18

CD14 and IL4R gene polymor-phisms modify the effect of day care attendance on serum IgE levels

Suzuki, Y., Hattori, S., Mashimo, Y., Funamizu, M.,
Kohno, Y., Okamoto, Y., Hata, A., Shimojo, N.

J Allergy Clin Immunol, 123, 1408-1411, e1401, 2009

2007, at the patient's request. After approximately 1½ years of normal eosinophil levels, the patient's eosinophilia worsened during late 2008 (Fig 1). A single osseous plasmacytoma was found and irradiated but progressed to multiple lytic lesions and multiple myelomas confirmed by means of bone marrow biopsy. The patient died in December 2008 as a result of this disease.

HES can occur as a myeloproliferative, lymphoproliferative, or, most frequently, idiopathic variant. Some myeloproliferative patients respond to imatinib mesylate[3] and possess a mutant gene on chromosome 4.[4] Deletion of approximately 800 kB on chromosome 4 results in a *FIP1L1/PDGFRA* fusion gene and formation of a kinase potently inhibited by imatinib mesylate. In our patient the *FIP1L1/PDGFRA* fusion gene was not detected, and imatinib mesylate failed to control eosinophilia. In the lymphoproliferative HES variant, T-lymphocyte clones produce cytokines, especially IL-5, that stimulate eosinophil production in the bone marrow. Mepolizumab, an anti–IL-5 drug, is used to inhibit eosinophil proliferation stimulated by IL-5; however, for unknown reasons, 16% of patients do not respond to mepolizumab.[5] In this case mepolizumab had no effect, suggesting that eosinophilia was not IL-5 dependent or that other cytokines, such as IL-3 or GM-CSF, were supporting eosinophil growth. Alternatively, the patient might have been producing so much IL-5 that the levels might have outpaced mepolizumab injections. IFN-α treatment is often effective because of a shift in the cytokine milieu from T_H2, which is supportive of eosinophil growth, to a T_H1-type response. In this case IFN-α caused a decrease in eosinophil counts, although not to normal levels, and the patient experienced the side effects of IFN-α.

Alemtuzumab is an anti-CD52 antibody that can bind to both eosinophils and T cells, potentially inhibiting either the myeloproliferative, lymphoproliferative, or idiopathic variant.[6] CD52 is a glycosylphosphatidylinositol-anchored molecule expressed on human eosinophils, lymphocytes, macrophages, and monocytes but not on neutrophils.[7] Alemtuzumab is approved for the treatment of B-cell chronic lymphocytic leukemia and is also used to treat small lymphocytic lymphoma and mantle cell lymphoma in conjunction with other treatments. Side effects include infusion reactions (often severe), lymphopenia, anemia, thrombocytopenia, and infections. Alemtuzumab was used successfully as a treatment for HES in 2 prior cases, 1 lymphoproliferative and 1 myeloproliferative, both of which did not respond to imatinib mesylate or IFN-α and that were not tested for the *FIP1L1/PDGFRA* fusion.[8,9] Alemtuzumab controlled our patient's eosinophilia for 1½ years, and the patient's quality of life appeared improved. Our patient most likely had the idiopathic HES variant. However, the occurrence of thromboembolism and an increased B12 level point to a possible myeloproliferative HES variant. The patient had a plasmacytoma and then multiple lytic lesions and multiple myelomas, suggesting involvement of 2 cell lineages by a single mutation or possibly independent mutations. Overall, the results in our patient and the previously reported cases suggest that alemtuzumab might be an effective treatment for the myeloproliferative, idiopathic, and lymphoproliferative HES variants.

Lori A. Wagner, PhD[a]
Stephen Speckart, MD[c]
Bruce Cutter, MD[c]
Gerald J. Gleich, MD[a,b]

From the Departments of [a]Dermatology and [b]Medicine, the University of Utah School of Medicine, Salt Lake City, Utah, and [c]Montana Oncology, Missoula, Mont. E-mail: lori.wagner@hsc.utah.edu.

Disclosure of potential conflict of interest: L. A. Wagner has received research support from the National Institutes of Health. G. J. Gleich has received research support from the National Institutes of Health and Dyax. The rest of the authors have declared that they have no conflict of interest.

REFERENCES

1. Klion AD, Bochner BS, Gleich GJ. Approaches to the treatment of hypereosinophilic syndromes: a workshop summary report. J Allergy Clin Immunol 2006;117:1292-302.
2. Roufosse F, Cogan E, Goldman M. The hypereosinophilic syndrome revisited. Annu Rev Med 2003;54:169-84.
3. Gleich GJ, Leiferman KM, Pardanani A, Tefferi A, Butterfield JH. Treatment of hypereosinophilic syndrome with imatinib mesilate. Lancet 2002;359:1577-8.
4. Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, et al. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. N Engl J Med 2003;348:1201-14.
5. Rothenberg ME, Klion AD, Roufosse FE, Kahn JE, Weller PF, Su H, et al. Treatment of patients with hypereosinophilic syndrome with mepolizumab. N Engl J Med 2008;358:1215-28.
6. Dumont FJ. Alemtuzumab (Millennium/Ilex). Curr Opin Investig Drugs 2001;2:139-60.
7. Elsner J, Hochstetter R, Spiekermann K, Kapp A. Surface and mRNA expression of the CD52 antigen by human eosinophils but not by neutrophils. Blood 1996;88:4684-93.
8. Pitini V, Teti D, Arrigo C, Righi M. Alemtuzumab therapy for refractory idiopathic hypereosinophilic syndrome with abnormal T cells: a case report. Br J Haematol 2004;127:477.
9. Sefcick A, Sowter D, Dasgupta E, Russell NH, Byrne JL. Alemtuzumab therapy for refractory idiopathic hypereosinophilic syndrome. Br J Haematol 2004;124:558-9.

Available online April 2, 2009.
doi:10.1016/j.jaci.2009.01.069

CD14 and *IL4R* gene polymorphisms modify the effect of day care attendance on serum IgE levels

To the Editor:

The cause of atopy is generally traced to the interplay of genetic and environmental factors.[1] Day care appears to be one of the most frequently investigated environmental factors. Although the results of studies investigating the association between day care attendance and atopy, as assessed by skin prick test responses, were inconsistent, all studies[2-4] measuring serum IgE levels have thus far shown a constant decreasing effect on serum IgE levels.

Among the genes that show a gene-environment interaction for the development of atopy or allergic diseases, the most frequently investigated is the *CD14* gene.[1] However, there is no report that investigates interaction of this gene and day care attendance. CD14 is a pattern-recognition receptor involved in the clearance of bacterial endotoxin and is also known as a receptor of respiratory syncytial virus. We investigated *CD14*−159C/T (rs2569190) and *CD14*−550C/T (rs5744455) polymorphisms in Japanese patients with severe respiratory syncytial virus–induced bronchiolitis and found that *CD14*−550C/T but not *CD14*−159C/T was significantly associated with the condition.[5]

The IL-4 receptor α gene *(IL4R)* is also one of the most frequently investigated genes and has been shown to be associated with atopy and atopic diseases.[6] The Ile50Val polymorphism (rs1805010) of the *IL4R* gene is a functional polymorphism and has been reported to be strongly associated with atopy and atopic asthma in the Japanese population. To date, only one study has reported the interaction of the *IL4R* Ile50Val polymorphism and day care attendance in the first year of life.[7] The result showed a

TABLE I. Characteristics of the subjects

Total no. of participants	473
Age (mo)	
Mean ± SD	111.1 ± 19.9
Range	76-147
Sex ratio (male:female)	1.00:1.01
Day care attendance before age 2 y (%)	14.5
Total IgE (IU/mL), mean ± SD	
Male	254 ± 340
Female	241 ± 469
Prevalence of atopy (%)	
Male	76.9
Female	68.0
Prevalence of allergic disorders (%)	
Asthma	
Male	14.1
Female	6.6
Atopic dermatitis	
Male	11.5
Female	9.7
Allergic rhinitis	
Male	42.1
Female	31.2
Food allergy	
Male	3.0
Female	3.4

significant gene-environment interaction for IFN-γ production at 1 year of age. However, it is not known whether this modified cytokine response affects the chance of having atopy or allergic diseases in the later period of life.

Here we report a relationship between serum total and specific IgE levels in Japanese elementary school children and day care attendance during earlier life. Our results suggest that day care attendance is associated with serum IgE levels, and this effect is modified by *CD14*−550C/T and *IL4R* Ile50Val polymorphisms. This is the first report that suggests an interaction between early-life day care attendance and genetic variations on IgE levels in later life.

Children attending an elementary school located in the central area of Chiba city (population of approximately 930,000) were recruited for this study. We first asked all (n = 843) children for participate in the survey. We then sent a detailed questionnaire to those who had a positive response (n = 582). Children with congenital heart diseases and lung diseases caused by immature birth were excluded. A total of 473 school children aged 6 to 12 years were enrolled. Blood samples were collected from 411 children on 2 separate days (July 3 and 12, 2006) for serum and DNA preparation. A complete set of information on total and 8 specific IgE levels, genotypes, and environmental factors was obtained from 375 children. All parents provided written informed consent. The study protocol was approved by the Ethics Committee of Chiba University Graduate School of Medicine.

The status of allergic diseases was evaluated by using questions based on the International Study of Asthma and Allergies in Childhood. We asked whether the child regularly attends a day care center where time is spent with other children at or before 2 years of age. For parents who responded yes to this question, the age of entry of their child to the day care center was obtained. The questionnaire also included the following items to assess possible confounding factors: number of siblings; number of older siblings; allergic diseases of parents and siblings (family history: scored as positive if parents, siblings, or both had any of 4 allergic diseases [asthma, allergic rhinitis, atopic eczema, and food allergy]); residential area (6 categories), type of house structure (5 categories), and floor type of bedroom (5 categories); yogurt/fermented food consumption; pet ownership; and smoking among family members.

Genotyping of the *CD14*−550C/T polymorphism was performed as described previously,[5] whereas that of the *IL4R* Ile50Val (rs1805010) polymorphism was carried out with the TaqMan allele-specific PCR method.[8] Primer sequences were as shown in this article's Online Repository at www.jacionline.org.

Table I shows the characteristics of the investigated population. The percentage of children who had regularly attended day care before 2 years of age was 14.5%. Atopy was defined as the presence of positive ($\geq$0.35 IU/mL) specific IgE level against at least 1 of the 8 allergens. Although the prevalences of asthma, atopic dermatitis, and food allergy were compatible with those in a recent large study,[9] prevalences of allergic rhinitis and atopy were about 10 to 20 points higher, suggesting that children who had allergic rhinitis were more likely to attend this study.

Table II shows the association between day care attendance and serum IgE levels or atopy after being stratified with the *CD14*−550C/T genotype. Day care significantly decreased total IgE levels ($P = 9.7 \times 10^{-5}$), mite-specific IgE levels ($P = .0016$), and rate of atopy ($P = .00041$) in individuals with the C/T or T/T genotype, whereas the effect of day care was not observed in those with the C/C genotype. Numbers of children with the C/T+T/T genotype and those with the C/C genotype were similar, suggesting that the difference is not likely due to the statistical power for detecting association. Multivariate analyses with confounding factors were performed to evaluate the significance of this gene-environment interaction. The interaction between the *CD14*−550C/T polymorphism and day care was significant for $\log_{10}$(total IgE) ($P = .0046$), mite-specific IgE classes ($P = .00047$), and atopy ($P = .0097$) after adjusting for age, sex, family history, and number of siblings.

Table III shows the association between day care attendance and serum IgE levels or atopy after being stratified with the *IL4R* Val50Ile genotype. The effects of day care on total and some specific IgE levels were significant in Val/Ile heterozygotes but not in Val/Val or Ile/Ile homozygotes. In Val/Ile individuals day care significantly decreased total IgE levels ($P = .0012$), mite-specific ($P = .011$) and cedar pollen–specific ($P = .034$) IgE levels, and rate of atopy ($P = .018$). No such trend was observed in Val/Val or Ile/Ile individuals. The numbers of Val/Val and Val/Ile individuals were similar. It is therefore unlikely that the lack of significant association in Val/Val individuals was due to smaller statistical power for detecting association. When the significance of gene-environment interaction was assessed with the confounding factors, the interaction term between *IL4R* and day care attendance was significant for $\log_{10}$(total IgE) ($P = .019$) and mite-specific ($P = .0025$) and cedar pollen–specific ($P = .040$) IgE classes but not for atopy.

Total IgE levels in 4 genotype groups (group 1: *CD14* C/C, *IL4R* Ile/Ile+Val/Val; group 2: *CD14* C/C, *IL4R* Val/Ile; group 3: *CD14* C/T+T/T, *IL4R* Ile/Ile+Val/Val; and group 4: *CD14* C/T+T/T, *IL4R* Val/Ile) were compared to evaluate the combined effect of 2 polymorphisms on total IgE levels. Fig 1 shows the box

TABLE II. Effects of day care attendance on IgE levels when stratified by *CD14*−550C/T genotype

	C/C				C/T + T/T				
	Day care attendance				Day care attendance				
	No	Yes	Effect size or odds ratio (95% CI)	*P* value	No	Yes	Effect size or odds ratio (95% CI)	*P* value	Gene-environment interaction *P* value*
No. of subjects	169	22			157	28			
Log_{10}(total IgE)									
Mean	1.88	1.98	0.094 (−0.21 to 0.39)¶	0.54†	2.09	1.58	−0.50 (−0.26 to −0.76)¶	**9.7 ×10^{-5}†**	**.0046****
SD	0.77	0.76			0.63	0.51			
Specific IgE (positive‡ rate)									
Mite	0.49	0.59	1.50 (0.61 to 3.69)#	.51§	0.61	0.32	0.30 (0.13 to 0.71)#	**.0016§**	**.00047††**
Cedar pollen	0.45	0.46	1.02 (0.42 to 2.45)#	.92§	0.57	0.32	0.35 (0.15 to 0.83)#	**.032§**	.116††
Atopy (rate)	0.77	0.68	1.60 (0.56 to 4.55)#	.38‖	0.81	0.50	0.24 (0.10 to 0.55)#	**.00041‖**	**.0097‡‡**

Boldface indicates statistically significant values.
*Adjusted for age, sex, number of siblings, and family history.
†Analysis of variance for log_{10}(total IgE [in international units per milliliter]).
‡Class ≥1 (≥0.35 IU/mL).
§Kruskal-Wallis test for IgE value (in international units per milliliter).
‖χ^2 Test of independence.
¶Effect size.
#Odds ratio.
**General liner model.
††Generalized linear model (Poisson distribution, log link function).
‡‡Logistic regression.

TABLE III. Effects of day care attendance on IgE levels when stratified by *IL4R* Val50Ile genotype

	Val/Val				Val/Ile				Ile/Ile				
	Day care attendance				Day care attendance				Day care attendance				
	No	Yes	Effect size or odds ratio (95% CI)	*P* value	No	Yes	Effect size or odds ratio (95% CI)	*P* value	No	Yes	Effect size of odds ratio (95% CI)	*P* value	Gene-environment interaction *P* value*
No. of subjects	125	18			152	27			49	5			
Log_{10}(total IgE)													
Mean	1.94	1.91	−0.058 (−0.38 to 0.27)¶	.72†	1.88	1.55	−0.44 (−0.71 to −0.18)¶	**.0012†**	1.99	2.32	0.33 (−0.31 to 0.97)¶	.12†	**.019****
SD	0.64	0.72			0.57	0.56			0.69	0.52			
Specific IgE (positive‡ rate)													
Mite	0.57	0.56	0.95 (0.35 to 2.57)#	.51§	0.52	0.30	0.39 (0.16 to 0.94)#	**.011§**	0.59	0.80	2.76 (0.29 to 26.5)#	.36§	**.0025††**
Cedar pollen	0.50	0.50	1.01 (0.38 to 2.73)#	.93§	0.51	0.30	0.41 (0.17 to 0.99)#	**.034§**	0.55	0.40	0.54 (0.083 to 3.54)#	.91§	**.040††**
Atopy (rate)	0.74	0.72	0.93 (0.31 to 2.82)#	.90‖	0.74	0.52	0.37 (0.16 to 0.86)#	**.018‖**	0.76	0.80	1.30 (0.13 to 12.8)#	.82‖	.118‡‡

Boldface indicates statistically significant values.
*Adjusted for age, sex, number of siblings, and family history.
†Analysis of variance for log_{10}(total IgE [in international units per milliliter]).
‡Class ≥1 (≥0.35 IU/mL).
§Kruskal-Wallis test for IgE value (in international units per milliliter).
‖χ^2 Test of independence.
¶Effect size.
#Odds ratio.
**General liner model.
††Generalized linear model (Poisson distribution, log link function).
‡‡Logistic regression.

plot of log_{10}(total IgE) in 4 genotype groups. Among children who attended day care compared with group 1, the mean log_{10} (total IgE) values of groups 2, 3, and 4 decreased by 0.41, 0.35, and 0.69, respectively. This magnitude of change suggests that the effects of *CD14* and *IL4R* were additive. The children in group 4 showed significantly (P = .0046) lower total IgE levels than those in group 1. On the other hand, among children who did not attend day care, the log_{10}(total IgE) levels of children in groups 3 (P = .031) and 4 (P = .036) were significantly higher than those of children in group 1. The *CD14* C/T and T/T genotypes appeared to show the opposite effect on the serum total IgE level in children who did not attend day care compared

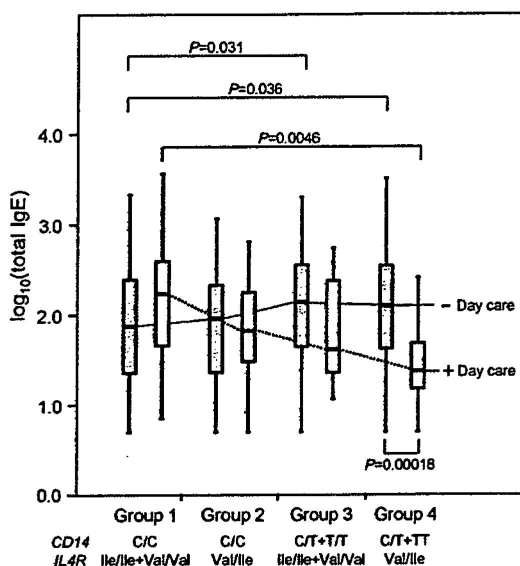

FIG 1. Total IgE levels in 4 groups of children classified based on a combination of *IL4R* and *CD14* genotypes. Box plot of $\log_{10}$(total IgE) values is shown for children who attended day care *(+ Day care)* and for those who did not *(− Day care)*. Results are presented as medians and interquartile ranges. Only significant *P* values (<.05) are shown.

with those who did attend day care. When we examined the effect of day care in each genotype group, the effect was not sufficiently large to show a significant change in IgE level in groups 2 and 3, in which individuals had only 1 IgE level–decreasing genotype. However, in group 4, in which individuals had 2 IgE level–decreasing genotypes, the effect was sufficiently large to show a significant difference ($P = .00018$). Significance of interaction between the *CD14* and *IL4R* genotypes was also evaluated by using general linear models in which age, sex, family history, number of siblings, and day care were included as variables. The interaction term of the 2 genes was not significant, suggesting an independent effect of the *CD14* and *IL4R* genes.

The interaction of the *CD14* gene with day care attendance suggests that the mechanism of the effect of day care involves at least in part a response to infection, environmental endotoxin exposure, or both. The interaction of the *IL4R* gene with day care attendance suggests that the mechanism also involves those related to T_H2 cell proliferation and IgE production. These results suggest that the complex nature of mechanisms underlies the effect of day care attendance on serum IgE levels.

Environmental factors investigated in the present study were determined based on a questionnaire on past day care attendance, and therefore recall bias can be a potential problem. The number of subjects investigated in this study was not so large and might be the acceptable minimum for investigating gene-environment interactions. The subjects evaluated were children who attended a single school and lived in a medium-populated city, thus representing those living in rather small regional environments in Japan. Nevertheless, these characteristics of the present sample might have contributed to minimizing the variances of background and outcome parameters and might have resulted in the positive findings obtained from a relatively small number of subjects. It is necessary to perform a cohort study to follow children with or without day care attendance until they reach school age to validate the current observations.

Yoichi Suzuki, MD, PhD[a]
Satoshi Hattori, MD[a]
Yoichi Mashimo, PhD[a]
Makiko Funamizu[a]
Yoichi Kohno, MD, PhD[b]
Yoshitaka Okamoto, MD, PhD[c]
Akira Hata, MD, PhD[a]
Naoki Shimojo, MD, PhD[b]

From the Departments of [a]Public Health, [b]Pediatrics, and [c]Otolaryngology, Graduate School of Medicine, Chiba University, Chiba, Japan. E-mail: ysuzuki@faculty.chiba-u.jp.

Supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and grants from the Ministry of Health, Labor, and Welfare, Japan.
Disclosure of potential conflict of interest: The authors have declared that they have no conflict of interest.

REFERENCES

1. Yang IA, Savarimuthu S, Kim ST, Holloway JW, Bell SC, Fong KM. Gene-environmental interaction in asthma. Curr Opin Allergy Clin Immunol 2007;7:75-82.
2. Celedon JC, Litonjua AA, Ryan L, Weiss ST, Gold DR. Day care attendance, respiratory tract illnesses, wheezing, asthma, and total serum IgE level in early childhood. Arch Pediatr Adolesc Med 2002;156:241-5.
3. Kramer U, Heinrich J, Wjst M, Wichmann HE. Age of entry to day nursery and allergy in later childhood. Lancet 1999;353:450-4.
4. Rothers J, Stern DA, Spangenberg A, Lohman IC, Halonen M, Wright AL. Influence of early day-care exposure on total IgE levels through age 3 years. J Allergy Clin Immunol 2007;120:1201-7.
5. Inoue Y, Shimojo N, Suzuki Y, Campos Alberto EJ, Yamaide A, Suzuki S, et al. CD14-550 C/T, which is related to the serum level of soluble CD14, is associated with the development of respiratory syncytial virus bronchiolitis in the Japanese population. J Infect Dis 2007;195:1618-24.
6. Ober C, Hoffjan S. Asthma genetics 2006: the long and winding road to gene discovery. Genes Immun 2006;7:95-100.
7. Hoffjan S, Nicolae D, Ostrovnaya I, Roberg K, Evans M, Mirel DB, et al. Gene-environment interaction effects on the development of immune responses in the 1st year of life. Am J Hum Genet 2005;76:696-704.
8. Fujii K, Matsubara Y, Akanuma J, Takahashi K, Kure S, Suzuki Y, et al. Mutation detection by TaqMan-allele specific amplification: application to molecular diagnosis of glycogen storage disease type Ia and medium-chain acyl-CoA dehydrogenase deficiency. Hum Mutat 2000;15:189-96.
9. Nisima S, Chisaka H, Fujiwara T. Surveys on the prevalence of pediatric bronchial asthma in Japan: a comparison between the 1982, 1992, and 2002 surveys conducted in the same region using the same methodology. Allergol Int 2009;58:37-53.

Available online May 18, 2009.
doi:10.1016/j.jaci.2009.03.035

Mucormycosis in chronic granulomatous disease: Association with iatrogenic immunosuppression

To the Editor:

Chronic granulomatous disease (CGD) results from mutations in either X-linked ($gp91^{phox}$) or autosomal ($p47^{phox}$, $p67^{phox}$, and $p22^{phox}$) genes encoding the phagocyte nicotinamide adenine dinucleotide phosphate (NADPH) oxidase. Impaired generation of reactive oxygen species predisposes to recurrent life-threatening bacterial and fungal infections. Septated hyaline molds (particularly *Aspergillus* species) are the primary fungal pathogens in CGD. Fungi of the order Mucorales (pauciseptated molds) are environmentally ubiquitous and cause mucormycosis in select immunocompromised patient populations, such as those with diabetic ketoacidosis and hematologic malignancy and recipients of transplants or deferoxamine. We investigated the prevalence of mucormycosis in patients with CGD.

PRIMERS FOR *IL4R* GENOTYPING (5′ TO 3′)

TaqMan probe (FAM-TACAGGTGACCAGCCTAACCCAGC CCCTGT-TAMRA); common primer (TGGAGGCATGTCCCG GACAC); Ile (A) allele primer (CGCCTCCGTTGTTCTCAG GGGT); and Val (G) allele primer (CGCCTCCGTTGTTCTC AGGGGC).

19

Replication of genetic association studies in asthma and related phenotypes

Undarmaa S, Mashimo Y, Hattori S, Shimojo N, Fujita K,
Miyatake A, Doi, S,Kohno Y, Okamoto Y, Hirota T, Tamari M,
Hata A, Suzuki Y.

J Hum Genet (e-pub) 2010

Journal of Human Genetics (2010), 1–8
© 2010 The Japan Society of Human Genetics All rights reserved 1434-5161/10 $32.00
www.nature.com/jhg

ORIGINAL ARTICLE

Replication of genetic association studies in asthma and related phenotypes

Siizkhuu Undarmaa[1], Yoichi Mashimo[1], Satoshi Hattori[1], Naoki Shimojo[2], Kimie Fujita[3], Akihiko Miyatake[4], Satoru Doi[5], Yoichi Kohno[2], Yoshitaka Okamoto[6], Tomomitsu Hirota[7], Mayumi Tamari[7], Akira Hata[1] and Yoichi Suzuki[1]

In asthma genetics, the association of highly replicated susceptibility genes lacks consistency across populations. To identify genuine associations, we investigated the reproducibility of the 23 most promising asthma and asthma-related candidate genes in a moderately sized sample from the Japanese population. We compared the frequency of 33 polymorphisms in unrelated cases and controls and tested for their association with asthma, atopy and serum total IgE levels using allele frequency, codominant, dominant and recessive genotype models. On the basis of the consistency of our findings with previous meta-analyses and large replication studies, *IL13*, *TNF*, *ADAM33*, *IL4RA* and *TBXA2R* might represent common major asthma and asthma-related trait genes. Individual gene assessment was extended to the interactions between two polymorphisms using our original method. Interactions between *TBXA2R* and *ADAM33*, and *IL4RA* and *C3* were suggested to increase the risk for childhood and all asthma (adult and childhood asthma combined). The confirmation of previously reported associations between gene polymorphisms and phenotypes was problematic when as few as several hundred samples per group were used. Stratification of the subjects by environmental factors or other confounding factors may be necessary to improve the sensitivity and reliability of association results.

Journal of Human Genetics advance online publication, 16 April 2010; doi:10.1038/jhg.2010.32

Keywords: association; asthma; atopy; polymorphism; replication

INTRODUCTION

Asthma is a heritable trait[1] and investigations to determine the genetic components underlying asthma using linkage mapping and the candidate gene approach have been carried out. By 2006, more than 100 genes were associated with asthma and asthma-related phenotypes;[2] 25 of these genes have been replicated in more than six independent association studies. In 2008, this list was complemented with an additional three genes, *FLG*, *NAT2* and *CCL15*.[3] However, no single polymorphic marker or gene locus has been unanimously labeled as a strong and independent genetic determinant of asthma, and the results for the highly replicated genes have been inconsistent across the tested populations.[3]

To identify true associations, it is of critical importance to comprehensively replicate the initial finding.[4] To this aim, we investigated whether the 23 most replicated genes for asthma and asthma-related phenotypes were positively associated with extrinsic childhood and adult asthma, atopy and total serum IgE levels in a moderately sized sample drawn from the Japanese population.

We also tested eight genes that were significantly associated with asthma in our subjects: *IL13*,[5] *TBXA2R*,[6] *GSTP1*,[7] *ADAM33*,[8] *MMP9*,[9] *IL12B*,[10] *C3*[11] and *SOCS1*.[12] The re-evaluation of these associations is conditioned by the limitations of the original reports in which childhood asthma included subjects with nonatopic asthma and those who were <4 years of age. The adult asthma cases also included nonatopic asthma in some of these reports. Moreover, the comparison of the childhood asthma group was with an adults-only control group. In this study, we redefined the atopic asthma patients, introduced age-matched child controls and re-evaluated the association of these genes with the asthma phenotypes, atopy and total serum IgE levels.

Further, we extended the assessment of individual genes to identify potential interactions between the genes, as increasing knowledge about biological pathways and gene networks implies that gene–gene interactions are important and must be taken into account when estimating the genetic risk of a disease.[13]

MATERIALS AND METHODS

Study population

The asthma population was restricted to extrinsic asthma patients with subsequent distinction between childhood and adult asthma by cutoff age of below or above 16 years old regardless of the age of the disease onset. We recruited 325 subjects with childhood atopic asthma, 367 adults with atopic

[1]Department of Public Health, Graduate School of Medicine, Chiba University, Chiba, Japan; [2]Department of Pediatrics, Graduate School of Medicine, Chiba University, Chiba, Japan; [3]School of Human Nursing, The University of Shiga Prefecture, Shiga, Japan; [4]Miyatake Asthma Clinic, Osaka, Japan; [5]Department of Pediatric Allergy, Osaka Prefectural Medical Center for Respiratory and Allergic Diseases, Osaka, Japan; [6]Department of Otolaryngology, Graduate School of Medicine, Chiba University, Chiba, Japan and [7]Laboratory for Respiratory Diseases, Center for Genomic Medicine, RIKEN, Kanagawa, Japan
Correspondence: Dr Y Suzuki, Department of Public Health, Chiba University Graduate School of Medicine, 1-8-1 Inohana, Chuo-ku, Chiba 260-8670, Japan.
E-mail: ysuzuki@faculty.chiba-u.jp
Received 30 November 2009; revised 15 March 2010; accepted 24 March 2010

asthma and 646 adult controls from Osaka City, Japan. The details of these subjects have been described elsewhere.[10,12,14] For childhood controls and the investigation of total and specific IgE levels, we recruited children attending an elementary school in Chiba City, Japan. The clinical characteristics of this population as well as inclusion and exclusion criteria have been described previously.[15] In brief, after the exclusion of questionnaire-assessed asthma and/or atopic dermatitis subjects and those with congenital heart diseases and lung diseases caused by premature birth, 336 children having a complete set of information on the total and eight specific IgE levels, genotypes and environmental factors were assigned to the child control group. The mean ages (range) of the four groups were as follows: childhood asthma, 9.9 (4–15); adult asthma, 45 (16–83); child controls, 9.3 (6–12); and adult control, 43.7 (20–75) years. Written informed consent was received from all participants and the study was approved by the ethics committees of Chiba University Graduate School of Medicine and RIKEN.

Gene and polymorphism selection

The list of candidate gene polymorphisms included in this study, their location within the gene and corresponding rs numbers are given in Supplementary Table S1. In this table, we also included the allele frequency in child and adult control populations.

The most replicated genes were selected based on the list created by Ober and Hoffjan.[2] From the 25 cited genes, we intentionally excluded *HLA-DRB1* and *HLA-DBP1* from our analysis due to the high number of variants linked to asthma and asthma-related phenotypes, the genotyping of which would surpass our capacity. The remaining 23 genes were tested for association based on the most positively reported polymorphisms and are represented by Group 1 (>10) and Group 2 (6–10), depending on the number of their replications (Supplementary Table S1). We included four genes (*MMP9, IL12B, C3* and *SOCS1*), which were not in the original 23 genes, but were found to be associated with asthma in our previous studies (Supplementary Table S1, Group 3). These were tested for association with the same positive polymorphisms as in the initial reports.

DNA extraction and genotyping

Genomic DNA was prepared from peripheral blood samples using the standard protocols. Whole genome amplification was carried out using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's instructions. Genotyping was carried out by means of allele-specific amplification, single-nucleotide primer extension reaction or fragment analysis of the PCR product. The genotyping methods used for each polymorphism are given in Supplementary Table S1 and the primer sequences are shown in Supplementary Tables S2, S3a, S3b and S4. The results obtained by these molecular assays were analyzed on ABI PRISM 3100 Genetic Analyzer and ABI PRISM 7000 Sequence Detector Systems (Applied Biosystems, Foster City, CA, USA), or by using Chromo4 Real-Times System (Bio-Rad Laboratories, Hercules, CA, USA). The detailed genotypic protocols are described in Supplementary Methods.

Statistical analysis

A case–control study design was used to analyze the associations between gene variants and the dichotomous phenotypes. We calculated allele frequencies and tested their agreement with Hardy–Weinberg equilibrium (HWE) using a χ^2 goodness-of-fit test. We compared differences in the allele frequencies and genotype distribution of each polymorphism between the case and control subjects by using a 2×2 (allele) or 2×3 contingency χ^2-test (dominant, codominant and recessive genotype models) with 1 or 2 degree of freedom. Total serum IgE values were logarithm transformed to approximate a normal distribution and analyzed as a quantitative trait using two different tests: linear regression analysis with age and sex as covariates; and analysis of variance (ANOVA). All statistical analyses were carried out with SPSS Statistics 17.0 (SPSS Japan Corporation, Tokyo, Japan). The statistical power of this study was calculated using SamplePower 2.0 (SPSS Corporation). *P*-values <0.05 were judged to be significant; as there was an *a priori* hypothesis with all tested polymorphisms, we did not adjust the significance levels for multiple testing.

Interactions between two polymorphisms were screened using a program written in Excel VBA (Microsoft Japan, Tokyo, Japan). This program identified every possible pairing of polymorphisms, calculated the χ^2 and *P*-values of contingency tables between the genotype of one polymorphism and the disease after stratification by the genotype of the other polymorphism. The program then calculated the χ^2 and *P*-values for the goodness-of-fit statistics of the distribution pattern of genotype vs disease table with stratification against no stratification. Single-nucleotide polymorphisms (SNPs) in the same gene were not tested for this interaction because most of them showed linkage disequilibrium.

RESULTS

Statistical power of this study

We estimated the frequency difference between the case and control groups, and odds ratio (OR) for our samples with different genotype/allele frequencies according to sample size. We set our detection power at 80% with an α level of 0.05.

Supplementary Table S5a shows the results of this calculation for the asthma cases, when estimation was carried out for given genotype/allele frequencies among their respective control peers. Adult asthma group had slightly better power than the childhood asthma group because of almost twice bigger number of adult controls compared with the child controls. In the case of adult asthma, enough (>80%) power could be expected for polymorphisms with an OR of 1.4 when the minor genotype frequency was from 30–60%. In the allele frequency comparison, there was slightly more detection power than for the genotype comparison, as the number of alleles was twice that of the genotype. When we estimated the genotype/allele frequencies in child asthmatics for given adult control genotype/allele frequencies (Supplementary Table S5b), as expected, some increase in power could be observed with ORs getting closer to those of childhood asthma vs child control estimation. In case of calculation based on the comparison between combined asthma subjects and combined controls (Supplementary Table S5b), ORs of 1.3 and 1.2 could be detected for minor genotype and allele frequencies of 30–60%, respectively.

The estimated power to detect an association between atopy and genetic polymorphisms (Supplementary Table S6) was comparatively inferior to that of asthma with an OR of 1.9.

Association of the 23 genes with childhood and adult atopic asthma

The genotype frequencies of all gene variants, except for *GSTM1* ins/del, were in Hardy–Weinberg equilibrium (goodness-of-fit χ^2-test, $P>0.01$) in adult and child controls. We could not test the *GSTM1* ins/del variant for Hardy–Weinberg equilibrium because the genotyping method did not distinguish between ins/del heterozygotes and homozygotes.

The summary of results for basic comparison between cases and controls is presented in Table 1, in which only polymorphisms with *P*-value <0.05 in at least one genetic model test are shown. When the frequency of polymorphisms in childhood asthmatics were compared with those of asthma free children, the most strong candidate variant was *IL13* −1112C>T, which was significantly associated in all but dominant genotype model; *LTC4S* was positively associated in the allele frequency and dominant model tests and *CCL5* −28C>G resulted in significant difference only in the dominant genotype model. In the adult samples, there were significant differences noted between the controls and the asthma patients for *TNF, ADAM33* and *NOS1*. *TNF* showed positive results in the allele frequency, codominant and dominant models. Association of *ADAM33* Met764Thr variant with adult asthma was significant for allele frequency and dominant genotype models. There was strong association of both *NOS1* variants with the disease in all tests except for GT repeat intron

Table 1 Association of genotyped polymorphisms with the asthma phenotypes (basic comparisons)

		Case–control study P*-value*											
		Childhood asthma vs child control				*Adult asthma vs adult control*				*All asthma vs all controls*			
Gene	*Polymorphism*	*Allele*	*Codom.*	*Dom.*	*Rec.*	*Allele*	*Codom.*	*Dom.*	*Rec.*	*Allele*	*Codom.*	*Dom.*	*Rec.*
Group 1													
IL13	Arg110Gln	×	×	×	×	×	×	×	×	×	×	×	×
	-1112C>T	0.014	0.011	×	0.003	×	×	×	×	×	×	×	×
TNF	-1037C>T	×	×	×	×	0.001	0.003	0.001	×	0.003	0.014	0.005	×
ADAM33	Met764Thr	×	×	×	×	0.008	×	0.005	×	0.029	×	0.035	×
	13236C>T	×	×	×	×	×	×	×	×	0.038	×	×	×
Group 2													
LTC4S	–444A>C	0.023	×	0.026	×	×	×	×	×	×	×	×	×
NOS1	GT repeat intron 2 (187 allele)	×	×	×	×	0.0003	0.002	0.001	0.011	0.026	×	×	×
	GT repeat intron 2 (183 allele)	×	×	×	×	0.007	0.022	0.007	×	0.045	×	×	×
CCL5	–403A>G	×	×	×	×	×	×	×	×	×	×	×	×
	–28C>G	×	×	0.048	×	×	×	×	×	×	×	×	×
Group 3													×
MMP9	2127G>T	×	×	×	×	×	×	×	×	×	×	×	×
	5546G>A	×	×	×	×	×	×	×	×	×	×	×	×
IL12B	–6415CTCTAA>GC	0.010	0.006	×	0.001	×	×	×	×	0.012	0.009	×	0.002
	1146 C> A	0.019	0.010	×	0.003	×	×	×	×	0.003	0.003	×	0.001
C3	Block 2 (haplotype 6)[a]	0.017	×	×	×	×	×	×	×	0.015	0.022	0.010	×
	Block 4 (haplotype 1)[a]	×	×	×	×	0.001	0.002	0.024	0.002	0.001	0.005	0.003	0.019
SOCS1	–1478CA>del	×	×	×	×	0.0005	0.002	0.002	0.015	0.002	0.005	0.006	0.017

Abbreviations: Allele, χ^2-test of allele frequency; codom., 2×3 codominant model genotype χ^2-test; dom., 2×2 dominant model genotype χ^2-test; rec., 2×2 recessive model genotype χ^2-test.
[a]For haplotype description please refer Inoue *et al.*[11]
Polymorphisms with a *P*-value ⩾0.05 in all association tests of the four genetic models are not shown.
×$P \geqslant 0.05$.

(183 allele) in the recessive genotype model. In the association test between the polymorphisms and all asthma, that is child and adult asthma combined, significant associations were observed for the same genes as in adult asthma. *TNF* and *ADAM33* Met764Thr remained positively associated with asthma in the allele frequency, codominant and dominant models alike in the single adult asthma group comparison. The association of *NOS1* variants became weaker with significance only in allele frequency model. Additional significant association was observed only for *ADAM33* 13236C>T in the allele frequency model.

Conditioned by the inherent characteristics of our case and control samples, we extended our analysis further (Table 2). In our child control group, we recognize the potential presence of asthma susceptibility gene variants carriers who might become asthmatics later in life, and consequently, to become a source of bias. Addressing this issue, we tested childhood asthmatics also with adult controls, considering the later ones as more reliable asthma 'free' subjects. Paradoxically, the previous significant associations found in the comparison for child asthma vs child control disappeared; instead, *TNF*, *IL4RA*, *ADAM33*, *GSTM1*, *AAA1* and *NOS1* showed significant results. The substantial difference in results between childhood asthma vs child control and childhood asthma vs adult control tests might be due to some confounding effects of hidden genetic and environmental heterogeneities between child and adult control groups.

Furthermore, as the natural history of asthma in some adults starts during childhood, we characterized the association results after stratifying adult asthma cases by age at the onset (Table 2). Using the same cutoff age (<16 years) as for discerning childhood asthma from adult asthma, we obtained 118 adults (32% of adult asthma), who have reported physician diagnosed asthma during their childhood, and 249 adults with newly onset bronchial asthma. To elucidate the genes associated with differential onset of asthma in children, we merged the child-onset adult subgroup with childhood asthma group (all L16 adult atopic asthma (BA)) and tested with either child or adult controls by applying the same strategy as used in the non-stratified adult asthma. In the association test with child control, besides the previously detected gene variants in the comparison of child asthma vs child control, *TNF* showed at least one significant association in four association tests, whereas the significance of *CCL5* disappeared. When the combined child-onset asthma group was compared with adult controls, the observed associations in child asthma vs adult control comparison could be confirmed for only *IL4RA* Ile50Val and *ADAM33*:13236C>T polymorphisms. We next examined the new-onset adult asthma, in which *TNF* and *ADAM33* Met764Thr have retained their significant associations found with the onset nonstratified adult asthma. In addition, *IL13*, *IL4RA* and *GSTM1* have emerged as genes associated with asthma establishment during adulthood. Interestingly, the two *ADAM33* polymorphisms were associated with different asthma phenotypes when comparisons were carried out with adult control. *ADAM33* Met764Thr was significantly associated with adult asthma, whereas *ADAM33* 13236C>T variant was significantly more frequent in childhood asthma and all child-onset asthma groups.

We also compared the distribution of polymorphisms between the two control groups; it showed a relative homogeneity with significant

Table 2 Summary of the association results between polymorphisms and asthma when case and control groups were compared in various combinations

		Two group comparison							
Gene	Polymorphism	Child BA vs child CO	Child BA vs adult CO	Adult BA vs adult CO	All L16 BA vs child CO	All L16 BA vs adult CO	M16 adult BA vs adult CO	Child CO vs adult CO	Child BA vs adult BA
Group 1									
IL13	Arg110Gln	×	×	×	×	×	×	×	×
	–1112C>T	°	×	×	°	×	°	°	×
TNF	–1037C>T	×	°	°	°	×	°	×	×
IL4RA	Ile50Val	×	°	×	×	°	°	×	×
	Gln551Arg	×	×	×	×	×	×	×	°
ADAM33	Met764Thr	×	×	°	×	×	°	×	°
	13236C>T	×	°	×	×	°	×	×	°
Group 2									
GSTM1	Ins/del	×	°	×	×	×	°	×	°
IL10	571C>A	×	×	×	×	×	×	×	°
LTC4S	–444A>C	°	×	×	°	×	×	°	×
AAA1	522363G>C	×	°	×	×	×	×	×	°
NOS1	GT repeat (187 allele)	×	°	°	×	×	×	×	°
	GT repeat (183 allele)	×	×	°	×	×	×	×	×
CCL5	–28C>G	°	×	×	×	×	×	×	×
Group 3									
MMP9	2127G>T	×	°	×	×	×	°	°	×
	5546G>A	×	°	×	×	°	×	°	×
IL12B	–6415ins/del	°	°	×	°	°	×	×	°
	1146 C> A	°	°	×	°	°	×	×	°
C3	Block 2 (haplotype 6)[a]	°	×	×	°	°	×	×	×
	Block 4 (haplotype 1)[a]	×	°	°	×	°	×	×	×
SOCS1	–1478CA>del	×	×	°	×	°	×	×	×

Abbreviations: Child BA, childhood atopic asthma; child CO, child controls; adult BA, adult atopic asthma; adult CO, adult controls; all L16 adult BA, childhood atopic asthma and adult atopic asthma at onset age <16 years combined; M16 adult BA, adult atopic asthma at onset age ⩾16.
[a]For haplotype description please refer Inoue *et al.*[11]
Polymorphisms with a *P*-value ⩾0.05 in all association tests of the four genetic models are not shown.
°*P*-value <0.05 in at least one association test of the four genetic models.
×None of the *P*-value in the four genetic models was <0.05.

difference only for *IL13* and *LTC4S*. It indicates that there would be little bias created using either sample as a control group for any of the case groups and justifies our combining of the two control groups to form a single one in the comparison for all asthma vs all control. In the comparison test for childhood asthma vs adult asthma, *IL4RA*, *ADAM33*, *GSTM1*, *IL10*, *AAA1* and *NOS1* genes were found significantly different, which is coherent with the case–control comparison results.

Association of the 23 genes with atopy

ADAM33 13236C>T, *NOS1* GT repeat 183 bp allele and *TBXA2R* 795T>C were significantly associated with atopy in the recessive or dominant models (Supplementary Table S7). *NOS1* and *TBXA2R* also showed significant association in the allele frequency test, whereas *ADAM33* did not.

Association of the 23 genes with total serum IgE

Supplementary Table S8 summarizes the relationship between the mean of $\log_{10}$-transformed total serum IgE level and gene variants. There was a significant association between *GSTP1* and total IgE level as determined by ANOVA that remained significant after adjusting for age and gender as covariates in linear regression analysis.

Re-evaluation of the associations previously discovered in our subjects

We confirmed the association of the *IL13*, *ADAM33*, *MMP9*, *IL12B*, *C3* and *SOCS1* polymorphisms with atopic asthma (Tables 1 and 2). The association of *TBXA2R* with asthma was not replicated; instead, we identified an association of *TBXA2R* with atopy (Supplementary Table S7). Comparatively to our previous studies, in this study we also investigated the polymorphisms for association with serum IgE level. Significant results were observed for *GSTP1* as described above and for *MMP9* 5546G>A (Supplementary Table S8).

Screening of gene–gene interactions

From the 40 genetic polymorphisms, polymorphism pairs with a *P*-value for association with the disease of <0.05 and a *P*-value for interaction <0.01 are listed in Table 3. A pair of *LTA* and *TNF* SNPs showed a very small *P*-value; however, these two genes are physically close and these SNPs are known to be in linkage disequilibrium. Thus, this combination was omitted from the table. Interactions between *TBXA2R* and *ADAM33* and between *IL4RA* and *C3* were suggested for childhood asthma and all asthma. These interactions were our most robust results.

Table 3 Screening of gene–gene interactions

SNP 1	SNP 1 genotype	SNP 2	Effect of SNP 1 on association between SNP 2 and disease[a]			Association between SNP 2 and disease after stratification by SNP 1		
			χ^2	d.f.	P-value	χ^2	d.f.	P-value
Childhood asthma								
LTA -735G>A	A/A	*IL12B* 1146C>A	7.12	1	0.0076	18.1	2	0.00012
IL4RA Ile50Val	Ile/Ile	*C3* block2 (haplotype 6)[b]	11.63	1	0.00064	13.42	1	0.00025
TBXA2R 924C>T	C/T	*ADAM33* 313236C>T	9.51	2	0.0086	14.02	2	0.0009
SPINK5 Lys420Glu	G/G	*ADAM33* 313236C>T	12.93	2	0.0016	12.24	2	0.0022
AAA1 522363G>C	C/C	*IL4* 33C>T	10.98	2	0.0041	8.88	2	0.012
AAA1 522363G>C	C/C	*IL4* -590T>C	9.31	2	0.0095	7.67	2	0.022
CCL5 -28C>G	G/G	*IL13* Arg110Gln	7.2	1	0.0073	4.69	1	0.03
CD14 -550C>T	T/T	*IL10* -571C>A	7.1	1	0.0077	4.11	1	0.043
Adult asthma								
SOCS1 -1478CA>del	CA/del	*CTLA4* -318C>T	7.38	1	0.0066	6.36	1	0.012
TNF -1037C>T	T/T	*CCL5* -403A>G	8.51	1	0.0035	5.25	1	0.022
LTA -753G>A	A/A	*SPINK5* Lys420Glu	8.65	1	0.0033	6.95	2	0.031
CD14 -550C>T	T/T	*IL4* -590T>C	11.01	2	0.0041	6.45	2	0.04
All asthma								
TBXA2R 924C>T	C/T	*ADAM33* 313236C>T	12.76	2	0.0017	11.83	2	0.0027
IL4RA Ile50Val	Ile/Ile	*C3* block2 (haplotype 6)[b]	7.29	1	0.0069	6.62	1	0.01
ADAM33 313236C>T	C/T	*TBXA2R* 924C>T	8.05	1	0.0046	8.99	2	0.011
TNF -1037C>T	T/T	*CD14* -159C>T	7.04	1	0.008	7.34	2	0.026
STAT6 GT repeat	172/172	*CC16* 38A>G	11.01	2	0.0041	6.9	2	0.036

Abbreviations: d.f., degree of freedom; SNP, single-nucleotide polymorphism.
[a]Polymorphism combination that demonstrates the interactions with $P<0.01$ and an association with $P<0.05$ are shown.
[b]For haplotype description please refer Inoue *et al.*[11]

DISCUSSION

We conducted our study in three phases. We first, under the assumption of the common disease–common variant hypothesis, investigated whether the 23 most promising asthma/atopy candidate genes retained their association in a Japanese population set. We considered the gene as the unit of our replication, and the gene was judged as positively replicated if demonstrated a statistically significant association with one or more phenotypes (atopic asthma, atopy and total serum IgE level) in at least one of four genetic models. In our samples, among the highly replicated genes (>10 positive associations), *IL13*, *TNF*, *IL4RA* and *ADAM33* maintained their reputation as robust asthma and asthma-related candidate genes. From the genes with a lower replication rate (6–10), we confirmed the associations of *GSTM1*, *GSTP1*, *LTC4S*, *AAA1*, *NOS1*, *CCL5* and *TBXA2R*. In the second phase of the study, we screened our initial significant asthma associations to *IL13*, *TBXA2R*, *GSTP1*, *ADAM33*, *MMP9*, *IL12B*, *C3* and *SOCS1*. These associations were replicated for all genes, except for *TBXA2R* and *GSTP1*, which were associated with atopy and total serum IgE level, respectively. In the final phase, we explored the potential multigenic effect of all 27 candidate genes (the three groups of genes combined) in the expression of asthma phenotypes based on a pairwise method.

If we summarize our findings from the replication study, interesting patterns of associations could be observed. Except for *IL13* and *IL4RA*, we found no association for the prominent genes implicated in innate immunity and immunoregulation with asthma or asthma-related phenotypes, that is, *CD14*, *IL10* and *TGFB1* as well as the Th2 cytokines and their receptors represented by *IL4*, *STAT6* and *MS4A2*. Instead, the genes secreted from airway epithelial cells (for example, *CCL5* and *AAA1*), and genes known to affect lung function, mediate inflammatory conditions and participate in airway remodeling (for example, *TNF*, *ADAM33*, *GSTM1*, *GSTP1*, *LTC4S*, *NOS1* and *TBXA2R*) demonstrated statistically significant associations. This observation may indicate the relatively higher predisposing effect that these two groups of genes exert on the development of asthma-related phenotypes in the Japanese population. Our postulation is supported by the hypothesis that not the dysregulated immune response, but the inherently abnormal respiratory epithelium of asthmatics and the reactivation of the epithelial–mesenchymal trophic unit leading to pathological airway wall remodeling has a major role in the disease.[16,17] Another interesting finding was the distinct partition of genes between adult and child asthma. The association of *LTC4S*, *AAA1* and *CCL5* specifically in the child samples might reflect the differing etiopathogenetic background of childhood asthma.[18,19] For example, *CCL5* is a key chemokine recruiting Th1 and Th2 proinflammatory cells, and its expression in epithelial cells is induced by the respiratory syncytial virus (RSV).[20,21] This is in line with the evidences that the epithelial barrier in young asthmatics is inherently abnormal[22] and that RSV bronchiolitis is a more important risk factor for the development of asthma and atopy up to the age of 7 years than heredity or environmental factors.[23] Moreover, the candidate genes showing significant association with both the phenotypes, when the case groups were compared with adult controls separately (childhood asthma vs adult control and adult asthma vs adult control), were the same as those observed in the combined analysis (all asthma vs all controls). Thus, *TNF*, *ADAM33* and *NOS1* might represent the common susceptibility gene for adult and childhood asthma. When we categorized asthma cases by the age at onset, the number of shared

significant genes increased (*IL13*, *TNF*, *IL4RA* and *ADAM33*), making in the overall no substantial differences in the genetic determinants between child-onset and adult-onset asthmas. The exception was for *LTC4S* and *GSTM1*, suggesting the former gene to be related with the development of asthma during childhood and the later one during adulthood. However, as *GSTM1* was significantly associated with childhood asthma in the comparison of childhood asthma vs adult control, it is difficult to conclude whether the polymorphism in this gene affected the susceptibility to new-onset adult asthma.

Our replication results were in agreement with several large-scale studies. A recent review of the literature revealed that five asthma candidate genes, *ADAM33*, *TNF*, *TBXA2R*, *CD14* and *LTC4S*, were the focus of several meta-analyses in which *ADAM33* and *TNF* had a modest association with asthma.[24] The first genome-wide replication study of 39 asthma candidate genes generated *IL4RA* results that were consistent with our observations.[25] In the most comprehensive replication study carried out to date, the reproducibility of 93 genes previously associated with asthma and/or asthma intermediate traits was tested.[26] *IL13* was associated with asthma, and *TBXA2R* was associated with atopy, as we also observed in this study. Our replication rate of 48% (11 genes out of 23; OR 1.15–1.62, if the outlier OR of 3.01 (95% confidence interval, 1.40–6.51) for *IL13* –1112C>T is excluded) was higher than that reported in previous large association studies (for OR see Table 4); the study of Daley *et al.*[26] (unrelated case–control sample of N=5565) and a genome-wide screen of 422 nuclear families using SNP arrays had low replication rates of 13% (12 out of 93 tested genes, OR <1.4) and 15.4% (6 out of 39 at SNP-level replication, OR 1.4–1.7), respectively. This better replication rate might be attributed to our sample size, as it is well documented that smaller studies have a tendency to have more favorable outcomes than larger ones.[27] Daley *et al.* concluded that many published associations for asthma and atopy may be false-positive results. Whereas Rogers *et al.*[25] suggested that the poor coverage of genome-wide association study genotyping platforms and lack of statistical power due to insufficient sample size were the main reasons for their low replication. We are more inclined to suspect the 'contextual' bias explaining our failure to replicate all candidate genes. By that we mean the confounding effect of the whole complex network of gene–gene and gene–environmental interactions. This can be seen from the controversy in the findings between this current study and our previous one. In this study, *CD14* –550C>T and *IL4RA* Ile50Val were not associated with total serum IgE level. Whereas, in our recent association study carried out on the same school children, these two gene variants had a modifying effect on the levels of total IgE later in life depending on the children's attendance of day care before 2 years of age.[15] This association could be detected because the day care attendance was taken in consideration as an environmental factor and the effect of a gene was investigated simultaneously with the effect of the other one.

In the gene–gene interaction analysis, we identified some statistical interactions that asserted the weak associations found in the individual gene assessment. Among them, significant interaction between *C3* and *IL4RA* and between *ADAM33* and *TBXA2R* were observed for both childhood and all asthma groups. Although straightforward functional evidences of such paired interactions are lacking, some plausibility can be inferred. C3 or complement component 3 is an important part of the innate immunity recognizing exogenous and endogenous molecular patterns. Some functions of its C3a subtype indicate a possible role for the complement system in asthma pathogenesis.[28] In allergen-sensitized mouse model of pulmonary allergy deficient in C3 or in its receptor *C3aR*, Drouin *et al.*[29,30] have observed that in the mutant mice the characteristic manifestations of asthma were significantly attenuated compared with wild-type animals and that in the lung the number of interleukin 4 (IL4)-producing cells was decreased; whereas Kawamoto *et al.*[31] showed that the absence of *C3aR* in mice results in significantly increased level of Th2 cytokines (IL4, IL5 and IL10). In spite of the fact that the two groups' results are contradictory calling for further examination, the observed functional relationship clearly indicates a modulator role of C3 on IL4 cytokine expression. IL4 signal transduction is mediated through the α subunit of the IL4 receptor (IL4RA), which is IL4 specific. Thus, the *C3-IL4-IL4RA* axis might be one of the plausible models for the interaction between *C3* and *IL4RA*. With regard to *ADAM33* and *TBXA2R*, one common feature that could indicate their putative interaction is their involvement in angiogenesis, a process frequently underestimated in the pathophysiology of asthma.[32] Novel findings on *ADAM33* showed that its catalytic domain promoted endothelial cell (EC) proliferation *in vitro*, and formation of new vessels *ex vivo* and *in vivo*.[33] *TBXA2R* is also known to be implicated in neovascularization but in an opposed way: suppresses EC migration and angiogenesis by inhibiting the effector pathways of the vascular endothelial growth factor (VEGF), a key angiogenic and chemotactic regulator of EC.[34] Although the exact mechanism by which *ADAM33* exerts its proangiogenic effect is yet to be elucidated, the involvement of VEGF is likely to take place. In that case, the above findings will suggest interactive effect of *ADAM33* and *TBXAR* on VEGF regulation and consequently on angiogenesis and microvascular remodeling of conductive airways in asthma.

Nevertheless, the significant results of our replication study as well as of the gene–gene interactions investigation should be interpreted with caution for inflation of type 1 errors. We have presented our findings based on the nominal α threshold of <0.05 without taking into account multiple testing. Relative to the replication study, this study is not an exploratory study aimed to find a 'significant' gene from multiple candidates but rather to test for confirmation of previously well-established hypotheses. Indeed, the genes from Group 1 and Group 2 are the top asthma and allergy related genes, each replicated in at least six or more independent populations, meaning they all have a high previous probability to show true associations even in the case of a relaxed threshold value for significance. However, if we adjusted for multiple comparisons by the Bonferroni method, none of our significant findings would survive this stringent level of correction. It is obvious that the power is enough to detect genetic effect with OR of around 1.4 with the current sample size, but not if we consider multiple testing. The same is for the results obtained from the screening of the interactions between two polymorphisms. If we strictly applied Bonferroni correction, the significant P-values would need to be in the order of 9.46×10^{-5} (0.05/528) because we carried out $33C^2$=528 tests for each phenotype; no P-value reached this value. Thus, our findings for the potential gene–gene interactions must be evaluated physiologically or by analyses of other sets of samples to validate these observations.

There are other limitations to this study. We focused on the effect of genetic polymorphisms on dichotomous phenotypes and ignored clinical severity and environmental factors. There was also a delay between the recruitment of child asthma cases and child control samples, which could be a source of bias due to differences in DNA processing as well as in environmental exposure. Although population stratification was not controlled in this study, we consider the confounding effect of this factor to be of a lesser extent in comparison to studies conducted on North American[35,36] or Western European[37,38] populations. From the genetic point of view, this

Table 4 Odds ratio and 95% CI of significant polymorphisms found in the basic association studies

Gene	Polymorphism	OR (95%CI)		
		Allele	Dom.	Rec.
Childhood asthma vs child control				
Group 1				
IL13	Arg110Gln	×	×	×
	−1112C>T	1.40 (1.07–1.84)	×	3.01 (1.40–6.51)
ADAM33	Met764Thr	×	×	×
	3236C>T	×	×	×
Group 2				
LTC4S	−444A>C	1.40 (1.05–1.88)	1.47 (1.05–2.06)	×
CCL5	−403A>G	×	×	×
	−28C>G	×	1.43 (1.00–2.05)	×
Group 3				
IL12B	−6415CTCTAA>GC	1.33 (1.07–1.66)	×	1.78 (1.25–2.55)
	1146 C> A	1.30 (1.04–1.61)	×	1.73 (1.21–2.48)
C3	Block 2 (haplotype 6)	1.92 (1.12–3.31)	×	×
	Block 4 (haplotype 1)	×	×	×
Adult asthma vs adult control				
Group 1				
TNF	−1037C>T	1.53 (1.20–1.96)	1.62 (1.22–2.16)	×
ADAM33	Met764Thr	1.47 (1.10–1.96)	1.57 (1.14–2.15)	×
	13236C>T	×	×	×
Group 2				
NOS1	GT repeat intron 2 (187allele)	1.42 (1.20–1.71)	1.55 (1.19–2.03)	1.56 (1.10–2.19)
	GT repeat intron 2 (183allele)	1.29 (1.07–1.55)	1.48 (1.11–1.97)	×
Group 3				
C3	Block 2 (haplotype 6)	×	×	×
	Block 4 (haplotype 1)	1.34 (1.14–1.64)	1.38 (1.04–1.84)	1.60 (1.19–2.14)
SOCS1	−1478CA>del	1.73 (1.27–2.36)	1.69 (1.20–2.37)	3.93 (1.20–12.86)
All asthma vs all controls				
Group 1				
TNF	−1037C>T	1.32 (1.10–1.59)	1.36 (1.10–1.68)	×
ADAM33	Met764Thr	1.26 (1.02–1.56)	1.28 (1.02–1.61)	×
	13236C>T	1.19 (1.01–1.40)	×	×
Group 2				
NOS1	GT repeat intron 2 (187allele)	1.17 (1.02–1.35)	×	×
	GT repeat intron 2 (183allele)	1.15 (1.00–1.33)	×	×
Group 3				
IL12B	−6415CTCTAA>GC	1.20 (1.04–1.37)	×	1.44 (1.14–1.81)
	1146 C> A	1.24 (1.08–1.42)	×	1.49 (1.18–1.88)
C3	Block 2 (haplotype 6)	1.53 (1.08–2.15)	1.58 (1.11–2.25)	×
	Block 4 (haplotype 1)[a]	1.27 (1.10–1.45)	1.39 (1.21–1.73)	1.31 (1.05–1.63)
SOCS1	−1478CA>del	1.47 (1.16–1.96)	1.43 (1.11–1.86)	3.09 (1.17–8.16)

Abbreviations: Allele, χ^2-test of allele frequency; CI, confidence interval; dom., 2×2 dominant model genotype χ^2-test; OR, odds ratio; rec., 2×2 recessive model genotype χ^2-test.
[a]For haplotype description please refer Inoue *et al.*[11]
Polymorphisms with a *P*-value ⩾0.05 in all association tests of the four genetic models are not shown.
×$P \geqslant 0.05$.

assumption is based on the fact that our control subjects were residents of the mainland of Japan, the population of which belongs to the genetically homogeneous Hondo cluster,[39] and also on the results of genomic control analysis[40] that showed the populations from the Kinki and Kanto regions (where we recruited our samples and controls) do not differ in the allele frequency of the null marker. In terms of stratification determined by an individual's socioeconomic position, we would refer to the specific egalitarian characteristic of the Japanese society in support of our claim.[41]

In conclusion, our findings and previous studies suggest that *IL13*, *TNF*, *IL4RA*, *ADAM33* and *TBXA2R* might represent the major asthma and asthma-related traits genes common across populations. *GSTM1*, *GSTP1*, *LTC4S*, *AAA1*, *NOS1* and *CCL5* along with *MMP9*, *IL12B*, *C3* and *SOCS1* might be additional susceptibility genes, which have stronger effects in the Japanese population. Despite our failure to replicate the other genes, our results were not strong enough to eliminate them from the candidate gene list because we did not investigate all known variations in these genes and we did not consider

the effects of environmental factors. Replication studies of genotype–phenotype associations with sample sizes ranging from several hundred to several thousand are not exempt from inconsistencies in findings and have low replication rates. Given the present limited availability of biobanks, methodologically irreproachable studies that integrate more detailed clinical information and that explore the effects of genes in their entirety by dissecting the direct and interactive effects from environmental factors and other genes are required to improve the power and reproducibility of genetic association studies.

ACKNOWLEDGEMENTS

We thank all patients and their families, the volunteers who served as controls and all staff members at the hospitals involved in this study. We also thank Kazuko Hatori, Rieko Yoshida, Yoshiko Hotta and Miyako Takano for their excellent technical assistance. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and grants from the Ministry of Health, Labor and Welfare, Japan. We also thank the two referees for their careful reviews, thoughtful comments and their helpful suggestions that greatly improved the paper.

1 Los, H., Koppelman, G. & Postma, D. The importance of genetic influences in asthma. *Eur. Respir. J.* **14**, 1210–1227 (1999).
2 Ober, C. & Hoffjan, S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun.* **7**, 95–100 (2006).
3 Vercelli, D. Discovering susceptibility genes for asthma and allergy. *Nat. Rev. Immunol.* **8**, 169–182 (2008).
4 Chanock, S. J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D. J., Thomas, G. *et al.* Replicating genotype-phenotype associations. *Nature* **447**, 655–660 (2007).
5 Heinzmann, A., Mao, X., Akaiwa, M., Kreomer, R. T., Gao, P., Ohshima, K. *et al.* Genetic variants of IL-13 signalling and human asthma and atopy. *Hum. Mol. Genet.* **9**, 549–559 (2000).
6 Unoki, M., Furuta, S., Onouchi, Y., Watanabe, O., Doi, S., Fujiwara, H. *et al.* Association studies of 33 single nucleotide polymorphisms (SNPs) in 29 candidate genes for bronchial asthma: positive association a T924C polymorphism in the thromboxane A2 receptor gene. *Hum. Genet.* **106**, 440–446 (2000).
7 Kamada, F., Mashimo, Y., Inoue, H., Shao, C., Hirota, T., Doi, S. *et al.* The GSTP1 gene is a susceptibility gene for childhood asthma and the GSTM1 gene is a modifier of the GSTP1 gene. *Int. Arch. Allergy Immunol.* **144**, 275–286 (2007).
8 Hirota, T., Hasegawa, K., Obara, K., Matsuda, A., Akahoshi, M., Nakashima, K. *et al.* Association between ADAM33 polymorphisms and adult asthma in the Japanese population. *Clin. Exp. Allergy* **36**, 884–891 (2006).
9 Nakashima, K., Hirota, T., Obara, K., Shimizu, M., Doi, S., Fujita, K. *et al.* A functional polymorphism in MMP-9 is associated with childhood atopic asthma. *Biochem. Biophys. Res. Commun.* **344**, 300–307 (2006).
10 Hirota, T., Suzuki, Y., Hasegawa, K., Obara, K., Matsuda, A., Akahoshi, M. *et al.* Functional haplotypes of IL-12B are associated with childhood atopic asthma. *J. Allergy Clin. Immunol.* **116**, 789–795 (2005).
11 Inoue, H., Mashimo, Y., Funamizu, M., Shimojo, N., Hasegawa, K., Hirota, T. *et al.* Association study of the C3 gene with adult and childhood asthma. *J. Hum. Genet.* **53**, 728–738 (2008).
12 Harada, M., Nakashima, K., Hirota, T., Shimizu, M., Doi, S., Fujita, K. *et al.* Functional polymorphism in the suppressor of cytokine signaling 1 gene associated with adult asthma. *Am. J. Respir. Cell Mol. Biol.* **36**, 491–496 (2007).
13 Phillips, P. C. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867 (2008).
14 Hasegawa, K., Tamari, M., Shao, C., Shimizu, M., Takahashi, N., Mao, X. Q. *et al.* Variations in the C3, C3a receptor, and C5 genes affect susceptibility to bronchial asthma. *Hum. Genet.* **115**, 295–301 (2004).
15 Suzuki, Y., Hattori, S., Mashimo, Y., Funamizu, M., Kohno, Y., Okamoto, Y. *et al.* CD14 and IL4R gene polymorphisms modify the effect of day care attendance on serum IgE levels. *J.Allergy Clin. Immunol.* **123**, 1408–1411.e1 (2009).
16 Holgate, S. T. Epithelium dysfunction in asthma. *J. Allergy Clin. Immunol.* **120**, 1233–1244 (2007).
17 Baldwin, L. & Roche, W. R. Does remodelling of the airway wall precede asthma? *Paediatr. Respir. Rev.* **3**, 315–320 (2002).
18 Jenkins, H. A., Cherniack, R., Szefler, S. J., Covar, R., Gelfand, E. W. & Spahn, J. D. A comparison of the clinical characteristics of children and adults with severe asthma. *Chest* **124**, 1318–1324 (2003).
19 Gelfand, E. W. Pediatric asthma: a different disease. *Proc. Am. Thorac. Soc.* **6**, 278–282 (2009).
20 Culley, F. J., Pennycook, A. M. J., Tregoning, J. S., Dodd, J. S., Walzl, G., Wells, T. N. *et al.* Role of CCL5 (RANTES) in viral lung disease. *J. Virol.* **80**, 8151–8157 (2006).
21 Murai, H., Terada, A., Mizuno, M., Asai, M., Hirabayashi, Y., Shimizu, S. *et al.* IL-10 and RANTES are elevated in nasopharyngeal secretions of children with respiratory syncytial virus infection. *Allergol. Int.* **56**, 157–163 (2007).
22 Martinez, F.D. The origins of asthma and chronic obstructive pulmonary disease in early life. *Proc. Am. Thorac. Soc.* **6**, 272–277 (2009).
23 Sigurs, N., Bjarnason, R., Sigurbergsson, F. & Kjellman, B. Respiratory syncytial virus bronchiolitis in infancy is an important risk factor for asthma and allergy at age 7. *Am. J. Respir. Crit. Care Med.* **161**, 1501–1507 (2000).
24 Contopoulos-Ioannidis, D. G., Kouri, I. N. & Ioannidis, J. P. Genetic predisposition to asthma and atopy. *Respiration* **74**, 8–12 (2007).
25 Rogers, A. J., Raby, B. A., Lasky-Su, J. A., Murphy, A., Lazarus, R., Klanderman, B. J. *et al.* Assessing the reproducibility of asthma candidate gene associations, using genome-wide data. *Am. J. Respir. Crit. Care Med.* **179**, 1084–1090 (2009).
26 Daley, D., Lemire, M., Akhabir, L., Chan-Yeung, M., He, J. Q., McDonald, T. *et al.* Analyses of associations with asthma in four asthma population samples from Canada and Australia. *Hum. Genet.* **125**, 445–459 (2009).
27 Ioannidis, J. P., Trikalinos, T. A., Ntzani, E. E. & Contopoulos-Ioannidis, D. G. Genetic associations in large versus small studies: an empirical assessment. *Lancet* **361**, 567–571 (2003).
28 Humbles, A. A., Lu, B., Nilsson, C. A., Lilly, C., Israel, E., Fujiwara, Y. *et al.* A role for the C3a anaphylatoxin receptor in the effector phase of asthma. *Nature* **406**, 998–1001 (2000).
29 Drouin, S. M., Corry, D. B., Kildsgaard, J. & Wetsel, R. A. Cutting edge: the absence of C3 demonstrates a role for complement in Th2 effector functions in a murine model of pulmonary allergy. *J. Immunol.* **167**, 4141–4145 (2001).
30 Drouin, S. M., Corry, D. B., Hollman, T. J., Kildsgaard, J. & Wetsel, R. A. Absence of the complement anaphylatoxin C3a receptor suppresses Th2 effector functions in a murine model of pulmonary allergy. *J. Immunol.* **169**, 5926–5933 (2002).
31 Kawamoto, S., Yalcindag, A., Laouini, D., Brodeur, S., Bryce, P., Lu, B. *et al.* The anaphylatoxin C3a downregulates the Th2 response to epicutaneously introduced antigen. *J. Clin. Invest.* **114**, 399–407 (2004).
32 Bischof, R. J., Bourke, J. E., Hirst, S. J., Meeusen, E. N. T., Snibson, K. J. & Van Der Velden, J. Measurement and impact of remodeling in the lung: airway neovascularization in asthma. *Proc. Am. Thorac. Soc.* **6**, 673–677 (2009).
33 Puxeddu, I., Pang, Y. Y., Harvey, A., Haitchi, H. M., Nicholas, B., Yoshisue, H. *et al.* The soluble form of a disintegrin and metalloprotease 33 promotes angiogenesis: implications for airway remodeling in asthma. *J. Allergy Clin. Immunol.* **121**, 1400–1406.e4 (2008).
34 Ashton, A. W. & Ware, J. A. Thromboxane A2 receptor signaling inhibits vascular endothelial growth factor-induced endothelial cell differentiation and migration. *Circ. Res.* **95**, 372–379 (2004).
35 Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R. *et al.* Estimating African American admixture proportions by use of population-specific alleles **63**, 1839–1851 (1998).
36 Adler, N.E. & Rehkopf, D. H. US disparities in health: descriptions, causes, and mechanisms. *Ann. Rev. Public Health* **29**, 235–252 (2008).
37 Tian, C., Plenge, R. M., Ransom, M., Lee, A., Villoslada, P., Selmi, C. *et al.* Analysis and application of European genetic substructure using 300 K SNP information. *PLoS Genet.* **4**, e4 (2008).
38 Mackenbach, J. P., Stirbu, I., Roskam, A.- J. R., Schaap, M. M., Menvielle, G., Leinsalu, M. *et al.* Socioeconomic inequalities in health in 22 European countries. *N. Engl. J. Med.* **358**, 2468–2481 (2008).
39 Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M. *et al.* Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
40 Imada, Y., Fujimoto, M., Hirata, K., Hirota, T., Suzuki, Y., Saito, H. *et al.* Large scale genotyping study for asthma in the Japanese population. *BMC Res. Notes* **2**, 54 (2009).
41 Kagamimori, S., Gaina, A. & Nasermoaddeli, A. Socioeconomic status and health in the Japanese population. *Soc. Sci. Med.* **68**, 2152–2160 (2009).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

20

Genome-wide association study of intracranial aneurysm identifies three new risk loci

Yasuno K, Bilguvar K, Bijlenga P, Low SK, Krischek B, Auburger G, Simon M, Krex D, Arlier Z, Nayak N, Ruigrok YM, Niemelä M, Tajima A, von und zu Fraunberg M, Dóczi T, Wirjatijasa F, Hata A, Blasco J, Oszvald A, Kasuya H, Zilani G, Schoch B, Singh P, Stüer C, Risselada R, Beck J, Sola T, Ricciardi F, Aromaa A, Illig T, Schreiber S, van Duijn CM, van den Berg LH, Perret C, Proust C, Roder C, Ozturk AK, Gaál E, Berg D, Geisen C, Friedrich CM, Summers P, Frangi AF, State MW, Wichmann HE, Breteler MM, Wijmenga C, Mane S, Peltonen L, Elio V, Sturkenboom MC, Lawford P, Byrne J, Macho J, Sandalcioglu EI, Meyer B, Raabe A, Steinmetz H, Rüfenacht D, Jääskeläinen JE, Hernesniemi J, Rinkel GJ, Zembutsu H, Inoue I, Palotie A, Cambien F, Nakamura Y, Lifton RP, Günel M.

Nat Genet ,May;42 (5) Epub 2010 Apr 4. 420-5, 2010

nature genetics

Genome-wide association study of intracranial aneurysm identifies three new risk loci

Katsuhito Yasuno[1–3*], Kaya Bilguvar[1–3], Philippe Bijlenga[4], Siew-Kee Low[5], Boris Krischek[6], Georg Auburger[7], Matthias Simon[8], Dietmar Krex[9], Zulfikar Arlier[1–3], Nikhil Nayak[1–3], Ynte M Ruigrok[10], Mika Niemelä[11], Atsushi Tajima[12], Mikael von und zu Fraunberg[13], Tamás Dóczi[14], Florentina Wirjatijasa[7], Akira Hata[15], Jordi Blasco[16], Agi Oszvald[17], Hidetoshi Kasuya[18], Gulam Zilani[19], Beate Schoch[20], Pankaj Singh[21,22], Carsten Stüer[23], Roelof Risselada[24], Jürgen Beck[17], Teresa Sola[25], Filomena Ricciardi[7], Arpo Aromaa[26], Thomas Illig[27], Stefan Schreiber[28], Cornelia M van Duijn[29,30], Leonard H van den Berg[10], Claire Perret[31], Carole Proust[31], Constantin Roder[6], Ali K Ozturk[1–3], Emília Gaál[1–3,11], Daniela Berg[32], Christof Geisen[33], Christoph M Friedrich[34], Paul Summers[19], Alejandro F Frangi[35–37], Matthew W State[3,38,39], H Erich Wichmann[27], Monique M B Breteler[29,30], Cisca Wijmenga[40], Shrikant Mane[41], Leena Peltonen[42,43], Vivas Elio[25], Miriam C J M Sturkenboom[24], Patricia Lawford[21], James Byrne[19], Juan Macho[16], Erol I Sandalcioglu[20], Bernhard Meyer[23], Andreas Raabe[17], Helmuth Steinmetz[7], Daniel Rüfenacht[4,44], Juha E Jääskeläinen[13], Juha Hernesniemi[11], Gabriel J E Rinkel[10], Hitoshi Zembutsu[5], Ituro Inoue[12], Aarno Palotie[42,43], François Cambien[31], Yusuke Nakamura[5], Richard P Lifton[3,45,46] & Murat Günel[1–3]

© 2010 Nature America, Inc. All rights reserved.

Saccular intracranial aneurysms are balloon-like dilations of the intracranial arterial wall; their hemorrhage commonly results in severe neurologic impairment and death. We report a second genome-wide association study with discovery and replication cohorts from Europe and Japan comprising 5,891 cases and 14,181 controls with ~832,000 genotyped and imputed SNPs across discovery cohorts. We identified three new loci showing strong evidence for association with intracranial aneurysms in the combined dataset, including intervals near *RBBP8* on 18q11.2 (odds ratio (OR) = 1.22, $P = 1.1 \times 10^{-12}$), *STARD13-KL* on 13q13.1 (OR = 1.20, $P = 2.5 \times 10^{-9}$) and a gene-rich region on 10q24.32 (OR = 1.29, $P = 1.2 \times 10^{-9}$). We also confirmed prior associations near *SOX17* (8q11.23–q12.1; OR = 1.28, $P = 1.3 \times 10^{-12}$) and *CDKN2A-CDKN2B* (9p21.3; OR = 1.31, $P = 1.5 \times 10^{-22}$). It is noteworthy that several putative risk genes play a role in cell-cycle progression, potentially affecting the proliferation and senescence of progenitor-cell populations that are responsible for vascular formation and repair.

Intracranial aneurysms affect approximately 2% of the general population and arise from the action of multiple genetic and environmental risk factors[1]. We previously reported the first genome-wide association study (GWAS) of intracranial aneurysms[2] that identified three risk loci on chromosomes 8q11.23–q12.1, 9p21.3 and 2q33.1 with $P < 5 \times 10^{-8}$. This previous study had limited power to detect loci imparting genotypic relative risk (GRR) <1.35 (Supplementary Table 1).

To increase the power to detect additional loci of similar or smaller effect, we ascertained and whole-genome genotyped two new European case cohorts (n = 1,616) and collected genotyping data from five additional European control cohorts (Supplementary Note, n = 11,955). We also increased the size of the original Japanese replication cohort and added a new one (2,282 affected individuals (cases) and 905 controls) (Table 1). The new combined cohort had nearly threefold more cases than the original cohort and increased our power to detect variants with modest effect sizes. For example, this study had 89% and 64% average power to detect common variants (minor allele frequencies (MAF) ≥ 10%) with GRR of 1.25 and 1.20, respectively (Supplementary Table 1).

All subjects were genotyped using the Illumina platform. The new as well as the previously analyzed genotyping data were subjected to well-established quality-control measures (Supplementary Table 2). We sought to eliminate potential confounding due to population stratification and gender[1,3] by matching cases and controls of the same gender based on inferred genetic ancestry. As previous studies demonstrated that the Finnish population forms an ancestry cluster distinct from other European populations similar to those included in this study[4,5], we analyzed our Finnish cohort independently from the others. To maximize opportunities for genetic matching and analytic power, we analyzed all subjects in the remaining European cohorts together. The resulting matched case-control data consisted of 808 cases and

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

Received 10 December 2009; accepted 9 March 2010; published online 4 April 2010; doi:10.1038/ng.563

Table 1 Overview of the study cohorts

	Cohort	Case (*n*)	Control (*n*)	Quality control-passed SNPs (*n*)	GIF
Discovery	Finland (FI)	808	4,393	1,303,876	1.074
	Combined European (CE)	1,972	8,122	905,906	1.094
	Total discovery	2,780	12,515	831,532	1.007
CE subcohorts	NL	708	3,954	905,906	1.108
	DE	789	2,228	905,906	1.059
	AN	475	1,940	905,906	1.057
Replication	Japan 1 (JP1)	829	761	12	
	Japan 2 (JP2)	2,282	905	13	
	Total replication	3,111	1,666	12	
	Total	5,891	14,181	12	

Combined European cohort consisted of all European subjects who were not ascertained in Finland. Sub-cohorts of the European cohort were defined on the basis of case series: NL, cases from The Netherlands with matched controls; DE, German cases with matched controls; AN, @neurIST cases with matched controls. NL, DE and AN were exclusive subsets of the European cohort (see also **Supplementary Table 3**). AN cases consisted of subjects from Germany, Great Britain, Hungary, The Netherlands, Switzerland and Spain. JP1 and JP2 were two independent Japanese case-control cohorts. Genomic inflation factors of the Finnish and European cohorts (as well as NL, DE and AN) were calculated for 1,303,876 and 905,906 SNPs, respectively. The genomic inflation factor of the discovery cohort (total discovery) was based on the meta-analysis result for 831,532 SNPs after correcting each cohort for genomic control. The discovery data (combined Finnish and European cohorts) was not corrected for genomic control. GIF, genomic inflation factor.

4,393 controls in the Finnish cohort and 1,972 cases and 8,122 controls in the rest of the combined European cohort (**Supplementary Table 3**). We used the genotype data that passed quality-control filters and phased chromosomes from the HapMap CEU sample to impute missing genotypes[6]. We based our further analyses on 831,534 SNPs that passed the quality-control filters both in the Finnish and European samples (**Table 1** and **Supplementary Table 2**).

We tested for association of each quality control–passed SNP with intracranial aneurysms using conditional logistic regression, assuming a log-additive effect of allele dosage. We corrected each cohort for residual overdispersion (**Table 1**) using genomic control[7] and combined the results from the Finnish and European cohorts to obtain *P* values, ORs and CIs for the discovery cohort of 2,780 cases and 12,515 controls using a fixed-effects model.

To evaluate the strength of association, in addition to obtaining *P* values, we employed a Bayesian approach[8]. We used the Bayes factor that represents the fold-change of the odds of association before and after observing the data[9] and the posterior probability of association (PPA), calculated through the Bayes factor, that provides a simple probabilistic measure of the evidence of association[8,10]. For every SNP, we assumed a uniform prior probability of association of 1/10,000 and set the prior of the logarithm of the per-allele OR as a normal distribution with a 95% probability for the OR to be between 0.67 and 1.5, with larger weights for smaller effect sizes[9,11].

From the discovery results, we eliminated two imputed SNPs that showed PPAs of 0.97 and 0.94 because their association signals were not supported by surrounding genotyped SNPs and because their genotypes were not confirmed by direct genotyping results (data not shown). This resulted in 831,532 SNPs that passed quality control (**Supplementary Table 2**).

We observed three regions that showed very high PPAs (>0.995; **Fig. 1a**) and also a substantial excess of SNPs with $P < 1 \times 10^{-3}$ (1,295 SNPs versus 831 SNPs expected by chance) even after excluding those within previously identified associated regions[2] (**Fig. 1b**). Moreover, we observed a strong correlation between the *P* values and Bayes factors for the upper tail of the distribution (**Fig. 1c**).

We focused on five genomic regions (**Fig. 1a**) that contained at least one SNP with PPA >0.5 for which the hypothesis of association with intracranial aneurysm was more likely than the null hypothesis of no association. The PPAs of the most highly associated SNPs in these intervals ranged from 0.6621 to >0.9999 and the *P* values ranged from 7.9×10^{-7} to 2.2×10^{-16} (**Supplementary Table 4**). The five chromosomal segments included three newly identified SNP clusters on 10q24.32, 13q13.1 and 18q11.2. The remaining two regions were previously identified loci on 8q11.23–q12.1 and 9p21.3 (**Fig. 2**; ref. 2). The third locus identified in our previous study, on 2q33, did not contain any SNPs with PPA >0.5. Furthermore, consistent with our previous results[2], detailed analysis of the 8q11.23–q12.1 region detected two independent association signals within the <100-kb interval that spans the *SOX17* locus (**Fig. 2** and **Supplementary Fig. 1**); these two signals are hereafter referred to as 5′-*SOX17* and 3′-*SOX17*. Thus, the five chromosomal segments comprised six independent association signals for follow-up.

We performed replication genotyping in two Japanese cohorts including 3,111 cases and 1,666 controls (JP1 and JP2, see **Table 1**). For each independent signal, we selected for replication the genotyped SNP with the highest PPA and added up to two additional SNPs per locus. For the 5′-*SOX17* region, we selected two SNPs analyzed previously, as they tag the most significant SNP in the current study (**Supplementary Fig. 1**).

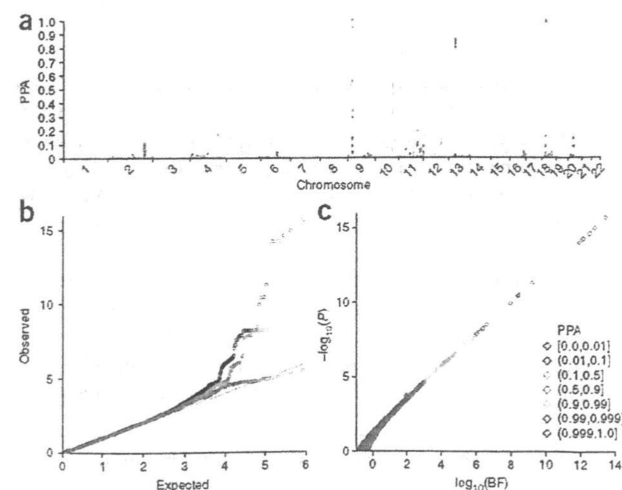

Figure 1 Genome-wide association analysis results in the discovery cohort. (**a**) The PPAs for 831,532 quality control–passed SNPs that were analyzed specifying a prior probability of association of 1/10,000 are plotted against genomic locations of SNPs. A gray horizontal line at PPA = 0.5 indicates the cutoff value for follow-up genotyping. (**b**) Quantile-quantile plots of *P* values ($-\log_{10}$ scale) are shown for all the SNPs analyzed (black; $n = 831{,}532$); for SNPs after excluding those within previously identified regions (red; $n = 830{,}907$); and for SNPs after excluding all within the final associated intervals (blue; $n = 830{,}158$). (**c**) A scatter plot of $-\log_{10} P$ versus $\log_{10}$ Bayes factors is shown with color for each point indicating the range of PPA values. There are very close relationships among the *P* values for association, the Bayes factor and the PPA value. Note that, given a uniform prior probability of association, the PPA increases as the Bayes factor increases. A vertical line indicates the minimum PPA threshold at 0.5 (Bayes factor = 1.0×10^4) for follow-up.

© 2010 Nature America, Inc. All rights reserved.

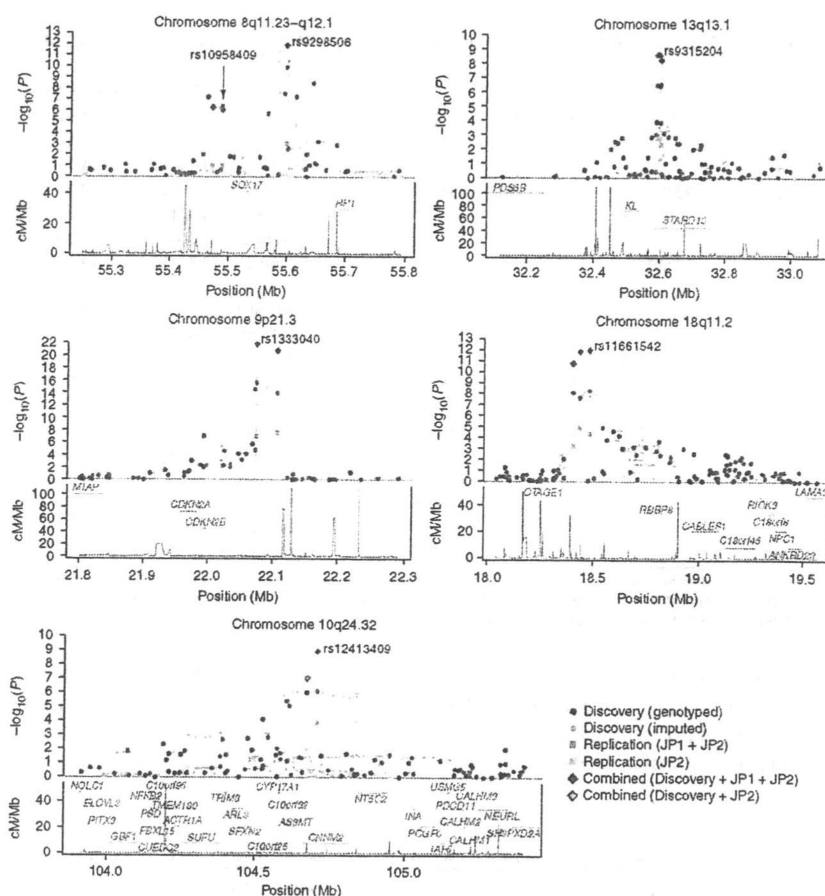

Figure 2 Regional plots for associated regions. For each chromosomal interval, $-\log_{10} P$ values for association are plotted against the genomic coordinates (NCBI build 36) in the panel above; the recombination rates obtained from the HapMap database and the RefSeq genes (hg18) within the regions are shown in the panel below. Above, rs identifiers of SNPs listed in **Table 2** are shown and their positions are indicated by gray vertical lines. Gray dashed lines indicate locations of other SNPs genotyped in the replication cohorts. Dark blue and light blue dots represent results of genotyped and imputed SNPs for the discovery cohort, respectively; orange and light orange squares represent association results for the replication cohort using JP1 combined with JP2 and also JP2-only, respectively; combined results for SNPs genotyped both in the discovery and the replication cohort using JP1 plus JP2 and JP2-only are shown by red and light red diamonds, respectively.

© 2010 Nature America, Inc. All rights reserved.

All but one of the SNPs (rs12411886 on 10q24.32 in JP1) were successfully genotyped and passed quality-control filters. We tested for association of each SNP with intracranial aneurysm using logistic regression stratified by gender, specifying the same model as for the discovery cohort (**Supplementary Table 5**). We combined results from JP1 and JP2 using a fixed-effects model (**Table 2** and **Supplementary Table 4**). We considered an association to be replicated if the Bayes factor increased the odds of association more than ten-fold after the replication data was observed.

Of the six candidate loci, all but the 5′-*SOX17* interval were replicated, with replication *P* values ranging from 0.0019 to 1.0×10^{-7}, and the odds of association with intracranial aneurysm increasing by 22.9-fold to 1.5×10^5-fold, yielding robust evidence for replication for each interval (**Table 2**).

We combined the discovery and replication results using a fixed-effects model. All of the five loci that replicated in the Japanese cohort surpassed the conventional threshold for genome-wide significance ($P < 5 \times 10^{-8}$), with *P* values ranging from 2.5×10^{-9} to 1.5×10^{-22}, and all also had PPAs ≥ 0.998 (**Table 2**).

In order to determine each cohort's contribution to the observed association and to assess the consistency of the effect size across groups, we analyzed each of them separately (**Table 1** and **Supplementary Table 5**) and then combined the results from the six cohorts using a random-effects model. The association results remained highly significant (**Fig. 3**). For the five loci that were replicated in the Japanese cohorts, we found no evidence of significant heterogeneity ($P > 0.1$). Every cohort had the same risk allele and provided support for association with the exception of the JP1 sample for the 3′-*SOX17* locus, consistent with our previous study[2] (**Fig. 3**).

The most significant association was detected in the previously reported[2] 9p21.3 region near *CDKN2A* and *CDKN2B* with $P = 1.5 \times 10^{-22}$ (OR = 1.32, PPA > 0.9999). All of the newly studied cohorts strongly supported this association with intracranial aneurysm (**Fig. 3**). The same allele is associated with coronary artery disease but not with type 2 diabetes[12]. Similarly, the previously reported 8q11.23–q12.1 region showed significant association. The 3′-*SOX17* interval (rs9298506) showed robust association with $P = 1.3 \times 10^{-12}$ (OR = 1.28, PPA > 0.9999) and all new cohorts supported the association of this SNP with intracranial aneurysm (**Fig. 3**). For the 5′-*SOX17* region (rs10958409), the new cohorts introduced a substantial heterogeneity, lowering the PPA to 0.016 (**Fig. 3**).

Among the newly identified loci, the strongest association was found at rs11661542 on 18q11.2 (OR = 1.22, $P = 1.1 \times 10^{-12}$, PPA > 0.9999). A cluster of SNPs that is associated with intracranial aneurysm spans the interval between 18.400 Mb and 18.509 Mb and is strongly correlated with rs11661542 (**Fig. 2**). A single gene, *RBBP8* (encoding the retinoblastoma binding protein 8), is located within an extended linkage disequilibrium interval (**Fig. 2**).

The second strongest new association was at rs12413409 on 10q24.32 (OR = 1.29, $P = 1.2 \times 10^{-9}$, PPA = 0.9990), which maps to intron 1 of *CNNM2* (encoding cyclin M2) (**Fig. 2**). A cluster of SNPs that are strongly correlated with rs12413409 and are located within a ~247-kb interval in the same linkage disequilibrium block supported the association (**Fig. 2**).

The third new locus is defined by rs9315204 at 13q13.1 (OR = 1.20, $P = 2.5 \times 10^{-9}$, PPA = 0.9981) in intron 7 of *STARD13* (encoding the StAR-related lipid transfer (START) domain containing 13) (**Fig. 2**). Two SNPs, rs1980781 and rs3742321, that are strongly correlated with rs9315204 ($r^2 > 0.9$) also showed significant association with intracranial aneurysm (**Fig. 2** and **Supplementary Table 4**). These two SNPs are missense (lysine to arginine) and synonymous coding variants of *STARD13*, respectively. Another gene that has been implicated in aging phenotypes, *KL* (encoding klotho), is located nearby[13].

Table 2 Representative SNPs analyzed both in the discovery and replication cohorts

Locus	SNP	Position	Genes	Risk allele	Cohort	*P* value	$\log_{10}$ (Bayes)	PPA	Per-allele OR (95% CI)	Control RAF	Case RAF
8q11.23	rs10958409	55,489,644	*SOX17*	A	Discovery	4.2×10^{-7}	4.64	0.8128	1.24 (1.14–1.35)	0.15, 0.19	0.18, 0.22
					Replication	0.12	-0.11		1.08 (0.98–1.20)	0.28	0.29
					Combined	9.0×10^{-7}	4.30	0.6685	1.17 (1.10–1.25)		
8q12.1	rs9298506	55,600,077	*SOX17*	A	Discovery	1.2×10^{-10}	7.94	0.9999	1.33 (1.22–1.45)	0.81, 0.76	0.85, 0.81
					Replication	0.0012	1.56		1.21 (1.08–1.36)	0.79	0.81
					Combined	1.3×10^{-12}	9.85	$1.0–1.4 \times 10^{-6}$	1.28 (1.20–1.38)		
9p21.3	rs1333040	22,073,404	*CDKN2A*, *CDKN2B*	T	Discovery	2.5×10^{-16}	13.41	$1.0–3.9 \times 10^{-10}$	1.32 (1.24–1.41)	0.56, 0.45	0.63, 0.53
					Replication	1.0×10^{-7}	5.18		1.31 (1.19–1.45)	0.66	0.72
					Combined	1.5×10^{-22}	19.48	$1.0–3.3 \times 10^{-15}$	1.32 (1.25–1.39)		
10q24.32	rs12413409	104,709,086	*CNNM2*	G	Discovery	7.9×10^{-7}	4.29	0.6621	1.38 (1.22–1.57)	0.91, 0.91	0.94, 0.93
					Replication	0.00014	2.34		1.23 (1.10–1.37)	0.74	0.77
					Combined	1.2×10^{-9}	7.00	0.9990	1.29 (1.19–1.40)		
13q13.1	rs9315204	32,591,837	*KL*, *STARD13*	T	Discovery	3.3×10^{-7}	4.73	0.8443	1.21 (1.13–1.31)	0.21, 0.33	0.24, 0.39
					Replication	0.0019	1.36		1.18 (1.06–1.31)	0.24	0.27
					Combined	2.5×10^{-9}	6.72	0.9981	1.20 (1.13–1.28)		
18q11.2	rs11661542	18,477,693	*RBBP8*	C	Discovery	5.6×10^{-9}	6.39	0.9959	1.21 (1.14–1.30)	0.49, 0.44	0.54, 0.47
					Replication	4.5×10^{-5}	2.79		1.22 (1.11–1.34)	0.61	0.65
					Combined	1.1×10^{-12}	9.92	$1.0–1.2 \times 10^{-5}$	1.22 (1.15–1.28)		

Genomic locations for SNPs are based on NCBI build 36, and risk alleles are aligned to the forward strand of the reference sequence. Control and case risk allele frequencies (RAFs) for the discovery cohort are shown in the form: RAF of European cohort, RAF of Finnish cohort. $\log_{10}$ (Bayes) indicates the logarithm of the Bayes factor in favor of association. PPA, posterior probability of association. Genes closest to the listed SNPs within the same LD regions are shown.

A search of the gene-expression database (eQTL browser, see URLs) for all the intracranial aneurysm-risk loci did not reveal any consistent pattern of association of intracranial aneurysm SNPs with variation in gene expression levels.

In this second GWAS of intracranial aneurysm, which included nearly three times as many cases as the initial study, we detected three new risk loci and obtained strong independent evidence for association of two previously identified loci. The evidence that these are

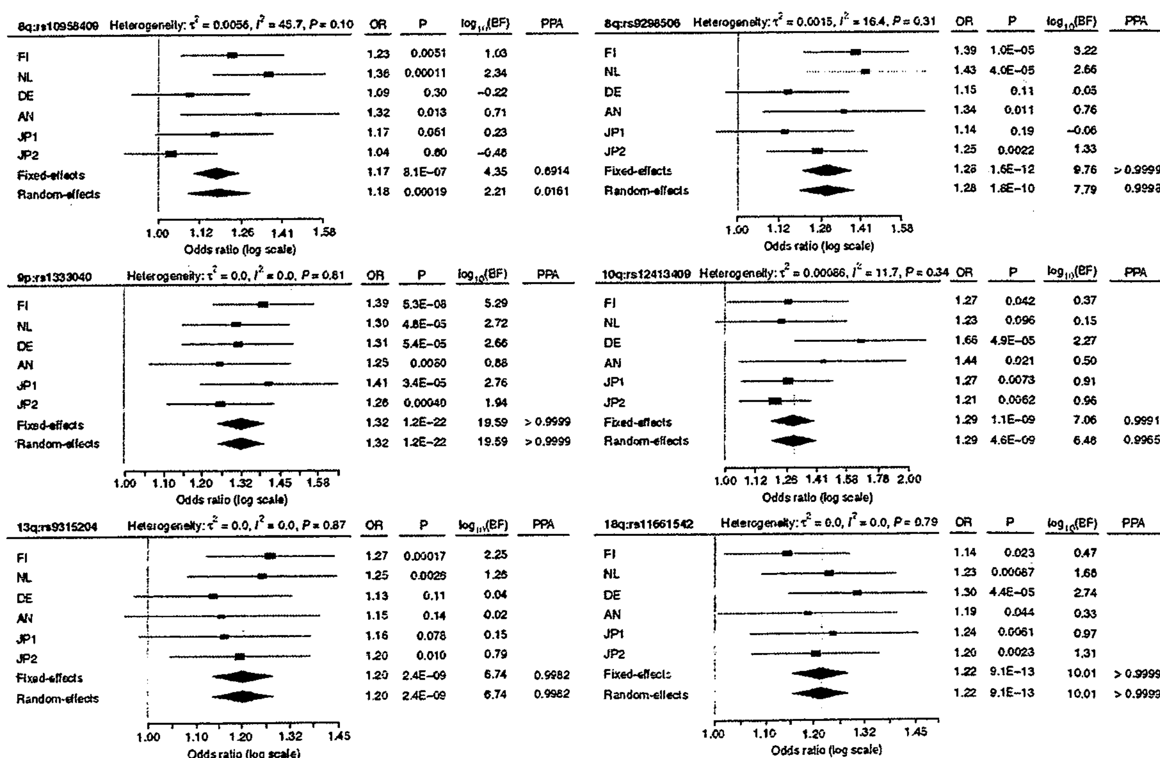

Figure 3 Consistency of association across cohorts. Forest plots are shown for meta-analysis of the SNPs listed in **Table 2**. Squares and horizontal segments represent estimated per-allele ORs and 95% CIs for individual cohorts. Diamonds represent the summary OR estimates and 95% CIs for the meta-analyses of six cohorts (using fixed- and random-effects models). $\log_{10}$ (Bayes factor) > 0 supports association with intracranial aneurysm, whereas $\log_{10}$ (Bayes factor) < 0 supports no association with intracranial aneurysm. Analyzing the results here as six distinct cohorts rather than four cohorts (as in the primary analysis) resulted in only minor differences due to different weights given to sub-cohorts of the combined European cohort.

© 2010 Nature America, Inc. All rights reserved.