*26:03*	9	2.3	29	3.5	0.64 (0.30-1.36)	NS	-
*A*31:01*	44	11.2	65	7.9	1.46 (0.98-2.19)	NS	-
*A*33:03*	10	2.5	65	7.9	0.30 (0.15-0.60)	2.70E-04	2.43E-03

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 6B. *HLA* allele association analysis of *HLA-C* with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-C alleles[a]	Allele frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*C*01:02*	64	16.2	140	17.0	0.94 (0.68-1.31)	NS	-
*C*03:03*	57	14.5	113	13.7	1.06 (0.75-1.50)	NS	-
*C*03:04*	58	14.7	102	12.4	1.22 (0.86-1.72)	NS	-
*C*04:01*	24	6.1	42	5.1	1.20 (0.72-2.02)	NS	-
*C*06:02*	2	0.5	7	0.9	0.59 (0.12-2.87)	NS	-
*C*07:02*	45	11.4	119	14.5	0.76 (0.53-1.10)	NS	-
*C*08:01*	41	10.4	48	5.8	1.87 (1.21-2.89)	4.22E-03	5.06E-02
*C*08:03*	4	1.0	12	1.5	0.69 (0.22-2.16)	NS	-
*C*12:02*	44	11.2	81	9.9	1.15 (0.78-1.70)	NS	-
*C*14:02*	37	9.4	49	6.0	1.63 (1.05-2.55)	2.90E-02	3.48E-01
*C*14:03*	4	1.0	60	7.3	0.13 (0.05-0.36)	4.37E-06	5.25E-05
*C*15:02*	9	2.3	31	3.8	0.60 (0.28-1.27)	NS	-

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 6C. *HLA* allele association analysis of *HLA-B* with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-B alleles[a]	Allele frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*B*07:02*	16	4.1	57	6.9	0.57 (0.32-1.00)	4.84E-02	1.06E+00
*B*13:01*	2	0.5	13	1.6	0.32 (0.07-1.41)	NS	-
*B*15:01*	38	9.6	72	8.8	1.11 (0.74-1.68)	NS	-
*B*15:07*	3	0.8	5	0.6	1.25 (0.30-5.27)	NS	-
*B*15:11*	3	0.8	5	0.6	1.25 (0.30-5.27)	NS	-
*B*15:18*	2	0.5	14	1.7	0.29 (0.07-1.30)	NS	-
*B*35:01*	33	8.4	67	8.2	1.03 (0.67-1.59)	NS	-
*B*37:01*	2	0.5	7	0.9	0.59 (0.12-2.87)	NS	-
*B*39:01*	15	3.8	34	4.1	0.92 (0.49-1.70)	NS	-
*B*40:01*	16	4.1	47	5.7	0.70 (0.39-1.25)	NS	-
*B*40:02*	45	11.4	58	7.1	1.70 (1.13-2.56)	1.05E-02	2.31E-01
*B*40:06*	32	8.1	33	4.0	2.11 (1.28-3.49)	2.88E-03	6.34E-02
*B*44:03*	4	1.0	59	7.2	0.13 (0.05-0.37)	5.69E-06	1.25E-04
*B*46:01*	19	4.8	38	4.6	1.05 (0.59-1.84)	NS	-
*B*48:01*	10	2.5	22	2.7	0.95 (0.44-2.02)	NS	-
*B*51:01*	42	10.7	70	8.5	1.28 (0.86-1.92)	NS	-
*B*52:01*	45	11.4	80	9.7	1.20 (0.81-1.76)	NS	-
*B*54:01*	22	5.6	65	7.9	0.69 (0.42-1.13)	NS	-
*B*55:02*	13	3.3	16	1.9	1.72 (0.82-3.61)	NS	-
*B*56:01*	5	1.3	5	0.6	2.10 (0.60-7.30)	NS	-
*B*59:01*	8	2.0	20	2.4	0.83 (0.36-1.90)	NS	-
*B*67:01*	9	2.3	11	1.3	1.72 (0.71-4.19)	NS	-

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 6D. *HLA* allele association analyses of *HLA-DRB1* with childhood SSNS in the discovery stage using HLA-imputation data.

	Allele frequency						
HLA-DRB1 alleles[a]	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DRB1*01:01*	15	3.8	57	6.9	0.53 (0.30-0.95)	3.06E-02	5.20E-01
*DRB1*04:03*	18	4.6	25	3.0	1.53 (0.82-2.83)	NS	-
*DRB1*04:05*	56	14.2	121	14.7	0.96 (0.68-1.35)	NS	-
*DRB1*04:06*	13	3.3	27	3.3	1.00 (0.51-1.97)	NS	-
*DRB1*08:02*	40	10.2	32	3.9	2.79 (1.72-4.51)	1.50E-05	2.56E-04
*DRB1*08:03*	27	6.9	63	7.7	0.89 (0.56-1.41)	NS	-
*DRB1*09:01*	73	18.5	126	15.3	1.26 (0.91-1.73)	NS	-
*DRB1*11:01*	14	3.6	23	2.8	1.28 (0.65-2.52)	NS	-
*DRB1*12:01*	23	5.8	30	3.6	1.64 (0.94-2.86)	NS	-
*DRB1*12:02*	3	0.8	18	2.2	0.34 (0.10-1.17)	NS	-
*DRB1*13:02*	5	1.3	55	6.7	0.18 (0.07-0.45)	4.40E-05	7.47E-04
*DRB1*14:03*	11	2.8	11	1.3	2.12 (0.91-4.93)	NS	-
*DRB1*14:05*	7	1.8	16	1.9	0.91 (0.37-2.23)	NS	-
*DRB1*14:06*	13	3.3	13	1.6	2.12 (0.97-4.62)	NS	-
*DRB1*14:54*	6	1.5	26	3.2	0.47 (0.19-1.16)	NS	-
*DRB1*15:01*	18	4.6	67	8.2	0.54 (0.32-0.92)	2.19E-02	3.72E-01
*DRB1*15:02*	44	11.2	70	8.5	1.35 (0.91-2.01)	NS	-

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 6E. *HLA* allele association analyses of *HLA-DQB1* with childhood SSNS in the discovery stage using HLA-imputation data.

	Allele frequency						
HLA-DQB1 alleles[a]	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DQB1*03:01*	62	15.7	99	12.0	1.36 (0.97-1.92)	NS	-
*DQB1*03:02*	68	17.3	75	9.1	2.08 (1.46-2.96)	3.77E-05	4.14E-04
*DQB1*03:03*	79	20.1	133	16.2	1.30 (0.95-1.77)	NS	-
*DQB1*04:01*	53	13.5	121	14.7	0.90 (0.64-1.27)	NS	-
*DQB1*04:02*	8	2.0	26	3.2	0.63 (0.28-1.41)	NS	-
*DQB1*05:01*	16	4.1	63	7.7	0.51 (0.29-0.89)	1.70E-02	1.87E-01
*DQB1*05:02*	6	1.5	17	2.1	0.73 (0.29-1.87)	NS	-
*DQB1*05:03*	11	2.8	29	3.5	0.79 (0.39-1.59)	NS	-
*DQB1*06:01*	71	18.0	133	16.2	1.14 (0.83-1.56)	NS	-
*DQB1*06:02*	14	3.6	65	7.9	0.43 (0.24-0.77)	3.94E-03	4.33E-02
*DQB1*06:04*	2	0.5	53	6.4	0.07 (0.02-0.31)	3.09E-06	3.40E-05

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 6F. *HLA* allele association analyses of *HLA-DPB1* with childhood SSNS in the discovery stage using HLA-imputation data.

	Allele frequency						
HLA-DPB1 alleles[a]	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DPB1*02:01*	102	25.9	210	25.5	1.02 (0.77-1.34)	NS	-
*DPB1*02:02*	10	2.5	35	4.3	0.59 (0.29-1.19)	NS	-
*DPB1*03:01*	8	2.0	36	4.4	0.45 (0.21-0.98)	4.01E-02	4.01E-01
*DPB1*04:01*	8	2.0	43	5.2	0.38 (0.17-0.81)	9.17E-03	9.17E-02
*DPB1*04:02*	30	7.6	83	10.1	0.73 (0.47-1.13)	NS	-
*DPB1*05:01*	188	47.7	316	38.4	1.46 (1.15-1.86)	2.13E-03	2.13E-02
*DPB1*06:01*	2	0.5	5	0.6	0.83 (0.16-4.32)	NS	-
*DPB1*09:01*	40	10.2	65	7.9	1.32 (0.87-1.99)	NS	-
*DPB1*13:01*	4	1.0	12	1.5	0.69 (0.22-2.16)	NS	-
*DPB1*14:01*	2	0.5	10	1.2	0.41 (0.09-1.90)	NS	-

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 7A. Association analysis of *HLA-A-C-B* haplotypes with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-A-C-B haplotypes[a]	Haplotype frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*A*02:01-C*01:02-B*54:01*	4	1.0	7	0.9	1.19 (0.35-4.10)	NS	-
*A*02:01-C*03:03-B*15:01*	3	0.8	8	1.0	0.78 (0.21-2.96)	NS	-
*A*02:01-C*03:04-B*40:01*	2	0.5	5	0.6	0.83 (0.16-4.32)	NS	-
*A*02:06-C*03:03-B*35:01*	4	1.0	11	1.3	0.76 (0.24-2.39)	NS	-
*A*02:06-C*07:02-B*39:01*	5	1.3	14	1.7	0.74 (0.27-2.07)	NS	-
*A*02:06-C*08:01-B*40:06*	16	4.1	9	1.1	3.82 (1.67-8.73)	6.47E-04	1.94E-02
*A*02:06-C*14:02-B*51:01*	2	0.5	7	0.9	0.59 (0.12-2.87)	NS	-
*A*02:07-C*01:02-B*46:01*	11	2.8	19	2.3	1.21 (0.57-2.58)	NS	-
*A*11:01-C*01:02-B*54:01*	2	0.5	17	2.1	0.24 (0.06-1.05)	4.00E-02	1.20E+00
*A*11:01-C*03:03-B*35:01*	2	0.5	5	0.6	0.83 (0.16-4.32)	NS	-
*A*11:01-C*04:01-B*15:01*	11	2.8	19	2.3	1.21 (0.57-2.58)	NS	-
*A*11:01-C*07:02-B*67:01*	3	0.8	6	0.7	1.04 (0.26-4.19)	NS	-
*A*24:02-C*01:02-B*54:01*	8	2.0	32	3.9	0.51 (0.23-1.12)	NS	-
*A*24:02-C*01:02-B*55:02*	3	0.8	7	0.9	0.89 (0.23-3.47)	NS	-
*A*24:02-C*01:02-B*59:01*	5	1.3	15	1.8	0.69 (0.25-1.92)	NS	-
*A*24:02-C*03:03-B*15:01*	7	1.8	5	0.6	2.96 (0.93-9.37)	NS	-
*A*24:02-C*03:03-B*35:01*	10	2.5	14	1.7	1.50 (0.66-3.41)	NS	-
*A*24:02-C*03:04-B*40:01*	3	0.8	11	1.3	0.57 (0.16-2.04)	NS	-
*A*24:02-C*03:04-B*40:02*	9	2.3	19	2.3	0.99 (0.44-2.20)	NS	-
*A*24:02-C*07:02-B*07:02*	10	2.5	45	5.5	0.45 (0.22-0.90)	2.11E-02	6.33E-01
*A*24:02-C*08:01-B*40:06*	4	1.0	6	0.7	1.39 (0.39-4.97)	NS	-
*A*24:02-C*12:02-B*52:01*	39	9.9	77	9.4	1.06 (0.71-1.59)	NS	-
*A*24:02-C*14:02-B*51:01*	8	2.0	14	1.7	1.20 (0.50-2.88)	NS	-
*A*26:01-C*03:03-B*15:01*	7	1.8	5	0.6	2.96 (0.93-9.37)	NS	-
*A*26:01-C*03:03-B*35:01*	5	1.3	11	1.3	0.95 (0.33-2.75)	NS	-
*A*26:01-C*03:04-B*40:02*	8	2.0	15	1.8	1.12 (0.47-2.65)	NS	-
*A*26:01-C*08:01-B*40:06*	4	1.0	6	0.7	1.39 (0.39-4.97)	NS	-
*A*31:01-C*03:04-B*40:02*	8	2.0	5	0.6	3.39 (1.10-10.42)	3.45E-02	1.04E+00
*A*31:01-C*14:02-B*51:01*	19	4.8	16	1.9	2.55 (1.30-5.02)	5.00E-03	1.50E-01
*A*33:03-C*14:03-B*44:03*	4	1.0	54	6.6	0.15 (0.05-0.41)	2.11E-05	6.33E-04

[a]Haplotypes: *HLA-A-C-B* haplotypes with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for *HLA* haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).

-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 7B. Association analysis of *HLA-DRB1-DQB1* haplotypes with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-DRB1-DQB1 haplotypes[a]	Haplotype frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DRB1*01:01-DQB1*05:01*	15	3.8	57	6.9	0.53 (0.30-0.95)	3.06E-02	6.12E-01
*DRB1*04:03-DQB1*03:02*	18	4.6	24	2.9	1.59 (0.85-2.97)	NS	-
*DRB1*04:05-DQB1*04:01*	53	13.5	120	14.6	0.91 (0.64-1.29)	NS	-
*DRB1*04:06-DQB1*03:02*	13	3.3	27	3.3	1.00 (0.51-1.97)	NS	-
*DRB1*08:02-DQB1*03:02*	33	8.4	17	2.1	4.33 (2.38-7.87)	2.17E-07	4.34E-06
*DRB1*08:02-DQB1*04:02*	7	1.8	15	1.8	0.97 (0.39-2.41)	NS	-
*DRB1*08:03-DQB1*06:01*	27	6.9	63	7.7	0.89 (0.56-1.41)	NS	-
*DRB1*09:01-DQB1*03:03*	72	18.3	123	15.0	1.27 (0.92-1.75)	NS	-
*DRB1*11:01-DQB1*03:01*	14	3.6	22	2.7	1.34 (0.68-2.65)	NS	-
*DRB1*12:01-DQB1*03:01*	16	4.1	22	2.7	1.54 (0.80-2.96)	NS	-
*DRB1*12:01-DQB1*03:03*	6	1.5	8	1.0	1.57 (0.54-4.57)	NS	-
*DRB1*12:02-DQB1*03:01*	3	0.8	18	2.2	0.34 (0.10-1.17)	NS	-
*DRB1*13:02-DQB1*06:04*	2	0.5	53	6.4	0.07 (0.02-0.31)	3.09E-06	6.17E-05
*DRB1*14:03-DQB1*03:01*	11	2.8	11	1.3	2.12 (0.91-4.93)	NS	-
*DRB1*14:05-DQB1*05:03*	7	1.8	15	1.8	0.97 (0.39-2.41)	NS	-
*DRB1*14:06-DQB1*03:01*	13	3.3	13	1.6	2.12 (0.97-4.62)	NS	-
*DRB1*14:54-DQB1*05:02*	2	0.5	13	1.6	0.32 (0.07-1.41)	NS	-
*DRB1*14:54-DQB1*05:03*	4	1.0	13	1.6	0.64 (0.21-1.97)	NS	-
*DRB1*15:01-DQB1*06:02*	14	3.6	65	7.9	0.43 (0.24-0.77)	3.94E-03	7.87E-02
*DRB1*15:02-DQB1*06:01*	44	11.2	70	8.5	1.35 (0.91-2.01)	NS	-

[a]Haplotypes: *HLA-DRB1-DQB1* haplotypes with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for *HLA* haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 7C. Association analysis of *HLA-DRB1-DQB1-DPB1* haplotypes with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-DRB1-DQB1-DPB1 haplotypes[a]	Haplotype frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DRB1*01:01-DQB1*05:01-DPB1*04:02*	11	2.8	49	6.0	0.45 (0.23-0.88)	1.69E-02	4.40E-01
*DRB1*04:03-DQB1*03:02-DPB1*02:01*	9	2.3	10	1.2	1.90 (0.77-4.71)	NS	-
*DRB1*04:03-DQB1*03:02-DPB1*05:01*	9	2.3	9	1.1	2.11 (0.83-5.36)	NS	-
*DRB1*04:05-DQB1*04:01-DPB1*02:01*	10	2.5	12	1.5	1.76 (0.75-4.10)	NS	-
*DRB1*04:05-DQB1*04:01-DPB1*05:01*	38	9.6	69	8.4	1.16 (0.77-1.76)	NS	-
*DRB1*04:06-DQB1*03:02-DPB1*02:01*	4	1.0	19	2.3	0.43 (0.15-1.28)	NS	-
*DRB1*08:02-DQB1*03:02-DPB1*05:01*	20	5.1	15	1.8	2.88 (1.46-5.68)	1.51E-03	3.91E-02
*DRB1*08:02-DQB1*04:02-DPB1*05:01*	4	1.0	7	0.9	1.19 (0.35-4.10)	NS	-
*DRB1*08:03-DQB1*06:01-DPB1*02:01*	6	1.5	16	1.9	0.78 (0.30-2.01)	NS	-
*DRB1*08:03-DQB1*06:01-DPB1*02:02*	3	0.8	16	1.9	0.39 (0.11-1.33)	NS	-
*DRB1*08:03-DQB1*06:01-DPB1*05:01*	16	4.1	22	2.7	1.54 (0.80-2.96)	NS	-
*DRB1*09:01-DQB1*03:03-DPB1*02:01*	40	10.2	32	3.9	2.79 (1.72-4.51)	1.50E-05	3.91E-04
*DRB1*09:01-DQB1*03:03-DPB1*05:01*	24	6.1	77	9.4	0.63 (0.39-1.01)	NS	-
*DRB1*11:01-DQB1*03:01-DPB1*02:01*	2	0.5	9	1.1	0.46 (0.10-2.14)	NS	-
*DRB1*11:01-DQB1*03:01-DPB1*05:01*	7	1.8	11	1.3	1.33 (0.51-3.47)	NS	-
*DRB1*12:01-DQB1*03:01-DPB1*05:01*	9	2.3	10	1.2	1.90 (0.77-4.71)	NS	-
*DRB1*12:01-DQB1*03:03-DPB1*05:01*	4	1.0	7	0.9	1.19 (0.35-4.10)	NS	-
*DRB1*12:02-DQB1*03:01-DPB1*05:01*	2	0.5	10	1.2	0.41 (0.09-1.90)	NS	-
*DRB1*13:02-DQB1*06:04-DPB1*04:01*	2	0.5	35	4.3	0.11 (0.03-0.48)	3.66E-04	9.52E-03
*DRB1*14:05-DQB1*05:03-DPB1*05:01*	4	1.0	11	1.3	0.76 (0.24-2.39)	NS	-
*DRB1*14:54-DQB1*05:03-DPB1*05:01*	4	1.0	7	0.9	1.19 (0.35-4.10)	NS	-
*DRB1*15:01-DQB1*06:02-DPB1*02:01*	3	0.8	39	4.7	0.15 (0.05-0.50)	3.71E-04	9.65E-03
*DRB1*15:01-DQB1*06:02-DPB1*04:02*	3	0.8	5	0.6	1.25 (0.30-5.27)	NS	-
*DRB1*15:01-DQB1*06:02-DPB1*05:01*	4	1.0	13	1.6	0.64 (0.21-1.97)	NS	-
*DRB1*15:02-DQB1*06:01-DPB1*05:01*	9	2.3	10	1.2	1.90 (0.77-4.71)	NS	-
*DRB1*15:02-DQB1*06:01-DPB1*09:01*	34	8.6	48	5.8	1.52 (0.96-2.40)	NS	-

[a]Haplotypes: *HLA-DRB1-DQB1-DPB1* haplotypes with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for *HLA* haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 7D. Association analysis of *HLA-A-C-B-DRB1-DQB1* haplotypes with childhood SSNS in the discovery stage using HLA-imputation data.

HLA-A-C-B-DRB1-DQB1 haplotypes[a]	Haplotype frequency						
	Cases (2n=394)		Controls (2n=822)		chi-square test		
	No	%	No	%	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*A*02:06-C*08:01-B*40:06-DRB1*09:01-DQB1*03:03*	12	3.0	5	0.6	5.13 (1.80-14.67)	7.04E-04	6.34E-03
*A*02:07-C*01:02-B*46:01-DRB1*08:03-DQB1*06:01*	9	2.3	16	1.9	1.18 (0.52-2.69)	NS	-
*A*11:01-C*04:01-B*15:01-DRB1*04:06-DQB1*03:02*	9	2.3	18	2.2	1.04 (0.46-2.35)	NS	-
*A*24:02-C*01:02-B*54:01-DRB1*04:05-DQB1*04:01*	5	1.3	24	2.9	0.43 (0.16-1.13)	NS	-
*A*24:02-C*01:02-B*59:01-DRB1*04:05-DQB1*04:01*	5	1.3	13	1.6	0.80 (0.28-2.26)	NS	-
*A*24:02-C*07:02-B*07:02-DRB1*01:01-DQB1*05:01*	9	2.3	39	4.7	0.47 (0.23-0.98)	3.92E-02	3.53E-01
*A*24:02-C*08:01-B*40:06-DRB1*09:01-DQB1*03:03*	2	0.5	5	0.6	0.83 (0.16-4.32)	NS	-
*A*24:02-C*12:02-B*52:01-DRB1*15:02-DQB1*06:01*	38	9.6	62	7.5	1.31 (0.86-2.00)	NS	-
*A*26:01-C*08:01-B*40:06-DRB1*09:01-DQB1*03:03*	2	0.5	6	0.7	0.69 (0.14-3.45)	NS	-

[a]Haplotypes: *HLA-A-C-B-DRB1-DQB1* haplotypes with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for *HLA* haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 8. *HLA* alleles and haplotypes significantly associated with childhood SSNS in the discovery stage using HLA-imputation data.

HLA alleles/haplotypes[a]	Allele frequency					
	Cases (2n=394)		Controls (2n=822)		chi-square test	
	No	%	No	%	OR (95% CI)[b]	*P*-corrected[c]
*A*02:06*	53	13.5	62	7.5	1.91 (1.29-2.81)	8.84E-03
*A*33:03*	10	2.5	65	7.9	0.30 (0.15-0.60)	2.43E-03
*C*14:03*	4	1.0	60	7.3	0.13 (0.05-0.36)	5.25E-05
*B*44:03*	4	1.0	59	7.2	0.13 (0.05-0.37)	1.25E-04
*DRB1*08:02*	40	10.2	32	3.9	2.79 (1.72-4.51)	2.56E-04
*DRB1*13:02*	5	1.3	55	6.7	0.18 (0.07-0.45)	7.47E-04
*DQB1*03:02*	68	17.3	75	9.1	2.08 (1.46-2.96)	4.14E-04
*DQB1*06:02*	14	3.6	65	7.9	0.43 (0.24-0.77)	4.33E-02
*DQB1*06:04*	2	0.5	53	6.4	0.07 (0.02-0.31)	3.40E-05
*DPB1*05:01*	188	47.7	316	38.4	1.46 (1.15-1.86)	2.13E-02
*A*33:03-B*44:03*	4	1.0	54	6.6	0.15 (0.05-0.41)	6.54E-04
*A*02:06-B*40:06*	17	4.3	10	1.2	3.66 (1.66-8.07)	1.86E-02
*B*44:03-C*14:03*	4	1.0	59	7.2	0.13 (0.05-0.37)	1.59E-04
*B*40:06-C*08:01*	28	7.1	24	2.9	2.54 (1.45-4.45)	2.05E-02
*A*02:06-C*08:01-B*40:06*	16	4.1	9	1.1	3.82 (1.67-8.73)	1.94E-02
*A*33:03-C*14:03-B*44:03*	4	1.0	54	6.6	0.15 (0.05-0.41)	6.33E-04
*DRB1*08:02-DQB1*03:02*	33	8.4	17	2.1	4.33 (2.38-7.87)	4.34E-06
*DRB1*13:02-DQB1*06:04*	2	0.5	53	6.4	0.07 (0.02-0.31)	6.17E-05
*DRB1*08:02-DQB1*03:02-DPB1*05:01*	20	5.1	15	1.8	2.88 (1.46-5.68)	3.91E-02
*DRB1*09:01-DQB1*03:03-DPB1*02:01*	40	10.2	32	3.9	2.79 (1.72-4.51)	3.91E-04
*DRB1*13:02-DQB1*06:04-DPB1*04:01*	2	0.5	35	4.3	0.11 (0.03-0.48)	9.52E-03
*DRB1*15:01-DQB1*06:02-DPB1*02:01*	3	0.8	39	4.7	0.15 (0.05-0.50)	9.65E-03
*A*02:06-C*08:01-B*40:06-DRB1*09:01-DQB1*03:03*	12	3.0	5	0.6	5.13 (1.80-14.67)	6.34E-03

[a]Alleles/haplotypes: *HLA* alleles/haplotypes with frequencies < 0.5% in cases or controls were excluded from the association analyses.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele/haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.

Table 9A. Replication and combined analysis of *HLA-DRB1* alleles with childhood SSNS.

HLA-DRB1 alleles[a]	Replication stage					Combined analysis				
	Cases (2n=426)	Controls (2n=1420)	chi-square test			Cases (2n=874)	Controls (2n=2238)	chi-square test		
	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DRB1*01:01*	11 (2.6%)	86 (6.1%)	0.41 (0.22-0.78)	4.82E-03	9.16E-02	26 (3.0%)	143 (6.4%)	0.45 (0.29-0.69)	1.58E-04	3.17E-03
*DRB1*04:01*	5 (1.2%)	11 (0.8%)	1.52 (0.53-4.40)	NS	-	6 (0.7%)	21 (0.9%)	0.73 (0.29-1.81)	NS	-
*DRB1*04:03*	10 (2.3%)	39 (2.7%)	0.85 (0.42-1.72)	NS	-	28 (3.2%)	63 (2.8%)	1.14 (0.73-1.80)	NS	-
*DRB1*04:05*	71 (16.7%)	158 (11.1%)	1.60 (1.18-2.16)	2.35E-03	4.46E-02	128 (14.6%)	279 (12.5%)	1.20 (0.96-1.51)	NS	-
*DRB1*04:06*	4 (0.9%)	56 (3.9%)	0.23 (0.08-0.64)	2.16E-03	4.10E-02	20 (2.3%)	84 (3.8%)	0.60 (0.37-0.98)	4.10E-02	8.20E-01
*DRB1*04:10*	4 (0.9%)	19 (1.3%)	0.70 (0.24-2.07)	NS	-	12 (1.4%)	31 (1.4%)	0.99 (0.51-1.94)	NS	-
*DRB1*08:02*	44 (10.3%)	62 (4.4%)	2.52 (1.69-3.77)	3.49E-06	6.64E-05	90 (10.3%)	94 (4.2%)	2.62 (1.94-3.54)	9.11E-11	1.82E-09
*DRB1*08:03*	35 (8.2%)	110 (7.7%)	1.07 (0.72-1.59)	NS	-	68 (7.8%)	173 (7.7%)	1.01 (0.75-1.35)	NS	-
*DRB1*09:01*	83 (19.5%)	196 (13.8%)	1.51 (1.14-2.01)	4.09E-03	7.77E-02	162 (18.5%)	319 (14.3%)	1.37 (1.11-1.68)	2.98E-03	5.97E-02
*DRB1*11:01*	11 (2.6%)	44 (3.1%)	0.83 (0.42-1.62)	NS	-	27 (3.1%)	67 (3.0%)	1.03 (0.66-1.63)	NS	-
*DRB1*12:01*	23 (5.4%)	43 (3.0%)	1.83 (1.09-3.07)	2.08E-02	3.95E-01	51 (5.8%)	73 (3.3%)	1.84 (1.27-2.65)	9.72E-04	1.94E-02
*DRB1*12:02*	5 (1.2%)	19 (1.3%)	0.88 (0.33-2.36)	NS	-	8 (0.9%)	37 (1.7%)	0.55 (0.25-1.18)	NS	-
*DRB1*13:02*	7 (1.6%)	125 (8.8%)	0.17 (0.08-0.37)	4.91E-07	9.32E-06	12 (1.4%)	180 (8.0%)	0.16 (0.09-0.29)	3.65E-12	7.31E-11
*DRB1*14:03*	11 (2.6%)	22 (1.5%)	1.68 (0.81-3.50)	NS	-	22 (2.5%)	33 (1.5%)	1.73 (1.00-2.98)	4.73E-02	9.45E-01
*DRB1*14:05*	7 (1.6%)	32 (2.3%)	0.72 (0.32-1.65)	NS	-	18 (2.1%)	47 (2.1%)	0.98 (0.57-1.70)	NS	-
*DRB1*14:06*	7 (1.6%)	19 (1.3%)	1.23 (0.51-2.95)	NS	-	22 (2.5%)	32 (1.4%)	1.78 (1.03-3.08)	3.68E-02	7.37E-01
*DRB1*14:54*	2 (0.4%)	26 (1.8%)	/	/	/	9 (1.0%)	52 (2.3%)	0.44 (0.21-0.89)	1.93E-02	3.86E-01
*DRB1*15:01*	19 (4.5%)	113 (8.0%)	0.54 (0.33-0.89)	1.40E-02	2.66E-01	40 (4.6%)	180 (8.0%)	0.55 (0.39-0.78)	6.98E-04	1.40E-02
*DRB1*15:02*	52 (12.2%)	178 (12.5%)	0.97 (0.70-1.35)	NS	-	97 (11.1%)	248 (11.1%)	1.00 (0.78-1.28)	NS	-
*DRB1*16:02*	5 (1.2%)	12 (0.8%)	1.39 (0.49-3.98)	NS	-	9 (1.0%)	14 (0.6%)	1.65 (0.71-3.83)	NS	-

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
/:The *HLA-DRB1* allele was excluded from association analysis because of low frequency (<0.5%) in the replication stage.
NS: Not significant (*P*-value>=0.05).

-: The *P*-value before multiple correction was not significant (*P*>=0.05), the *P*-corrected was omitted.

Table 9B. Replication and combined analysis of *HLA-DQB1* alleles with childhood SSNS.

HLA-DQB1 alleles[a]	Replication stage					Combined analysis				
	Cases (2n=426)	Controls (2n=1420)	chi-square test			Cases (2n=874)	Controls (2n=2238)	chi-square test		
	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DQB1*03:01*	69 (16.2%)	151 (10.6%)	1.62 (1.19-2.21)	1.88E-03	2.07E-02	140 (16.0%)	250 (11.2%)	1.52 (1.21-1.90)	2.42E-04	2.66E-03
*DQB1*03:02*	48 (11.3%)	136 (9.6%)	1.20 (0.85-1.70)	NS	-	128 (14.6%)	213 (9.5%)	1.63 (1.29-2.06)	3.86E-05	4.25E-04
*DQB1*03:03*	82 (19.2%)	211 (14.9%)	1.37 (1.03-1.81)	2.97E-02	3.26E-01	173 (19.8%)	340 (15.2%)	1.38 (1.13-1.69)	1.87E-03	2.06E-02
*DQB1*04:01*	70 (16.4%)	158 (11.1%)	1.57 (1.16-2.13)	3.51E-03	3.86E-02	123 (14.1%)	279 (12.5%)	1.15 (0.92-1.44)	NS	-
*DQB1*04:02*	19 (4.5%)	50 (3.5%)	1.28 (0.75-2.19)	NS	-	36 (4.1%)	76 (3.4%)	1.22 (0.82-1.83)	NS	-
*DQB1*05:01*	12 (2.8%)	91 (6.4%)	0.42 (0.23-0.78)	4.62E-03	5.08E-02	28 (3.2%)	154 (6.9%)	0.45 (0.30-0.68)	8.53E-05	9.38E-04
*DQB1*05:02*	6 (1.4%)	32 (2.3%)	0.62 (0.26-1.49)	NS	-	12 (1.4%)	49 (2.2%)	0.62 (0.33-1.17)	NS	-
*DQB1*05:03*	8 (1.9%)	62 (4.4%)	0.42 (0.20-0.88)	1.84E-02	2.02E-01	22 (2.5%)	90 (4.0%)	0.62 (0.38-0.99)	4.29E-02	4.72E-01
*DQB1*06:01*	87 (20.4%)	284 (20.0%)	1.03 (0.78-1.34)	NS	-	163 (18.6%)	417 (18.6%)	1.00 (0.82-1.22)	NS	-
*DQB1*06:02*	14 (3.3%)	108 (7.6%)	0.41 (0.23-0.73)	1.65E-03	1.81E-02	29 (3.3%)	173 (7.7%)	0.41 (0.27-0.61)	7.13E-06	7.84E-05
*DQB1*06:04*	5 (1.2%)	117 (8.2%)	0.13 (0.05-0.33)	2.63E-07	2.89E-06	7 (0.8%)	170 (7.6%)	0.10 (0.05-0.21)	1.90E-13	2.09E-12

[a]Alleles: *HLA* alleles with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for allele frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 9C. Replication and combined analysis of *HLA-DRB1-DQB1* haplotypes with childhood SSNS.

HLA-DRB1-DQB1 haplotypes[a]	Replication stage					Combined analysis				
	Cases (2n=426)	Controls (2n=1420)	chi-square test			Cases (2n=874)	Controls (2n=2238)	chi-square test		
	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]	No (%)	No (%)	OR (95% CI)[b]	*P*-value	*P*-corrected[c]
*DRB1*01:01-DQB1*05:01*	11 (2.6%)	85 (6.0%)	0.42 (0.22-0.79)	5.52E-03	1.10E-01	26 (3.0%)	142 (6.3%)	0.45 (0.30-0.69)	1.85E-04	4.07E-03
*DRB1*04:01-DQB1*03:01*	5 (1.2%)	11 (0.8%)	1.52 (0.53-4.40)	NS	-	6 (0.7%)	21 (0.9%)	0.73 (0.29-1.81)	NS	-
*DRB1*04:03-DQB1*03:02*	10 (2.3%)	38 (2.7%)	0.87 (0.43-1.77)	NS	-	28 (3.2%)	61 (2.7%)	1.18 (0.75-1.86)	NS	-
*DRB1*04:05-DQB1*04:01*	70 (16.4%)	155 (10.9%)	1.60 (1.18-2.18)	2.27E-03	4.54E-02	123 (14.1%)	275 (12.3%)	1.17 (0.93-1.47)	NS	-
*DRB1*04:06-DQB1*03:02*	4 (0.9%)	56 (3.9%)	0.23 (0.08-0.64)	2.16E-03	4.32E-02	20 (2.3%)	84 (3.8%)	0.60 (0.37-0.98)	4.10E-02	9.02E-01
*DRB1*04:10-DQB1*04:02*	4 (0.9%)	17 (1.2%)	0.78 (0.26-2.34)	NS	-	12 (1.4%)	28 (1.3%)	1.10 (0.56-2.17)	NS	-
*DRB1*08:02-DQB1*03:02*	28 (6.6%)	31 (2.2%)	3.15 (1.87-5.32)	6.25E-06	1.25E-04	64 (7.3%)	48 (2.1%)	3.60 (2.46-5.29)	3.19E-12	7.01E-11
*DRB1*08:02-DQB1*04:02*	15 (3.5%)	30 (2.1%)	1.69 (0.90-3.17)	NS	-	24 (2.7%)	45 (2.0%)	1.38 (0.83-2.27)	NS	-
*DRB1*08:03-DQB1*06:01*	35 (8.2%)	106 (7.5%)	1.11 (0.74-1.65)	NS	-	66 (7.6%)	169 (7.6%)	1.00 (0.74-1.34)	NS	-
*DRB1*09:01-DQB1*03:03*	80 (18.8%)	193 (13.6%)	1.47 (1.10-1.96)	8.16E-03	1.63E-01	158 (18.1%)	312 (13.9%)	1.36 (1.10-1.68)	3.78E-03	8.31E-02
*DRB1*11:01-DQB1*03:01*	11 (2.6%)	43 (3.0%)	0.85 (0.43-1.66)	NS	-	25 (2.9%)	64 (2.9%)	1.00 (0.63-1.60)	NS	-
*DRB1*12:01-DQB1*03:01*	20 (4.7%)	26 (1.8%)	2.64 (1.46-4.78)	8.82E-04	1.76E-02	38 (4.3%)	48 (2.1%)	2.07 (1.35-3.20)	7.53E-04	1.66E-02
*DRB1*12:01-DQB1*03:03*	2 (0.4%)	17 (1.1%)	/	/	/	8 (0.9%)	23 (1.0%)	0.89 (0.40-2.00)	NS	-
*DRB1*12:02-DQB1*03:01*	5 (1.2%)	19 (1.3%)	0.88 (0.33-2.36)	NS	-	8 (0.9%)	37 (1.7%)	0.55 (0.25-1.18)	NS	-
*DRB1*13:02-DQB1*06:04*	5 (1.2%)	117 (8.2%)	0.13 (0.05-0.33)	2.63E-07	5.25E-06	7 (0.8%)	170 (7.6%)	0.10 (0.05-0.21)	1.90E-13	4.18E-12
*DRB1*14:03-DQB1*03:01*	11 (2.6%)	22 (1.5%)	1.68 (0.81-3.50)	NS	-	22 (2.5%)	33 (1.5%)	1.73 (1.00-2.98)	4.73E-02	1.04E+00
*DRB1*14:05-DQB1*05:03*	7 (1.6%)	29 (2.0%)	0.80 (0.35-1.84)	NS	-	16 (1.8%)	43 (1.9%)	0.95 (0.53-1.70)	NS	-
*DRB1*14:06-DQB1*03:01*	7 (1.6%)	19 (1.3%)	1.23 (0.51-2.95)	NS	-	22 (2.5%)	32 (1.4%)	1.78 (1.03-3.08)	3.68E-02	8.10E-01
*DRB1*14:54-DQB1*05:03*	1 (0.2%)	17 (1.2%)	/	/	/	6 (0.7%)	30 (1.3%)	0.51 (0.21-1.23)	NS	-
*DRB1*15:01-DQB1*06:02*	14 (3.3%)	107 (7.5%)	0.42 (0.24-0.74)	1.89E-03	3.77E-02	29 (3.3%)	172 (7.7%)	0.41 (0.28-0.62)	8.41E-06	1.85E-04
*DRB1*15:02-DQB1*06:01*	52 (12.2%)	177 (12.5%)	0.98 (0.70-1.36)	NS	-	97 (11.1%)	247 (11.0%)	1.01 (0.78-1.29)	NS	-
*DRB1*16:02-DQB1*05:02*	5 (1.2%)	12 (0.8%)	1.39 (0.49-3.98)	NS	-	9 (1.0%)	14 (0.6%)	1.65 (0.71-3.83)	NS	-

[a]Haplotypes: *HLA-DRB1-DQB1* haplotypes with frequencies < 0.5% in cases or controls are omitted.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.

[c]*P*-corrected: *P*-values for *HLA* haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
/:The *HLA-DRB1-DQB1* haplotype was excluded from association analysis because of low frequency (<0.5%) in the replication stage.
NS: Not significant (*P*-value>=0.05).
-: The *P*-value before multiple correction was not significant (*P*-value>=0.05), the *P*-corrected was omitted.

Table 10. Validation, replication and combined analyses of the significant *HLA-DRB1/-DQB1* alleles and *HLA-DRB1-DQB1* haplotypes detected in the discovery stage.

HLA alleles/ haplotypes[a]	Validation				Replication				Combined analysis			
	Cases (2n=448)	Controls (2n=818)	Chi-square test		Cases (2n=426)	Controls (2n=1,420)	Chi-square test		Cases (2n=874)	Controls (2n=2,238)	Chi-square test	
	No (%)	No (%)	OR (95% CI)[b]	*P*-corrected[c]	No (%)	No (%)	OR (95% CI)[b]	*P*-corrected[c]	No (%)	No (%)	OR (95% CI)[b]	*P*-corrected[c]
*DRB1*08:02*	46 (10.3%)	32 (3.9%)	2.81 (1.76-4.48)	1.24E-04	44 (10.3%)	62 (4.4%)	2.52 (1.69-3.77)	6.64E-05	90 (10.3%)	94 (4.2%)	2.62 (1.94-3.54)	1.82E-09
*DRB1*13:02*	5 (1.1%)	55 (6.7%)	0.16 (0.06-0.39)	1.28E-04	7 (1.6%)	125 (8.8%)	0.17 (0.08-0.37)	9.32E-06	12 (1.4%)	180 (8.0%)	0.16 (0.09-0.29)	7.31E-11
*DQB1*03:02*	80 (17.9%)	77 (9.4%)	2.09 (1.49-2.93)	1.44E-04	48 (11.3%)	136 (9.6%)	1.20 (0.85-1.70)	NS	128 (14.6%)	213 (9.5%)	1.63 (1.29-2.06)	4.25E-04
*DQB1*06:02*	15 (3.4%)	65 (8.0%)	0.40 (0.23-0.71)	1.43E-02	14 (3.3%)	108 (7.6%)	0.41 (0.23-0.73)	1.81E-02	29 (3.3%)	173 (7.7%)	0.41 (0.27-0.61)	7.84E-05
*DQB1*06:04*	2 (0.5%)	53 (6.5%)	0.06 (0.02-0.27)	5.26E-06	5 (1.2%)	117 (8.2%)	0.13 (0.05-0.33)	2.89E-07	7 (0.8%)	170 (7.6%)	0.10 (0.05-0.21)	2.09E-12
*DRB1*08:02-DQB1*03:02*	36 (8.0%)	17 (2.1%)	4.12 (2.28-7.42)	8.76E-06	28 (6.6%)	31 (2.2%)	3.15 (1.87-5.32)	1.25E-04	64 (7.3%)	48 (2.1%)	3.60 (2.46-5.29)	7.01E-11
*DRB1*13:02-DQB1*06:04*	2 (0.5%)	53 (6.5%)	0.06 (0.02-0.27)	1.00E-05	5 (1.2%)	117 (8.2%)	0.13 (0.05-0.33)	5.25E-06	7 (0.8%)	170 (7.6%)	0.10 (0.05-0.21)	4.18E-12
*DRB1*15:01-DQB1*06:02*	15 (3.4%)	65 (8.0%)	0.40 (0.23-0.71)	2.74E-02	14 (3.3%)	107 (7.5%)	0.42 (0.24-0.74)	3.77E-02	29 (3.3%)	172 (7.7%)	0.41 (0.28-0.62)	1.85E-04

[a]Alleles/haplotypes: The significant *HLA-DRB1/-DQB1* alleles and *HLA-DRB1-DQB1* haplotypes identified in the discovery stage using HLA-imputation data.
[b]OR: Odds ratio; 95% CI: lower and upper limits of confidence interval at 95%.
[c]*P*-corrected: *P*-values for haplotype frequency comparisons between cases and controls using the Pearson's chi-square test or Fisher's exact test and then corrected for the multiplicity of testing by the number of comparisons.
NS: Not significant (*P*-corrected>=0.05).

Table 11. *HLA* allele reciprocal conditional analyses.

HLA allele	TEST	OR (95% CI)	*P*
*HLA-DRB1*08:02*	Logistic regression	2.70 (1.98-3.67)	2.88E-10
	Condition on *DQB1*03:02*	2.41 (1.72-3.37)	3.13E-07
*HLA-DQB1*03:02*	Logistic regression	1.62 (1.28-2.04)	5.62E-05
	Condition on *DRB1*08:02*	1.25 (0.96-1.62)	0.10 (NS)
*HLA-DRB1*13:02*	Logistic regression	0.17 (0.09-0.30)	2.32E-09
	Condition on *DQB1*06:04*	1.68 (0.53-5.30)	0.37 (NS)
*HLA-DQB1*06:04*	Logistic regression	0.10 (0.05-0.21)	2.63E-09
	Condition on *DRB1*13:02*	0.06 (0.01-0.24)	7.22E-05
*HLA-DRB1*08:02*	Logistic regression	2.70 (1.98-3.67)	2.88E-10
	Condition on *DQB1*06:04*	2.48 (1.82-3.39)	1.14E-08
*HLA-DQB1*06:04*	Logistic regression	0.10 (0.05-0.21)	2.63E-09
	Condition on *DRB1*08:02*	0.11 (0.05-0.23)	8.47E-09

NS: Not significant (*P*>=0.05).

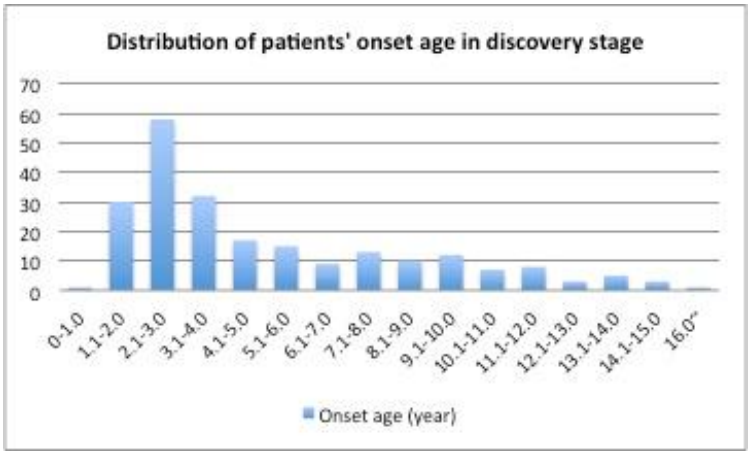
Table 12. Single-tag SNP (rs3129888) for capturing *HLA-DRB1*08:02* and *HLA-DRB1*08:02-DQB1*03:02.*

Tag SNP	*HLA-DRB1*08:02*	Other *HLA-DRB1* sub-alleles
rs3129888 (G)	177	63
rs3129888 (A)	7	2865
Total	184	2928

Tag SNP	*DRB1*08:02-DQB1*03:02*	Other *HLA-DRB1-DQB1* haplotypes
rs3129888 (G)	110	130
rs3129888 (A)	2	2870
Total	112	3000

Figure 1. Distribution of patients' onset age: No bias of onset-age between the discovery case samples and the replication case samples (P=0.12).
(A) Distribution of patients' onset age in discovery stage.
(B) Distribution of patients' onset age in replication stage.

A.



B.

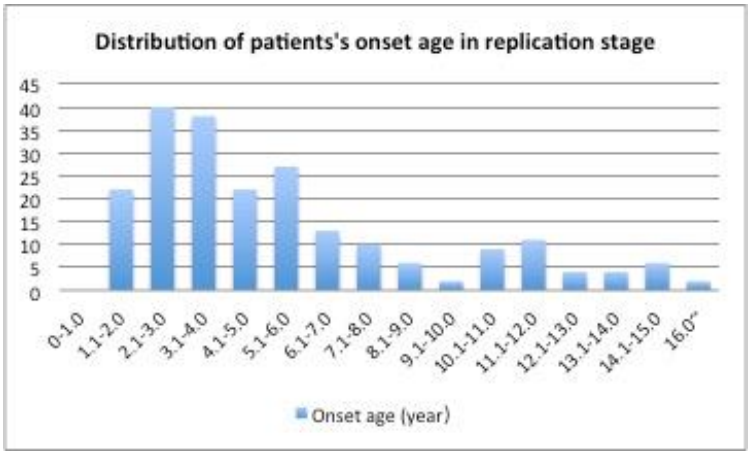
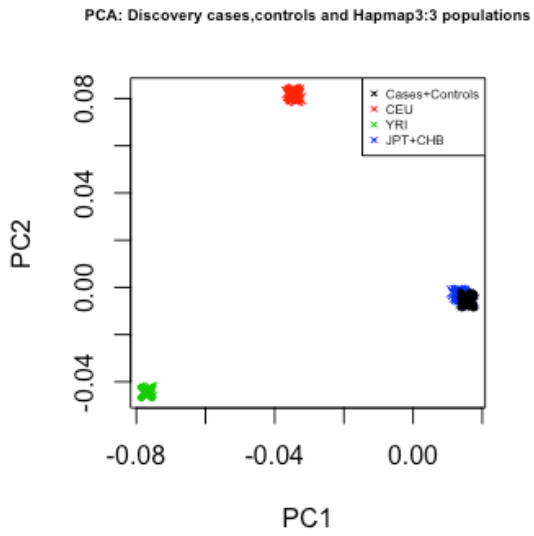
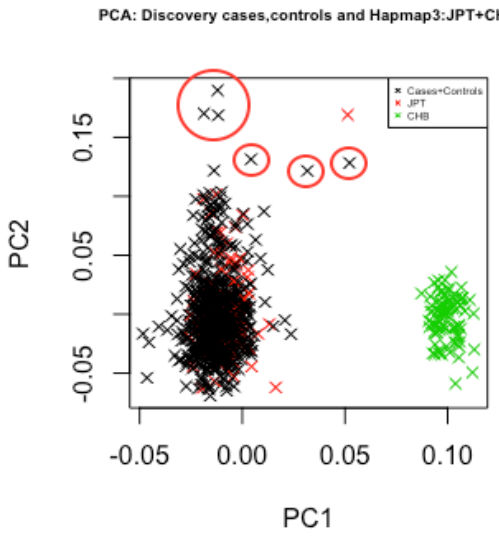


Figure 2. Principal component analysis of 642 samples (224 cases and 418 healthy controls) from the discovery stage of GWAS and HapMap Phase III samples (113 Utah residents with ancestry from northern and western Europe [CEU], 113 Yoruba in Ibadan [YRI], 84 Han Chinese in Beijing [CHB], and 86 Japanese in Tokyo [JPT]).
(**A**) Plotted all samples with global reference data (CEU, YRI and JPT+CHB), no outlier was detected.
(**B**) Focused on Asian samples (JPT and CHB): 6 controls around the JPT cluster were identified as outliers (highlighted by circles in red), and then excluded.
(**C**) 224 cases and 412 controls for association analyses, after the removal of outliers based on the results of PCA.

A.



B.



C.

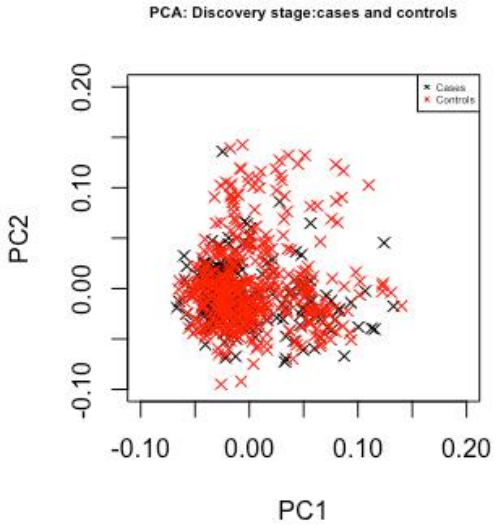


Figure 3. Quantile-quantile (Q-Q) plot of *P*-values for each SNP calculated using Cochran-Armitage trend test in the initial GWAS (224 cases with childhood SSNS and 412 healthy controls). The inflation factor, λ, was estimated to be 1.045 for all tested SNPs.

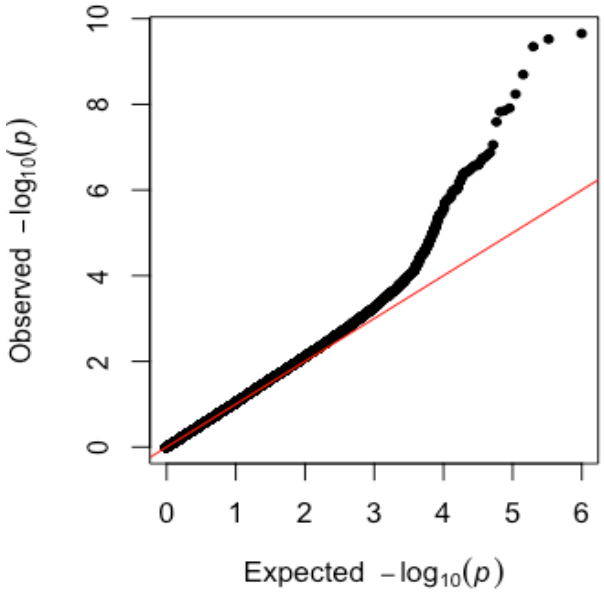


Figure 4. Results of the genome-wide association study (Manhattan plot) for childhood SSNS in the Japanese population.
P-values were calculated using Cochran-Armitage trend test in 224 cases and 412 controls in the discovery stage and adjusted by genomic control (GC-corrected *P*). The horizontal red and blue lines indicate the genome-wide significant threshold ($P=5\times10^{-8}$) and suggestive threshold ($P=1\times10^{-5}$), respectively.

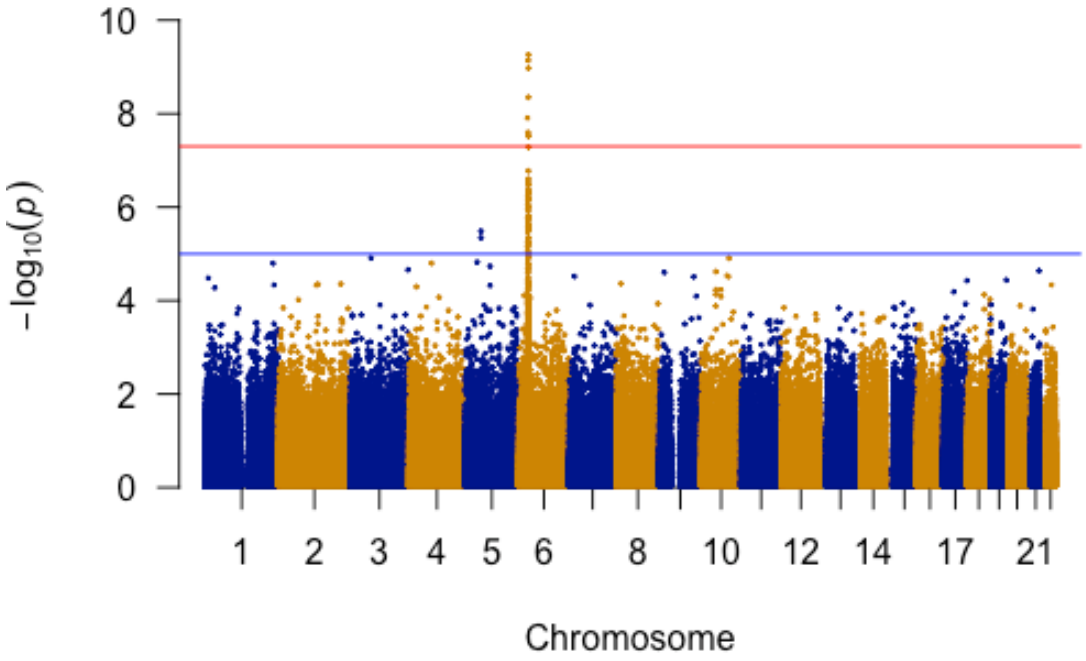


Figure 5. Regional plot of top signal in *HLA-DR/DQ* region in discovery GWAS.
The associations were calculated using Cochran-Armitage trend test and adjusted by genomic control (GC-corrected *P*). The purple dot indicates the top single nucleotide polymorphism in the association analysis (rs4642516, GC-corrected $P=5.44\times10^{-10}$).

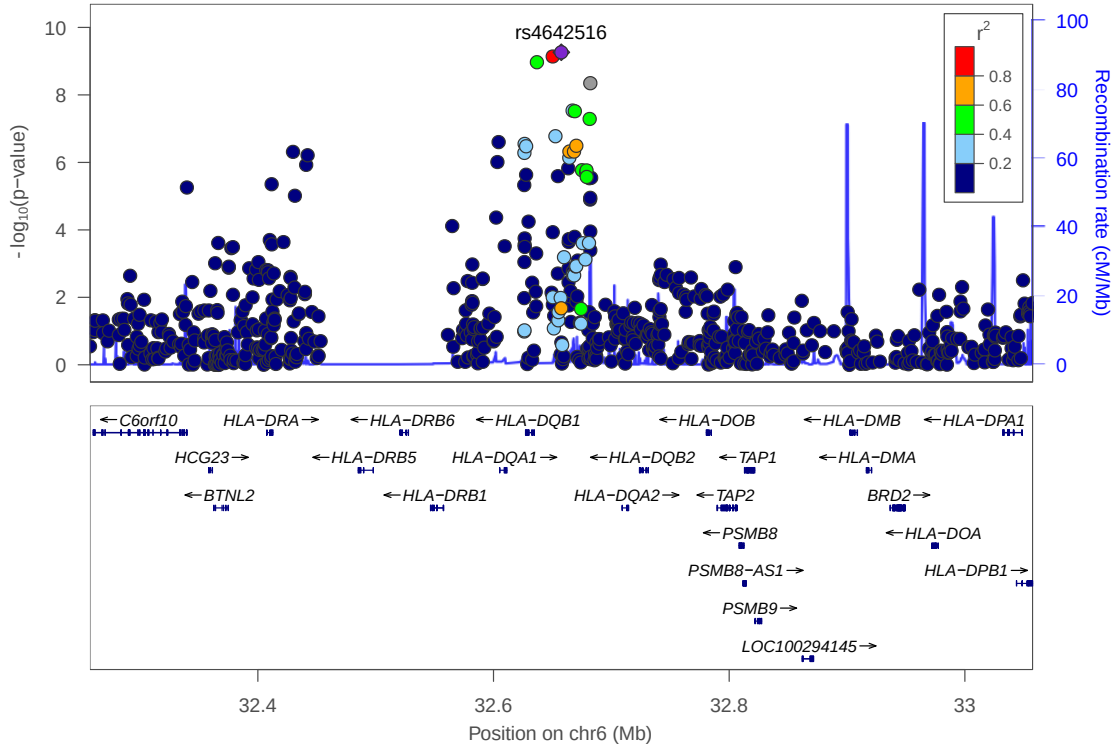
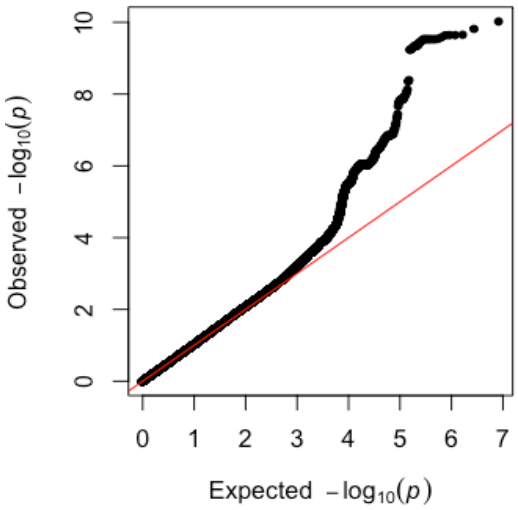


Figure 6. Quantile-quantile (Q-Q) plots of *P*-values for each variant calculated using Cochran-Armitage trend test after whole-genome imputation (224 cases with childhood SSNS and 412 healthy controls).
The inflation factor, λ, was estimated to be (**A**) 1.023 for all tested variations, including those in the *HLA* region, and (**B**) 1.019 when SNPs and INDELs in the *HLA* region (Hg19: chr6: 29,691,116–33,054,976) were excluded.

A.



B.

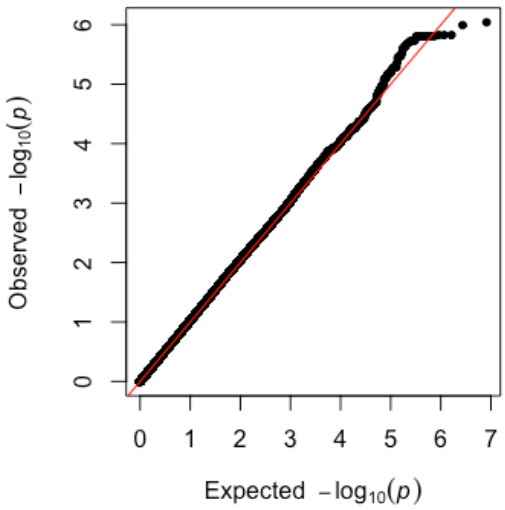


Figure 7. Whole-genome imputation was performed using the 2KJPN panel.
The associations were conducted using Cochran-Armitage trend test, *P*-values were corrected by genomic control. The horizontal red and blue lines indicate the genome-wide significant threshold ($P=5\times10^{-8}$) and suggestive threshold ($P=1\times10^{-5}$), respectively.

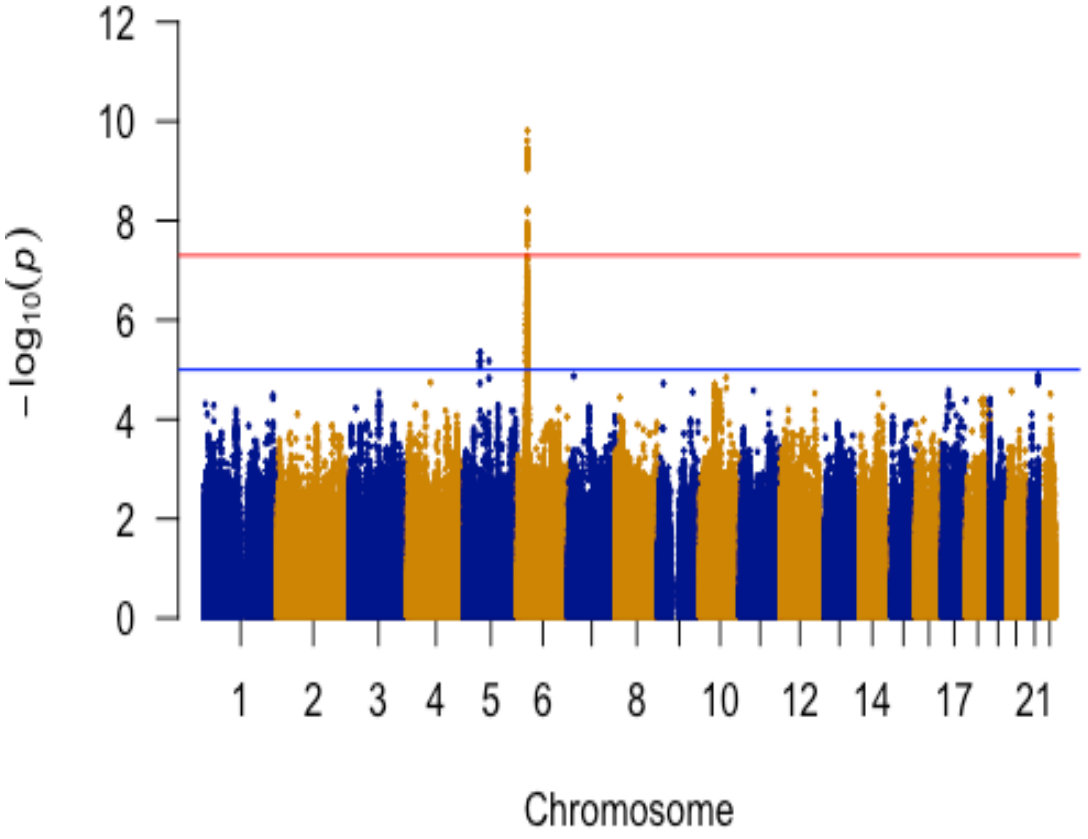


Figure 8. Regional plot of *HLA* region after conditioning on top typed SNP (rs4642516) using logistic regression.

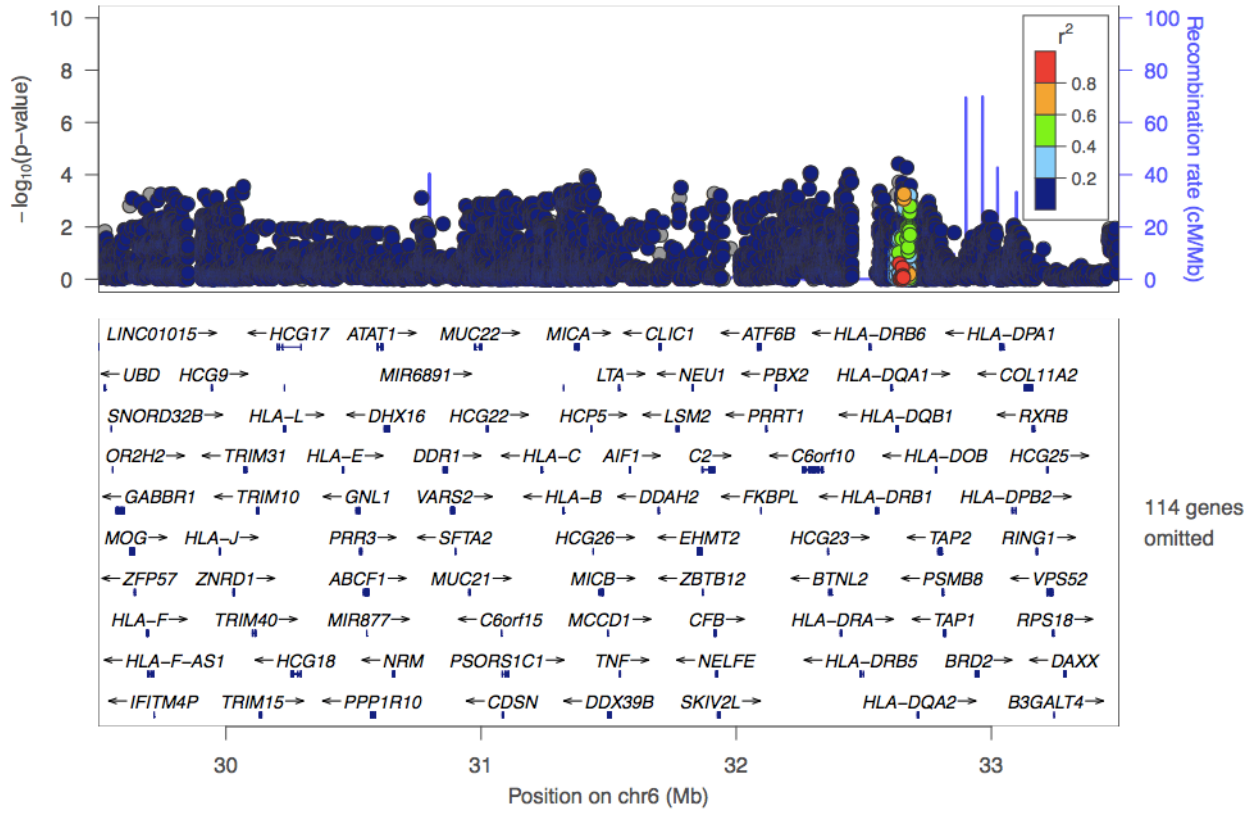


Figure 9. The study power of the current GWAS.
The power of our two-stage GWAS exceeded 80% to detect common alleles (minor allele frequency (MAF) >= 5%) with genotypic relative risk (RR) above 2.95, at a significant *P*-value threshold of 5×10^{-8} under additive model.