HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2010 January 01.

Published in final edited form as:

Nat Genet. 2009 July ; 41(7): 820–823. doi:10.1038/ng.395.

REL, a member of the NF-κB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis

Peter K. Gregersen[1,†], **Chistopher I Amos**[2,†], **Annette T. Lee**[1], **Emily Lu**[2], **Elaine F. Remmers**[3], **Daniel L. Kastner**[3], **Michael F. Seldin**[4], **Lindsey A Criswell**[5], **Robert M. Plenge**[6], **V. Michael Holers**[7], **Ted Mikuls**[8], **Tuulikki Sokka**[9], **Larry W. Moreland**[10], **S. Louis Bridges**[11], **Gang Xie**[12], **Ann B. Begovich**[13], and **Katherine A. Siminovitch**[12]

[1)]The Feinstein Institute for Medical Research, North Shore-Long Island Jewish Health System, Manhasset, New York 11030, USA [2)]University of Texas M.D. Anderson Cancer Center, Houston, USA [3)]Genetics and Genomics Branch, NIAMS, NIH, Bethesda, Maryland 20892, USA [4)]Rowe Program in Genetics, University of California at Davis, Davis, California 95616, USA [5)]Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, University of California, San Francisco, USA [6)]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115 USA [7)]University of Colorado Denver School of Medicine [8)]University of Nebraska Medical Center [9)]Jyväskylä Central Hospital, Jyväskylä, Finland [10)]University of Pittsburgh Medical Center, Pittsburgh, PA 15261 [11)]University of Alabama at Birmingham, Birmingham, AL [12)]Mount Sinai Hospital and University Health Network, Toronto, Ontario M5G 1X5 Canada [13)]Celera, Alameda, California, USA

A genome-wide association study of rheumatoid arthritis in 2418 cases and 4504 controls from North America identified an association at the *REL* locus, encoding c-Rel, on chromosome 2p13 (rs13031237, $p=6.01 \times 10^{-10}$). Replication in independent case-control datasets comprising 2604 cases and 2882 controls confirmed this association, yielding an

Correspondence to Peter K. Gregersen MD Robert S. Boas Center for Genomics and Human Genetics The Feinstein Institute for Medical Research North Shore LIJ Health System 350 Community Drive Manhasset, New York 11030 Email: peterg@nshs.edu Telephone: 516-562-1542 FAX: 516-562-1153. Katherine Siminovitch PhD Samuel Lunenfeld Research Institute Mount Sinai Hospital, Room 656A 600 University Avenue Toronto, Ontario, M5G 1X5 CANADA e-mail: ksimin@mshri.on.ca Tel: (416) 586-8723 Fax: (416) 586-8731.

[†]Authors contributed equally

Author Contributions PKG organized and designed the study, supervised genotyping of U.S samples, contributed to statistical analysis and is the primary author of the manuscript.

CIA coordinated the design of the study, carried out and supervised all of the statistical analyses, and contributed to the writing of the manuscript.

ATL performed GWAS on all U.S. samples and organized samples for distribution to carry out replication studies.

EL carried out statistical analysis

EFM contributed to study design, carried out replication genotyping and participated in preparation of the manuscript.

DLK contributed to study design and participated in review of the manuscript

MFS contributed to study design and data analysis, and participated in manuscript preparation

LAC contributed to study design, contributed patients for study, and participated in manuscript preparation

RMP contributed to study design and participated in manuscript preparation

VMH, TM, TS, SLB, LWM contribute patient populations for study and contributed to manuscript preparation.

GX carried out genotyping and sample preparation of Canadian samples

ABB contributed samples for study, carried out replication genotyping and participated in manuscript preparation.

KAS participated in study design and organization, supervised all genotyping of Canadian samples and participated in preparation of the manuscript.

allelic OR=1.25 (95%CI 1.177-1.318, $p = 3.08 \times 10^{-14}$) for marker rs13031237 and an allelic OR= 1.21 (95%CI 1.150-1.282, $p = 2.60 \times 10^{-11}$) for marker rs13017599 in the combined dataset. The combined dataset also provides definitive support for associations at both *CTLA4* (rs231735; OR=0.85, 95% CI 0.81-0.90; $p= 6.25 \times 10^{-9}$) and *BLK* (rs2736340; OR= 1.19, 95% CI 1.125-1.268; $p=5.69 \times 10^{-9}$). C-Rel is an NF-kappaB family member with distinct functional properties in hematopoietic cells, and its association with RA suggests disease pathways that involve other recently identified RA susceptibility genes including CD40, TRAF1, TNFAIP3, and PRKCQ1,2.

Rheumatoid arthritis [RA (MIM 180300)] is a common autoimmune disorder affecting approximately 1% of populations of European origin and whose predominant manifestation is inflammation with bone and cartilage destruction in diarthrodial joints. The genetic basis for RA is complex, with at least six genes generally accepted as associated with disease in populations of European origin, including *HLA-DRB1*, *PTPN22*, *STAT4*, *TRAF1*, and *TNFAIP3*1. A number of additional loci have recently been reported as a result of expanded genome-wide association studies3 and metanalyses2. Many of these are likely to reflect true associations, although a convincing demonstration often requires very large samples sizes, given that many of the associations at these loci are quite modest. In most cases the causative allele(s) have not been identified, and therefore the actual contribution to disease risk at these loci is unknown.

The diagnosis of rheumatoid arthritis is based on clinical criteria established over two decades ago4. However, these criteria do not yet include antibody reactivity to cyclic citrullinated peptides (CCP), the presence of which is a highly sensitive and specific marker for the diagnosis of RA; between 50 and 80% of patients meeting standard criteria for RA exhibit anti-CCP antibodies5. Remarkably, the classical HLA-DRB1 associations with RA are entirely restricted to this phenotypic subgroup6, as are many of the other reported genetic associations. In addition to this phenotypic heterogeneity, there is evidence for genetic heterogeneity in risk for RA among different racial groups7, and this has complicated efforts at replication. Given these considerations, it is apparent that additional risk genes for RA remain to be discovered. For these reasons we undertook an expansion of our previous genome-wide association study of rheumatoid arthritis8, restricted to North American cases of European origin that were overwhelmingly (~95%) anti-CCP antibody positive. We also assembled a large case-control population for replication studies.

All new genotyping of case samples for this study was performed on Illumina HapMap370 BeadArray typing platforms, and after quality control filtering (see Supplementary Methods), a combined dataset of 2418 RA cases and 4504 controls was available for WGA analysis that had been genotyped on 278502 SNPs that passed all quality control filters applied to each set of data. The cases are derived in part from affected sibling pair families of the North American Rheumatoid Arthritis Consortium (NARAC) previously reported8 (one case per family), as well as new collections from both the U.S. and Canada (see Supplementary Table 1A). The genome-wide lambda value inflation of chi-square values was calculated to be 1.21, allowing for the large sample size. Structured association was therefore applied to correct for population stratification by matching homogeneous clusters

of cases and controls, as implemented in PLINK. The genome-wide lambda value after adjustment was calculated to be 1.06.

Figure 1 displays a graphical summary of the results of the genome-wide analysis as implemented in PLINK, after conditioning on clusters. As noted in previous studies, the largest association signal is in an extended region within the MHC achieving a maximal level significance of $p=9.50 \times 10^{-104}$ in this study. (The y axis is truncated at - log p=27 in Figure 1). In addition, the previously reported associations8 are confirmed at PTPN22 (rs2476601, $p= 1.62 \times 10^{-21}$) and the TRAF1/C5 regions (maximal association at rs881375, $p= 4.09\times10^{-8}$). However, after *PTPN22* the second most significant association is observed on chromosome 2 in the region of *REL*, where two markers provide highly significant evidence of association (rs13031237, allele specific $p=6.01 \times 10^{-10}$ and rs13017599, allele specific $p=9.05 \times 10^{-9}$), as shown in Table 1. This is a novel finding, since *REL* has not been brought forward as a candidate region for RA risk by any of the previous association studies in RA, although a recent publication has reported an association with Coeliac disease9. At lower levels of significance, we also noted evidence of association (Table 1) with *CTLA4* (rs 6748358, $p= 8.24 \times 10^{-5}$) and a SNP marker in the region of the *BLK* locus (rs2736340, $p= 6.06 \times 10^{-7}$). A complete list of results at significance level $p<0.01$ for the entire dataset is provided (Supplementary Table 2).

In order to establish that these results are robust to various analytic approaches, we also performed analysis using Eigenstrat. Before performing the PCA, we removed markers in the region of chromosome 8p that shows inversions in Northern Europeans (8.135-11.936 Mb) and markers in the centromeric region of chromosome 17q21.31 (40-43 Mb) that is polymorphic in European origin populations10. We also removed markers around the HLA region that are related to European ancestry and also rheumatoid arthritis risk (from 24-36 Mb). These HLA-related markers were also removed for analyses that estimate inflation of the genome-wide inflation of test statistics that can arise from differences in population ancestry, since the HLA region includes many hundreds of markers that are associated with RA risk6. The association tests with all markers (including those on chromosomes 6p, 8p, and 17p) were performed adjusting for the eigenvectors derived using Eigenstrat to remove population admixture effects. The lambda value was 1.20 prior to PCA correction and this value reduced to a lambda value of 1.06 after PCA correction. Results from Eigenstrat analysis along with trend tests from PLINK are presented in supplementary table 3. Despite some genome-wide excess in the expected number of positive results from tests for association, the specific findings for *REL* did not appear to be influenced substantially by population structure. As shown in supplementary table 3, we found that without and with adjustment for population structure the p-values and odds ratios associating the SNPs we queried were quite similar. For example, for rs13031237, the odds ratio and p-value without adjustment for population structure in the combined genome-wide association analysis was 1.278 (p=5.2 × 10-11) versus an odds ratio of 1.268 with adjustment (p=6.0 × 10-10).

In order to confirm the associations with *REL*, we carried out a replication study on independent sets of 2604 cases and 2882 controls from the U.S. and Canada (Supplementary Table 1B). While complete serologic data were not available for all subjects, the majority of cases were seropositive (either rheumatoid factor or CCP+) in the replication datasets. In

addition to selected SNPs from the *REL* locus, we also included candidate SNPs from the *CTLA4* and *BLK* regions. For technical reasons, slightly different SNP panels were utilized for the replication studies in the Canadian and U.S. samples, based on tagging LD provided in the HapMap, and the results from the separate datasets are provided in Supplementary Table 2. The combined results with common SNPs markers across all the datasets are shown in Table 1 for *REL*, *CTLA4* and *BLK*. Two highly associated SNPs at the *REL* locus are observed in the combined data using a Cochrane-Mantel-Haenszel analysis to allow for stratification among populations (rs13031237, OR=1.24, $p = 3.08 \times 10^{-14}$ and rs13017599, OR=1.21, $p= 2.06 \times 10^{-12}$). A graphical representation of the association results across the *REL* locus is shown in Figure 2. In addition, the data in Table 1 (and supplemental Table 2) provide definitive evidence for the previously suggested association with *CTLA4* (rs 231735, OR=0.86, $p=6.25 \times 10^{-9}$). The data also support *BLK* as a new RA risk locus (rs2736340, OR 1.19, $p=5.69 \times 10^{-9}$), a finding of some interest given the recent association of this locus with systemic lupus11,12. Supplementary table 4 presents genotype-specific results, which show codominance for all the loci except *CTLA4*, which is nearly dominant.

The nuclear factor-κB (NF-κB)/REL family of transcription factors contain five members including c-Rel, p65/Rel-A, Rel-B, p50/NFkB-1 and p52/NFkB-2. These factors have a central role in coordinating the expression of a wide variety of genes that control immune responses and autoimmunity13. Therefore, the identification of *REL*, encoding c-Rel, as a new risk locus for RA has provoked us to consider how this observation may fit in with pathways suggested by the complex emerging landscape of genetic susceptibility for RA. While the various NF-κB subunits have complex overlapping functions, current data suggest some distinct roles for c-Rel. The production of IL12 and IL23 subunits by macrophages and dendritic cells are critically dependent on c-Rel14,15. Thus, c-Rel knockout animals exhibit deficiencies in Th1-type immune responses, although intrinsic T cell defects may also contribute to this phenotype16. Intriguingly, both c-Rel and another recently identified RA risk gene, *PRKCQ*, are specifically involved in the survival of activated CD8 cells, at least in part through the regulation of IL2 production by these cells. In addition, a variety of genes in T cells are regulated by c-Rel, including *CD40* and *TNFAIP3*, both of which are now accepted RA susceptibility loci17.

In addition to effects on T cell and antigen presenting cell function, c-Rel has been shown to have a role in B cell proliferation and survival that particularly involves CD40 signaling pathways. Specifically, c-rel deficient B cells are susceptible to BCR induced apoptosis that cannot be prevented by activation through CD4018. This is due to reduced expression of the anti-apoptotic protein *Bcl-X_L*, a gene that is known to be regulated by c-Rel. Interestingly, rescue from Fas induced apoptosis is normal in these cells, demonstrating the existence of distinct CD40 signaling pathways that are at least in part distinguished by the involvement of c-rel. A similar c-rel associated difference in CD40 signaling has been seen in patients with ectodermal dysplasia and hyperIgM syndrome due to mutations in the NF-κB essential modulator, NEMO, where c-Rel dependent IL4 responses are also impaired19. C-Rel is the only NF-kappaB family member with oncogenic activity and the gene is amplified in some B cell lymphomas. Intriguingly, in both tumors and normal B cells it has recently been reported that the CD40 and c-Rel proteins can physically interact and form a heterodimer

that is translocated to the nucleus20 and has transcriptional regulatory activity for known c-rel target genes, including *CD154*, *BLyS/BAFF*, and *Bfl-1/A1*.

The associations of *CD40*2, *REL*, *TRAF1*8 and *TNFAIP3* 21are consistent with an important role for CD40 signaling pathways in RA pathogenesis. Indeed, CD40-CD40 ligand interactions have previously been identified as a potential target for therapy in autoimmune disease22, and clinical trials in lupus have shown promise23. Unfortunately, the clinical development of monoclonal antibody inhibitors of CD40L was cut short by adverse effects on the platelet function associated with the development of thromboembolic complications. Nevertheless, this pathway remains viable as a therapeutic target, and the current results mandate a thorough analysis of all the genes in this pathway to search for additional susceptibility alleles

METHODS

Subjects

The cases and controls utilized in this analysis were taken from a variety of collections in both the U.S. and Canada, as detailed below. The genome-wide association data from 868 RA cases and 1194 controls from North America have been previously reported in a study of susceptibility loci for rheumatoid arthritis8; all the additional genome-wide association data (1550 cases and 3310 controls) and all the replication sample data (2604 cases and 2882 controls) have not been previously reported in a RA association study. Informed consent was obtained from all subjects and these studies were approved by the Institutional Review Board of the North Shore LIJ Health System.

Subjects - U.S Cohorts

For our WGA analysis, we included WGA data on 868 North American RA cases samples reported previously8. We refer to these samples as “NARAC I”. All cases included in the current analysis are anti-CCP positive.

Details for patient collections used to compile the “NARAC II” dataset (N=952) used for final analysis on the Illumina HapMap370 chip are as follows.

1. RA probands in NARAC collection of affected sibling pair families in whom siblings did not meet criteria for RA (N=17), as well as probands of NARAC trio families (N=158). Collection criteria are previously published 8. No affected members of these families were included in our previously published GWAS 8. Only CCP+ patients were studied in the current analysis.
2. The Veterans Affairs Rheumatoid Arthritis Registry (VARA). VARA subjects, all meeting classification criteria for RA are currently being enrolled from six VA centers (Dallas, Denver, Jackson, Omaha, Salt Lake City, and Washington, DC). VARA is a longitudinal observational study that was started by Dr. Ted Mikuls at the University of Omaha. As of May 1, 2008, over 1,100 subjects have been enrolled. Of the 332 patients included in this analysis, 301 were documented to be CCP+ (N=332).

3. The Studies of Etiologies of Rheumatoid Arthritis (SERA) cohort. This NIH-funded (R01 AR051394) cohort study is co-directed by Drs. Michael Holers and Jill Norris at the University of Colorado Denver. SERA is a multi-center prospective cohort study designed to investigate genetic and epidemiologic associations with RA-related autoimmunity during the pre-clinical period of RA development. For the SERA study, probands with RA are recruited by Rheumatologists from clinics at the University of Colorado Denver School of Medicine (UCDSOM), Cedars-Sinai Medical Center, Los Angeles, California, the Rheumatoid Arthritis Investigation Network (RAIN) centered in Omaha, Nebraska, North Shore-Long Island Jewish Health System/Feinstein Institute for Medical Research, and the Benaroya Research Institute at Virginia Mason, Seattle, Washington. Additionally, probands are recruited through community outreach efforts and advertisements posted in the Arthritis Foundation Newsletter. Once a proband with RA has expressed interest in the study and consents for release of medical information, their medical charts are reviewed by a SERA Rheumatologist to ensure a diagnosis of RA. Probands are considered to have RA if they meet 4 of 7 of the ACR 1987 Revised Classification Criteria for RA.(Arnett et al) *or* if they have been determined to have RA based on a clinical evaluation by a board-certified Rheumatologist. Only CCP+ patients were studied in the current analysis (N=160).

4. Multiple Autoimmune Disease Genetics Consortium (MADGC). The MADGC collection of multiplex families includes RA patients enrolled who meet 1987 ACR criteria for disease. The details of the MADGC collection are previously published24. Only CCP+ patients were studied in the current analysis (N=105).

5. UCSF Rheumatoid Arthritis Genetics Project. Participants in the UCSF Rheumatoid Arthritis Genetics Project collection were recruited from UCSF Arthritis Clinics and private rheumatology practices in northern California as well as by nation wide outreach according to a protocol approved by the University of California, San Francisco Institutional Review Board. All patients fulfilled 1987 ACR criteria for RA. All patients are of European origin based on self-report, including grandparental countries of origin. Only CCP+ patients were studied in the current analysis (N=86).

6. Early Rheumatoid Arthritis Treatment Evaluation Registry (ERATER). 452 patients with RA of less than 3 years duration were enrolled in the ERATER study in Nashville, Tennessee between February 2001 and August 2004, including 336 patients from a private practice of Arthritis Specialists of Nashville, 48 from Vanderbilt University Rheumatology Clinics, and 68 from other sites. A detailed decription of this cohort has been published 25. Only CCP+ patients were studied in the current analysis (N=94).

Controls for the replication WGA were taken from control datasets that are publicly available in the Ilumina iControl datasbase. (www.illumina.com/iControlDB, Illumina, San Diego, CA). Additional control genotypes were from the neurodevelopmental control group obtained from the NIH Laboratory of Neurogenetics (http://neurogenetics.nia.nih.gov/

paperdata/public/). These control genotypes were selected from the entire set of European American genotypes available from these resources based on the following data filters: 1) > 90% complete genotyping data; 2) HW > 10^{-4}; and 3) >90% European continental ancestry. The European continental ancestry was determined using ancestry informative markers as previously described11. In addition we utilized control derived from a recent GWA study of lung cancer26 None of these controls overlap with controls utilized for the previously reported NARAC WGA study8. A third set of 1137 controls samples from M.D. Anderson Cancer Center Lung Cancer Study were used. These healthy individuals were seen for routine care at Kelsey-Seybold Clinics in the Houston Metropolitan area; all control subjects were required to have been current or former smokers.

The "NARAC II Replication" dataset contained RA patients from the following sources.

1. Treatment of Early Arthritis (TEAR) Trial. The TEAR trial is an investigator-initiated multicenter trial comparing two different strategies (early intensive therapy versus step-up therapy) and two combinations of medications (etanercept plus methotrexate [MTX] versus MTX plus hydroxychloroquine [HCQ] plus sulfasalazine [SSZ]). Patients under study are those with early RA and an aggressive clinical phenotype defined by presence of active synovitis of multiple joints, with positive RF or baseline erosions as assessed by radiographs of hands or feet. 347 RF+ patients and 50 RF- patients were studied (N=397).
2. The National Data Bank for Rheumatic Diseases (NDBRD). The National Databank is a longitudinal cohort of RA patients which has been described previously27. There is no overlap between the samples used in the previous WGA and the samples used for replication reported here. Of the 312 patients included for study, 258 were documented to be anti-CCP+, 15 were anti-CCP-, and 39 subjects were not tested (Total N=312)
3. The National Inception Cohort of Rheumatoid Arthritis (NICRA). The NICRA enrolled patients within 6 months of clinical diagnosis and has been described in previous publications28. Only CCP+ patients were studied in the current analysis. . There is no overlap between the samples used in the previous WGA and the samples used for replication reported here. (N=52)
4. Study of New Onset Rheumatoid Arthritis (SONORA). The SONORA enrolled RA patients and patients with polyarthritis within 12 months after clinical diagnosis. 29. Only patients meeting 1987 ACR critieria for RA and who were CCP+ were included in the current study. There is no overlap between the samples used in the previous WGA and the samples used for replication reported here. (N=184)
5. The GCI (Genomics Collaborative, Inc.) sample set consisted of 475 RA cases and 475 individually-matched controls. All case samples met the 1987 American College of Rheumatology diagnostic criteria for RA. All case samples were white North Americans of European descent who were rheumatoid factor positive. Control samples were healthy white individuals with no medical history of RA, also of European descent. A single control was matched to each case on the basis of

sex, age (+/- 5years), and self-reported ethnic background. Informed written consent was obtained from every subject.

Note that for analytic purposes, the NDRBD, NICRA, and SONORA cohorts were combined into a case group designated “OPA3” as shown in Supplementary Table 1B and Table 2.

Control subjects for replication “NARAC II- replication studies were taken in part from the New York Cancer Project 30 collection using subjects with self reported European ancestry (N=1163). Random matching of these controls (Supplementary Table 1B) with the TEAR and OPA3 replication datasets simply reflects the logistics of how replication genotyping was batched and has been used as a convenient way of grouping the analysis, shown in Supplementary Table 2. Additional controls were taken from the Controls (N=474) for the GCI collection are described above.

Subjects - Canadian cohorts

1. Unrelated RA probands (N = 598)for the WGA study were recruited from the Toronto area, the diagnosis of RA based on clinical, serological and radiological data in accordance with 1987 American College of Rheumatology criteria for RA. Subjects diagnosed with RA at an age of 16 years or younger were excluded from the study. Control samples (N=413) were individuals of European origin from the Toronto area who had no history of rheumatoid arthritis or other inflammatory disease. Informed consent, demographic data and blood samples for genomic DNA extraction were obtained from every subject following Institutional ethics committee approval.
2. A second , independent set of RA patients and healthy controls of European origin for replication were recruited from the Toronto(601) and Halifax(564) areas based on the criteria described above. Controls for the replication study were recruited in Toronto (N=921) and Halifax (N=324) again based on the criteria described above.

Genotyping and Quality Control Filtering

The NARAC-II and Canadian case control collections were genotyped using 373,400 SNPs on the Illumina HapMap370 BeadChip. Genotyping of the NARAC-II dataset was carried out at the Feinstein Institute for Medical Research while the Canadian samples were genotyped at Illumina in San Diego, CA. In order to organize the data to permit integration of results among studies, all genotypes were called using BeadStudio Software according to Top designation (http://www.illumina.com/downloads/TOPBOT_technote27Jun06.pdf). Each data set was subjected to quality control filtering based on SNP genotype call rates (>95% completeness), minor allele frequency (>0.01), and the Hardy-Weinberg equilibrium (P>1×10-4). Subjects with more than 5% missing genotype data or showing evidence of non-European ancestry were excluded. In addition samples showing evidence of relatedness to another sample, or possible DNA contamination, were also excluded. After filtering, genotypes derived from SNP markers that were common to both datasets were merged into a single file for analysis. A summary of all samples utilized for analysis after data cleaning is given in Supplementary Table 1. These filters were applied to each data set independently

and SNPs that passed quality control filters in each data set were then merged. In all a total of 278502 SNPs passed all filters

Replication genotyping of selected markers of interest was carried out using Sequenom iPlex technology at the Analytic Genetics Technology Centre in Toronto (Canadian samples) or within the Genetics and Genomics Branch of the National Institutes of Arthritis and Musculoskeletal and Skin Diseases (NARAC II-Replication dataset). The GCI cohort was genotyped using a kinetic PCR assay at Celera Diagnostics. Slightly different panels of SNPs were selected for replication in each dataset, based on technical considerations for efficient multiplexing of markers. The patterns of linkage disequilibrium reported in the HapMap informed the selection of replacement markers to develop marker sets for the replication studies. For rs13031237 and rs2736340, the PCR reaction provided results on the alternate strand (T/G for A/C and T/C for A/G, respectively for rs13031237 and rs2736340).

To allow for potential effects of population substructure and heterogeneity among populations we employed several techniques. In initial analyses of genome-wide data, we combined marker genotypes of all individuals and performed analyses to identify clusters of individuals who had similar genotypes using the program PLINK, with the clustering criterion set at PPC set at 0.0001. We also applied the program Eigenstrat to perform association analysis adjusting for correlations among the subjects according to marker similarities. For both clustering analyses in PLINK and construction of Eigenvectors using Eigenstrat we first removed SNPs on chromosomes 6p near the HLA region, 8p, and 17p that contain large polymorphic inversions. We removed these SNPs as the HLA region contains many SNPs relating to case-control status and including them could reduce our power to detect true associations. On the other hand the SNPs in chromosomes 8p and 17p show different LD patterns compared with other SNPs throughout the genome and we removed these regions to avoid influencing the results due to the potential chance variations caused by effects from many markers in these regions that are not of interest for correcting for population structure but that may show stronger correlations among subjects than expected.

To check for gene-gene interactions we evaluated all possible interactions between SNPs in CTLA4, Rel, and BLK were performed for the 157 most significantly associated non-HLA region SNPs with rheumatoid arthritis. Results showed no interaction effects with significance levels below $p < 0.001$. Given that 471 tests were performed, no significant interaction effects were identified, after correcting for multiple testing.

Lambda values indicating population structure varied among the populations being studied and according to the test conducted. For NARAC, the lambda value after correcting to a sample of 1000 cases and 1000 controls using PLINK was 1.148 and for Eigenstrat lambda was 1.029. For the Canadian sample, lambdas were 0.996 for PLINK and 1.010 for Eigenstrat. For the combined sample, the lambda value was 1.077 for PLINK and 1.023 for Eigenstrat.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by grants from the National Institutes of Health NO1-AR-2-2263 (PKG), RO1 AR44422 (PKG) and by the Eileen Ludwig Greenland Center for Rheumatoid Arthritis and the Muriel Fusfeld Foundation. The work was also supported in part by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases and by by grants from the Canadian institutes for Health Research (MOP79321) and Ontario Research Fund (RE01061)and a Canada Research Chair to KAS.

REFERENCES

1. Coenen MJ, Gregersen PK. Rheumatoid arthritis: a view of the current genetic landscape. Genes Immun. 2008
2. Raychaudhuri S, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet. 2008; 40:1216–23. [PubMed: 18794853]
3. Barton A, et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat Genet. 2008; 40:1156–9. [PubMed: 18794857]
4. Arnett FC, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988; 31:315–24. [PubMed: 3358796]
5. Schellekens GA, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum. 2000; 43:155–63. [PubMed: 10643712]
6. Ding B, et al. Different patterns of associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in the extended major histocompatibility complex region. Arthritis Rheum. 2009; 60:30–8. [PubMed: 19116921]
7. Lee HS, et al. Genetic risk factors for rheumatoid arthritis differ in Caucasian and Korean populations. Arthritis Rheum. 2009; 60:364–71. [PubMed: 19180477]
8. Plenge RM, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis--a genomewide study. N Engl J Med. 2007; 357:1199–209. [PubMed: 17804836]
9. Trynka G, et al. Coeliac disease associated risk variants in TNFAIP3 and REL implicate altered NF-{kappa}B signalling. Gut. 2009
10. Pennisi E. Genetics. 17q21.31: not your average genomic address. Science. 2008; 322:842–5. [PubMed: 18988819]
11. Hom G, et al. Association of Systemic Lupus Erythematosus with C8orf13-BLK and ITGAM-ITGAX. N Engl J Med. 2008
12. Harley JB, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet. 2008; 40:204–10. [PubMed: 18204446]
13. Hayden MS, West AP, Ghosh S. NF-kappaB and the immune response. Oncogene. 2006; 25:6758–80. [PubMed: 17072327]
14. Mise-Omata S, et al. A proximal kappaB site in the IL-23 p19 promoter is responsible for RelA- and c-Rel-dependent transcription. J Immunol. 2007; 179:6596–603. [PubMed: 17982049]
15. Grumont R, et al. c-Rel regulates interleukin 12 p70 expression in CD8(+) dendritic cells by specifically inducing p35 gene transcription. J Exp Med. 2001; 194:1021–32. [PubMed: 11602633]
16. Mason NJ, Liou HC, Hunter CA. T cell-intrinsic expression of c-Rel regulates Th1 cell responses essential for resistance to Toxoplasma gondii. J Immunol. 2004; 172:3704–11. [PubMed: 15004174]
17. Bunting K, et al. Genome-wide analysis of gene expression in T cells to identify targets of the NF-kappa B transcription factor c-Rel. J Immunol. 2007; 178:7097–109. [PubMed: 17513759]
18. Owyang AM, et al. c-Rel is required for the protection of B cells from antigen receptor-mediated, but not Fas-mediated, apoptosis. J Immunol. 2001; 167:4948–56. [PubMed: 11673501]

19. Lu KT, Sinquett FL, Dryer RL, Song C, Covey LR. c-Rel plays a key role in deficient activation of B cells from a non-X-linked hyper-IgM patient. Blood. 2006; 108:3769–76. [PubMed: 16896156]
20. Zhou HJ, et al. Nuclear CD40 interacts with c-Rel and enhances proliferation in aggressive B-cell lymphoma. Blood. 2007; 110:2121–7. [PubMed: 17567982]
21. Plenge RM, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet. 2007
22. Toubi E, Shoenfeld Y. The role of CD40-CD154 interactions in autoimmunity and the benefit of disrupting this pathway. Autoimmunity. 2004; 37:457–64. [PubMed: 15621572]
23. Boumpas DT, et al. A short course of BG9588 (anti-CD40 ligand antibody) improves serologic activity and decreases hematuria in patients with proliferative lupus glomerulonephritis. Arthritis Rheum. 2003; 48:719–27. [PubMed: 12632425]
24. Criswell LA, et al. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. Am J Hum Genet. 2005; 76:561–71. [PubMed: 15719322]
25. Sokka T, Pincus T. An Early Rheumatoid Arthritis Treatment Evaluation Registry (ERATER) in the United States. Clin Exp Rheumatol. 2005; 23:S178–81. [PubMed: 16273804]
26. Wang Y, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet. 2008
27. Wolfe F, Michaud K, Gefeller O, Choi HK. Predicting mortality in patients with rheumatoid arthritis. Arthritis Rheum. 2003; 48:1530–42. [PubMed: 12794820]
28. Fries JF, et al. HLA-DRB1 genotype associations in 793 white patients from a rheumatoid arthritis inception cohort: frequency, severity, and treatment bias. Arthritis Rheum. 2002; 46:2320–9. [PubMed: 12355479]
29. Weisman M, et al. Analysis at one year of an inception cohort of early rheumatoid arthritis(RA): The SONORA study. Arthritis and Rheumatism. 2003; 48:S119.
30. Mitchell MK, Gregersen PK, Johnson S, Parsons R, Vlahov D. The New York Cancer Project: rationale, organization, design, and baseline characteristics. J Urban Health. 2004; 81:301–10. [PubMed: 15136663]

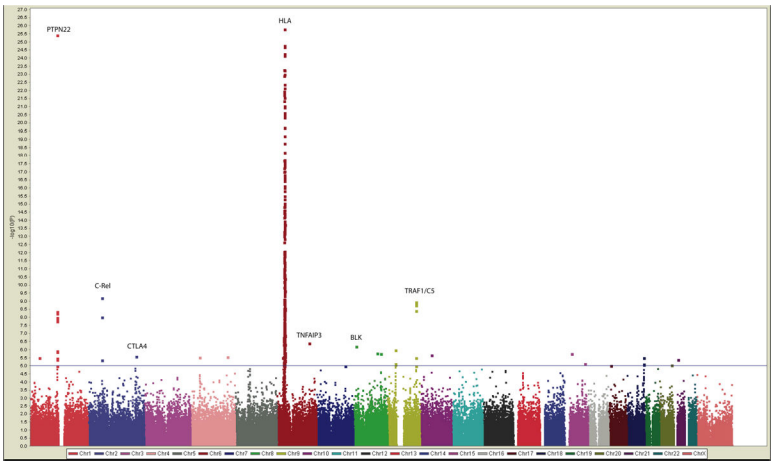

Figure 1.

Summary of genome-wide association scan results for 2418 cases and 4504 controls. The -$\log_{10}$ of the trend test P values after homogeneous clustering and corrected for the residual genome wide inflation of chi square are plotted against position on each chromosome. Chromosomes are shown in alternating colors for clarity. The blue horizontal line indicates SNPs that are significant at a level ($P = 5 \times 10^{-5}$). The results for markers in the HLA region are truncated.

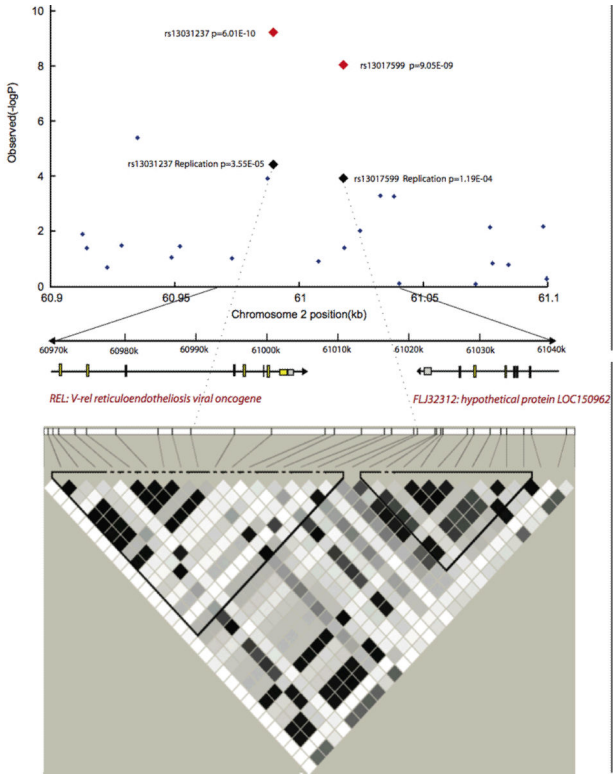

Figure 2.

Association localization plots in the region around *REL* following discovery and replication phases. The p values for all samples in the GWA scan are shown as small blue diamonds, with the exception of the results for two SNPs, rs13031237 and rs13017599 ($P = 9.05 \times 10^{-9}$ and $p = 6.01 \times 10^{-10}$) in c-Rel , which are presented as red diamonds. The black diamonds show the independent results for these two SNPs in the replication samples.

Table 1

Summary of association results for *REL*, *CTLA4* and *BLK*.

Gene	SNP	Study	Cases	Controls	Minor Allele Frequency		OR	95 % CI	p-value
(alleles)					Cases	Controls			
REL *	rs13031237	GWA	1875/2747	3086/5776	0.406	0.348	1.268	1.176-1.367	6.01×10^{-10}
A/C		Replication	1536/2494	1579/3085	0.381	0.339	1.207	1.104-1.320	3.55×10^{-5}
		Combined	3411/5241	4665/8861	0.394	0.345	1.246	1.177-1.318	3.08×10^{-14}
REL	rs13017599	GWA	1837/2751	3071/5773	0.400	0.347	1.248	1.157-1.345	9.05×10^{-9}
A/G		Replication	1890/3158	1920/3736	0.374	0.340	1.171	1.080-1.269	1.19×10^{-4}
		Combined	3727/5909	4991/9509	0.387	0.344	1.214	1.150-1.282	2.60×10^{-12}
CTLA4	rs231735	GWA	2057/2563	4319/4541	0.445	0.488	0.863	0.802-0.929	8.24×10^{-5}
G/T		Replication	2154/2750	2706/2930	0.439	0.480	0.869	0.804-0.939	3.81×10^{-4}
		Combined	4211/5313	7025/7471	0.442	0.485	0.855	0.812-0.902	6.25×10^{-9}
BLK	rs2736340	GWA	1307/3315	2113/6747	0.283	0.239	1.234	1.136-1.341	6.06×10^{-7}
A/G		Replication	1338/3732	1372/4276	0.264	0.243	1.122	1.027-1.225	1.08×10^{-2}
		Combined	2645/7047	3485/11023	0.273	0.240	1.194	1.125-1.268	5.69×10^{-9}

* GWA results: plink cmh test stratified by population-based identity by state clusters; Replication: cmh test stratified by 5 centers including Halifax, Toronto, OPA3, TEAR and GCI (except rs13031237 which was not genotyped in the GCI collection); combined: cmh test stratified by 8 centers. Including NARACI, NARAC II, Canada + lung cancer controls, Halifax, Toronto, OPA3, TEAR and GCI (except rs13031237 which was not genotyped in the GCI collection)