RHEUMATOLOGY

Rheumatology 2013;52:1769–1774
doi:10.1093/rheumatology/ket241
Advance Access publication 19 July 2013

Concise report

Association of Takayasu arteritis with HLA-B*67:01 and two amino acids in HLA-B protein

Chikashi Terao[1,2], Hajime Yoshifuji[2], Koichiro Ohmura[2], Kosaku Murakami[3], Daisuke Kawabata[2], Kimiko Yurugi[4], Junichi Tazaki[5], Hideyuki Kinoshita[6], Akinori Kimura[7], Masashi Akizuki[8], Yasushi Kawaguchi[9], Hisashi Yamanaka[9], Yasuo Miura[4], Taira Maekawa[4], Hiroo Saji[10], Tsuneyo Mimori[2] and Fumihiko Matsuda[1]

Abstract

Objective. Takayasu arteritis (TAK) is a rare autoimmune arteritis that affects large arteries. Although the association between TAK and HLA-B*52:01 is established, the other susceptibility *HLA-B* alleles are not fully known. We performed genetic association studies to determine independent *HLA-B* susceptibility alleles other than HLA-B*52:01 and to identify important amino acids of HLA-B protein in TAK susceptibility.

Methods. One hundred patients with TAK and 1000 unrelated healthy controls were genotyped for *HLA-B* alleles in the first set, followed by a replication set containing 73 patients with TAK and 1000 controls to compare the frequencies of *HLA-B* alleles. Step-up logistic regression analysis was performed to identify susceptibility amino acids of HLA-B protein.

Results. Strong associations of susceptibility to TAK with HLA-B*52:01 and HLA-B*67:01 were observed ($P = 1.0 \times 10^{-16}$ and 9.5×10^{-6}, respectively). An independent susceptibility effect of HLA-B*67:01 from HLA-B*52:01 was also detected ($P = 1.8 \times 10^{-7}$). Amino acid residues of histidine at position 171 and phenylalanine at position 67, both of which are located in antigen binding grooves of the HLA-B protein, were associated with TAK susceptibility ($P \leqslant 3.8 \times 10^{-5}$) with a significant difference from other amino acid variations ($\Delta AIC \geqslant 9.65$).

Conclusion. HLA-B*67:01 is associated with TAK independently from HLA-B*52:01. Two amino acids in HLA-B protein are strongly associated with TAK susceptibility.

Key words: Takayasu arteritis, genetic association study, HLA-B, aortitis, vasculitis.

[1]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, [2]Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, [3]Department of Rheumatology and Clinical Immunology, Osaka Red Cross Hospital, Osaka, [4]Department of Transfusion Medicine and Cell Therapy, Kyoto University Hospital, Kyoto, [5]Department of Cardiovascular Medicine, Kyoto University Graduate School of Medicine, Kyoto, [6]Department of Medicine and Clinical Science, Kyoto University Graduate School of Medicine, Kyoto, [7]Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, [8]Akizuki Clinic, Yokohama, [9]Institute of Rheumatology, Tokyo Women's Medical University, Tokyo and [10]HLA Laboratory, Kyoto, Japan.

Submitted 29 January 2013; revised version accepted 21 May 2013.

Correspondence to: Chikashi Terao, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan. E-mail: a0001101@kuhp.kyoto-u.ac.jp

Introduction

Takayasu arteritis (TAK) is a relatively rare systemic arteritis that affects mainly large branches of the arteries, including the aorta, carotid artery, subclavian artery and coronary artery. TAK affects mainly young females. Both environmental and genetic components have been shown to be involved with the onset of TAK [1]. Among genetic components, HLA-B52, mainly B*52:01, has shown a strong association with TAK with odds ratios (ORs) of ~3 [2, 3]. The association with HLA-B52 has been repeatedly shown in various studies and the association is established beyond ethnicity [4, 5]. Other associations between TAK and HLA alleles have not been confirmed due to the low prevalence of the disease and the lack of

large-scale comprehensive genetic analysis using TAK patients. While a previous study showed that HLA-B39 is also associated with TAK [6], another recent study failed to replicate the association [7]. The recent study also suggested that a relatively rare *HLA-B* allele, HLA-B*67:01, was associated with TAK susceptibility (corrected $P = 0.023$) [7], but there are no replication studies. No studies have addressed independent associations of *HLA-B* alleles from HLA-B*52:01. Moreover, no large studies have ever been performed to analyse which amino acid residues are important for TAK susceptibility. As some TAK patients develop autoantibodies against aortic endocardium [8], the detection of susceptibility amino acids to TAK would lead to identification of a possible antigen that provokes an autoimmune response in TAK patients.

Materials and methods

Study subjects

DNA samples of 100 Japanese patients with TAK were collected at Kyoto University Hospital and Tokyo Women's Medical University. DNA samples of 73 patients with TAK were also collected at Kyoto University Hospital. Two patients in the first set are a parent and a child. Each of 1000 control DNA samples in the first and replication sets were collected at the HLA laboratory from unrelated healthy individuals. All the patients were diagnosed with TAK based on ACR criteria [9] or guidelines of the Japanese Circulation Society [10] or were registered to the Japanese national registry for rare and intractable diseases (http://www.nanbyou.or.jp/english/index.htm). Information on classifications of TAK based on criteria by Hata *et al.* [11] and of complications of aortic regurgitation (AR) were obtained from 75 and 85 patients, respectively, by reviewing clinical charts from the Kyoto University Hospital. A summary of the subjects is shown in supplementary Table S1, available at *Rheumatology* Online. This study was approved by the local ethics committees at each institution (Kyoto University Graduate School and Faculty of Medicine Ethics Committee, Tokyo Women's Medical University Genome Ethics Committee and Tokyo Women's Medical University Ethics Committee) and written informed consent was obtained from all subjects.

HLA-B genotyping

The cases and controls were genotyped for *HLA-B* alleles in four-digit resolution using the WAKFlow system at the Kyoto University Hospital and HLA laboratory, respectively.

Amino acid sequences of *HLA-B* alleles

Amino acid sequences were obtained for each *HLA-B* allele for four-digit resolution from the IMGT database (http://www.ebi.ac.uk/ipd/imgt/hla/). Amino acid variations were searched across all the *HLA-B* alleles contained in the study population. In total, 99 variants of amino acids over 53 positions were determined in the *HLA-B* alleles. A three-dimensional structure analysis of the HLA-B protein was performed using UCSF Chimera software (University of California, San Francisco, San Francisco, CA, USA) [12].

Statistical analysis

The association between TAK and *HLA-B* alleles was investigated using 2×2 contingency tables and tested for statistical significance using the χ^2 test or Fisher's exact test for alleles with a frequency >1% in either the cases or the controls. Heterogeneity between the two studies was analysed using the Breslow–Day test. The ORs and 95% CIs were also calculated. The relative predispositional effect (RPE) method [13] was applied to identify the associations of more than one *HLA-B* allele sequentially according to their strength. The associations between clinical phenotypes and HLA-B*67:01 were assessed by the Wilcoxon rank sum test for age at onset or Fisher's exact test for other clinical phenotypes. Crude *P*-values were indicated unless we mentioned the use of corrected *P*-values.

The 99 amino acids over 53 positions were used as independent variables in logistic regression analysis (supplementary Table S2, available at *Rheumatology* Online). When we found only two amino acids in the same position, we set one independent variable in this position. When we found more than three amino acids in the same position, we distinguished all amino acids as independent variables. Step-up multiple logistic regression analyses for TAK susceptibility were performed to identify susceptibility amino acids in *HLA-B* in the combined study until an amino acid whose *P*-value did not reach significance was chosen. The Akaike information criterion (AIC) was also calculated.

Significant levels were set according to Bonferroni's correction, namely, 0.0028 for *HLA-B* alleles, 0.05 for clinical phenotypes and 0.00051 for amino acids. Amino acid variations were considered significant over other variations when the logistic regression model containing the variation showed the smallest AIC with the difference in AIC >7 in comparison with those containing other alleles (ΔAIC > 7) [14]. Permutation tests were performed 10 000 times to assess the probability that the observed improvements of deviance in multiple logistic regression analysis were obtained by chance. Deviances of the best-fitting model using one, two and three amino acids were calculated in each permutation test and sequential improvements of deviance by the best three amino acids were compared with the observed improvements. Deviance of logistic regression analysis is defined as $-2 \times \log$ likelihood of logistic models with degrees of freedom of k, where $k-1$ is the number of alleles in the models.

Results

The strong association of TAK with HLA-B*52:01 was confirmed in the first set ($P = 2.6 \times 10^{-13}$). A significant association with HLA-B*67:01 was also observed ($P = 0.00081$; Table 1), and no other *HLA-B* alleles showed significant associations with TAK. The replication

TABLE 1 Associations of HLA-B alleles with TAK susceptibility

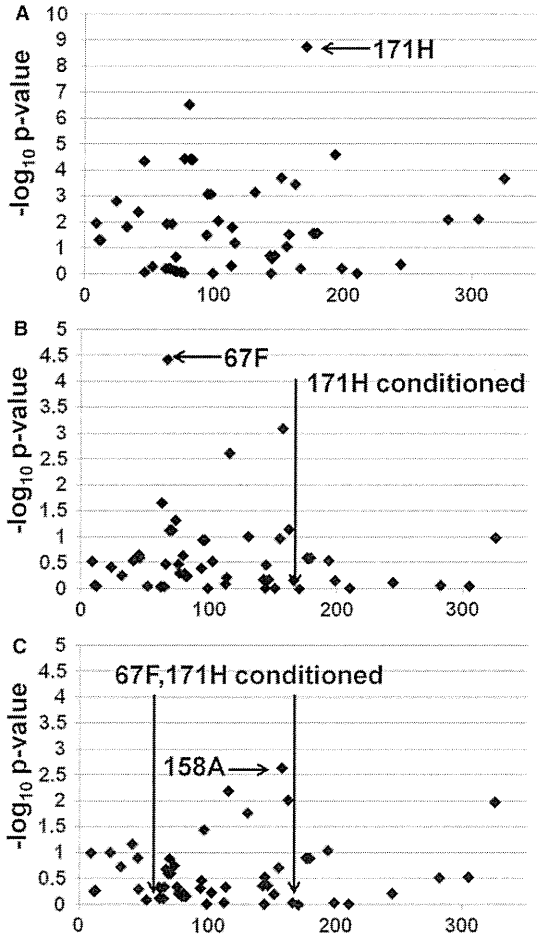
	First set				Replication set				Combined study		
HLA-B allele[a]	Case	Control	*P*	OR (95% CI)	Case	Control	*P*	OR (95% CI)	*P*	OR (95% CI)	RPE *P*
B*07:02	7	122	0.14	0.56 (0.26, 1.21)	6	119	0.36	0.68 (0.29, 1.57)	0.085	0.61 (0.34, 1.08)	
B*13:01	1	28	0.51	0.35 (0.0086, 2.17)	1	27	1.00	0.50 (0.012, 3.10)	0.32	0.42 (0.049, 1.59)	
B*15:01	16	153	0.86	1.05 (0.61, 1.80)	11	152	0.98	0.99 (0.52, 1.87)	0.90	1.03 (0.68, 1.54)	
B*15:18	1	27	0.31	0.51 (0.0089, 2.25)	0	35	0.17	0 (0, 1.52)	0.059	0.18 (0.0046, 1.08)	
B*35:01	10	160	0.13	0.61 (0.31, 1.17)	9	164	0.38	0.74 (0.37, 1.47)	0.084	0.66 (0.41, 1.06)	
B*39:01	3	68	0.20	0.43 (0.086, 1.34)	8	61	0.11	1.84 (0.86, 3.93)	0.96	0.99 (0.53, 1.84)	
B*40:01	6	108	0.14	0.54 (0.24, 1.25)	9	115	0.84	1.08 (0.53, 2.17)	0.33	0.77 (0.45, 1.31)	
B*40:02	14	177	0.38	0.78 (0.44, 1.36)	6	153	0.11	0.52 (0.22, 1.19)	0.11	0.68 (0.43, 1.09)	
B*40:06	8	92	0.70	0.86 (0.41, 1.81)	1	102	0.0085	0.13 (0.0032, 0.74)	0.057	0.52 (0.27, 1.03)	
B*44:03	13	125	0.89	1.04 (0.58, 1.88)	6	115	0.41	0.7 (0.30, 1.62)	0.70	0.91 (0.56, 1.47)	
B*46:01	11	87	0.45	1.28 (0.67, 2.44)	6	102	0.6	0.8 (0.34, 1.85)	0.87	1.04 (0.63, 1.73)	
B*48:01	5	60	0.69	0.83 (0.33, 2.09)	3	61	0.80	0.67 (0.13, 2.08)	0.45	0.76 (0.37, 1.57)	
B*51:01	10	161	0.12	0.60 (0.31, 1.16)	14	169	0.63	1.15 (0.65, 2.04)	0.39	0.83 (0.54, 1.27)	
B*52:01	60	232	$2.6\text{x}10^{-13}$	3.27 (2.34, 4.55)	34	235	4.8×10^{-5}	2.28 (1.52, 3.43)	1.6×10^{-16}	2.82 (2.19, 3.64)	1.6×10^{-16}
B*54:01	12	133	0.72	0.90 (0.49, 1.65)	9	149	0.57	0.82 (0.41, 1.63)	0.57	0.85 (0.54, 1.35)	
B*55:02	2	47	0.31	0.42 (0.049, 1.62)	5	58	0.72	1.19 (0.47, 3.01)	0.72	0.77 (0.35, 1.66)	
B*59:01	2	41	0.43	0.48 (0.056, 1.88)	1	39	0.52	0.35 (0.0085, 2.08)	0.21	0.43 (0.086, 1.31)	
B*67:01	9	27	0.00081	3.44 (1.60, 7.43)	5	19	0.0061	3.70 (1.36, 10.05)	9.5×10^{-6}	3.62 (1.97, 6.66)	1.8×10^{-7}

[a]Alleles with a frequency of >1% in either the case or control are indicated.

study was performed to confirm the findings in the first set. As a result, the association of HLA-B*67:01 with TAK was observed ($P = 0.0061$; Table 1), as well as that of HLA-B*52:01 ($P = 4.8 \times 10^{-5}$). HLA-B*40:06 showed a protective association against TAK ($P = 0.0085$). In the combined analysis, solid evidence of associations of HLA-B*67:01 and HLA-B*52:01 with TAK was obtained [$P = 9.5 \times 10^{-6}$ and 1.0×10^{-16}, OR = 3.62 (95% CI 1.97, 6.66) and 2.82 (95% CI 2.19, 3.64), respectively; Table 1]. Any signs of heterogeneity between the two studies were observed for these two alleles ($P \geqslant 0.18$). No other alleles, including HLA-B*40:06, showed significant associations. To assess the independence of HLA-B*67:01's association from HLA-B*52:01 and to detect further candidates of independent susceptibility *HLA-B* alleles, RPE analysis was performed in the combined study. As a result, HLA-B*67:01 showed a significant association with TAK susceptibility that was independent of HLA-B*52:01 ($P = 1.8 \times 10^{-7}$; Table 1). The analysis did not show further candidates with independent associations with TAK ($P \geqslant 0.17$). Since previous studies suggested associations between *HLA-B* alleles and disease phenotypes [15], we analysed the associations between HLA-B*67:01 and age at onset, female ratio, classifications of TAK or complications of AR. As a result, we did not observe significant associations ($P \geqslant 0.61$; supplementary Tables S3 and S4, available at *Rheumatology* Online).

Next, which amino acid residues contribute to TAK susceptibility was analysed by a step-up multiple logistic regression analysis in the combined study, using 99 amino acid variations over the HLA-B protein (see the Materials and methods section). Logistic regression analysis revealed the strongest association of histidine residue at the position of the 171st amino acid residue ($P = 1.8 \times 10^{-9}$; Fig. 1 and supplementary Table S5, available at *Rheumatology* Online), with a ΔAIC of 11.17 in comparisons with other amino acids. In the case of conditioning with the 171st histidine, phenylalanine at the 67th amino acid residue showed the strongest association ($P = 3.8 \times 10^{-5}$; Fig. 1), with an ΔAIC of 9.65. An alanine residue at the 158th position showed a suggestive association in conditioning with these two amino acids ($P = 0.0024$; Fig. 1) without significant ΔAIC (0.13), indicating uncertainty of the selection of this amino acid as the third susceptibility amino acid. We did not find any further amino acid variations with significant associations in the condition with these three amino acids ($P \geqslant 0.065$). A total of 10 000 permutation tests revealed that improvements in deviance from these three amino acids were less likely to be obtained by chance (permutation *P*-value 0.0001, 0.0001 and 0.0024, respectively). The 171st and 67th amino acid residues are located in peptide binding grooves, implying that an antigen binding capacity conferred by these amino acids might influence TAK susceptibility (supplementary Fig. S1, available at *Rheumatology* Online). When HLA-B*52:01 was excluded from the association studies, the two amino acid variations did not show significant associations (data not shown), indicating that the associations of the two amino acids were brought by HLA-B*52:01.

FIG. 1 Associations of amino acids of the HLA-B protein with TAK susceptibility.



P-values are plotted according to the amino acid positions in (**A**) single logistic regression analysis, (**B**) in the condition with histidine at position 171 and (**C**) with 171 histidine and 67 phenylalanine.

Discussion

This is the first study to provide solid evidence of an *HLA-B* TAK susceptibility allele independent of HLA-B*52:01 and to report on TAK susceptibility amino acids in HLA-B protein. As the top two residues of 171 and 67 were located at peptide binding grooves and have shown their significance in peptide binding [16, 17], the susceptibility effects on TAK appear to be reasonable.

HLA-B39 did not show an association in the current study, or in another recent study [7]. Thus association between HLA-B39 and TAK is not likely. Although the previous study suggested the association of TAK with HLA-B*67:01, the association was not conclusive [7]. The current study has clearly revealed that HLA-B*67:01 is associated with TAK independently of HLA-B*52:01. The association and independence of

HLA-B*67:01 were observed in the two independent sets. The two relatives in our study did not have HLA-B*67:01, thus having no effect on the conclusion. When we analysed the previous data in which only serological *HLA-B* type was available [6], HLA-B67 showed a suggestive association with TAK after excluding HLA-B52 (95 cases and 232 controls) [6]. As HLA-B*67:01 is present specifically in East Asians, this association should be specific to East Asians. The RPE analysis did not reveal an association of HLA-B*40:06. This indicates that the protective association of HLA-B*40:06 in the replication set was conferred by positive associations of HLA-B*52:01 and HLA-B*67:01. While we detected the association of HLA-B*67:01 with TAK, HLA-B*67:01 did not show significant associations with age at onset, female ratio, TAK classifications and AR. Based on the fact that the *P*-values are far from significant, it is not very likely that HLA-B*67:01 has a strong effect on clinical phenotype. Further detailed clinical information, including affected branches of the aorta, and disease activity were available for only a limited number of patients in our study, so we could not analyse the association between the susceptibility allele and detailed clinical phenotypes.

The 171st histidine was found to be a susceptibility amino acid and the 67th phenylalanine was found to be a protective amino acid. While the permutation *P*-value supported the existence of the third susceptibility amino acid, alanine at position 158 did not show a significant ΔAIC in comparison with other amino acids. Thus we cannot specify the third susceptibility amino acid to TAK. The significant associations of the top two amino acids were mainly attributed to HLA-B*52:01. No other alleles share the two amino acids with HLA-B*52:01. Although HLA-B*51:01 has similar amino acid sequences to HLA-B*52:01 with two different amino acid residues, it is not associated with TAK [4, 7]. This lack of association can be explained by our findings because HLA-B*51:01 includes the 67th phenylalanine, the protective allele against TAK. HLA-B*67:01 did not have histidine in position 171. The effects of amino acids in HLA-B*67:01 were not very apparent due to its low frequency. These two amino acids did not explain TAK susceptibility more efficiently than HLA-B*52:01 (ΔAIC over the null model = 53.94 and 66.01, respectively), indicating that these amino acids cannot explain all the susceptibility effects of HLA-B*52:01. This suggests the existence of a haplotypic effect of amino acids in HLA-B*52:01 or further susceptibility amino acids.

The 171st tyrosine forming an A pocket in the α2 domain is one of the highly conserved amino acid residues among *HLA-B* alleles. The conversion into histidine has been shown to modulate the binding groove, coordinating the N-terminus of the binding peptides in HLA-B14 and HLA-B*51:01 [17, 18]. The 67th serine in the B pocket of the peptide binding groove in HLA-B27 has been shown to influence peptide presentation [16]. These reports suggest important roles for the two amino acids in TAK susceptibility by modulating peptide binding. A previous Mexican study, which included 19 patients with TAK, suggested that serine at 67 and glutamic acid at 63 was associated with TAK [19]. The previous report seemed to support our results, as the current study showed non-phenylalanine amino acid residues at position 67, including serine, were positively associated with TAK. As the current study did not specify the third susceptibility amino acid, the specific association of alanine at position 158 with TAK is inconclusive. Thus these combinations of amino acids should be regarded as incomplete and further investigation is required.

As previous studies surveying susceptibility HLA alleles beyond *HLA-B* did not find significant alleles beyond the *HLA-B* alleles [6, 7], it is not very likely that amino acids in other HLA alleles linked with HLA-B*52:01 and HLA-B*67:01 can explain the TAK susceptibility of the HLA locus. Nevertheless, it is undeniable that other genes in the HLA locus, without belonging to HLA alleles, play a critical role in TAK susceptibility. Amino acid residues in the HLA locus should be analysed to grasp the whole picture of the relationship between the HLA locus and TAK susceptibility. It would be interesting to determine what kinds of protein bind to the alleles containing the susceptibility amino acid residues.

Rheumatology key messages

- HLA-B*67:01 is an independent susceptibility allele to TAK from HLA-B*52:01.
- Amino acids at positions 171 and 67 of the HLA-B protein show TAK susceptibility effects .

Acknowledgements

We would like to thank Ms Kayo Umemoto, who coordinated the acquisition of blood samples. We would like to thank all the patients who gave blood samples and the medical staff members who helped us with this study. C.T. and H. Yoshifuji participated in conception and design of this study. C.T., H. Yoshifuji, K.O., K.M., D.K., J.T., H.K., K.A., K.Y., Y.M., T. Maekawa, H.S., M.A., Y.K., H. Yamanaka, F.M. and T. Mimori participated in acquisition of data. C.T. performed analysis and interpretation of the data. C.T. wrote the initial draft and all the authors reviewed and approved the final manuscript.

Funding: This study was supported by a Kyoto University Step-up grant as well as grants-in-aid from the Ministry of Health, Labor, and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Disclosure statement: The authors declare no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Arnaud L, Haroche J, Mathian A *et al*. Pathogenesis of Takayasu's arteritis: a 2011 update. Autoimmun Rev 2011; 11:61–7.

2 Yoshida M, Kimura A, Katsuragi K *et al*. DNA typing of HLA-B gene in Takayasu's arteritis. Tissue Antigens 1993; 42:87–90.

3 Yajima M, Numano F, Park YB *et al*. Comparative studies of patients with Takayasu arteritis in Japan, Korea and India—comparison of clinical manifestations, angiography and HLA-B antigen. Jpn Circ J 1994;58:9–14.

4 Sahin Z, Bicakcigil M, Aksu K *et al*. Takayasu's arteritis is associated with HLA-B*52, but not with HLA-B*51, in Turkey. Arthritis Res Ther 2012;14:R27.

5 Vargas-Alarcon G, Flores-Dominguez C, Hernandez-Pacheco G *et al*. Immunogenetics and clinical aspects of Takayasu's arteritis patients in a Mexican Mestizo population. Clin Exp Rheumatol 2001;19:439–43.

6 Kimura A, Kitamura H, Date Y *et al*. Comprehensive analysis of HLA genes in Takayasu arteritis in Japan. Int J Cardiol 1996;54(Suppl):S61–9.

7 Takamura C, Ohhigashi H, Ebana Y *et al*. New human leukocyte antigen risk allele in Japanese patients with Takayasu arteritis. Circ J 2012;76:1697–702.

8 Eichhorn J, Sima D, Thiele B *et al*. Anti-endothelial cell antibodies in Takayasu arteritis. Circulation 1996;94: 2396–401.

9 Arend WP, Michel BA, Bloch DA *et al*. The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum 1990; 33:1129–34.

10 Japanese Circulation Society Joint Working Group. Guideline for management of vasculitis syndrome (JCS 2008). Circ J 2011;75:474–503.

11 Hata A, Noda M, Moriwaki R *et al*. Angiographic findings of Takayasu arteritis: new classification. Int J Cardiol 1996; 54(Suppl):S155–63.

12 Pettersen EF, Goddard TD, Huang CC *et al*. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004;25: 1605–12.

13 Payami H, Joe S, Farid NR *et al*. Relative predispositional effects (RPEs) of marker alleles with disease: HLA-DR alleles and Graves disease. Am J Hum Genet 1989;45: 541–6.

14 Burnham KP, Anderson, David R. Model selection and multi-model inference: a practical information-theoretic approach. 2nd edn. New York: Springer-Verlag, 2002.

15 Kitamura H, Kobayashi Y, Kimura A *et al*. Association of clinical manifestations with HLA-B alleles in Takayasu arteritis. Int J Cardiol 1998;66(Suppl 1): S121–6.

16 Appel H, Kuon W, Kuhne M *et al*. The solvent-inaccessible Cys67 residue of HLA-B27 contributes to T cell recognition of HLA-B27/peptide complexes. J Immunol 2004;173:6564–73.

17 Maenaka K, Maenaka T, Tomiyama H *et al*. Nonstandard peptide binding revealed by crystal structures of HLA-B*5101 complexed with HIV immunodominant epitopes. J Immunol 2000;165:3260–7.

18 Kumar P, Vahedi-Faridi A, Saenger W *et al*. Structural basis for T cell alloreactivity among three HLA-B14 and HLA-B27 antigens. J Biol Chem 2009;284: 29784–97.

19 Vargas-Alarcon G, Hernandez-Pacheco G, Soto ME *et al*. Comparative study of the residues 63 and 67 on the HLA-B molecule in patients with Takayasu's arteritis. Immunol Lett 2005;96:225–9.

Two Susceptibility Loci to Takayasu Arteritis Reveal a Synergistic Role of the *IL12B* and *HLA-B* Regions in a Japanese Population

Chikashi Terao,[1,2,*] Hajime Yoshifuji,[2] Akinori Kimura,[3] Takayoshi Matsumura,[4] Koichiro Ohmura,[2] Meiko Takahashi,[1] Masakazu Shimizu,[1] Takahisa Kawaguchi,[1] Zhiyong Chen,[3] Taeko K. Naruse,[3] Aiko Sato-Otsubo,[5] Yusuke Ebana,[6] Yasuhiro Maejima,[7] Hideyuki Kinoshita,[8] Kosaku Murakami,[9] Daisuke Kawabata,[2] Yoko Wada,[10] Ichiei Narita,[10] Junichi Tazaki,[11] Yasushi Kawaguchi,[12] Hisashi Yamanaka,[12] Kimiko Yurugi,[13] Yasuo Miura,[13] Taira Maekawa,[13] Seishi Ogawa,[5] Issei Komuro,[4] Ryozo Nagai,[14] Ryo Yamada,[1] Yasuharu Tabara,[1] Mitsuaki Isobe,[7] Tsuneyo Mimori,[2] and Fumihiko Matsuda[1]

Takayasu arteritis (TAK) is an autoimmune systemic vasculitis of unknown etiology. Although previous studies have revealed that HLA-B*52:01 has an effect on TAK susceptibility, no other genetic determinants have been established so far. Here, we performed genome scanning of 167 TAK cases and 663 healthy controls via Illumina Infinium Human Exome BeadChip arrays, followed by a replication study consisting of 212 TAK cases and 1,322 controls. As a result, we found that the *IL12B* region on chromosome 5 (rs6871626, overall $p = 1.7 \times 10^{-13}$, OR = 1.75, 95% CI 1.42–2.16) and the *MLX* region on chromosome 17 (rs665268, overall $p = 5.2 \times 10^{-7}$, OR = 1.50, 95% CI 1.28–1.76) as well as the *HLA-B* region (rs9263739, a proxy of HLA-B*52:01, overall $p = 2.8 \times 10^{-21}$, OR = 2.44, 95% CI 2.03–2.93) exhibited significant associations. A significant synergistic effect of rs6871626 and rs9263739 was found with a relative excess risk of 3.45, attributable proportion of 0.58, and synergy index of 3.24 ($p \leq 0.00028$) in addition to a suggestive synergistic effect between rs665268 and rs926379 ($p \leq 0.027$). We also found that rs6871626 showed a significant association with clinical manifestations of TAK, including increased risk and severity of aortic regurgitation, a representative severe complication of TAK. Detection of these susceptibility loci will provide new insights to the basic mechanisms of TAK pathogenesis. Our findings indicate that *IL12B* plays a fundamental role on the pathophysiology of TAK in combination with HLA-B*52:01 and that common autoimmune mechanisms underlie the pathology of TAK and other autoimmune disorders such as psoriasis and inflammatory bowel diseases in which *IL12B* is involved as a genetic predisposing factor.

Introduction

Takayasu arteritis (TAK [MIM 207600]) is an autoimmune systemic vasculitis that was first reported from Japan.[1] It is estimated that TAK affects around 0.004% of the population in Japan, especially young women aged between 15 and 35. Although TAK was originally thought to affect individuals of mainly Asian origin, individuals with TAK have been identified worldwide, though with lower prevalence compared to Asia.[2] TAK is characterized by the involvement of large arteries, especially the aorta and its large branches, and is grouped into "vasculitis affecting large vessels" according to the Chapel Hill classification.[3] Individuals with TAK develop a wide range of symptoms such as fatigue, syncope, and lowering of vision in addition to its characteristic complications including aortic regurgitation (AR), pulselessness, and difference of blood pressure between right and left upper limbs. Previous studies have revealed that genetic components are involved in the pathogenesis of TAK, and HLA-B*52:01 is so far the only established genetic factor across the world.[4–7] Other genetic components especially outside of the HLA locus have not been confirmed to date. Establishment of association with non-HLA regions would lead to a deeper understanding of the basics of TAK pathology and the development of a novel therapy for this vasculitis. Here, we performed a genome-scanning study of TAK to identify the genetic predisposing factors for TAK.

Subjects and Methods

Study Subjects

A total of 379 TAK cases and 1,985 controls were enrolled in this study. All the cases were diagnosed based on the criteria of

[1]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan; [2]Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan; [3]Department of Molecular Pathogenesis, Medical Research Institute, Tokyo Medical and Dental University, Tokyo 113-8510, Japan; [4]Department of Cardiovascular Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan; [5]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo 113-8655, Japan; [6]Department of Bioinformational Pharmacology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo 113-8510, Japan; [7]Department of Cardiovascular Medicine, Tokyo Medical and Dental University, Tokyo 113-8510, Japan; [8]Department of Medicine and Clinical Science, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan; [9]Osaka Red Cross Hospital, Osaka 543-8555, Japan; [10]Division of Clinical Nephrology and Rheumatology, Niigata University Graduate School of Medical and Dental Sciences, Niigata 951-8510, Japan; [11]Department of Cardiovascular Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan; [12]Institute of Rheumatology, Tokyo Women's Medical University, Tokyo 162-0054, Japan; [13]Department of Transfusion Medicine and Cell Therapy, Kyoto University Hospital, Kyoto 606-8507, Japan; [14]Jichi Medical University, Tochigi 329-0498, Japan
*Correspondence: a0001101@kuhp.kyoto-u.ac.jp
http://dx.doi.org/10.1016/j.ajhg.2013.05.024.

Table 1. Summary of Study Subjects

	Case	Control
Genome Scanning		
Number	167	663
Age[a]	45.7 ± 15.2	53.5 ± 13.5
Female ratio	0.92	0.74
Age at onset[a]	30.5 ± 14.5	NA
Genotyping	Illumina Infinium Human-Exome BeadChip	Illumina Infinium Human-Exome BeadChip
Subjects with clinical information	AR:87; CRP:89	NA
Institutions	Kyoto University; Tokyo Women's Medical University	Kyoto University
Replication Study		
Number	212	1,322
Age[a]	46.6 ± 17.6	53.3 ± 13.4
Female ratio	0.94	0.62
Age at onset[a]	27.0 ± 11.8	NA
Genotyping	Taqman assay	Illumina Infinium Human Omni 2.5-4 BeadChip, Illumina Infinium Human Omni 2.5-8 BeadChip
Subjects with clinical information	AR:102; CRP:None	NA
Institutions	Tokyo Medical and Dental University; Kyoto University; Niigata University	Kyoto University

Abbreviations are as follows: NA, not applicable; AR, aortic regurgitation; CRP, C-reactive protein.
[a]Mean ± standard deviation (SD).

American College of Rheumatology[8] or guideline provided by Japanese Circulation Society.[9] The control subjects were collected as a part of the Nagahama Prospective Genome Cohort for Comprehensive Human Bioscience (The Nagahama Study), a community-based prospective multiomics cohort study conducted by Kyoto University.[10] This study was approved by the local ethical committees at each institution, and written informed consent was obtained from each subject involved in the study.

Genome Scanning

Illumina Infinium Human Exome BeadChip arrays (Illumina) were used for genome scanning of the cases and the controls. The genome scanning was conducted in Center for Genomic Medicine, Kyoto University Graduate School of Medicine.

Quality Control of Genome Scanning

Polymorphisms showing success rates less than 0.95 in either cases or controls, departure from Hardy-Weinberg equilibrium (HWE) ($p < 1.0 \times 10^{-5}$), or minor allele frequencies less than 0.05 in both cases and controls were excluded from the analysis. Subjects who showed success rates less than 0.95 or evidence of relatedness with other subjects were also excluded. Kinship between study subjects were estimated by PLINK.[11] Quantile-quantile plot (QQ plot) was used to assess the population stratification of the study. Because 1,827 markers over 24,487 were located in the HLA locus in which polymorphisms are very closely linked with each other, the 22,660 markers in the non-HLA regions were used for QQ plot.

Replication Study

The SNPs with p values less than 1.0×10^{-5} in the genome scanning were selected for the replication study. Because the association found in the *HLA-B* region (MIM 142830) was largely attributable to HLA-B*52:01, rs9263739, a proxy of HLA-B*52:01, was selected as a representative of the HLA locus. In the replication study, case samples were genotyped by Taqman Assay (Applied Biosystems) and control genotypes were extracted from array data (Table 1).

Combined Study and Association Study for Genotypes

Association studies of genotypes were performed by chi-square test based on 2 × 2 contingency tables. Combined study of the two studies was performed by inverse-variance method, assuming a fixed-effects model from the effect size (logarithm of odds ratio [OR]) in each study. A significant level for detecting susceptibility genes was set as 2.0×10^{-6}, which was obtained by Bonferroni's correction. A stringent cut-off level of 5.0×10^{-8} was also applied to assess overall significance.

Imputation of Genotypes

Mach dat2 software[12] was used for imputation of the whole genomes based on the results of genome scans with the use of the East Asian panel of HapMap phase II data as reference. SNPs with low imputation scores (Rsq < 0.3) were excluded from the analysis.

Calculation of Linkage Disequilibrium

LD between SNPs in the Illumina Infinium Human Exome BeadChip was assessed based on the genome-scanning data. HapMap project phase II data was used when SNPs were not contained in the array. LD between HLA-B*52:01 and SNPs was calculated by combining our previous HLA-genotyping data of the 173 TAK cases (C.T., unpublished data) by WAKFlow system (Wakunaga Pharmaceutical) with the genome-scanning data.

Estimation of Interaction

We used the method for evaluation of interaction proposed by Andersson et al.[13] Gene-gene interaction was defined as departure from additivity of two loci and measured by three indices based on calculation of relative risk (RR); relative excess risk due to interaction (RERI), attributable proportion (AP), and synergy index (SI). We considered an interaction as significant only when both RERI and AP were different from 0 and additionally SI was more than 1. The very low prevalence of TAK justifies to approximate OR by RR. For instance, when we assessed the interaction between rs9263739 and rs6871626 through these three indices, the subjects were classified into four groups: negative for both rs9263739 T allele and rs6871626 A allele, positive for rs9263739 T allele and negative for rs6871626 A allele, negative for rs9263739 T allele and positive for rs6871626 A allele, and positive for both rs9263739 T allele and rs6871626 A allele. Logistic models were used to calculate the indices.

In Silico Analysis of Association between the Gene Expression and rs6871626

We used two methods to assess the effect of rs6871626 on the *IL12B* (MIM 161561) expression. Gene expression data for *IL12B*

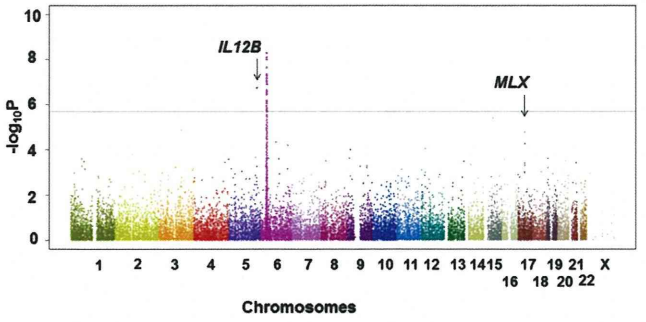


Figure 1. Manhattan Plot of Genome Scanning
The horizontal line indicates the significant level based on Bonferroni's correction. The HLA locus on chromosome 6 and the *IL12B* region on chromosome 5 reached the significant level.

in lymphoblastoid cells were obtained from GEO database (accession number GSE6536)[14] and analyzed for association with genotypes of rs6871626 obtained from HapMap project. Genevar software was used for analyzing the *IL12B* expression in adipose and skin in association with the rs6871626 genotypes.[15] Associations between genotypes and gene expression were evaluated by a linear regression analysis.

Associations between Genotypes and Clinical Phenotypes of TAK

Data of age at onset were analyzed for the association with the susceptibility alleles. AR, ischemic heart disease, and pulmonary infarction were selected for the association with genotypes as representative complications of TAK because cardiovascular event was the major cause of death in TAK individuals[16] and it was previously demonstrated that these phenotypes were associated with HLA-B*52:01,[17] suggesting that genetic backgrounds were at least partly responsible for these clinical manifestations. Data of the clinical manifestations were collected in Kyoto University Hospital or Tokyo Medical and Dental University by medical doctors who were blinded to genotype data reviewing clinical charts. Although AR evaluated by transthoracic echocardiography or angiography was positive for 44% of cases, other complications were found in less than 16%. Only AR was analyzed because of lack of power for other manifestations. Data for severity of AR assessed by the three categories[18] (mild, moderate, and severe) were also collected. C-reactive protein (CRP) was focused on as a biomarker reflecting disease activity. We calculated time-averaged CRP and dosage of prednisolone. Individuals who had visited hospitals for less than 500 days were excluded from the analysis of CRP. The associations between genotypes and clinical phenotypes were assessed by logistic regression analysis for existence of AR or linear regression analysis for severity of AR, time-averaged CRP, and age at onset. Time-averaged CRP was analyzed in condition with time-averaged dosage of prednisolone alone or in combination with rs3093059 genotypes in the *CRP* (MIM 123260) region. Associations between genotypes and clinical manifestations with p values less than 0.05 were regarded as significant.

Statistical Analysis

Statistical analyses were performed by PLINK v.1.07, R statistical software, or SPSS v.18.0.

Results

A summary of basic information of the subjects in our study is shown in Table 1. DNA samples from 167 cases and 663 healthy controls were genome scanned with the use of Illumina Human-Exome arrays containing 247,730 SNPs. One sample of the TAK cases and six samples in controls with success rates of less than 0.95 or with evidence of relatedness with other subjects (PI_HAT > 0.2 calculated by PLINK, see Subjects and Methods) were excluded from further analysis. The genotyping revealed that more than 80% of the markers in the array were monomorphic and 9% of the markers showed low minor allele frequency (<0.05) in the Japanese population, respectively. A total of 24,487 markers remained after filtering of SNPs that showed success rates of less than 0.95, deviation from HWE ($p < 1 \times 10^{-5}$) in either cases or controls, or minor allele frequencies of less than 0.05 in both cases and controls. The mean success rate of individuals was 0.999 after filtering.

Association studies were performed by chi-square test to compare allele frequencies between cases and controls. Population stratification was evaluated by QQ plot. The results indicated a lambda value of 1.05 in the QQ plot, indicating no excess population stratification in our study. Manhattan plot revealed that a region on chromosome 5 as well as the HLA locus showed significant associations that satisfied the genome-wide significant threshold obtained by Bonferroni's correction ($p = 2.0 \times 10^{-6}$; Figure 1). The associations were also confirmed by the imputed results (Figure S1 available online). rs4947248 in the *HLA-B* region, which is a known susceptibility gene to TAK, showed the strongest association ($p = 5.1 \times 10^{-9}$, OR = 2.17, 95% CI 1.67–2.82). rs9263739, a proxy of HLA-B*52:01 ($r^2 = 0.94$), similarly showed a significant association ($p = 8.0 \times 10^{-9}$, OR = 2.30, 95% CI 1.72–3.07; Table 2) and in moderate LD with rs4947248 ($D' = 0.95$, $r^2 = 0.58$). Because rs4947248 did not show evidence of an independent association from rs9263739 in logistic regression analysis ($p = 0.04$), we assumed that the top association in the HLA locus was attributable to HLA-B*52:01. rs6871626 in the *IL12B* region on chromosome 5 also showed a significant association ($p = 1.8 \times 10^{-7}$, OR = 1.90, 95% CI 1.49–2.42; Table 2 and Figure 2A). Four other loci showed suggestive associations in our study ($p < 5.0 \times 10^{-5}$; Table 2). No departure from HWE was observed for these six SNPs ($p \geq 0.041$).

A replication study was performed with the use of DNA samples from 212 cases and 1,322 controls. The six SNPs with p values less than 5.0×10^{-5} in the genome scanning were genotyped in the replication study. rs9263739 was selected as a representative of the associations in the HLA locus. As a result, the significant associations of TAK with rs6871626 and rs665268 in the *MLX* (MAX dimerization protein [MIM 602976]) region on chromosome 17 as well as rs9263739 were replicated ($p = 1.1 \times 10^{-7}$, 0.0032, and 6.0×10^{-15}, respectively; Table 2, Figures 2A

Table 2. Results of Association Studies for TAK Susceptibility

						Genome Scan			Replication			Meta-analysis	
SNP	Chr	Position	Gene	Ref(A1)	Var(A2)[a]	Case A2freq	Cont A2freq	p	Case A2freq	Cont A2freq	p	p	OR (95% CI)
rs10934853	3	129521063	*EEFSEC*	A	C	0.59	0.45	1.3×10^{-5}	0.52	0.47	0.066	2.6×10^{-5}	1.40 (1.20–1.64)
rs6871626	5	158759370	*IL12B*	C	A	0.53	0.37	1.8×10^{-7}	0.53	0.39	1.1×10^{-7}	1.7×10^{-13}	1.75 (1.42–2.16)
rs9263739	6	31219335	*CCHCR1*	C	T	0.27	0.14	8.0×10^{-9}	0.30	0.14	6.0×10^{-15}	2.8×10^{-21}	2.44 (2.03–2.93)
rs1570843	6	84577239	*RIPPLY2*	C	T	0.62	0.50	4.6×10^{-5}	0.54	0.51	0.19	3.1×10^{-4}	1.34 (1.14–1.57)
rs12102203	15	49578851	*DMXL2*	G	A	0.64	0.49	3.8×10^{-6}	0.53	0.54	0.71	0.0081	1.24 (1.06–1.46)
rs665268	17	37975555	*MLX*	A	G	0.58	0.44	1.7×10^{-5}	0.49	0.42	0.0032	5.2×10^{-7}	1.50 (1.28–1.76)

Abbreviations are as follows: chr, chromosome; ref, reference allele; var, variant allele; CaseA2freq, variant allele frequency in cases; ContA2freq, variant allele frequency in controls; OR, odds ratio; CI, confidence interval. Positions are according to National Center for Biotechnology Information (NCBI) build 36.
[a]Risk alleles for TAK based on the results of the genome scanning are set as variant alleles.

and 2B). The suggestive association on chromosome 15 (Figures S1 and 1) was not replicated. Again, no departure from HWE was observed ($p \geq 0.11$).

A combined study in which the associations in the two studies were integrated by inverse-variance method demonstrated that rs6871626, rs665268, and rs9263739 showed significant associations ($p = 1.7 \times 10^{-13}$, 5.2×10^{-7}, and 2.8×10^{-21}; OR = 1.75, 1.50, and 2.44; 95% CI 1.42–2.16, 1.28–1.76, and 2.03–2.93, respectively; Table 2) satisfying the significance obtained by Bonferroni's correction. rs6871626 and rs9263739 satisfied the more stringent, widely accepted genome-wide significance ($p = 5.0 \times 10^{-8}$).

Because it was suggested that genetic components had influence on the manifestations of the disease,[17] we analyzed whether the variant of the *IL12B* region had clinical effects on the disease course or severity. Age at onset was not associated with rs6871626 ($p = 0.36$), whereas a significant association between rs6871626 and development of AR was observed in a recessive model ($p = 0.0046$; Figure 3A). Focusing on the cases with AR, a significant association between rs6871626 and severity of AR was observed in the recessive model ($p = 0.0018$; Figure 3B). Risk allele of rs6871626 (A allele) also demonstrated a significant association with increased level of time-averaged CRP, which was a representative marker of the disease activity ($p = 0.021$; Figure 3C). The association between rs6871626 and CRP levels was independent from rs3093059 in the *CRP* region ($p = 0.029$), which showed the strongest association with circulating CRP levels in Japanese.[19] These associations between rs6871626 and clinical manifestations were independent from rs9263739 (conditioned p value of rs6871626 ≤ 0.020). Although rs665268 also demonstrated a significant association with development of AR in a dominant model ($p = 0.0089$; Figure S2A), the association was not significant in condition with rs9263739 ($p = 0.080$). No significant associations were observed between rs665268 and other clinical phenotypes (Figures S2B and S2C).

Next, we investigated the interaction between the *IL12B* and *HLA-B* loci to TAK susceptibility. The risk of TAK in the population positive for both rs6871626 A allele and rs9263739 T allele surpassed the product and sum of the risk in those who were positive for either rs6871626 A allele or rs9263739 T allele alone (Figure 4). The analysis revealed that those who were positive for both had OR of 6.00 (95% CI 4.22–8.55), whereas those who were positive for either rs9263739 T allele or rs6871626 A allele showed OR of 1.80 (95% CI 1.11–2.93) or 1.74 (95% CI 1.23–2.47), respectively. Interaction measures revealed RER of 3.46 ($p = 1.4 \times 10^{-5}$, 95% CI 1.90–5.01), AP of 0.58 ($p = 1.0 \times 10^{-12}$, 95% CI 0.42–0.73), and SI of 3.24 ($p = 0.00028$, 95% CI 1.72–6.11). This significant interaction between *IL12B* and *HLA-B* on TAK susceptibility could be observed in both studies (Table 3). The synergistic interaction effects between rs6871626 and rs9263739 were not evident in the clinical manifestations associated with rs6871626 (Figure S3). When we analyzed the interaction between the *MLX* and *HLA-B* regions, we observed suggestive interaction with RER of 1.73, AP of 0.43, and SI of 2.29 ($p \leq 0.027$; Figure S4 and Table S1). The associations between the interaction and clinical manifestations were not significant (Figure S5).

IL12B encodes a common subunit of the IL12 and IL23 protein, known as p40. Because previous studies showed that the *IL23R/IL12RB2* (MIM 607562/601642) region was associated with Behçet disease[20] (MIM 109650), another connective tissue disease where vasculitis is involved in its pathology, we investigated this region for the possible associations in the current study. As a result, no suggestive association was found, either in our study or in the imputed results (Figure S6).

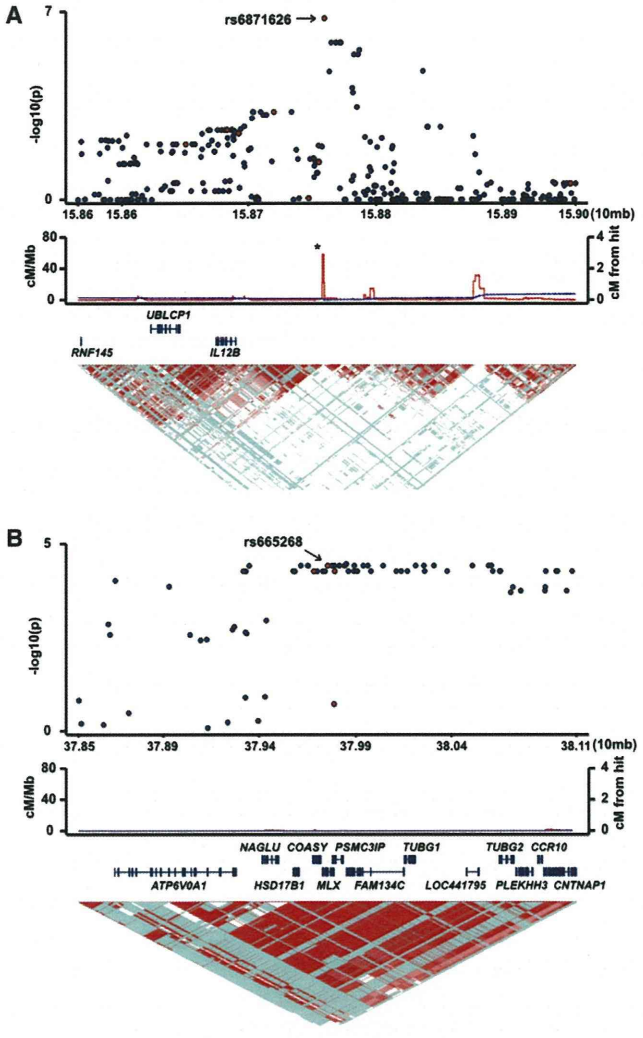


Figure 2. Associations of the *IL12B* and *MLX* Regions with the Susceptibility to TAK

Associations of SNPs in the (A) *IL12B* and (B) *MLX* regions in the genome scanning are plotted according to the position of the markers. Red circles indicate results of the current genome scanning. Blue circles indicate results of the imputation analysis based on the current results. The middle panel indicates recombination rates. The lower panel indicates LD of markers. Asterisk indicates a recombination hot spot in the *IL12B* region.

Discussion

This study provides a convincing evidence of associations between non-HLA genes and TAK susceptibility along with a synergistic role of susceptibility genes to TAK. The lack of evidence for associations of non-HLA genes with TAK so far is attributable to the lack of GWASs of TAK performed to date. Low prevalence of this disease had made it difficult to collect DNA samples to obtain sufficient power to detect susceptibility genes and perform a GWAS. Previous studies have revealed that the *IL12B* region was associated with a wide variety of autoimmune disorders and infectious diseases, including psoriasis[21–23] (MIM 177900), ankylosing spondylitis[24] (MIM 106300), Crohn disease[25] (CD [MIM 266600]), ulcerative colitis[26] (UC [MIM 191390]), and leprosy[27] (MIM 609888). rs6871626 showed a significant association with UC and leprosy over the genome-wide significance. Notably, rs6871626 A allele is susceptible to UC but protective against leprosy. A previous study from Turkey reported a suggestive association of TAK with rs3212227 in the 3′ UTR of the *IL12B* region.[28] rs3212227 is not in strong LD with rs6871626 in the Japanese population ($r^2 = 0.11$) and in Europeans ($r^2 = 0.06$) because of a recombination hot spot adjacent to rs6871626 (Figure 2A). In fact, an imputed association of rs3212227 with TAK in the current study resulted in only a suggestive association ($p = 0.0027$). There is a possibility that rs6871626 was responsible for the suggestive association between rs3212227 and TAK reported in the Turkish population. The association between gene expression and SNPs in the *IL12B* region appears to be complicated and inconsistent across different studies. rs3212227 in the 3′ UTR and rs17860508, an ins/del polymorphism in the promoter region of *IL12B*, were shown to have potential effects on the gene expression.[29,30] However, the previous studies showed that the association patterns varied according to the cell type and the protocol used for stimulation.[31–33] No previous report analyzed the effects of rs6871626 on the gene expression of *IL12B*. Although our in silico analysis failed to show the effects of rs6871626 on *IL12B* expression (data not shown, see Subjects and Methods), specific cell types or stimulus could lead to a significant association. Because a recent study showed that a haplotype of SNPs in the *IL12B* region could influence the gene and protein expression of *IL12B*,[22] a combination of rs6871626 and other SNPs in the *IL12B* region might lead to consistent results.

The associations between rs6871626 and clinical manifestations of TAK suggest the fundamental effects of IL-12p40 protein on TAK progression as well as TAK onset. We found that HLA-B*52:01 was associated with AR as reported previously ($p = 0.00014$). This finding supported the accuracy of our data. Although the risk allele of rs6871626 was associated with a significant dose-dependent increase in risk and severity of AR and in circulating CRP levels ($p = 0.013$, 0.030, and 0.023, respectively), these associations were more evident in a recessive manner. This raises a possibility that those who are homozygous for rs6871626 have strong disease activity that exceeds the additive disease activity of cases with single risk alleles, leading to severe destruction of aortic valve. Genetic variations in *IL12B* are known to influence the risk of psoriasis[21–23] and CD.[25] Because ustekinumab, a monoclonal antibody against IL-12p40, is an effective treatment for both diseases,[34,35] our findings would raise a possibility of its therapeutic use for TAK by targeting the IL-12/23 pathway. A previous study reported that the level of IL-12 protein was increased in TAK cases compared to healthy populations,[36] whereas there have been no reports addressing the circulating levels of IL-23 in TAK cases. IL-12 directly leads to type 1 helper T cell proliferation[37] and IL-23 upregulates IL-17 production and

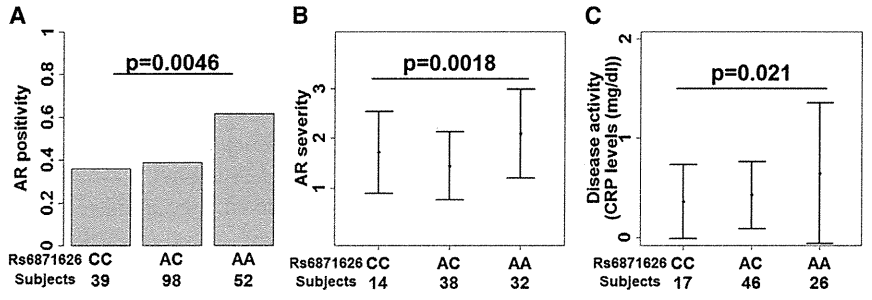


Figure 3. Associations between rs6871626 Genotypes and Clinical Manifestations of TAK

An association between rs6871626 genotypes and (A) development of AR, (B) severity of AR, and (C) time-averaged CRP levels in TAK cases. The p value was calculated by (A) logistic regression analysis, (B) linear regression analysis, and (C) linear regression analysis with time-averaged dosage of prednisolone as covariate. The recessive model is applied to all calculations. Severity of 1 to 3 in AR corresponds to mild, moderate, and severe, respectively. Mean ± SD are indicated for (B) and (C).

supports survival of activated Th17 cells.[38] Further analyses addressing circulating T cells in individuals with TAK or cell types infiltrating the artery specimen obtained from cases would provide clues to specify a critical pathway in TAK pathology.

The synergistic effect between rs6871626 and HLA-B*52:01 was notable. Those carrying both risk alleles had OR of 6.00 in comparison with those not carrying any risk alleles. Combination of rs6871626 and HLA-B*52:01 showed tendency of severe clinical phenotypes. Thus, we assume that increase of subjects and extraction of subjects who are homozygous for rs6871626 and positive for HLA-B*52:01 would provide evidence for significant effects of the combination on the disease phenotypes. The synergistic effect of these two loci raises a possibility that immune-related cells that recognize a yet-to-be-determined antigen through HLA-B*52:01 can be overactivated by IL-12/23 whose expression or function is modulated by rs6871626. In vitro analysis of immune-related cells from cases with TAK or healthy individuals would provide functional evidence of this synergistic role in the TAK pathogenesis.

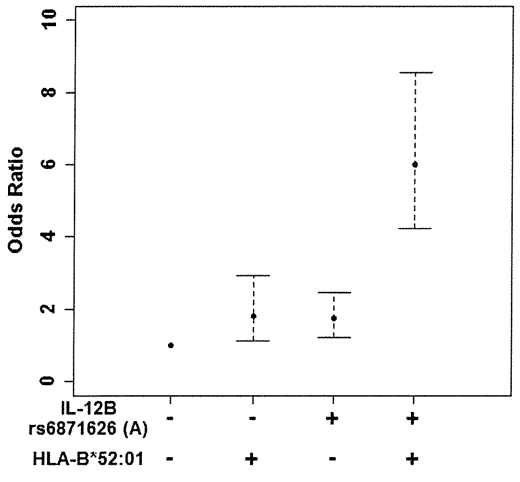


Figure 4. A Synergistic Effect between *IL12B* and HLA-B*52:01 on TAK Susceptibility

ORs are shown for the four strata of subjects according to combination of rs6871626 and rs9263739 genotypes. Those who are negative for rs9263739 T allele, a proxy of HLA-B*52:01, and rs6871626 A allele are used as reference. ORs and 95% CI are indicated.

rs665268 is a missense mutation of *MLX* that alters the 139th glutamine to arginine (Gln139Arg). MLX is a member of the basic helix-loop-helix leucine zipper (bHLH-Zip) transcription factor family and regulates gene expression by forming heterodimers with Mad protein.[39] The 17q21 region, whose associations with other autoimmune diseases including psoriasis[40] and CD[41] were shown, contains a number of genes including immune-related genes and polymorphisms that are in strong LD with each other (Figure 2B), so the corresponding gene to TAK susceptibility was inconclusive. Because risk allele frequency of rs665268 is comparable to that of rs6871626, the lack of associations between rs665268 and clinical manifestations and the weaker interaction between rs665268 and HLA-B*52:01 compared to rs6871626 might be a reflection of the milder effect of rs665268 on TAK progression. No interaction was observed between rs665268 and rs6871626 (data not shown).

We set the relatively low cut-off value of imputation score (Rsq $\geq$ 0.3) in the imputation analysis to increase sensitivity at the expense of specificity, but we failed to find other candidates of susceptibility loci. Another imputation analysis based on the data from the 1000 Genomes Project[42] revealed the same signals as the current study (data not shown). However, because the array used in the current study focused on SNPs in exons or nearby genes, it did not fully cover the whole genome with dense markers even in imputation analysis. There is a possibility that other SNPs not tagged by the markers on the array are associated with TAK. When the associations in the HLA locus were conditioned by rs9263739 or rs4947248, the most significantly associated SNPs, suggestive association signals in this locus could still be observed (the smallest p value = 5.5×10^{-5}, data not shown). The use of arrays with denser markers especially in intergene regions and using an increased number of cases could lead to the discovery of other susceptibility regions or independent associations in the HLA locus. Considering that both of the non-HLA susceptibility loci to TAK found in the current study are also associated with psoriasis and inflammatory bowel diseases, further analysis of TAK susceptibility genes would reveal other overlapping loci and common autoimmune mechanisms between TAK and other autoimmune diseases. It is feasible to

Table 3. Synergistic Effects between *IL12B* and HLA-B*52:01 in Each Study

	RERI		AP		SI	
Study	**(95% CI)**	**p**	**(95% CI)**	**p**	**(95% CI)**	**p**
Genome scanning	2.90 (0.60–5.20)	0.014	0.50 (0.23–0.78)	0.00034	2.57 (1.08–6.09)	0.032
Replication study	3.87 (1.70–6.05)	0.00049	0.62 (0.42–0.81)	4.7×10^{-10}	3.76 (1.51–9.32)	0.0043
Combined study	3.46 (1.90–5.02)	1.4×10^{-5}	0.58 (0.42–0.73)	1.0×10^{-12}	3.24 (1.72–6.11)	0.00028

Abbreviations are as follows: RERI, relative excess risk; AP, attributable proportion; SI, synergy index; CI, confidence interval.

analyze whether these two loci are associated with TAK and whether the interactions are observed in other populations.

Taken together, the current study identified two susceptibility genes to TAK and provided evidence of a common immunological pathway exerted by the *IL12B* region that is involved in the etiology of TAK and other autoimmune disorders and of its synergistic role with HLA in the susceptibility to TAK.

Supplemental Data

Supplemental Data include six figures and one table and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

We'd like to thank all the individuals with TAK who gave their blood samples and medical staffs to help us for this study; Miki Kokubo for her excellent technique to extract DNA; Kayo Umemoto for coordination of meetings to obtain blood samples; and Masashi Akizuki for his help to collect DNA samples. This study is supported by Kyoto University Step-up grant, Grants-in-aid from Research on rare and intractable diseases, the Ministry of Health, Labor, and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, grants from SENSHIN Medical Research Foundation (to T. Matsumura), and the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program) from the Japan Society for the Promotion of Science (to R.N.).

Received: March 23, 2013
Revised: May 18, 2013
Accepted: May 28, 2013
Published: July 3, 2013

Web Resources

The URLs for data presented herein are as follows:

Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/geo/
Genevar (Gene Expression Variation), http://www.sanger.ac.uk/resources/software/genevar/
International HapMap Project, http://hapmap.ncbi.nlm.nih.gov/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/
R statistical software, http://www.r-project.org/

References

1. Isobe, M. (2013). Takayasu arteritis revisited: Current diagnosis and treatment. Int. J. Cardiol. Published online February 13, 2013. http://dx.doi.org/10.1016/j.ijcard.2013.01.022.
2. Dreyer, L., Faurschou, M., and Baslund, B. (2011). A population-based study of Takayasu's arteritis in eastern Denmark. Clin. Exp. Rheumatol. *29*(1, Suppl 64), S40–S42.
3. Jennette, J.C., Falk, R.J., Bacon, P.A., Basu, N., Cid, M.C., Ferrario, F., Flores-Suarez, L.F., Gross, W.L., Guillevin, L., Hagen, E.C., et al. (2013). 2012 revised International Chapel Hill Consensus Conference Nomenclature of Vasculitides. Arthritis Rheum. *65*, 1–11.
4. Yoshida, M., Kimura, A., Katsuragi, K., Numano, F., and Sasazuki, T. (1993). DNA typing of HLA-B gene in Takayasu's arteritis. Tissue Antigens *42*, 87–90.
5. Yajima, M., Numano, F., Park, Y.B., and Sagar, S. (1994). Comparative studies of patients with Takayasu arteritis in Japan, Korea and India—comparison of clinical manifestations, angiography and HLA-B antigen. Jpn. Circ. J. *58*, 9–14.
6. Sahin, Z., Bıcakcıgil, M., Aksu, K., Kamali, S., Akar, S., Onen, F., Karadag, O., Ozbalkan, Z., Ates, A., Ozer, H.T., et al.; Turkish Takayasu Study Group. (2012). Takayasu's arteritis is associated with HLA-B*52, but not with HLA-B*51, in Turkey. Arthritis Res. Ther. *14*, R27.
7. Vargas-Alarcón, G., Flores-Domínguez, C., Hernández-Pacheco, G., Zuñiga, J., Gamboa, R., Soto, M.E., Granados, J., and Reyes, P.A. (2001). Immunogenetics and clinical aspects of Takayasu's arteritis patients in a Mexican Mestizo population. Clin. Exp. Rheumatol. *19*, 439–443.
8. Arend, W.P., Michel, B.A., Bloch, D.A., Hunder, G.G., Calabrese, L.H., Edworthy, S.M., Fauci, A.S., Leavitt, R.Y., Lie, J.T., Lightfoot, R.W., Jr., et al. (1990). The American College of Rheumatology 1990 criteria for the classification of Takayasu arteritis. Arthritis Rheum. *33*, 1129–1134.
9. JCS Joint Working Group; Japanese Circulation Society. (2011). Guideline for management of vasculitis syndrome (JCS 2008). Circ. J. *75*, 474–503.
10. Yoshimura, K., Nakayama, T., Sekine, A., Matsuda, F., Kosugi, S., Sugino, Y., Yoshimura, K., and Ogawa, O.; Nagahama Cohort Research Group. (2013). Prevalence of postmicturition urinary incontinence in Japanese men: Comparison with other types of incontinence. Int. J. Urol. Published online January 10, 2013. http://dx.doi.org/10.1111/iju.12074.
11. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575.

12. Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet. Epidemiol. *34*, 816–834.
13. Andersson, T., Alfredsson, L., Källberg, H., Zdravkovic, S., and Ahlbom, A. (2005). Calculating measures of biological interaction. Eur. J. Epidemiol. *20*, 575–579.
14. Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., et al. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science *315*, 848–853.
15. Yang, T.P., Beazley, C., Montgomery, S.B., Dimas, A.S., Gutierrez-Arcelus, M., Stranger, B.E., Deloukas, P., and Dermitzakis, E.T. (2010). Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. Bioinformatics *26*, 2474–2476.
16. Andrews, J., and Mason, J.C. (2007). Takayasu's arteritis—recent advances in imaging offer promise. Rheumatology (Oxford) *46*, 6–15.
17. Kitamura, H., Kobayashi, Y., Kimura, A., and Numano, F. (1998). Association of clinical manifestations with HLA-B alleles in Takayasu arteritis. Int. J. Cardiol. *66(Suppl 1)*, S121–S126.
18. Warnes, C.A., Williams, R.G., Bashore, T.M., Child, J.S., Connolly, H.M., Dearani, J.A., del Nido, P., Fasules, J.W., Graham, T.P., Jr., Hijazi, Z.M., et al.; American College of Cardiology; American Heart Association Task Force on Practice Guidelines (Writing Committee to Develop Guidelines on the Management of Adults With Congenital Heart Disease); American Society of Echocardiography; Heart Rhythm Society; International Society for Adult Congenital Heart Disease; Society for Cardiovascular Angiography and Interventions; Society of Thoracic Surgeons. (2008). ACC/AHA 2008 guidelines for the management of adults with congenital heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Develop Guidelines on the Management of Adults With Congenital Heart Disease). Developed in Collaboration With the American Society of Echocardiography, Heart Rhythm Society, International Society for Adult Congenital Heart Disease, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J. Am. Coll. Cardiol. *52*, e143–e263.
19. Okada, Y., Takahashi, A., Ohmiya, H., Kumasaka, N., Kamatani, Y., Hosono, N., Tsunoda, T., Matsuda, K., Tanaka, T., Kubo, M., et al. (2011). Genome-wide association study for C-reactive protein levels identified pleiotropic associations in the IL6 locus. Hum. Mol. Genet. *20*, 1224–1231.
20. Mizuki, N., Meguro, A., Ota, M., Ohno, S., Shiota, T., Kawagoe, T., Ito, N., Kera, J., Okada, E., Yatsu, K., et al. (2010). Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. Nat. Genet. *42*, 703–706.
21. Tsunemi, Y., Saeki, H., Nakamura, K., Sekiya, T., Hirai, K., Fujita, H., Asano, N., Kishimoto, M., Tanida, Y., Kakinuma, T., et al. (2002). Interleukin-12 p40 gene (IL12B) 3′-untranslated region polymorphism is associated with susceptibility to atopic dermatitis and psoriasis vulgaris. J. Dermatol. Sci. *30*, 161–166.
22. Johnston, A., Xing, X., Swindell, W.R., Kochkodan, J., Riblett, M., Nair, R.P., Stuart, P.E., Ding, J., Voorhees, J.J., Elder, J.T., and Gudjonsson, J.E. (2013). Susceptibility-associated genetic variation at IL12B enhances Th1 polarization in psoriasis. Hum. Mol. Genet. *22*, 1807–1815.
23. Cargill, M., Schrodi, S.J., Chang, M., Garcia, V.E., Brandon, R., Callis, K.P., Matsunami, N., Ardlie, K.G., Civello, D., Catanese, J.J., et al. (2007). A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. Am. J. Hum. Genet. *80*, 273–290.
24. Evans, D.M., Spencer, C.C., Pointon, J.J., Su, Z., Harvey, D., Kochan, G., Oppermann, U., Dilthey, A., Pirinen, M., Stone, M.A., et al.; Spondyloarthritis Research Consortium of Canada (SPARCC); Australo-Anglo-American Spondyloarthritis Consortium (TASC); Wellcome Trust Case Control Consortium 2 (WTCCC2). (2011). Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat. Genet. *43*, 761–767.
25. Parkes, M., Barrett, J.C., Prescott, N.J., Tremelling, M., Anderson, C.A., Fisher, S.A., Roberts, R.G., Nimmo, E.R., Cummings, F.R., Soars, D., et al.; Wellcome Trust Case Control Consortium. (2007). Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat. Genet. *39*, 830–832.
26. Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A., et al. (2011). Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat. Genet. *43*, 246–252.
27. Liu, H., Irwanto, A., Tian, H., Fu, X., Yu, Y., Yu, G., Low, H., Chu, T., Li, Y., Shi, B., et al. (2012). Identification of IL18-RAP/IL18R1 and IL12B as leprosy risk genes demonstrates shared pathogenesis between inflammation and infectious diseases. Am. J. Hum. Genet. *91*, 935–941.
28. Saruhan-Direskeneli, G., Biçakçigil, M., Yilmaz, V., Kamali, S., Aksu, K., Fresko, I., Akkoç, N., Kiraz, S., Ozer, H.T., Tunç, E., et al.; Rheumatology Education and Research Society Vasculitis Study Group. (2006). Interleukin (IL)-12, IL-2, and IL-6 gene polymorphisms in Takayasu's arteritis from Turkey. Hum. Immunol. *67*, 735–740.
29. Morahan, G., Huang, D., Wu, M., Holt, B.J., White, G.P., Kendall, G.E., Sly, P.D., and Holt, P.G. (2002). Association of IL12B promoter polymorphism with severity of atopic and non-atopic asthma in children. Lancet *360*, 455–459.
30. Stanilova, S., and Miteva, L. (2005). Taq-I polymorphism in 3′UTR of the IL-12B and association with IL-12p40 production from human PBMC. Genes Immun. *6*, 364–366.
31. Basu, M., Das, T., Ghosh, A., Majumder, S., Maji, A.K., Kanjilal, S.D., Mukhopadhyay, I., Roychowdhury, S., Banerjee, S., and Sengupta, S. (2012). Gene-gene interaction and functional impact of polymorphisms on innate immune genes in controlling *Plasmodium falciparum* blood infection level. PLoS ONE *7*, e46441.
32. Yanagihori, H., Oyama, N., Nakamura, K., Mizuki, N., Oguma, K., and Kaneko, F. (2006). Role of IL-12B promoter polymorphism in Adamantiades-Behcet's disease susceptibility: An involvement of Th1 immunoreactivity against *Streptococcus Sanguinis* antigen. J. Invest. Dermatol. *126*, 1534–1540.
33. Yilmaz, V., Yentür, S.P., and Saruhan-Direskeneli, G. (2005). IL-12 and IL-10 polymorphisms and their effects on cytokine production. Cytokine *30*, 188–194.
34. Yeilding, N., Szapary, P., Brodmerkel, C., Benson, J., Plotnick, M., Zhou, H., Goyal, K., Schenkel, B., Giles-Komar, J., Mascelli, M.A., and Guzzo, C. (2012). Development of the IL-12/23

antagonist ustekinumab in psoriasis: past, present, and future perspectives—an update. Ann. N Y Acad. Sci. *1263*, 1–12.

35. Sandborn, W.J., Gasink, C., Gao, L.L., Blank, M.A., Johanns, J., Guzzo, C., Sands, B.E., Hanauer, S.B., Targan, S., Rutgeerts, P., et al.; CERTIFI Study Group. (2012). Ustekinumab induction and maintenance therapy in refractory Crohn's disease. N. Engl. J. Med. *367*, 1519–1528.
36. Verma, D.K., Tripathy, N.K., Verma, N.S., and Tiwari, S. (2005). Interleukin 12 in Takayasu's arteritis: plasma concentrations and relationship with disease activity. J. Rheumatol. *32*, 2361–2363.
37. Murphy, K.M., and Reiner, S.L. (2002). The lineage decisions of helper T cells. Nat. Rev. Immunol. *2*, 933–944.
38. Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T.B., Oukka, M., Weiner, H.L., and Kuchroo, V.K. (2006). Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. Nature *441*, 235–238.
39. Cairo, S., Merla, G., Urbinati, F., Ballabio, A., and Reymond, A. (2001). WBSCR14, a gene mapping to the Williams-Beuren syndrome deleted region, is a new member of the Mlx transcription factor network. Hum. Mol. Genet. *10*, 617–627.
40. Tsoi, L.C., Spain, S.L., Knight, J., Ellinghaus, E., Stuart, P.E., Capon, F., Ding, J., Li, Y., Tejasvi, T., Gudjonsson, J.E., et al.; Collaborative Association Study of Psoriasis (CASP); Genetic Analysis of Psoriasis Consortium; Psoriasis Association Genetics Extension; Wellcome Trust Case Control Consortium 2. (2012). Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat. Genet. *44*, 1341–1348.
41. Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat. Genet. *42*, 1118–1125.
42. Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A.; 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061–1073.

OPEN

Crowdsourcing genetic prediction of clinical utility in the Rheumatoid Arthritis Responder Challenge

To the Editor:

The ability to translate large-scale genetics and genomics data into biological knowledge has not kept pace with our ability to generate these data sets. As a consequence, a major bottleneck in biomedical research has become access to data within a computational workspace that allows for robust, collaborative analyses. One innovative solution is to bring together scientific data, code, tools and disease models into an open commons or workspace, for example, the Synapse platform of Sage Bionetworks[1]. This environment allows for real-time sharing of large genomic data sets, continuous peer review and rapid learning within a system constructed to provide data access in a manner aligned with the informed consent provided by patients and research participants.

This crowdsourcing approach has been used to predict breast cancer survival from clinical and omics data[2] and was suggested as a way to find new drugs[3] by soliciting contributions from a large online community collaborating or competing to answer an inherently difficult but important question[4]. Researchers initiating an open challenge invite solutions but also incentivize the process by offering new data, a process in which the participants' methods can be assessed by testing their predictions against previously unseen data sets. This year, Sage and DREAM (Dialogue for Reverse Engineering Assessments and Methods) are running four open challenges (http://www.sagebase.org/challenges-overview/2013-dream-challenges/).

Here we announce the challenge to develop genetic predictors of response to immunosuppressive therapy in a common autoimmune disease, rheumatoid arthritis (RA). Disease-modifying antirheumatic drugs such as those that block the inflammatory cytokine tumor necrosis factor-α (known as anti-TNF therapy) are not effective in all patients with RA, with up to one-third of such patients failing to enter clinical remission after a standard course of therapy[5]. Moreover, the biological mechanisms underlying this failure are unknown, limiting the development of clinical biomarkers to guide either this therapy or the development of new drugs to target refractory cases.

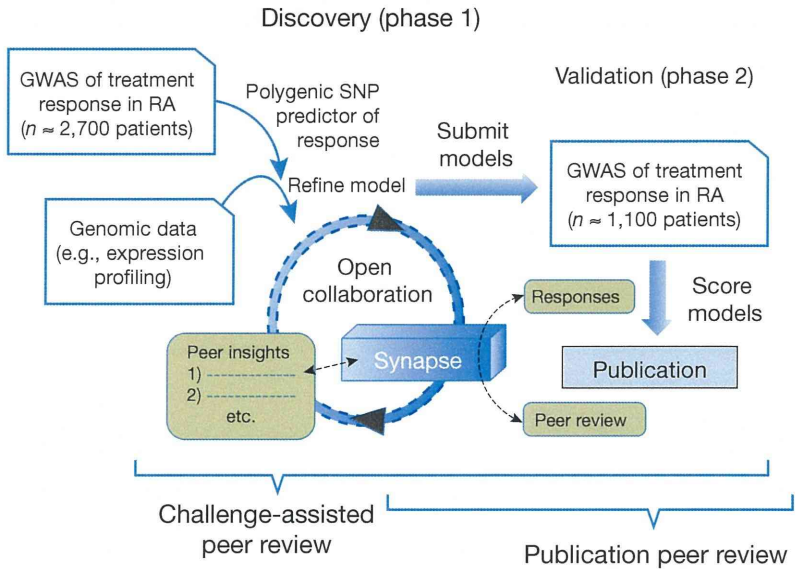


Figure 1 Overview of the Rheumatoid Arthritis Responder Challenge. There are two phases to the challenge. In phase 1 (discovery), analysts build genetic models of response to anti-TNF therapy using SNP data from a GWAS of ~2,700 patients with RA. To facilitate model building, additional genomic data will be made available. In a model of open collaboration, participants will use Synapse to post code, share insights and engage in rapid learning prepublication. In phase 2 (validation), models will be posted, tested and scored in an independent GWAS data set of ~1,100 patients with RA treated with anti-TNF therapy. To complement challenge-assisted peer review (which occurs in both the discovery and validation phases), conventional peer review will have access to Synapse to understand the iterative process of model building. Synapse will allow study investigators to respond to peer-review critiques and resubmit versions of their models and studies.

The Rheumatoid Arthritis Responder Challenge is for teams to build the best genetic predictor of response to anti-TNF therapy. There are two phases to the challenge: discovery and validation (**Fig. 1**). In the discovery phase, teams will utilize genomic data sets—several of which will be generated for the purposes of this challenge—and a variety of analytical methods to build predictive polygenic models of treatment response. We recently published a genome-wide association study (GWAS) in ~2,700 patients with RA treated with anti-TNF therapy[6]. Our GWAS data indicate that the genetic architecture of the anti-TNF response is probably highly polygenic, similar to what has been observed for other complex traits, such as risk of RA[7]. Importantly, our challenge will incorporate a new GWAS data set, which will be used in the validation phase, in which models built in the discovery phase are tested. The data set of ~1,100 patients with RA treated with anti-TNF therapy will be made available though

a public-private partnership between the Consortium of Rheumatology Researchers of North America, Inc. (CORRONA) and the Pharmacogenomics Research Network (PGRN) sponsored by the National Institute of General Medical Sciences (NIGMS) and the US National Institutes of Health (NIH).

A unique component of our Rheumatoid Arthritis Responder Challenge is the diversity of participation across a number of groups from academic institutions, private foundations and for-profit companies. In addition to support from CORRONA and PGRN, we received funding from pharmaceutical companies (see complete list on our website; link below) and a private foundation (the Arthritis Foundation) to support the public commons. We also received support from the Arthritis Internet Registry (AIR) and the Broad Institute to generate new genomic data sets, as well as in-kind support from a large number of academic collaborators from across the world to make GWAS data available in the discovery phase. We anticipate that a winning classifier could enable a follow-on prospective clinical trial within the group of appropriately consented patients in AIR.

Through Synapse, analysts who are inclined to establish collaborations will have the opportunity to see in real time the models that others are using so that each team can learn from the others (**Fig. 1**). A leaderboard will show the relative performance ranking of the different teams on the basis of a crossvalidation strategy designed to minimize overfitting. During the discovery phase, teams that choose to collaborate with each other will have the opportunity to check each other's algorithms for readability, speed and reproducibility. Then, during the validation phase, each team will submit computer code, which the Sage-DREAM team (http://www.sagebase.org/) will test in Synapse to establish whether it runs as expected to predict if a subject is an anti-TNF therapy responder or nonresponder on the basis of the GWAS data. Predefined performance metrics will be used to objectively determine the accuracy of the predictions, their statistical significance and the final performance ranking of the participating teams. The team that develops the most highly predictive model will be deemed the 'winner', with precise attribution of contributor roles going to all members of teams that contributed to building the final consensus model.

The best-performing models, therefore, will have passed a test of performance that is outside the realm of, and complements, traditional peer review. Indeed, this stringent test of method performance can be used as an enhanced way of publication vetting, what we call 'challenge-assisted peer review'. Traditional peer review is essential for ensuring the clarity, originality, contextualization and logical thread of a discrete set of work that is ready to be used by researchers in the form of a published article. However, the complexity of working with omics data—entailing multiple analytical decisions, computational simulations and statistical calculations—means that referees are challenged to follow and check the components of even a traditional research paper. In our Rheumatoid Arthritis Responder Challenge, we will explore the feasibility of enhancing the reliability and transparency of conventional peer review in partnership with *Nature Genetics*. This can be achieved if the referees and authors of the paper reporting on the best-performing methods in the challenge are willing to leave their comments openly (yet anonymously) on the Synapse platform (**Fig. 1**). We anticipate that the challenge-based assessment of accuracy will provide an objective metric of performance and a comparison with state-of-the-art analytical methodologies that will greatly enhance the task of refereeing a body of work with more quality control than is currently provided by conventional peer review.

In conclusion, we believe that the Rheumatoid Arthritis Responder Challenge is an apt use of crowdsourcing in human genetics to gain insight into clinical prediction and disease biology. Details of the challenge, including the rules by which the models will be judged, can be found at https://synapse.prod.sagebase.org/#!Synapse:syn1734172.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Robert M Plenge[1,2], Jeffrey D Greenberg[3], Lara M Mangravite[4], Jonathan M J Derry[4], Eli A Stahl[5], Marieke J H Coenen[6], Anne Barton[7], Leonid Padyukov[8], Lars Klareskog[8], Peter K Gregersen[9], Xavier Mariette[10], Larry W Moreland[11], S Louis Bridges Jr[12], Niek de Vries[13], Tom W J Huizinga[14], Henk-Jan Guchelaar[15], International Rheumatoid Arthritis Consortium (INTERACT)[16], Stephen H Friend[4,17] & Gustavo Stolovitzky[17,18].

[1]Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [2]Medical and Population Genetics Program, Chemical Biology Program, Broad Institute, Cambridge, Massachusetts, USA. [3]Division of Rheumatology, New York University School of Medicine, New York, New York, USA. [4]Sage Bionetworks, Seattle, Washington, USA. [5]Division of Psychiatric Genomics and Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mt. Sinai, New York, New York, USA. [6]Department of Human Genetics, Radboud University Medical Centre, Nijmegen, The Netherlands. [7]Arthritis Research UK Epidemiology Unit, Musculoskeletal Research Group, University of Manchester, Manchester Academic Health Sciences Centre, Manchester, UK. [8]Rheumatology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden. [9]The Feinstein Institute for Medical Research, North Shore–Long Island Jewish Health System, Manhasset, New York, USA. [10]Université Paris-Sud, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Institut National de la Santé et de la Recherche Médicale (INSERM) U1012, Rheumatology, le Kremlin-Bicêtre, France. [11]Division of Rheumatology and Clinical Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [12]Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, USA. [13]Department of Clinical Immunology and Rheumatology, Academic Medical Center (AMC), University of Amsterdam, Amsterdam, The Netherlands. [14]Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands. [15]Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Leiden, The Netherlands. [16]A full list of members is provided at https://synapse.prod.sagebase.org/#!Synapse:syn1734172. [17]Sage/DREAM Project, Sage Bionetworks, Seattle, Washington, USA. [18]IBM Computational Biology Center, T.J. Watson Research Center, Yorktown Heights, New York, USA.
e-mail: rplenge@partners.org.

1. Derry, J.M.J. *et al. Nat. Genet.* **44**, 127–130 (2012).
2. Margolin, A.A. *et al. Sci. Transl. Med.* **5**, 181re1 (2013).
3. Wadman, M. *Nature* **490**, 15 (2012).
4. Costello, J.C. & Stolovitzky, G. *Clin. Pharmacol. Ther.* **93**, 396–398 (2013)
5. McInnes, I.B. & Schett, G. *N. Engl. J. Med.* **365**, 2205–2219 (2011).
6. Cui, J. *et al. PLoS Genet.* published online, http://dx.doi.org/10.1371/journal.pgen.1003394 (28 March 2013).
7. Stahl, E.A. *et al. Nat. Genet.* **44**, 483–489 (2012).

PLOS | GENETICS

Genome-Wide Association Study and Gene Expression Analysis Identifies *CD84* as a Predictor of Response to Etanercept Therapy in Rheumatoid Arthritis

Jing Cui[1☯], Eli A. Stahl[1,2,3☯¤a], Saedis Saevarsdottir[4,5], Corinne Miceli[6,7], Dorothee Diogo[1,2,3], Gosia Trynka[1,2,3], Towfique Raj[2,3,8], Maša Umićević Mirkov[9], Helena Canhao[1,10,11], Katsunori Ikari[12], Chikashi Terao[13,14], Yukinori Okada[1,2,3], Sara Wedrén[4,5], Johan Askling[4,15], Hisashi Yamanaka[12], Shigeki Momohara[12], Atsuo Taniguchi[12], Koichiro Ohmura[13], Fumihiko Matsuda[14], Tsuneyo Mimori[13], Namrata Gupta[3], Manik Kuchroo[3,8], Ann W. Morgan[16], John D. Isaacs[17], Anthony G. Wilson[18], Kimme L. Hyrich[19], Marieke Herenius[20], Marieke E. Doorenspleet[20], Paul-Peter Tak[20¤b], J. Bart A. Crusius[21], Irene E. van der Horst-Bruinsma[22], Gert Jan Wolbink[23,24,25], Piet L. C. M. van Riel[9], Mart van de Laar[26], Henk-Jan Guchelaar[27], Nancy A. Shadick[1], Cornelia F. Allaart[28], Tom W. J. Huizinga[28], Rene E. M. Toes[28], Robert P. Kimberly[29], S. Louis Bridges Jr.[29], Lindsey A. Criswell[30], Larry W. Moreland[31], João Eurico Fonseca[10,11], Niek de Vries[20], Barbara E. Stranger[2,3], Philip L. De Jager[2,3,7], Soumya Raychaudhuri[1,2,3,32], Michael E. Weinblatt[1], Peter K. Gregersen[33], Xavier Mariette[6,7], Anne Barton[34], Leonid Padyukov[5], Marieke J. H. Coenen[9], Elizabeth W. Karlson[1], Robert M. Plenge[1,2,3]*

1 Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **3** Medical and Population Genetics Program, Chemical Biology Program, Broad Institute, Cambridge, Massachusetts, United States of America, **4** Rheumatology Unit, Department of Medicine, Karolinska Institutet and Karolinska University Hospital Solna, Stockholm, Sweden, **5** Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden, **6** Université Paris-Sud, Orsay, France, **7** APHP–Hôpital Bicêtre, INSERM U1012, Le Kremlin Bicêtre, Paris, France, **8** Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **9** Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, **10** Rheumatology Research Unit, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal, **11** Rheumatology Department, Santa Maria Hospital–CHLN, Lisbon, Portugal, **12** Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan, **13** Department of Rheumatology and Clinical Immunology, Kyoto University Graduate School of Medicine, Kyoto, Japan, **14** Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **15** Clinical Epidemiology Unit, Department of Medicine, Karolinska Institute/Karolinska University Hospital, Stockholm, Sweden, **16** NIHR–Leeds Musculoskeletal Biomedical Research Unit and Leeds Institute of Molecular Medicine, University of Leeds, Leeds, United Kingdom, **17** Musculoskeletal Research Group, Institute of Cellular Medicine, Newcastle Upon Tyne, United Kingdom, **18** Rheumatology Unit, Medical School, University of Sheffield, Sheffield, United Kingdom, **19** School of Translational Medicine, Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester, United Kingdom, **20** Department of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, Amsterdam, The Netherlands, **21** Laboratory of Immunogenetics, Department of Pathology, Vrije Universiteit Medical Center, Amsterdam, The Netherlands, **22** Department of Rheumatology, Vrije Universiteit University Medical Center, Amsterdam, The Netherlands, **23** Sanquin Research Landsteiner Laboratory, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, **24** School of Medicine and Biomedical Sciences, Sheffield University, Sheffield, United Kingdom, **25** Jan van Breemen Institute, Amsterdam, The Netherlands, **26** Arthritis Center Twente, University Twente and Medisch Spectrum Twente, Enschede, The Netherlands, **27** Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Leiden, The Netherlands, **28** Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands, **29** Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, United States of America, **30** Rosalind Russell Medical Research Center for Arthritis, Division of Rheumatology, Department of Medicine, University of California San Francisco, San Francisco, California, United States of America, **31** Division of Rheumatology and Clinical Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **32** NIHR Manchester Musculoskeletal Biomedical Research Unit, Central Manchester NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester, United Kingdom, **33** The Feinstein Institute for Medical Research, North Shore–Long Island Jewish Health System, Manhasset, New York, United States of America, **34** Arthritis Research UK Epidemiology Unit, Musculoskeletal Research Group, University of Manchester, Manchester Academic Health Sciences Centre, Manchester, United Kingdom

Abstract

Anti-tumor necrosis factor alpha (anti-TNF) biologic therapy is a widely used treatment for rheumatoid arthritis (RA). It is unknown why some RA patients fail to respond adequately to anti-TNF therapy, which limits the development of clinical biomarkers to predict response or new drugs to target refractory cases. To understand the biological basis of response to anti-TNF therapy, we conducted a genome-wide association study (GWAS) meta-analysis of more than 2 million common variants in 2,706 RA patients from 13 different collections. Patients were treated with one of three anti-TNF medications: etanercept (n = 733), infliximab (n = 894), or adalimumab (n = 1,071). We identified a SNP (rs6427528) at the *1q23* locus that was associated with change in disease activity score (ΔDAS) in the etanercept subset of patients ($P = 8 \times 10^{-8}$), but not in the infliximab or adalimumab subsets ($P > 0.05$). The SNP is predicted to disrupt transcription factor binding site motifs in the 3′ UTR of an immune-related gene, *CD84*, and the allele associated with better response to etanercept was associated with higher *CD84* gene expression in peripheral blood mononuclear cells ($P = 1 \times 10^{-11}$ in 228 non-RA patients and $P = 0.004$ in 132 RA patients). Consistent with the genetic findings, higher *CD84* gene expression correlated with lower cross-sectional DAS ($P = 0.02$, n = 210) and showed a non-significant trend for better ΔDAS in a subset of RA patients with gene expression data (n = 31, etanercept-treated). A small, multi-ethnic replication showed a non-significant trend towards an association among etanercept-treated RA

patients of Portuguese ancestry (n = 139, $P = 0.4$), but no association among patients of Japanese ancestry (n = 151, $P = 0.8$). Our study demonstrates that an allele associated with response to etanercept therapy is also associated with *CD84* gene expression, and further that *CD84* expression correlates with disease activity. These findings support a model in which *CD84* genotypes and/or expression may serve as a useful biomarker for response to etanercept treatment in RA patients of European ancestry.

Citation: Cui J, Stahl EA, Saevarsdottir S, Miceli C, Diogo D, et al. (2013) Genome-Wide Association Study and Gene Expression Analysis Identifies *CD84* as a Predictor of Response to Etanercept Therapy in Rheumatoid Arthritis. PLoS Genet 9(3): e1003394. doi:10.1371/journal.pgen.1003394

Editor: Alison Motsinger-Reif, North Carolina State University, United States of America

Received August 13, 2012; **Accepted** January 13, 2013; **Published** March 28, 2013

Funding: JC is supported by grants from ACR REF HPNIA award and the NIH (R01-AR059073, P60-AR047782, U01-GM092691, R01-AR049880). RMP is supported by grants from the NIH (R01-AR057108, R01-AR056768, U01-GM092691, R01-AR059648) and holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund. The American College of Rheumatology Research and Education Foundation provided funding to support this project. The eRA study was supported by R01 AI/AR47487. SS's work was supported by a clinical research fund from Stockholm County (ALF fund). LP's work was supported by the Swedish Rheumatism Association and the Swedish Medical Research Council. NdV and MED were sponsored by CTMM, the Center for Translational Molecular Medicine, and the Dutch Arthritis Foundation project TRACER (grant 04I-202). The Swedish studies (EIRA, Karolinska) were supported by grants from the Swedish Medical Research Council, the Stockholm County Council, the Swedish Council for Working Life and Social Research, King Gustaf V's 80-year foundation, the Swedish Rheumatism Association, the Swedish Foundation for Strategic Research, the Swedish COMBINE project, and the IMI funded BTCure project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rplenge@partners.org

¤a Current address: Division of Psychiatric Genomics, Mt. Sinai School of Medicine, New York, New York, United States of America
¤b Current address: GlaxoSmithKline, Stevenage, United Kingdom

☯ These authors contributed equally to this work.

Introduction

Rheumatoid arthritis (RA) is an autoimmune disease characterized by chronic inflammation of the synovial lining of the joint [1]. If left untreated, outcome varies from self-limited disease in a small proportion of RA patients to severe disease resulting in profound structural damage, excess morbidity and disability, and early mortality [2]. In the last twenty years, disease activity has been controlled in many patients by treatment with disease-modifying anti-rheumatic drugs (DMARDs), such as methotrexate, and the more recently developed biologic DMARDs that block inflammatory cytokines such as tumor necrosis factor-alpha (TNFa) [3]. Unfortunately, these medications are not effective in all RA patients, with up to one-third of patients failing to respond to any single DMARD [1–3]. Moreover, the biological mechanisms underlying treatment failure are unknown, which limits the development of clinical biomarkers to guide DMARD therapy or the development of new drugs to target refractory cases.

There are two classes of anti-TNF therapy: the TNF receptor fusion protein (etanercept), which acts as a soluble receptor to bind circulating cytokine and prevent TNF from binding to its cell surface receptor, and monoclonal antibodies that bind TNF (adalimumab, infliximab, certolizumab, and golimumab). There are undoubtedly shared mechanisms between the two drug classes (e.g., downstream signaling factors), as illustrated by similar effects on the change in inflammatory cytokines, complement activation, lymphocyte trafficking, and apoptosis [4,5,6]. Similarly, there are likely to be different biological factors that influence response: infliximab and adalimumab are approved for treatment of Crohn's disease; infliximab and adalimumab bind to transmembrane TNF on the surface of activated immune cells, whereas etanercept only binds soluble TNF [7]; and etanercept also binds a related molecule, lymphotoxin alpha (LTA), whereas infliximab/adalimumab do not [8].

Pharmacogenetics of response to anti-TNF therapy in RA remains in its early stages, with no single variant reaching an unambiguous level of statistical significance. Candidate gene studies suggest associations of TNFa or TNF receptor alleles, RA risk alleles or other SNPs with response to anti-TNF therapy [9,10,11]. Two GWAS in small sample sets (largest was 566 patients) have been performed, which identified loci with suggestive evidence for association [12,13]. Therefore, GWAS of large sample sizes may yet uncover genetic factors associated with response to anti-TNF therapy in RA, and larger cohorts enable separate analyses of the different types of anti-TNF drugs.

Here we report a GWAS of 2,706 samples with anti-TNF treatment response data collected from an international collaboration, including previously published GWAS data [12,13]. Our primary outcome measure was the change in disease activity score based on a joint count in 28 joints (DAS28) from baseline to 3–12 months after initiating anti-TNF therapy. Our secondary outcome measure was European League Against Rheumatism (EULAR) responder status [14,15], where patients are classified as EULAR good responders, moderate responders or non-responders based on follow up DAS28 after treatment and overall change in DAS28. We found a highly significant association for a variant that we also show is also a strong expression quantitative trait locus (eQTL) for the *CD84* gene. Our findings suggest that *CD84* genotype and/or expression may prove to be a biomarker for etanercept response in RA patients.

Results

Genome-wide association study

Clinical and GWAS data were compiled for 2,706 individuals of European ancestry from 13 collections as part of an international collaboration. Table 1 shows sample sizes, phenotypes and clinical variables for the four collections that were the units of analysis (additional details are shown in Table S1). Disease activity score based on a 28-joint count (DAS28) were collected at baseline and at one time point after anti-TNF therapy administration (mean 3.7 months, range 3–12 months). We defined our primary phenotype

Author Summary

There are no genetic predictors of response to one of the most widely used classes of drugs in the treatment of rheumatoid arthritis—biological modifiers of the inflammatory cytokine tumor necrosis factor-alpha (or anti-TNF therapy). To identify genetic predictors, we performed the largest genome-wide association study (GWAS) to date as part of an international collaboration. In our study, which included 2,706 RA patients treated with one of three anti-TNF drugs, the most significant finding was restricted to RA patients treated with etanercept ($P=8\times10^{-8}$), a drug that acts as a soluble receptor to bind circulating cytokine and prevents TNF from binding to its cell surface receptor. The associated variant influences expression of a nearby immune-related gene, *CD84*, whose expression is correlated with disease activity in RA patients. Together, our data support a model in which genomic factors related to *CD84* expression serve as a predictor of disease activity and response to etanercept therapy among RA patients of European ancestry, but not anti-TNF therapies that act through different biological mechanisms or potentially in RA patients of other genetic ancestries.

as a change in DAS28 (ΔDAS) from baseline (so that greater ΔDAS corresponded with better response to therapy; overall mean and standard deviation of 2.1±1.3), adjusted for baseline DAS. A secondary phenotype was used based on European League Against Rheumatism (EULAR) response criteria. EULAR 'good response' was defined as ending DAS<3.2 and ΔDAS>1.2; 'non-response' was defined as ΔDAS <0.6 or ΔDAS≤1.2, and ending DAS >5.1; and 'moderate response' is in between [15]. We limited our secondary analysis to a dichotomous outcome, EULAR good responders (n = 998 for all patients treated with anti-TNF therapy) versus EULAR non-responders (n = 655), excluding the moderate category based on the hypothesis that a more extreme phenotype of response would yield improved discrimination.

Clinical variables were examined for association with phenotype, and therefore possible confounding in genetic association tests. In multivariate models (Table S2), only baseline DAS was strongly associated with the ΔDAS phenotype. As previously shown [11], age and gender showed univariate associations that were attenuated in the multivariate analysis. Accordingly, we used only baseline DAS as a clinical covariate, as this allowed us to maximize sample size given clinical variable missing data in some cohorts.

We performed quality control (QC) filtering and data processing of GWAS data for each of eleven genotyping batches. Genotyping array platforms are described in the Methods. HapMap2 imputation allowed us to test for association at >2 M SNPs with imputation quality scores >0.5. Genotype data were merged across several genotype batches to create four collections for genome-wide association testing. We performed linear regression association tests using baseline DAS and three principal components as covariates, and performed inverse-variance weighted meta-analysis to combine results across the four collections. Quantile-quantile plots with genomic control λ_{GC} values are shown in Figure S1. We found no evidence of systematic inflation of association test results, and no evidence of deflation for imputed versus genotyped SNPs. As a final filter, we excluded SNPs that

Table 1. Samples and clinical data.

Collection (analysis batch):	REF	BRAGGSS	DREAM	ReAct	Total
Sample sizes	959*	595	880*	272	2706
Drug subsets					
etanercept	365	259	109	0	733
infliximab	415	268	211	0	894
adalimumab	174	68	557	272	1071
EULAR Reponse categories					
Good responder	432**	161	313	92	998
Moderate responder	243	258	359	131	991
Non-responder	322	176	208	49	755
Genotype platform	mixed	Affy 500K	Illu550K +650K	Illumina OmniExpress	
Clinical variables					
Age, yr; mean (SD)	53.6 (12.7)	57.4 (10.9)	54.8 (12.9)	53.9 (10.8)	
Disease duration, yr; mean (SD)	6.7 (9.4)	14 (9.8)	9.6 (9.5)	12 (9.1)	
Gender, female %	75.6	77.3	68.3	77.9	
Seropositive, %	87	78	80	70	
MTX co-therapy, %	65.6	85.6	76.0	50.0	
Baseline DAS, mean (SD)	5.5 (1.2)	6.7 (0.9)	5.5 (1.2)	5.9 (1.0)	
ΔDAS, mean (SD)	1.9 (1.6)	2.5 (1.5)	1.9 (1.3)	2.2 (1.3)	
Mean treatment duration	4.6	5.6	3	3	
Study design	All***	Observational	Observational	Observational	

*8 patients had no TNF drug information.
**38 patients had only EULAR response (good, moderate or none) clinical data.
***ABCoN, GENRA are prospective cohorts, BeSt, eRA and TEAR are randomized controlled trial (RCT), and rest of REF group are observational cohorts.
doi:10.1371/journal.pgen.1003394.t001