Integrative Analyses for Functional Mechanisms Underlying Associations for Rheumatoid Arthritis

Fei-Yan Deng, Shu-Feng Lei, Hong Zhu, Yong-Hong Zhang, and Zeng-Li Zhang

ABSTRACT. ***Objective.*** Extensive association analyses including genome-wide association studies (GWAS) and powerful metaanalysis studies have identified a long list of loci associated with rheumatoid arthritis (RA) in very large populations, but most of them established statistical associations of genetic markers and RA only at the DNA level, without supporting evidence of functional relevance. Our study serves as a trial to detect the functional mechanisms underlying associations for RA by searching publicly available datasets and results.

Methods. Based on publicly available datasets and results, we performed integrative analyses (gene relationships across implicated loci analysis, differential gene expression analysis, and functional annotation clustering analysis) and combined them with the expression quantitative trait locus (eQTL) results to dissect functional mechanisms underlying the associations for RA.

Results. By searching 2 GWAS, Integrator and PheGenI, we selected 98 RA association results (p < 10^{-5}). Among these associations, we found that 8 single-nucleotide polymorphisms (SNP; rs1600249, rs2736340, rs3093023, rs3093024, rs4810485, rs615672, rs660895, and rs9272219) serve as cis-effect regulators of the corresponding eQTL genes (BLK and CD4 in non-HLA region; CCR6, HLA-DQA1, and HLA-DQB1 in HLA region) that also were differentially expressed in RA-related cell groups. These 5 genes are closely related with immune response in function.

Conclusion. Our results showed the functional mechanisms underlying the associations of 8 SNP and the corresponding genes. This study is an example of mining publicly available datasets and results in validation of significant disease-association results. Using public data resources for integrative analyses may provide insights into the molecular genetic mechanisms underlying human diseases. (First Release May 15 2013; J Rheumatol 2013;40:1063–8; doi:10.3899/jrheum.121119)

Key Indexing Terms:
RHEUMATOID ARTHRITIS   EXPRESSION QUANTITATIVE TRAIT LOCI   GENOME-WIDE ASSOCIATION STUDIES

Rheumatoid arthritis (RA) is a common systemic inflammatory arthritis characterized by chronic inflammation of multiple peripheral joints, which leads to destruction of cartilage and bones, progressive deformity, and severe disability. Identifying genetic factors for RA is challenging because of the complex nature of genetic determination, including polygenic determinations and gene-by-gene and gene-by-environment interactions. Recently, there has been great progress in characterization of genetic susceptibility to RA[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17]. Extensive genetic analyses including hypothesis-independent genome-wide association studies (GWAS)[1,3,5,10,11,16] and metaanalysis studies[8,12,15,18] identified a long list of loci associated with RA with relatively high statistical power. However, most of those studies established the statistical associations of genetic markers and RA only at the DNA level without exploring functional mechanisms underlying the associations. Such established associations usually did not provide immediate insights into functions of genes or regulation of gene expression that bridges gene code information and RA phenotype directly.

Currently, numerous public gene expression datasets (e.g., National Center for Biotechnology Information, US National Library of Medicine; website: www.ncbi.nlm.nih.gov/geo) are available to identify genes relevant to RA in the original studies. These datasets will serve as important resources for further data-mining to provide supplementary functional evidence to bridge the relationship of genetic markers and RA.

From the Center for Genetic Epidemiology and Genomics, and the Department of Epidemiology, School of Public Health, Soochow University, Suzhou, Jiangsu, China.

F-Y. Deng, PhD, Center for Genetic Epidemiology and Genomics, School of Public Health; S-F. Lei, PhD, Center for Genetic Epidemiology and Genomics, and the Department of Epidemiology, School of Public Health; H. Zhu, MS, Center for Genetic Epidemiology and Genomics, School of Public Health; Y-H. Zhang, PhD, Center for Genetic Epidemiology and Genomics, and the Department of Epidemiology, School of Public Health; Z-L. Zhang, PhD, Department of Epidemiology, School of Public Health, Soochow University.

Address correspondence to Dr. S-F. Lei, Center for Genetic Epidemiology and Genomics, School of Public Health, Soochow University, Suzhou, Jiangsu 215123, China. E-mail: leisf@suda.edu.cn

Based on the available datasets and results, our study aimed to detect functional mechanisms underlying the associations identified in previous studies by performing integrative analyses [gene relationships among implicated loci (GRAIL), differential gene expression analysis, and functional annotation clustering analysis] and by combining the findings of expression quantitative trait loci (eQTL).

MATERIALS AND METHODS

Selection of RA-associated single-nucleotide polymorphisms (SNP). The GWAS Integrator[19] and Phenotype-Genotype Integrator (PheGenI; www.ncbi.nlm.nih.gov/gap/PheGenI/) are 2 bioinformatics resources that provide robust lookup and analytic functionalities for published GWAS and metaanalysis studies. We used the phenotype "rheumatoid arthritis" to search the 2 databases and found 98 interesting association results with p < 10[-5] (data not shown, available from author upon request) between 85 unique SNP and RA[1,2,3,4,5,6,7,8,9,10,11,12,13,14].

In most situations, significant SNP means at least 1 susceptible gene or functional locus is located within the nearby region of the SNP. A common approach for identification of candidate genes is to select nearby genes located at the 2 sides of each associated SNP as candidate genes for further functional studies. To increase the chances of finding susceptible genes, we used the GRAIL[20] to prioritize and identify the susceptible candidate genes near RA-associated SNP. This approach automatically assesses the degree of relatedness of implicated genes using 250,000 PubMed abstracts, then prioritizes and identifies the candidate gene around associated SNP.

Analysis through eQTL. As is commonly recognized, variants at the DNA level may lead to variations in quantifiable intermediate phenotypes (such as at the RNA level), which subsequently result in variations in endpoint phenotypes, such as susceptibility to RA. Analysis with eQTL is an important approach in detection of functional mechanisms underlying association by testing whether identified variants at the DNA level may lead to variations in messenger RNA (mRNA) expression of nearby genes. Several recent large-scale studies have identified thousands of putative eQTL in multiple tissues or cells including liver, lymphoblastoid cell lines (LCL), monocytes, T cells, brain, and fibroblasts[21,22,23,24,25,26,27], which were summarized and archived in databases for quick and easy searching and comparison (eQTL: http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/; GTEx eQTL: www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi). To investigate the functional mechanisms underlying the above-selected associations, we first searched the 2 databases for eQTL to evaluate whether the above SNP influence transcript levels of genes as cis-effect regulators (eSNP) in multiple RA-related tissues or cells.

Differential expression analysis for RA-associated genes. We performed differential expression analyses for nearby genes physically located at the 2 sides of each associated SNP and the genes detected by GRAIL. First, we downloaded 5 publicly available expression datasets from GEO Datasets (www.ncbi.nlm.nih.gov/geo; GSE10500, GSE1402, GSE12021, GSE15573, and GSE23561). These studies were performed with the original purpose of identifying genes underlying RA in multiple tissues or cells, e.g., peripheral blood, peripheral blood mononuclear cells (PBMC), synovial membrane, synovial macrophages, and synovial fluid mononuclear cells (SFMC). The experiment procedures and data analyses including normalization of raw data were detailed in the original publications[28,29,30,31,32]. Using the normalized data available on the public databases, t tests were performed to identify differential expression genes by comparing means of the gene expression signals in RA cases and controls.

Functional annotation clustering analysis. Functional annotation clustering analyses were performed to explore functional similarity of the associated genes to cluster them in particular biological gene ontology (GO) terms and functional pathways as defined by the Gene Ontology project and the KEGG gene database, respectively. The significantly associated genes were functionally annotated using the Database for Annotation, Visualization and Integrated Discovery (DAVID) integrated database query tools (http://david.abcc.ncifcrf.gov/)[33,34]. Fisher's exact test was performed to determine whether the GO term or pathway annotates a specified list of genes at a frequency greater than would be expected by chance. Usually, p value is ≤ 0.05 to be considered strongly enriched in the annotation categories. Bonferroni correction was also adopted for multiple testing.

RESULTS

By searching the GWAS Integrator and PheGenI, we selected 98 RA association results (p < 10[-5]) between 85 unique SNP and RA (data not shown, available from author upon request). These association results come from 14 published RA association studies[1,2,3,4,5,6,7,8,9,10,11,12,13,14]. Some of these studies have relatively large sample sizes, e.g., the study performed by Stahl, *et al*[12] has 5539 cases and 20,169 controls and 6768 cases and 8806 controls in initial and replication analyses, respectively. The associations were distributed across the genome except for chromosomes 13 and 19, but were relatively tightly clustered into the HLA region (chromosome 6p21.3; 24 associations; data not shown, available from author upon request).

About 39 "susceptible" genes detected by GRAIL analysis are not overlapped with the genes identified according to the physical positions (data not shown, available from author upon request). A total of 123 unique candidate genes were identified using the above 2 methods.

Among the 85 unique SNP, we found that about 21 SNP have potential eQTL effects (Table 1, and data not shown, available from the author upon request) in monocytes, T cells, and LCL, which are 3 types of cells closely correlated with immune response. Among them, 18 associations were detected (Table 1) between 11 cis-effect SNP and the expression of 9 genes from the 123 "identified" genes. Specifically, SNP (rs2736340 and rs9272219) showed strong associations with expression levels of 2 nearby genes of both sides (FAM167A and BLK for rs2736340; HLA-DRB1 and HLA-DQA1 for rs9272219), as evidenced in both monocytes and LCL. Five SNP are located at an HLA region and 6 at a non-HLA region.

The 123 "identified" genes might be involved in the pathology of RA. In the peripheral blood cells, PBMC, synovial membrane, synovial macrophages, and SFMC (Table 2, and data not shown, available from author upon request) in multiple RA-related cells, t tests showed that a total of 96 genes among the 123 "identified" genes have differential expression signals (p < 0.05), in at least 1 functional study. Among the above-identified 9 eQTL genes, 5 genes (BLK and CD4 in the non-HLA region; CCR6, HLA-DQA1, and HLA-DQB1 in the HLA region) were differentially expressed in samples S1 (synovial macrophages), S2 (PBMC/SFMC), or S5 (peripheral blood cells; Table 2).

To test the probability of our "identified" genes clustering into a particular biological pathway and GO term, we performed a functional annotation clustering analysis

Table 1. Expression quantitative trait locus (eQTL) analysis results between 11 single-nucleotide polymorphisms (SNP) and the "identified" 9 genes.

SNP ID	Chr	Loc.	Allele	Role	Assn p	Gene 1	Gene 2	Implicated Gene	PMID for Assn	eQTL Gene	Effect	Score	Tgt.	PMID for eQTL
rs1600249	8	11359638	A/C	Intron	5.00E-06	BLK	BLK	BLK	21452313	BLK	eQTL	31.6	LCL	Database#
rs2736340	8	11343973	C/T	Interg.	6.00E-09	FAM167A	BLK	C8orf13	19503088	BLK	eQTL	14.5	Mono	20502693
										FAM167A	trans-QTL	6.5	LCL	20220756
										FAM167A	exon-QTL	6.4	LCL	20220756
										BLK	eQTL	59.5	LCL	Database
rs3093023	6	167534290	A/G	Intron	2.00E-11	CCR6	CCR6	CCR6	20453842	CCR6	eQTL	48.3	Mono	20502693
rs3093024	6	167532793	A/G	Intron	8.00E-19	CCR6	CCR6	CCR6	20453841	CCR6	eQTL	50.9	Mono	20502693
rs4810485	20	44747947	G/T	Intron	3.00E-09	CD40	CD40	CD40	20453842	CD40	eQTL	34.4	Mono	20502693
rs615672	6	32574171	C/G	Interg.	8.00E-27	HLA-DRB1	HLQ-DQA1	BTNL2	17554300	HLA-DQA1	eQTL	102.9	Mono	20502693
rs660895	6	32577380	A/G	Interg.	1.00E-108	HLA-DRB1	HLA-DQA1	BTNL2	17804836	HLA-DQA1	eQTL	10.2	LCL	18846210
rs9272219	6	32602269	G/T	Interg.	1.00E-45	HLA-DRB1	HLA-DQA1	BTNL2	21653640	HLA-DQA1	eQTL	191.2	Mono	20502693
										HLA-DQB1	eQTL	83.2	Mono	20502693
										HLA-DQA1	eQTL	6.9	LCL	18846210
										HLA-DQA1	sQTL	13.3	LCL	20220758
										HLA-DRB1	eQTL	72.6	Mono	20502693
rs3816587	4	25417244	C/T	Intron	9.00E-06	ANAPC4	ANAPC4	P14K2B	17554300	ANAPC4	eQTL	25.3	Mono	20502693
rs3890745	1	2553624	A/G	Intron	1.00E-07	MMEL1	MMEL1	TNFRSF14	18794853	MELL1	eQTL	52.8	Mono	20502693
rs9268853	6	32429643	C/T	Interg.	5.00E-109	HLA-DRB9	HLA-DRB5	BTNL2	21653640	HLA-DRB5	eQTL	8.0	LCL	17873874

All eQTL have cis-effects. Database#: http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl. "Gene 1" and "Gene 2" are the nearby genes physically located at 2 sides of SNP. "Score":−log10 (P); LCL: lymphoblastoid cell lines; trans-QTL: transcript expression levels against SNP; exon-QTL: quantified reads for known exons by RNA sequencing against SNPs; sQTL: splicing QTL; PMID for Association: Pubmed ID for association results; PMID for eQTL: PubMed ID for eQTL results; Loc.: location; Assn: association; Tgt.: target; interg.: intergenic; trans: transcript; mono: monocytes.

using the functional annotation tool of the DAVID database. We found significant GO terms and KEGG pathways even after Bonferroni correction. For the 123 "identified" genes, genes tend to enrich in immune-related pathways (such as "autoimmune thyroid disease," "intestinal immune network for IgA production," and "immune response antigen processing and presentation"; data not shown, available from author upon request). Most importantly, we found a significant clustering (Bonferroni correction p = 4.25E-09) of genes involved in immune response — genes very closely relevant to RA. Five eQTL genes (HLA-DQB1, HLA-DRB1, HLA-DRB5, CD40, and HLA-DQA1) were significantly clustered in multiple immune-related KEGG pathways or GO terms (data not shown, available from author upon request).

This functional evidence, taken together, highlighted the significance of 8 RA-associated SNP (rs1600249, rs2736340, rs3093023, rs3093024, rs4810485, rs615672, rs660895, and rs9272219) listed in Table 1, because these SNP serve as cis-effect regulators of RA-associated genes that were differentially expressed in RA-related cell groups.

DISCUSSION

Using publicly available datasets, we performed integrative analyses and combined them with eQTL results to detect functional mechanisms underlying the associations for RA. We discovered 11 SNP acting as cis-effect regulators on 9 "identified" genes, of which 5 (BLK, CCR6, CD40, HLA-DQA1, and HLA-DQB1) were differentially expressed. The 123 "identified" genes tend to enrich in KEGG pathways (intestinal immune network for IgA production, and immune response antigen processing and presentation) and are closely relevant to RA.

In biologic systems, genetic information carried by DNA is passed on to RNA molecules by transcription and then to protein molecules through translation. Sequence variants at the DNA level represent a class of heritable molecules and contribute to variability of complex traits in a population. Functional mechanisms underlying associations between variants at the DNA level and RA may be that DNA sequence variants lead to variation in quantifiable intermediate phenotypes (such as at the RNA level), which subsequently lead to variation of susceptibility to RA.

Table 2. Differential expression analyses for the detected eQTL genes in multiple RA-related cells.

		Sample			
	S1	S2	S3	S4	S5
Disease	RA	JRA	RA	RA	RA
Target cells	Synovial macrophages	PBMC/SFMC	Synovial membrane	PBMC	PBC
Sample size	5:3	37:20#	12:9	18:15	6:9
Platform	Affymetrix Human Genome U95 Version 2 Array	Affymetrix Human Genome U95 Version 2 Array	Affymetrix Human Genome U133A Array	Illumina Human-6 v2.0 expression Beadchip	Human 50K Exonic Evidence-Based Oligonucleotide Array
PMID Ref.	18345002	15150433	18721452	19710928	21368773
GSE No.	GSE10500	GSE1402	GSE12021	GSE15573	GSE23561

Gene Symbol	Probe ID	T test p	RA/ cont.	Probe ID	T test	PBMC/ SFMC	Probe ID	T test p	RA/ cont.	Probe ID	T test p	RA/ cont.	Probe ID	T test p	RA/ cont.
BLK	854_at	2.98E-01	0.27	854_at	3.32E-07	2.89	210934_at	6.59E-02	1.7	ILMN_1668277	9.65E-01	1	11488	4.10E-02	0.61
CCR6	458_at	1.66E-01	0.09	458_at	2.42E-01	0.78	206983_at	7.72E-02	2.48	ILMN_1690907	4.24E-01	0.93	14199	5.62E-04	0.41
CD40	35149_at	4.37E-02	0.19	35150_at	3.14E-06	0.65	205153_s_at	1.66E-02	1.47	ILMN_1779257	1.15E-01	0.93	14306	7.97E-03	0.46
HLA-DQA1	32773_at	3.51E-01	0.16	32773_at	5.39E-05	0.23	236203_at	3.44E-01	5.99	ILMN_1808405	3.63E-01	0.88	22797	4.33E-04	0.31
HLA-DQB1	36108_at	6.01E-02	0.22	36878_f_at	3.35E-17	0.29	212998_x_at	3.04E-03	2.09	ILMN_1661266	5.27E-01	0.88	22801	8.50E-03	0.34

#All RA cases. If 1 gene has multiple detected probes, only the most significant expression results of probes are listed. PBMC: peripheral blood mononuclear cell; PBC: peripheral blood cells; SFMC: synovial fluid mononuclear cell; PMID: PubMed ID; GSE no.: Gene Expression Omnibus Number, http://www.ncbi.nlm.nih.gov/geo/; RA: rheumatoid arthritis; JRA: juvenile rheumatoid arthritis; eQTL: expression quantitative trait loci; cont: control.

Therefore, integrating substantial evidence from multiple levels (i.e., DNA, RNA) could ascertain the potential functioning mechanism of genes and their contribution to variation in susceptibility to RA.

The functional evidence from our study suggests potential regulatory mechanisms underlying the associations between RA and 8 SNP (rs1600249, rs2736340, rs3093023, rs3093024, rs4810485, rs615672, rs660895, and rs9272219) in a general way. Briefly, different genotypes of the 8 SNP in the corresponding genes (BLK, CCR6, CD40, HLA-DQA1, and HLA-DQB1), by regulating differential mRNA transcriptions of the 5 genes and thus differential protein expression or enzyme activity, consequently have an effect on variation of susceptibility to RA in a population. However, we could not establish the functional link between variants and RA in an exact way. We still do not know which allele has the risk effect, and whether the risk allele of a variant leads to downregulation or upregulation of mRNA transcription, because the databases archive only the p value, not the direction and effect of the associations (i.e., risk allele and OR). Despite this, the supportive functional evidence may still be useful for future in-depth functional validation of the involvement of variants in RA.

It is noteworthy that failure to find functional evidence underlying the selected associations in publicly available data and results does not necessarily exclude the importance of the associations in RA, given that (1) the absence of association between RA-associated SNP and mRNA expression may result from limited statistical power due to the small sample size or small effect of the SNP; (2) the associated genes may exert effects on RA through other tissues or cells not studied here; and (3) the sequence variant in a gene may lead to RA by other mechanisms (e.g., epigenetic regulation) instead of regulating gene expression. In-depth functional studies are required to provide evidence for the mechanisms underlying the detected associations. Other bioinformatics resources are available for in-depth functional studies, such as the ENCODE database, which holds additional information about noncoding human enhancer, transcription factor binding, etc., and the FASTSNP program, used to analyze and predict potential functional effects of causal variants[35].

The classic MHC loci have long been regarded as the major contributors to the pathogenesis of RA. However, recent GWAS have identified a series of candidate genes outside the MHC region, suggesting that non-MHC RA susceptibility loci may also be important in RA pathogenesis and thus may attract researchers' interest. BLK and CD40 are 2 strongly and consistently RA-associated loci outside the MHC region. B cell lymphocyte kinase (BLK) gene is a member of the Src kinase family, which may influence the proliferation and differentiation of cells. BLK plays a role in B cell receptor signaling and B cell development. Studies have identified several RA-associated variants at the DNA level in both white and Asian populations (e.g., rs922483, rs2248932, rs2736340, and rs1600249)[3,4,36]. A study

performed to detect functional mechanisms underlying the associations identified that the rs922483 may serve as a cis-regulatory locus in regulating mRNA and protein expression of BLK prominently observed in B cells early in the development of BLK[37]. Our study provided functional evidence for 2 SNP (rs2736340 and rs1600249) that contribute to variation of susceptibility to RA in the population. CD40 is a member of the TNF-receptor super-family. This gene plays an essential role in mediating a broad variety of immune and inflammatory responses. The SNP rs4810485 is a well established variant that confers risk of RA[11,38,39] and influences the rate of joint destruction in RA[40]. However, to our knowledge, no followup functional studies have investigated the underlying mechanisms for the association. Our study attempted to provide suggestive functional evidence to support the association.

Although we gained some functional evidence for the associations between RA and 8 SNP, based on the current available evidence, it is still challenging to discriminate causal variants, which have potential functional effects, from nearby marker SNP, especially for SNP at the HLA region, where there is a well known strong linkage disequilibrium (LD). About 10 SNP extending a relatively long distance (460 kb) were consistently associated with RA (data not shown, available from author upon request). The HLA region may contain only a few true functional variants, but because of the strong LD between the markers and functional variants, about 10 significant SNP have been detected. A study was performed to identify functional and potential causal variants within the HLA region[41].

The functional evidence found in our study supports the associations between RA and 8 SNP, because these SNP serve as cis-effect regulators, and the corresponding eQTL genes were also differentially expressed. Use of public data resources for integrative analyses may provide insights into the molecular genetic mechanisms underlying human diseases.

REFERENCES

1. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661-78.
2. Eleftherohorinou H, Hoggart CJ, Wright VJ, Levin M, Coin LJ. Pathway-driven gene stability selection of two rheumatoid arthritis GWAS identifies and validates new susceptibility genes in receptor mediated signalling pathways. Hum Mol Genet 2011;20:3494-506.
3. Freudenberg J, Lee HS, Han BG, Shin HD, Kang YM, Sung YK, et al. Genome-wide association study of rheumatoid arthritis in Koreans: Population-specific loci as well as overlap with European susceptibility loci. Arthritis Rheum 2011;63:884-93.
4. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, et al. REL, encoding a member of the NF-kappa B family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat Genet 2009;41:820-3.
5. Julia A, Ballina J, Canete JD, Balsa A, Tornero-Molina J, Naranjo A, et al. Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. Arthritis Rheum 2008;58:2275-86.
6. Kochi Y, Okada Y, Suzuki A, Ikari K, Terao C, Takahashi A, et al. A regulatory variant in CCR6 is associated with rheumatoid arthritis susceptibility. Nat Genet 2010;42:515-9.
7. Liu C, Batliwalla F, Li W, Lee A, Roubenoff R, Beckman E, et al. Genome-wide association scan identifies candidate polymorphisms associated with differential response to anti-TNF treatment in rheumatoid arthritis. Mol Med 2008;14:575-81.
8. Okada Y, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, et al. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. Nat Genet 2012;44:511-6.
9. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477-82.
10. Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis — A genomewide study. N Engl J Med 2007;357:1199-209.
11. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008;40:1216-23.
12. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 2010;42:508-14.
13. Terao C, Yamada R, Ohmura K, Takahashi M, Kawaguchi T, Kochi Y, et al. The human AIRE gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. Hum Mol Genet 2011;20:2680-5.
14. Hu HJ, Jin EH, Yim SH, Yang SY, Jung SH, Shin SH, et al. Common variants at the promoter region of the APOM confer a risk of rheumatoid arthritis. Exp Mol Med 2011;43:613-21.
15. Zhernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. PLoS Genet 2011;7:e1002004.
16. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nat Genet 2009;41:1313-8.
17. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 2012;44:1336-40.
18. Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF, et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006;119:634-41.
19. Yu W, Yesupriya A, Wulf A, Hindorff LA, Dowling N, Khoury MJ, et al. GWAS Integrator: A bioinformatics tool to explore human genetic associations reported in published genome-wide association studies. Eur J Hum Genet 2011;19:1095-9.
20. Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P, et al. Identifying relationships among genomic disease regions: Predicting genes at pathogenic SNP associations and rare deletions. PLoS Genet 2009;5:e1000534.
21. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. PLoS Biol 2008;6:e107.
22. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond — The transcriptome of human monocytes and disease susceptibility. PloS One 2010;5:e10693.
23. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet 2008;4:e1000214.
24. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. Nat Genet 2007;39:1217-24.
25. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C,

Attar-Cohen H, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 2009;325:1246-50.

26. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. Nature 2010;464:773-7.

27. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. Nature 2010;464:768-72.

28. Grayson BL, Wang L, Aune TM. Peripheral blood gene expression profiles in metabolic syndrome, coronary artery disease and type 2 diabetes. Genes Immun 2011;12:341-51.

29. Teixeira VH, Olaso R, Martin-Magniette ML, Lasbleiz S, Jacq L, Oliveira CR, et al. Transcriptome analysis describing new immunity and defense genes in peripheral blood mononuclear cells of rheumatoid arthritis patients. PloS One 2009;4:e6803.

30. Huber R, Hummert C, Gausmann U, Pohlers D, Koczan D, Guthke R, et al. Identification of intra-group, inter-individual, and gene-specific variances in mRNA expression profiles in the rheumatoid arthritis synovial membrane. Arthritis Res Ther 2008;10:R98.

31. Barnes MG, Aronow BJ, Luyrink LK, Moroldo MB, Pavlidis P, Passo MH, et al. Gene expression in juvenile arthritis and spondyloarthropathy: Pro-angiogenic ELR+ chemokine genes relate to course of arthritis. Rheumatology 2004;43:973-9.

32. Yarilina A, Park-Min KH, Antoniv T, Hu X, Ivashkiv LB. TNF activates an IRF1-dependent autocrine loop leading to sustained expression of chemokines and STAT1-dependent type I interferon-response genes. Nat Immunol 2008;9:378-87.

33. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.

34. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1-13.

35. Yuan HY, Chiou JJ, Tseng WH, Liu CH, Liu CK, Lin YJ, et al. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. Nucleic Acids Res 2006;34:W635-41.

36. Zhang H, Wang L, Huang Y, Zhuang C, Zhao G, Liu R, et al. Influence of BLK polymorphisms on the risk of rheumatoid arthritis. Mol Biol Rep 2012;39:9965-70.

37. Simpfendorfer KR, Olsson LM, Manjarrez Orduno N, Khalili H, Simeone AM, Katz MS, et al. The autoimmunity-associated BLK haplotype exhibits cis-regulatory effects on mRNA and protein expression that are prominently observed in B cells early in development. Hum Mol Genet 2012;21:3918-25.

38. Plant D, Flynn E, Mbarek H, Dieude P, Cornelis F, Arlestig L, et al. Investigation of potential non-HLA rheumatoid arthritis susceptibility loci in a European cohort increases the evidence for nine markers. Ann Rheum Dis 2010;69:1548-53.

39. Garcia-Bermudez M, Gonzalez-Juanatey C, Lopez-Mejias R, Teruel M, Corrales A, Miranda-Filloy JA, et al. Study of association of CD40-CD154 gene polymorphisms with disease susceptibility and cardiovascular risk in Spanish rheumatoid arthritis patients. PloS One 2012;7:e49214.

40. van der Linden MP, Feitsma AL, le Cessie S, Kern M, Olsson LM, Raychaudhuri S, et al. Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. Arthritis Rheum 2009;60:2242-7.

41. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 2012;44:291-6.