EXTENDED REPORT

Preferential association of a functional variant in complement receptor 2 with antibodies to double-stranded DNA

Jian Zhao,1 Brendan M Giles,2 Rhonda L Taylor,3 Gabriel A Yette,2 Kara M Lough,2 Han Leng Ng,3 Lawrence J Abraham,3 Hui Wu,1 Jennifer A Kelly,4 Stuart B Glenn,4 Adam J Adler,4 Adrienne H Williams,5 Mary E Comeau,5 Julie T Ziegler,5 Miranda Marion,5 Marta E Alarcón-Riquelme4,6 for the BIOLUPUS and GENLES Networks, Graciela S Alarcón,7 Juan-Manuel Anaya,8 Sang-Cheol Bae,9 Dam Kim,9 Hye-Soon Lee,9 Lindsey A Criswell,10 Barry I Freedman,11 Gary S Gilkeson,12 Joel M Guthridge,4 Chaim O Jacob,13 Judith A James,4,14,15 Diane L Kamen,12 Joan T Merrill,16 Kathy Moser Sivils,4,14 Timothy B Niewold,17 Michelle A Petri,18 Rosalind Ramsey-Goldman,19 John D Reveille,20 R Hal Scofield,4,15,21 Anne M Stevens,22,23 Luis M Vilá,24 Timothy J Vyse,25 Kenneth M Kaufman,26,27 John B Harley,26,27 Carl D Langefeld,5 Patrick M Gaffney,4 Elizabeth E Brown,7,28 Jeffrey C Edberg,7 Robert P Kimberly,7 Daniela Ulgiati,3 Betty P Tsao,1 Susan A Boackle2,29

Handling editor Tore K Kvien

ABSTRACT

Objectives Systemic lupus erythematosus (SLE; OMIM 152700) is characterised by the production of antibodies to nuclear antigens. We previously identified variants in complement receptor 2 (CR2/CD21) that were associated with decreased risk of SLE. This study aimed to identify the causal variant for this association.

Methods Genotyped and imputed genetic variants spanning CR2 were assessed for association with SLE in 15 750 case-control subjects from four ancestral groups. Allele-specific functional effects of associated variants were determined using quantitative real-time PCR, quantitative flow cytometry, electrophoretic mobility shift assay (EMSA) and chromatin immunoprecipitation (ChIP)-PCR.

Results The strongest association signal was detected at rs1876453 in intron 1 of CR2 (\(p_{\text{meta}}=4.2\times10^{-4}\), OR 0.85), specifically when subjects were stratified based on the presence of dsDNA autoantibodies (case-control \(p_{\text{meta}}=7.6\times10^{-7}\), OR 0.71; case-only \(p_{\text{meta}}=1.9\times10^{-4}\), OR 0.75). Although allele-specific effects on B cell CR2 mRNA or protein levels were not identified, levels of complement receptor 1 (CR1/CD35) mRNA and protein were significantly higher on B cells of subjects harbouring the minor allele (\(p=0.0248\) and \(p=0.0006\), respectively). The minor allele altered the formation of several DNA protein complexes by EMSA, including one containing CCCTC-binding factor (CTCF), an effect that was confirmed by ChIP-PCR.

Conclusions These data suggest that rs1876453 in CR2 has long-range effects on gene regulation that decrease susceptibility to lupus. Since the minor allele at rs1876453 is preferentially associated with reduced risk of the highly specific dsDNA autoantibodies that are present in preclinical, active and severe lupus, understanding its mechanisms will have important therapeutic implications.

INTRODUCTION

Systemic lupus erythematosus (SLE (OMIM 152700)) is a heterogeneous autoimmune disease with a strong genetic component modified by environmental exposures. The complement system has been linked to its pathogenesis since low serum complement levels were first demonstrated in patients with active disease.1,2 Although complement activation resulting in tissue damage is an important feature of lupus, deficiencies of early classical complement pathway components are paradoxically strongly associated with lupus susceptibility. Deficiencies or altered function of complement receptors have also been associated with lupus in humans and murine models of disease.3–5 These effects have been attributed to alterations in antigen clearance, antigen processing, tolerance induction and cell activation, but the exact mechanisms remain poorly understood.

SLE is characterised by the production of class-switched autoantibodies against nuclear antigens that have undergone affinity maturation, suggesting their generation in germinal centre reactions. Complement receptor 2 (CR2/CD21) is primarily expressed on mature B cells and follicular dendritic cells, two major components of germinal centres. We first showed Cr2, which encodes both CR2 and complement receptor 1 (CR1/CD35) in the mouse, to be a candidate gene for lupus susceptibility in the NZM2410 model of lupus based on structural and functional alterations in its gene.
Regarding how CR2 contributes to disease development.

manifestations of lupus in order to generate new hypotheses previously analysed.8 Subjects for functional studies were geno-

ciers removed. Final clean data were from European Americans analysis12 and ADMIXMAP13 as described14 and genetic out-

Global ancestry was estimated based on the genotype of ancestors identical by descent >0.4 or gender mismatch were removed. 

adults without family history of autoimmune disease at the Clinical Immunology laboratory. Samples for functional analyses 

data were collected by chart review or testing in the OMRF College of Rheumatology revised classification criteria.9 Clinical 

In this study, we fine-mapped the region spanning CR2 in 15 750 subjects from four ancestral groups to identify potential 

causal variant(s) for these associations with lupus. Additionally, we explored the association of CR2 polymorphisms with clinical 

manifestations of lupus in order to generate new hypotheses regarding how CR2 contributes to disease development.

METHODS

Subjects

DNA from individuals recruited from multiple sites was processed at the Oklahoma Medical Research Foundation (OMRF; Large Lupus Association Study 2) with institutional review board approval. All patients with SLE met the 1997 American College of Rheumatology revised classification criteria.5 Clinical data were collected by chart review or testing in the OMRF Clinical Immunology laboratory. Samples for functional analyses were from healthy non-smoking 18-year-old to 60-year-old adults without family history of autoimmune disease at the University of Colorado School of Medicine.

Genotyping

Genotyping was performed on the OMRF Illuma iSelect platform.9 11 Subjects with missing genotype rate >10%, shared identical by descent >0.4 or gender mismatch were removed. Global ancestry was estimated based on the genotype of ancestry informative markers (AIMs), using principal components analysis15 and ADMIXMAP13 as described14 and genetic outliers removed. Final clean data were from European Americans (EA), African Americans (AA; 7.5% Gullahs), Asians (AS; 74.6% Koreans, 16.1% Chinese, 9.3% Japanese and Singaporeans) and Hispanics (HS) enriched for Amerindian–European admixture, 2001 EA cases and 2153 EA controls were previously analysed.6 Subjects for functional studies were genotyped using a Taqman SNP Genotyping Assay.

Imputation

SNP and insertion-deletion (INDEL) genotypes of 379 Europeans, 246 Africans, 286 Asians and 181 Americans from the 1000 Genomes Project (V3, Phase 1 integrated data, March 2012 release) were references in imputation for EA, AA, AS and HS subjects, respectively. Imputation was performed using IMPUTE 2.1.215; genotypes with information scores >0.9 and minor allele frequency (MAF)>0.01 were further analysed.

Association tests

SNPs and INDELS showing biased Hardy–Weinberg equilibrium (p<0.01 in controls, p<0.0001 in cases), missing genotype rate >5% or different missing genotype rates between cases and controls (rate >2% and p<0.05) were excluded. Variants were assessed for association with SLE under a logistic regression model, and haplotypic and haplotype-based conditional association tests were performed, adjusting for gender and the first three principal components estimated using AIMs. For trans- cestral meta-analysis, a fixed effect model was applied if Cochran’s Q statistic showed no evidence of genetic heterogeneity among ORs (p>0.05); otherwise, a random effect model was used. Analyses were performed using PLINK V1.07.16

Sample preparation

Peripheral blood mononuclear cells (PBMC) were isolated over Ficoll–Paque (Sigma–Aldrich). DNA was purified using the QIAamp DNA Mini Kit (QIAgen). B cells were purified using the Easy Sep Human B Cell Enrichment Kit (StemCell Technologies). RNA was processed using the RNeasy Plus Mini Kit (QIAgen). Quality check and quantification of RNA was performed using the Agilent 2100 Bioanalyzer. RNA and DNA were stored at -70°C. Epstein–Barr virus (EBV)-transformed B cell lines were generated by incubating PBMCs with supernatant from cell line GM7404A and cyclosporine A.

qPCR and quantitative flow cytomtery

qPCR of primary B cell transcripts was performed using cDNA transcribed using random primers and MultiScribe reverse transcriptase (Applied Biosystems), customised CR2 primers (5’-CGGAAAGTATATCTGCTGTGACATCA-3’, 5’-CTAATC AATFATTGCCGTGATC-3’) and probe (6FAM-AACCTG TGTGTCCTCA-MGBNFQ), Taqman assays for CR1 (4331182) and β-actin (4332935E), and the Applied Biosysyms 7500 Real-Time PCR System. Relative expression levels of CR2 and CR1, normalised to β-actin, were calculated using the comparative C_T method.17

Quantitative flow cytometry was performed using Quantum Simply Cellular microbeads (Bangs Laboratories). Purified B cells were incubated with ethidium monoazide to exclude dead cells and mouse IgG to block Fc receptors and stained with bioin
tinylated anti-human CD19 (clone HD37) followed by AF405-labelled streptavidin (Molecular Devices) and either AF488-labelled anti-human CR1 (clone 6B1.H12) or AF488-labelled anti-human CR2 (clone THBS5, Harlan). Data acquired on fixed cells within 24 h using the LSR II and CellQuest software (BD Biosciences) were analysed with FlowJo (Tree Star).

Electrophoretic mobility shift assay

39-mer oligonucleotides (5’-GGGCCAAAAGCGAGACGGTG[A]G [AGGGGACGTCTGCAGCAGC]–3’, 5’-CGTCAACGTGACGTCCCTCC [CT]ACCGTGCGCTTGGCCC–3’) were synthesised, HPLC-purified and biotin-labelled; 25 fmol labelled oligonucleotides were incubated with 2–4 μg nuclear extracts (NE) or 50 mM potassium chloride (for unbound controls) in 4% Ficoll, 20 mM N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid (HEPES), 1 mM EDTA, 0.5 mM dithiothreitol and 1 μg poly di:dc for 30 min on ice. For competition and blocking experiments, NE were pre-incubated with unlabelled oligonucleotide or anti-CCCTC-binding factor (CTCF) (sc-15914; Santa Cruz Biotechnology) for 10 or 30 min on ice. Reactions were electrophoresed on 6% DNA retardation gels (Life Technologies), transferred and cross-linked using ultraviolet light to nylon membranes and visualised using the Chemiluminescent Nucleic Acid Detection Module (Thermo Fisher Scientific).

Chromatin immunoprecipitation-PCR

Chromatin immunoprecipitation (ChIP) was performed using the Upstate Biotechnology ChIP assay kit with minor modifications.2 1.1×10^7 cells were incubated with 1% formaldehyde for 15 min and 0.125 M glycine added for 5 min. Washed cells
<table>
<thead>
<tr>
<th>Subphenotypes</th>
<th>N, number in sample</th>
<th>MAF, minor allele frequency</th>
<th>Pos, subphenotype positive</th>
<th>Neg, subphenotype negative</th>
<th>p value, p value and OR were calculated using a logistic regression model</th>
<th>OR (95% CI)</th>
<th>Pos vs Neg</th>
<th>OR (95% CI)</th>
<th>Met analyses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malar rash</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.89 (0.78 to 1.03)</td>
<td>0.92</td>
<td>1.01</td>
<td>0.85 (0.83 to 1.21)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>685</td>
<td>0.08</td>
<td>0.066</td>
<td>0.086</td>
<td>0.24</td>
<td>0.2</td>
<td>1.24</td>
<td>0.92 (0.95 to 1.65)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>836</td>
<td>0.03</td>
<td>0.032</td>
<td>0.051</td>
<td>3.0 × 10^-3</td>
<td>0.52</td>
<td>1.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discoid rash</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.62</td>
<td>0.51</td>
<td>0.92</td>
<td>0.57 (0.57 to 1.40)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>485</td>
<td>0.08</td>
<td>0.065</td>
<td>0.086</td>
<td>0.81</td>
<td>0.43</td>
<td>1.14</td>
<td>0.37 (0.37 to 1.53)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>180</td>
<td>0.02</td>
<td>0.036</td>
<td>0.051</td>
<td>0.032</td>
<td>0.3</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photosensitivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.26</td>
<td>0.26</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral ulcers</td>
<td>115</td>
<td>0.10</td>
<td>0.09</td>
<td>0.099</td>
<td>0.11</td>
<td>0.2</td>
<td>1.14</td>
<td>0.95 (0.95 to 1.36)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>495</td>
<td>0.059</td>
<td>0.081</td>
<td>0.086</td>
<td>1.7 × 10^-3</td>
<td>0.59</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>562</td>
<td>0.036</td>
<td>0.027</td>
<td>0.051</td>
<td>0.094</td>
<td>0.2</td>
<td>1.33</td>
<td>0.85 (0.85 to 2.08)</td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.96</td>
<td>1.02</td>
<td>0.62 (0.62 to 1.66)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>1166</td>
<td>0.076</td>
<td>0.059</td>
<td>0.086</td>
<td>0.048</td>
<td>0.42</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>1007</td>
<td>0.031</td>
<td>0.031</td>
<td>0.051</td>
<td>3.7 × 10^-3</td>
<td>0.96</td>
<td>1.01</td>
<td>0.62 (0.62 to 1.66)</td>
<td></td>
</tr>
<tr>
<td>Serositis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.96</td>
<td>1.02</td>
<td>0.62 (0.62 to 1.66)</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>653</td>
<td>0.07</td>
<td>0.076</td>
<td>0.086</td>
<td>0.025</td>
<td>0.55</td>
<td>0.92</td>
<td>0.69 (0.69 to 1.22)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>384</td>
<td>0.029</td>
<td>0.03</td>
<td>0.051</td>
<td>0.022</td>
<td>0.21</td>
<td>0.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.59</td>
<td>0.96</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>697</td>
<td>0.065</td>
<td>0.081</td>
<td>0.086</td>
<td>6.1 × 10^-3</td>
<td>0.59</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>644</td>
<td>0.032</td>
<td>0.028</td>
<td>0.051</td>
<td>0.029</td>
<td>0.4</td>
<td>1.21</td>
<td>0.77 (0.77 to 1.91)</td>
<td></td>
</tr>
<tr>
<td>Haematological disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.96</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>1011</td>
<td>0.07</td>
<td>0.079</td>
<td>0.086</td>
<td>0.012</td>
<td>0.45</td>
<td>0.89</td>
<td>0.65 (0.65 to 1.22)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>811</td>
<td>0.031</td>
<td>0.028</td>
<td>0.051</td>
<td>9.6 × 10^-3</td>
<td>0.79</td>
<td>1.07</td>
<td>0.65 (0.65 to 1.77)</td>
<td></td>
</tr>
<tr>
<td>Anti-dsDNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.80</td>
<td>0.96</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>675</td>
<td>0.067</td>
<td>0.083</td>
<td>0.086</td>
<td>9.7 × 10^-3</td>
<td>0.69</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>726</td>
<td>0.03</td>
<td>0.028</td>
<td>0.051</td>
<td>7.1 × 10^-3</td>
<td>0.81</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-Sm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.2 × 10^-3</td>
<td>0.68</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>430</td>
<td>0.087</td>
<td>0.076</td>
<td>0.086</td>
<td>0.78</td>
<td>0.44</td>
<td>0.89</td>
<td>0.65 (0.65 to 1.22)</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>294</td>
<td>0.03</td>
<td>0.029</td>
<td>0.051</td>
<td>0.069</td>
<td>0.26</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-RNP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0 × 10^-3</td>
<td>0.91</td>
<td>1.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>378</td>
<td>0.078</td>
<td>0.077</td>
<td>0.086</td>
<td>0.34</td>
<td>0.7</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>171</td>
<td>0.04</td>
<td>0.04</td>
<td>0.051</td>
<td>0.033</td>
<td>0.15</td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-SSA/Ro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.86</td>
<td>0.86</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>450</td>
<td>0.086</td>
<td>0.093</td>
<td>0.099</td>
<td>0.28</td>
<td>0.49</td>
<td>1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>292</td>
<td>0.062</td>
<td>0.08</td>
<td>0.086</td>
<td>0.023</td>
<td>0.21</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-SSB/La</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.85</td>
<td>0.87</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>81</td>
<td>0.037</td>
<td>0.08</td>
<td>0.086</td>
<td>0.025</td>
<td>0.53</td>
<td>1.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>71</td>
<td>0.042</td>
<td>0.033</td>
<td>0.051</td>
<td>0.63</td>
<td>0.53</td>
<td>1.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.85</td>
<td>0.87</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N, number in sample; MAF, minor allele frequency; Pos, subphenotype positive; Neg, subphenotype negative; p value, p value and OR were calculated using a logistic regression model adjusted for gender and global ancestry; EA, European-American; AA, African-American; HS, Hispanic; Meta, Meta-analysis; SLE, systemic lupus erythematosus. p Values ≤0.05 are marked with bold italics; significant p values after Bonferroni correction were ≤3.6 × 10^-3. ** doi:10.1136/annrheumdis-2014-205584**
were incubated in NP-40 lysis buffer and nuclei pelleted, resuspended in sodium dodecyl sulfate lysis buffer and sonicated five times on ice. Precleared soluble chromatin was incubated with anti-CTCF, an isotype control, or no antibody followed by Protein G agarose/salmon sperm DNA. qPCR was performed on immune complex-associated DNA using primers spanning rs1876453 (5′-GGAAAGTTTCTGTGCCGCGA-3′, 5′-GACATTCTAGGCACCACTTCTCCCTGCG-3′), SYBR Green-based detection, and the Illumina Eco Real-Time PCR System. A standard curve was constructed using a chromatin input control.

Statistical analysis
A two-tailed unpaired Student t test detected differences between groups. Spearman rank correlation determined correlation within a group. A p value of <0.05 was considered significant. Statistics and graphs were generated using GraphPad Prism software.

RESULTS
Association of rs1876453 with decreased risk of lupus
To localise causal variants of the CR2 gene, we genotyped 56 SNPs in CR2 and CR1 and 347 AIMs in 15 750 unrelated case-control subjects from four ancestral groups: EA (3872 cases vs 4402 controls), AA (2111 cases vs 2112 controls), HS (1300 cases vs 1301 controls) and EA (3871 cases vs 4403 controls), SYBR Green-based detection, and the Eco Real-Time PCR System (Illumina). Transcript abundance was determined by comparison with a standard curve and normalised to β-actin.

Statistical analysis
A two-tailed unpaired Student t test detected differences between groups. Spearman rank correlation determined correlation within a group. A p value of <0.05 was considered significant. Statistics and graphs were generated using GraphPad Prism software.

RESULTS
Association of rs1876453 with decreased risk of lupus
To localise causal variants of the CR2 gene, we genotyped 56 SNPs in CR2 and CR1 and 347 AIMs in 15 750 unrelated case-control subjects from four ancestral groups: EA (3872 cases vs 4402 controls), AA (2111 cases vs 2112 controls), HS (1300 cases vs 1301 controls) and EA (3871 cases vs 4403 controls), SYBR Green-based detection, and the Eco Real-Time PCR System (Illumina). Transcript abundance was determined by comparison with a standard curve and normalised to β-actin.

Statistical analysis
A two-tailed unpaired Student t test detected differences between groups. Spearman rank correlation determined correlation within a group. A p value of <0.05 was considered significant. Statistics and graphs were generated using GraphPad Prism software.

RESULTS
Association of rs1876453 with decreased risk of lupus
To localise causal variants of the CR2 gene, we genotyped 56 SNPs in CR2 and CR1 and 347 AIMs in 15 750 unrelated case-control subjects from four ancestral groups: EA (3872 cases vs 4402 controls), AA (2111 cases vs 2112 controls), HS (1300 cases vs 1301 controls) and EA (3871 cases vs 4403 controls), SYBR Green-based detection, and the Eco Real-Time PCR System (Illumina). Transcript abundance was determined by comparison with a standard curve and normalised to β-actin.
Controls), AA (1676 vs 1929), AS (1265 vs 1260) and HS (1492 vs 807). Genotypes for additional variants (SNPs and INDELs) were imputed using reference data from the 1000 Genomes Project. After applying quality control measures, 138 EA, 167 AA, 102 AS and 133 HS genetic variants deeply covering a 57.6 kB region 5′ upstream of \(\text{CR2}\) to intron 1 of \(\text{CR1}\) were assessed for association with SLE (see online supplementary figure S1 and table S1). Only rs1876453, in intron 1 of \(\text{CR2}\), was consistently associated with SLE in EA (MAF 0.086 in cases vs 0.099 in controls, \(p=0.045\), OR 0.89 (95% CI 0.79 to 0.99)), AA (0.075 vs 0.086, \(p=0.045\), OR 0.83 (95% CI 0.70 to 0.92)) and HS (0.032 vs 0.051, \(p=3.5\times10^{-3}\), OR 0.62 (95% CI 0.46 to 0.86)). rs1876453 was not associated with SLE in AS given the low MAF in this group (0.0004 vs 0.0008, \(p=0.68\), OR 0.60 (95% CI 0.45 to 0.79)). In transancestral meta-analysis of 75 genetic variants assessed in EA, AA, AS and HS (see online supplementary figure S1 and table S1), the strongest association was detected at rs1876453 (\(p=6.6\times10^{-4}\), OR 0.73 (95% CI 0.61 to 0.88); \(p_{\text{AA}}=9.7\times10^{-3}\), OR 0.72 (95% CI 0.56 to 0.92); \(p_{\text{HS}}=7.1\times10^{-3}\), OR 0.58 (95% CI 0.39 to 0.86); \(p_{\text{meta}}=7.6\times10^{-7}\), OR 0.71) (table 1). In cases only, the minor allele of rs1876453 was associated with decreased risk of dsDNA autoantibodies (\(p_{\text{meta}}=1.9\times10^{-4}\), OR 0.75) and renal disorder (\(p_{\text{meta}}=1.3\times10^{-3}\), OR 0.83) (table 1), but only the association with dsDNA autoantibodies remained significant after correction for multiple comparisons.

Increased signal associated with rs1876453 after subphenotype stratification

We next assessed the association of rs1876453 with 14 lupus subphenotypes. In case controls, the minor allele was consistently associated with decreased risk of serositis, renal disorder, haematological disorder and dsDNA autoantibodies in EA, AA and HS, of which the strongest association was detected with dsDNA autoantibodies (\(p_{\text{EA}}=6.6\times10^{-4}\), OR 0.73 (95% CI 0.61 to 0.88); \(p_{\text{AA}}=9.7\times10^{-3}\), OR 0.72 (95% CI 0.56 to 0.92); \(p_{\text{HS}}=7.1\times10^{-3}\), OR 0.58 (95% CI 0.39 to 0.86); \(p_{\text{meta}}=7.6\times10^{-7}\), OR 0.71) (table 1). In cases only, the minor allele of rs1876453 was associated with decreased risk of dsDNA autoantibodies (\(p_{\text{meta}}=1.9\times10^{-4}\), OR 0.75) and renal disorder (\(p_{\text{meta}}=1.3\times10^{-3}\), OR 0.83) (table 1), but only the association with dsDNA autoantibodies remained significant after correction for multiple comparisons.

rs1876453 tags the protective EA haplotype

No other SNP in this region exhibited stronger association with dsDNA autoantibodies in EA, AA and HS than rs1876453

Figure 2  Allele-specific effects of rs1876453 on CR1 and CR2 expression. Relative amounts of CR2 (A) or CR1 (C) RNA transcripts in primary B cells from 35 or 40 healthy donors respectively were measured by qPCR using the comparative Ct method.\(^{17}\) Levels of surface CR2 (B) and CR1 (D) on primary B cells from 131 healthy donors were determined by quantitative flow cytometry (ABC, antibody binding capacity). Each point represents a unique subject, and the line and error bars represent the mean SD for each group; \(p\) values were determined using a two-tailed Student t test, and a \(p\) value of <0.05 was considered significant. (E) Levels of surface CR1 and CR2 were plotted and subjected to correlation analysis and linear regression. Each point represents a unique subject (minor allele, squares; major allele, triangles) and the lines represent the line of best fit for each allele (minor allele, dashed line; major allele, solid line). (F) Levels of surface CR1 and CR2 were also used to calculate the ratio of CR1 to CR2. Each point represents a unique subject, and the line and error bars represent the mean±SD for each group; \(p\) values were determined using a two-tailed Student t test and a \(p\) value of <0.05 was considered significant.
(figure 1A–C, see online supplementary table S2). When conditioning on rs1876453, the association signals at all other SNPs were eliminated (see online supplementary table S2). The H1 haplotype, constructed using rs1876453 and previously associated or tightly linked CR2 SNPs, was equivalent to the protective EA SLE haplotype\(^8\) and was perfectly tagged in EA by the minor allele of rs1876453 (figure 1D). In all three ancestral groups, it was associated with decreased risk of SLE (\(p_{EA}=0.037, p_{AA}=0.043\) and \(p_{HS}=0.042\)) and dsDNA autoantibodies (\(p_{EA}=6.0\times 10^{-4}, p_{AA}=7.8\times 10^{-3}\) and \(p_{HS}=0.034\)).

The minor allele at rs1876453 is associated with altered expression of B cell CR1

We did not detect allele-specific differences in CR2 mRNA or protein levels in primary B cells from healthy controls (figure 2A, B). However, CR1 mRNA and protein levels were significantly higher in subjects with the minor allele (figure 2C, D). Although B cell CR1 and CR2 levels were positively correlated in both groups (major allele Spearman \(r=0.6342, p<0.0001\); minor allele Spearman \(r=0.7998, p<0.0001\), figure 2E), the ratio of CR1 to CR2 on B cells was higher in subjects with the minor allele (\(p=0.0002\); figure 2F). Allele-specific differences in CR1 expression were not detected on other peripheral blood cells (see online supplementary figure S2). These data suggest that rs1876453 has long-range effects on the regulation of expression of CR1, which lies directly 3' of CR2, that are either B cell-specific or dependent on coexpression of CR2.

The minor allele at rs1876453 alters transcription factor binding

rs1876453 is 97 nucleotides from the 5' end of the first intron of CR2. This \(\sim 12\) kb intron contains a conserved silencing domain that controls CR2 expression in a cell type-specific and developmentally regulated manner.\(^{19–22}\) The ENCODE\(\text{Of DNA Elements (ENCODE) database}\)\(^{23}\) reports the in vivo interaction of the region surrounding rs1876453 with multiple transcription factors in B cells, including Pax5, Oct2 and CTCF (figure 3). We confirmed the formation of multiple protein-DNA complexes containing the sequence surrounding rs1876453 by electrophoretic mobility shift assay (EMSA) using NE from several cell lines (K562, erythroid, CR1+CR2+; Reh, pre-B, CR1+CR2+; Ramos, mature B, CR1+CR2+, see online supplementary figure S3). In the presence of the minor A allele at rs1876453, transcription factor binding was reduced or

![Figure 3](http://example.com/figure3.png)

**Figure 3** The ENCODE\(\text{Of DNA Elements (ENCODE) Project} data surrounding rs1876453. (A) The first exon and 5' end of the first intron of the CR2 gene. The 5' untranslated region (5' UTR) is shown before the methionine start codon, in green. These data are derived from the University of California Santa Cruz (UCSC) Genes Track. (B) The location of rs1876453 (highlighted in yellow) and previously reported systemic lupus erythematosus-associated rs1839346 (in blue font). These data are derived from the Common Single-Nucleotide Polymorphisms (SNP) (138) Track (ft.ncbi.nih.gov/snp).\(^39\) (C) DNase hypersensitive sites in the GM12878 Epstein–Barr virus (EBV)-transformed B cell line and in primary CD20+ B cells derived by DNase-seq. These data are derived from the UCSC Uniform DNase HS Track. Signal values are shown as grayscale-coloured items where higher signal values correspond to darker-coloured blocks. Primary B cells contain an additional hypersensitivity site that overlies rs1876453. (D) Histone marks surrounding rs1876453, as determined by chromatin immunoprecipitation (ChIP)-seq. These data were derived from the Layered H3K4Me3, H3K4Me1 and H3K27Ac Tracks. The H3K4Me3 histone mark is associated with poised or active promoters, the H3K4me1 histone mark is associated with enhancers and with DNA regions downstream of transcription sites and the H3K27Ac histone mark may enhance transcription by blocking the spread of the repressive histone mark H3K27Me3. Data shown are for the GM12878 EBV-transformed B cell line. (E) Transcription factor binding sites determined by ChIP-seq are shown as grey boxes that encompass the peaks of transcription factor occupancy, with the darkness of the box proportional to the maximal signal strength observed in any cell line. The name to the left of the box is the transcription factor, and includes in parentheses the antibody used for ChIP. The letters to the right of the box indicate the cell lines in which a signal is detected, with the darkness of the letter proportional to the signal strength in the cell line. Data are derived from the Transcription Factor ChIP Track. CTCF-binding factor (CTCF) binding was seen in multiple EBV-transformed B cell lines (G, g) as well as a variety of other cell lines. (F) CTCF binding to primary CD20+ B cells by ChIP-seq. Peak occupancy lies over exon 1 and the 5' UTR. Data are derived from the Broad Histone Track. (G) Transcription levels for several cell types assayed by high-throughput sequencing of polyadenylated RNA (RNA-seq). Each cell line is associated with a particular colour; the GM12878 cell line is shown in pink. This figure was obtained from the UCSC Genome Browser (Human Feb 2009 (GRCh37/hg19) Assembly; http://genome.ucsc.edu).\(^{40}\)
ablated and less cold competitor was required to out-compete protein-DNA complexes (figure 4A). Addition of an antibody to CTCF blocked the formation of complex C (figure 4B), and specific enrichment of the region surrounding rs1876453 was three-fold higher in the presence of the major allele by ChIP-PCR (p=0.0178; figure 5A–D), suggesting that CTCF binds the region surrounding rs1876453 in an allele-specific manner. CTCF abundance was comparable between the two cell lines (figure 5E), and transcript abundance of CR1 relative to CR2 was higher in the cell line homozygous for the minor allele (p=0.023; figure 5F), consistent with the data from primary B cells.

DISCUSSION

The minor allele of rs1876453, located in intron 1 of CR2, was associated with decreased risk of lupus in three of four ancestral groups studied, with the strongest association when cases were stratified based on the presence of dsDNA autoantibodies. This allele altered the formation of multiple DNA-protein complexes, including one containing CTCF, which has been termed the master regulator of chromatin organisation. Furthermore, it was associated with increased B cell expression of CR1, suggesting long-range effects of rs1876453 on gene regulation and providing a plausible mechanism by which it may alter lupus susceptibility.

We previously identified the SLE-associated haplotype tagged by the minor allele of rs1876453 (H1 as shown in figure 1D) in EA,8 and show here that it is found also in AA and HS (figure 1D). Although it is difficult to prove that a SNP is causal based on association testing, our transancestral study, which analysed a high density of SNP markers and capitalised on the weak linkage disequilibrium around rs1876453 in AA (figure 1B), provides compelling evidence that rs1876453 rather than the previously identified SNPs (rs3813946, rs1048971, rs17615 and rs4308977) or SNPs tightly linked to rs1876453 in EA (rs17258955 and rs61821130 in intron 1, rs61735651 in exon 14 and rs7549152 in intron 15) best explains the association signal. Modest association signals of rs1876453 increased after stratification of subjects based on specific clinical manifestations of lupus, with the strongest signals in the case-control analysis when subjects were stratified based on the presence of dsDNA autoantibodies. Although increased signals were seen with other
Clinical manifestations, only dsDNA autoantibodies were significantly associated in both case-control and case-only analyses, suggesting that rs1876453 has a preferential effect on the production or presence of dsDNA autoantibodies. The mechanism underlying this dsDNA specificity is unclear, and characterising it will lead to a greater understanding of the etiopathogenesis of lupus, since dsDNA autoantibodies are highly specific for SLE (reviewed in refs 24, 25), and can predict onset and exacerbation of disease.26–28 Anti-dsDNA antibody levels in the serum of genotyped patients with SLE did not correlate with rs1876453 genotype (data not shown), but therapy and other disease-related factors may obscure an association. Future studies with longitudinal analyses of patients will address this more stringently.

rs1876453 is located in a transcriptional regulatory region defined by DNase hypersensitivity and containing histone marks associated with active enhancers (H3K27ac, H3K4me1 and H3K4me3) (Figure 3). We demonstrated that the minor allele at rs1876453 alters the formation of multiple protein-DNA complexes (Figure 4A) and that CTCF is a component of one of these complexes (Figures 4B and 5). CTCF, a highly conserved and ubiquitously expressed zinc finger protein, is a master regulator of genome spatial organisation that mediates chromatin loops within the genome.29 By binding to insulators and boundary elements, it demarcates chromatin into regulatory regions and blocks communication between promoters and enhancers to regulate gene expression. CTCF interacted with the region surrounding rs1876453 in both lymphoblastoid B cell lines and

Figure 5  CCCTC-binding factor (CTCF) interacts with CR2 intron 1 in vivo and demonstrates differential affinity for rs1876453 alleles. (A) Chromatin immunoprecipitation performed using an antibody specific for CTCF yielded allele-specific enrichment of the region surrounding rs1876453 from Epstein–Barr virus (EBV)-transformed B cells homozygous for the major or minor allele at rs1876453. The qPCR products were visualised by ethidium bromide staining of a 1.8% agarose gel and sized using a PCR marker (New England Biolabs). A non-specific IgG control (IgG) and a control without antibody (No Ab) were included to measure background enrichment. Decreasing amounts of enrichment were observed with serially diluted input samples (Lanes 5–8; 13–16). NTC, no template control. (B) A representative qPCR amplification plot. (C) The percentage enrichment at the CR2 promoter was determined by quantification against the standard curve. CTCF enrichment was normalised to the background enrichment generated by a non-specific IgG. (D) CTCF enrichment normalised to the minor allele at rs1876453. (E) CTCF transcript abundance for each homozygous cell line after normalisation to β-actin. (F) Transcript abundance of CR1 relative to CR2 in each homozygous cell line, with each transcript normalised to β-actin. Data shown are the mean±SEM for three independent experiments.
primary B cells, with peak occupancy directly over rs1876453 (figure 3), and our data demonstrate an allele-dependent effect on its binding in vivo (figure 5) with less binding of CTCF to the minor A allele, which is associated with higher CR1 expression and is protective, and more binding to the major G risk allele, which is associated with lower CR1 expression. These data suggest that CTCF may have a repressive effect on CR1 transcription when bound to this locus. Efforts to identify the other transcription factors differentially binding this region as a result of the allele at rs1876453 are currently underway.

CR1 expression was increased in B cells of subjects harbouring the minor allele at rs1876453 at both the mRNA and protein level (figure 2). Eukaryotic transcription is tightly controlled by regulatory elements that, though perhaps distant from their target genes, can be brought into close proximity with the general transcriptional machinery at transcription start sites through the formation of chromatin loops. Therefore, altered general transcriptional machinery at transcription start sites their target genes, can be brought into close proximity with the CR1 sequences. The minor allele at rs1876453 introduces a variation of mutant from wild type (68.75%) and is predicted to result of the allele at rs1876453 are currently underway.

Acknowledgements We thank the volunteers who participated in this study; Agnessa Gadiyia, Carissa Homme, Najjuk Jain and Lauren Kuhlman (University of Colorado School of Medicine, Aurora, Colorado, USA) for recruiting subjects, processing samples and technical assistance; Krista Bean (OMRF, Oklahoma City, Oklahoma, USA) for assistance with data analysis; Dr Henry Marsh and Celldec Therapeutics, Boston, Massachusetts, USA for providing the anti-CR1 monoclonal antibody 6B1H12; Dr Ellen Vitetta from University of Texas-Southwestern, Dallas, Texas, USA for providing the anti-C1q monoclonal antibody HD37; and Drs V. Michael Holers, Tasha E Fingerlin, and Richard A Spritz (University of Colorado School of Medicine, Aurora, Colorado, USA) for helpul discussions. The BIOLUPUS network is composed of Johanne Frostegård, MD, PhD (Huddinge, Sweden), Lennart Truedsson, MD, PhD (Lund, Sweden), Enrique de Ramón, MD PhD (Málaga, Spain), José M Sabio, MD, PhD (Granada, Spain), Maria F González-Escobar, PhD (Sevilla, Spain), Javier Martín, MD, PhD (Granada, Spain), Norberto Ortega-Centeno (Granada, Spain), José Luis Callejas MD (Granada, Spain), Julio Sánchez-Román, MD (Sevilla, Spain), Sandra D’Alfonso, PhD (Novara, Italy), Sergio Migliarese MD (Napoli, Italy), Gian-Domenico Sebastiani MD (Rome, Italy), Mauro Galeazzi MD (Siena, Italy), Torsten Witte, MD, PhD (Hannover, Germany), Bernard R Lauerweys, MD, PhD (Louvain, Belgium), Emoke Endreiffy, PhD (Szeged, Hungary), László Kovics, MD, PhD (Szeged, Hungary), Carlos Vasconcelos, MD, PhD (Porto, Portugal) and Berta Martins da Silva, PhD (Porto, Portugal). The members of GENES Network are Hugo R Scherbarth, Pilar C Marino, Estela L Motta, Susana Garron, Cristina Drenkard, Emilia Menso, Alberto Allievi, Guillermo A Tate, Jose L Presas, Simon A Palatnik, Marcelo Abdala, Mariela Bearzotti, Alejandro Alvarezes, Francisco Caeiro, Ana Bertoli, Sergio Paia, Susana Rovasano, Cesar G Graf, Estela Bertero, Cesar Caprunalo, Griselda Buchanan, Carolina Guillén, Sebastián Grimaudo, Jorge Mamm, Luis J Catoggio, Enrique R Soriano, Carlos D Santos, Cristina Pignone, Fernando A Ramos, Sandra M Navarro, Guillermo A Berboto, Marisa Jorfen, Elisa J Romero, Mercedes A Garcia, Juan C Marcos, Ana I Marcos, Carlos E Perandones, Alicia Eimon, Sanatorio Parque and Cristina G. Battagliotti in Argentina; Eduardo Acquedo and Mariano Cacho in Peru; Ignacio García de la Torre, Mario Cardiel Ríos, José Francisco Motezuma and Marco Maradiaga Ceceria in Mexico.
Basic and translational research

The UCSC Genes Track from ENCODE was produced at UCSC using a computational pipeline developed by Jim Kent, Chuck Sugnet, Melissa Cline and Mark Diekhans. It is based on data from NCBI RefSeq, UniProt (including TrEMBL and TrEMBL-NEW), CCDS and GenBank, as well as data from Rfam and the Todd Lopez track. The UCSC Genome Database project is hosted by the Genome Research Laboratory (broadinstitute.org) and is supported by grants from the National Institutes of Health (NIH) grants P30 DK57516, R01AR042460 (K.M.S.), P01AR043274 (K.M.S.), P30 GM103510 (J.A.J.), P60AR30692 (J.B.H.), R37AI024717 (J.B.H.), RC1AR058621 (M.E.A.R.), U01AI101934 (J.A.J and J.M.G.), P30AR055385 (E.E.B.), LAC, JCE, the Wold Lab at Caltech.

Contributors  

Make sure that you provide clear and concise explanations of the contributions of each person.

Funding  

The funding section should clearly state the sources of funding that were used to collect and analyze the data, and to prepare the report. This is an important section for transparency and to ensure that the research is not influenced by financial interests.

Competing interests  

This section should list any financial relationships or potential conflicts of interest that could be seen as influencing the results or conclusions of the study.

Provenance and peer review  

This section should provide information on the provenance of the manuscript and the peer review process. It is important to state whether the manuscript was peer-reviewed, and if so, by whom.

REFERENCES  

This section should list all the references cited in the text. Each reference should be cited at least once in the main text, and the format should follow the guidelines provided by the journal or conference.

Contact  

Name:  

Affiliation:  

E-mail:  

Website:  

All authors approved the final draft of the manuscript.


Preferential association of a functional variant in complement receptor 2 with antibodies to double-stranded DNA


Ann Rheum Dis published online September 1, 2014

Updated information and services can be found at:
http://ard.bmj.com/content/early/2015/05/15/annrheumdis-2014-205584

These include:

Supplementary Material
Supplementary material can be found at:
http://ard.bmj.com/content/suppl/2014/09/01/annrheumdis-2014-205584.DC1

References
This article cites 39 articles, 15 of which you can access for free at:
http://ard.bmj.com/content/early/2015/05/15/annrheumdis-2014-205584#BIBL

Open Access
This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See:
http://creativecommons.org/licenses/by/4.0/

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/