HHS Public Access

Author manuscript

Arthritis Rheum. Author manuscript; available in PMC 2014 April 15.

Published in final edited form as:

Arthritis Rheum. 2013 December ; 65(12): 3014–3016. doi:10.1002/art.38111.

Ethnic-Specific Genetic Analyses in Rheumatoid Arthritis: Incremental Gains but Valuable Contributions to the Big Picture

Maria I. Danila, MD, MSc[1], **Richard J. Reynolds, PhD**[1], **Hemant K. Tiwari, PhD**[2], and **S. Louis Bridges Jr., MD, PhD**[1,*]

[1]UAB Division of Clinical Immunology and Rheumatology; Department of Medicine; University of Alabama at Birmingham

[2]UAB Section on Statistical Genetics, Department of Biostatistics; University of Alabama at Birmingham

Rheumatoid arthritis (RA) is a common autoimmune disease characterized by chronic inflammation of the synovial membrane, which can lead to joint damage and a variety of other clinical manifestations (1). The number of genes reported to be associated with susceptibility to RA continues to grow, now totaling ~60 (2). Although RA affects individuals from diverse ethnic backgrounds, genome-wide association (GWA) studies of RA have been performed only in subjects of European or east Asian (particularly Japanese) ancestry (2). In addition, as is the case with most complex diseases, a substantial proportion of "missing heritability" remains to be identified in RA (3). The difference between RA heritability estimates from family-based studies and the variance explained by SNPs from GWA studies has not been investigated in non-European populations.

In this issue of *Arthritis & Rheumatism*, Negi et al. (4) report a novel association of the gene *ARL15* (ADP-ribosylation factor-like 15) from the first GWA study of RA in North Indians. Although India's population of 1.2 billion has thousands of ethnic groups characterized by differences in language, customs, and religion, two ancient populations are ancestral to most present-day Indians: ancestral North Indians (ANI) and ancestral South Indians (5). Of the two ancestral groups, ANI are genetically closer to Middle Easterners, Central Asians, and Europeans. ANI ancestry ranges from 39-71% in most Indian groups and the degree of ANI ancestry is reportedly higher in traditionally upper caste and Indo-European speakers (5). There is a clear sub-structure in the ANI population in the study by Negi et al., with three clusters identified using multidimensional scaling (see Figure 1b), but single marker tests were adjusted for the resulting genomic inflation (Figure 2).

The association of *ARL15* with RA reported by Negi et al. underscores the importance of the search for genetic underpinnings of RA and other immune-mediated diseases among

*Corresponding Author: S. Louis Bridges, Jr., MD, PhD *Marguerite Jones Harbert-Gene V. Ball, MD* Professor of Medicine Director, Division of Clinical Immunology and Rheumatology University of Alabama at Birmingham 1720 Second Avenue South, SHEL 178 Birmingham, Alabama 35294-2182 Telephone 205 934-4616 LBridges@uab.edu.

Financial Support Information:

None of the authors have received financial support or other benefits from commercial sources for the work reported on in this manuscript, or any other financial interests which could create a potential conflict of interest or the appearance of a conflict of interest with regard to the work.

different ethnicities. There are several reasons to study RA genetics among different ethnicities – to look for different, ethnic-specific risk factors (as was the focus of Negi et al.); to look for risk factors shared across multiple ethnicities; and to leverage divergent linkage disequilibrium patterns to refine the distance between the tag polymorphism and causal variant.

The presence of a significant number of ethnic-specific risk alleles for RA might lead to the development of ethnic-specific diagnostic and therapeutic tools. However, associations that are limited to one or more, but not all ethnicities, such as that of *PTPN22* with RA in Europeans (6), but not in African-Americans (7), appear to be the exception rather than the rule. The paucity of ethnic-specific risk alleles points to the potential value of large-scale genetic association studies across RA patients of different ethnicities, namely enabling the conduct of trans-ethnic analyses. By analyzing large groups of RA patients of different ethnicities together, there is the capability of performing fine mapping of causal variants and increased statistical power to identify new genetic associations (8). Such an approach has been used to identify new loci associated with RA (9) as well as broader phenotypes such as serum protein levels (10).

The findings of Negi et al. underscore the importance of adipocytokine pathways in RA. As reviewed by Müller-Ladner et al. (11), adipocytokines such as adiponectin are produced by synovial fibroblasts; are present in substantial amounts in the serum and joints of patients with inflammatory joint diseases; and can up-regulate pro-inflammatory pathways and RANKL-dependent osteoclast activation. In addition, serum adiponectin levels are associated with radiographic damage in RA (12). While the association between RA and cardiovascular (CV) disease is known, there are conflicting data on whether circulating adiponectin levels are associated with CV disease. Negi et al. found that the allele (risk) C of the rs255758 influences adiponectin levels in RA patients. There is a link between RA and *ARL15*, and between *ARL15* and adiponectin; it is interesting to speculate that *ARL15* variants may contribute to the CV disease phenotype in RA through the adiponectin pathway.

In addition to conventional association analysis, the investigators used a machine learning approach, namely support vector machines (SVM), to identify novel susceptibility genetic loci for RA. The SVM approach is a well-developed machine-learning technique used in computer science for pattern recognition which utilizes a set of training data in order to learn how to classify objects (13). As applied to the problem of disease risk prediction, the SVM approach attempts to find an optimal set of genetic variants that can accurately classify a set of data between cases and controls. This is a different question than asking if individual markers explain variation in case-control status. Therefore it is not surprising that 6 additional loci, not including *ARL15*, were found to be most informative for RA risk prediction in ANI. SVMs have been used to predict genetic susceptibility based on genotype data in complex diseases such as type II diabetes (14). Another interesting application of the SVM methodology would be to preselect known RA risk variants from other or all ethnicities and test their predictive ability in the ANI sample set. This analysis would then help to answer the question of the extent to which ANI and other ethnicities share RA risk factors.

Use of different statistical analytic approaches may lead to additional novel genetic associations of RA in ANI. For example, in the replication phase of the current study, at least 20 SNPs flagging the SNP with lowest p-values could have been used rather than 1 SNP per region to provide more detailed data. In addition, the threshold used for significance for multiple testing (5×10^{-8}) could have been lowered if the spectral decomposition method had been used to find the effective number of SNPs for association. Finally, during the quality control analysis, using a p value of 1×10^{-7} as a threshold to exclude SNPs not in Hardy-Weinberg equilibrium (HWE) may be too stringent; all SNPs with deviation from HWE could have been included in the analysis with subsequent use of zoom plots to investigate clustering of genotypes and ensure correct calling by the algorithm.

In summary, the finding of an association of *ARL15* with RA in ancestral North Indians illustrates the utility of such genetic studies in populations of different ethnicity. By performing larger analyses, including trans-ethnic meta-analyses, the totality of the genetic contributions to this disease may finally start to become apparent. These studies have important implications on the full delineation of the multiple pathways involved in RA. Genetic profiles may allow identification of subsets of patients with different dominant pathologic pathways. This stratification may in turn lead to tailored approaches to early diagnosis or targeted therapy. For example, subjects with variants in *CTLA4*, which is associated with RA, may be more (or less) likely to respond to abatacept (CTLA4Ig). Through incremental gains, the big picture of RA becomes clearer and hopefully will lead improved outcomes for this chronic disease.

Acknowledgements

MID is supported by NIH K23 AR062100; RJR is supported by NIH K01 AR060848; SLB and HKT are supported by NIH R01 AR057202; SLB is supported by NIH R01 AR062376.

References

1. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. The New England journal of medicine. Dec 8. 2011; 365(23):2205–19. PubMed PMID: 22150039. [PubMed: 22150039]
2. Viatte S, Plant D, Raychaudhuri S. Genetics and epigenetics of rheumatoid arthritis. Nature reviews Rheumatology. Mar. 2013; 9(3):141–53. PubMed PMID: 23381558. Epub 2013/02/06. eng.
3. Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, Do R, Voight BF, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet. May. 2012; 44(5):483–9. PubMed PMID: 22446960. Pubmed Central PMCID: 3454878. [PubMed: 22446960]
4. Negi SJ,G, Senapati S, Prasad P, Gupta A, Singh S, Kashyap S, Kumar A, Kumar U, Gupta R, Kaur S, Agrawal S, Aggarwal A, Ott J, Jain S, Juyal RC, Thelma BK. A genome-wide association study reveals ARL15, a novel non-HLA susceptibility gene for Rheumatoid arthritis in north Indians. Arthritis & Rheumatism.
5. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature. Sep 24. 2009; 461(7263):489–94. PubMed PMID: 19779445. Pubmed Central PMCID: 2842210. [PubMed: 19779445]
6. Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. AmJHumGenet. 2004; 75(2):330–7. 8/2004.
7. Hughes LB, Reynolds RJ, Brown EE, Kelley JM, Thomson B, Conn DL, et al. Most common single-nucleotide polymorphisms associated with rheumatoid arthritis in persons of European

ancestry confer risk of rheumatoid arthritis in African Americans. Arthritis Rheum. 2010; 62(12): 3547–53. 12/2010. [PubMed: 21120996]

8. Morris AP. Transethnic meta-analysis of genomewide association studies. Genet Epidemiol. Dec. 2011; 35(8):809–22. PubMed PMID: 22125221. Pubmed Central PMCID: PMC3460225. Epub 2011/11/30. eng. [PubMed: 22125221]
9. Kurreeman FA, Stahl EA, Okada Y, Liao K, Diogo D, Raychaudhuri S, et al. Use of a multiethnic approach to identify rheumatoid- arthritis-susceptibility loci, 1p36 and 17q12. AmJHumGenet. 2012; 90(3):524–32. 3/9/2012.
10. Franceschini N, van Rooij FJ, Prins BP, Feitosa MF, Karakas M, Eckfeldt JH, et al. Discovery and fine mapping of serum protein loci through transethnic meta-analysis. Am J Hum Genet. Oct 5. 2012; 91(4):744–53. PubMed PMID: 23022100. Pubmed Central PMCID: PMC3484648. Epub 2012/10/02. eng. [PubMed: 23022100]
11. Muller-Ladner U, Neumann E. Rheumatoid arthritis: the multifaceted role of adiponectin in inflammatory joint disease. Nature reviews Rheumatology. Dec. 2009; 5(12):659–60. PubMed PMID: 19946293. Epub 2009/12/01. eng.
12. Giles JT, van der Heijde DM, Bathon JM. Association of circulating adiponectin levels with progression of radiographic joint destruction in rheumatoid arthritis. Ann Rheum Dis. Sep. 2011; 70(9):1562–8. PubMed PMID: 21571734. Pubmed Central PMCID: 3543946. [PubMed: 21571734]
13. Noble WS. What is a support vector machine? Nat Biotechnol. Dec. 2006; 24(12):1565–7. PubMed PMID: 17160063. [PubMed: 17160063]
14. Ban HJ, Heo JY, Oh KS, Park KJ. Identification of type 2 diabetes-associated combination of SNPs using support vector machine. BMC genetics. 2010; 11:26. PubMed PMID: 20416077. Pubmed Central PMCID: PMC2875201. Epub 2010/04/27.eng. [PubMed: 20416077]