

Asthma Phenotypes in Japanese Adults - Their Associations with the *CCL5* and *ADRB2* Genotypes

Yoshiko Kaneko¹, Hironori Masuko¹, Tohru Sakamoto¹, Hiroaki Iijima², Takashi Naito², Yohei Yatagai¹, Hideyasu Yamada¹, Satoshi Konno³, Masaharu Nishimura³, Emiko Noguchi⁴ and Nobuyuki Hizawa¹

ABSTRACT

Background: Cluster analyses were previously performed to identify asthma phenotypes underlying asthma syndrome. Although a large number of patients with asthma develop the disease later in life, these previous cluster analyses focused mainly patients with younger-onset asthma.

Methods: Cluster analysis examined the existence of distinct phenotypes of late-onset asthma in Japanese patients with adult asthma. We then associated genotypes at the *CCL5*, *TSLP*, *IL4*, and *ADRB2* genes with the clusters of asthma identified.

Results: Using the 8 variables of age, sex, age at onset of the disease, smoking status, total serum IgE, %FEV₁, FEV₁/FVC, and specific IgE responsiveness to common inhaled allergens, two-step cluster analysis of 880 Japanese adult asthma patients identified 6 phenotypes: cluster A (*n* = 155): older age at onset, no airflow obstruction; cluster B (*n* = 170): childhood onset, normal-to-mild airflow obstruction; cluster C (*n* = 119): childhood onset, the longest disease duration, and moderate-to-severe airflow obstruction; cluster D (*n* = 108): older age at onset, severe airflow obstruction; cluster E (*n* = 130): middle-age at onset, no airflow obstruction; and cluster F (*n* = 198): older age at onset, mild-to-moderate airflow obstruction. The *CCL5*-28C>G genotype was significantly associated with clusters A, B and D (OR 1.65, *p* = 0.0021; 1.67, 0.018; and 1.74, 0.011, respectively). The *ADRB2* Arg16Gly genotype was also associated with clusters B and D (OR 0.47, *p* = 0.0004; and 0.63, 0.034, respectively).

Conclusions: The current cluster analysis identified meaningful adult asthma phenotypes linked to the functional *CCL5* and *ADRB2* genotypes. Genetic and phenotypic data have the potential to elucidate the phenotypic heterogeneity and pathophysiology of asthma.

KEY WORDS

ADRB2, age at onset, asthma phenotype, *CCL5*, lung function

INTRODUCTION

Asthma describes a clinical syndrome that can be caused by a number of different pathologies, rather than a single disease. The phenotype is described as the clinical, physiological, morphological, and biochemical characteristics of a subtype of the disease that result from the interaction between genes and the environment. Clinical or physiological pheno-

types relevant to asthma include those defined by the presence of allergic diathesis, presence of chronic airflow obstruction, frequency of exacerbations, age of asthma onset, and response to asthma therapy. Asthma phenotypes may also be identified as clusters of measurements from different dimensions of the disease (ie, the clinical features, physiology, immunology, pathology, hereditary components, environmental influences, response to treatment, and so

¹Department of Pulmonary Medicine, ⁴Department of Medical Genetics, Faculty of Medicine, University of Tsukuba, ²Tsukuba Medical Center, Ibaraki and ³First Department of Medicine, Hokkaido University School of Medicine, Hokkaido, Japan.
Conflict of interest: No potential conflict of interest was disclosed.
Correspondence: Nobuyuki Hizawa, Department of Pulmonary

Medicine, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan.

Email: nhizawa@md.tsukuba.ac.jp

Received 17 May 2012. Accepted for publication 7 September 2012.

©0000 Japanese Society of Allergology

forth). A large-scale cluster analysis was performed on 726 patients with asthma in the Severe Asthma Research Program (SARP).¹ This study used 628 variables compressed into 34 weighted variables and identified 5 separate clusters. The strongest predictors were FEV₁% and age at onset. Another cluster analysis cohort revealed similar findings.² Latent class analysis, a clustering model-based method, applied in 2 large epidemiological studies conducted in patients with adult asthma, also identified 4 asthma phenotypes.³

Although these previous cluster analyses focused mainly on patients with younger-onset asthma (age of asthma onset <20 years), asthma can develop at any stage of life and asthma phenotypes show age-related variations,⁴ suggesting that adult-onset asthma may constitute some distinct phenotypes of asthma. In fact, we previously reported that the gain-of-function -28G allele of the promoter SNP (rs2280788: -28C>G) in the *CCL5* gene is associated with susceptibility to late-onset asthma in patients who developed asthma at age >40 years⁵ and with variable expression of emphysema in patients with COPD.⁶ The development of asthma in older adults is also influenced by smoking; compared with younger adults with asthma (18-34 years), older adults with asthma were more likely to report a history of smoking (54% vs. 34%).⁷ Although previously reported cluster analyses did not include patients with asthma who have smoked, possible pathophysiological mechanisms underlying smoking asthma deserves further medical and research attention.

Unsupervised cluster analysis is a useful tool for identifying phenotypes; however, it has not been applied to adult-onset patients with persistent asthma across a wide range of severities. In addition, a cluster, which describes “observable characteristics,” does not necessarily relate to or give any insight into the underlying disease processes that are defined by a distinct functional or pathophysiological mechanism. Therefore, given the previous cluster analyses that emphasized the importance of age of asthma onset in distinguishing asthma clusters, we here investigated the presence of distinct adult-onset asthma phenotypes by applying a clustering method to a well-characterized asthma cohort of Japanese adults. We then searched for the phenotype-genotype correlations by examining the influence of functional genotypes at the *CCL5*, *TSLP*, *IL4*, and *ADRB2* genes on the expression of the particular clusters of asthma identified.

METHODS

STUDY SUBJECTS

This study was a cross-sectional analysis to identify asthma phenotypes, comprising 880 patients with asthma and 1329 healthy controls. Asthma was defined on the basis of recurrent episodes of at least 2

of 3 symptoms (cough, wheeze, or dyspnea) associated with demonstrable reversible airflow limitation (15% variability in FEV₁ or in peak expiratory flow rate either spontaneously or with an inhaled, short-acting β 2-agonist) or with increased airway responsiveness to methacholine, or with both, as described elsewhere.⁸ All patients (*N* = 880) were diagnosed as having asthma by respiratory diseases physicians at University of Tsukuba Hospital, Hokkaido University Hospital, or their affiliated hospitals. All patients had asthma symptoms deemed by their physician to require initiation of asthma-controller therapy. Patients with asthma in the study had persistent asthma with variable airflow limitation and their pulmonary functions with asthma medication were measured at baseline. Because smoking prevalence remains high in Japan and the prevalence rates of smoking in subjects with asthma have also been reported to be similar to those in the general population,^{9,10} the current study included asthma patients with a history of smoking so as not to exclude the diverse patient populations seen in clinical practice. The data for age at onset of asthma were self-reported. To judge the age at onset of asthma as accurately as possible, patients were asked about episodes of dyspnea, wheezing, or cough they had experienced during childhood and puberty. In cases of uncertainty, the time of the earliest respiratory symptoms was designated as the age at onset of asthma symptoms, and used for the calculation of the duration of asthma. Of the 880 asthmatic patients, 298 were previously also studied in our case-control study of the *CCL5* -28C>G SNP.⁵ The remaining patients were newly recruited.

Details of the non-asthmatic healthy subject recruitment (*N* = 1329) have been reported previously.¹¹ Briefly, detailed information on respiratory health, lifestyle, and exposure conditions to environmental irritants such as tobacco smoke, allergens, and air pollution was collected. Lung function data, chest radiographic findings, and results of a standard clinical examination were also available. On the basis of a detailed questionnaire on pulmonary symptoms, a physical examination, chest roentgenogram and lung function data, we carefully excluded the presence of pulmonary diseases such as asthma and COPD. Atopy was assessed by measurement of specific IgE responsiveness to 14 common inhaled allergens including *Dermatophagoides farinae*, grass pollen, animal dander, and molds. We defined atopy as a positive response to at least 1 of the 14 allergens.

The protocol was approved by the ethics committees of all the participating hospitals, and written informed consent was obtained from all the participants.

ALLELE-SPECIFIC PCR AND DETECTION OF FLUORESCENCE-LABELED PCR FRAGMENTS

Five common SNPs at the *CCL5* (rs2280788), *TSLP*

(rs3806933 and rs2289276), *IL4* (rs2070874) and *ADRB2* (rs1042713, Arg16Gly) genes were genotyped using a TaqMan assay (Applied Biosystems, Foster City, CA, USA) combined with a kinetic allele-specific polymerase chain reaction, as described previously.¹¹ These 5 SNPs have been demonstrated to be of functional relevance, and have been consistently associated with asthma or asthma-related phenotypes.^{5,12-15}

STATISTICAL ANALYSIS

Statistical analysis was carried out with SPSS version 19 software (SPSS, Chicago, IL, USA) for the cluster, discriminate, and logistic regression analyses. Cluster analysis is a technique that defines the distances of subjects from each other based on the combined values multidimensional vector of the subjects' measured characteristics. It can be used to describe the phenotypes of a disease without the need for historical or arbitrary *a priori* assumptions about classification. We applied a uniform cluster analysis methodology to the asthma population ($N = 880$) using a two-step approach. In the first step, hierarchical cluster analysis using Ward's method generated a dendrogram for estimation of the number of likely clusters within the studied population. This estimate was pre-specified in a *k*-means cluster analysis that was used as the principal clustering technique.¹⁶ On the basis of previous reports that consistently revealed the strongest discriminatory variables for cluster assignment including age at onset, lung function, and IgE responsiveness,¹⁻³ we decided to apply a rather biased hypothesis-driven approach to identify asthma phenotypes by selecting 8 factors for estimation that may contribute to characterizing asthma phenotypes. Complete data for age, sex, smoking status (current, past or never), age at onset, total IgE levels, atopic status (specific IgE responsiveness to common inhaled allergens), baseline FEV₁, and FEV₁/FVC were available for all of the 880 patients with asthma. We did not include BMI in the cluster analysis because obese female patients with asthma are comparatively rare in Japan, and we also wanted to avoid subject drop-out due to an incomplete dataset for BMI. The natural logarithm of serum IgE was standardized using the Z score.

To compare differences between 1329 healthy controls and 880 patients with asthma, and to compare between clusters, one-way analysis of variance (ANOVA), the Kruskal-Wallis test, and the χ^2 test were used for parametric continuous, nonparametric continuous and categorical variables, respectively. To determine the strongest predictors of cluster assignment, stepwise discriminant analysis of the cluster variables was performed using Wilks' lambda as the statistic, an *F* value of 3.84 as the criterion for determining entry into the model, and an *F* value of 2.71 for determining removal from the model.

For the classification and regression tree (CART) analysis,¹⁷ 2 strongest discriminators (age at onset and %FEV₁) were included in the determination of the model. CART analysis involves binary recursive partitioning, which splits a single parent node into two daughter nodes based on the predictor variable that best stratifies the population into groups with and without the outcome of interest. Each daughter node can be split further into two nodes, and so on. CART was generated through the rpart package [version 3.1-54] incorporated in the R package [version 2.14.0].

To search for phenotype-genotype correlations, we studied 4 candidate genes that could affect the development of distinct asthma phenotypes. These candidate genes include *CCL5*, *TSLP*, *IL4*, and *ADRB2*. We examined the genetic effects of 5 functional SNPs at these genes on the development of each asthma phenotype using multinomial logistic regression analysis. For these case-control analyses, we used 1329 healthy adults as a reference. We adjusted the analysis for possible confounding factors such as age, sex, smoking status (never, past, or current), and log total serum IgE.

RESULTS

Our dataset included 880 asthmatics ranging in age from 16 to 84 years (Table 1). In the first step, hierarchical clustering using Ward's minimum-variance method generated a dendrogram, which indicated 6 clusters within the studied population. As a second step, *k*-means cluster analysis was performed using this estimated number of clusters. The demographic characteristics of each cluster are shown in Table 2. Age, sex, age at onset, duration of the disease, total serum IgE, atopic status, FEV₁/FVC, %FEV₁, and smoking status were all significantly different ($p < 0.0001$) among these 6 clusters.

Cluster A ($n = 155$, median age 65.0 years) was characterized by the oldest age at onset, the shortest duration, normal lung function, the lowest levels of total IgE, the lowest sensitizations to inhaled allergens (atopy), and the lowest prevalence of smokers (late-onset mild less-atopic asthma). Cluster B was the second largest group ($n = 170$, 28.5 years). It was characterized by young atopic individuals and childhood onset with long disease duration (early-onset mild atopic asthma). Despite the rather long disease duration (median, 20 years) and large number of smokers (34.7%), this group was distinguished by retained lung function (%FEV₁, 85.7%; FEV₁/FVC, 79.3%). Cluster C ($n = 119$, 37.0 years) was a male-dominant group characterized by early-onset atopic asthma with the longest disease duration (22 years) and the largest number of smokers (42.0%). Compared with cluster B, this group had lower %FEV₁ (59.3%) and FEV₁/FVC (58.4%) (early-onset moderate-to-severe atopic asthma). Cluster D ($n = 108$, 61.0 years) was

Table 1 Characteristics of the study population

	Healthy Control Subjects	Patients with Asthma	<i>p</i> value*
Number of subjects	1329	880	
Sex (female, %)	713 (53.6)	515 (58.5)	0.013
Age, y (range)	51.0 (22-78)	54.0 (16-84)	0.029
Smoker (past or current, %)	494 (37.2)	319 (36.2)	0.108
† Smoking pack-year: 0	835 (62.8)	561 (76.0)	<0.0001
(%) : 0-10	165 (12.4)	81 (11.0)	
: >10	329 (24.8)	96 (13.0)	
‡ Atopy (%)	786 (59.1)	624 (70.9)	<0.01
Serum IgE (log, SD)	1.78 (0.58)	2.30 (0.64)	<0.0001
FEV ₁ % pred (% , SD)	91.6 (13.0)	80.6 (23.4)	<0.0001
FEV ₁ /FVC (% , SD)	83.0 (5.8)	70.3 (12.9)	<0.0001
§ BMI (kg/m ² , SD)	23.0 (3.1)	23.4 (3.8)	0.008

* χ^2 test, analysis of variance, or Kruskal-Wallis test was used where appropriate.

† Information about smoking pack-year was available for 738 patients with asthma.

‡ Atopy: specific IgE responsiveness to common inhaled allergens.

§ Information about BMI was available for 649 patients with asthma.

FEV₁, forced expiratory volume; FVC, forced vital capacity; % pred, % predicted.

Table 2 Characteristics of the asthma population

	Late-Onset Mild Less-Atopic	Early-Onset Mild Atopic	Early-Onset Moderate-to-Severe Atopic	Late-Onset Severe	Middle-Age Onset Female-Dominant	Late-Onset Moderate Less Atopic	
	cluster A	cluster B	cluster C	cluster D	cluster E	cluster F	<i>p</i> value*
Number of subjects	155	170	119	108	130	198	
Sex (female, %)	66.5	51.2	39.5	55.6	70	64.1	<0.000
Age (y) (range)	65.5 (50-84)	31.2 (16-78)	40.5 (17-79)	62.8 (43-84)	41.7 (21-73)	60.0 (39-83)	<0.000
† BMI	23.9	23.7	23.8	22.6	23.0	23.4	0.10
BMI >30 (%)	5.1	11.1	5.3	2.3	5.3	5.1	0.26
Age onset (y) (range)	58.0 (20-80)	5.0 (0-30)	13.0 (0-36)	49.0 (20-78)	33.0 (5-50)	52.5 (25-82)	<0.000
Duration (y) (range)	5.0 (0-60)	20.0 (0-71)	22.0 (0-73)	10.0 (0-61)	4.0 (0-63)	5.0 (0-41)	<0.000
FEV ₁ % pred (%) (SD)	108.2 (12.6)	85.7 (12.5)	53.9 (12.6)	48.3 (10.0)	98.3 (13.0)	76.5 (8.8)	<0.000
FEV ₁ /FVC (%) (SD)	75.9 (7.8)	79.3 (8.1)	58.4 (9.1)	52.4 (9.1)	80.7 (7.6)	68.1 (8.3)	<0.000
Serum IgE (log)	2.02	2.57	2.43	2.26	2.20	2.31	<0.000
Atopy (%)	52.3	95.3	86.6	61.1	74.6	56.9	<0.000
Ever smoker (%)	23.2	34.7	42	40.8	40	40.4	<0.000
‡ pack-year: 0	119 (81.0)	111 (82.8)	69 (77.5)	65 (71.4)	78 (72.9)	119 (70.0)	
(%) : 0-10	11 (7.5)	16 (11.9)	13 (14.6)	7 (7.7)	17 (15.9)	17 (10.0)	
: >10	17 (11.5)	7 (5.2)	7 (7.9)	19 (20.9)	12 (11.2)	34 (20.0)	<0.010

* χ^2 test, analysis of variance, or Kruskal-Wallis test was used where appropriate.

† Information about BMI was available for 649 patients with asthma.

‡ Information about smoking pack year was available for 738 patients with asthma.

another late-onset group characterized by the worst lung function, intermediate duration of the disease (10 years) (late-onset severe asthma). Cluster E (*n* = 130, 40.5 years) was a female-dominant group charac-

terized by middle-age onset, short disease duration, normal lung function, and lower levels of total serum IgE (middle-age onset female-dominant asthma). Cluster F (*n* = 198, 60.0 years) was the largest cluster,

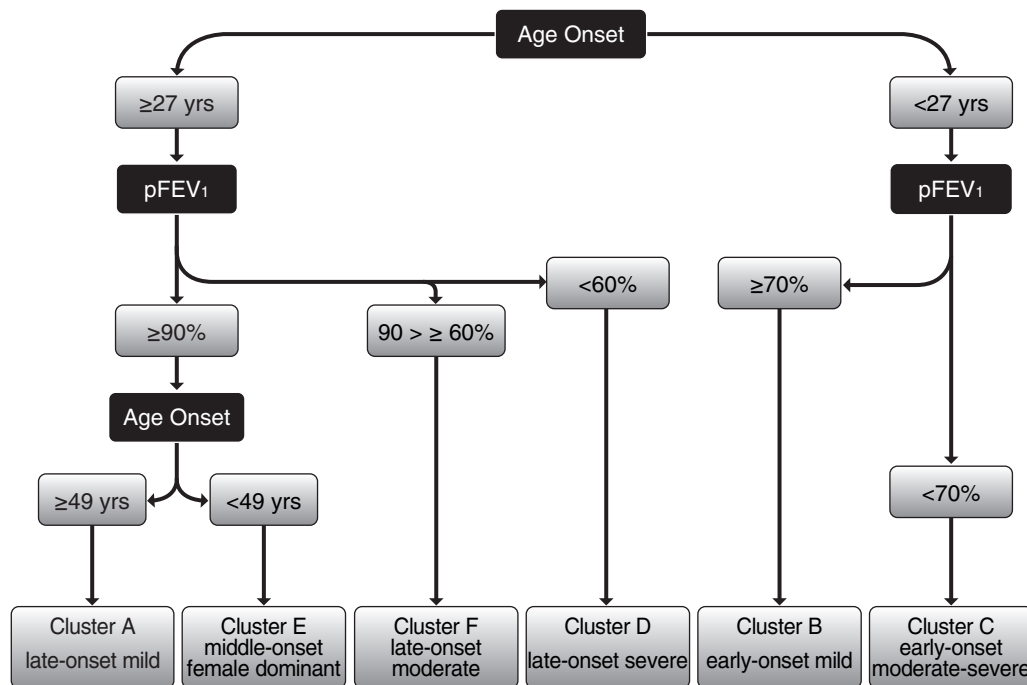


Fig. 1 Tree analysis. Using two variables (age at onset of asthma and baseline pFEV₁), patients with adult asthma can be assigned to the 6 clusters; 80.8% of patients are assigned to the correct cluster of asthma.

characterized by late onset, mildly decreased lung function, higher prevalence of smokers, and short disease duration (late-onset moderate less-atopic asthma).

A discriminant analysis using the same 8 variables revealed that the 5 strongest discriminatory variables for assignment were age at onset of asthma (F value, 219.9), %FEV₁ (206.7), FEV₁/FVC (36.4), age (33.9), and total serum IgE levels (4.2). Using these 5 variables, 95.6% of the subjects in the current samples were assigned to the appropriate cluster. We developed a CART model, shown in Figure 1. Using the variables of age at onset and %FEV₁, 80.8% of the asthmatic patients in the current study were assigned to the appropriate cluster.

The overall success rate for genotyping was 99.0%, 96.7%, 95.2%, 98.0% and 96.8% for the *CCL5* (rs2280788), *TSLP* (rs3806933 and rs2289276), *IL4* (rs2070874), and *ADRB2* (rs1042713) SNPs, respectively. In the control group ($N = 1329$), none of the SNPs deviated from the Hardy-Weinberg equilibrium (χ^2 , $p > 0.05$ for all SNPs). The -28G allele at the *CCL5* promoter polymorphism is significantly associated with late-onset asthma ($n = 532$, age at onset > 40 years) (OR = 1.38 [95%CI: 1.03, 1.85], $p = 0.031$), replicating our previous finding of the association of this allele with susceptibility to the development of adult-onset asthma.⁵ When the association of this SNP was evaluated by multinomial logistic regression analysis according to each of the 6 clusters identified in this

study, significant associations of the -28G allele were found with clusters A (OR 1.65, $p = 0.0021$), D (OR 1.74, $p = 0.011$), and B (OR 1.67, $p = 0.018$) (Table 3). In addition, the presence of the *ADRB2* Gly16 allele was inversely associated with clusters B (OR 0.45, $p = 0.00040$) and D (OR 0.63, $p = 0.034$). In contrast, neither the *TSLP* nor the *IL4* SNPs was associated with any of the asthma clusters, although *TSLP* rs2289276 was associated with the presence of asthma in the whole study population (OR = 1.25, $p = 0.019$).

DISCUSSION

The median age of the subjects in the current study was 54.0 years, making this study distinct from previous cluster analyses that studied much younger asthma populations. Involving adult asthma patients who developed the disease later in life allowed us to cover larger ranges of asthma phenotypes. Although the substantial differences in baseline characteristics and methodology make it hard to compare the phenotypes produced directly, it is interesting to compare our study with the 3 other studies that also used clustering approaches to characterize adult asthma (summarized in Table 4). In those previous studies, there were no clear clusters that corresponded to the clusters A, D and F identified in the current study, which were characterized by patients who developed asthma at an older age. In contrast, it is remarkable that the clusters described in the previous studies showed significant overlaps with clusters B and C,

Table 3A Association of the *CCL5* promoter polymorphism with asthma according to clusters

				OR [†] (95% CI)	Adjusted OR (95% CI)	p value*
-28C/G Genotype	CC	CG	GG			
Control, <i>N</i> = 1329	984	310	35	reference		
Late onset clusters						
Cluster A; late-onset mild less-atopic, <i>n</i> = 151	99	47	5	1.50 (1.05-2.14)	1.65 (1.20-2.28)	0.0021
Cluster D; late-onset severe, <i>n</i> = 106	71	30	5	1.41 (0.92-2.15)	1.74 (1.13-2.66)	0.011
Cluster F; late-onset moderate less atopic, <i>n</i> = 193	146	43	4	0.92 (0.65-1.30)	1.06 (0.74-1.53)	0.75
Early onset clusters						
Cluster B; early-onset mild atopic, <i>n</i> = 165	116	45	4	1.21 (0.84-1.72)	1.67 (1.09-2.55)	0.018
Cluster C; early-onset moderate-to-severe atopic, <i>n</i> = 116	95	19	2	0.63 (0.39-1.03)	0.68 (0.45-1.05)	0.084
Middle-age onset cluster						
Cluster E; middle-age onset female-dominant, <i>n</i> = 126	98	26	2	0.82 (0.53-1.26)	0.99 (0.72-1.37)	0.97

Definition of abbreviations: CI, confidence interval; OR, odds ratio.

*p value was calculated for adjusted OR.

[†]Odds ratios (95% CIs) were calculated for the presence of the -28G allele using a multinomial logistic regression model.

In addition to the polymorphism, sex, age, smoking status and log [total serum IgE] were adjusted in the multinomial logistic regression model.

Some patients with asthma could not be genotyped for technical reasons. The overall success rate for the SNP was 99.0%.

Table 3B Association of the *ADRB2* Arg16Gly polymorphism with asthma according to clusters

				OR [†] (95% CI)	Adjusted OR (95% CI)	p value*
<i>ADRB2</i> Arg16Gly Genotype	ArgArg	ArgGly	GlyGly			
Control, <i>N</i> = 1329	287	677	322	reference		
Late onset clusters						
Cluster A; late-onset mild less-atopic, <i>n</i> = 151	32	83	35	1.03 (0.68-1.54)	1.06 (0.74-1.52)	0.74
Cluster D; late-onset severe, <i>n</i> = 106	33	48	27	0.64 (0.41-0.98)	0.63 (0.41-0.97)	0.034
Cluster F; late-onset moderate less atopic, <i>n</i> = 193	48	92	54	0.89 (0.62-1.26)	0.85 (0.59-1.23)	0.40
Early onset clusters						
Cluster B; early-onset mild atopic, <i>n</i> = 165	54	71	42	0.59 (0.42-0.84)	0.45 (0.29-0.70)	0.00040
Cluster C; early-onset moderate-to-severe atopic, <i>n</i> = 116	25	67	26	1.15 (0.71-1.85)	0.93 (0.61-1.43)	0.76
Middle-age onset cluster						
Cluster E; middle-age onset female-dominant, <i>n</i> = 126	28	69	28	0.98 (0.63-1.53)	0.82 (0.59-1.14)	0.24

Definition of abbreviations: CI, confidence interval; OR, odds ratio.

*p value was calculated for adjusted OR.

[†]Odds ratios (95% CIs) were calculated for the presence of the Gly allele using a multinomial logistic regression model.

In addition to the polymorphism, sex, age, smoking status and log [total serum IgE] were adjusted in the multinomial logistic regression model.

Some participants could not be genotyped for technical reasons. The overall success ratio for the SNP was 96.8%.

which were characterized by individuals who developed the disease in early life (early-onset persistent atopic asthma). Our study is also clearly characterized by inclusion of many smokers, reflecting the complexity and diversity of asthma in the actual clinical setting. Cluster D, characterized by aged smokers with severe airflow obstruction and a higher than 60% prevalence of atopy, may demonstrate "overlap" features of asthma and COPD. Conversely, cluster F was

characterized by late-onset asthma with rather preserved lung function. Because patients in clusters D and F had similar intensities of smoking exposure, patients in cluster F may have been resistant to smoking effects. Although our study population showed no difference in BMI according to the clusters, cluster E, which was characterized by middle-age onset, less atopy, and female dominance, might share some pathobiological features with a group of nonatopic,

Table 4 Summary of previous studies

First author (ref.)		Our study	Moore (1)		Haldar (2)	Siroux (3)
Severity		persistent asthma	severe asthma	mild asthma	refractory asthma	mild asthma [†]
Subjects (N)		880	728 [‡]	184	187	1895 [§] /641 [¶]
Smoking (%)		36.2	-	-	-	-
BMI (mean)		23.4	29	27.5	28.5	-
Age at onset	Lung function					
Young onset (<16 y)	mild	cluster B	cluster 1	cluster 1		phenotype C/G
	moderate		cluster 2			
	severe	cluster C	cluster 4		cluster 1/cluster 3	phenotype A/E
Middle-age onset	mild	cluster E		cluster 3		phenotype D/H
	severe		cluster 5		cluster 4	phenotype B/F
Older onset (40< y)	mild	cluster A				
	moderate	cluster F				
	severe	cluster D				
Middle-age onset, obese, female-dominant			cluster 3 (age at onset, 42 y)	cluster 2 (age at onset, 35.3 y)	cluster 2 (age at onset, 15.4 y)	

[†] Subjects who had a wheeze experience were recruited.

[‡] Severe Asthma Research Program (SARP)

[§] European Community and Respiratory Health Survey II (ECRHS II)

[¶] Epidemiological Study on the Genetics and Environment of asthma 2 (EGEA2)

primarily obese women that has been consistently identified in previous studies.^{1,2}

In the current study, the *CCL5* -28G allele was associated with 3 distinct asthma phenotypes (cluster A, B and D), whereas the *ADRB2* Arg16Gly genotype influenced the development of 2 distinct asthma phenotypes (clusters B and D). Our previous study identified a significant association of the *CCL5* -28C>G SNP with asthma only in individuals who developed the disease after the age of 40 years⁵ and with non-emphysematous COPD.⁶ However, genetic associations between early-onset atopic asthma and the *CCL5* -28C>G genotype have also been reported in Asian populations,^{18,19} possibly reflecting the genetic heterogeneity among different ethnicities and different asthma severities. In the *ADRB2* gene, Arg16Gly and Glu27Gln are the two most extensively studied polymorphisms for asthma-related phenotypes. Neither polymorphic site has been associated with asthma per se,^{20,21} but among asthmatic subjects, the Gly16 allele has been associated with corticosteroid use,²⁰ nocturnal asthma²² and severe asthma.²³ Subjects who were homozygous for Arg16 were reported to have lower lung function.²¹ In the British 1958 birth cohort study, among the members (*n* = 8018) who contributed DNA samples at the 45-year follow-up, the common polymorphisms Arg16Gly and Gln27Glu were significantly associated with persistence of asthmatic symptoms from childhood to middle age. Collectively, both the *CCL5* -28C>G and the *ADRB2* Arg16Gly genotypes could play important roles in the expression of the distinct characteristics of the

asthma clusters.

In this study, the same allele at *CCL5* was associated with distinct clusters (clusters A, B and D). The phenotypic differences between these clusters could be caused by additional independent genetic association including the *ADRB2* Arg16Gly genotype. Environmental influences also play a role. Among clusters A, B, and D, cluster B was characterized by a highly atopic condition that is the strongest risk factor for asthma. In contrast, cluster D was characterized by higher frequency of ever-smokers with relatively high pack-years. Environmental exposure to tobacco smoke is also known to be an independent risk factor for asthma other than respiratory infection and atopy.²⁴ Therefore, given that allergic diathesis or smoking could have an enormous influence on the genetic determinants of asthma, the presence or absence of an interaction of the *CCL5* -28G allele with the *ADRB2* genotype, allergic diathesis, or smoking could have had a significant impact on shaping these phenotypic characteristics in clusters A, B, or cluster D. Conversely, cluster F, which was characterized by late-onset moderate asthma, was associated with neither the *CCL5* -28G allele nor the *ADRB2* Gly16 allele. This indicates that late-onset asthma may itself comprise more than 1 clinical phenotype, which encompasses different disease variants with different genetic components.

It is generally believed that patients with longstanding asthma have more severe airflow limitation than do those with late-onset asthma. Indeed, in our study, when all patients were analyzed together, linear re-

gression analysis adjusted for sex, age, and smoking showed a strong association between duration of the disease and severity of airflow obstruction (β , -0.148; t , -4.804; $p < 0.0001$). However, our study clearly identified both a group of asthma patients with severely impaired lung function but rather short duration of the disease (cluster D) and a group of asthma patients with normal lung function and long term duration of the disease (cluster B). Although the asthma in some of the patients in cluster D could have started as mild childhood asthma that was never appropriately diagnosed or was clinically silent, we believe that our finding suggests that analyzing risk factors for impaired lung function in asthmatic patients by using the overall asthma population may be misleading and thus emphasizes the importance of careful asthma sub-typing in epidemiology studies, which will have important implications for understanding the pathogenesis of asthma.

Well-established clinical and physiological phenotypes relevant to asthma include those defined by the presence of atopy, the presence of airflow obstruction, and the age of asthma onset.⁴ The results of recent studies on defining asthma phenotypes by use of exploratory methods such as a clustering approach are consistent with observations from clinical practice, showing the importance of age at onset, airflow obstruction, and allergic diathesis in phenotyping asthma.¹⁻³ Therefore, the decisions made here about which phenotypic variables to include were made based on clinical judgment of relevance. Rather than using *a priori* cutoff values for age at onset, baseline lung function, and levels of total serum IgE, we applied cluster analysis, and thus the identified clusters conform to shared phenotypical features. However, we used a very limited range of variables to describe asthma heterogeneity, and we recognize that the choice of variables could have affected the outcome and thus is not immune to *a priori* assumptions. It is, therefore, more appropriate to consider our selection approach as one that is hypothesis-driven rather than unbiased. Because it remains to be seen whether one of these approaches, an unbiased approach such as cluster analysis or a biased approach relying on clinical characteristics, is superior in identifying subtypes or whether these approaches are separately informative and complementary, an attempt to validate our findings of specific subtypes using these or similar methods in other well-phenotyped asthma cohorts should be performed.

Obesity has a significant impact on asthma risk, severity, and control.²⁵ Moore *et al.* reported that an unsupervised hierarchical cluster analysis identified a group of nonatopic women who were primarily obese as a distinct phenotype of asthma.¹ Thirty-seven percent of the total cohort of their study had a BMI >30. In contrast, only 5.7% of participants in the current study had a BMI >30. In addition, data from a nation-

wide population-based cross-sectional survey of adult asthma prevalence in Japan ($N = 22,962$; age range, 20-79 years) showed that the prevalence of obesity (BMI ≥ 30.0) was very low (men, 3.0%; women, 2.3%).²⁶ Therefore, we believe that, at least in the Japanese population, the asthma phenotype related to obese women does not constitute a large proportion of the disease heterogeneity of asthma.

In our study, to judge the age at onset of asthma as accurately as possible, a detailed history of each patient's episodes of dyspnea, wheezing, or cough during childhood and puberty was taken. In case of uncertainty, the earliest respiratory symptoms were taken into account and the age at onset of asthma symptoms was used for the calculation of the duration of asthma. A study that examined age at onset of asthma in 4335 Japanese adult asthmatics found that 26.3% of those individuals developed the disease before the age of 20 years, whereas 73.8% developed it at the age of 20 years or older.²⁷ This variable age at onset of adult asthma is very similar to those of our study, which may validate the data on age at onset of asthma based on self-reporting in our study. In addition, a previous analysis revealed that adults reported important childhood events with high consistency regardless of symptom status.²⁸ However, it should be acknowledged that it is virtually impossible to distinguish with any certainty recrudescence from incident asthma cases in an adult population. Even though a number of asthma patients whom we designated as having late-onset asthma might have developed the disease in childhood, it seems unlikely that the particular genotypes examined in the current study contributed to this recall bias. We therefore believe that the observed difference in the frequency of the *CCL5* -28G allele or the *ADRB2* Gly16 allele among the subgroups of asthma cannot be attributed to recall bias.

In conclusion, our clinic-based cluster analysis identified meaningful adult asthma phenotypes linked to the functional *CCL5* -28C>G and/or *ADRB2* Arg16 Gly SNP. Our analysis is best seen as hypothesis-generating rather than hypothesis-solving, and an attempt to validate these findings of specific subtypes using these or similar methods in other well-phenotyped asthma cohorts should be performed. The classification of patients with asthma by phenotype will facilitate future research to identify biomarkers for disease phenotypes and to test novel therapeutic targets and phenotype-specific treatments.

ACKNOWLEDGEMENTS

We thank all the physicians who participated in recruiting the study subjects and collecting data. Mrs. Flaminia Miyamasu, an associate professor at Faculty of Medicine, University of Tsukuba, proofread and commented on this paper. Mrs. Yuka Yoshikawa and Mrs. Takako Nakamura gave technical assistance on DNA extraction from clinical samples.

This study was partly supported by a Grant-in-Aid for Scientific Research (B), No. 21390254, from the Japan Society for the Promotion of Science.

REFERENCES

- Moore WC, Meyers DA, Wenzel SE *et al.* Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010;**181**:315-23.
- Haldar P, Pavord ID, Shaw DE *et al.* Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008;**178**:218-24.
- Siroux V, Basagaña X, Boudier A *et al.* Identifying adult asthma phenotypes using a clustering approach. *Eur Respir J* 2011;**38**:310-7.
- Wenzel SE. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006;**368**:804-13.
- Hizawa N, Yamaguchi E, Konno S, Tanino Y, Jinushi E, Nishimura M. A functional polymorphism in the RANTES gene promoter is associated with the development of late-onset asthma. *Am J Respir Crit Care Med* 2002;**166**:686-90.
- Hizawa N, Makita H, Nasuhara Y *et al.* Functional single nucleotide polymorphisms of the CCL5 gene and nonemphysematous phenotype in COPD patients. *Eur Respir J* 2008;**32**:372-8.
- Diette GB, Krishnan JA, Dominici F *et al.* Asthma in older patients: factors associated with hospitalization. *Arch Intern Med* 2002;**162**:1123-32.
- Hizawa N, Yamaguchi E, Takahashi D, Nishihira J, Nishimura M. Functional polymorphisms in the promoter region of macrophage migration inhibitory factor and atopy. *Am J Respir Crit Care Med* 2004;**169**:1014-8.
- Price D, Musgrave SD, Shepstone L *et al.* Leukotriene antagonists as first-line or add-on asthma-controller therapy. *N Engl J Med* 2011;**364**:1695-707.
- Spears M, Cameron E, Chaudhuri R, Thomson NC. Challenges of treating asthma in people who smoke. *Expert Rev Clin Immunol* 2010;**6**:257-68.
- Masuko H, Sakamoto T, Kaneko Y *et al.* Lower FEV1 in non-COPD, nonasthmatic subjects: association with smoking, annual decline in FEV1, total IgE levels, and TSLP genotypes. *Int J Chron Obstruct Pulmon Dis* 2011;**6**:181-9.
- Hirota T, Takahashi A, Kubo M *et al.* Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* 2011;**43**:893-6.
- Harada M, Hirota T, Jodo AI *et al.* Thymic stromal lymphopoietin gene promoter polymorphisms are associated with susceptibility to bronchial asthma. *Am J Respir Cell Mol Biol* 2011;**44**:787-93.
- Suzuki I, Hizawa N, Yamaguchi E, Kawakami Y. Association between a C+33T polymorphism in the IL-4 promoter region and total serum IgE levels. *Clin Exp Allergy* 2000;**30**:1746-9.
- Marsh DG, Neely JD, Breazeale DR *et al.* Linkage analysis of IL4 and other chromosome 5q31.1 markers and total serum immunoglobulin E concentrations. *Science* 1994;**264**:1152-6.
- Ball GH, Hall DJ. A clustering technique for summarizing multivariate data. *Behav Sci* 1967;**12**:153-5.
- Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. In: *Boca Raton*. Chapman and Hall/CRC, 1984.
- Li X, Zhang Y, Zhang J *et al.* Asthma susceptible genes in Chinese population: a meta-analysis. *Respir Res* 2010;**11**:129.
- Sohn MH, Kim SH, Kim KW, Jee HM, Park HS, Kim KE. RANTES gene promoter polymorphisms are associated with bronchial hyperresponsiveness in Korean children with asthma. *Lung* 2008;**186**:37-43.
- Reihnsaus E, Innis M, MacIntyre N, Liggett SB. Mutations in the gene encoding for the β_2 adrenergic receptor in normal and asthmatic subjects. *Am J Respir Cell Mol Biol* 1993;**8**:334-9.
- Summerhill E, Leavitt SA, Gidley H, Parry R, Solway J, Ober C. Beta(2)-adrenergic receptor Arg16/Arg16 genotype is associated with reduced lung function, but not with asthma, in the Hutterites. *Am J Respir Crit Care Med* 2000;**162**:599-602.
- Turki J, Pak J, Green SA, Martin RJ, Liggett SB. Genetic polymorphisms of the β_2 -adrenergic receptor in nocturnal and nonnocturnal asthma: evidence that Gly16 correlates with the nocturnal phenotype. *J Clin Invest* 1995;**95**:1635-41.
- Holloway JW, Dunbar PR, Riley GA *et al.* Association of β_2 -adrenergic receptor polymorphisms with severe asthma. *Clin Exp Allergy* 2000;**30**:1097-103.
- Tamimi A, Serdarevic D, Hanania NA. The effects of cigarette smoke on airway inflammation in asthma and COPD: therapeutic implications. *Respir Med* 2012;**106**:319-28.
- Stream AR, Sutherland ER. Obesity and asthma disease phenotypes. *Curr Opin Allergy Clin Immunol* 2012;**12**:76-81.
- Fukutomi Y, Taniguchi M, Nakamura H *et al.* Association between body mass index and asthma among Japanese adults: risk within the normal weight range. *Int Arch Allergy Immunol* 2012;**157**:281-7.
- Akiyama K, Takahashi K, Yanagawa H. Nation-wide survey of asthma patients in Japan [abstract]. *Allergy Clin Immunol Int* 1997;**9**:87.
- Svanes C, Sunyer S, Zock J *et al.* Long-term reliability in reporting of childhood pets by adults interviewed twice, nine years apart. Results from the European Community Respiratory Health Survey I and II. *Indoor Air* 2008;**18**:84-92.