Mol Biol Rep
DOI 10.1007/s11033-012-1952-x

Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis

Young Ho Lee · Sang-Cheol Bae · Sung Jae Choi · Jong Dae Ji · Gwan Gyu Song

Received: 8 December 2011 / Accepted: 1 October 2012

Abstract The aim of this study was to explore candidate single nucleotide polymorphisms (SNPs) and candidate mechanisms of systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA). Two SLE genome-wide association studies (GWASs) datasets were included in this study. Meta-analysis was conducted using 737,984 SNPs in 1,527 SLE cases and 3,421 controls of European ancestry, and 4,429 SNPs that met a threshold of $p < 0.01$ in a Korean RA GWAS dataset was used. ICSNPathway (identify candidate causal SNPs and pathways) analysis was applied to the meta-analysis results of the SLE GWAS datasets, and a RA GWAS dataset. The most significant result of SLE GWAS meta-analysis concerned rs2051549 in the human leukocyte antigen (HLA) region ($p = 3.36\text{E}{-22}$). In the non-HLA region, meta-analysis identified 6 SNPs associated with SLE with genome-wide significance (STAT4, TNPO3, BLK, FAM167A, and IRF5). ICSNPathway identified five candidate causal SNPs and 13 candidate causal pathways. This pathway-based analysis provides three hypotheses of the biological mechanism involved. First, rs8084 and rs7192 → HLA-DRA → bystander B cell activation. Second, rs1800629 → TNF → cytokine network. Third, rs1150752 and rs185819 → TNXB → collagen metabolic process. ICSNPathway analysis identified three candidate causal non-HLA SNPs and four candidate causal pathways involving the PADI4, MTR, PADI2, and TPH2 genes of RA. We identified five candidate SNPs and thirteen pathways, involving bystander B cell activation, cytokine network, and collagen metabolic processing, which may contribute to SLE susceptibility, and we revealed candidate causal non-HLA SNPs, genes, and pathways of RA.

Keywords Genome-wide association studies · Meta-analysis · Pathway-based analysis · Systemic lupus erythematosus · Rheumatoid arthritis

Introduction

Systemic lupus erythematosus (SLE) is a prototypic systemic autoimmune disease, characterized by autoantibody production and immune complex formation leading to intense inflammation and multiple organ damage [1, 2]. Rheumatoid arthritis (RA) is a chronic inflammatory disease of predominantly synovial joints, and currently affects up to 1 % of the world's population [3]. Although the etiopathology of SLE and RA remains largely unknown, a strong genetic component is known to influence disease susceptibility and to modify its clinical manifestations [4].

Genome-wide association studies (GWASs) provide a powerful means of identifying genes associated with complex diseases. Furthermore, the number of GWASs conducted is growing rapidly and this has resulted in the discovery and replication of new disease-related genes [5]. Although large-scale GWASs have been carried out on

Electronic supplementary material The online version of this article (doi:10.1007/s11033-012-1952-x) contains supplementary material, which is available to authorized users.

Y. H. Lee (✉) · S. J. Choi · J. D. Ji · G. G. Song
Division of Rheumatology, Department of Internal Medicine, Korea University Anam Hospital, Korea University College of Medicine, 126-1, 5-ga, Anam-dong, Seongbuk-gu, Seoul 136-705, Korea
e-mail: lyhcgh@korea.ac.kr

S.-C. Bae
Division of Rheumatology, Department of Internal Medicine, The Hospital for Rheumatic Diseases, Hanyang University Medical Center, Seoul, Korea

Published online: 07 October 2012

complex diseases like SLE, much of the genetic component of SLE remains unexplained. Meta-analysis of studies conducted in same populations enhance power to detect further novel loci by increasing sample size. In particular, meta-analysis of GWAS datasets increases power to detect association signals by increasing sample size and by allowing the examination of more variants [6].

The availability of increasing numbers of GWAS datasets provides powerful opportunities to conduct research. Researchers can use datasets from a centralized GWAS database like dbGAP at the NCBI, which contains voluntarily registered data and results (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap). Although SLE GWAS data show that the human leukocyte antigen (HLA) region on chromosome 6p plays a key role in SLE susceptibility, other genes are known to contribute [7–9]. In fact, many genes or genetic variants make small contributions by interacting with each other to cause SLE. However, genetic signal has been examined at single marker level in the SLE GWASs and biological mechanisms are unresolved. One of the key challenges of GWAS data interpretation is the identification of causal SNPs and the provision of evidence and hypotheses regarding the mechanism through which they act. Using new methods to study to probe existing GWAS data may provide additional biological insights and highlight new candidate genes. ICSNPathway (identify candidate causal SNPs and pathways) was developed to identify candidate causal SNPs and their corresponding candidate causal pathways from GWAS data by integrating linkage disequilibrium (LD) analysis, functional SNP annotation, and pathway-based analysis (PBA) [10].

In the present study, we performed meta-analysis on available SLE GWAS datasets. For functional analysis, we applied ICSNPathway analysis to the meta-analysis data of SLE GWAS datasets, and a RA GWAS dataset to explore candidate SNPs and candidate mechanisms of SLE and RA, and to generate a hypothesis for SNP → gene → pathways in SLE and RA.

Methods

Study populations

We used two SLE GWAS datasets extracted from the NCBI dbGap (http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap) (Table 1). The study conducted by Hom et al. [11] included genotypes of 489,876 SNPs from the Illumina HumanHap550 Genotyping Bead-Chip (Illumina Inc., San Diego, CA, USA). This dataset comprises 1,311 SLE cases and 3,583 controls, and all subjects were European. In the second study, Harley et al. [12] genotyped 258,402 SNPs on the Illumina HumanHap 300 platform in 216 SLE cases and 81 controls of European ancestry. Each dataset was filtered to remove individuals with $p < 0.0001$ for Hardy–Weinberg violation and a call rate <95 % to reduce the impacts of genotyping errors. These 488,678 and 249,306 SNPs passed quality control filters in each dataset. We used a publicly available Korean RA GWAS dataset derived by Freudenberg et al. [13], which included the genotypes of 441,398 SNPs obtained from ILLUMINA 550 k Genotyping Chips. This dataset comprised 801 RA cases and 757 controls; all subjects were Korean. And, we used all 4,429 SNPs that satisfied a p value threshold of <0.01.

GWAS meta-analysis and statistics

Point estimates of risks, odds ratios (ORs), and 95 % confidence intervals (CIs) were estimated for each study. In addition, within- and between-study variations and heterogeneities were assessed using Cochran's Q-statistic. This heterogeneity test assesses the null hypothesis that all studies evaluated the same effect. The effect of heterogeneity was quantified using I^2, which ranges between 0 and 100 %, and represents the proportion of between-study variability attributable to heterogeneity rather than chance [14]. The fixed-effects model assumes that a genetic factor has a similar effect on SLE susceptibility across all studies investigated, and that observed variations among studies are caused by chance alone [15]. The random-effects model assumes that different studies show substantial diversity and assesses both within-study sampling error and between-study variance [16]. When study groups are homogeneous, the two models are similar, but if this is not the case, the random-effects model usually provides wider CIs than the fixed effects model. The random-effects model is best used in the presence of significant between-study heterogeneity [16]. We used the random-effects model for this GWAS meta-analysis since it allows for the heterogeneity of allelic effects between studies. Statistical manipulations were undertaken using Genome-Wide Association Meta-Analysis (GWAMA) program (company, city, state if USA, country) [6]. For each SNP, GWAMA outputs summary information and statistics. Outputs were used with R scripts, supplied with the software, to generate Quantile–Quantile and Manhattan plots that summarized the genome-wide meta-analysis. The genome-wide significance threshold was defined a priori as $p < 5 \times 10^{-8}$, the Bonferroni adjustment for 1 million independent tests.

When a candidate SNP was not present on a particular genotyping array, proxy SNPs in LD with that candidate SNP can be identified based on observed LD patterns in HapMap. Thus, SNAP, a tool for identifying and annotating proxy SNPs using HapMap, was used [17].

Table 1 Characteristics of individual SLE GWAS included in the meta-analysis

GWAS dataset [ref]	Ethnicity	Case	Control	Number of SNPs	Number of SNPs (post QC)	Chip platform
Hom et al. [11]	European	1,311	3,340	489,876	488,678	HumanHap550
Harley et al. [12]	European	216	81	258,402	249,306	HumanHap300

GWAS genome-wide association study, *SNP* single nucleotide polymorphism

Identification of candidate causal SNPs and pathways

We applied ICSNPathway analysis to the meta-analysis results obtained from the two SLE GWASs, and a Korean RA GWAS. ICSNPathway implements a two-stage analysis. The first stage involves pre-selecting candidate causal SNPs by LD analysis followed by functional SNP annotation based on the most significant SNPs found. The second stage involves the annotation of biological mechanisms of the pre-selected candidate causal SNPs using PBA. We inputted a full list of GWAS SNP p values to analyze the ICSNPathway. One was within gene, meaning that only p-values of SNPs located within genes were utilized in the PBA, and the other was a false discovery rate cutoff (0.05) for multiple testing correction of PBA. To avoid stochastic bias and the testing to general biological processes, we discarded pathways that contained <5 or >20 genes.

Results

SLE GWAS meta-analysis

Meta-analysis was conducted using 737,984 SNPs in 1,527 SLE cases and 3,421 controls of European ancestry. A comparison of the observed meta-analysis p-values and p-values for a null distribution is shown in Fig. 1a. Significant deviation from the null distribution was observed at the tail of the distribution, suggesting the presence of true positive associations. The most significant result of SLE GWAS meta-analysis was obtained for rs2051549 in the HLA region ($p = 3.36\text{E}{-22}$), though an additional 103 SNPs had observed p-values of less than 5×10^{-8} (Fig. 1b). All SNPs were located in chromosome 6, except for 6 (Table 2; Fig. 2). Outside the HLA region, we found 6 SNPs associated with SLE with genome-wide significance. Genome-wide significant associations between STAT4, TNPO3, BLK, FAM167A, and IRF5 and SLE were also observed (Table 2).

Candidate causal SNPs and pathways derived from meta-analysis of the two SLE GWASs

Utilizing the 492,420 GWAS SNPs p-values as input and the 104 most significant SNPs $p < 5 \times 10^{-8}$, ICSNPathway analysis identified five candidate causal SNPs and 13 candidate causal pathways (Tables 3, 4). SNP rs8084 was found to be in LD with rs3130320 ($r^2 = 0.808$), which had genome-wide significance in the original GWAS ($p = 4.44\text{E}{-11}$). Rs7192, which did not reach genome-wide significance in the original GWAS ($p = 0.788$), was found to be in LD with rs9268832 ($r^2 = 0.961$), which had genome-wide significance in the original GWAS ($p = 2.38\text{E}{-10}$). Rs1800629 and rs1150752, which were not represented in the original GWAS, were found to be in LD with rs2857595 and rs7775397 ($r^2 = 0.841,\ 0.899$), which had genome-wide significance in the original GWAS ($p = 8.07\text{E}{-15}$ and $2.44\text{E}{-19}$). Finally, rs185819, which reached genome-wide significance in the original GWAS ($p = 6.32\text{E}{-14}$), was not found to be in LD with any SNP (Table 3).

These five candidate SNPs and 13 candidate pathways provide three hypothetical biological mechanisms. First, rs8084 (essential spice site and intronic) and rs7192 (nonsynonymous coding) → HLA-DRA → bystander B cell activation, antigen dependent B cell activation, the role of eosinophils in the chemokine network of allergy, Th1/Th2 differentiation, activation of Csk by cAMP-dependent protein kinase, IL-5 signaling pathway, B lymphocyte cell surface molecules, and Lck and Fyn tyrosine kinases in the initiation of TCR activation. Second, rs1800629 (nonsynonymous coding) → TNF → cytokine network, regulation of transcriptional activity by PML, stress induction of HSP regulation, and Msp/Ron receptor signaling pathway. Third, rs1150752 [nonsynonymous coding (deleterious) and rs185819 (nonsynonymous coding (deleterious)] → TNXB → collagen metabolic process.

Candidate causal non-HLA SNPs and pathways of RA

Utilizing the 4,429 GWAS SNPs p-values reaching a threshold of $p < 0.01$ as input, ICSNPathway analysis identified three candidate causal SNPs and four candidate causal pathways (Supplementary Tables 1 and 2, Supplementary Fig. 1). SNP rs874881, which was not represented in the original GWAS, was found to be in LD with rs11203366 ($r^2 = 0.887$), which was significant in the original GWAS ($p = 1.57\text{E}{-07}$). Rs11203367, which was not represented in the original GWAS, was found to be in LD with rs11203368 ($r^2 = 0.846$), which was significant in the original GWAS ($p = 2.61\text{E}{-08}$). Rs11203366, which

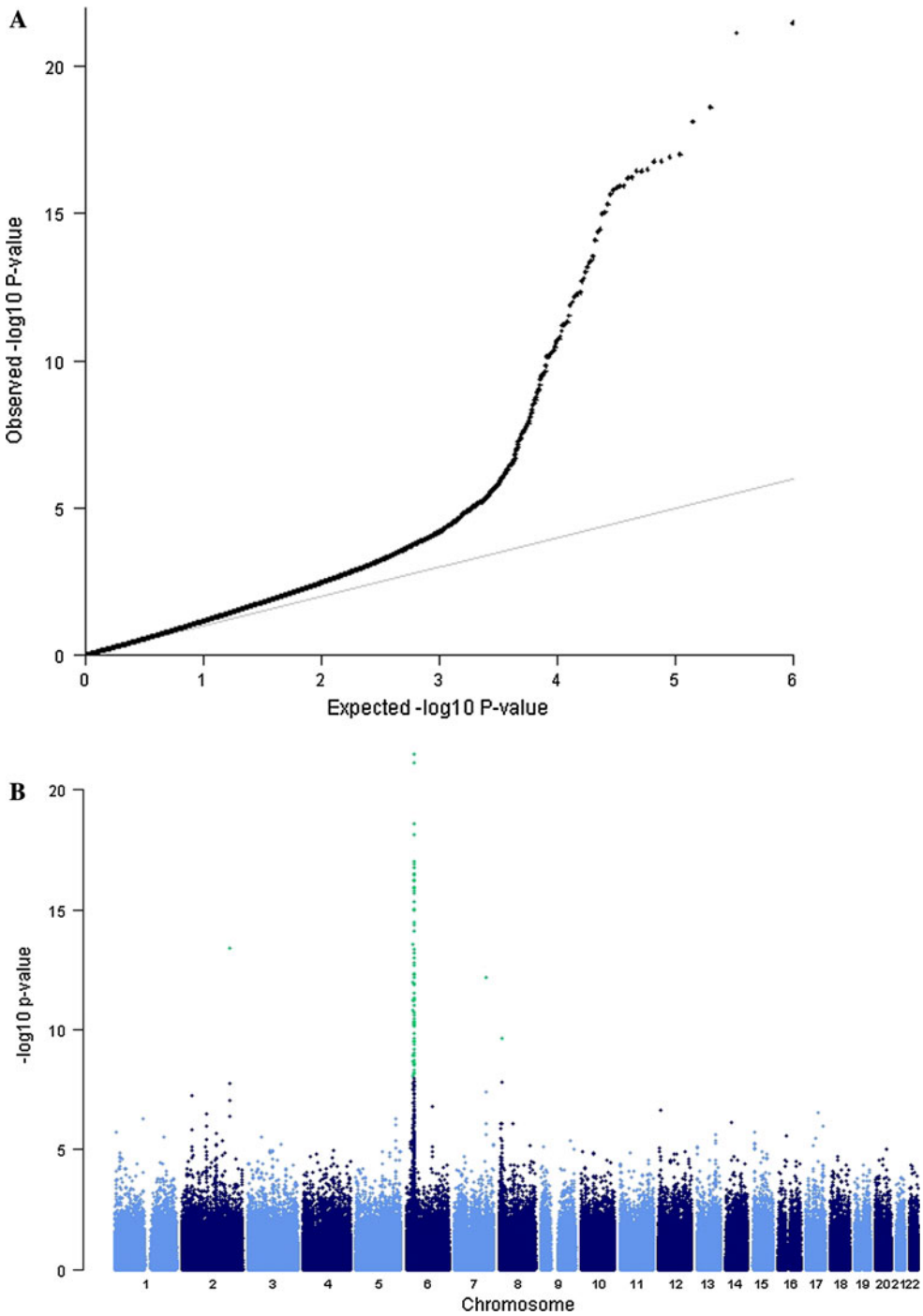

Fig. 1 Quantile–Quantile plot showing the expected and observed distributions of *p*-values in the meta-analysis (**a**), and a Manhattan plot showing *p*-values by chromosome for the meta-analyses (**b**)

was represented in the original GWAS ($p = 1.57E{-}07$), was not found to be in LD with any SNPs.

These three candidate causal SNPs and four candidate causal pathways provided one hypothetical biological mechanism, namely, rs874881 (nonsynonymous coding), rs11203367 (nonsynonymous coding (deleterious)), and rs11203366 (nonsynonymous coding) → peptidylarginine deiminases (PADI)4 → amino acid metabolic process, chromosome organization and biogenesis, nitrogen compound biosynthetic process, and carboxylic acid metabolic process. We then sought to identify genes that play roles in the four identified pathways. Distinct clustering of genes involved in the four candidate causal pathways was observed (Supplementary Tables 3–6). Further examination of the gene contents of these four pathways revealed some overlap, and it was observed that the pathways were mostly driven by the PADI4 gene, which was involved in all four pathways. MTR, PADI2, and TPH2 were involved in three of

the four pathways, and BCKDHA, CARS2, DCT, FARS2, GFPT2, MSRA, PADI3, TDO2 in 2, and many genes, such as, ACADM, ALOX5, ARID1B, ATF3, CARM1, CDYL2, CHD9, CKMT2, CSGALNACT1, CYP4F8, and EHMT1 were involved in only one pathway only (Table 6 in Supplementary material). Genes without an obvious immunological function are of interest as they could lead to the discovery of novel mechanisms. The present analysis suggests roles for pathways and genes in RA susceptibility (summarized in Supplementary Tables 3–5).

Discussion

GWAS provide a successful means of identifying novel common genetic variants that contribute to susceptibility to complex diseases [5]. However, a single GWAS has limited power to identify new loci, because a limited set of variants are genotyped, and the reported variant is unlikely to be the causal variant, but rather to be in LD with relevant variants. Reported loci are only those that meet stringent statistical "genome-wide" significance criteria. Thus, combined analysis of GWASs, an extremely large GWAS, or PBA is required to identify further new loci associated with susceptibility to complex diseases like SLE [18]. In particular, combining the results of multiple GWAS datasets may strengthen associations for previously identified loci and suggest new disease loci, or pathways [19]. Furthermore, multiple related genes in the same pathway may work together to confer disease susceptibility, and some genes may not reach genome-wide significance in any GWAS.

It is well known that SLE is caused by interactions between multiple genetic factors and environmental factors, and that complex molecular networks and cellular pathways play key roles in development of SLE. Pathways represent the combined genetic effect and candidate causal SNPs are supported to have higher confidence to be true. If a specific pathway is relevant to disease susceptibility, association signals would be expected to be overrepresented among SNPs in genes involved in the pathway.

In the present study, we identified five candidate causal SNPs and 13 pathways with significant nominal *p*-values by ICSNPathway. Our meta-analysis of two GWASs confirmed evidence for the majority of known susceptibility loci in HLA-DR. The strongest associations found in this meta-analysis were for known loci in HLA-DR. In addition, associations with STAT4, IRF5, TNPO3, BLK, and FAM167A reached genome-wide significance. It is known that STAT4, BLK, FAM167A, and IRF5 play key roles in susceptibility to SLE [20–22], but the contribution made by TNPO3, which is located next to IRF5, has not been studied yet in SLE. Modest associations can be detected by analyzing biological meaningful pathways [23]. Given the limited power of GWASs to detect single SNP associations, we utilized a pathway-based approach, because it takes into account biological interplay among genes. Furthermore, this approach provides insights as to how multiple genes might contribute to the pathogeneses of diseases [24]. Although there are several ways of annotating pathways, we chose; the Kyoto Encyclopedia of Genes and Genomes (KEGG) [25], BioCarta, GO biological process, and GO molecular function [26] to ensure comprehensive coverage of pathways and high-quality information for well-defined pathways. In this genome-wide search for pathways associated with SLE, the most strongly associated pathway was related to bystander B cell activation, which is consistent with the well-known role of B cells in the pathogenesis of SLE.

ICSNPathway analysis identified five candidate causal SNPs and three hypothetical mechanisms showing strong evidence for the majority of known susceptibility loci in HLA-DR (rs8084, rs7192). In addition, associations with TNF (rs1800629) and TNXB (rs1150752 and rs185819) reached genome-wide significance. It is well known that TNF and TNXB play key roles in susceptibility to SLE. Three hypothetical mechanisms included three gens such as HLA-DR, TNF and TNXB. SLE-autoantibodies are

Table 2 Non-HLA SNPs reaching $p < 5 \times 10^{-8}$ in the SLE GWAS meat-analysis

SNPs	Chromosome	Position (bp)	Nearby gene	Allele		OR	95 % CI	*p*-value
				Effect	Other			
rs7574865	2	191672878	STAT4	T	G	1.477	1.335–1.634	4.06E−14
rs12531711	7	128404702	TNPO3	A	G	1.593	1.403–1.808	6.41E−13
rs13277113	8	11386595	BLK	A	G	1.391	1.256–1.540	2.28E−10
rs12680762	8	11369436	FAM167A	A	G	1.335	1.208–1.475	1.45E−08
rs10931481	2	191663097	STAT4	T	C	1.312	1.194–1.442	1.74E−08
rs729302	7	128356196	IRF5	A	C	0.774	0.707–0.848	3.93E−08

HLA human leukocyte antigen, *SNP* single nucleotide polymorphism, *SLE* systemic lupus erythematosus, *GWAS* genome-wide association study, *OR* odds ratio, *CI* confidence interval

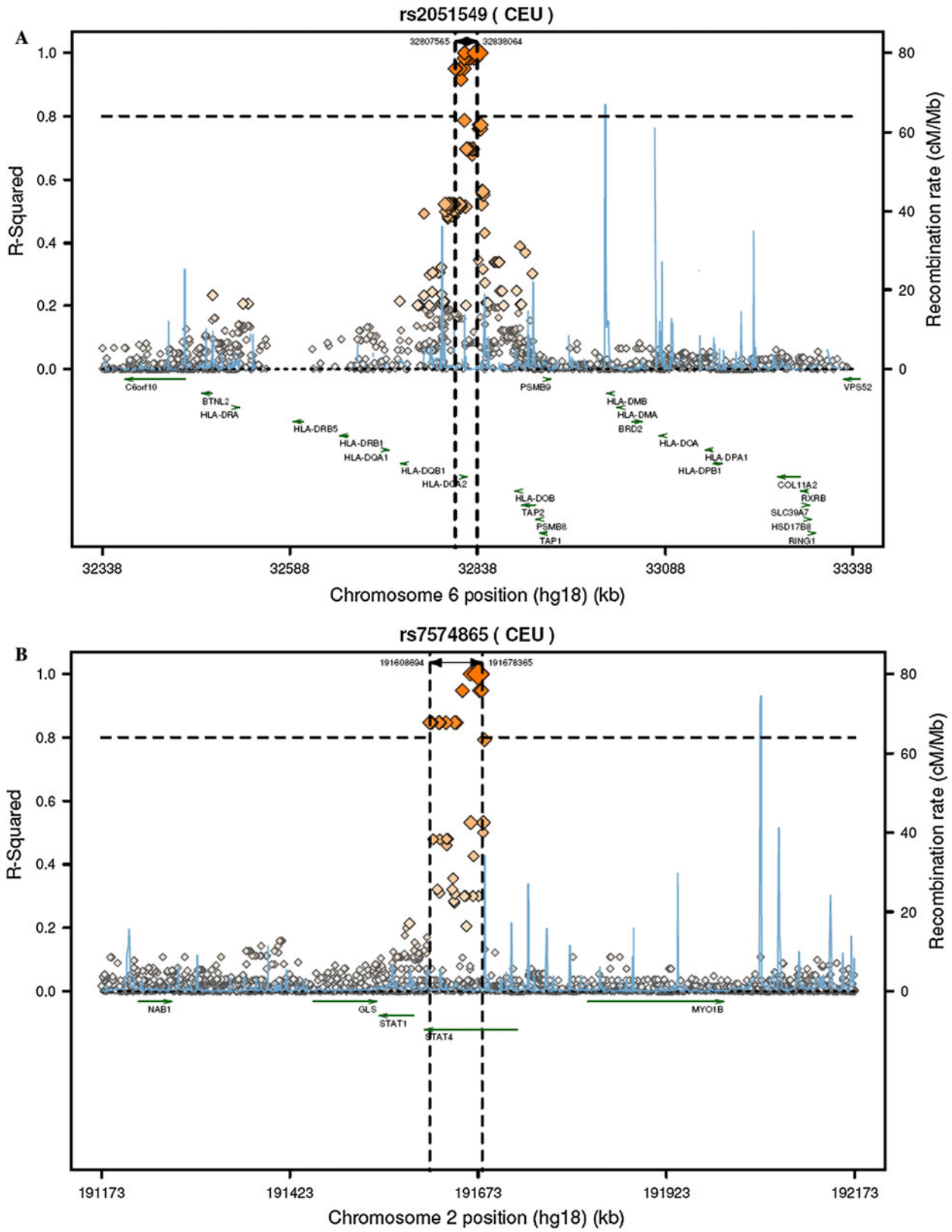

Fig. 2 Regional LD plots for the SNPs; rs 2051549 (HLA-DQA2) (**a**), rs7574865 (STAT4) (**b**), rs729302 (IRF5) (**c**), and rs13277113 (BLK) (**d**). SNPs are plotted along with their proxies (based on HAPMAP CEU) as a function of genomic location, and are annotated by recombination rate across the locus (*light-blue line*). On the y-axis, pairwise r^2 is provided for each proxy SNP, using color shading

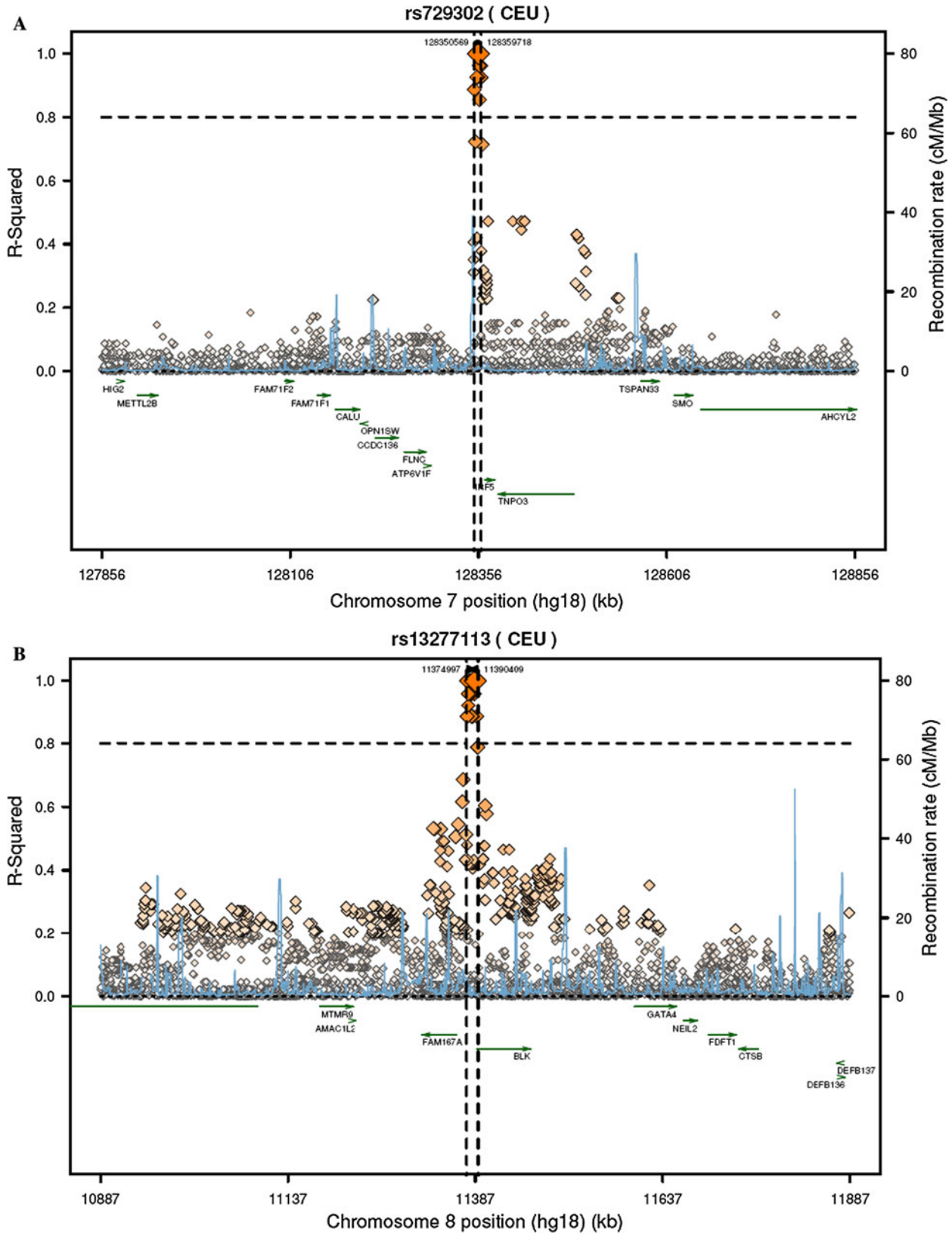

Fig. 2 continued

Table 3 Candidate SNPs of SLE

Candidate causal SNP	Functional class	Gene	Candidate causal pathway[a]	$-\log_{10}(P)$[b]	In LD with	r^2	D'	$-\log_{10}(P)$[c]
rs8084	Essential splice site & intronic	*HLA-DRA*	1 4 5 8 10 11 12 13	–	rs3130320	0.808	0.917	10.353
rs7192	Nonsynonymous coding	*HLA-DRA*	1 4 5 8 10 11 12 13	0.103	rs9268832	0.961	1.0	9.623
rs1800629	Nonsynonymous coding	*TNF*	2 3 6 7	–	rs2857595	0.841	0.943	14.093
rs1150752	Nonsynonymous coding (deleterious)	*TNXB*	9	–	rs7775397	0.899	1.0	18.613
rs185819	Nonsynonymous coding (deleterious)	*TNXB*	9	13.199	rs185819	–	–	13.199

SNP single nucleotide polymorphism, *SLE* systemic lupus erythematosus, *LD* linkage disequilibrium, *GWAS* genome-wide association study

[a] The number indicates the index of pathways, which are ranked by their statistical significance (false discovery rate)

[b] $-\log_{10}(P)$ for candidate causal SNP in original GWAS. –, denotes that this SNP is not represented in the original GWAS

[c] $-\log_{10}(P)$ for the SNP (which candidate causal SNP is in LD with) in original GWAS

Table 4 Candidate pathways of SLE

Index	Candidate causal pathway	Nominal *p*	FDR
1	Bystander B cell activation	<0.001	<0.001
2	Cytokine network	0.001	0.003
3	Regulation of transcriptional activity by PML	<0.001	0.003
4	Antigen dependent B cell activation	<0.001	0.003
5	The role of eosinophils in the chemokine network of allergy	0.001	0.004
6	Stress induction of HSP regulation	0.001	0.004
7	Msp/Ron receptor signaling pathway	0.018	0.013
8	Th1/Th2 differentiation	0.002	0.014
9	Collagen metabolic process	0.004	0.015
10	Activation of Csk by cAMP-dependent protein kinase	0.005	0.023
11	IL-5 signaling pathway	0.008	0.023
12	B lymphocyte cell surface molecules	0.007	0.031
13	Lck and Fyn tyrosine kinases in initiation of TCR activation	0.004	0.032

SLE systemic lupus erythematosus, *FDR* false discovery rate

initiated by environmental T cell epitope mimics of the SLE-related autoantigens in hosts with susceptible HLA-DR alleles. TNF is a pleiotropic cytokine that produces different stimuli in various physiological and pathological conditions. TNF contributes importantly to the development of T cells, B cells, and dendritic cells. The TNF gene polymorphism has been involved in the susceptibility of SLE. TNXB plays role in interactions between cell and extracellular matrix, antiadhesive effect, inhibiting cell migration, and in maintaining homeostasis of the extracellular matrix. However, the role of TNXB in immune system remains unclear. Further studies are needed to clarify the role of TNXB in SLE.

The ICSNPathway and other pathway-based approaches were developed to address the challenge posed by GWAS data interpretation. However, there is no accepted way of comparing the various pathway analysis methods. Furthermore, the incomplete annotation of the human genome represents an important limitation of the use of the pathway-based approach for GWAS analysis. The limitation of the ICSNPathway should be discussed. The ICSNPathway is not intended to predict true causal SNPs and pathways, because of limited understanding of their genetic basis in complex diseases [10]. There is no concrete evidence to be used to establish the predictive properties for ICSNPathway [10]. Furthermore, some human genes are uncharacterized and have not been mapped to predicted pathways, so their effects cannot be accounted for in pathway association analysis. Accordingly, pathway analysis for GWAS data is under developed, and thus, results should be interpreted with caution. Nevertheless, pathway-based approaches can play a complementary role in the identification of novel genes that confer disease susceptibility. The goal of pathway-based approaches is not to replace conventional single-maker analysis but to play a complementary part in identifying novel gene or sets of genes that confer disease susceptibility [27]. By taking into account prior biological knowledge about genes and pathways, we may have a better chance to identify the genes and mechanisms that are

involved in disease pathogenesis. An important application of the ICSNPathway results is to allow investigators to test a hypothesis concerning pathways by using candidate causal SNPs. Our results obtained using pathway association approaches using ICSNPathway could result in the formulation of new hypotheses.

In conclusion, we performed meta-analysis on the results of two SLE GWAS datasets to identify genetic associations with SLE at the SNP and pathway levels. The pathway-based approach used in the present study may complement the single-marker analysis of GWAS data. We identified five candidate SNPs and thirteen pathways involved in bystander B cell activation, the cytokine network, and collagen metabolic process that may contribute to SLE susceptibility. We identified three candidate non-HLA causal SNPs, many genes, and four pathways, involving PADI4, MTR, PADI2, and TPH2, which might contribute to RA susceptibility. Further studies are needed to confirm and explore genetic variations of these pathways in SLE and RA.

Acknowledgments This study is supported by a Grant from the Korea Healthcare Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (A102065).

Conflict of interest We have no financial and non-financial conflicts of interest.

References

1. Tsokos GC (2011) Systemic lupus erythematosus. N Engl J Med 365:2110–2121
2. Kang SH, Chung BH, Choi SR et al (2011) Comparison of clinical outcomes by different renal replacement therapy in patients with end-stage renal disease secondary to lupus nephritis. Korean J Intern Med 26:60–67
3. Harris ED Jr (1990) Rheumatoid arthritis. Pathophysiology and implications for therapy. N Engl J Med 322:1277–1289
4. Lee YH, Nath SK (2005) Systemic lupus erythematosus susceptibility loci defined by genome scan meta-analysis. Hum Genet 118:434–443
5. Johnson AD, O'Donnell CJ (2009) An open access database of genome-wide association results. BMC Med Genet 10:6
6. Egger M, Davey Smith G, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. BMJ 315:629–634
7. Lee YH, Witte T, Momot T et al (2005) The mannose-binding lectin gene polymorphisms and systemic lupus erythematosus: two case-control studies and a meta-analysis. Arthritis Rheum 52:3966–3974
8. Lee YH, Harley JB, Nath SK (2006) Meta-analysis of TNF-alpha promoter −308 A/G polymorphism and SLE susceptibility. Eur J Hum Genet 14:364–371
9. Lee YH, Ji JD, Song GG (2009) Fcgamma receptor IIB and IIIB polymorphisms and susceptibility to systemic lupus erythematosus and lupus nephritis: a meta-analysis. Lupus 18:727–734
10. Zhang K, Chang S, Cui S, Guo L, Zhang L, Wang J (2011) ICSNPathway: identify candidate causal SNPs and pathways from genome-wide association study by one analytical framework. Nucleic Acids Res 39:W437–W443
11. Hom G, Graham RR, Modrek B et al (2008) Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med 358:900–909
12. Harley JB, Alarcon-Riquelme ME, Criswell LA et al (2008) Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 40:204–210
13. Freudenberg J, Lee HS, Han BG et al (2011) Genome-wide association study of rheumatoid arthritis in Koreans: population-specific loci as well as overlap with European susceptibility loci. Arthritis Rheum 63:884–893
14. Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. Stat Med 21:1539–1558
15. Egger M, Smith GD, Phillips AN (1997) Meta-analysis: principles and procedures. BMJ 315:1533–1537
16. DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7:177–188
17. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI (2008) SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 24:2938–2939
18. Wang K, Li M, Bucan M (2007) Pathway-based approaches for analysis of genomewide association studies. Am J Hum Genet 81:1278–1283
19. Willer CJ, Speliotes EK, Loos RJ et al (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41:25–34
20. Ji JD, Lee WJ, Kong KA et al (2010) Association of STAT4 polymorphism with rheumatoid arthritis and systemic lupus erythematosus: a meta-analysis. Mol Biol Rep 37:141–147
21. Lee YH, Song GG (2009) Association between the rs2004640 functional polymorphism of interferon regulatory factor 5 and systemic lupus erythematosus: a meta-analysis. Rheumatol Int 29:1137–1142
22. Yang W, Ng P, Zhao M et al (2009) Population differences in SLE susceptibility genes: STAT4 and BLK, but not PXK, are associated with systemic lupus erythematosus in Hong Kong Chinese. Genes Immun 10:219–226
23. Elbers CC, van Eijk KR, Franke L et al (2009) Using genome-wide pathway analysis to unravel the etiology of complex diseases. Genet Epidemiol 33:419–431
24. Hong MG, Pawitan Y, Magnusson PK, Prince JA (2009) Strategies and issues in the detection of pathway enrichment in genome-wide association studies. Hum Genet 126:289–301
25. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 38:D355–D360
26. Ashburner M, Ball CA, Blake JA et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29
27. Wang K, Li M, Hakonarson H (2010) Analysing biological pathways in genome-wide association studies. Nat Rev Genet 11:843–854