Association between copy number variations of HLA-DQA1 and ankylosing spondylitis in the Chinese Han population

J Wang, Y Yang, S Guo, Y Chen, C Yang, H Ji, X Song, F Zhang, Z Jiang, Y Ma, Y Li, A Du, L Jin, JD Reveille, H Zou

INTRODUCTION

Ankylosing spondylitis (AS) is an immune-mediated complex disease with inflammation of the spine and extra-skeletal sites such as the peripheral joints, entheses and eyes. Although its etiopathogenesis is not fully understood, genetic predisposition as the peripheral joints, entheses and eyes. Although its immunodeficiency virus acquisition and progression and Alzheimer’s disease. Like many other human complex diseases, sequence variations of genetic risk may explain only a proportion of trait heritability to AS. The aim of the studies herein was to explore specific CNVs that potentially contribute to genetic susceptibility to AS.

RESULTS

Characteristics of study subjects

The median and interquartile range (IQR) of age of all AS cases and controls were 32 years (IQR = 16) and 60.5 years (IQR = 17), respectively. In the first and second cohorts, they were 33 (IQR = 18) and 32 (IQR = 16) in cases, and 64 (IQR = 11) and 58 (IQR = 17) in controls, respectively.

Male to female ratios were 3.6/1 (77.3 vs 19.5%) in cases and 0.81/1 (44.7 vs 55.2%) in controls. In the first and second cohorts, there were 137/43 and 243/61 male/female cases, and 126/99 and 228/338 male/female controls, respectively.

The HLA-B27 allele was observed in 90.6% of AS cases. There is no significant association between HLA-B27 positivity and age, gender or CNVs of the tested genomic region.

Analysis of CNV and AS

The CNV studies using comparative genomic hybridization (CGH) microarrays showed aberrant copy number (CN) at 6p21.32 in the majority of patients. In particular, aberrant CN of a 42-kb fragment in the region containing HLA-DRB5 was seen in four of five AS patients, and that of a 12-kb fragment in the region containing HLA-DQA1 was seen in three of five AS patients. The CGH results for the CNV of the DRB5 and DQA1 regions are shown in Figure 1.

Specific probes were designed for validation assays with the AccuCopy technology. Two cohorts were examined in validation.
In considering the heterogeneity of gender in the study population, we further tested the associations between the CN and AS using a mixed model for logistic regression analysis and a linear mixed model with gender as covariates. The results showed that AS association with the CN of the HLA-DQA1 allele became stronger (logistic regression analysis: \( P = 6.99 \times 10^{-6}, \) OR = 8.72, 95% CI = 3.97–19.17; linear mixed model: \( P < 10^{-8}, \beta = 1.5, \) 95% CI = 1.3–1.74). The association between AS and the CN of the HLA-DRB5 allele remained insignificant. In addition, we performed analysis on male and female subgroups for the CN of HLA-DQA1’s association with AS. The results showed that in male and female subgroups the achieved \( P \) values were 4.75 \( \times 10^{-5} \) (OR = 6.55, 95% CI = 2.65–16.19) and 1.88 \( \times 10^{-4} \) (OR = 19.77, 95% CI = 4.13–94.6), respectively.

DISCUSSION

In studies of genetic predisposition to AS, sequence variations have been extensively investigated, and from which multiple genetic susceptibility loci for AS have been identified.\(^9\) However, sequence variations may not fully cover the genetic contribution to AS. CNV is another significant source of human genetic variation and the attributed cause for disease and population diversity. Recent studies have indicated that CNVs of the HLA region are abundant in human genome.\(^12,13\) Our studies demonstrated that aberrant CN of the HLA-DQA1 allele is strongly associated with AS, which provided first evidence that CNV may also have important roles in the heritability of AS.

HLA-DQA1 is a HLA class II gene encoding an alpha chain of the HLA-DQ molecule, along with a beta chain (HLA-DQB) to form a heterodimer anchored in the membrane of antigen-presenting cells. Like other HLA class II molecules, HLA-DQ plays a central role in immune response to foreign antigens by presenting specific antigenic peptides to T cells. Genetic variations including sequence and CN of HLA genes contribute to enhance the recognition repertoire of the immune system, as well as to a wide
Association between copy number variations and AS
J Wang et al

AIMS AND METHODS

Subjects
A total of 533 AS patients and 792 unrelated controls of Chinese Han were examined in the studies. AS patients and controls were enrolled from Shanghai and Taizhou cities in China. The first 205 patients were enrolled from 2010 to 2011, and were studied earlier as the first cohort. The remaining 328 patients were enrolled from 2012 to 2013, and were examined later as the second cohort. All patients met the modified New York diagnostic criteria for AS. Controls were selected on the basis of a small number of cases screened by the CGH arrays should be cautious. The findings observed in the studies are limited to AS patients of the Chinese Han population. Further confirmation studies, especially in other ethnic populations, are necessary. In addition, biological functions of the CNV of the HLA-DQA1 have not been studied. Whether it impacts on presenting specific antigenic peptides to T cells, and induces or accelerates an immune response would be interesting to know.

Nonetheless, this is the first report based on studies of CNVs in AS. The results suggest that the CN of the HLA-DQA1 allele is strongly associated with susceptibility to AS in the Chinese Han population. Further replication and functional studies of this finding will be necessary.

Genome-wide CN variation (CNV) analysis
Five AS patients were examined by using Agilent SurePrint G3 Human CGH Microarrays (1 \times 1M) (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer’s protocol. Commercial genomic DNA (Promega, Madison, WI, USA) was used as the internal control. Briefly, genomic DNA of each subject was mixed with restriction digestion, and then was labeled with ULS-Cy5 (for patients) and ULS-Cy3 (for sex-matched controls). The labeled products of one patient and one control were mixed and hybridized to the array for 40 h at 60°C. Then, the array was washed and scanned on an Agilent Microarray Scanner. The data were extracted by Agilent Feature Extraction 10.7.3.1 and analyzed by Agilent Workbench 7.0. ADM-2 was used as statistical algorithm with a P-value threshold of 0.05. CN gains or losses of at least five consecutive oligomers on the array were selected for further analysis. The genome-wide common CNVs in human populations were adopted from the public data via the Database of Genomic Variants (http://projects.tcag.ca/ variation/) and our private Chinese CNV data. Genes present in these common aberrations regions (referred to as CNV regions) were identified using the human genome browser at UCSC.

AccuCopy technology for CNV validation
Two sets of primers were designed to examine the CNVs of two regions (HLA-DRB5 and HLA-DQA1) at 6p21.32 selected from the results of the CGH arrays. The primer sequences for these two regions are: 5’-GAGGGGG TTGGGTTCCTGC-3’/5’-ACAGGCACGTTGGGGAGTAC-3’ and 5’-GGTCACAGTGT TITTTAAGCTCTCC-3’/5’-ATGGGCGACGGTGAGGACAGT-3’. For the HLA-DRB5 and HLA-DQA1 regions, the probes cover human chromosomes (genome assembly of UCSC hg19) chr6: 32489824–32489901 and chr6: 32609759–32609998, respectively.

A total of 205 AS cases and 225 controls were examined as the first cohort, and then 328 cases and 567 controls were examined as the second cohort for validation of specific CNVs identified from the array-based CGH studies with the AccuCopy assay following the manufacturer’s protocol. Briefly, the genomic DNA of each subject was mixed with fluorescence-labeled specific primers, PCR Master mix and a competitive DNA with known CN for a multiple competitive real-time PCR reaction. The PCR products were diluted, and were then loaded on an ABI 3730XL sequencer for quantitation analysis. Raw data were analyzed by Gene Mapper 4.0. HG19 was used for the genome build for the genomic coordinates. The peak ratio between sample DNA and the corresponding competitive DNA (S/C) was calculated and then normalized to the median of four preset two-copy reference genes, respectively. Two normalized S/C ratios were further normalized to the median value in all samples for each reference gene and then averaged. The CN of each target fragment was determined by the average S/C ratio times two. Cases and controls were examined and read at the same time to minimize non-random errors.

HLA-B sequence typing
HLA-B genotyping was performed using the sequence-based typing method using SeCore Kits (Life Technologies, Grand Island, NY, USA). Briefly, allele-specific PCRs were performed using primers supplied in the SeCore kits, and were then followed by sequencing exons 2 and 3 of the HLA-B gene. The HLA sequence-based typing uTYPE 6.0 program (Life Technologies) was used in sequencing analysis and assigning HLA-B alleles.

Association analysis
Median and interquartile range (IQR) were used to describe the distribution of the epidemiological variables in the studies. The distributions of CNs between patients and controls after CN assignment according to the predefined threshold were compared using \( \chi^2 \) test for trend in proportions with R. Logistic regression models were constructed to determine the OR and 95% CI after adjusting for gender using SPSS (version 17.0, SPSS Inc., Chicago, IL, USA). Thresholds for deletions and duplications were empirically set at <1.75 and >2.35, respectively, in the above CNV validation assays according to the manufacturer’s instruction. All samples were tested in duplicates.

CONFLICT OF INTEREST

The authors declare no conflict of interest.
ACKNOWLEDGEMENTS
These studies were supported by research grants from the National Basic Research Program (2012CB944600), International S&T Cooperation Program of China (2013DFA30870), the Science and Technology Committee of Shanghai Municipality (11410701800, 11DJ1400102), Ministry of Science and Technology (2011BAI09800), Ministry of Health (201002007) and the US NIH NIAID U01, 1U01AI09090.

REFERENCES
1 Reveille JD, Brown MA. Epidemiology of ankylosing spondylitis: IGAS 2009. J Rheumatol 2010; 37: 2624–2625.
2 Thomas GP, Brown MA. Genomics of ankylosing spondylitis. Discov Med 2010; 10: 263–271.
3 Brown MA. Genomewide screens in ankylosing spondylitis. Adv Exp Med Biol 2009; 649: 148–158.
4 Li J, Yang T, Wang L, Yan H, Zhang Y, Guo Y et al. Whole genome distribution and ethnic differentiation of copy number variation in Caucasian and Asian populations. PLoS One 2009; 4: e7958.
5 Fanciulli M, Norrisworthy PJ, Petretto E, Dong R, Harper L, Kamesh L et al. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. Nat Genet 2007; 39: 721–723.
6 Grayson BL, Smith ME, Thomas JW, Dexheimer P, Jeffrey J et al. Genome-wide analysis of copy number variation in type 1 diabetes. PLoS One 2010; 5: e15393.
7 Shostakovich-Koretskaya L, Catano G, Chykarenko ZA, He W, Gornalusse G, Mummidi S et al. Combinatorial content of CCL3L and CCL4L gene copy numbers influence HIV-AIDS susceptibility in Ukrainian children. AIDS 2009; 23: 679–688.
8 Swaminathan S, Kim S, Shen L, Risacher SL, Foroud T, Pankratz N et al. Genomic copy number analysis in Alzheimer’s disease and mild cognitive impairment: an ADNI study. Int J Alzheimers Dis 2011; 2011: 729478.
9 Lin Z, Bei JX, Shen M, Li Q, Liao Z, Zhang Y et al. A genome-wide association study in Han Chinese identifies new susceptibility loci for ankylosing spondylitis. Nat Genet 2012; 44: 73–77.
10 Australo-Anglo-American Spondyloarthritis Consortium (TASC)Reveille JD, Sims AM, Danoy P, Evans DM, Leo P et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. Nat Genet 2010; 42: 123–127.
11 Brown MA, Pile KD, Kennedy LG, Campbell D, Andrew L, March R et al. A genomewide screen for susceptibility loci in ankylosing spondylitis. Arthritis Rheum 1998; 41: 588–595.
12 Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD et al. Global variation in copy number in the human genome. Nature 2006; 444: 444–454.
13 McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. Nat Genet 2007; 39: 537–542.
14 Solid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ, alpha/beta heterodimer. J Exp Med 1989; 169: 345–350.
15 Murray JA, Moore SB, Van Dyke CT, Lahr BD, Dierkhising RA, Zinzmeister RR et al. HLA DQ gene dosage and risk and severity of celiac disease. Clin Gastroenterol Hepatol 2007; 5: 1146–11412.
16 Petronzelli F, Maltagliati F, Ferrante P, Bonamico M, Rabuffo G, Campea L. Different dose effect of HLA-DQ alpha beta heterodimers in insulin-dependent diabetes mellitus and celiac disease susceptibility. Hum Immunol 1993; 36: 156–162.
17 Todd JA, Fukui Y, Kitagawa T, Sasazuki T. The A3 allele of the HLA-DQA1 locus is associated with susceptibility to type 1 diabetes in Japanese. Proc Natl Acad Sci USA 1990; 87: 1094–1098.
18 Wen L, Wang FS, Tang J, Chen NY, Altierei M, David C et al. In vivo evidence for the contribution of human histocompatibility leukocyte antigen (HLA)-DQ molecules to the development of diabetes. J Exp Med 2000; 191: 97–104.
19 Britten AC, Mijovic CH, Barnett AH, Kelly MA. Differential expression of HLA-DQ alleles in peripheral blood mononuclear cells: alleles associated with susceptibility to and protection from autoimmune type 1 diabetes. Int J Immunogenet 2009; 36: 47–57.
20 Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W et al. Evaluation of early diagnostic criteria for ankylosing spondylitis in a 10 year follow-up. Z Rheumatol 1990; 49: 82–87.
21 Iafraje AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y et al. Detection of large-scale variation in the human genome. Nat Genet 2004; 36: 949–951.
22 Lou H, Li S, Yang Y, Kang L, Zhang X, Jin W et al. A map of copy number variations in Chinese populations. PLoS One 2011; 6: e27341.
23 Zhang J, Feuk L, Duggan GE, Khaja R, Scherer SW. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. Cytogenet Genome Res 2006; 115: 205–214.
24 Du R, Lu C, Jiang Z, Li S, Ma R, An H et al. Efficient typing of copy number variations in a segmental duplication-mediated rearrangement hotspot using multiplex competitive amplification. J Hum Genet 2012; 57: 545–551.