

16. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol*. 2000;27:261–3.
17. Takeuchi T, Yamanaka H, Inoue E, Nagasawa H, Nawata M, Ikari K, Saito K, Sekiguchi N, Sato E, Kameda H, Iwata S, Mochizuki T, Amano K, Tanaka Y. Retrospective clinical study on the notable efficacy and related factors of infliximab therapy in a rheumatoid arthritis management group in Japan: one-year outcome of joint destruction (RECONFIRM-2 J). *Mod Rheumatol*. 2008;18:447–54.
18. Smolen JS, Aletaha D, Grisar JC, Stamm TA, Sharp JT. Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials. *Ann Rheum Dis*. 2010;69:1058–64.
19. Aletaha D, Funovits J, Breedveld FC, Sharp J, Segurado O, Smolen JS. Rheumatoid arthritis joint progression in sustained remission is determined by disease activity levels preceding the period of radiographic assessment. *Arthritis Rheum*. 2009;60:1242–9.
20. Aletaha D, Funovits J, Smolen JS. Physical disability in rheumatoid arthritis is associated with cartilage damage rather than bone destruction. *Ann Rheum Dis*. 2011;70:733–9.

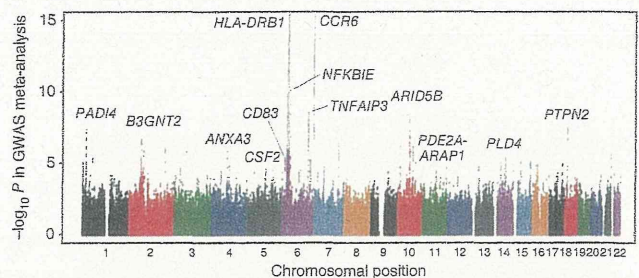
## Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population

Yukinori Okada<sup>1-3,40</sup>, Chikashi Terao<sup>4,5,40</sup>, Katsunori Ikari<sup>6,40</sup>, Yuta Kochi<sup>1,2,40</sup>, Koichiro Ohmura<sup>5</sup>, Akari Suzuki<sup>1</sup>, Takahisa Kawaguchi<sup>4</sup>, Eli A Stahl<sup>7,8</sup>, Fina A S Kurreeman<sup>7-9</sup>, Nao Nishida<sup>10</sup>, Hiroko Ohmiya<sup>3</sup>, Keiko Myouzen<sup>1</sup>, Meiko Takahashi<sup>4</sup>, Tetsuji Sawada<sup>11</sup>, Yuichi Nishioka<sup>12</sup>, Masao Yukioka<sup>13</sup>, Tsukasa Matsubara<sup>14</sup>, Shigeyuki Wakitani<sup>15</sup>, Ryota Teshima<sup>16</sup>, Shigeto Tohma<sup>17</sup>, Kiyoshi Takasugi<sup>18</sup>, Kota Shimada<sup>17</sup>, Aldira Murasawa<sup>19</sup>, Shigeru Honjo<sup>20</sup>, Keitaro Matsuo<sup>21</sup>, Hideo Tanaka<sup>21</sup>, Kazuo Tajima<sup>22</sup>, Taku Suzuki<sup>6,23</sup>, Takuji Iwamoto<sup>6,23</sup>, Yoshiya Kawamura<sup>24</sup>, Hisashi Tani<sup>25</sup>, Yuji Okazaki<sup>26</sup>, Tsukasa Sasaki<sup>27</sup>, Peter K Gregersen<sup>28</sup>, Leonid Padyukov<sup>29</sup>, Jane Worthington<sup>30</sup>, Katherine A Siminovitch<sup>31</sup>, Mark Lathrop<sup>32,33</sup>, Atsuo Taniguchi<sup>6</sup>, Atsushi Takahashi<sup>3</sup>, Katsushi Tokunaga<sup>10</sup>, Michiaki Kubo<sup>34</sup>, Yusuke Nakamura<sup>35</sup>, Naoyuki Kamatani<sup>36</sup>, Tsuneyo Mimori<sup>5</sup>, Robert M Plenge<sup>7,8</sup>, Hisashi Yamanaka<sup>6</sup>, Shigeki Momohara<sup>6,41</sup>, Ryo Yamada<sup>37,41</sup>, Fumihiko Matsuda<sup>4,38,39,41</sup> & Kazuhiko Yamamoto<sup>1,2,41</sup>

**Rheumatoid arthritis is a common autoimmune disease characterized by chronic inflammation. We report a meta-analysis of genome-wide association studies (GWAS) in a Japanese population including 4,074 individuals with rheumatoid arthritis (cases) and 16,891 controls, followed by a replication in 5,277 rheumatoid arthritis cases and 21,684 controls. Our study identified nine loci newly associated with rheumatoid arthritis at a threshold of  $P < 5.0 \times 10^{-8}$ , including *B3GNT2*, *ANXA3*, *CSF2*, *CD83*, *NFKBIE*, *ARID5B*, *PDE2A-ARAP1*, *PLD4* and *PTPN2*. *ANXA3* was also associated with susceptibility to systemic lupus erythematosus ( $P = 0.0040$ ), and *B3GNT2* and *ARID5B* were associated with Graves' disease ( $P = 3.5 \times 10^{-4}$  and  $2.9 \times 10^{-4}$ , respectively). We conducted a multi-ancestry comparative analysis with a previous meta-analysis in individuals of European descent (5,539 rheumatoid arthritis cases and 20,169 controls). This provided evidence of shared genetic risks of rheumatoid arthritis between the populations.**

Rheumatoid arthritis is a complex autoimmune disease characterized by inflammation and the destruction of synovial joints and affects up to 1% of the population worldwide. To date, more than 35 rheumatoid arthritis susceptibility loci, including *HLA-DRB1*, *PTPN22*, *PADI4*, *STAT4*, *TNFAIP3* and *CCR6*, among others, have been identified by GWAS in multiple populations<sup>1-12</sup> and by several meta-analyses of the original GWAS<sup>13-16</sup>. In particular, each meta-analysis of these GWAS uncovered a number of loci that were not identified in the single GWAS, leading to recognition of the enormous power of the meta-analysis approach for detecting causal genes in disease. However, these previous meta-analyses have been performed solely in European populations<sup>13-16</sup> and not in

Asian ones. As multi-ancestry studies on validated rheumatoid arthritis susceptibility loci showed the existence of both population-specific and shared genetic components of rheumatoid arthritis<sup>10,17</sup>, additional studies in Asian populations might provide useful insight into the underlying genetic architecture of rheumatoid arthritis, which would otherwise be difficult to capture using the studies in a single population. Here, we report a meta-analysis of GWAS and a replication study for rheumatoid arthritis in a Japanese population that was conducted by the Genetics and Allied research in Rheumatic diseases NETWORKING (GARNET) consortium<sup>10,12</sup>. We subsequently performed a multi-ancestry comparative analysis that incorporated results from a previously conducted meta-analysis of individuals of European ancestry<sup>15</sup>.



**Figure 1** Manhattan plots of the GWAS meta-analysis for rheumatoid arthritis in the Japanese population. The genetic loci that satisfied the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$  (gray line) in the meta-analysis or in the combined study of the meta-analysis and the replication study are presented. The y axis shows the  $-\log_{10} P$  values of the SNPs in the meta-analysis. The SNPs for which the  $P$  values were smaller than  $1.0 \times 10^{-15}$  are indicated at the upper limit of the plot.

A full list of author affiliations appears at the end of the paper.

Received 24 October 2011; accepted 1 March 2012; published online 25 March 2012; doi:10.1038/ng.2231

**Table 1 Results of the GWAS meta-analysis and the replication studies for rheumatoid arthritis**

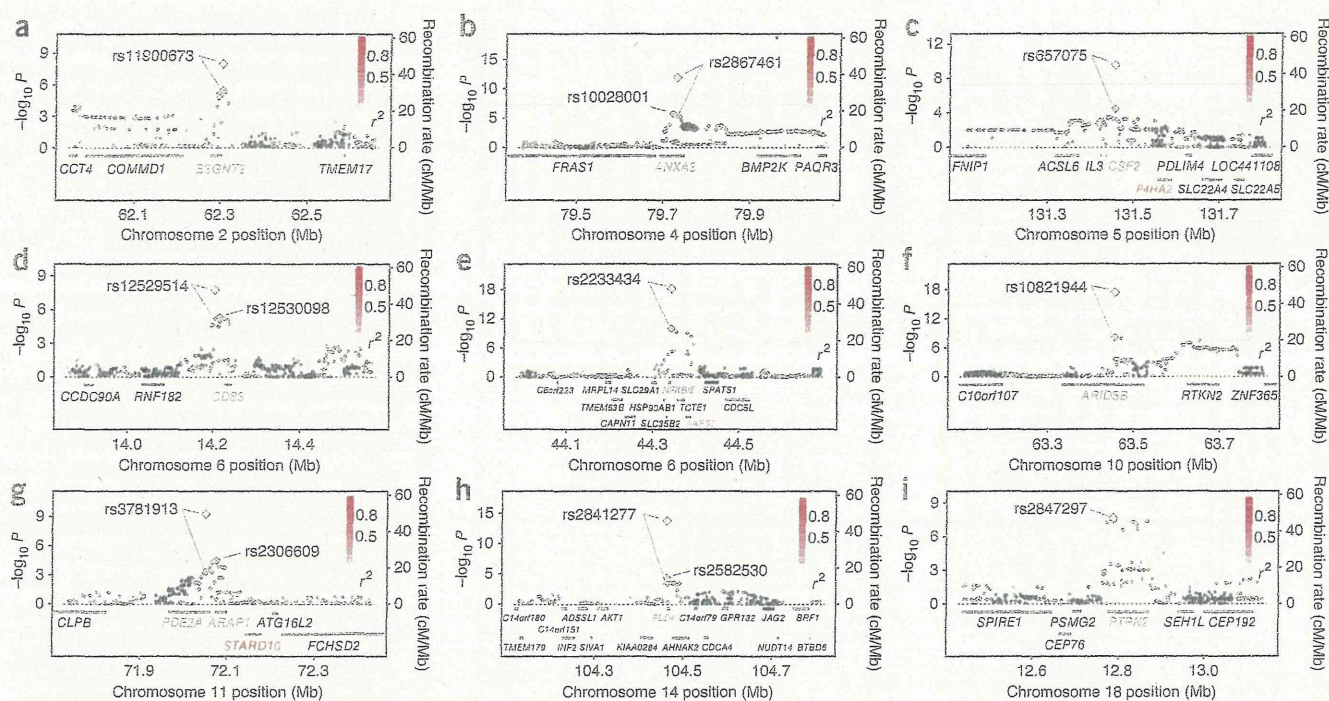
rsID <sup>a</sup>	Chr.	Position (bp)	Cytoband	Gene(s)	Allele	GWAS meta-analysis				Replication study				Associations in Europeans <sup>b</sup>			
						Allele 1 freq.		OR (95% CI) <sup>b</sup>	P	Allele 1 freq.		OR (95% CI) <sup>b</sup>	P	Allele 1 Freq.		OR (95% CI) <sup>b</sup>	P
						1/2	RA			Control	RA			Control	RA		
SNPs with significant associations ( $P < 5.0 \times 10^{-8}$ in the combined study)																	
rs11900673	2	62306165	2p15	B3GNW2	T/C	0.31	0.28	1.15 (1.08–1.21)	$3.5 \times 10^{-6}$	1.09 (1.04–1.14)	$6.0 \times 10^{-4}$	1.11 (1.07–1.15)	$1.1 \times 10^{-8}$	0.13	0.13	1.05 (0.98–1.13)	0.17
rs2867461	4	79732239	4q21	ANXA3	A/G	0.46	0.44	1.13 (1.08–1.19)	$4.7 \times 10^{-6}$	1.12 (1.08–1.17)	$1.2 \times 10^{-7}$	1.13 (1.09–1.17)	$1.2 \times 10^{-12}$	0.37	0.37	0.98 (0.92–1.04)	0.52
rs657075	5	131458017	5q31	CSF2	A/G	0.38	0.36	1.12 (1.06–1.18)	$3.2 \times 10^{-5}$	1.11 (1.06–1.16)	$3.8 \times 10^{-6}$	1.12 (1.08–1.15)	$2.8 \times 10^{-10}$	0.10	0.10	1.04 (0.95–1.13)	0.37
rs12529514	6	14204637	6p23	CD83	C/T	0.16	0.14	1.19 (1.10–1.27)	$6.8 \times 10^{-6}$	1.11 (1.05–1.18)	$6.0 \times 10^{-4}$	1.14 (1.09–1.19)	$2.0 \times 10^{-8}$	0.055	0.053	1.11 (0.99–1.24)	0.074
rs2233434	6	44340898	6p21.1	NFKBIE	G/A	0.24	0.21	1.23 (1.16–1.31)	$9.2 \times 10^{-11}$	1.17 (1.11–1.23)	$2.2 \times 10^{-9}$	1.19 (1.15–1.24)	$5.8 \times 10^{-19}$	0.059	0.040	1.57 (1.11–2.21)	0.0099
rs10821944	10	63455095	10q21	ARID5B	G/T	0.39	0.36	1.17 (1.11–1.23)	$1.0 \times 10^{-8}$	1.15 (1.10–1.20)	$3.0 \times 10^{-10}$	1.16 (1.12–1.20)	$5.5 \times 10^{-18}$	0.29	0.26	1.11 (1.05–1.17)	$1.9 \times 10^{-4}$
rs3781913	11	72051144	11q13	PDE2A-ARAF1	T/G	0.71	0.69	1.11 (1.05–1.17)	$3.2 \times 10^{-4}$	1.13 (1.08–1.18)	$6.7 \times 10^{-7}$	1.12 (1.08–1.16)	$5.8 \times 10^{-10}$	0.45	0.43	1.04 (0.99–1.09)	0.13
rs2841277	14	104462050	14q32	PLD4	T/C	0.72	0.69	1.11 (1.05–1.18)	$2.8 \times 10^{-4}$	1.18 (1.13–1.24)	$7.0 \times 10^{-12}$	1.15 (1.11–1.19)	$1.9 \times 10^{-14}$	0.47	0.46	1.02 (0.96–1.09)	0.54
rs2847297	18	12787694	18p11	PTPN2	G/A	0.37	0.33	1.16 (1.11–1.23)	$3.5 \times 10^{-8}$	1.06 (1.01–1.11)	0.013	1.10 (1.07–1.14)	$2.2 \times 10^{-8}$	0.36	0.34	1.10 (1.05–1.15)	$9.2 \times 10^{-5}$
SNPs with suggestive associations ( $5.0 \times 10^{-8} \leq P < 5.0 \times 10^{-6}$ in the combined study)																	
rs4937362	11	127997949	11q24	ETS1-FLJ1	T/C	0.71	0.68	1.13 (1.07–1.19)	$2.0 \times 10^{-5}$	1.07 (1.02–1.12)	0.0061	1.09 (1.06–1.13)	$7.5 \times 10^{-7}$	0.46	0.44	1.06 (1.01–1.11)	0.015
rs3783637	14	54417868	14q22	GCH1	C/T	0.76	0.74	1.13 (1.07–1.20)	$6.5 \times 10^{-5}$	1.07 (1.02–1.13)	0.0062	1.10 (1.06–1.14)	$2.0 \times 10^{-6}$	0.88	0.88	0.99 (0.88–1.11)	0.87
rs1957895	14	60978085	14q23	PRKCH	G/T	0.40	0.39	1.12 (1.06–1.18)	$4.1 \times 10^{-5}$	1.07 (1.02–1.12)	0.0020	1.09 (1.05–1.13)	$3.6 \times 10^{-7}$	0.093	0.089	1.01 (0.95–1.07)	0.73
rs6496667	15	88694672	15q26	ZNF774	A/C	0.38	0.35	1.13 (1.07–1.19)	$4.7 \times 10^{-5}$	1.07 (1.02–1.11)	0.0050	1.09 (1.05–1.13)	$1.4 \times 10^{-6}$	0.21	0.20	1.07 (1.01–1.13)	0.031
rs7404928	16	23796341	16p12	PRKCB1	T/C	0.65	0.62	1.13 (1.07–1.19)	$1.5 \times 10^{-5}$	1.05 (1.01–1.10)	0.026	1.08 (1.05–1.12)	$4.0 \times 10^{-6}$	0.75	0.75	1.01 (0.94–1.09)	0.79
rs2280381	16	84576134	16q24	IRF8	T/C	0.86	0.84	1.16 (1.08–1.25)	$1.0 \times 10^{-4}$	1.09 (1.03–1.15)	0.0049	1.12 (1.07–1.17)	$2.4 \times 10^{-6}$	0.62	0.60	1.05 (0.99–1.11)	0.081
SNPs in previously reported rheumatoid arthritis susceptibility loci ( $P < 5.0 \times 10^{-8}$ in the GWAS)																	
rs766449	1	17547439	1p36	PADI4	T/C	0.44	0.40	1.17 (1.11–1.24)	$4.6 \times 10^{-8}$	—	—	—	—	0.38	0.37	1.09 (1.03–1.05)	0.0022
rs2157337	6	32609122	6p21.3	HLA-DRB1	C/T	0.59	0.44	1.99 (1.88–2.11)	$2.6 \times 10^{-118}$	—	—	—	—	0.69	0.46	2.50 (2.39–2.62)	$< 1.0 \times 10^{-300}$
rs6932056	6	138284130	6q23	TNFAIP3	C/T	0.092	0.073	1.35 (1.23–1.49)	$3.2 \times 10^{-9}$	—	—	—	—	0.044	0.034	1.41 (1.24–1.60)	$1.3 \times 10^{-7}$
rs1571878	6	167460832	6q27	CCR6	C/T	0.54	0.48	1.31 (1.24–1.39)	$3.2 \times 10^{-19}$	—	—	—	—	0.47	0.43	1.13 (1.08–1.19)	$5.9 \times 10^{-7}$

Chr., chromosome; Freq., frequency; RA, rheumatoid arthritis; OR, odds ratio; CI, confidence interval. <sup>a</sup>SNPs with  $P < 5.0 \times 10^{-6}$  in the combined study of the GWAS meta-analysis and the replication study or SNPs with  $P < 5.0 \times 10^{-8}$  in the GWAS meta-analysis are annotated according to forward strand and NCBI Build 36.3. Full results of the replication study are provided in Supplementary Table 3. <sup>b</sup>Odds ratio of allele 1. <sup>c</sup>Associations in the previous meta-analysis in European populations<sup>15</sup>.

The meta-analysis included 4,074 rheumatoid arthritis cases (with 81.4% and 80.4% of the subjects being positive for antibody to cyclic citrullinated peptide (anti-CCP) and rheumatoid factor, respectively) and 16,891 controls from three GWAS of Japanese subjects (from the BioBank Japan Project<sup>10,18</sup>, Kyoto University<sup>12</sup> and the Institute of Rheumatology Rheumatoid Arthritis (IORRA)<sup>19</sup>; Supplementary Table 1). After the application of stringent quality control criteria, including principal-component analysis (PCA; Supplementary Fig. 1) for each GWAS, the meta-analysis was conducted by evaluating ~2.0 million autosomal SNPs with minor allele frequencies (MAFs)  $\geq 0.01$ , which were obtained through whole-genome imputation of genotypes on the basis of the HapMap Phase 2 East Asian panels (Japanese in Tokyo (JPT) and Han Chinese in Beijing (CHB)). The inflation factor of the test statistics in the meta-analysis  $\lambda_{GC}$  was as low as 1.036, suggesting no substantial effects of population structure (Supplementary Table 2). The quantile-quantile plot of  $P$  values showed a marked discrepancy in the values in its tail from those anticipated under the null hypothesis that there is no association—even after removal of the SNPs located in the human leukocyte antigen (HLA) region, the major rheumatoid arthritis susceptibility locus—thereby showing the presence of significant associations in the meta-analysis (Supplementary Fig. 2).

We identified seven loci in the current meta-analysis that satisfied the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$ . These included previously known rheumatoid arthritis susceptibility loci, such as *PADI4* at 1p36, *HLA-DRB1* at 6p21.3, *TNFAIP3* at 6q23 and *CCR6* at 6q27 (refs. 1,3,6,10,15) (the smallest  $P = 2.6 \times 10^{-118}$  was found at the *HLA-DRB1* locus; Fig. 1 and Table 1). To our knowledge, the other three loci identified, *NFKBIE* at 6p21.1, *ARID5B* at 10q21 and *PTPN2* at 18p11, are newly associated ( $P = 9.2 \times 10^{-11}$ ,  $1.0 \times 10^{-8}$  and  $3.5 \times 10^{-8}$ , respectively).

To validate the associations identified in the meta-analysis, we conducted a replication study of two independent Japanese rheumatoid arthritis case-control cohorts (cohort 1: 3,830 rheumatoid arthritis cases and 17,920 controls, cohort 2: 1,447 rheumatoid arthritis cases and 3,764 controls; Supplementary Table 1). To increase the number of subjects and enhance statistical power, genotype data obtained from other GWAS projects conducted for non-autoimmune diseases in Japanese using Illumina platforms were used for the replication control panels. For each of the 46 loci that exhibited  $P < 5.0 \times 10^{-4}$  in



**Figure 2** Regional plots of the loci newly associated with rheumatoid arthritis at the genome-wide significance threshold of  $P < 5.0 \times 10^{-8}$  in the combined study of the meta-analysis and the replication study. (a–i) Regional plots are shown at *B3GNT2* (a), *ANXA3* (b), *CSF2* (c), *CD83* (d), *NFKBIE* (e), *ARID5B* (f), *PDE2A-ARAP1* (g), *PLD4* (h) and *PTPN2* (i). Diamonds represent the  $-\log_{10} P$  values of the SNPs, and the red diamonds represent the  $-\log_{10} P$  values of the SNPs in the meta-analysis. Red color for the smaller circles represents the  $r^2$  value with the most significantly associated SNP (larger red circle). The purple circle represents the  $P$  value in the combined study. The blue line shows the recombination rates given by the HapMap Phase 2 east Asian populations (release 22). RefSeq genes at the loci are indicated below. Genes nearest to the marker SNPs at the loci are colored blue (Supplementary Note), and genes implicated in eQTL analysis are colored red (Supplementary Table 4). At 11q13, two genes (*PDE2A* and *ARAP1*) that are nearest to the SNP selected for the replication study and the most significant SNP in the meta-analysis are highlighted. The plots were drawn using SNP Annotation and Proxy Search (SNAP) version 2.2.

the meta-analysis and had not been reported as rheumatoid arthritis susceptibility loci<sup>1–16</sup>, we selected a marker SNP for the replication study (Online Methods and Supplementary Table 3).

In the combined analyses of the meta-analysis and the replication study, including a total of 9,351 rheumatoid arthritis cases and 38,575 controls, we identified six newly associated loci, in addition to the *NFKBIE*, *ARID5B* and *PTPN2* loci, that satisfied the significance threshold of  $P < 5.0 \times 10^{-8}$ , including *B3GNT2* at 2p15, *ANXA3* at 4q21, *CSF2* at 5q31, *CD83* at 6p23, *PDE2A-ARAP1* at 11q13 and *PLD4* at 14q32 (Figs. 1 and 2 and Table 1). Of these loci, *NFKBIE* had the smallest  $P$  value ( $5.8 \times 10^{-19}$ ). Although association with rheumatoid arthritis has been described for the *CSF2* and *PTPN2* loci<sup>11,15,16,20,21</sup>, ours is the first report to our knowledge validating these associations with a threshold of  $P < 5.0 \times 10^{-8}$ . Suggestive associations were also observed in *ETS1-FLI1* at 11q24, *GCH1* at 14q22, *PRKCH* at 14q23, *ZNF774* at 15q26, *PRKCB1* at 16p12 and *IRF8* at 16q24 ( $5.0 \times 10^{-8} \leq P < 5.0 \times 10^{-6}$ ). A summary of the genes in the newly associated loci and the results of *cis* expression quantitative trait locus (*cis* eQTL) analysis of the marker SNPs are provided (Supplementary Table 4 and Supplementary Note).

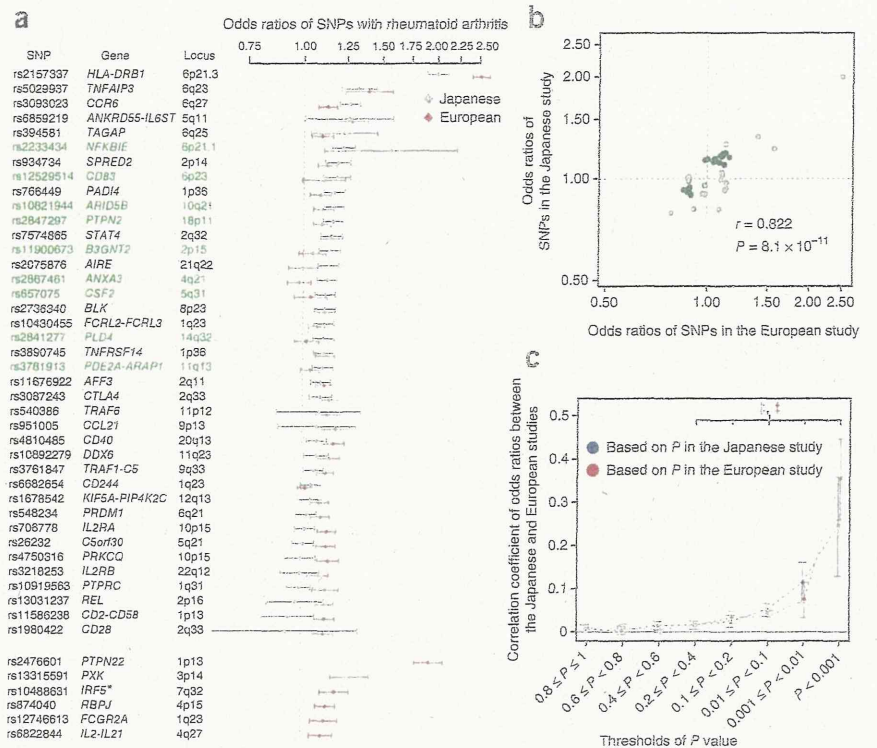
Previous studies have reported associations of rheumatoid arthritis susceptibility loci with other autoimmune diseases<sup>4,10,15,16</sup>. Therefore, we assessed the association of these newly identified susceptibility loci with systemic lupus erythematosus (SLE) by examining the results of an SLE GWAS in the Japanese population (891 cases and 3,384 controls)<sup>22</sup> and in Graves' disease by genotyping 1,783 cases<sup>10</sup> (the controls from the SLE analysis were used for testing for Graves'

disease). We observed significant associations of the *ANXA3* locus with SLE and of the *B3GNT2* and *ARID5B* loci with Graves' disease, which showed the same directional effects of the alleles as in rheumatoid arthritis ( $P < 0.05/9 = 0.0056$ , Bonferroni correction of the number of loci; Supplementary Table 5). It should be noted that relatively small sample sizes in the SLE and Graves' disease cohorts might yield limited statistical power, and further evaluations enrolling larger numbers of subjects would be desirable.

To highlight genetic backgrounds of rheumatoid arthritis that are common and divergent in different ancestry groups, we conducted a multi-ancestry comparative analysis of the present study in Japanese and a previous GWAS meta-analysis in Europeans that included 5,539 rheumatoid arthritis cases and 20,169 controls<sup>15</sup> (Fig. 3a–c). First, we compared associations in the reported<sup>1–16</sup> or newly identified rheumatoid arthritis susceptibility loci (Fig. 3a and Supplementary Table 6). Of the 46 rheumatoid arthritis risk variants evaluated, 6 were monomorphic in Japanese, and all were polymorphic in Europeans. We observed significant associations at 22 loci in Japanese and at 36 loci in Europeans (false discovery rate (FDR)  $< 0.05$ ,  $P < 0.0030$ ), with 14 loci being shared between the populations. Of the newly associated rheumatoid arthritis susceptibility loci identified in our Japanese meta-analysis, significant associations were also observed in the European meta-analysis at the *ARID5B* and *PTPN2* loci ( $P = 1.9 \times 10^{-4}$  and  $9.2 \times 10^{-5}$ , respectively; Table 1). Significant positive correlation of odds ratios was observed between the studies ( $r = 0.822$ ,  $P = 8.1 \times 10^{-11}$ ; Fig. 3b), suggesting that a substantial proportion of genetic factors are shared between

# LETTERS

**Figure 3** Overlap of the associations with rheumatoid arthritis between Japanese and European populations. (a) Forest plots of SNPs in the rheumatoid arthritis susceptibility loci (Supplementary Table 6). We selected the genetic loci that have been validated to be associated with rheumatoid arthritis susceptibility by showing associations in the reports of multiple cohorts or satisfying the genome-wide significant threshold ( $P < 5.0 \times 10^{-8}$ ) in previous studies, including in the meta-analysis and replication phases<sup>1-16</sup>. For each of the loci, the most significant SNP among those reported in the previous or present study were selected<sup>1-16</sup>. SNPs in the newly identified rheumatoid arthritis susceptibility loci are colored green. Odds ratios and 95% confidence interval (CI) values are based on rheumatoid arthritis risk alleles, and the SNPs are ordered according to the odds ratios in the Japanese study. Several SNPs were monomorphic in the Japanese population. The odds ratios of these SNPs in the European study are presented below. The asterisk indicates that an association of another variant at the *IRF5* locus was reported in the Japanese population<sup>24</sup>. (b) Correlation of the odds ratios of the SNPs in the validated rheumatoid arthritis susceptibility loci between the two populations. SNPs that were polymorphic in both populations were used; odds ratios were based on the minor allele in the Japanese population. (c) Correlation of the odds ratios of the genome-wide SNPs, excluding the rheumatoid arthritis susceptibility loci. Correlations were evaluated for sets of SNPs stratified by the thresholds based on the meta-analysis *P* values in each population after pruning of the SNPs by LD ( $r^2 < 0.3$ ). Correlation coefficient and 95% CI are indicated on the y axis. Significant correlation of the odds ratios was observed ( $\ddagger$ ,  $P < 0.005$ ), even for the SNPs that showed moderate associations with rheumatoid arthritis (meta-analysis  $P < 0.4$  in each population).



the two ancestry groups<sup>17</sup>. When the rheumatoid arthritis cases of the Japanese GWAS meta-analysis were stratified into anti-CCP-positive or rheumatoid factor-positive cases ( $n = 3,209$ ) and controls ( $n = 16,891$ ), similar results were observed (data not shown). Nevertheless, most of the SNPs assessed here are not necessarily causal variants, and further fine mapping of the loci is warranted to precisely evaluate the shared genetic predisposition between the populations.

Next, we compared regional associations within each of the loci and identified unique patterns in the *ARID5B* locus at 10q21 (Supplementary Fig. 3). In Japanese, three peaks of association were observed ( $P = 1.0 \times 10^{-8}$  at rs10821944,  $P = 5.7 \times 10^{-8}$  at rs10740069 and  $P = 8.5 \times 10^{-6}$  at rs224311). These three variants were in weak linkage disequilibrium (LD) in Japanese ( $r^2 < 0.10$ ), indicating independent associations with each of the other SNPs that satisfied a region-wide significance threshold of  $P < 3.5 \times 10^{-5}$  (conditional  $P = 4.3 \times 10^{-6}$ ,  $1.7 \times 10^{-5}$  and  $1.8 \times 10^{-5}$ , respectively) (Supplementary Fig. 3). In contrast, there was only one peak of association in Europeans ( $P = 1.2 \times 10^{-6}$  at rs12764378;  $r^2 = 0.59$  with rs10821944 in Europeans), and no additional association was observed in conditional analysis with rs12764378 (the smallest conditional  $P = 2.2 \times 10^{-4}$ ), suggesting that the number of independent associations may be different at this locus in the two populations.

Finally, we conducted polygenic assessment for common variants showing modest associations to rheumatoid arthritis (those not meeting the genome-wide association threshold). This approach has been recognized to be a means to explain a substantial proportion of genetic risk<sup>23</sup>. For the SNPs that were shared between the two meta-analyses but not included in the validated rheumatoid arthritis

susceptibility loci, we adopted LD pruning of the SNPs ( $r^2 < 0.3$ ). We then evaluated the correlation of odds ratios of the SNPs between the two meta-analyses and observed a significant positive correlation ( $r = 0.023$ ,  $P < 1.0 \times 10^{-300}$ ). When the SNPs were stratified according to the *P* values in each meta-analysis, significant positive correlations of odds ratios were observed for the SNPs, even for those showing modest association ( $P < 0.4$  in the meta-analysis of Japanese or Europeans;  $r = 0.014-0.36$  for each *P* value range,  $P < 0.005$  for each correlation test) (Fig. 3c). Correlations ( $r$ ) of odds ratios observed herein suggest substantial overlap of the genetic risk of rheumatoid arthritis between the two populations, not only in the validated rheumatoid arthritis susceptibility loci but also at the loci showing nonsignificant associations. This suggests the usefulness of a meta-analysis approach involving multiple ancestry groups in identifying additional susceptibility loci.

In summary, we identified multiple new loci associated with rheumatoid arthritis through a large-scale meta-analysis of GWAS in Japanese. Multi-ancestry comparative analysis provided evidence of significant overlap in the genetic risks of rheumatoid arthritis between Japanese and Europeans. Thus, findings from the present study should contribute to the further understanding of the etiology of rheumatoid arthritis.

URLs. GARNET consortium, <http://www.twmu.ac.jp/IOR/garnet/home.html>; The BioBank Japan Project (in Japanese), <http://biobankjp.org/>; International HapMap Project, <http://www.hapmap.org/>; PLINK, <http://pngu.mgh.harvard.edu/~purcell/plink/>; EIGENSTRAT, <http://genepath.med.harvard.edu/~reich/Software.htm>; MACH and mach2dat, <http://www.sph.umich.edu/csg/abecasis/MACH/index>.

html; R statistical software, <http://cran.r-project.org/>; SNAP, <http://www.broadinstitute.org/mpg/snap/index.php>; NCBI GEO database, <http://www.ncbi.nlm.nih.gov/geo/>.

## METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

*Note: Supplementary information is available on the Nature Genetics website.*

## ACKNOWLEDGMENTS

The authors acknowledge the essential role of the GARNET consortium in developing the study. In this study, the following GARNET members are included: CGM of RIKEN, University of Tokyo, the BioBank Japan Project, Kyoto University and IORRA. We would like to thank all the doctors and staff who participated in sample collection for the RIKEN cohort and the BioBank Japan Project. We thank K. Kobayashi and M. Kitazato for their technical assistance. We thank T. Raj for calculation of composite of multiple signals (CMS). We thank M. Kokubo for DNA extraction, GWAS genotyping and secretarial assistance. We would also like to thank H. Yoshifuji, N. Yukawa, D. Kawabata, T. Nojima, T. Usui and T. Fujii for collecting DNA samples. We thank Y. Katagiri for her technical efforts. We also appreciate the contribution of E. Inoue and other members of the Institute of Rheumatology, Tokyo Women's Medical University, for their efforts on the IORRA cohort. This study was supported in part by grants-in-aid from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan, the Ministry of Health, Labour and Welfare (MHLW) in Japan, the Japan Society for the Promotion of Science (JSPS), Core Research for Evolutional Science and Technology (CREST), Solution-Oriented Research for Science and Technology (SORST), INSERM and the Okawa Foundation for Information and Telecommunications.

## AUTHOR CONTRIBUTIONS

Y. Okada, C.T., K.I., Y. Kochi and K.O. designed the study and drafted the manuscript. Y. Okada, C.T., K.I., T.K., H.O., N.N., M.T., M.L., K. Tokunaga and M.K. managed genotyping and manipulation of GWAS data. Y. Okada, Y. Kochi, C.T. and K.I. managed genotyping of replication cohorts. Y. Okada, T.K., H.O., E.A.S., A. Takahashi and R.Y. performed statistical analysis. Y. Kochi, A.S., K. Myouzen, T. Sawada, Y. Nishoka, M.Y., T. Matsubara, S.W., R.T. and S.T. collected samples and managed phenotype data for the rheumatoid arthritis cohorts from the BioBank Japan Project and CGM, RIKEN. C.T., K.O., T.K., M.T., K. Takasugi, K.S., A.M., S.H., K. Matsuo, H. Tanaka, K. Tajima and M.L. collected samples and managed phenotype data for the rheumatoid arthritis cohorts from Kyoto University. K.I., T. Suzuki, T.I., Y. Kawamura, H. Tani, Y. Okazaki and T. Sakaki collected samples and managed phenotype data for the rheumatoid arthritis cohorts from IORRA. Y. Kochi managed the data for the SLE and Graves' disease cohorts. A.S., C.T. and K.I. analyzed the sera of subjects with rheumatoid arthritis. E.A.S., F.A.S.K., P.K.G., J.W., K.A.S., L.P. and R.M.P. managed the data for the rheumatoid arthritis cohorts in European populations. A. Taniguchi, A. Takahashi, K. Tokunaga, M.K., Y. Nakamura, N.K., T. Minori, R.M.P., H.Y., S.M., R.Y., F.M. and K.Y. supervised the overall study.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Suzuki, A. *et al.* Functional haplotypes of *PADI4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat. Genet.* **34**, 395–402 (2003).
- Kochi, Y. *et al.* A functional variant in *FCRL3*, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nat. Genet.* **37**, 478–485 (2005).
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
- Remmers, E.F. *et al.* *STAT4* and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N. Engl. J. Med.* **357**, 977–986 (2007).
- Plenge, R.M. *et al.* *TRAF1-C5* as a risk locus for rheumatoid arthritis—a genome-wide study. *N. Engl. J. Med.* **357**, 1199–1209 (2007).
- Plenge, R.M. *et al.* Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat. Genet.* **39**, 1477–1482 (2007).
- Barton, A. *et al.* Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. *Nat. Genet.* **40**, 1156–1159 (2008).
- Suzuki, A. *et al.* Functional SNPs in *CD244* increase the risk of rheumatoid arthritis in a Japanese population. *Nat. Genet.* **40**, 1224–1229 (2008).
- Gregersen, P.K. *et al.* *REL*, encoding a member of the NF- $\kappa$ B family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat. Genet.* **41**, 820–823 (2009).
- Kochi, Y. *et al.* A regulatory variant in *CCR6* is associated with rheumatoid arthritis susceptibility. *Nat. Genet.* **42**, 515–519 (2010).
- Freudenberg, J. *et al.* Genome-wide association study of rheumatoid arthritis in Koreans: population-specific loci as well as overlap with European susceptibility loci. *Arthritis Rheum.* **63**, 884–893 (2011).
- Terao, C. *et al.* The human *AIRE* gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. *Hum. Mol. Genet.* **20**, 2680–2685 (2011).
- Raychaudhuri, S. *et al.* Common variants at *CD40* and other loci confer risk of rheumatoid arthritis. *Nat. Genet.* **40**, 1216–1223 (2008).
- Raychaudhuri, S. *et al.* Genetic variants at *CD28*, *PRDM1* and *CD2/CD58* are associated with rheumatoid arthritis risk. *Nat. Genet.* **41**, 1313–1318 (2009).
- Stahl, E.A. *et al.* Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.* **42**, 508–514 (2010).
- Zhernakova, A. *et al.* Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet.* **7**, e1002004 (2011).
- Kurzeeman, F. *et al.* Genetic basis of autoantibody positive and negative rheumatoid arthritis risk in a multi-ethnic cohort derived from electronic health records. *Am. J. Hum. Genet.* **88**, 57–69 (2011).
- Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).
- Yamanaka, H. *et al.* Influence of methotrexate dose on its efficacy and safety in rheumatoid arthritis patients: evidence based on the variety of prescribing approaches among practicing Japanese rheumatologists in a single institute-based large observational cohort (IORRA). *Mod. Rheumatol.* **17**, 98–105 (2007).
- Yamada, R. *et al.* Association between a single-nucleotide polymorphism in the promoter of the human interleukin-3 gene and rheumatoid arthritis in Japanese patients, and maximum-likelihood estimation of combinatorial effect that two genetic loci have on susceptibility to the disease. *Am. J. Hum. Genet.* **68**, 674–685 (2001).
- Tokuhiro, S. *et al.* An intronic SNP in a *RUNX1* binding site of *SLC22A4*, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat. Genet.* **35**, 341–348 (2003).
- Okada, Y. *et al.* A genome-wide association study identified *AFF1* as a susceptibility locus for systemic lupus erythematosus in Japanese. *PLoS Genet.* **8**, e1002455 (2012).
- Stranger, B.E., Stahl, E.A. & Raj, T. Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* **187**, 367–383 (2011).
- Shimane, K. *et al.* A single nucleotide polymorphism in the *IRF5* promoter region is associated with susceptibility to rheumatoid arthritis in the Japanese patients. *Ann. Rheum. Dis.* **68**, 377–383 (2009).

<sup>1</sup>Laboratory for Autoimmune Diseases, Center for Genomic Medicine (CGM), RIKEN, Yokohama, Japan. <sup>2</sup>Department of Allergy and Rheumatology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. <sup>3</sup>Laboratory for Statistical Analysis, CGM, RIKEN, Yokohama, Japan. <sup>4</sup>Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan. <sup>5</sup>Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. <sup>6</sup>Institute of Rheumatology, Tokyo Women's Medical University, Tokyo, Japan. <sup>7</sup>Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. <sup>8</sup>Broad Institute, Cambridge, Massachusetts, USA. <sup>9</sup>Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands. <sup>10</sup>Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. <sup>11</sup>Department of Rheumatology, Tokyo Medical University Hospital, Tokyo, Japan. <sup>12</sup>Yamanashi Prefectural Central Hospital, Yamanashi, Japan. <sup>13</sup>Department of Orthopaedic Surgery, Yukioka Hospital, Osaka, Japan. <sup>14</sup>Matsubara Mayflower Hospital, Hyogo, Japan. <sup>15</sup>Osaka Minami National Hospital, Osaka, Japan. <sup>16</sup>Department of Orthopedic Surgery, Tottori University, Tottori, Japan. <sup>17</sup>Department of Rheumatology, National Hospital Organization, Sagami Hospital, Kanagawa, Japan. <sup>18</sup>Center for Rheumatic Diseases, Dohgo Spa Hospital, Ehime, Japan. <sup>19</sup>Department of Rheumatology, Niigata Rheumatic Center, Niigata, Japan. <sup>20</sup>Saiseikai Takaoka Hospital, Toyama, Japan. <sup>21</sup>Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Aichi, Japan. <sup>22</sup>Aichi Cancer Center Hospital and Research Institute, Aichi, Japan. <sup>23</sup>Department of Orthopaedic Surgery, Keio University, Tokyo, Japan. <sup>24</sup>Yokohama Clinic, Warakukai Medical Corporation, Yokohama, Japan. <sup>25</sup>Department of Psychiatry, Mie University School of Medicine, Mie, Japan. <sup>26</sup>Metropolitan Matsuzawa Hospital, Tokyo, Japan. <sup>27</sup>Graduate School of Education, University of Tokyo, Tokyo, Japan. <sup>28</sup>The Feinstein Institute for Medical Research, North Shore–Long Island Jewish Health System, Manhasset, New York, USA. <sup>29</sup>Rheumatology Unit,

## LETTERS

Department of Medicine in Solna, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden. <sup>30</sup>Arthritis Research Campaign–Epidemiology Unit, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. <sup>31</sup>Division of Medicine, University of Toronto, Mount Sinai Hospital and University Health Network, Toronto, Ontario, Canada. <sup>32</sup>Commissariat à l’Energie Atomique (CEA), Institut Genomique, Centre National de Genotypage, Evry, France. <sup>33</sup>Fondation Jean Dausset, Centre d’Etude du Polymorphisme Humain, Paris, France. <sup>34</sup>Laboratory for Genotyping Development, CGM, RIKEN, Yokohama, Japan. <sup>35</sup>Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. <sup>36</sup>Laboratory for International Alliance, CGM, RIKEN, Yokohama, Japan. <sup>37</sup>Unit of Statistical Genetics, Center for Genomic Medicine Graduate School of Medicine Kyoto University, Kyoto, Japan. <sup>38</sup>Core Research for Evolutional Science and Technology (CREST) Program, Japan Science and Technology Agency, Kawaguchi, Japan. <sup>39</sup>Institut National de la Santé et de la Recherche Médicale (INSERM), Unité U852, Kyoto University Graduate School of Medicine, Kyoto, Japan. <sup>40</sup>These authors contributed equally to this work. <sup>41</sup>These authors jointly directed this work. Correspondence should be addressed to Y.K. (ykochi@src.riken.jp) or K.O. (ohmurako@kuhp.kyoto-u.ac.jp).



## ONLINE METHODS

**Subjects.** The Japanese participants in the meta-analysis (4,074 rheumatoid arthritis cases and 16,891 controls) and the replication study (5,277 rheumatoid arthritis cases and 21,684 controls) were obtained through the collaborations of the GARNET consortium (Supplementary Table 1)<sup>10,12</sup>. The meta-analysis was conducted on three independent GWAS (from the BioBank Japan Project<sup>18</sup> with 2,414 rheumatoid arthritis cases and 14,245 controls<sup>10</sup>, Kyoto University with 1,237 rheumatoid arthritis cases and 2,087 controls<sup>12</sup> and IORRA<sup>19</sup> with 423 rheumatoid arthritis cases and 559 controls). The replication study consisted of two independent cohorts (cohort 1 included 3,830 rheumatoid arthritis cases and 17,920 controls, and cohort 2 included 1,447 rheumatoid arthritis cases and 3,764 controls). We employed a case-control cohort of SLE (891 cases and 3,384 controls)<sup>22</sup> and 1,783 cases with Graves' disease<sup>10</sup>. Details of 5,539 rheumatoid arthritis cases and 20,169 controls included in the meta-analysis in European populations were described elsewhere<sup>15</sup>. All participants provided written informed consent for participation in the study, as approved by the ethical committees of the institutional review boards. Detailed descriptions of the participating subjects are provided (Supplementary Note).

**Genotyping and quality control in the GWAS.** Genotyping platforms and quality control criteria for the GWAS, including cutoff values for sample call rates, SNP call rates, MAF and Hardy-Weinberg *P* values, are given (Supplementary Table 2). For the subjects enrolled in each of three GWAS, we excluded closely related subjects with first- or second-degree kinship, which was estimated using PLINK version 1.06 (see URLs). We also excluded the subjects determined to be ancestry outliers from East Asian populations using PCA performed by EIGENSTRAT version 2.0 (see URLs) along with HapMap Phase 2 panels (release 24; Supplementary Fig. 1). Genotype imputation was performed on the basis of the HapMap Phase 2 East Asian populations, using MACH version 1.0.16 (see URLs) in a two-step procedure as described elsewhere<sup>25</sup>. We excluded imputed SNPs with MAF < 0.01 or *Rsq* < 0.5 from each of the GWAS. Associations of the SNPs with rheumatoid arthritis were assessed by logistic regression models assuming additive effects of the allele dosages of the SNPs using mach2dat software (see URLs).

**Meta-analysis.** We included 1,948,139 autosomal SNPs that satisfied quality control criteria in all three GWAS (Supplementary Table 2). SNP information was based on a forward strand of the NCBI build 36.3 reference sequence. The meta-analysis was performed using an inverse variance method assuming a fixed-effects model from the study-specific effect sizes (logarithm of odds ratio) and the standard errors of the coded alleles of the SNPs determined with the Java source code implemented by the authors<sup>25</sup>. Genomic control corrections<sup>26</sup> were carried out on test statistics of the GWAS using the study-specific inflation factor ( $\lambda_{GC}$ ) and was applied or reapplied to the results of our current meta-analysis (Supplementary Fig. 2).

**Replication study.** We selected a SNP for the replication study from each of the loci that exhibited  $P < 5.0 \times 10^{-4}$  in the meta-analysis that had not previously been reported as rheumatoid arthritis susceptibility loci<sup>1-16</sup> (Supplementary Table 3). For control subjects, we used genotype data obtained from additional GWAS for non-autoimmune diseases or healthy controls, genotyped using Illumina HumanHap550 BeadChips or HumanHap610-Quad BeadChips, and

the cases for rheumatoid arthritis and Graves' disease were genotyped with the TaqMan genotyping system (Applied Biosystems; Supplementary Table 1). Selection of the SNP was conducted according to the following criteria: if the SNP with the most significant association in the locus was genotyped in the replication control panel, then that SNP was selected; otherwise, a tag SNP in the replication control panel with the strongest LD was selected (mean  $r^2 = 0.89$ ). For the three SNPs that yielded low call rates (<90%), we alternatively selected proxy SNPs with the second strongest LD. As a result, average genotyping call rates of the SNPs were 99.9% and 99.0% for the controls and cases, respectively. We then evaluated concordance rates between the assayed genotypes by applying these two different methods to samples from 376 subjects who were randomly selected. This procedure yielded high concordance rates of  $\geq 99.9\%$ . Associations of the SNPs were evaluated using logistic regression assuming an additive-effects model of genotypes in R statistical software version 2.11.0 (see URLs). The combined study of the meta-analysis and replication study was performed using an inverse variance method assuming a fixed-effects model<sup>25</sup>.

**Cis eQTL analysis.** For each marker SNP of the newly identified rheumatoid arthritis susceptibility locus, correlations between SNP genotypes and expression levels of genes located 300 kb upstream or downstream of the SNP measured in B-lymphoblastoid cell lines (GSE6536) were evaluated using data from the HapMap Phase 2 east Asian populations<sup>27</sup>.

**Multi-ancestry analysis of the meta-analyses in Japanese and Europeans.** We evaluated the associations of the variants in the validated rheumatoid arthritis susceptibility loci by comparing the results from the current meta-analysis in Japanese with those from a previous meta-analysis in Europeans<sup>15</sup>. We assessed two variants in the *IRF5* locus, where different causal variants were identified in the two populations<sup>24</sup>. For the conditional analysis of the regional associations in the *ARID5B* locus (Supplementary Fig. 3), we repeated the meta-analysis at that locus by incorporating genotypes of the referenced SNP(s) as additional covariate(s). For comparison of the odds ratios of the SNPs, we first selected SNPs that were shared between the meta-analyses in Japanese and Europeans. Next, we removed the SNPs located more than 1 Mb away from each of the marker SNPs in the validated rheumatoid arthritis susceptibility loci, except for in the HLA region, where we removed the SNPs located between 24,000,000 bp to 36,000,000 bp on chromosome 6 because of the existence of long-range haplotypes with rheumatoid arthritis susceptibility in this region<sup>28</sup>. LD pruning of the SNPs was conducted for the SNP pairs that were in LD ( $r^2 \geq 0.3$ ) in both HapMap Phase 2 East Asian and Utah residents of Northern and Western European ancestry (CEU) populations (release 24). Correlations of the odds ratios were evaluated using R statistical software version 2.11.0.

25. Okada, Y. *et al.* Identification of nine novel loci associated with white blood cell subtypes in a Japanese population. *PLoS Genet.* **7**, e1002067 (2011).
26. de Bakker, P.I. *et al.* Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* **17**, R122-R128 (2008).
27. Stranger, B.E. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39**, 1217-1224 (2007).
28. Okada, Y. *et al.* Contribution of a haplotype in the HLA region to anti-cyclic citrullinated peptide antibody positivity in rheumatoid arthritis, independently of HLA-DRB1. *Arthritis Rheum.* **60**, 3582-3590 (2009).





## 第5章 リウマチ性疾患の個別化医療

鎌谷直之\*1, 谷口敦夫\*2

### 1 要旨

個別化医療は患者すべてに同じ治療法を適用するのではなく、個人個人に合った治療法を行う試みである。リウマチ性疾患は全身の運動器に痛みとこわばりを来たす疾患の総称である。リウマチ性疾患の中で変形性関節症、関節リウマチ、痛風、全身性エリテマトーデスなどの発症に関連するゲノム多型が多数報告されている。また、関節リウマチに最も多く用いられる薬物、メトトレキサートの効果と副作用に関連する SNP、および同様に有効な薬剤であるスルファサラジンの副作用に関連するハプロタイプが知られている。更には、日本人の関節リウマチの死因として重要な二次性アミロイドーシスのリスクに関係する SNP が知られている。我々は、それぞれの証拠の validation (確認) を行った後、それらを利用した関節リウマチを対象とした個別化医療を東京女子医科大学において 2005 年から開始した。既に、その有効性を示唆する結果が得られている。ゲノム多型に基づいた個別化医療が実施されているのはまだ限られた疾患の限られた施設においてであるが、これが広く応用されれば重症副作用は大幅に減少し、今より効果的な医療が行われると予想される。しかし、ゲノム多型と表現型との関連の正否や、その臨床介入の是非については疫学、統計学に基づく洗練された判断が要求され、信頼性の乏しい手法が氾濫すれば、かえって医療界を混乱させると予想される。我々が提唱する、遺伝的多型を臨床応用するためのロードマップを紹介する。

### 2 リウマチ性疾患とバイオマーカー

リウマチ性疾患は、全身の運動器（骨、筋肉、関節など）に痛みとこわばりを来たす疾患の総称である<sup>1)</sup>。変形性関節症、関節リウマチ、痛風、膠原病などを含む。変形性関節症は老人に多く、程度の差はあれ変形性関節症に罹患している老人は多い。関節リウマチは、日本に 40~70 万の患者が居ると推定されている。次に多い疾患は痛風であり、日本における推定患者数は 30

---

\*1 Naoyuki Kamatani 東京女子医科大学 附属膠原病リウマチ痛風センター 所長；教授

\*2 Atsuo Taniguchi 東京女子医科大学 附属膠原病リウマチ痛風センター 助教授

～40万人である。膠原病は色々の疾患の総称であり、一つ一つの疾患に罹患している患者の数は関節リウマチや痛風ほど多くないが、ほとんどは原因不明の疾患であり、生命予後に関係することも多いので問題は小さくない。全身性エリテマトーデス、全身性硬化症（強皮症）、皮膚筋炎、多発性動脈炎などは膠原病に含まれる。また、ウェジナー肉芽腫症、顕微鏡的血管炎などはバイオマーカーが有用な膠原病として知られる。

それぞれの疾患の、疾患バイオマーカーは多く知られている。例えば、関節リウマチ患者の70～80%はリウマトイド因子という、特定の抗原に反応する抗体を持っている<sup>1)</sup>。痛風は血清尿酸値の上昇が特徴的であり、これが原因として重要である。膠原病にはそれぞれの疾患に特異的とまでは言えないが、診断に参考となる自己抗体が存在する。例えば、全身性エリテマトーデスには抗二本鎖DNA抗体、Sm抗体などが比較的特異的として知られる。また、前述のウェジナー肉芽腫症にはPR3-ANCA、顕微鏡的血管炎にはMPO-ANCAが比較的特異的な抗体として知られ、これらの抗体は治療により軽症になると低下することが知られており、治療経過のバイオマーカーとして重要である。

### 3 リウマチ性疾患感受性とゲノム多型

いくつかのリウマチ性疾患では疾患に罹患しやすいゲノム多型が知られている。例えば、関節リウマチではHLA-DRB1座位の特定の多型が疾患感受性を増加させることが古くから知られている<sup>2)</sup>。また、最近のゲノム研究の進歩により、PADI4遺伝子<sup>3)</sup>、FCRL3遺伝子<sup>4)</sup>、NFKBIL1遺伝子<sup>5)</sup>、PTPN8遺伝子<sup>6)</sup>、MHC2TA遺伝子<sup>7)</sup>の多型が関節リウマチ発症と関連していることが示された。このうち、PADI4遺伝子、FCRL3遺伝子の多型は全ゲノムに配置された極めて多くのSNPの検索により発見されたものである。

また、痛風については以前からHPRT遺伝子、PRPP合成酵素遺伝子などが疾患感受性遺伝子として知られていたが、連鎖解析<sup>8)</sup>によりUMOD遺伝子が発見され<sup>9)</sup>、腎尿細管における尿酸の輸送を担当するURAT1をコードするSLC22A12遺伝子の多型<sup>10)</sup>が痛風発症を抑制することもわかっている<sup>11)</sup>。変形性関節症についてはFRZB遺伝子<sup>12)</sup>、ASPN遺伝子<sup>13)</sup>、CILP遺伝子<sup>14)</sup>、CALM1遺伝子<sup>15)</sup>などが疾患感受性遺伝子として報告されている。全身性エリテマトーデスについては、FCGR2A遺伝子<sup>16)</sup>、PDCD1遺伝子<sup>17)</sup>、PTPN22遺伝子<sup>18)</sup>、CTLA-4遺伝子<sup>19)</sup>、TNF $\alpha$ 遺伝子<sup>20)</sup>などとの関連が報告されている。さらに若年性関節炎の疾患感受性遺伝子としてIL-6遺伝子<sup>21)</sup>、MIF遺伝子<sup>22)</sup>、IL-18遺伝子<sup>23)</sup>が報告されている。さらに最近、関節リウマチについてマイクロサテライトを用いた全ゲノム検索から多くの感受性多型の候補が報告された<sup>24)</sup>。

歴史的に最も古くは、強直性脊椎炎がHLA-B座位のB27アレルと強く関連していることが

知られている<sup>25)</sup>。

#### 4 Wright のモデル

一般に、これらの多型と疾患の関係については感受性のゲノム遺伝子型の頻度と遺伝的効果の間に図1のような関係が知られている。即ち、遺伝的効果の大きいゲノム多型ほど、感受性のゲノム遺伝子型頻度が低い傾向がある。Wrightはこのような傾向を認め、表現型に関連するゲノム多型を3つに分類している。主要遺伝子 (major gene)、稀少遺伝子 (oligogene)、ポリジーン (polygene) である。主要遺伝子はメンデル型遺伝に従って遺伝継承することがほとんどであり、疾患関連多型を持っているほとんどの個体が発症する。これらの多型の頻度は極めて低いのが特徴的である。稀少遺伝子は比較的頻度が低く、遺伝的効果も中間的である。ポリジーンは頻度は高いが遺伝的効果は極めて低い(図1)。上記の遺伝子のうち、痛風の HPRT 遺伝子, PRPP 合成酵素遺伝子, UMOD 遺伝子などは疾患のリスクの原因となる多型の頻度は極めて低く、それらのリスクアレルを持ち、感受性のある個体のほとんどは疾患を発症する。そして感受性があるかどうかはリスク座位に関する遺伝子型によることが特徴で、優性、劣性などの遺伝継承モードに従う。いわゆる遺伝病に属する疾患である。それに比較し、それ以外のほとんどの遺伝子は上記の稀少遺伝子またはポリジーンに属する。リスク型の遺伝子型を持っている個体の頻度は高いものの、持っても発症する確率は低い。これらの表現型は複雑な形質 (complex trait)、または多因子疾患と呼ばれる。

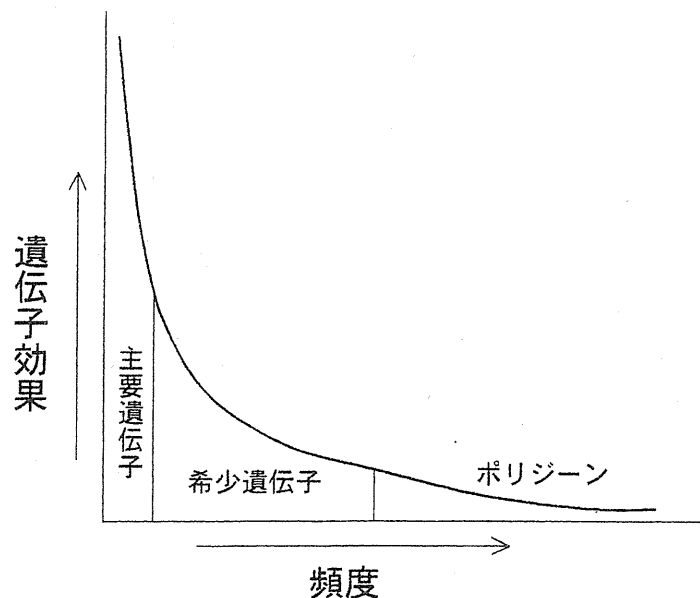


図1 遺伝子効果の大きさによる Wright の遺伝子の分類

遺伝的効果の大きさを示す指標の一つが遺伝子型相対リスク (genotype relative risk) である。これはリスクの高い遺伝子型とそれ以外の遺伝子型の個体の浸透率 (特定の表現型を取る確率) の比であり、これが高いほど遺伝的効果が高い。リウマチ性疾患で発見されている上記の稀少遺伝子、ポリジーンのほとんどは遺伝子型相対リスクが 1.5 を超えない。即ち、リスク型の遺伝子型を持っていても、それ以外の人々よりせいぜい 1.5 倍発症する確率が高いにすぎない。

さらに、これらの弱い効果の遺伝的多型については、関連の科学的妥当性についても注意して解釈する必要がある。一般に、これらの遺伝的多型と疾患との関連はケース・コントロール研究と言う研究形式によりなされることがほとんどである。即ち、多数の患者と多数の患者以外の人々を集め、遺伝子型を決定して特定の (例えばリスク型の) 遺伝子型や、特定のアレル (例えばリスクアレル) の頻度の差を統計的に検定する。強い遺伝的効果があればその差は大きく、検定結果に問題は無い。しかし、弱い遺伝的効果の場合はその差は微妙であり、一つの研究の結論のみで関連を断定することは困難である。このような場合は、再現研究 (replication study) で結論が再現される<sup>26~29)</sup>、メタ解析で差が証明される<sup>30)</sup>などの証拠がないと関連が確実と断定することは困難である。これらの複雑な形質においては、遺伝的多型が長い年月を通して疾患を引き起こしたり、引き起こさなかったりする。そもそも、これらの多型に関連して極めて大きな生物学的差異が実験により証明できるという保証は無い。多型が分子生物学や細胞レベルの機能の差に関連していることを生物学的実験により証明することも重要であるが、それだけでは関連が証明されたとはいえず、むしろこれから再現研究やメタ解析で確実に遺伝的に差があることを確認すること (validation) の重要性が高まるであろう。

## 5 重要な重症化関連多型と薬物反応性関連多型

遺伝的効果の弱い疾患感受性の稀少遺伝子、ポリジーンについては直接の臨床的有効性は低い。例えば、関節リウマチは発症確率が 1% に満たない疾患であるが、特定の遺伝子型の場合、一般人より 1.5 倍発症しやすいとわかっても、それほど参考にはならないからである。しかも、既に関節リウマチに罹患した患者にとって、疾患感受性遺伝子を保有しているかどうかは、あまり興味のないことである。一般の人々にとっても、1% の確率で発症する疾患の確率が 1.5% になっても、それほど有用な情報とはいえない。

しかし、直接の臨床的有効性に乏しい多型であっても、創薬の標的として有望であることが多い。疾患の発症に関連している分子生物学的要因は治療に有効な薬物反応に関連していることがしばしばあるからである。

疾患感受性に関連する遺伝的多型以外にも、医学に関連した遺伝的多型が知られている。中で

も重症化関連多型と薬物反応性関連多型が興味深い。

重症化関連多型とは、特定の遺伝子型を保有する患者において疾患が重症化しやすい、あるいは合併症を来しやすいような多型である。薬物反応性関連多型とは、特定の遺伝子型を保有する患者が特定の薬物を服用した場合、副作用がでやすい、あるいは効果がでにくいような多型である。このような多型であれば、既に疾患に罹患した患者にとって有用である。例えば、合併症がでやすい遺伝子型を持つ場合は合併症を防ぐ対策をより強く取る必要があるだろう。副作用が出やすいような遺伝子型を持つ場合はより慎重な投薬や、用量を減らす必要がある可能性がある。効果が出にくい場合は、他の薬物の選択や、用量の増量などに参考になる可能性がある。

例えば、前述の HLA-DRB 座位の関節リウマチの疾患感受性の遺伝子型の保有者は、それ以外の患者より重症化しやすい傾向がある事がわかっている<sup>31,32)</sup>。また、関節リウマチの重症の合併症である二次性アミロイドーシスと SAA1 遺伝子多型との関連が明らかになった<sup>33~37)</sup>。最近、興味深い報告が痛風の治療薬アロプリノールの副作用についてなされた。即ち、アロプリノールによる重症皮膚副作用と HLA-B\*5801 アレルの関連が報告された<sup>38)</sup>。

## 6 薬物の副作用の問題化

薬物に副作用があることは常識であったが、その実態は不明であった。個々の症例におきる障害については、それが薬物療法の副作用なのか、あるいは疾患の合併症によるものか不明の事が多い。1998年に JAMA に発表されたメタ分析の結果は衝撃的なものであった。この論文は 1966年から 1996年に発表された 39の米国の病院からの前向き研究の結果をカナダの動物学教室の研究者がメタ解析して発表したものである。それによると米国の入院患者の中で重症の薬物副作用の率は 6.7% (95% CI 5.2~8.2%)、致死的な副作用の率は 0.32% (95% CI 0.23~0.41%) と見積もられた。合衆国における 1994年の入院患者の副作用の発生人数は 2,216,000人、致死的な副作用は 106,000人で、死因の 4~6番目と推定された<sup>39)</sup>。交通事故による死亡と同じか、それより多いという予測である。この論文は世界の医学、薬物学の分野に大きな衝撃をもたらした。薬物治療の良い面だけでなく、悪い面も無視できないことを世界に示した論文であった。

また、薬物が無効の例も無視できないほど多いことも問題となっている。実際に抗リウマチ薬の有効性は高いとは言えず、効果の無い患者には深刻な問題である。しかも重要なことは、情報公開が進んできたということである。薬物による副作用や無効の問題は医師の側からはある程度は予測はついても、これまで患者や家族の側からそれを明らかに知ることは困難であった。しかし、米国を中心にインターネットなどで情報公開が進んだ結果、患者や家族側にも、このような薬物治療の問題点が明らかになった。

以上のような状況により、特に米国において、医学界は薬物による副作用と無効を、できる限り減らす研究を行うべきであるという考えが支配的になった。もちろん、全員に効果があり、副作用が全く無い化合物が開発されれば理想の薬物となるが、純系ではなく遺伝的に雑多な人間において、一つの化合物が全く同じ作用を示すことはありえないという考えが理解されるようになった (One size does not fit all)。これが個別化医療の推進に大きく影響している。

## 7 ゲノム情報の急速な解明

ヒトの約 30 億個の全ゲノム配列の解明が 2003 年までにほぼ終了し (Human genome project), SNP (single nucleotide polymorphism) の情報も蓄積されている (JSNP ; <http://snp.ims.u-tokyo.ac.jp/dbSNP> ; <http://www.ncbi.nlm.nih.gov/SNP/>)。さらに 3 人種のハプロタイプの情報も蓄積された (International HapMap project)<sup>40)</sup>。個体は受精卵から始まり、受精卵の情報はほとんど遺伝情報だけなのだが、そのすべてが個人ごとにわかる時代がすぐ来そうである。個人の身体的情報の基礎はゲノム配列である。それはわかっていたのだが、これまで医療はゲノム配列を観察せずに行ってきた。それが可能な新たな時代に我々は入ろうとしている。以上のような技術開発に支えられ、個別化医療が急速に進展するという期待が広がっている。

米国では行政を中心に既に個別化医療が動き始めた。米国 FDA は薬物の承認のために FDA にゲノムデータを提出するためのガイダンスのドラフトを発表し、治験にゲノムデータを含むことを推奨している (<http://www.fda.gov/cder/guidance/5900dft.pdf>)。これにより、海外の薬品会社では治験において個人の遺伝子検索を行う事が日常的になりつつある。

## 8 ヒトゲノム多型を臨床応用するための問題点

ヒトゲノム多型を臨床応用するための三つの問題点は、

- ① 正確な遺伝子型タイピング
- ② 統計的エビデンスの確実性
- ③ 臨床応用の有効性

である。①の、正確な遺伝子型タイピングについて、現在、DNA チップによる方法、インベーター法、ビーズ法、TaqMan 法など多数の方法が存在するがいずれも 100% 正確ではない。数少ない多型座位のタイピングであれば問題とならないが、多数の座位のタイピングを多数の個体について行うとすると、微細なタイピングエラーでも問題となる。タイピングの正確性を更に技術的に向上させるとともに、遺伝統計学的にタイピングミスを検出し、誤った結論に至らないよ

うに注意する必要がある。

②の問題はゲノム多型と表現型の関連に関するエビデンスの質の問題である。タイピングが完全に正確だとしても、表現型とゲノム多型の関連は統計的な証拠に基づかざるを得ないので、統計的な見地から正しい結論を導かなければならない。ゲノム多型により分子生物学的レベルで、あるいは細胞レベルで機能的に妥当な差があることを実験で証明することは望ましい。しかし、機能的な研究で結論が確実になるわけではない。実際に、多くの疾患は先天性異常と違って、何十年もかけて発症するのである。短時間の実験で証明できるような明確な違いがあると期待するのは常識的ではない。むしろ、遺伝統計的な違いが本当のものであるかどうかを確認することが(validation) 大切である。このようなゲノム多型を利用した関連解析の一般的注意として、(a) 臨床データの質、(b) 十分なサンプルサイズ、(c) 洗練された統計解析が不可欠である。これらのいずれが欠けても信頼できる結論を出すことはできない。

特に、膨大な数の多型座位(多くの場合はSNP)を用いて関連解析を行う場合は十分な注意が必要である。一つのSNPに関する関連の有無はFisherの正確法やPearsonの $\chi^2$ 分布を利用した独立性の検定で行われることが多いが、それらの検定によりP値が0.05より小さければ通常関連ありと結論される。しかし、多数のSNP座位について同様の関連解析を行えば、すべての座位が関連が無くても平均5%の割合でP値は0.05以下となる。これが多重比較(多重検定)の問題である。この問題の解決は容易ではないが、最も保守的な(有意と言う結論が出にくい)Bonferroniの補正法が通常用いられる。これは例えば $n$ 個のSNPについて検定を行うとすると、個々のSNPの検定について有意水準を $0.05/n$ として検定を行うというものである。しかし、実際にはBonferroniの補正を行うと過度に保守的な結論が得られることが知られている。即ち、実際に関連がある場合にも有意差が出ない場合が多い。これを修正するために、FDR(false-discovery rate)を用いる方法などが提唱されている<sup>41)</sup>。

また、複数のSNPが連鎖の関係にあると、そのSNPを用いた関連解析にはハプロタイプを基礎にした解析が必要である。実際に関節リウマチ治療薬スルファサラジンの副作用にハプロタイプが関係することがわかっている<sup>42,43)</sup>。多重検定の問題を解決したり、ハプロタイプに基づいた解析をするためには洗練された情報、統計処理の技術と知識が必要である。このような情報処理、統計解析の体制が日本では極端に弱いことが問題である。我々はそのような問題に対処するための新たなアルゴリズムを発表している(LDSUPPORT<sup>44)</sup>、LDPOOLED<sup>45)</sup>、PENHAPLO<sup>46)</sup>、QTLHAPLO<sup>47)</sup>)。多数の薬物関連遺伝子の信頼できる頻度を伴った多型情報の収集とハプロタイプ解析は世界に先駆けて日本で行われた<sup>48)</sup>。1,000名以上のボランティアより試料を採取し、752人の200以上の薬物関連遺伝子について4,000以上のSNP座位の情報が頻度付きデータベースとして提供されている(東京女子医科大学、理化学研究所、ファルマ・スニップ・コンソーシア

ム、[http://www.jpma.or.jp/psc/11\\_data/index.html](http://www.jpma.or.jp/psc/11_data/index.html))。ハプロタイプ解析の内容についても発表されている<sup>48)</sup>。

更に③の臨床応用の有効性の考察が極めて重要である。(a)と(b)の問題が解決し、遺伝的多型と表現型との関連が確実にあったとしても、それが本当に臨床応用の価値があるかを考察する必要がある。前述のように、疾患感受性ゲノム多型を既に疾患に罹患した患者にテストするメリットは考えにくい。健康診断などで疾患感受性ゲノム多型をテストする事業が今後出現する可能性があるが、当人が極めて低い確率で罹患する可能性のある疾患について、その感受性をテストするには慎重でなければならない。前述のように、1%の確率で罹患する可能性のある疾患に1.5%の確率で罹患することがわかったとしても当人にメリットはほとんど無いからである。そのテストを臨床応用することによりどの程度の被験者のメリットが期待できるかを客観的手法により評価する必要がある。

## 9 ヒトゲノム多型データを臨床応用するためのロードマップ

我々は上記のヒトゲノム多型を臨床応用するための三つの問題点を解決するため、ヒトゲノム多型データを臨床応用するためのロードマップを提唱している。

即ち、

- (i) 正当な統計的手法を用い、仮説検定により遺伝情報により表現型が異なることが証明されること
- (ii) その結果が独立のサンプルで再確認されること
- (iii) それに基づきゲノム情報を治療に応用するアルゴリズムが構成でき、医療介入の結果が患者に有利であると予測されること

である<sup>49)</sup>。このような基準は、不確かなエビデンスが臨床応用されることを防ぐために、さらにはエビデンスはあっても患者に利益の無い医療介入を防ぐためにも是非必要である。

## 10 個別化医療に関する我々の試み

膠原病リウマチ痛風センターでは2005年から以下の4項目について関節リウマチ患者に対して個別化医療を開始した。項目は、

- (a) aMTHFR 遺伝子の C 677 T 多型によりメトトレキサート (MTX) の副作用を予測する
- (b) MTHFR 遺伝子の A 1298 C 多型により MTX の必要用量を予測する
- (c) NAT 2 遺伝子のハプロタイプによりスルファサラジンの副作用を予測する



(d) SAA1 遺伝子の-13 T/C 多型によりアミロイドーシスの発症を予測する  
である。

前の節で、ヒトゲノム多型データを臨床応用するためのロードマップとして、(i) (ii) (iii) の三つの必要条件を挙げた。我々はこの三つの条件をクリアするため研究を進めてきて、それが満足できるようになったため現実に個別化医療の臨床応用を始めたのである。

まず、我々は (a) の C 677 T 多型と MTX の副作用の関係、(b) の A 1298 C 多型と MTX の必要容量の関係を報告した<sup>49)</sup>。また、(c) の NAT 2 遺伝子のハプロタイプの保有とスルファサラジンの副作用の関連について報告した<sup>42)</sup>。SAA 1 遺伝子と続発性アミロイドーシスの合併の関連についても報告した<sup>35,37)</sup>。これにより、(i) の条件は (a) ~ (d) のすべての項目についてクリアした。次に、我々は、(a) (b) (c) の3項目について、独立のサンプルで再確認することに成功した<sup>43)</sup>。しかも、これらについては我々以外のグループからも同様の報告がなされている<sup>50,51)</sup>。(d) については、日本での別のグループによる研究で再確認されている<sup>36)</sup>。最後の (iii) の項目についての詳細については述べないが、例えば、(c) の NAT 2 遺伝子のハプロタイプとスルファサラジンの副作用との関連については、「ハプロタイプ推定アルゴリズム<sup>44,46)</sup>により、NAT 2 のリスクタイプのハプロタイプと判定された場合、スルファサラジンの投与開始をしない」というアルゴリズムを設定している。

これまで、関節リウマチの個別化医療を開始した後、約 10 か月になるが、その有用性が強く示唆されている。例えば、(c) のスルファサラジンの副作用を NAT 2 遺伝子のハプロタイプで予測する、という項目について、このアルゴリズムを厳密に適用すればスルファサラジンの重症副作用を半分以下にできることが示唆された。

## 11 まとめ

いくつかのリウマチ性疾患で、疾患感受性に関連するゲノム多型が報告されている。また、薬物反応性に関連するゲノム多型や、合併症発症に関連するゲノム多型が発表されている。我々はゲノム多型情報に基づいた個別化医療を開始するための必要条件を規定したロードマップを提唱している。それに基づいて、東京女子医科大学では実際にメトトレキサートの副作用と必要容量、スルファサラジンの副作用、合併症アミロイドーシスに関連するゲノム多型を利用した個別化医療を開始した。ゲノム多型に基づいた個別化医療の応用により、副作用の低減や効果の増大などの極めて多くのメリットが期待できる。

文 献

- 1) 鎌谷直之, 慢性関節リウマチ, 鎌谷直之編 Evidence Based Medicine を活かす膠原病・リウマチ診療, pp.22-51, メディカルビュー社 (2000)
- 2) HLA and RA 鎌谷直之編 Evidence Based Medicine を活かす膠原病・リウマチ診療, メディカルビュー社 (2000)
- 3) A. Suzuki *et al.*, *Nat. Genet.*, 34, 395 (2003)
- 4) Y. Kochi *et al.*, *Nat. Genet.*, 37, 478 (2005)
- 5) K. Okamoto *et al.*, *Am. J. Hum. Genet.*, 72, 303 (2003)
- 6) N. Bottini *Nat. Genet.*, 36, 337 (2004)
- 7) M. Swanberg *et al.*, *Nat. Genet.*, 37, 486 (2005)
- 8) N. Kamatani *et al.*, *Arthritis Rheum.*, 43, 925 (2000)
- 9) T.C. Hart *et al.*, *J. Med. Genet.*, 39, 882 (2002)
- 10) A. Enomoto *et al.*, *Nature*, 417, 447 (2002)
- 11) A. Taniguchi *et al.*, *Arthritis Rheum.*, 52, 2576 (2005)
- 12) J. Loughli *et al.*, *Proc. Nat. Acad. Sci.*, 101, 9757 (2004)
- 13) H. Kizawa *et al.*, *Nat. Genet.*, 37, 138 (2005)
- 14) S. Seki *et al.*, *Nat. Genet.*, 37, 607 (2005)
- 15) H. Mototani *et al.*, *Hum. Mol. Genet.*, 14, 1009 (2005)
- 16) J. E. Salmon *et al.*, *J. Clin. Invest.*, 97, 1348 (1996)
- 17) L. Prokunina *et al.*, *Nat. Genet.*, 32, 666 (2002)
- 18) C. Kyogoku *et al.*, *Am. J. Hum. Genet.*, 75, 504 (2004)
- 19) L. L. Hudson *et al.*, *Hum. Genet.*, 111, 452 (2002)
- 20) Y. H. Lee *et al.*, *Europ. J. Hum. Genet.*, 14, 364 (2006)
- 21) D. Fishman *et al.*, *J. Clin. Invest.*, 102, 1369 (1998)
- 22) R. P. Donn *et al.*, *Arthritis Rheum.*, 44, 1782 (2001)
- 23) T. Sugiura *et al.*, *Arthritis Res Ther.*, 8, R 60 (2006)
- 24) G. Tamiya *et al.*, *Hum Mol Genet.*, 14, 2305 (2005)
- 25) S. Sengupta *et al.*, *Lancet* I, 1209 (1977)
- 26) K. Ikari *et al.*, *Arthritis Rheum.*, 52, 3054 (2005)
- 27) K. Ikari *et al.*, *Ann Rheum Dis.*, 65, 671 (2006)
- 28) S. Ikegawa *et al.*, *Arthritis Res Ther.*, 8, 403 (2006)
- 29) K. Ikari *et al.*, *Rheumatology (Oxford)* Epub ahead of print (2006)
- 30) T. Iwamoto *et al.*, *Rheumatology (Oxford)*, 45, 804 (2006)
- 31) C. Seidl *et al.*, *Br J Rheumatol.*, 36, 941 (1997)
- 32) T. Mottonen *et al.*, *Ann Rheum Dis.*, 57, 533 (1998)
- 33) S. Baba *et al.*, *Hum Mol Genet.*, 4, 1083 (1995)
- 34) M. Moriguchi *et al.*, *Hum Genet.*, 105, 360 (1999)
- 35) M. Moriguchi *et al.*, *Arthritis Rheum.*, 44, 1266 (2001)
- 36) T. Yamada *et al.*, *Amyloid.*, 10, 7 (2003)

第5章 リウマチ性疾患の個別化医療

- 37) M Moriguchi *et al.*, *Amyloid.*, 12, 26 (2005)
- 38) SI Hung *et al.*, *Proc Natl Acad Sci USA.*, 102, 4134 (2005)
- 39) J Lazarou *et al.*, *JAMA*, 279, 1200 (1998)
- 40) International HapMap Consortium: Integrating ethics and science in the International HapMap Project. *Nat Rev Genet.*, 5, 467 (2004)
- 41) D Yekutieli *et al.*, *J. Statistical Planning and Inference*, 82, 171 (1999)
- 42) E Tanaka *et al.*, *J Rheumatol.*, 29, 2492 (2002)
- 43) A Taniguchi *et al.*, *Pharmacogenet Genomics* in press.
- 44) Y Kitamura *et al.*, *Ann Hum Genet.*, 66, 183 (2002)
- 45) T Ito *et al.*, *Am J Hum Genet.*, 72, 384 (2003)
- 46) T Ito *et al.*, *Genetics*, 168, 2339 (2004)
- 47) K Shibata *et al.*, *Genetics*, 168, 525 (2004)
- 48) N Kamatani *et al.*, *Am J Hum Genet.*, 75, 190-203 (2004)
- 49) W Urano *et al.*, *Pharmacogenetics*, 12, 183 (2002)
- 50) AE Van Ede *et al.*, *Arthritis Rheum.*, 44, 2525 (2001)
- 51) H Schmeling *et al.*, *J Rheumatol.*, 32, 1832 (2005)

# 1. ゲノム情報

## —その先制医療への展望—

[株式会社スタージェン情報解析研究所] 鎌谷直之

### 1 ゲノムと健康、医療との関係

ゲノム(genome)とは、もともと Winkler が gene(遺伝子)と chromosome(染色体)を結合させた造語として作成したものであるが、現在では特定の種の、遺伝する情報のワンセットを意味することが多い。初期には染色体レベルで議論されたが、最近では多くの種の全ゲノム配列が明らかになってきたので、ゲノム配列そのものをゲノムということも多くなっている。最近ゲノムということばが多用されるようになった背景には、遺伝子ということばで、遺伝する情報のすべてを含むのは困難であるとわかったことがある。蛋白をコードしている領域(エクソン)は全ゲノム配列の2%以下にすぎず、コードしていない領域(イントロンなど)も含んだ遺伝子の領域も20~30%程度である。そのため、遺伝子以外の遺伝する領域を含んだ全体をゲノムとよぶことが多くなった<sup>1)</sup>。

ゲノムということばは両親由来の情報である生殖細胞系系(germline)だけではなく、体細胞の保有する、ゲノム配列に対応する情報にも適用される。たとえば癌細胞のゲノム配列は本人の生殖細胞系系ゲノム配列に比較して数千~数万個も異なるが<sup>1)</sup>、そのように世代を超えて伝わらない情報もゲノム情報とよばれる(たとえば、国際癌ゲノムコンソーシアム(ICGC))<sup>2)</sup>。しかし、DNAであっても機械で合成したDNAや、mRNAを用いて逆転写酵素により合成した、たとえばcDNAはゲノムとはよばない。

ゲノム配列はATCGの4種類の文字で、ヒトでは約30億文字ある。この30億文字は23種類の染色体(22個の常染色体、1つの性染色体)に分かれて書かれている。これがいかに膨大かを知るためには、この文字を通常の大きさ(1cmに3文字)で書くと東京からニューヨークに達することを知れば十分である<sup>3)</sup>。ゲノム情報は父親から約30億個(ハプロイドゲノム)、母親から約30億個のヌクレオチド情報が子に伝わる。そして子の体内で組み換えが起き、変異(mutation)が起きたうえで次の世代に30億個が配偶子(精子、卵)に入って伝えられる。

精子と卵の受精により受精卵ができ、この中に30億個×2のゲノム情報が入っている。細胞分裂により体細胞となったものは受精卵とはほとんど同じゲノム情報をもつが、次第に変異を蓄積し癌となる細胞も出る。

ゲノム情報の特殊性は、それがほとんど唯一、次世代に伝わる情報ということである。たとえばmRNA、蛋白などは次世代に伝わることはない。ただし、ゲノム情報の中にmRNAや蛋白の情報が含まれるため、次世代でも同様のmRNA、蛋白が合成される。

ゲノム情報はこのように種において基本的な情報であるため、医療や健康維持において様々な有益な情報を提供すると期待されている。ゲノム情報が身体に影響を及ぼすためにはmRNA、蛋白、小分子、細胞、臓器、個体という各レベルを通じた効果が必要であると考えられる。たとえば、1型糖尿病は膵臓のランゲル