SMOKING

Association of *CYP2A6* deletion polymorphism with smoking habit and development of pulmonary emphysema

N Minematsu, H Nakamura, M Iwata, H Tateno, T Nakajima, S Takahashi, S Fujishima, K Yamaguchi

Thorax 2003;**58**:623–628

See end of article for authors' affiliations

Correspondence to: Dr H Nakamura, Cardiopulmonary Division, Department of Medicine, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan; hnakamura@cpnet.med.keio.ac.jp

Revised version received 20 March 2003
Accepted for publication 2 April 2003

Background: Nicotine is responsible for smoking dependence and is mainly metabolised by CYP2A6. Several types of genetic polymorphism of *CYP2A6* have been reported, but their relation to smoking habit and chronic obstructive pulmonary disease (COPD) phenotypes has not been fully clarified.
Methods: 203 current or ex-smokers (lifelong cigarette consumption (CC) ⩾10 pack years) with subclinical and established COPD phenotypes were clinically evaluated and pulmonary function tests and a chest CT scan were performed (smoker group). The non-smoker group consisted of 123 healthy volunteers. *CYP2A6* genotypes were determined in both groups.
Results: The percentage of subjects with a *CYP2A6del* allele (genotype D) was lower in heavy smokers (20.5%, n=88, CC ⩾60 pack years) than in light smokers (37.4%, n=115, CC 10–59 pack years, χ^2=6.8, p=0.01) or non-smokers (36.1%, n=122, χ^2=6.0, p=0.01); lower in ex-smokers (20.7%, n=111) than in current smokers (41.3%, n=92, χ^2=10.1, p<0.01); and lower in smokers with a high LAA (low attenuation area) score on the chest CT scan (18.4%, n=76, LAA ⩾8.0) than in those with a low LAA score (37.0%, n=127, LAA <8.0, χ^2=7.8, p<0.01).
Conclusions: Subjects with the *CYP2A6del* allele tend not to be heavy habitual smokers but can be light habitual smokers. The *CYP2A6del* polymorphism may inhibit smokers from giving up smoking, but appears to function as a protective factor against the development of pulmonary emphysema independent of smoking habit.

Cigarette smoking is a primary risk factor for chronic obstructive pulmonary disease (COPD) and lung cancer, and cessation of smoking to reduce the risk of these diseases is a major medical concern. Since smoking dependence is significantly associated with the serum concentration of nicotine, it is important to understand nicotine metabolism in order to solve smoking related problems. Racial differences in serum levels of nicotine metabolites have recently been reported, suggesting the presence of genetic variance in enzymes related to nicotine metabolism.[1 2] Nicotine is mainly metabolised by CYP2A6, a member of cytochrome P450. At least three types of genetic polymorphism of *CYP2A6* have been reported: *CYP2A6*2* T to A transition in exon 3,[3 4] *CYP2A6*3* gene conversions in exons 3, 6, and 8 between *CYP2A6* and *CYP2A7*,[4 5] and *CYP2A6del*, a whole gene deletion.[6] The enzymatic activity of *CYP2A6* has been found to be impaired in subjects with *CYP2A6*2* or *CYP2A6del* allele,[3 6] and a relationship between the genotypes and smoking habit has recently been proposed.[7 8] These studies suggested the possibility that the presence of the mutant alleles could reduce the consumption of cigarettes. However, the detailed relationship between *CYP2A6* polymorphism and smoking habit—including lifelong cigarette consumption (CC), daily CC, and smoking duration—has not been reliably established because of problems with the specificity of the genotyping method[5 7] and the low frequency of defective alleles[8] reported in these studies.

COPD develops primarily in heavy smokers, suggesting that smoking behaviour regulated by the serum nicotine level may be associated with the development of COPD. In addition, it is difficult to exclude the possibility that substances metabolised or activated by CYP2A6, including nicotine and procarcinogens, may be related to the pathogenesis of COPD. However, no authentic studies have shown an association between *CYP2A6* genotypes and COPD manifestations. A study was undertaken to clarify the effects of the *CYP2A6* genotype on smoking habit in a Japanese population in which the frequency of the *CYP2A6del* allele is expected to be high,[9] and to elucidate the relationship between the *CYP2A6* genotypes and COPD phenotypes.

METHODS

Participants

Genomic DNA was isolated from peripheral blood obtained from Japanese subjects with a significant smoking history who visited the outpatient clinic of Keio University Hospital between 1998 and 2002 for the diagnosis or treatment of COPD (smoker group, n=203, 189 men). The inclusion criteria for the smoker group were lifelong CC ⩾10 pack years and age ⩾50 years. Two hundred and fifteen patients from about 1500 outpatients with respiratory diseases thought to be COPD met the inclusion criteria. Twelve subjects were excluded because of accompanying giant bullae (n=2), pulmonary fibrosis (n=3), diffuse bronchiectasis (n=1), bronchial asthma (n=4), and lung cancer compromising pulmonary function (n=2), leaving 203 subjects who were judged eligible for further analysis. The subjects in this group were either current smokers (n=92) or ex-smokers (n=111), defined as those who had stopped smoking for at least 6 months. All subjects were directly interviewed and the number of cigarettes consumed per day, the duration of smoking, and the time elapsed since quitting smoking were recorded as accurately as possible. Close attention was paid to the change in the daily number of cigarettes smoked. CC was calculated from mean daily consumption of cigarettes (packs/day) and duration of smoking (years) and used for further analysis.

Table 1 *CYP2A6* genotype and smoking habit

Group	Genotype	n	Age	CC	Packs/day	Duration
All smokers	W	142	67 (9)	65 (34)	1.61 (0.75)	41 (10)
(n=203)	D	61	65 (10)	52 (30)*	1.36 (0.63)*	39 (11)
Current smokers	W	54	64 (9)	65 (28)	1.62 (0.66)	41 (10)
(n=92)	D	38	64 (8)	53 (26)*	1.31 (0.59)*	41 (9)
Ex-smokers	W	88	69 (9)	65 (37)	1.61 (0.80)	41 (10)
(n=111)	D	23	68 (13)	51 (35)	1.39 (0.71)	36 (14)

Values are mean (SD).
CC=lifelong cigarette consumption (pack years); duration=duration of smoking period (years);
W=homozygous wild type; D=heterozygote or homozygous mutant for *CYP2A6del* allele.
*p<0.05 *v* W.

Genomic DNA was also obtained from Japanese non-smoking healthy volunteers with few respiratory symptoms screened by a written questionnaire and few abnormalities on their chest radiograph (non-smoker group, n=123, 109 men). The inclusion criteria for the non-smoker group were lifelong CC <100 cigarettes and age ⩾50 years. The smoking history of the subjects in this group was self-reported. Of 132 volunteers who met the inclusion criteria, nine were excluded because of respiratory symptoms (n=5) and chest radiographic abnormalities (n=4). This group included subjects who had no chance to smoke and who had tried smoking but did not continue. No subjects who smoked a pipe or cigars were included in either the smoker or non-smoker group.

Informed consent was obtained from each subject and the study protocol was approved by the ethical committee of Keio University Hospital.

Assessment of pulmonary function and emphysematous changes on chest CT scan

Vital capacity (VC), forced vital capacity (FVC), and forced expiratory volume in 1 second ($FEV_{1.0}$) were measured in all subjects in the smoker group using an electronic spirometer (MFR-8200; Nihon Koden, Tokyo, Japan). Carbon monoxide transfer factor (T_{LCO}) was estimated by 10 second breath holding in 170 subjects in the smoker group (Chestac-55V; Chest, Tokyo, Japan). The lung volume at the time of T_{LCO} measurement (V_A) was simultaneously determined and used for calculating T_{LCO}/V_A, the carbon monoxide transfer coefficient (K_{CO}). All pulmonary function tests were performed at our pulmonary function laboratory by three expert technicians in a manner consistent with the criteria recommended by the American Thoracic Society.[10] The functional impairment in smokers was assessed on the basis of either %K_{CO} or %$FEV_{1.0}$. The reference values of these pulmonary function parameters were obtained from data reported for healthy Japanese subjects.[11 12]

A chest CT scan was performed in all subjects in the smoker group (Proseed, GE Yokogawa Medical Systems, Tokyo, Japan) under the following conditions: 120 kVp, 200 mA, 1 second scan time, and 5 mm collimation. Low attenuation area (LAA) was visually assessed by the method of Goddard *et al*.[13] Briefly, the whole lung was divided into six zones (left and right zones in the upper, middle, and lower lung fields). Low attenuation areas in each image section were scored on a scale from 0 to 4 where 0=no LAA, 1=1–25%, 2=26–50%, 3=51–75%, and 4=76–100%. The total (0–24) was defined as the LAA score. Evaluation of the LAA score was made by three pulmonologists in a blinded manner, and the mean score was used as a quantitative indicator of emphysematous change. The standard deviation (SD) in LAA scores for each subject determined by three examiners averaged 1.0 point, and we did not rectify the scatter in LAA scores between the examiners.

Genotyping of *CYP2A6*

The *CYP2A6del* allele was detected by two different methods: (1) the two step polymerase chain reaction (PCR) method reported by Oscarson *et al*[14] and (2) the restriction fragment length polymorphism (RFLP) method described by Nunoya *et al*.[15] (3) *CYP2A6*2* and *CYP2A6*3* were identified by the RFLP method reported by Chen *et al*.[5] Oligonucleotide primer sequences and restriction enzymes used for these analyses were as follows: (1) 2A6ex7F: 5′-GGCCAACA TGCCCTACATG-3′; 2A6ex8F: 5′-CACTTCCTGAATGAG-3′; 2A7ex8F: 5′-CATTT CCTGGATGAC-3′; 2A6R1: 5′-GCACTTATGTTTTGTGAGACAT CAGAGACAA-3′; 2A6R2: 5′-AAAATGGGCATGAACGCC-3′, (2) 2A6-B6: 5′-CGATGGAAAAGGGCACAAAAGCA-3′; 2A6-8S: 5′-CACCGAAGTGTACCCTATGCTG-3′; 2A7-B1: 5′-CACCGAA GTGTTCCCTATGCTG-3′; *Bst* NI, (3) CYP2A6F03: 5′-CTG ATCGACTA GGCGTGGTA-3′; CYP2A6R06: 5′-CGTCCTGGGT GTTTTCCTTC-3′; *Xcm* I and *Dde* I.

Data analysis

Values are presented as mean (SD). Genotype frequencies were compared between groups using the χ^2 test. Mean values of age, CC, packs/day, smoking duration, LAA score, %K_{CO}, and %$FEV_{1.0}$ were compared between groups using the unpaired *t* test. Logistic regression analysis was performed to examine the effects of *CYP2A6del* genotype on CC, cessation of smoking, the severity of emphysematous changes, K_{CO}, and airflow obstruction independent of other factors including age. A p value of <0.05 was considered significant.

RESULTS

*CYP2A6*3* and *CYP2A6del* allele frequencies in non-smokers and smokers

Three *CYP2A6* alleles (*CYP2A6*1* (wild type allele), *CYP2A6*3*, and *CYP2A6del*) were identified in the Japanese subjects used in the study; no subject had *CYP2A6*2*. Based on these genetic results, smoking and non-smoking subjects were categorised into four subgroups, **1/*1* (n=220), **1/del* (n=95), *del/del* (n=10), and **1/*3* (n=1). The allele frequencies of *CYP2A6*1*, *CYP2A6*3*, and *CYP2A6del* were 0.821, 0.002, and 0.176, respectively, which are qualitatively consistent with the values previously reported in Japanese or Chinese populations.[5 9 14 15] Since only one non-smoking subject had the **1/*3* genotype, this subject was excluded from further analysis and **1/*1*, **1/del*, and *del/del* genotypes were designated as the wild type, heterozygote, and homozygous mutant, respectively; subjects with the **1/*1* genotype were defined as the W (wild type) group and those with the **1/del* or *del/del* genotypes were defined as the D (deletion) group.

CYP2A6 genotype and smoking habit

Differences in smoking habit between subjects with the *CYP2A6del* genotypes are shown in table 1. CC and number of cigarettes/day (packs/day) in the D group were significantly lower than in the W group in all smokers and in current smokers. Although the same trend was also observed for ex-smokers, the difference did not reach statistical significance. Duration of smoking did not differ between the genotypes in the groups.

Table 2 Smoking habit and CYP2A6 genotype frequencies

	All smokers			Current smokers			Ex-smokers			Non-smokers		
	n	Age	%D	n	Age	%D	n	Age	%D	n	Age	%D
Total	203	66 (9)	30.0	92	64 (8)	41.3	111	69 (10)	20.7††	122	56 (3)	36.1
CC (pack years)												
10–59	115	65 (10)	37.4	52	63 (9)	50.0	63	67 (11)	27.0			
60+	88	68 (9)	20.5**†	40	65 (8)	30.0	48	70 (8)	12.5††			
Packs/day												
0.5–1.49	99	68 (9)	38.4	44	65 (8)	56.8†	55	70 (10)	23.6			
1.5+	104	65 (9)	22.1*†	48	62 (8)	27.1**	56	67 (10)	17.9†			
Duration (years)												
10–39	80	59 (8)	33.8	39	57 (6)	38.5	41	61 (10)	29.3			
40+	123	71 (10)**	27.6	53	69 (6)**	43.4	70	73 (7)**	15.7††			

CC=lifelong cigarette consumption; duration=duration of smoking period; %D=percentage of subjects with genotype D.
*p<0.05, **p<0.01*v* the rest of the subjects in each group.
†p<0.05, ††p<0.01 *v* non-smokers.
Non-smokers were younger than the subgroups of smokers except those in the 10–39 years duration group of current smokers.

Table 3 Contribution of D genotype to smoking habit in logistic regression analysis

		All smokers (n=203)				Current smokers (n=92)				Ex-smokers (n=111)			
Definition	Factor	χ^2	p value	OR	95% CI	χ^2	p value	OR	95% CI	χ^2	p value	OR	95% CI
CC ⩾60	Genotype D	6.21	0.01*	0.44	0.23 to 0.84	3.71	0.05	0.42	0.18 to 1.02	3.37	0.07	0.38	0.13 to 1.07
	Age (10 years)	2.97	0.08	1.31	0.96 to 1.77	1.17	0.28	1.32	0.80 to 2.19	2.61	0.11	1.40	0.93 to 2.12
Packs/day ⩾1.5	Genotype D	7.16	<0.01**	0.42	0.23 to 0.80	8.26	<0.01**	0.27	0.11 to 0.66	0.62	0.43	0.69	0.27 to 1.75
	Age (10 years)	5.45	0.02*	0.69	0.51 to 0.94	3.22	0.07	0.62	0.36 to 1.05	1.93	0.17	0.76	0.51 to 1.12
Duration ⩾40	Genotype D	0.09	0.77	0.89	0.41 to 1.93	0.03	0.85	1.11	0.36 to 3.48	2.84	0.09	0.34	0.09 to 1.20
	Age (10 years)	51.2	<0.01**	6.85	4.04 to 11.61	23.2	<0.01**	15.4	5.06 to 46.9	26.1	<0.01**	5.52	2.87 to 10.6

CC=lifelong cigarette consumption (pack years); duration=duration of smoking period (years); OR=odds ratio; CI=confidence interval.
*p<0.05, **p<0.01.
OR and 95% CI correspond to having genotype D or aging every 10 years.

The D group consisted of 53 subjects with the heterozygote and eight with the homozygous mutant in all smokers. When the W group and the heterozygote and homozygous mutant groups in all smokers were compared, CC was lower in the homozygous mutant group (40 (10) pack years) than in the W group (65 (34) pack years, p<0.01, Kruskal-Wallis rank test and Games-Howell test) or the heterozygote group (54 (31) pack years, p<0.05). The homozygous mutant group also smoked fewer packs/day (0.96 (0.11)) than the W (1.61 (0.75), p<0.01) or the heterozygote groups (1.42 (0.66), p<0.01).

CYP2A6del genotype frequency in subjects with different smoking habits

The relationship between lifelong CC, number of cigarettes/day, duration of smoking and *CYP2A6del* genotypes is shown in table 2. To elucidate the difference in *CYP2A6del* genotype frequency between the subjects with different smoking habits, we selected thresholds for these smoking related parameters around the mean values obtained from all smokers studied. Since there are no authentic criteria to define the thresholds for these parameters, the analysis was made by setting three cut off points for each parameter (CC: 40, 60, 80 pack years; packs/day: 1.25, 1.5, 1.75; duration: 35, 40, 45 years). There was no difference in the percentage of subjects with genotype D (%D) between non-smokers and all smokers. %D in heavy smokers with CC ⩾60 or ⩾80 pack years (17.3%) was significantly lower than in light smokers with CC 10–59 or 10–79 pack years (34.3%, p<0.05), respectively. However, the difference was not statistically significant when the threshold for CC was set at 40 pack years, although the tendency was similar to that for a cut off point of 60 or 80 pack years. When the smokers were divided into current and ex-smokers, %D was much lower in current smokers with a relatively heavy smoking history of >80 pack years (16.7%) than in current smokers with a light smoking history (CC 10–79 pack years, 50.0%, p<0.01). In ex-smokers %D was lower in subjects with a relatively heavy smoking history (CC ⩾40 pack years, 16.0%) than in light smokers (CC 10–39 pack years, 33.3%, p<0.05).

When all smokers were included in the analysis, %D in the smoker group with a daily CC of ⩾1.5 packs/day was lower than in those with a lower daily CC (0.5–1.49 packs/day) and than the non-smoker group. This tendency was seen more clearly in current smokers than in ex-smokers. In current

Table 4 Clinical profiles of current and ex-smokers

	n	%D	Age	CC	Packs/day	Duration	LAA	%Kco†	%$FEV_{1.0}$
Current smokers	92	41.3	64 (8)	60 (27)	1.50 (0.64)	41 (10)	4.5 (4.5)	84 (29)	76 (24)
Ex-smokers	111	20.7	69 (10)	62 (37)	1.56 (0.79)	40 (11)	7.3 (5.8)	76 (32)	66 (30)
p value		<0.01##	<0.01**	0.62	0.59	0.59	<0.01**	0.10	0.01*

Values are mean (SD).
CC=lifelong cigarette consumption (pack years); duration=duration of smoking period (years); %D=percentages of subjects with genotype D; LAA=low attenuation area score on CT scan; Kco=carbon monoxide transfer coefficient; $FEV_{1.0}$=forced expiratory volume in 1 second.
*p<0.05, **p<0.01, by *t* test.
##p<0.01, χ^2=10.1, by χ^2 test.
†n=76 current smokers and 94 ex-smokers in this analysis.

Table 5 Contribution of D genotype to cessation of smoking in logistic regression analysis

Independent factor	χ^2	p value	Odds ratio	95% CI
Genotype D	9.41	<0.01**	0.36	0.19 to 0.69
Age (10 years)	11.46	<0.01**	1.76	1.27 to 2.45
CC (10 pack years)	0.48	0.49	0.97	0.88 to 1.06
Genotype D	7.73	<0.01**	0.39	0.20 to 0.76
Age (10 years)	7.29	<0.01**	1.59	1.14 to 2.23
CC (10 pack years)	1.01	0.32	0.95	0.87 to 1.05
LAA	6.71	<0.01**	1.08	1.02 to 1.15
Genotype D	8.97	<0.01**	0.37	0.19 to 0.71
Age (10 years)	7.93	<0.01**	1.64	1.16 to 2.32
CC (10 pack years)	0.79	0.38	0.96	0.88 to 1.05
$\%FEV_{1.0}$ (10%)	1.64	0.20	0.93	0.83 to 1.04

CI=confidence interval; LAA=low attenuation area score; CC=lifelong cigarette consumption; $FEV_{1.0}$=forced expiratory volume in 1 second.
*p<0.05, **p<0.01.
Odds ratio and 95% CI correspond to having genotype D, aging every 10 years, 10 pack years increase in CC, one point increase in LAA scores, or 10% increase in $\%FEV_{1.0}$.

Table 6 Association of *CYP2A6* genotype with LAA score, %Kco, and $\%FEV_{1.0}$

	n	Age	CC	%D
LAA score				
<8.0	127	65 (9)	59 (32)	37.0
⩾8.0	76	69 (9)¶	66 (34)	18.4**
%Kco				
⩾65	115	65 (9)	59 (29)	34.8
<65	55	71 (8)‡	70 (39)‡‡	18.2†
$\%FEV_{1.0}$				
⩾50	146	65 (10)	59 (31)	32.2
<50	57	71 (8)§	69 (38)§§	24.6

LAA=low attenuation area score; Kco=carbon monoxide transfer coefficient; $FEV_{1.0}$=forced expiratory volume in 1 second; CC=lifelong cigarette consumption (pack years); %D=percentage of subjects with genotype D.
**χ^2=7.8, p=0.005 *v* LAA <8.0.
†χ^2=4.9, p=0.03 *v* %Kco ⩾65.
‡p<0.01, ‡‡<0.05 *v* %Kco ⩾65.
§p<0.01, §§p<0.05 *v* $\%FEV_{1.0}$ ⩾50.
¶p<0.01 *v* LAA <8.0.

smokers %D was significantly lower in subjects with a higher daily CC than in those with a lower daily CC for all three thresholds examined (1.25 packs/day: 32.7% *v* 54.1%, p<0.05; 1.75 packs/day: 27.3% *v* 49.2%, p<0.05). %D was much higher in current smokers with a light smoking history (0.5–1.49 packs/day) than in non-smokers.

There was no difference in %D with duration of smoking when all smokers were included, or when current smokers were analysed separately. However, %D tended to be lower in ex-smokers whose duration of smoking was relatively longer (35 years: 15.2% *v* 34.3%, p<0.05). %D in ex-smokers with a relatively heavy smoking history for both lifelong and daily CC and duration of smoking was lower than in non-smokers.

The number of subjects with the wild type, heterozygote, and homozygous mutant was 78, 42, and 2, respectively, in the non-smoker group, and 142, 53, and 8, respectively, in the smoker group. The genotype frequency distributions in these groups did not deviate from the Hardy-Weinberg equilibrium.

The age adjusted effect of %D on smoking habit determined by logistic regression analysis is shown in table 3. Genotype D was confirmed to function as an independent factor, reducing daily and lifelong CC but not the duration of smoking in all smokers or in current smokers. This tendency was not evident in ex-smokers. Aging was found to be an important factor in determining the duration of smoking in both current and ex-smokers.

Contribution of genotype D to cessation of smoking

As shown in table 4, the percentage of subjects with genotype D was significantly higher in current smokers than in ex-smokers. There were significant differences in LAA score, $\%FEV_{1.0}$, and age between the groups, while no difference was seen in %Kco, CC, packs/day, or duration of smoking.

Logistic regression analysis revealed that genotype D functioned as an independent factor inhibiting the cessation of smoking irrespective of age, CC, LAA score, or $\%FEV_{1.0}$ (table 5). Aging and augmented LAA score were also shown to promote cessation of smoking.

CYP2A6del genotype frequency and severity of emphysema, impairment of Kco, and airflow obstruction

The relationship between the *CYP2A6del* genotype and severity of emphysematous changes examined by LAA scores on CT images, impairment of Kco, and airflow obstruction is shown in table 6. The morphological severity of emphysema was assessed by dividing the smokers into two groups with LAA scores of ⩾8.0 or <8.0. Since an LAA score of 24 implies that most of the lung fields have emphysematous changes, an LAA score of 8 indicates that about one third of the total lung fields is occupied by areas of emphysema. %D in smokers with LAA scores of ⩾8.0 was significantly lower than in those with LAA scores of <8.0. These two groups were matched for CC but not for age distribution.

Table 7 Contribution of D genotype to LAA score, Kco, and $FEV_{1.0}$ independent of age and cigarette consumption in logistic regression analysis

Definition	Independent factor	χ^2	p value	Odds ratio	95% CI
LAA ⩾8	Genotype D	6.41	0.01*	0.40	0.20 to 0.81
(76/203, 37%)	Age (10 years)	7.48	0.01*	1.58	1.14 to 2.18
	CC (10 pack years)	0.29	0.59	1.03	0.94 to 1.12
%Kco <65	Genotype D	4.57	0.03*	0.40	0.17 to 0.93
(55/170, 32%)	Age (10 years)	14.46	<0.01**	2.19	1.46 to 3.29
	CC (10 pack years)	0.85	0.36	1.05	0.95 to 1.17
$\%FEV_{1.0}$ <50	Genotype D	0.43	0.51	0.78	0.37 to 1.64
(57/203, 28%)	Age (10 years)	14.34	<0.01**	2.04	1.41 to 2.95
	CC (10 pack years)	1.78	0.18	1.07	0.97 to 1.17

CI= confidence interval; LAA=low attenuation area score; CC=lifelong cigarette consumption; Kco=carbon monoxide transfer coefficient; $FEV_{1.0}$=forced expiratory volume in 1 second.
*p<0.05, **p<0.01.
Odds ratio and 95% CI correspond to having genotype D, aging every 10 years, or 10 pack years increase in CC.

The functional severity of emphysema was assessed from %KCO. The threshold for KCO was taken to be 65% of the reference value—that is, subjects with %KCO of <65% were assumed to have significant impairment in TLCO between the alveolar gas phase and pulmonary microcirculation. The severity of airflow obstruction was estimated by $\%FEV_1$ with a threshold of 50% of the reference value. %D in smokers with %KCO of <65% was lower than that in those with %KCO of ≥65%, the difference being qualitatively similar to that obtained with LAA scores. In the analysis of %KCO, however, age distribution and CC were not matched between the groups. There was no significant difference in %D between the groups divided by $\%FEV_{1.0}$.

Logistic regression analysis was performed to examine whether the D genotype functioned as an independent factor to modify the development of emphysema and airflow obstruction apart from aging and the amount of cigarettes smoked (table 7). The D genotype was found to reduce the risk of enhancement of the LAA score and impairment of %KCO, but it was not related to the extent of airflow obstruction. Unlike the D genotype, aging was found to be an important factor in promoting emphysematous changes (LAA score and %KCO) and deteriorating airflow obstruction ($\%FEV_{1.0}$).

DISCUSSION

Contribution of *CYP2A6del* genotype to smoking habit

Analysis of the results of lifelong CC and *CYP2A6del* genotype frequency showed that the presence of this deletion allele prevented subjects from becoming heavy smokers (CC ≥60 pack years, table 2). The association of *CYP2A6del* genotypes with lifelong CC, number of cigarettes per day, and duration of smoking suggested that the deletion allele limited the number of cigarettes/day rather than the duration of smoking, so that packs/day nearly paralleled lifelong CC, especially in current smokers (tables 1–3). It is reasonable to speculate that the number of cigarettes consumed per day is limited by the increased serum levels of nicotine in subjects with the deletion allele. These observations are consistent with previous findings that impaired function of *CYP2A6* could reduce cigarette consumption.[7 8] However, this is only true for heavy smokers with a lifelong CC of ≥60 pack years, and does not apply to relatively light smokers with a CC of 10–59 pack years. The frequency of the *CYP2A6del* genotype in these light smokers was comparable to that in non-smokers (table 2), suggesting that this deletion allele could not always protect subjects from becoming habitual smokers. It is interesting that the frequency of the D genotype was higher in current smokers with a relatively light CC of 0.5–1.49 packs/day than in non-smokers, which suggests that this genotype may promote some subjects to become habitual smokers.

These findings have not been suggested in previous reports, probably because of the small number of subjects with defective alleles.[7] However, it is also possible that our results were affected by linkage with other functional polymorphisms in *CYP2A6* which have recently been reported.[8 16]

It should be noted that many of the non-smokers in this study had not tried to smoke, and they should not be regarded as those who cannot smoke but rather as those who did not become habitual smokers. The effect of this defective polymorphism on addiction to smoking cannot therefore be fully discussed by comparing the genotype frequencies in the groups studied. The similar frequency of the *CYP2A6del* genotype in light smokers and in non-smokers may be important in understanding the essential role of *CYP2A6* genotypes in smoking related diseases, because the *CYP2A6del* allele does not protect subjects from becoming habitual smokers in whom the amount of CC does not exceed 60 pack years or 1.5 packs/day (table 2), which may be sufficient to exacerbate various smoking related disorders including respiratory and cardiovascular disease.

The contribution of genotype D to the decrease in the rate of ex-smokers was independent of age and CC (table 5). The prolonged presence of nicotine in the circulation may inhibit subjects with this defective allele from withdrawing their dependence on nicotine when they try to quit smoking. In future, the *CYP2A6* genotype should be determined when nicotine replacement therapy is considered because the nicotine concentration in the blood is expected to differ in smokers with different genotypes.[6 8] However, our analysis of this issue was limited because cessation of smoking was not confirmed by objective methods such as measurement of plasma cotinine levels. Furthermore, people quit smoking for various reasons such as aging, impaired lung function, diagnosis of COPD, and other environmental and genetic factors. In addition to the fact that there were far fewer subjects with genotype D among the ex-smokers than the current smokers (table 4), such heterogeneity of ex-smokers may result in obscuring the effects of the genotype on the parameters of smoking (tables 1–3).

Effect of *CYP2A6del* genotype on development of smoking related COPD

The severity of COPD has been considered to be associated with age, cigarette consumption, and sensitivity to smoking.[17] It is interesting that the *CYP2A6del* genotype occurred significantly more frequently in smokers with a low LAA score than in those with a high LAA score, although these groups were matched for lifelong CC (table 6). These observations were confirmed by the same tendency for %KCO which is known to be closely correlated with the severity of pulmonary emphysema. Furthermore, the protective effects of the D genotype against emphysema were confirmed by logistic regression analysis adjusted for age and cigarette consumption (table 7). These findings suggest that the *CYP2A6del* mutation directly acts as an intrinsic factor against the development of emphysematous changes, independent of the function limiting the amount of CC. In the present analysis we have used lifleong CC (pack years) but not daily CC (packs/day) as an overall measure of exposure to smoking because, in terms of potential to promote emphysematous changes leading to impaired lung function, lifelong CC is better than packs/day as the latter does not take into account duration of smoking, although CC is significantly influenced by age. We have eliminated the effect of age contained implicitly in CC by applying logistic regression analysis in which the effect of age and that of CC can be assessed separately.

There are a number of other confounding factors related to the development of emphysema and airflow obstruction besides age and CC such as race, sex, air pollution, and social status. The contribution of the genotype D to the severity of emphysema and impaired KCO was not as much as that of age, indicating that further investigations are needed to clarify the relationship between this polymorphism and the development of pulmonary emphysema.

The protective effect of the genotype D against the development of emphysema was obscured when $\%FEV_{1.0}$ was used as an indicator to judge the severity of COPD (tables 6 and 7). The severity of airflow obstruction represented by $\%FEV_{1.0}$ is known to be determined by various factors including emphysematous changes, airway wall thickening, and airway hyperreactivity.[18] The absence of a definite association between the *CYP2A6* mutation and $\%FEV_{1.0}$ may not therefore be unreasonable. Our observations suggest that unspecified substances metabolised by *CYP2A6* may be responsible for the difference in LAA score between smokers with different genotypes, and are related to (or modify) the inflammatory processes specific to emphysematous changes such as elastase-antielastase imbalance and interaction between oxidants and antioxidants. Miyamoto *et al*[9] found that the *CYP2A6del* allele reduced the risk of lung cancer in Japanese

populations, which may be explained by inhibition of the activation of procarcinogens in subjects with the deletion allele. Smoking habit including CC was not considered in that report, although CC is also important as a risk factor for lung cancer. Cantlay *et al*[19] suggested a possible association between a genetic polymorphism in *CYP1A1*, which is also known to activate procarcinogens, and the development of both lung cancer and pulmonary emphysema. However, the mechanisms causing the difference in emphysematous changes between the genotypes were not described.

Critique of methods

One criticism of studies of CC is that there is no reliable threshold for distinguishing between heavy and light smokers, which should be determined not only from lifelong and/or daily CC but also by the duration of smoking. We therefore evaluated the difference in *CYP2A6del* genotype frequency at different thresholds of lifelong CC, daily CC, and duration of smoking, and investigated whether the essential tendency would be changed qualitatively when the threshold was moved to other values. Since the frequency of the *CYP2A6del* genotype was consistent at most of the different thresholds studied, we adopted thresholds for lifelong CC, daily CC, and duration of smoking of 60 pack years, 1.5 packs/day, and 40 years, respectively, and these were used in subsequent analyses (table 3).

Another criticism is that there are no authentic criteria for defining the severity of emphysema from the LAA score and Kco. We have previously shown that the 95% confidence limit of the relative area of LAA (%LAA) examined in non-smoking controls with no signs of pulmonary disease averaged 25%.[20] Furthermore, Mishima *et al*[21] found that the %LAA never exceeded 30% in smokers with a small amount of airflow obstruction. These findings suggest that pathological emphysema can be judged to be present when 25% LAA is taken as the threshold, while the extent of morphological severity of emphysema can be approximately assessed when 30% is used as the threshold. We therefore assumed that an LAA threshold of 8 (corresponding to 33% LAA) would distinguish smokers with significant emphysema from those with less emphysema.

We assumed that Kco was significantly impaired when it was lower than 65% of the reference value. Based on the SD of Kco reported for healthy Japanese subjects,[11] we found that the 95% confidence limit of this parameter corresponded to 68% of the reference value. We therefore took 65% of the reference value as the threshold for judging the impairment in Kco. Significant correlation was found between %Kco and the LAA score, expressed by the equation: LAA score $= 16.05 - 0.120 \times$ %Kco ($r=0.70$, $p<0.0001$). Based on this equation, the LAA score corresponding to 65% of the reference value of Kco was found to be 8.3, which is consistent with the threshold used for judging the morphological severity of emphysema on CT images.

The severity of airflow obstruction in smokers was evaluated according to the criteria recommended by the GOLD guideline.[25] Using a threshold value of 50% of predicted FEV_1, smokers were divided into two groups ($\geq$50% and $<$50% FEV_1); $\%FEV_1$ was not closely correlated with LAA score ($n=203$, $r=0.50$, $p<0.0001$).

We conclude that (1) the *CYP2A6del* allele appears to restrict the amount of CC in heavy smokers but not in light smokers, (2) the *CYP2A6del* allele may inhibit smokers from quitting smoking, and (3) the protective role of the *CYP2A6* deletion allele in pulmonary emphysema seems to be derived from impaired activity of this enzyme, independent of its effect on regulating CC. These findings suggest that determination of the genotype will be useful in efficiently withdrawing patients from nicotine dependence in smoking cessation protocols with nicotine containing materials, and will give a new insight into the pathogenesis of smoking induced pulmonary emphysema.

ACKNOWLEDGEMENT

This study was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science, and Culture of Japan (12470138 and 12770301) and Keio University grant-in-aid for encouragement of young medical scientists.

Authors' affiliations
N Minematsu, H Nakamura, M Iwata, H Tateno, T Nakajima, S Takahashi, S Fujishima, K Yamaguchi, Department of Medicine, School of Medicine, Keio University, Tokyo, Japan

REFERENCES

1 **Caraballo RS**, Giovino GA, Pechacek TF, *et al*. Racial and ethnic differences in serum cotinine levels of cigarette smokers. *JAMA* 1998;**280**:135–9.
2 **Sellers EM**. Pharmacogenetics and ethnoracial differences in smoking. *JAMA* 1998;**280**:179–80.
3 **Yamano S**, Tatsuno J, Gonzalez FJ. The CYP2A3 gene product catalyzes coumarin 7-hydroxylation in human liver microsomes. *Biochemistry* 1990;**29**:1322–9.
4 **Fernandez-Salguero P**, Hoffman SM, Cholerton S, *et al*. A genetic polymorphism in coumarin 7-hydroxylation: sequence of the human *CYP2A* genes and identification of variant *CYP2A6* alleles. *Am J Hum Genet* 1995;**57**:651–60.
5 **Chen GF**, Tang YM, Green B, *et al*. Low frequency of *CYP2A6* gene polymorphism as revealed by a one-step polymerase chain reaction method. *Pharmacogenetics* 1999;**9**:327–32.
6 **Kitagawa K**, Kunugita N, Katoh T, *et al*. The significance of the homozygous *CYP2A6* deletion on nicotine metabolism: a new genotyping method of *CYP2A6* using a single PCR-RFLP. *Biochem Biophys Res Commun* 1999;**262**:146–51.
7 **Pianezza ML**, Sellers EM, Tyndale RF. Nicotine metabolism defect reduces smoking. *Nature* 1998;**393**:750.
8 **Rao Y**, Hoffmann W, Zia M, *et al*. Duplications and defects in the *CYP2A6* gene: identification, genotyping, and in vivo effects on smoking. *Mol Pharmacol* 2000;**58**:747–55.
9 **Miyamoto M**, Umetsu Y, Dosaka-Akita H, *et al*. *CYP2A6* gene deletion reduces susceptibility to lung cancer. *Biochem Biophys Res Commun* 1999;**261**:658–60.
10 **American Thoracic Society**. Standardization of spirometry: 1994 update. *Am J Respir Crit Care Med* 1995; **152**:1107–36.
11 **Nishida O**, Kambe M, Sewake N, *et al*. Pulmonary function in healthy subjects and its prediction. 5: Pulmonary diffusing capacity in adults. *Rinsho Byori* 1976;**24**:941–7.
12 **Japanese Respiratory Society**. Reference values of spirogram and arterial blood gas levels in Japanese. *Nihon Kokyuki Gakkai Zasshi* 2001;**39**:S1–17.
13 **Goddard PR**, Nicholson EM, Laszlo G, *et al*. Computed tomography in pulmonary emphysema. *Clin Radiol* 1982;**33**:379–87.
14 **Oscarson M**, McLellan RA, Gullstén H, *et al*. Characterisation and PCR-based detection of a *CYP2A6* gene deletion found at a high frequency in a Chinese population. *FEBS Lett* 1999;**448**:105–10.
15 **Nunoya K**, Yokoi T, Kimura K, *et al*. A new deleted allele in the human cytochrome P450 2A6 (*CYP2A6*) gene found in individuals showing poor metabolic capacity to coumarin and (+)-*cis*-3,5-dimethyl-2-(3-pyridyl) thiazolidin-4-one hydrochloride (SM-12502). *Pharmacogenetics* 1998;**8**:239–49.
16 **Kitagawa K**, Kunugita N, Kitagawa M, *et al*. *CYP2A6*6*, a novel polymorphism in cytochrome p450 2A6, a single amino acid substitution (R128Q) that inactivates enzymatic activity. *J Biol Chem* 2001;**276**:17830–5.
17 **Fletcher C**, Peto R. The natural history of chronic airway obstruction. *BMJ* 1977;**25**:1645–8.
18 **Nakano Y**, Muro S, Sakai H, *et al*. Computed tomographic measurements of airway dimensions and emphysema in smokers. *Am J Respir Crit Care Med* 2000;**162**:1102–8.
19 **Cantlay AM**, Lamb D, Gillooly M, *et al*. Association between the CYP1A1 gene polymorphism and susceptibility to emphysema and lung cancer. *J Clin Pathol Clin Mol Pathol* 1995;**48**:M210–4.
20 **Soejima K**, Yamaguchi K, Kohda E, *et al*. Longitudinal follow-up study of smoking-induced lung-density changes by high-resolution computed tomography. *Am J Respir Crit Care Med* 2000;**161**:1264–73.
21 **Mishima M**, Hirai T, Itoh H, *et al*. Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. *Proc Natl Acad Sci USA* 1999;**96**:8829–34.
22 **Roberts CM**, MacRae KD, Winning AJ, *et al*. Reference values and prediction equations for normal lung function in a non-smoking white urban population. *Thorax* 1991;**46**:643–50.
23 **Robson AG**, Innes JA. Short term variability of single breath carbon monoxide transfer factor. *Thorax* 2001;**56**:358–61.
24 **Clark KD**, Wardrobe-Wong N, Elliott JJ, *et al*. Patterns of lung disease in a "normal" smoking population: are emphysema and airflow obstruction found together? *Chest* 2001;**120**:743–7.
25 **Global Initiative for Chronic Obstructive Lung Disease**. *Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease*. Publication no 2701. NHLBI/WHO Workshop Report. National Institute of Health, National Heart, Lung, and Blood Institute, 2001:7–8.