J Hum Genet (2005) 50:264–266
DOI 10.1007/s10038-005-0246-8

SHORT COMMUNICATION

**Mikako Mori · Ryo Yamada · Kyoko Kobayashi
Reimi Kawaida · Kazuhiko Yamamoto**

Ethnic differences in allele frequency of autoimmune-disease-associated SNPs

Received: 28 December 2004 / Accepted: 14 March 2005 / Published online: 10 May 2005

Abstract Several multiple, large-scale, genetic studies on autoimmune-disease-associated SNPs have been reported recently: peptidylarginine deiminase type 4 (*PADI4*) in rheumatoid arthritis (RA); solute carrier family 22 members 4 and 5 (*SLC22A4* and *5*) in RA and Crohn's disease (CD); programmed cell death 1 (*PDCD1*) in systemic lupus erythematosus (SLE), type 1 diabetes mellitus (T1D), and RA; and protein tyrosine phosphatase non-receptor type 22 (*PTPN22*) in T1D, RA, and SLE. Because these reports on association were not always evaluated in multiple ethnic groups and because ethnic difference in allele frequency of the variants has been also reported, we investigated allele frequencies of nine SNPs in four autoimmune-disease-associated loci in Caucasian, African-descent, and Japanese populations. Although SNPs in *PADI4* had similar allele frequency among three groups [maximal difference 11%; ($P > 0.05$)], the other three loci revealed statistically significant allele frequency differences (maximal difference 39% ($P < 0.00001$), 13% ($P < 0.00001$), and 8% ($P < 0.00001$) in *SLC22A4*, *PDCD1*, and *PTPN22*, respectively). Of note, three SNPs in the three loci that had allele frequency more than 8% in the Caucasian population were either not polymorphic at all or extremely rare in the Japanese population. Our data suggest that ethnic variations of polymorphisms should be evaluated in detail, and differences should be incorporated into investigations of susceptibility variants for common diseases.

Keywords SNP · Ethnicity · Autoimmune disease · PADI4 · SLC22A4 · SLC22A5 · PTPN22 · PDCD1

M. Mori · R. Yamada (✉) · K. Kobayashi
R. Kawaida · K. Yamamoto
Laboratory for Rheumatic Diseases, SNP Research Center,
The Institute of Physical and Chemical Research,
1-7-22 Suehiro-cho, Tsurumi-ku,
Yokohama, Kanagawa 230-0045, Japan
E-mail: ryamada@src.riken.go.jp
Tel.: +81-45-5039569
Fax: +81-45-5039590

Introduction

Autoimmune diseases are common disorders, affecting more than 5% of the world population. They share aspects of aberrant immunological tolerance toward self-antigens, and all are known to have genetic components. Several multiple, large-scale, genetic studies on autoimmune-disease-associated SNPs have been reported recently: peptidylarginine deiminase type 4 (*PADI4*) in rheumatoid arthritis (RA) (Suzuki et al. 2003); solute carrier family 22 members 4 and 5 (*SLC22A4* and *5*) in RA (Tokuhiro et al. 2003) and Crohn's disease (CD) (Peltekova et al. 2004); programmed cell death 1 (*PDCD1*) in systemic lupus erythematosus (SLE) (Prokunina et al. 2002), type 1 diabetes mellitus (T1D) (Nielsen et al. 2003), and RA (Prokunina et al. 2004); and *PTPN22* in T1D (Bottini et al. 2004), RA (Begovich et al. 2004), and SLE (Kyogoku et al. 2004).

Interestingly, the allele frequencies of some of these autoimmune-disease-associated SNPs vary substantially in different ethnic groups, as reported for CD-associated SNPs in the Caucasian and Japanese populations (Yamazaki et al. 2004). Information on ethnic differences in allele frequency of disease-associated variants is important for better understanding of the pathologic mechanisms of polymorphisms. We therefore investigated the allele frequency of nine SNPs in four autoimmune-disease-associated genes/loci in Caucasian, African-descent, and Japanese populations.

Subjects and methods

Japanese volunteers were recruited, and their informed consent to the study was obtained, as required by our ethical committee and as previously described (Suzuki et al. 2003). Human variation panels for African American and Caucasian samples were obtained from Coriell Cell Repositories. We genotyped 376 Japanese,

94 African American, and 94 Caucasian subjects for the following nine SNPs in four loci: padi_89, 90, 92, and 104 in *PADI4* (Suzuki et al. 2003); slc2F1 and slc2F2 (Tokuhiro et al. 2003) and L503F in *SLC22A4* (Peltekova et al. 2004); G-207C in *SLC22A5* (Peltekova et al. 2004); PD-1.3A in *PDCD1* (Prokunina et al. 2002), and R620W in *PTPN22* (Bottini et al. 2004). Japanese data for the four SNPs in *PADI4* were identical to those reported previously (Suzuki et al. 2003). Genotyping was performed using the Invader assay, TaqMan assay, or direct sequencing, and haplotype frequency was inferred with an EM algorithm and/or Haplotyper (Niu et al. 2002). The statistical significance of differences in allele frequency was determined (P value in Fisher's exact test), and the genetic distance indicated by differences in allele frequency was quantitated as F_{ST}.

Results

Allele frequencies of the nine SNPs in the three ethnic groups are shown in Table 1. The largest difference in allele frequency among ethnic groups was 39% (L503F, C versus J). Three pairwise comparisons Caucasian versus Japanese (C versus J), African American versus Japanese (A versus J), and African American versus Caucasian (A versus C), were performed for the allele frequency of each SNP and CD-susceptible haplotype in the SLC22A4/5 locus. Of the nine SNPs in four loci, a significant difference in allele frequency ($P < 0.01$) was observed: five SNPs in three loci were significant in the C versus J comparison, five SNPs in three loci in the A versus J comparison, and two SNPs in two loci in the A versus C comparison. The F_{ST} values of more than 0.1 were obtained for four SNPs in two loci: one SNP in one locus was significant in the C versus J comparison, two SNPs in one locus in the A versus J comparison, and two SNPs in two loci in the A versus C comparison.

Discussion

The allele frequencies of the four SNPs in *PADI4* do not differ among the three ethnic groups, as indicated by a low F_{ST} value and an insignificant P value. However, SNPs in all the other three loci had allele frequencies that were significantly different in at least one of the comparisons among the three ethnic groups. The most remarkable difference was observed for the L503F SNP and haplotypes containing the SNP in *SLC22A4/5* in a comparison between the Caucasian and Japanese populations. The SNP and the corresponding haplotypes were much rarer in the Japanese population and were also rare in African Americans. Although other differ-

Table 1 Allele frequency of autoimmune-disease-associated variants in three ethnic groups

Gene name/ variant name	Caucasian	Japanese	African American	Caucasian versus Japanese		African American versus Japanese		African American versus Caucasian	
				P value	F_{ST}	P value	F_{ST}	P value	F_{ST}
PADI4									
padi_89	0.38 (0.40[a])	0.40[j]	0.49	0.86	0.00034	0.10	0.0079	0.078	0.12
padi_90	0.42 (0.41[a])	0.40[j]	0.46	0.86	0.00052	0.51	0.0037	0.88	0.0014
padi_92	0.38 (0.42[a])	0.39[j]	0.48	0.84	0.000053	0.023	0.0090	0.31	0.010
padi_104	0.31 (0.31[a])	0.33[j]	0.35	0.66	0.0056	0.70	0.00065	0.73	0.0024
SLC22A4									
slc2F1	0.09 (0.07[b])	0.29	0.04	< 0.00001*	0.064	< 0.00001*	0.12	0.14	0.012
slc2F2	0.09 (0.05[b])	0.29	0.04	< 0.00001*	0.066	< 0.00001*	0.12	0.14	0.012
L503F (C1672T)	0.39 (0.43[b])	0.001 (0.00[k])	0.04	< 0.00001*	0.24	< 0.00001*	0.018	< 0.00001*	0.19
SLC22A4/A5									
Haplotype TC	0.39 (0.42[c])	0.00 (0.00[k])	0.02	< 0.00001*	0.24	< 0.00001*	0.011	< 0.00001*	0.037
PDCD1									
PD-1.3A	0.13 (0.07[d]) 0.13 (0.07[e]) 0.13 (0.07[f])	0.00	0.03	< 0.00001*	0.068	0.00040*	0.014	0.0039*	0.036
PTPN22									
R620W	0.08 (0.09[g]) 0.08 (0.09[h]) 0.08 (0.12[i])	0.00	0.02	< 0.00001*	0.040	0.0032*	0.011	0.16	0.016

Data from other groups shown in *parenthesis*:
[a]British data derived from Barton et al. (2004)
[b]Canadian data derived from Newman et al. (2005)
[c]European data derived from Peltekova et al.(2004)
[d]Swedish data derived from Prokunina et al. (2004) and Prokunina et al. (2002)
[e]European American data derived from Prokunina et al. (2004)
[f]Danish data derived from Nielsen et al. (2003)
[g]North American data derived from Begovich et al. (2004)
[h]North American data derived from Kyogoku et al. (2004)
[i]North Amerian data derived from Bottini et al. (2004)
[j]Japanese data from our group, reported by Suzuki et al. (2003)
[k]Japanese data derived from Yamazaki et al. (2004)
* P value calculated by Fisher's exact test: $P < 0.01$

ences in allele frequency less outstanding than L503F in *SLC22A4/5*, it was of note that PD-1.3A in *PDCD1* and R620W in *PTPN22* were only polymorphic in Caucasians and in the African-descent population, as for L503F in *SLC22A4/5*. Populations of African descent, Caucasians, and Japanese (an ethnic subgroup from far-east islands), should have different population histories of mutation, migration, isolation, and genetic drift, and there is no doubt that there should be substantial differences in allele frequencies among these ethnic groups. However, an explanation of the remarkable difference seen for L503F in *SLC22A4/5* and others appears to require an additional mechanism. We note that Stefansson et al. (2005) reported that a 900-kb inversion polymorphism is frequently present in Caucasians but is rare in Africans and almost absent in East Asians, similar to the L503F SNP, and that prevalence of the variation in Caucasians was due to a positive effect in reproduction.

Beside the population genetics perspective, the difference in allele frequency among ethnic groups is important for association studies. Ioannidis et al. (2004) recently summarized multiple meta-analyses on common disease-susceptible polymorphisms from the perspective of ethnic variability. They concluded that the relative risk produced by susceptible variants seemed to be more similar among various ethnic groups compared with variability in their allele frequency. These findings would be correct in general but may not be applicable for polymorphisms that are common for some ethnic groups but absent in other groups. Ethnic uniformity and diversity in the identity of common disease-susceptible variants and their genetic contribution need to be extensively investigated before ethnic variations of polymorphisms are understood. Therefore, the details and differences in ethnic variations need to be incorporated into investigations of susceptibility variants for common diseases.

References

Barton A, Bowes J, Eyre S, Spreckley K, Hinks A, John S, Worthington J (2004) A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. Arthritis Rheum 50:1117–1121

Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, Huang Q, Smith AM, Spoerke JM, Conn MT, Chang M, Chang SY, Saiki RK, Catanese JJ, Leong DU, Garcia VE, McAllister LB, Jeffery DA, Lee AT, Batliwalla F, Remmers E, Criswell LA, Seldin MF, Kastner DL, Amos CI, Sninsky JJ, Gregersen PK (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 75:330–337

Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, Eisenbarth GS, Comings D, Mustelin T (2004) A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat Genet 36:337–338

Ioannidis JP, Ntzani EE, Trikalinos TA (2004) 'Racial' differences in genetic effects for complex diseases. Nat Genet 36:1312–1318

Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, Chang M, Ramos P, Baechler EC, Batliwalla FM, Novitzke J, Williams AH, Gillett C, Rodine P, Graham RR, Ardlie KG, Gaffney PM, Moser KL, Petri M, Begovich AB, Gregersen PK, Behrens TW (2004) Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am J Hum Genet 75:504–507

Newman B, Wintle RF, van Oene M, Yazdanpanah M, Owen J, Johnson B, Gu X, Amos CL, Keystone E, Rubin LA, Siminovitch KA (2005) SLC22A4 polymorphisms implicated in rheumatoid arthritis in a Canadian Caucasian population. Arthritis Rheum 52:425–429

Nielsen C, Hansen D, Husby S, Jacobsen BB, Lillevang ST (2003) Association of a putative regulatory polymorphism in the PD-1 gene with susceptibility to type 1 diabetes. Tissue Antigens 62:492–497

Niu T, Qin ZS, Xu X, Liu JS (2002) Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. Am J Hum Genet 70:157–169

Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, Griffiths AM, St George-Hyslop PH, Siminovitch KA (2004) Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet 36:471–475

Prokunina L, Castillejo-Lopez C, Oberg F, Gunnarsson I, Berg L, Magnusson V, Brookes AJ, Tentler D, Kristjansdottir H, Grondal G, Bolstad AI, Svenungsson E, Lundberg I, Sturfelt G, Jonssen A, Truedsson L, Lima G, Alcocer-Varela J, Jonsson R, Gyllensten UB, Harley JB, Alarcon-Segovia D, Steinsson K, Alarcon-Riquelme ME (2002) A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. Nat Genet 32:666–669

Prokunina L, Padyukov L, Bennet A, de Faire U, Wiman B, Prince J et al (2004) Association of the PD-1.3A allele of the PDCD1 gene in patients with rheumatoid arthritis negative for rheumatoid factor and the shared epitope. Arthritis Rheum 50:1770–1773

Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, Baker A, Jonasdottir A, Ingason A, Gudnadottir VG, Desnica N, Hicks A, Gylfason A, Gudbjartsson DF, Jonsdottir GM, Sainz J, Agnarsson K, Birgisdottir B, Ghosh S, Olafsdottir A, Cazier JB, Kristjansson K, Frigge ML, Thorgeirsson TE, Gulcher JR, Kong A, Stefansson K (2005) A common inversion under selection in Europeans. Nat Genet 37:129–137

Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 34:395–402

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) 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

Yamazaki K, Takazoe M, Tanaka T, Ichimori T, Saito S, Iida A, Onouchi Y, Hata A, Nakamura Y (2004) Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. J Hum Genet 49:664–668