



Brief Report: Identification of BACH2 and RAD51B as Rheumatoid Arthritis Susceptibility Loci in a Meta-Analysis of Genome-Wide Data

Citation

McAllister, K., A. Yarwood, J. Bowes, G. Orozco, S. Viatte, D. Diogo, L. J. Hocking, et al. 2013. "Brief Report: Identification of BACH2 and RAD51B as Rheumatoid Arthritis Susceptibility Loci in a Meta-Analysis of Genome-Wide Data." *Arthritis and Rheumatism* 65 (12): 3058-3062. doi:10.1002/art.38183. <http://dx.doi.org/10.1002/art.38183>.

Published Version

doi:10.1002/art.38183

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:12406772>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

BRIEF REPORT

Identification of *BACH2* and *RAD51B* as Rheumatoid Arthritis Susceptibility Loci in a Meta-Analysis of Genome-Wide Data

Kate McAllister,¹ Annie Yarwood,¹ John Bowes,¹ Gisela Orozco,¹ Sebastian Viatte,¹ Dorothée Diogo,² Lynne J. Hocking,³ Sophia Steer,⁴ Paul Wordsworth,⁵ A. G. Wilson,⁶ Ann W. Morgan,⁷ UK Rheumatoid Arthritis Genetics Consortium, Rheumatoid Arthritis Consortium International, Joel M. Kremer,⁸ Dimitrios Pappas,⁹ Peter Gregersen,¹⁰ Lars Klareskog,¹¹ Robert Plenge,² Anne Barton,¹ Jeffrey Greenberg,¹² Jane Worthington,¹ and Stephen Eyre¹

Objective. A recent high-density fine-mapping (ImmunoChip) study of genetic associations in rheumatoid arthritis (RA) identified 14 risk loci with validated genome-wide significance, as well as a number of loci showing associations suggestive of significance ($P = 5 \times 10^{-5} < 5 \times 10^{-8}$), but these have yet to be replicated. The aim of this study was to determine whether these potentially significant loci are involved in the pathogenesis of RA, and to explore whether any of the loci are associated with a specific RA serotype.

Methods. A total of 16 single-nucleotide polymorphisms (SNPs) were selected for genotyping and association analyses in 2 independent validation cohorts, comprising 6,106 RA cases and 4,290 controls. A meta-analysis of the data from the original ImmunoChip

discovery cohort and from both validation cohorts was carried out, for a combined total of 17,581 RA cases and 20,160 controls. In addition, stratified analysis of patient subsets, defined according to their anti-cyclic citrullinated peptide (anti-CCP) antibody status, was performed.

Results. A significant association with RA risk ($P < 0.05$) was replicated for 6 of the SNPs assessed in the validation cohorts. All SNPs in the validation study had odds ratios (ORs) for RA susceptibility in the same direction as those in the ImmunoChip discovery study. One SNP, rs72928038, mapping to an intron of *BACH2*,

Genotyping of the UK Rheumatoid Arthritis Group samples was supported by Arthritis Research UK (grant 17552). Dr. Yarwood's work was supported by the Innovative Medicines Initiative (Joint Undertaking project BeTheCure; contract 115142-2). Dr. Orozco's work was supported by the Wellcome Trust (Research Career Development Fellowship 095684/Z/11/A). Dr. Viatte's work was supported by the Swiss Foundation for Medical-Biological Scholarships, which is managed by the Swiss National Science Foundation (SSMBS grant PASMP3_134380, funded by a donation from Novartis to the Foundation). Dr. Morgan's work was supported by the National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit. Drs. Barton and Worthington's work was supported by the NIHR Manchester Musculoskeletal Biomedical Research Unit.

¹Kate McAllister, BSc, Annie Yarwood, PhD, John Bowes, PhD, Gisela Orozco, PhD, Sebastian Viatte, MD, PhD, Anne Barton, PhD, FRCP, Jane Worthington, PhD, Stephen Eyre, PhD: University of Manchester, NIHR Manchester Musculoskeletal Biomedical Research Unit, and Manchester Academic Health Science Centre, Manchester, UK; ²Dorothée Diogo, PhD, Robert Plenge, MD, PhD: Harvard Medical School and Brigham and Women's Hospital, Boston, Massachusetts, and Broad Institute, Cambridge, Massachusetts; ³Lynne J. Hocking, PhD: University of Aberdeen, Aberdeen, UK; ⁴Sophia Steer, PhD, MRCP: Kings College Hospital National Health Service Foundation Trust, London, UK; ⁵Paul Wordsworth, MA, MB, BS, FRCP: NIHR Oxford Musculoskeletal Research Unit and Botnar

Research Centre, Nuffield Orthopaedic Centre, Oxford, UK; ⁶A. G. Wilson, MB, PhD, FRCP, DCH: University of Sheffield, Sheffield, UK; ⁷Ann W. Morgan, PhD, FRCP: University of Leeds, Leeds, UK; ⁸Joel M. Kremer, MD: Albany Medical College and the Center for Rheumatology, Albany, and Consortium of Rheumatology Researchers of North America, New York, New York; ⁹Dimitrios Pappas, MD, MPH: Columbia University, and Consortium of Rheumatology Researchers of North America, New York, New York; ¹⁰Peter Gregersen, MD: The Feinstein Institute for Medical Research and North Shore-LIJ Health System, Manhasset, New York; ¹¹Lars Klareskog, MD, PhD: Karolinska Institute, Stockholm, Sweden; ¹²Jeffrey Greenberg, MD, MPH: New York University and New York University Hospital for Joint Diseases, New York, New York.

Dr. Kremer has received consulting fees, speaking fees, and/or honoraria from AbbVie, Bristol-Myers Squibb, Genentech, Lilly, and Pfizer (less than \$10,000 each) and is an employee of the Consortium of Rheumatology Researchers of North America (CORRONA). Dr. Pappas has received consulting fees, speaking fees, and/or honoraria from Novartis (less than \$10,000) and is an employee of CORRONA. Dr. Barton has received consulting fees, speaking fees, and/or honoraria from Eli Lilly and Pfizer (less than \$10,000 each).

Address correspondence to Stephen Eyre, PhD, University of Manchester, Arthritis Research UK Epidemiology Unit, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, Stopford Building, Oxford Road, Manchester M13 9PT, UK. E-mail: steve.eyre@manchester.ac.uk.

Submitted for publication May 29, 2013; accepted in revised form August 29, 2013.

achieved genome-wide significance in the meta-analysis ($P = 1.2 \times 10^{-8}$, OR 1.12), and a second SNP, rs911263, mapping to an intron of *RAD51B*, was significantly associated in the anti-CCP-positive RA subgroup ($P = 4 \times 10^{-8}$, OR 0.89), confirming that both are RA susceptibility loci.

Conclusion. This study provides robust evidence for an association of RA susceptibility with genes involved in B cell differentiation (*BACH2*) and DNA repair (*RAD51B*). The finding that the *RAD51B* gene exhibited different associations based on serologic subtype adds to the expanding knowledge base in defining subgroups of RA.

Rheumatoid arthritis (RA) is a complex, chronic autoimmune disease that affects ~1% of the adult population worldwide (1). In addition to inflammation of the synovial joints, RA is characterized by systemic inflammation and the presence of serum autoantibodies against citrullinated peptides (anti-citrullinated protein antibodies [ACPAs]), as defined by positive findings on the anti-cyclic citrullinated peptide (anti-CCP) antibody test (2). Genome-wide association studies have been successful in determining many loci associated with complex diseases, including RA (3). The utility of susceptibility variants within single genetic loci, in isolation, is likely to be limited, with evidence emerging that linking multiple associated genes in pathways will lead to the better understanding of differing disease mechanisms (4,5). To achieve robust pathway analysis, a comprehensive list of associated loci must be defined. Currently, 46 loci have been confirmed to be associated with RA susceptibility in Caucasians, at accepted levels of genome-wide significance ($P < 5 \times 10^{-8}$), including 14 loci newly identified in a recent high-density, fine-mapping (ImmunoChip) study of RA (6).

In studies of inflammatory bowel disease (IBD), the findings have become much more informative, implicating risk pathways that have not been previously recognized as important to this disease, and thus increasing the number of susceptibility markers from 92 to 163. The increased number of susceptibility loci for IBD has also enabled much more informative investigation of disease overlap. Genetic studies of RA to date, albeit successful, have not yet delivered validated evidence of novel pathways. Moreover, disease overlap studies have been limited, thus emphasizing the continuing need for discovery of disease susceptibility markers in RA.

RA is currently divided into 2 groups based on serologic subtypes, which are defined according to the presence or absence of anti-CCP antibodies, although it is still unclear whether there are biologic pathways that

are common or distinct to each group (2). Determining the genetic predisposition to each serologic subtype has the potential to better define the mechanism underlying each form of disease, enabling progress toward more focused clinical management.

The most recent study aimed at identifying RA susceptibility loci (6) used a custom Illumina array (ImmunoChip), designed to interrogate 196,524 single-nucleotide polymorphisms (SNPs) for 186 loci that have been previously shown to be associated with a number of autoimmune diseases. The study by Eyre and colleagues was the first to be powered to analyze the subgroups of seronegative RA and seropositive RA separately. Genotyping in 11,475 RA cases and 15,870 controls provided evidence for 14 novel SNPs that achieved genome-wide significance, with a further 16 SNPs putatively associated with RA in a second tier of significance ($P = 5 \times 10^{-5} < 5 \times 10^{-8}$), either in an unstratified analysis or in stratified analyses of the anti-CCP antibody subgroups. We therefore tested the 16 SNPs for which there was suggestive evidence of association with RA risk in 2 independent cohorts, comprising 6,106 RA cases and 4,290 controls, and performed a meta-analysis in which we combined our results with existing data to enhance the power to identify associations in the whole data set and in subgroups stratified by anti-CCP antibody status.

PATIENTS AND METHODS

ImmunoChip discovery cohort. The characteristics of the patients in the original ImmunoChip discovery cohort (from the UK Rheumatoid Arthritis Consortium International [RACI]), as well as the genotyping and quality control procedures used for the ImmunoChip analysis, have been previously described (6). Individual-level data were available for each participant, and these data were used in the present analysis.

Validation cohorts. UK cohort. The UK Rheumatoid Arthritis Group (UKRAG) cohort of patients with RA was recruited from 6 centers across the UK (Table 1), as previously described (7). Genotyping was performed using the Sequenom platform.

US cohort. Samples from patients with RA in the US Consortium of Rheumatology Researchers of North America (CORRONA) collection and Informatics for Integrating Biology and the Bedside (I2B2) program were also tested using the ImmunoChip custom genotyping SNP array (Table 1). Principal components analysis was performed using the EigenSoft program (version 29) with HapMap phase III samples, to exclude individuals of non-European ancestry. To check relatedness between the CORRONA/I2B2 samples and the RACI ImmunoChip samples, estimates of identity-by-descent and identity-by-state allele-sharing proportions were performed in Plink, using a set of SNPs selected for high quality (missingness $P < 0.002$, minor allele frequency [MAF] > 0.1) and pruned for LD ($r^2 < 0.2$; samples were removed if the PI_HAT value was > 0.2 [$n = 1$ sample per pair]).

Table 1. Distribution of rheumatoid arthritis (RA) cases and controls in the ImmunoChip discovery cohort and the 2 validation cohorts*

Collection	RA cases				Controls	
	All	Female, %	Anti-CCP positive	Anti-CCP negative	All	Female, %
ImmunoChip						
US	2,536	75	1,803	593	2,134	65
Swedish EIRA	2,762	70	1,762	987	1,940	73
Swedish Umea	852	70	524	242	963	69
Dutch	648	66	330	301	2,004	42
Spanish	807	74	397	216	399	65
UK	3,870	74	2,406	1,000	8,430	53
Total	11,475	72	7,222	3,339	15,870	61
Validation						
CORRONA/I2B2 (US)	2,206	78	1,410	446	1,863	75
UKRAG (UK)	3,900	73	536	274	2,427	74
Total	6,106	76	1,946	720	4,290	75
Combined	17,581	73	9,168	4,059	20,160	65

* Except where indicated otherwise, values are the number of participants. Anti-CCP = anti-cyclic citrullinated peptide; EIRA = Epidemiological Investigation of Rheumatoid Arthritis; CORRONA = Consortium of Rheumatology Researchers of North America; I2B2 = Informatics for Integrating Biology and the Bedside; UKRAG = UK Rheumatoid Arthritis Group.

SNP selection/prioritization. Sixteen SNPs were selected for genotyping, based on suggestive evidence of a significant association with RA susceptibility ($P > 5 \times 10^{-8}$ to $P < 5 \times 10^{-5}$) in either the overall ImmunoChip genotyping analysis or in subgroups defined by the presence or absence of anti-CCP antibodies. Seven SNPs were selected from the full ImmunoChip cohort, 5 from the anti-CCP-positive subgroup, and 4 from the anti-CCP-negative subgroup.

Statistical analysis. Stage one: validation analysis. SNPs were included in the validation analysis if they passed quality control in both the UKRAG and US cohorts. Samples or SNPs with a call rate of $<98\%$ were removed. In addition, SNPs that did not conform to Hardy-Weinberg equilibrium ($P < 5 \times 10^{-7}$) as well as SNPs with differential missingness ($P > 0.01$) or an MAF <0.01 were also removed. After applying all of the quality control filters, 4 SNPs (from 944 RA cases [20%] and 244 controls [9%]) failed quality control.

We tested for the association of each SNP with RA in each study independently, using logistic regression under an additive model, and then combined the results in a fixed-effects meta-analysis with inverse variance weighting (performed in Plink). The top 10 principal components were incorporated as covariates in the logistic regression analysis of the US CORRONA data set, to correct for population stratification. P values less than 0.05 were considered as evidence of a significant association in the validation cohorts.

Stage two: meta-analysis of ImmunoChip and validation data. The association of each SNP with RA in the original ImmunoChip analysis and in the validation analysis was tested in a fixed-effects meta-analysis with inverse variance weighting (performed in Plink). A $P_{\text{meta-analysis}}$ threshold of less than 5×10^{-8} was considered significant, i.e., indicative of the discovery of a novel RA risk gene. All P values reported are 2-tailed. Heterogeneity between studies was assessed using Cochran's Q statistic test and the I^2 test. A P value for heterogeneity of less than 0.05 is suggestive of heterogeneity. The I^2 test estimates the extent of heterogeneity and takes values between 0% and 100%. I^2 values of 0–25% indicate low heterogeneity,

25–50% moderate heterogeneity, 50–75% high heterogeneity, and 75–100% extreme heterogeneity (8).

Anti-CCP subset analysis. In separate analyses of the data from the validation study and the data from the meta-analysis, we compared associations of all SNPs between controls and either anti-CCP-positive or anti-CCP-negative RA patients. In the validation study, there were 1,946 anti-CCP-positive RA patients, 720 anti-CCP-negative RA patients, and 4,290 controls. For the meta-analysis, there were 9,169 anti-CCP-positive RA patients, 4,059 anti-CCP-negative RA patients, and 20,160 controls.

Power calculations. All power calculations were undertaken using Quanto. The model assumed was a log additive model, and population baseline risk estimate was 1%.

RESULTS

In the validation analysis, 6 loci showed an association with RA susceptibility that reached significance ($P < 0.05$): *COG6*, *PVT1*, *PTPN2*, *TNFSF4*, *RAD51B*, and *BACH2* (Table 2). In the meta-analysis, 2 SNPs achieved genome-wide significance levels ($P_{\text{meta-analysis}} < 5 \times 10^{-8}$): *BACH2* in the full meta-analysis (Figure 1A), and *RAD51B* in the anti-CCP-positive subgroup (Figure 1B). These findings in the meta-analysis indicate that *BACH2* and *RAD51B* are novel validated RA risk alleles. For the remaining SNPs, all effect sizes were in the same direction as in the original ImmunoChip analysis.

In the meta-analysis, 1 SNP, rs7535176, showed evidence of high heterogeneity between studies ($I^2 = 57\%$), and was therefore removed from further analysis. In power calculations, the average power to detect the likelihood of an association with RA at an OR of 1.12 ($P = 5 \times 10^{-8}$) was 81% in the overall validation study,

Table 2. Associations of single-nucleotide polymorphisms (SNPs) with rheumatoid arthritis susceptibility in the ImmunoChip discovery cohort, validation cohorts, and meta-analysis*

Chr	SNP	Position (hg19)	Gene	Minor allele	ImmunoChip		Validation		Meta-analysis					
					Study <i>P</i>	Study OR	Study <i>P</i> †	Study OR	Study <i>P</i>	Study OR	Anti-CCP+		Anti-CCP-	
											<i>P</i>	OR	<i>P</i>	OR
1	rs61828284‡	173,299,743	TNFSF4	A	5.5×10^{-6}	0.83	0.01	0.81	1.5×10^{-5}	0.88	2.1×10^{-7}	0.82	0.48	0.97
2	rs888427§	172,368,120	CYBRD1	A	5.2×10^{-5}	1.12	0.76	1.02	0.00021	1.06	0.017	1.05	1.6×10^{-4}	1.1
3	rs1875463¶	143,165,502	SLC9A9	A	1.5×10^{-6}	0.90	0.69	0.99	2.0×10^{-5}	0.92	1.2×10^{-4}	0.92	8.7×10^{-5}	0.89
6	rs72928038¶	90,976,768	BACH2	A	8.2×10^{-7}	1.13	0.004	1.11	1.1×10^{-8}	1.12	8.5×10^{-6}	1.12	2.5×10^{-4}	1.13
7	rs59472144¶	37,369,908	ELMO1	A	3.5×10^{-6}	1.12	0.78	1.01	5.1×10^{-5}	1.09	9.3×10^{-5}	1.10	0.04	1.07
8	rs6651252¶	129,567,181	PVT1	G	2.0×10^{-5}	0.89	0.004	0.88	2.4×10^{-7}	0.89	5.3×10^{-5}	0.89	0.001	0.88
9	rs7857530§	102,747,941	TXNDC4	G	2.5×10^{-5}	0.89	0.94	1.00	0.022	0.96	0.40	0.98	1.7×10^{-4}	0.91
13	rs7993214‡	40,350,912	COG6	A	4.6×10^{-6}	0.90	0.05	0.92	9.1×10^{-7}	0.92	6.0×10^{-7}	0.9	0.15	0.96
13	rs17230016¶	82,338,338	SPRY2	G	9.6×10^{-6}	0.88	0.56	0.97	4.7×10^{-5}	0.91	0.003	0.92	0.001	0.88
14	rs911263‡	68,753,593	RAD51B	G	2.3×10^{-6}	0.89	0.005	0.88	4.0×10^{-5}	0.93	4.0×10^{-8}	0.89	0.44	0.98
17	rs8074003¶	76,364,493	SOCS3	A	1.4×10^{-5}	0.81	0.44	0.95	9.2×10^{-5}	0.86	1.9×10^{-3}	0.86	9.2×10^{-4}	0.80
18	rs62097857¶	12,857,758	PTPN2	A	4.5×10^{-6}	1.22	0.01	1.20	1.4×10^{-7}	1.22	3.9×10^{-6}	1.23	6.2×10^{-5}	1.26

* Association analyses of SNPs were performed using case and control data from the original ImmunoChip discovery cohort, the validation cohorts (UK and US), and the meta-analysis (ImmunoChip and validation cohorts). OR = odds ratio.

† Validation *P* value is presented for each SNP in the same stratification as that in the discovery study.

‡ Anti-cyclic citrullinated peptide (anti-CCP)-positive RA subgroup.

§ Anti-CCP-negative RA subgroup.

¶ Full RA group.

41% in the anti-CCP-positive RA subset, and 6% in the anti-CCP-negative RA subset, based on an average MAF of 0.19.

DISCUSSION

Of the 12 loci assessed for an association with susceptibility to RA in the validation study, 2 reached accepted genome-wide significance thresholds either in the RA group as a whole or in the subset with anti-CCP antibodies, indicating that these 2 loci (*BACH2*, *RAD51B*) represent novel RA susceptibility loci. A further 4 SNPs showed evidence for validation of an association at a significance level of $P < 0.05$ (*COG6*, *PVT1*, *PTPN2*, and *TNFSF4*). All of the loci assessed showed effect sizes in the same direction as those in the original ImmunoChip discovery study. Of the newly confirmed RA susceptibility SNPs, rs72928038 maps to an intron of *BACH2* (6q15), which has previously been

associated with type 1 diabetes (9), Crohn's disease (10), and celiac disease (11). *BACH2* encodes BTB and CNC homology 1, basic leucine zipper transcription factor 2, a B cell-specific transcription factor that has been shown to regulate the *BLIMP1* gene (also known as *PRDM1*, a validated RA susceptibility locus) in mice, which in turn influences plasma cell differentiation and promotes the antibody class switch (12). Rituximab, an effective RA treatment, inhibits CD20+ B cells, a cell subtype in which *BACH2* is highly expressed.

The second novel SNP associated with RA, rs911263, was found to be significantly associated with RA risk in the anti-CCP-positive subgroup ($P_{\text{anti-CCP+ meta-analysis}} = 4 \times 10^{-8}$) but not in the anti-CCP-negative subgroup ($P_{\text{anti-CCP- meta-analysis}} = 0.45$). This SNP maps to an intron in *RAD51B* (14q24), which has recently been found to be associated with primary biliary cirrhosis (13). *RAD51B* is a gene in-

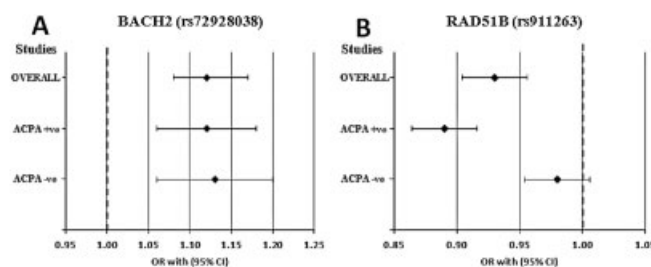


Figure 1. Forest plots of the association of rheumatoid arthritis (RA) susceptibility with the 2 single-nucleotide polymorphisms showing genome-wide significance, *BACH2* (A) and *RAD51B* (B), in the meta-analyses. Results are presented as the odds ratio (OR) with 95% confidence interval (95% CI) for the overall meta-analysis and for the subgroups of anti-citrullinated protein antibody-positive (ACPA+ve) RA and anti-ACPA-negative (ACPA-ve) RA (determined by anti-cyclic citrullinated peptide test). The meta-analyses are based on a fixed-effects model.

volved in the homologous recombination repair pathway of double-stranded DNA breaks. It has been demonstrated that anti-CCP-positive and anti-CCP-negative disease have differing allelic associations at the HLA locus, and identification of an association of *RAD51B* with anti-CCP-positive RA may add to the list of genes that differentiate between the disease subgroups, although this could not be formally proven in the present study because of the reduced power in the anti-CCP-negative cohort.

An association of 4 loci (*COG6*, *PVT1*, *PTPN2*, and *TNFSF4*) with RA risk was replicated in the validation study, at $P < 0.05$, but none reached genome-wide significance in the overall analysis. Given that these loci are associated with multiple autoimmune diseases, there is strong a priori evidence that these are likely to be RA susceptibility loci. Indeed, an association of RA risk with *PTPN2* has now reached genome-wide significance levels, as demonstrated in an expanded, independent analysis in a European cohort (14).

The number of SNPs with suggestive evidence from the ImmunoChip analysis suggests that there are many additional undiscovered risk alleles with modest effect sizes that have yet to be formally confirmed at genome-wide significance levels of association. Power calculations showed the necessity of larger sample sizes to detect the more modest effect sizes detected in this second tier of significance, indicating that there is still great value in increasing study size to identify novel susceptibility loci. Indeed, combined sample sizes of more than 75,000 cases and controls were required to identify the 163 loci now confirmed to be associated with IBD. Sample size and power therefore remain a limitation in our current study, in particular with respect to the lack of data on the ACPA status of patients in the validation cohorts. A further limitation in this study was the targeted SNP genotyping, chosen as a cost-effective strategy to confirm putative associations. By adopting this strategy, our capacity to fine map the newly identified loci was restricted to data generated in the ImmunoChip experiment.

In conclusion, we obtained convincing evidence of 2 new RA susceptibility loci, *BACH2* and *RAD51B*, in a combined meta-analysis of 17,581 cases and 20,160 controls, bringing the total number of novel RA susceptibility loci identified in Caucasians to 48. These findings implicate 2 distinct biologic pathways, those of B cell differentiation and DNA repair, in serologic subtypes of RA. Identification of all genetic variants that predispose to RA will increase our understanding of the molecular mechanisms involved and better annotate biologic pathway analyses.

ACKNOWLEDGMENTS

We would like to thank Edward Flynn for preparing and genotyping the UKRAG samples. We thank the RACI, the UKRAG, the CORRONA, and the I2B2 project groups for making genetic data available for this meta-analysis.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Eyre had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. McAllister, Wordsworth, Morgan, Gregersen, Klareskog, Plenge, Barton, Worthington, Eyre.

Acquisition of data. McAllister, Diogo, Hocking, Steer, Wordsworth, Wilson, Kremer, Pappas, Gregersen, Plenge, Greenberg, Eyre.

Analysis and interpretation of data. McAllister, Yarwood, Bowes, Orozco, Viatte, Klareskog, Plenge, Barton, Eyre.

REFERENCES

1. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res* 2002;4 Suppl 3:S265–72.
2. Van der Helm-van Mil AH, Huizinga TW. Advances in the genetics of rheumatoid arthritis point to subclassification into distinct disease subsets. *Arthritis Res Ther* 2008;10:205.
3. Stranger BE, Stahl EA, Raj T. Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 2011;187:367–83.
4. Bakir-Gungor B, Sezerman OU. A new methodology to associate SNPs with human diseases according to their pathway related context. *PLoS One* 2011;6:e26277.
5. Song GG, Bae SC, Lee YH. Pathway analysis of genome-wide association studies on rheumatoid arthritis. *Clin Exp Rheumatol* 2013;31:566–74.
6. 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.
7. Orozco G, Hinks A, Eyre S, Ke X, Gibbons LJ, Bowes J, et al. Combined effects of three independent SNPs greatly increase the risk estimate for RA at 6q23 [published erratum appears in *Hum Mol Genet* 2010;19:4544]. *Hum Mol Genet* 2009;18:2693–9.
8. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
9. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al, and the Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;41:703–7.
10. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40:955–62.
11. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet* 2008;40:395–402.
12. Muto A, Ochiai K, Kimura Y, Itoh-Nakadai A, Calame KL, Ikebe D, Tashiro S, et al. Bach2 represses plasma cell gene regulatory network in B cells to promote antibody class switch. *EMBO J* 2010;29:4048–61.
13. Liu JZ, Almarri MA, Gaffney DJ, Mells GF, Jostins L, Cordell HJ, et al. Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet* 2012;44:1137–41.
14. Cobb J, Plant P, Flynn E, et al. Identification of the tyrosine-protein phosphatase non-receptor type 2 as a rheumatoid arthritis susceptibility locus in Europeans. *PLoS One* 2013;8:e66456.