extracting hidden combinations that are statistically significant. The main principle of multiple stratifications is the same as the former "decision tree," but this method constructs plural tree models in contrast to the one-tree model by regular decision tree algorithm.

In this article, we introduce an example of application of this new exploratory data mining algorithm to our SNPs data of stroke patients and healthy controls focusing upon the combination extraction.

Materials and methods

The patients and healthy controls analyzed here were from the same population published in our previous studies (Ishii et al. 2004; Ito et al. 2002; Oguchi et al. 2000; Sonoda et al. 2000 etc.). We recruited 235 unrelated Japanese patients under 70 years of age with symptomatic ischemic cerebrovascular disease (CVD) from Keio University Hospital and 189 age- and gender-matched healthy controls. CVD patients with cardioembolic cerebral infarction and cerebral hemorrhage were excluded. Control subjects were enlisted from people who came to the hospital for regular checkups; those who had a clinical history of CVD or myocardial infarction or peripheral vascular diseases were excluded. Informed consent was obtained from all subjects. Brain CT and/or MRI were performed in all CVD patients. MR angiography and/or extracranial duplex ultrasonography were available for more than 80% of CVD patients. On the basis of classification of CVDs III (ad hoc committee of the National Institute of Neurological Disorders and Stroke 1990), the CVD patients were divided into three clinical categories: an atherothrombotic infarction group (69 patients), a lacunar infarction group (142 patients), and a transient ischemic attack group (24 patients). In this study, we focused on the atherothrombotic infarction group to clarify the differences from healthy controls by limiting the cases to one pathophysiological state. We typed 14 genetic polymorphism markers in 69 atherothrombotic stroke patients and 189 healthy controls. These 14 markers were chosen because they were already known to be related to stroke by effecting blood coagulation, platelet function, and lipid metabolism (target gene approach). These genetic polymorphism markers and their abbreviations used in this analysis are shown in Table 1, and the characteristics of the stroke patients and healthy controls are compared in Table 2.

For conventional statistical analysis, we used SPSS 11.0 and Clementine 7.1 (Chicago, IL, USA). For advanced statistical analysis, we devised the exact tree method based on decision tree analysis. The procedure of the exact tree method is described below. First, we made cross-tabs of genotypes between the groups of stroke patients and healthy controls. Then, we compared the prevalence of these alleles between the groups and calculated the P value of Fischer's exact test (See "Appendix"). We defined these P values as splitting indices of the data set to make the decision tree models. The smaller the P value, the more influential the allele was assumed to be. In regular decision tree methods, we adopt the best splitting point that shows the best index score (e.g., split the first data set by the marker that shows the smallest P value). But this method discards the second-best or the third-best onward, even when they are also meaningful. In the exact tree method, we also adopt the second-best or third-best onward as well as the first-best, as the occasion demands. How many best points we decide to adopt is the number of the tree models obtained later. For the effectiveness in data exploration, we must decide the number of the tree analytically. One simple approach is counting the variables that show the P value at an arbitrary statistical significance level depending on the number of variables analyzed. Another approach we adopted here was to make the talus plot of P values of each variable. The talus plot is a line graph of P values and is plotted in ascending order (Fig. 1). Using this plot, we could figure out the point where the slope changed sharply. If we chose the variables plotted on the left side of the large slope change, we adopted the variables that contained relatively larger amounts of information efficiently. With reference to Fig. 1, we chose the first five variables visually and made five tree models. In each tree model, we adopted the genetic polymorphism marker that had the smallest P value as the second splitting point and grew the trees with the next smallest P values continuously. Such a repetitive process needs stopping rules. As is common for all decision tree methods, the more the root node was split deeply, the smaller the number of subjects becomes in each stratified node, and it tends to be difficult to be interpreted. How many times the tree model should be split depends on the number of study subjects. In this report, we empirically defined the depth of tree as four, which is the level less than

Table 1 List of genetic polymorphism markers and abbreviations used for analysis. Genetic polymorphism markers are listed in alphabetical order by abbreviations

Genetic polymorphism markers	Abbreviations
5HT2A receptor (Exon1 102 T/C)	5HT2A
Angiotensin-converting enzyme (Exon 16)	ACE
Human apolipoprotein E	APO-E
Cholesteryl ester transfer protein	CETP
Fibrinogen β	FIB-B
Factor XII gene (Exon1 46 C/T)	F-XII
Glycoprotein Ib (145 Thr/Met)	GPIb (145Thr/Met)
Glycoprotein Ib kozak sequence (5 T/C)	GPIb (kozak)
Mitochondrial DNA mutation (5178 A/C)	Mitochondria
Methylene tetra hydoro folate reductase (677 C/T)	MTHFR
Endothelial nitric oxide synthase (Intron 4)	NOS
NADPH oxidase p22phox	p22phox
Plasminogen activator inhibitor-1 (promotor region)	PAI-1
Paraoxonase (192 Arg/Gln)	PON

Table 2 Comparative table of characteristics

	All participants ($n=257$)	Stroke patients ($n=68$)	Healthy controls ($n=189$)
Demographic			
Age (years)[a] **	60.3 (5.2)	63.4 (7.7)	59.5 (3.7)
Male/female	201/56	56/12	145/44
Clinical			
Body-mass index (kg/m^2)[a] *	22.8 (2.7)	23.7 (2.6)	22.6 (2.7)
Smoking (%) never/ex or current **	65.4/34.6	42.1/57.9	72.5/27.5
Hypertension, yes/no *	82/173	30/38	52/135
Diabetes, yes/no **	35/222	22/46	13/176
Hyperlipemia, yes/no **	93/138	25/17	68/121
Family history of cerebellar vein disease, yes/no	37/141	13/39	24/102
Family history of cardiac artery disease, yes/no	17/163	4/51	13/112
Biochemical			
Fast blood sugar (mg/dl)[a] *	101.1 (18.8)	109.1 (31.7)	99.2 (13.7)
Total cholesterol (mg/dl)[a]	203.4 (32.7)	201.1 (37.1)	204.1 (31.1)
Triglyceride (mg/dl)[a] **	117.3 (77.4)	160.0 (90.8)	108.5 (71.5)
Serum uric acid (mg/dll)[a]	5.7 (1.3)	5.6 (1.4)	5.7 (1.3)
Platelet count (μl)[a]	235.5 (70.0)	229.4 (101.4)	237.5 (55.4)
High density lipoprotein (mg/dl)[a] **	56.2 (15.3)	44.0 (11.6)	57.5 (15.1)
CETP density (μg/ml)[a]	2.5 (0.6)	2.4 (0.6)	2.5 (0.6)

[a]mean (SD)
*$P<0.05$, *t*-test; **$P<0.01$, chi-square test

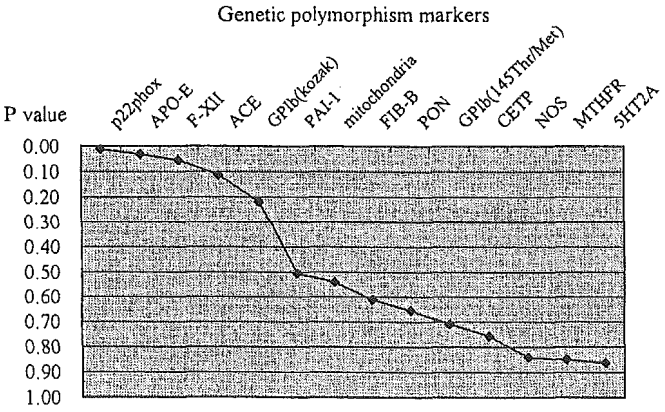


Fig. 1 Talus plot of *P* values for Fischer's exact test

half of the terminal nodes that consist of 30 subjects. Before reaching the defined depth, the extension was terminated when the number of cases or controls in the node became zero.

In addition, we calculated the odds ratios (OR) of each node with 95% confidence intervals (Figs. 2, 3, 4, 5, 6). Each OR stands for odds ratio of each node to "root node," which is the whole population we analyzed shown as Node #0 in each figure. It thus makes it possible to remove the influence of the sampling ratio of cases to controls on the estimated OR. We added the *P* value of each OR in addition to the 95% confidence interval of each OR. The OR shows the estimate of the risk ratio of the disease, so the odds ratio in this study indicates the risk of stroke. We circled the node with statistically significant odds ratio in each tree. These processes (which we call the "exact tree" method) enabled us to appraise risks quantitatively and extract meaningful combinations from the tree models constructed.

Results

The alleles used as variables in this analysis are listed in Table 3. The *P* values according to Fischer's exact test in the case-control tables are listed in Table 4, and the talus plot of these *P* values, which was plotted in ascending order, is shown in Fig. 1. From the view point of *P* value, p22phox (*P* value = 0.012) and APO-E (*P* value = 0.031) had statistically significant differences in allele frequencies between stroke patients and healthy controls. Based on the talus plots, we adopted F-XII (*P* value = 0.056), ACE (*P* value = 0.112), and GPIb (kozak) (*P* value = 0.220) as useful genetic polymorphism markers for the first node splitting in the tree models, since it is more likely to be statistically significant in some combinations than markers with lower *P* values. By exact tree method described above, we constructed five exact tree models (Figs. 2, 3, 4, 5, 6). As to the second or later splitting point in each exact tree, we adopted the variables that had the smallest *P* value continuously. Subsequently, we calculated the OR of each node of the trees and drew circles around nodes whose 95% confidence interval of the odds ratio was not across 1 (i.e., significant at 5% alpha-error level). In the first exact tree (Fig. 2), which is the usual decision tree in itself, we could not extract any significant combination, since the only circled Node #2 indicated the elevated risk [Odds ratio: 2.34 (1.14–4.81)] in p22phox t allele-positive people (single polymorphism). Significant combination extraction needs more than twice-split circled node. In the second exact tree (Fig. 3), we observed three significant combinations in three nodes (Node #5, #8 and #11). In Node #5, for example, when with a combination of APO-E E4 allele-positive and PAI-1 5G allele-negative, the odds ratio became 3.24 (1.05–9.99). In Node #8,

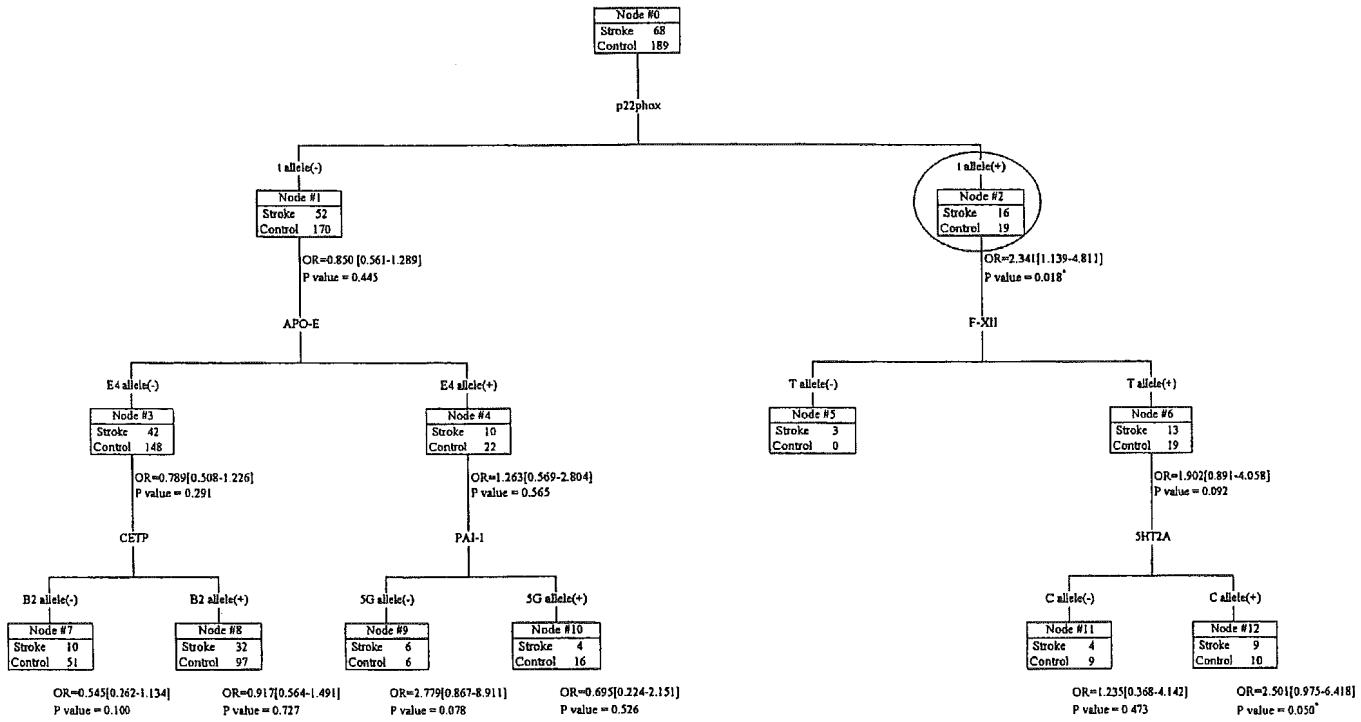


Fig. 2 Exact tree model #1 (p22phox). *Circle* statistically significant combination, *OR* odds ratio of each node to root node (Node #0) and its 95% confidence interval. $*P < 0.05$, chi-square test

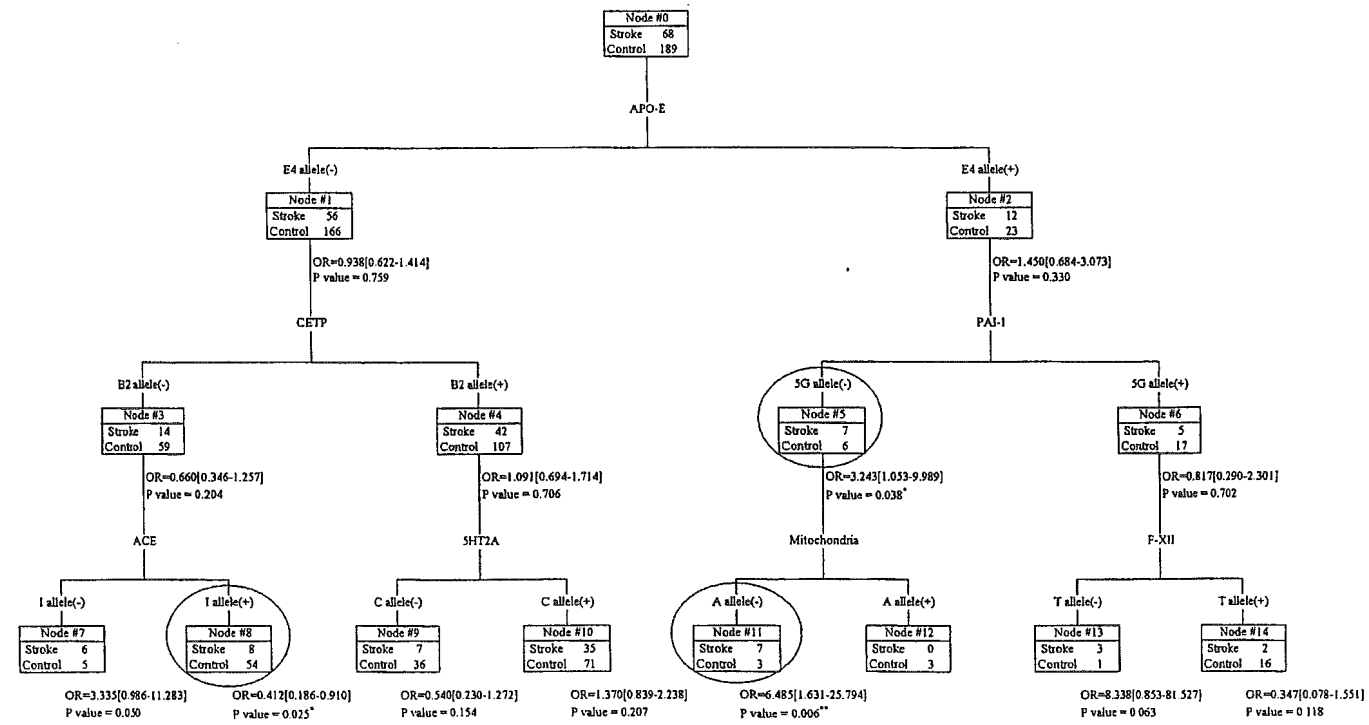


Fig. 3 Exact tree model #2 (APO-E). *Circle* statistically significant combination, *OR* odds ratio of each node to root node (Node #0) and its 95% confidence interval. $*P < 0.05$ and $**P < 0.01$, chi-square test

with a combination of APO-E E4 allele-positive, PAI-1 5G allele-negative and Mitochondria A allele-negative, the odds ratio became 6.49 (1.63–25.79). And in Node #11, with a combination of APO-E E4 allele-negative, CETP B2 allele-negative, and ACE I allele-positive, the odds ratio became 0.41 (0.19–0.91), which can be a protective combination against stroke. In the same way, we can extract another one, three, and three statistically significant combinations from the tree models in Figs. 4, 5, and 6, respectively.

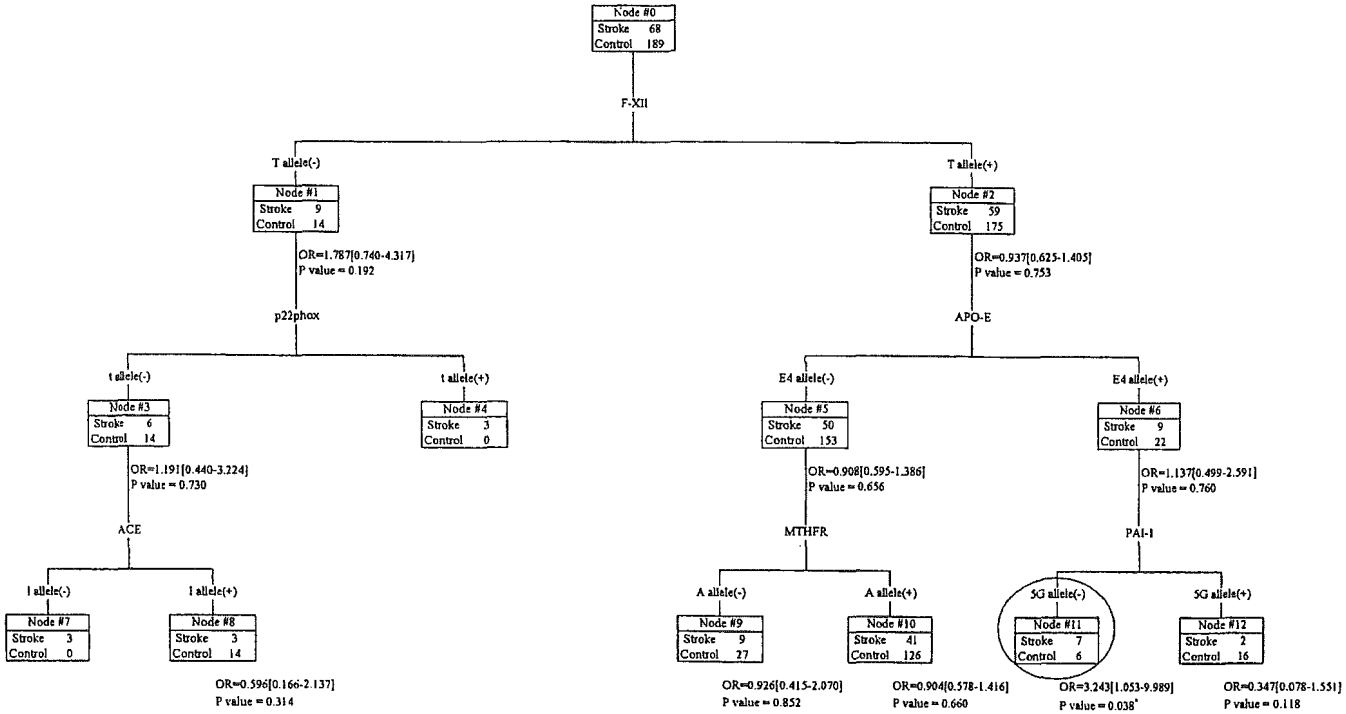


Fig. 4 Exact tree model #3 (F-XII). *Circle* statistically significant combination, *OR* odds ratio of each node to root node (Node #0) and its 95% confidence interval. $*P < 0.05$, chi-square test

Discussion

Although not shown in the Figures, we extended Nodes #7, #8, #9, and #10 in Fig. 2 (which are the same structure as Nodes #3, #4, #5, and #6 in Fig. 3) one more step down and found the same split in Nodes #7, #9, and #10. Node #8 was split by APO-E E2 allele instead of 5HT2A C allele. Because the left-side branch in Fig. 2 started from 222 subgroup [Node #1, p22phox t allele(−)], it is not surprising that the later splits are not

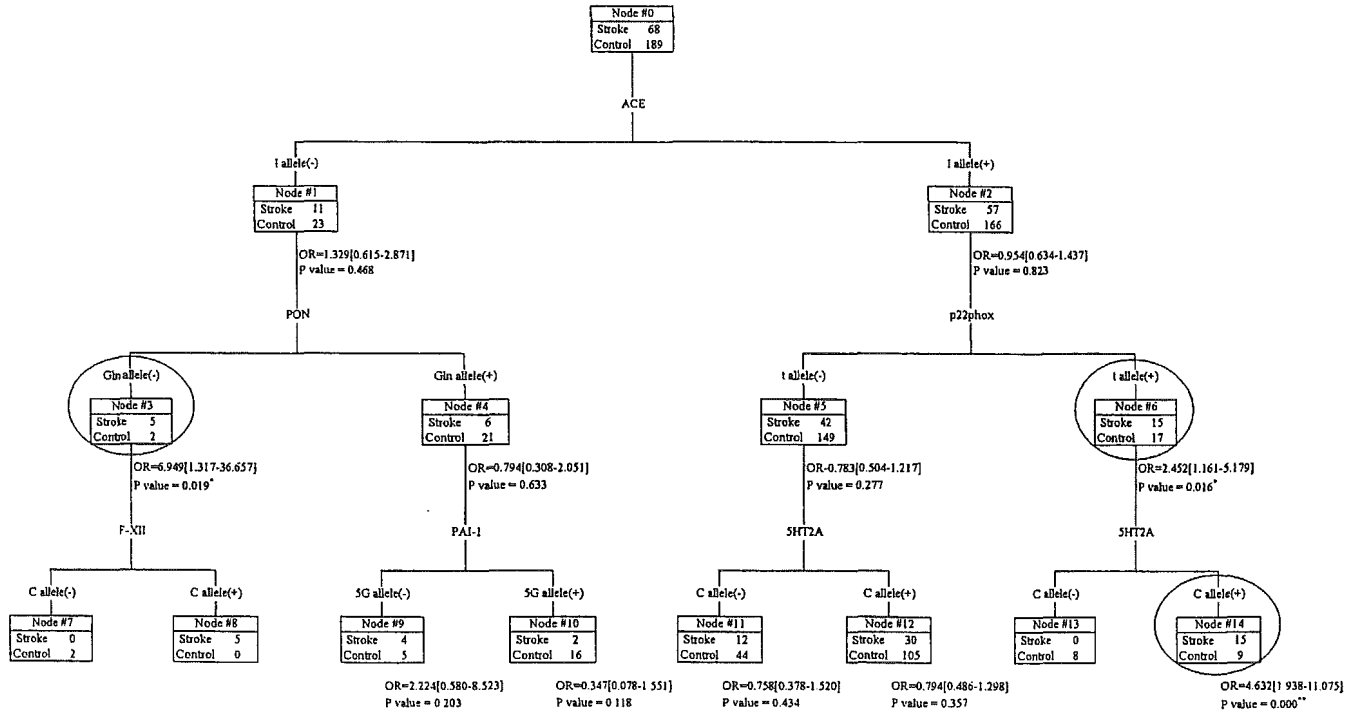


Fig. 5 Exact tree model #4 (ACE). *Circle* statistically significant combination, *OR* odds ratio of each node to root node (Node #0) and its 95% confidence interval. $*P < 0.05$ and $**P < 0.01$, chi-square test

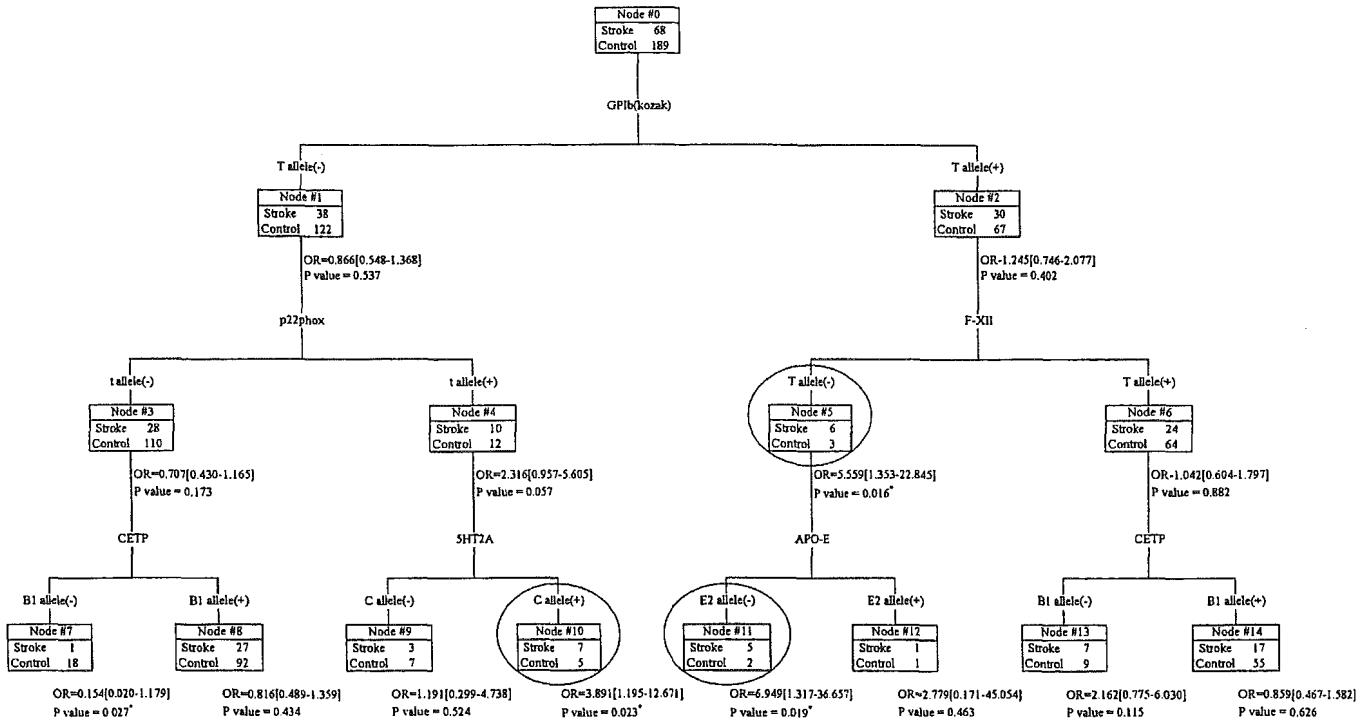


Fig. 6 Exact tree model #5 [GPIb (kozak)]. *Circle* statistically significant combination, *OR* odds ratio of each node to root node (Node #0) and its 95% confidence interval. $*P < 0.05$, chi-square test

Table 3 List of alleles used as variates for this analysis. Genetic polymorphism markers are listed in alphabetical order

Genetic polymorphism markers	Variates		
5HT2A	C allele	T allele	
ACE	D allele	I allele	
APO-E	E2 allele	E3 allele	E4 allele
CETP	B1 allele	B2 allele	
FIB-B	H1 allele	H2 allele	
F-XII	C allele	T allele	
GPIb (145Thr/Met)	M allele	T allele	
GPIb (kozak)	C allele	T allele	
Mitochondria	A allele	T allele	
MTHFR	A allele	V allele	
NOS	a allele	b allele	
P22phox	c allele	t allele	
PAI-1	4G allele	5G allele	
PON	Arg allele	Gln allele	

Table 4 List of acquired *P* values by Fischer's exact test in case-control tables. Genetic polymorphism markers are listed in ascending order by *P* value

Genetic polymorphism markers	Variates	*P* values
P22phox	t allele*	0.012
APO-E	E4 allele*	0.031
F-XII	T allele	0.056
ACE	I allele	0.112
GPIb (kozak)	T allele	0.220
PAI-1	5G allele	0.506
Mitochondria	A allele	0.541
FIB-B	H1 allele	0.611
PON	gln allele	0.654
GPIb (145Thr/Met)	M allele	0.707
CETP	B2 allele	0.756
NOS	a allele	0.840
MTHFR	A allele	0.849
5HT2A	C allele	0.863

* $P < 0.05$, Fischer's exact test between stroke patients and healthy controls

completely consistent with the result from 257 whole population cases. When a similar structure was observed such as between Fig. 3 and the left-side branch in Fig. 2, we could speculate the root split (p22phox t allele in this case) influence little on the later splits. However, it does not mean the root split is unnecessary for making the tree, since the other side branch can be meaningful (unfortunately meaningless in this case). Such a consideration can be made because the number of tree models is plural in the exact tree method. Regular decision tree algorithm makes only one tree, and structural similarity cannot be evaluated. This is one of the strong points the exact tree algorithm has over the regular decision tree algorithm.

As shown in Fig. 1, 5HT2A did not show a high score (i.e., low *P* value) by itself, but was useful for making trees (Figs. 2, 3, 5, 6). It is speculated that this polymorphism has a low risk by itself but has a high risk when combined with the other polymorphisms. Such a combination is the main target the exact tree method aims for.

From the same population, the allele frequencies observed can differ in each sampling time. However,

these differences can be controlled within reasonable statistical error rates. Of course even statistically significant combinations may contain some false positives, but this problem might be out of scope of exploratory methods by nature. Other verificatory approaches are needed for explored hidden combinations, and the exact tree method will work efficiently in the former steps.

In this paper, we adopted the P value of Fischer's exact test as the splitting criterion for the decision tree. This splitting criterion can be substituted for the chi-square of CHAID (Kass 1980), the improvement score of CART (Breiman et al. 1980), the gain ratio of C4.5 (Quinlan 1993), and so on, by using the exact tree method we proposed. We have already tested the effectiveness of the exact tree method with the odds ratio as the splitting criterion (data not shown). It is impossible to state which splitting criterion is the best in decision tree modeling, because the each criterion is the best in each mathematical modeling theory. So we adopted the P value of Fischer's exact test because it is easily interpreted as the probability and is familiar to biologists and medical researchers. Moreover, we can use the commonly used alpha error rate (0.05 or 0.01) as the standard for judging the magnitude of the value.

Here we need to clarify the advantage of the exact tree method over the regular decision tree method. If we adopt the regular decision tree method, we can get the best (i.e., only one) classification tree model (Fig. 2). On the contrary, the exact tree method makes plural tree models, including the best classification tree model (Figs. 2, 3, 4, 5, 6). In the real data analysis, while the regular decision tree shows us no significant combination (Fig. 2), five tree models constructed by the exact tree method showed us ten significant combinations (Figs. 2, 3, 4, 5, 6). Because the output of the exact tree method always includes the best tree model, it is never inferior to the regular decision tree method as to the extraction of hidden significant combinations. In addition, the amount of computation increases not exponentially but only several times from that of regular decision tree analysis. That is why we report this new method as an efficient and feasible exploratory technique for handling such large-scale genomic data as round-robin approach, which otherwise is impossible.

For reference, we executed the logistic regression analysis in which 14 polymorphisms are incorporated as independent variables. APO-E E4 allele and Factor XII gene t allele are statistically significant polymorphisms by both compulsory and stepwise methods. In the exact tree method, these two polymorphisms are selected as the first splitting variables in Figs. 3 and 4, and four statistically significant combinations are extracted. Regular decision tree method failed to find any combinations of the two polymorphisms above. This finding supports the advantage of the exact tree method over regular decision tree methods.

Since this method is an exploratory screening method, we do not insist that the individual results are solid evidences by themselves. We want to emphasize our success in extracting combinations of three polymorphisms using real genetic data. From the obtained tree models, we succeeded in extracting ten statistically significant combinations that may elevate or decrease the risk of stroke.

Here we handle a very small data set (only 14 genetic markers) in this pilot analysis, since round-robin explorations of multiple combinations are executable in ordinary computer workstations. So the advantage of the exact tree method is obscure, but the situation will change in handling extremely large-scale data set. For example, when we handle 100,000 genetic markers, which can be real in some whole genome approach researches, we need 10,000,000,000 operations to check all the combinations of two markers. In this situation, we need 1,000,000,000,000,000 operations to check all the combinations of three markers. Such an acute increase of the amount of computation at an exponential rate makes it impossible to extract combinations of large numbers of markers. With this new data mining method, however, the increase of the amount of computation is not exponential, and it can extract multiple combinations efficiently without computational explosion.

To assess the degree of the reduction in calculation task quantitatively, we did the reduction rate estimation trial. Suppose the number of genetic polymorphisms is N and we want to know the statistically significant disease-associated combinations of three polymorphisms. In the round-robin approach, we need to execute chi-square test $N \times (N-1) \times (N-2)$ times. On the other hand, the 3-step splitting exact tree method (i.e., the depth of the tree is 4) demands $3 \times n \times N$ times, where n is the number of tree models we make, which is less than N and usually less than 10. So the degree of the reduction in calculation task is evaluated as $\frac{3 \times n \times N}{N \times (N-1) \times (N-2)}$ and this reduction ratio simply decreases as N increases. And the number of tree models (n) is usually rather smaller than the total number of polymorphisms in consideration (N), this reduction ratio decreases exponentially. In spite of the rather small number of polymorphisms we demonstrated in this paper ($N=14$, $n=5$), this reduction rate is estimated to be 0.096. This estimation suggests the exact tree method works effectively enough in handling rather small genetic data, and this efficiency will increase when handling larger data. This new data mining method will facilitate the extraction of combinations from the huge genetic data that we are now confronted with and may provide a good foothold for further verificatory research.

Appendix

The exact probability of observing a table with cells (a, b, c, d) is

when

$$P(a,b,c,d) = \frac{{}_{m_1}C_a \cdot {}_{m_2}C_b}{{}_{N}C_{n_1}} = \frac{n_1!n_2!m_1!m_2!}{N!}\frac{1}{a!b!c!d!}$$

	Disease (+)	Disease (−)	
Polymorphism (+)	a	b	n_1
Polymorphism (−)	c	d	n_2
	m_1	m_2	N (total)

This probability distribution is known as the hypergeometric distribution (Bernard 2000).

Suppose the probability that a man had polymorphism (+) given that he had disease (+) $=p_1$, and the probability that a man had polymorphism (+) given that he had disease (−) $=p_2$. Here we wish to test the hypothesis H_0: $p_1=p_2$ versus H_1: p_1 not $=p_2$.

General procedure and computation of P value using Fischer's exact test are as follows:

1. Enumerate all possible tables with the same row and column margins (i.e., n_1, n_2, m_1, m_2) as the observed table
2. Compute the exact probability of each table enumerated in previous step
3. Suppose the observed table is the x table and the last table enumerated is the k table. To test the hypothesis H_0: $p_1=p_2$ versus H_1: $p_1<p_2$, the P value is computed as $P(0)+P(1)+\cdots+P(x)$ (left-hand tail area). To test the hypothesis H_0: $p_1=p_2$ versus H_1: $p_1>p_2$, the P value is computed as $P(x)+P(x+1)+\cdots+P(k)$ (right-hand tail area). To test the hypothesis H_0: $p_1=p_2$ versus H_1: p_1 not $=p_2$, the P value is computed as $2\times\min[P(0)+P(1)+\cdots+P(x), P(x)+P(x+1)+\cdots+P(k), 0.5]$.

This P value can be interpreted as the probability of obtaining a table as extreme as the observed table.

References

Ad Hoc Committee of National Institute of Neurological Disorders and Stroke (1990) Classification of cerebrovascular disease III. Stroke 21:637–676

Bernard R (2000) Fundamentals of biostatistics, 5th edn. Duxbury, Pacific Grove

Breiman L, Friedman JH, Olshen RA, Stone CJ (1980) Classification and regression trees. Wadsworth, Belmont

Ishii K, Murata M, Oguchi S, Takeshita E, Ito D, Tanahashi N, Fukuuchi Y, Saitou I, Ikeda Y, Watanabe K (2004) Genetic risk factors for ischemic cerebrovascular disease of candidate prothrombotic gene polymorphisms in the Japanese population. Rinsho Byori 52:22–27

Ito D, Tanahashi N, Murata M, Sato H, Saito I, Watanabe K, Fukuuchi Y (2002) Notch3 gene polymorphism and ischemic cerebrovascular disease. J Neurol Neurosurg Psychiat 72:382–384

Kass G (1980) An exploratory technique for investigating large quantities of categorical data. Appl Stat 29:119–127

Miyaki K, Takei I, Watanabe K, Nakashima H, Watanabe K, Omae K (2002) Novel statistical classification model of type 2 diabetes mellitus patients for tailor-made prevention using data mining algorithm. J Epidemiol 12:243–248

Oguchi S, Ito D, Murata M, Yoshida T, Tanahashi N, Fukuuchi Y, Ikeda Y, Watanabe K (2000) Genotype distribution of the 46C/T polymorphism of coagulation factor XII in the Japanese population: absence of its association with ischemic cerebrovascular disease. Thromb Haemost 83:178–179

Quinlan JR (1993) C4.5: Programs for machine learning. Morgan Kaufmann, Los Altos

Sonoda A, Murata M, Ito D, Tanahashi N, Oota A, Tada-Yatabe Y, Takeshita E, Yoshida T, Saito I, Yamamoto M, Ikeda Y, Fukuuchi Y, Watanabe K (2000) Association between platelet glycoprotein Ibα genotype and ischemic cerebrovascular disease. Stroke 31:493–497

Brief report

Identification of novel mutations in *ADAMTS13* in an adult patient with congenital thrombotic thrombocytopenic purpura

Toshihiro Uchida, Hideo Wada, Minoru Mizutani, Miho Iwashita, Hiroaki Ishihara, Toshiro Shibano, Misako Suzuki, Yumiko Matsubara, Kenji Soejima, Masanori Matsumoto, Yoshihiro Fujimura, Yasuo Ikeda, and Mitsuru Murata, for The Research Project on Genetics of Thrombosis

Congenital thrombotic thrombocytopenic purpura/hemolytic uremic syndrome (TTP/HUS) is associated with an inherited von Willebrand factor–cleaving protease (ADAMTS13 [a disintegrin and metalloprotease with thrombospondin type I domains 13]) deficiency. In this study, we identified novel mutations in the *ADAMTS13* gene in a patient with TTP. The patient was a 51-year-old Japanese male who exhibited TTP symptoms at frequent intervals. The ADAMTS13 activity during acute episodes was less than 3% that of normal. The enzyme activities of the patient's father and mother were both 46%, and both parents were asymptomatic. Genetic analysis revealed that the patient was a compound heterozygote for 2 mutations. One mutation was a missense mutation in the metalloprotease domain (A250V, exon 7), and the other was a guanine to adenine substitution at the 5′ end of intron 3 (intron 3 G→A). In vitro expression studies revealed that the A250V mutation markedly reduced ADAMTS13 activity and the intron 3 G→A mutation caused abnormal mRNA synthesis. (Blood. 2004;104: 2081-2083)

Introduction

ADAMTS13 (a disintegrin and metalloprotease with thrombospondin type I domains 13) is a metalloprotease that specifically cleaves the multimeric von Willebrand factor (VWF).[1-5] VWF is a large glycoprotein essential for high-shear stress-associated platelet adhesion and aggregation.[6] VWF is synthesized in vascular endothelial cells[7-9] and released into the plasma as unusually large multimeric forms (UL-VWFM).[10] In the presence of excess amounts of UL-VWFM, the more effective platelet interaction with VWF multimers is thought to foster thrombus formation in the microcirculation.[11]

Severely deficient ADAMTS13 activity (less than 5% of that in normal plasma) caused either by mutations of the *ADAMTS13* gene[2,12-15] or by inhibitory antibodies against ADAMTS13[16-18] is linked to thrombotic thrombocytopenic purpura/hemolytic uremic syndrome (TTP/HUS). In this article, we report 2 novel ADAMTS13 mutations in a Japanese patient with TTP who had reduced plasma ADAMTS13 activity. In vitro studies indicated that the 2 mutations affected the enzymatic activity and mRNA splicing, which would account for the patient's phenotype.

Study design

Patient

All experiments were performed with the permission of the ethics committees of both the sample-collecting hospital (Yamada Red Cross Hospital, Watarai-gun, Japan) and the gene-analyzing institute (Toyobo, Osaka, Japan and TAKARA BIO, Otsu, Japan). The patient was a 51-year-old Japanese male. He had a past history of thrombocytopenia of unknown etiology, but now it is impossible to identify whether he had such episodes during his early childhood. He developed convulsive seizures after surgery for hemorrhoids in March 2000. He was diagnosed with TTP because of progressing renal failure and decreased platelet counts. Fresh-frozen plasma (FFP) infusion was effective. After recovery, he was observed in an outpatient clinic with no maintenance FFP infusion. In March 2001, TTP symptoms again developed without any triggering episodes. He was treated with FFP for several consecutive days. After recovery, however, the patient stopped visiting the hospital. In August 2001, he was hospitalized because of the recurrence of TTP. No apparent triggering infection or trauma was documented for this episode. Laboratory data on admission were as follows: red blood cells (RBCs), 2.12×10^{12}/L (212×10^4/μL); hemoglobin (Hb), 68 g/L (6.8 g/dL), white blood cells (WBCs), 3.6×10^9/L (3600/μL); platelets (PLTs), 70×10^9/L (7.0×10^4/μL); reticulocytes, 130×10^9/L (13×10^4/μL); lactate dehydrogenase (LDH), 414 U/L; total bilirubin (T-bilirubin), 37.62 μmol/L (2.2 mg/dL); blood urea nitrogen (BUN), 44 mg/dL; and creatinine, 335.92 μmol/L (3.8 mg/dL). FFP infusion was again effective, and the patient regularly received 4 units of FFP infusion every week at the outpatient clinic, by which hemolysis and thrombocytopenia were well controlled. In January 2002, he was hospitalized because of abdominal pain with gallstone disease. No apparent biliary tract infection was documented at this point, but cholecystectomy was performed to control pain. After surgery, thrombotic microangiopathy with severe thrombocytopenia developed, which was refractory to FFP infusion. Gastrointestinal bleeding occurred, and the patient died of renal failure in July 2002. The plasma ADAMTS13 activity of this patient during an acute episode was less than 3% of normal, and ADAMTS13

From New Product Research Laboratories II, Daiichi Pharmaceutical, Tokyo, Japan; Department of Laboratory Medicine, Mie University School of Medicine, Tsu, Japan; Department of Internal Medicine, Yamada Red Cross Hospital, Watarai-gun, Japan; Department of Medicine, Keio University, Tokyo, Japan; First Research Department, The Chemo-Sero-Therapeutic Research Institute, Kumamoto, Japan; and Department of Blood Transfusion, Nara Medical University, Kashihara, Japan.

Submitted February 26, 2004; accepted April 15, 2004. Prepublished online as *Blood* First Edition Paper, May 4, 2004; DOI 10.1182/blood-2004-02-0715.

Reprints: Mitsuru Murata, Department of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo, Japan; e-mail: murata@sc.itc.keio.ac.jp.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

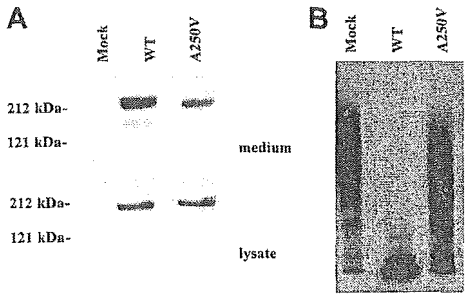


Figure 1. Effect of the A250V mutation on ADAMTS13 activity. The wild-type (WT) and A250V mutant ADAMTS13 were expressed in HEK293 cells. (A) The culture media and cell lysate were analyzed by Western blot. The molecular sizes as judged by protein standards are indicated at the left. (B) After normalization of ADAMTS13 protein concentration, the culture media were analyzed for the enzymatic activity by measuring the degradation of VWF multimers. The A250V mutant had markedly reduced activity compared with wild type.

inhibitor was not detected. The ADAMTS13 activities of his parents were both 46%, and both were asymptomatic.

ADAMTS13 gene analysis

Human genomic DNA was isolated from peripheral leukocytes. The 29 exons and exon-intron boundary sites of the *ADAMTS13* gene were analyzed using an ABI3100 sequencer (Applied Biosystems, Foster City, CA). The frequencies of newly discovered mutations were investigated using genomic DNA from the general Japanese population without a history of TTP/HUS.

Expression of recombinant ADAMTS13

ADAMTS13 cDNA (GenBank accession no. NM139025) was polymerase chain reaction (PCR) amplified from a human fetal liver cDNA library (BioChain Institute, Hayward, CA) and cloned into pcDNA3.1/myc-His (Invitrogen, Carlsbad, CA). Construction of mutants was performed using PCR-based mutagenesis. Each expression vector was transfected into HEK293 cells using FuGENE6 (Roche Molecular Biochemicals, Indianapolis, IN). After a 24-hour incubation, the medium was changed to serum-free OPTI-MEM I (Invitrogen), and the cells were cultured for 48 hours. The media were analyzed for protein expression with Western blot using anti-ADAMTS13 monoclonal antibodies.[19] Wild-type (WT) and A250V mutant culture media were concentrated appropriately so that the 2 culture media had an equal concentration of ADAMTS13 protein. Assays of ADAMTS13 activity were performed using the method of Furlan et al[16,20] with a slight modification as described previously.[21]

Minigene expression system

The exon 3, intron 3, and exon 4 of *ADAMTS13* genomic DNA was cloned into pcDNA3.1 (Invitrogen). The vector was transfected into HEK293 cells. After 48 hours of incubation, we extracted total RNA and performed reverse transcription (RT)–PCR.

Results and discussion

Genetic analysis of the *ADAMTS13* gene

We analyzed the *ADAMTS13* gene in the patient with TTP and his mother. Two novel mutations, A250V and intron 3 G→A, were identified. A250V localizes in the metalloprotease domain, caused by a missense mutation in exon 7. Intron 3 G→A was a guanine to adenine substitution at the 5′ end of intron 3 in the *ADAMTS13* gene. The patient was heterozygous for both mutations. His mother had the heterozygous intron 3 G→A mutation but did not have the A250V mutation. Genomic DNA of his father was not available. These results suggested that the patient was a compound heterozygote for A250V and intron 3 G→A. The frequencies of these mutations were investigated in the general Japanese population. Although 66 individuals were analyzed for A250V and 772 individuals were analyzed for intron 3 G→A, we found no individuals with these mutations.

Effect of A250V mutation on ADAMTS13 activity

To determine the effect of the A250V mutation on ADAMTS13 activity, WT and A250V were expressed in HEK293 cells. Lysate of A250V-transfected cells contained basically the same amount of ADAMTS13 as WT (Figure 1A). On the other hand, in culture media, the amount of A250V antigen was less than that of WT, indicating that the mutation affected secretion or stability of the protein. After normalization of protein concentration for WT and A250V, enzymatic activities were measured. A250V had markedly reduced activity as compared with WT (Figure 1B). A250 localizes in the Zn^{2+} binding domain, which is highly conserved in ADAMTS families. In particular, the M249 residue is essential for Zn^{2+} binding.[22,23] The A250V mutation might influence the structure of the Zn^{2+} binding domain, resulting in reduced protease activity.

Effect of intron 3 G→A mutation on RNA splicing

The intron 3 G→A mutation was found at the 5′ border of intron 3 in the *ADAMTS13* gene. The guanine at this position is highly conserved in general exon-intron boundary sites, and most RNA splicing occurs due to recognition of the GT-AG motif at both ends of the intron. It was expected that this mutation would affect RNA splicing. We constructed the minigene expression system because

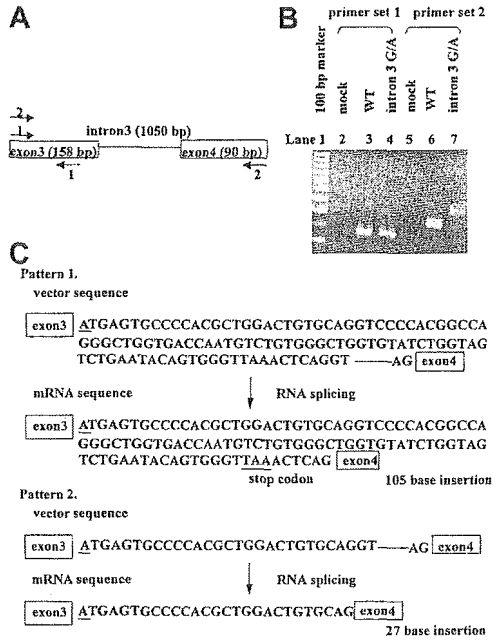


Figure 2. The effect of the intron 3 G→A mutation on RNA splicing. To evaluate the effect of the intron 3 G→A mutation on RNA splicing, we transfected minigene expression vectors into HEK293 cells. Total RNA was purified, which was subjected to RT-PCR analysis. (A) Primer sets for the PCR analysis are shown. Set 1 amplified exon 3 (158 bp), whereas set 2 amplified exon 3 and exon 4 (248 bp). (B) Results of RT-PCR. Primer set 2 yielded the 248-bp band from WT-transfected cells (lane 6), and primer set 2 yielded different products from intron 3 G→A mutant–transfected cells (lane 7). (C) Sequence analysis of the transcript from intron 3 G→A mutant–transfected cells. Underlines indicate the position of the G→A mutation. Forty-nine clones were analyzed: 38 clones had a 105-bp insert, and 11 clones had a 27-bp insert. Each variant was spliced with recognition at the GT-AG motif.

mRNA of the patient was not available. The transcription products were analyzed by RT-PCR using 2 primer sets (Figure 2A). The WT-expressing cells produced normally spliced RNA, whereas the intron 3 G→A mutant produced splicing variants (Figure 2B). To further analyze the aberrantly spliced RNA products, RT-PCR products of intron 3 G→A mutant were subcloned into pDrive cloning vector (Qiagen, Valencia, CA), and 49 colonies were picked for sequence analysis. As a result, 2 splicing variants were detected (Figure 2C). One variant in 38 of 49 clones had a 105-nucleotide insertion containing an in-frame stop codon between exon 3 and exon 4, which would result in the synthesis of a truncated form that stops at the metalloprotease domain. The remaining 11 clones had a 27-bp insertion, which would lead to a 9–amino acid polypeptide-inserted phenotype.

These results indicate that A250V and intron 3 G→A mutations might be the cause of the decreased plasma ADAMTS13 activity in this patient, the former abolishing ADAMTS13 enzymatic activity and the latter impairing normal mRNA synthesis.

References

1. Soejima K, Mimura N, Hirashima M, et al. A novel human metalloprotease synthesized in the liver and secreted into the blood: possibly, the von Willebrand factor-cleaving protease? J Biochem. 2001;130:475-480.
2. Levy GG, Nichols WC, Lian EC, et al. Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura. Nature. 2001;413:488-494.
3. Gerritsen HE, Robles R, Lämmle B, Furlan M. Partial amino acid sequence of purified von Willebrand factor-cleaving protease. Blood. 2001;98:1654-1661.
4. Fujikawa K, Suzuki H, McMullen B, Chung D. Purification of human von Willebrand factor-cleaving protease and its identification as a new member of the metalloproteinase family. Blood. 2001;98:1662-1666.
5. Zheng X, Chung D, Takayama TK, et al. Structure of von Willebrand factor-cleaving protease (ADAMTS13), a metalloprotease involved in thrombotic thrombocytopenic purpura. J Biol Chem. 2001;276:41059-41063.
6. Ruggeri ZM. Structure and function of von Willebrand factor. Thromb Haemost. 1999;82:576-584.
7. Furlan M. Von Willebrand factor: molecular size and functional activity. Ann Hematol. 1996;72:341-348.
8. Ruggeri ZM. Von Willebrand factor. J Clin Invest. 1997;99:559-564.
9. Sadler JE. Biochemistry and genetics of von Willebrand factor. Annu Rev Biochem. 1998;67:395-424.
10. Federici AB, Bader R, Pagani S, et al. Binding of von Willebrand factor to glycoproteins Ib and IIb/IIIa complex: affinity is related to multimeric size. Br J Haematol. 1989;73:93-99.
11. Moake JL, Turner NA, Stathopoulos NA, et al. Involvement of large plasma von Willebrand factor (vWF) multimers and unusually large vWF forms derived from endothelial cells in shear stress-induced platelet aggregation. J Clin Invest. 1986;78:1456-1461.
12. Kokame K, Matsumoto M, Soejima K, et al. Mutations and common polymorphisms in ADAMTS13 gene responsible for von Willebrand factor-cleaving protease activity. Proc Natl Acad Sci U S A. 2002;99:11902-11907.
13. Antonie G, Zimmermann K, Plaimauer B, et al. ADAMTS13 gene defects in two brothers with constitutional thrombotic thrombocytopenic purpura and normalization of von Willebrand factor-cleaving protease activity by recombinant human ADAMTS13. Br J Haematol. 2003;120:821-824.
14. Schneppenheim R, Budde U, Oyen F, et al. von Willebrand factor cleaving protease and ADAMTS13 mutations in childhood TTP. Blood. 2003;101:1845-1850.
15. Matsumoto M, Kokame K, Soejima K, et al. Molecular characterization of ADAMTS13 gene mutations in Japanese patients with Upshaw-Schulman syndrome. Blood. 2004;103:1305-1310.
16. Furlan M, Robles R, Galbusera M, et al. von Willebrand factor-cleaving protease in thrombotic thrombocytopenic purpura and the hemolytic-uremic syndrome. N Engl J Med. 1998;339:1578-1584.
17. Tsai HM, Lian EC. Antibodies to von Willebrand factor-cleaving protease in acute thrombotic thrombocytopenic purpura. N Engl J Med. 1998;339:1585-1594.
18. Tsai HM. High titers of von Willebrand factor-cleaving metalloprotease in a fetal case of acute thrombotic thrombocytopenic purpura. Am J Hematol. 2000;65:251-255.
19. Suzuki M, Murata M, Matsubara Y, et al. Detection of von Willebrand factor-cleaving protease (ADAMTS-13) in human platelets. Biochem Biophys Res Commun. 2004;313:212-216.
20. Furlan M, Robles R, Lamie B. Partial purification and characterization of a protease from human plasma cleaving von Willebrand factor to fragments produced by in vivo proteolysis. Blood. 1996;87:4223-4234.
21. Kinoshita S, Yoshioka A, Park YD, et al. Upshaw-Schulman syndrome revisited: a concept of congenital thrombotic thrombocytopenic purpura. Int J Hematol. 2001;74:101-108.
22. Bode W, Gomis-Ruth FX, Stockler W. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. FEBS Lett. 1993;331:134-140.
23. Stocker W, Grams F, Baumann U, et al. The metzincins—topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. Prot Sci. 1995;4:823-840.

Genetic analyses and expression studies identified a novel mutation (W486C) as a molecular basis of congenital coagulation factor XII deficiency

Keiko Ishii[a], Shuji Oguchi[a], Takanori Moriki[b], Yoko Yatabe[a], Eiko Takeshita[a], Mitsuru Murata[a,c], Yasuo Ikeda[c] and Kiyoaki Watanabe[a]

We analyzed the factor XII (FXII) gene of a patient with congenital FXII deficiency and identified a novel amino acid substitution (W486C) in the catalytic domain. The proband was an asymptomatic 49-year-old Japanese female with abnormal coagulation test, discovered by chance. The FXII activity and antigen level were both under 10%, suggesting a cross-reacting material-negative FXII deficiency. Sequence analysis of the proband's FXII gene revealed a homozygous nucleotide substitution G → C in exon 12, resulting in the amino acid substitution W486C in the catalytic domain. We constructed the mutant FXII cDNA in an expression plasmid vector and transfected it into Chinese hamster ovary cells. The recombinant wild-type FXII antigen was detected in the culture medium by immunoprecipitation assay, but the mutant FXII (W486C) was not observed. On the other hand, both the wild-type FXII and W486C cell lysates contained FXII antigen and FXII mRNA, as estimated by western blotting and quantitative reverse transcriptase-polymerase chain reaction. These findings suggest that the W486C substitution of FXII impairs intracellular processing of the protein and/or transport system. *Blood Coagul Fibrinolysis* **15:367–373 © 2004 Lippincott Williams & Wilkins.**

Blood Coagulation and Fibrinolysis 2004, 15:367–373

Keywords: blood coagulation, factor XII, deficiency, expression study, mutation, genetics

[a]Department of Laboratory Medicine, [b]Health Center and [c]Department of Internal Medicine, Division of Hematology, School of Medicine, Keio University, Tokyo, Japan.

Correspondence and requests for reprints to Mitsuru Murata, MD, Department of Internal Medicine, Division of Hematology, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan.
Tel: +81 3 5363 3785; fax: +81 3 3353 3515; e-mail: murata@sc.itc.keio.ac.jp

Received 8 July 2003 **Revised** 19 January 2004
Accepted 26 January 2004

Introduction

Human blood coagulation factor XII (FXII) (Hageman factor) is a serine protease precursor, circulating in the plasma as a zymogen. FXII is a single-chain 80-kDa glycoprotein, composed of 596 amino acids, that is primarily produced by hepatocytes and secreted to the plasma at a level of approximately 30 μg/ml (range 15–47 μg/ml). FXII has a role in the coagulation pathway, fibrinolysis, and the complement system. The protein structure is composed of a COOH terminal catalytic domain, type I and type II domains, two growth factor domains, and a kringle domain. On the negatively charged surface, cleavage by kallikrein at Arg353-Val354 generates activated factor XII (FXIIa), thereby initiating the rapid intrinsic coagulation pathway by converting factor XI to activated factor XI [1–4]. The FXII gene is 12 kb, consists of 14 exons, and is located at 5q33-qter [5,6]. FXII mRNA is approximately 2 kb long, including 150 bases of 3′-non-coding sequences. The transcription is initiated 49 bases upstream from the translation initiation ATG codon [7,8].

Congenital FXII deficiency is a rare coagulation disorder, found in many cases by chance through a prolonged activated partial thromboplastin time prior to surgery. The patients demonstrate no obvious clinical symptoms such as a bleeding tendency, although some reports describe the possibility of a thromboembolic predisposition [9–11]. FXII deficiency is basically divided into two categories, a cross-reacting material (CRM)-negative group (negative FXII antigen detection) and a CRM-positive group (positive FXII antigen detection). In CRM-negative deficiency, Y34C (FXII Tenri) [12] and R123P are found in homozygous patients and Q421K in heterozygous patients [13]. C571S (FXII Washington DC) [14] and R353P (FXII Locarno) [15] are reported as a CRM-positive deficiency. A common genetic polymorphism, 46 C → T in exon 1 at the 5′-untranslated region, four bases upstream of initiation ATG codon, is associated with a low translation efficiency and a decrease in the plasma FXII level [16].

Here, we present a case of CRM-negative FXII deficiency and describe the results of genetic and expression studies that revealed a molecular etiology of the disorder.

0957-5235 © 2004 Lippincott Williams & Wilkins

DOI: 10.1097/01.mbc.0000114447.59147.d1

Materials and methods

Patient characteristics

The proband was a 49-year-old Japanese woman, who was identified by chance to have prolonged activated partial thromboplastin time. No bleeding or thrombotic tendencies were observed. Laboratory tests indicated that the FXII antigen level and FXII clotting activity were both less than 10%, and circulating and lupus anticoagulant were both negative, indicating a diagnosis of CRM-negative FXII deficiency. Her two children had reduced FXII antigen levels of 28 and 43%.

FXII gene analysis

Peripheral venous blood was collected from the proband and her family members after obtaining written informed consent. DNA fragments corresponding to all exons and exon/intron boundaries of the FXII gene were amplified directly from whole blood by polymerase chain reaction (PCR) using Ampdirect (Shimadzu, Kyoto, Japan) [17]. The PCR primers are presented in Table 1. In a total volume of 50 μl, 0.5 μl whole blood, 1.3 μl sense and antisense primers, 10 μl Ampdirect, 10 μl AmpAddition, 4 μl of 10 mmol/l dNTPs, and 0.3 μl Taq polymerase were mixed and amplified. Five microliters of 5% dimethylsulfoxide/0.1% Triton was added to exons 8, 9, 10, 11, and 12, and 10 μl GC-Melt (BD Biosciences Clontech, Palo Alto, California, USA) was added to enhance the reaction for exon 10. The annealing temperature was 62°C, except for exon 9 (64°C). The PCR products were gel-purified and direct DNA sequence analysis was performed with BigDye Terminator using the ABI PRISM 377 DNA Sequencer (Applied Biosystems, Foster City, California, USA).

PCR-restriction fragment length polymorphism

After PCR amplification of the DNA fragments (547 bp) including the site of one point mutation in exon 12 (primers are presented in Table 1), the products were digested with restriction enzyme Cfr13I and electrophoresed in an agarose gel. In addition to the proband and her two children, normal subjects ($n = 95$) were simultaneously analyzed using the same procedure.

Plasmid construction

Full-length human FXII cDNA was kindly provided by Dr Ross T. A. MacGillivray (University of British Columbia, Vancouver, British Columbia, Canada). Wild-type (WT) FXII cDNA was cloned into pBluescript (Stratagene, La Jolla, California, USA) after removing the *Kpn*I site from the multiple cloning site. PCR-based mutagenesis was performed by mutagenic primers (sense, 5′-AGCGCGGAACTGGGGACTGG; antisense, 5′-AGGAACGGTACCTGCGCCTCCTGC AGGAAGCTGGCATATTCCTCCGCCCCCTCGAAC TGGTGGCCGCAGCCGGCCA, underlined G is a mutated base) using Platinum Pfx DNA polymerase with enhancer (Gibco BRL, Rockville, Maryland, USA). The constructed mutant FXII cDNA (W486C) was excised at the *Not*I site and cloned into the expression vector PCI^{neo} (Promega, Madison, Wisconsin, USA) and the DNA sequence was confirmed.

Table 1 Polymerase chain reaction (PCR) primers for amplification of exons and exon/intron boundaries of FXII gene

Exon	Sequence (5′-3′)	Nucleotide Number	PCR products bp
1*	CCA GTC CCA CTA TCT AGA AAA G	−195−−174	369
	ATG GCT CAT GGC TGT GAT AG	154−173	
2	GGC AGA GGA GAC AGA CAA CCA GAC	4511−4544	568
	GGA GAA GCT GGG GAA ACT GAG GAA	5055−5078	
3,4	TCA GAG AGT GTG TTG TCC CTG CAG	7767−7790	411
	AGG AGA GTA ATG AGG CGG GAG GAG	8154−8177	
5	TAG CAA GGT CAC AAG GCA AGT AGG	8322−8345	292
	CTC CGA GTA TCC AGC AAC CTA TTC	8590−8613	
6	TAG GTT GCT GGA TAC TCG	8593−8610	223
	TCC CAA CCA TCT GCT CAC	8798−8815	
7	CAG GGA GGA GCG TCA GGA	8822−8839	239
	GCC CCA GAC CCT CAC TCA	9043−9060	
8	GTG AGT GAG GGT CTG GGG CAA GCA	9042−9065	434
	GCA CGA AGC ACC ACG GGC GGA TGT	9452−9475	
9	GGA ACC CGG ACA ACG ACA	9437−9454	375
	AGT GGG CCG TTC CTG GTC	9794−9811	
10**	(1st) AGC AAG CGC GGA ACT GGG GAC TGG	9310−9332	795 (1st)
	(1st) AGA AGG GTG GGG CGG GAG AAG GTA	10081−10104	
	(2nd) GAA GGA GGA GCC GAG AGG	9673−9690	432 (2nd)
	(2nd) AGA AGG GTG GGG CGG GAG AAG GTA	10081−10104	
11,12	CAC GGG ATT GGG GTT CGG GAG CAG	10144−10167	547
	GGC ACC CGG AAC GAT ACC AAA GTC	10667−10690	
13,14	GAA GGG CAT GAG TGG GTT TAC AAG	11047−11070	549
	GCG GAA TCA CCA AGG AGG GAA AGA	11572−11595	

*Kanaji *et al.* [16].
**Nested PCR.

Transient expression of FXII WT and W486C in Chinese hamster ovary cells

Chinese hamster ovary cells were grown in 10% fetal calf serum/Dulbecco's modified Eagle's medium (Gibco BRL) under 5% CO_2, and transfected using FuGENE6 (Roche Diagnostics, Basel, Switzerland) and 1 μg plasmid DNA (PCI^{neo}, PCI^{neo}/FXII-WT, PCI^{neo}/FXII-W486C) in six-well culture dishes. For transient studies, culture medium was replaced with serum-free Dulbecco's modified Eagle's medium after 24 h of transfection and the supernatant was collected after 24 h incubation. Cell lysates were also collected after solubilization in sodium dodecyl sulfate (SDS) sample buffer [62.5 mmol/l Tris–HCl (pH 6.8), 2% SDS, 10% glycerol, 0.01% Bromophenol Blue].

Immunoprecipitation and western blotting

Recombinant FXII secreted into serum-free medium was immunoprecipitated using three FXII antibodies (gout anti-FXII CL20055A, Cedarlane Laboratories, Ontario, Canada; GAFXIIIG, Affinity Biology Inc., Ontario, Canada; mouse IgG1 (B7C9) MAB2084, Chemicon International Inc., Temecula, California, USA) with Immobilized Protein A beads (ImmunoPure; Pierce, Rockford, Illinois, USA). In a total volume of 1000 μl, collected serum-free medium was mixed with 2.5 μl each antibody and 100 μl Immobilized Protein A beads solution, and rotated at 4°C for 60 min. The combined beads were collected, washed three times with ice-cold washing buffer [1% Triton X-100, 10 mmol/l Tris–HCl (pH 7.5), 150 mmol/l NaCl, 0.1% SDS], and mixed with 100 μl SDS sample buffer (with 1% 2-Mercaptoethanol). The transfected cells were directly solubilized with SDS sample buffer. After boiling for 5 min, samples were applied onto 5–20% gradient SDS polyacrylamide gels and electrophoresed. The separated proteins were transferred to polyvinylidene difluoride membranes (BioRad, Hercules, California, USA), blocked in 5% skim milk/0.05% Tween-20/TBS (20 mmol/l Tris–HCl, 500 mmol/l NaCl; pH 7.4) and incubated with horseradish peroxidase-conjugated anti-FXII antibody (CL20055HP; Cedarlane Laboratories) in 1:1000 dilution. The membranes were washed three times with Tween-20/TBS and labeled bands were detected using ECL Western Blotting Detection Reagents (Amersham, Little Chalfont, UK).

Quantitative reverse transcriptase-PCR for FXII mRNA expression

Transfected cells were subjected to FXII mRNA expression level analysis by quantitative reverse transcriptase (RT)-PCR with a TaqMan probe. Total RNA was purified from transiently transfected cells using an RNeasy Mini Kit with QIAshredder spin column and an RNase-Free DNase Set (Qiagen, Hilden, Germany). The purity of RNA was confirmed by Agilent Bioanalyzer (Agilent Technologies, Palo Alto, California, USA). The appropriate PCR primers (FXII-TaqF, 5′-CAG CTGTACCACAAATGTACCCAC; FXII-TaqR, 5′-AA ACAGTATCCCCATCGCTGG) and the TaqMan probe (TaqMan-FXII, TGTGCTACCACCCCCAACT TTGATCAG) were designed by the manufacture's software (Applied Biosystems). Quantitative RT-PCR was performed using a Platinum quantitative RT-PCR ThermoScript one-step system (Invitrogen, Carlsbad, California, USA) and ABI PRISM 7700 (Applied Biosystems) according to the manufacturer's protocol. Total RNA from WT was serially diluted and used as a template for quantitative RT-PCR to generate the standard curve. An adequate fluorescent intensity was chosen and the PCR cycles required to reach that point were determined as the threshold cycle. Results were from three independent transient transfection assays with triplicate RNA samples. The data were statistically analyzed using Student's *t* test.

Results and discussion

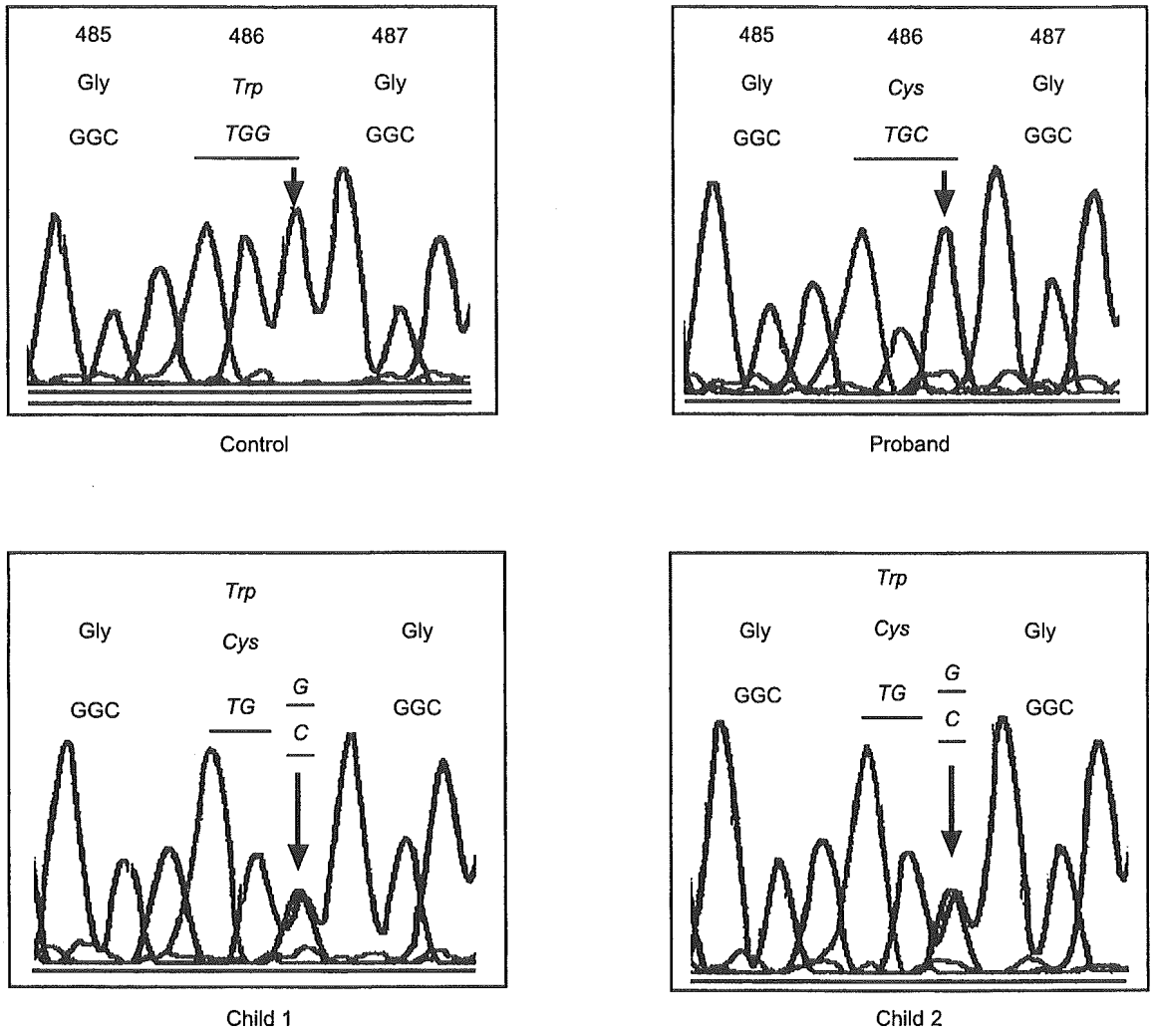
Sequence analysis of all exons and exon/intron boundaries of the FXII gene revealed that the proband was homozygous for one base substitution (G → C) in exon 12, which results in the substitution Trp486 (TGG) to Cys (TGC) (Fig. 1). The proband's two children were thought to be heterozygous for this substitution. To confirm the detected substitution as a mutation, PCR products including the mutation site from the proband, proband's two children, and control subjects ($n = 95$) were digested with restriction enzyme Cfr13I for restriction fragment length polymorphism analysis. Cfr13I digested the 547-bp PCR product into 177, 170, 112, and 89 bp in the controls and 282, 177, and 89 bp in the proband. The electrophoresis patterns of the digested fragments reconfirmed the proband's homozygosity and the children's heterozygosity for the mutation (Fig. 2).

This was a novel mutation not yet reported in congenital FXII deficiency. In addition, we checked the polymorphism of the FXII gene promotor site, 46C/T, which is considered to affect the plasma FXII antigen level. The proband and the two children were 46T/T type, common in Japanese, which is associated with a lower plasma FXII antigen level [16]. During this study, however, another Japanese family with FXII deficiency of the same mutation was reported [18].

The mutation results in the amino acid substitution Trp → Cys at position 486 (W486C), which is located in the catalytic domain of FXII protein. As a member of serine proteases, the structure formed by the catalytic triad His393-Asp442-Ser544 is critical for the enzyme activity [19].

We therefore performed expression analysis to elucidate the molecular basis of the proband's disorder. After introducing the mutation into FXIIcDNA, the

Fig. 1



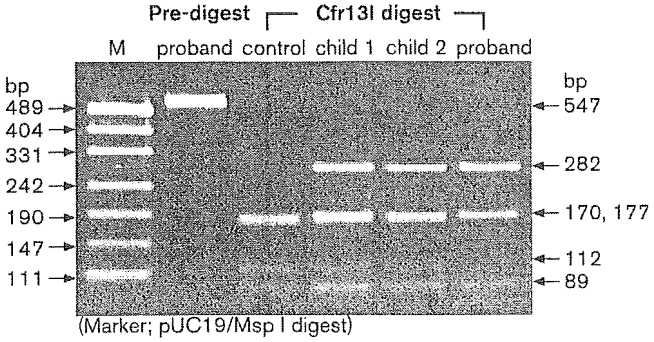
Sequence analysis of the FXII gene mutation in the proband and proband's children. Samples from healthy subjects were used as controls. The proband was homozygous for a G → C mutation in exon 12, which results in the amino acid substitution Trp486 (TGG) to Cys (TGC). The proband's two children were thought to be heterozygous for the mutation.

constructs (designated W486C) were inserted into the expression plasmid vector pCI^{neo} and transfected into CHO cells. Plasmid vector (negative control) and WT were simultaneously transfected.

First, we directly examined the FXII antigen secreted to serum free supernatant by enzyme-linked immunosorbent assay. However, the antigen was not detected in WT medium from transient or stable transfectants (data not shown). Because the FXII antigen level secreted from transfectants was considered to be low, we performed immunoprecipitation assays to compare the secreted FXII antigen level between WT and W486C. Using three anti-FXII antibodies, FXII antigen secreted into the supernatant of the culture medium was immunoprecipitated with protein A beads and the transfected cells were simultaneously solubilized. The immunoprecipitates and cell lysates were separated by SDS-polyacrylamide gel electrophoresis, transferred to polyvinylidene difluoride membranes, and used for western blotting (Fig. 3). In the supernatant, the FXII antigen was detected only in WT but not in W486C, a finding compatible with the CRM-negative phenotype. In cell lysates, we observed the FXII antigen both in WT and W486C. The difference in recombinant FXII molecular weights between supernatants and cell lysates is probably due to a post-translational modification, such as carbohydrate addition.

We then investigated the mRNA expression levels of WT and W486C in transient transfectants by quantitative RT-PCR. The difference in the threshold cycle

Fig. 2

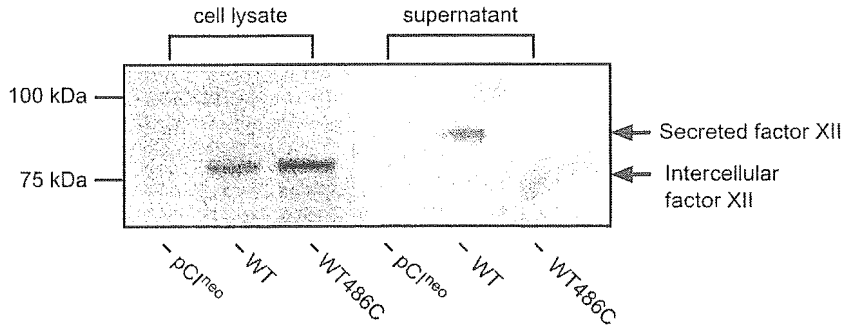


Polymerase chain reaction (PCR)-restriction fragment length polymorphism. PCR was performed on genomic DNA of the proband, the proband's two children, and normal subjects ($n = 95$) to amplify a DNA fragment containing the mutation site in exon 12. PCR products were then digested with restriction enzyme Cfr13I. Pre-digested PCR product (547 bp) of the proband is shown in the second lane, which is representative of the pre-digested PCR product of all subjects shown here. Cfr13I digested the 547 bp PCR product into 177, 170, 112, and 89 bp in the controls and 282, 177, and 89 bp in the proband, respectively. The electrophoresis patterns of digested fragments indicated that the proband was homozygous and two children were heterozygous for the mutation. None of the normal subjects showed the mutant pattern.

between WT and W486C was approximately 0.9, indicating a slightly increased mRNA expression level of W486C (1.7-fold) compared with that of WT ($P < 0.0001$, Student's t test; Fig. 4). These observations suggest that the W486C-expressing cells did not have less intracellular mRNA than the WT transfectants, and that the W486C protein is normally synthesized at the endoplasmic reticulum. Because our *in vitro* studies utilized a transient expression system, however, the meaning and significance of the slightly elevated mRNA expression level in W486C should not be overestimated.

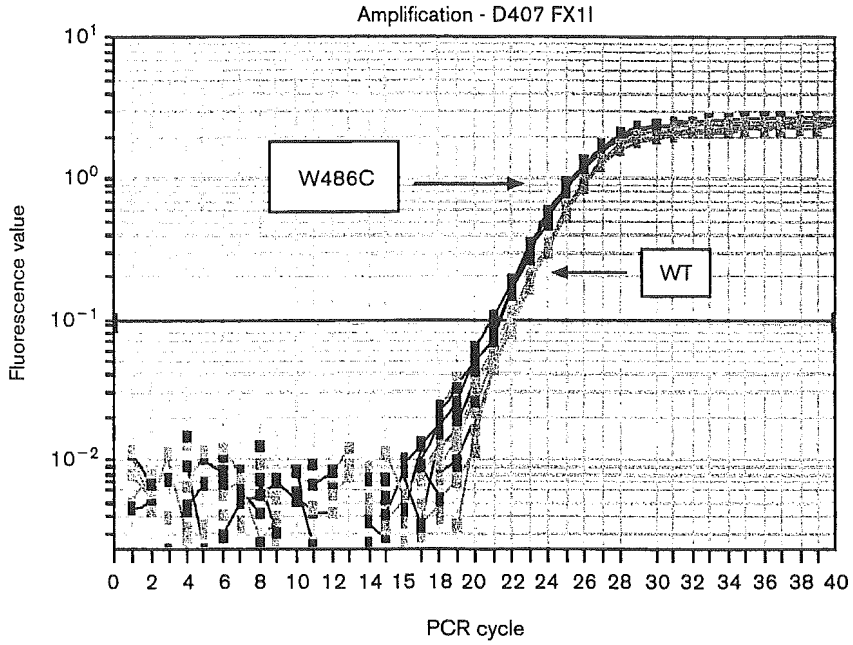
Several gene mutations have been reported as molecular bases of CRM-negative FXII deficiency. Screening of mutations in the human FXII gene of 31 unrelated FXII-deficient patients identified six mutations associated with CRM-negative FXII deficiency [19], although expression analysis was not performed. Three CRM-negative mutations were reported with expression analysis. Among them, FXII Tenri [12] exhibited a Y34C substitution. It is interesting to compare the protein characteristics between FXII Tenri and the present case because both mutants exhibited a replacement by cysteine. In FXII Tenri, the additional Cys in the NH_2-terminal type II domain is related mostly to proteasome-mediated degradation, and a small amount of the residual FXII Tenri was secreted to the plasma combined with α_1-microglobulin, which was connected to the mutated Cys34. There are 20 tentative disulfide bonds in FXII protein based on homologies with other proteins and an additional unpaired Cys residue would result in misfolding of the protein during protein processing in the endoplasmic reticulum. Two other mutations, Q421K and R123P, were examined by expression analysis of CRM-negative FXII [13]. Q421K is likely to accumulate in the endoplasmic reticulum due to an impaired transport system, whereas R123P is susceptible to proteasome degradation. Interestingly, FXII mutations in the catalytic domain could have both CRM-negative and CRM-positive phenotypes. FXII Washington DC (C571S) [14], FXII Locarno (R353P) [15], FXII D442N, and G570R are related to the CRM-positive phenotype [19]. Whether the mutants were degraded or secreted

Fig. 3



Western blot of recombinant factor XII (FXII) detected in the culture medium and cell lysates. Plasmid constructs (PCI^{neo}, PCI^{neo}/FXII-WT, PCI^{neo}/FXII-W486C) were transiently transfected into an equal number of Chinese hamster ovary cells. The supernatants containing secreted recombinant FXII were immunoprecipitated by anti-FXII antibodies. Transfected cells were solubilized with sodium dodecyl sulfate (SDS) sample buffer. All samples were subjected to SDS-polyacrylamide gel electrophoresis, transferred onto polyvinylidene difluoride membranes, and blotted by anti-FXII antibody. Only FXII-wild-type (WT) was detected in the supernatant. In contrast, both FXII-WT and W486C were detected in the cell lysates.

Fig. 4



Comparison of mRNA expression levels between wild-type (WT) and W486C. Total RNA was purified from Chinese hamster ovary cells transiently transfected by plasmid constructs (PCI^{neo}, PCI^{neo}/FXII-WT, PCI^{neo}/FXII-W486C). Quantitative reverse transcriptase-polymerase chain reaction (RT-PCR) was performed using a TaqMan probe and appropriate PCR primers using the Platinum quantitative RT-PCR ThermoScript one-step system (Invitrogen) and the ABI PRISM 7700 (Applied Biosystems). Total RNA from WT was serially diluted and used as a template for quantitative RT-PCR to generate the standard curve. An adequate fluorescent intensity was chosen and the PCR cycles to reach the point were determined as the threshold cycle. A representative result is shown from three independent transient transfection assays. The difference in the threshold cycle between WT and W486C was approximately 0.9, estimating a slightly increased mRNA expression level of W486C (1.7-fold) compared with that of WT ($P < 0.0001$, Student's t test).

would depend on the misfolded conformation induced by the mutation, which would be recognized by the protein quality control system in the endoplasmic reticulum [20–22].

In conclusion, we identified a novel mutation, FXII W486C, in congenital FXII deficiency and clarified the mutation as an independent molecular basis of the CRM-negative phenotype. Elucidation of the mechanism of the protein degradation requires further investigation of the intracellular protein quality control system.

Acknowledgement

The authors thank Dr Ross T.A. MacGillivray (University of British Columbia, Vancouver, British Columbia, Canada) for providing full-length human FXII cDNA, and Dr Taisuke Kanaji (Kyusyu University, Fukuoka, Japan) for providing plasmid constructs and useful technical advice.

References

1 Kluft C, Dooijewaard G, Emeis JJ. Role of the contact system in fibrinolysis. *Semin Thromb Hemost* 1987; **13**:50–68.
2 Magnusson S, Sottrup-Jensen L, Petersen TE, Dudeke-Wojciechowska G, Claeys H. Homologous kringle structures common to plasminogen and prothrombin. Substrate specificity of enzymes activating prothrombin and plasminogen. In: Ribbons DW, Brew K (editors): *Proteolysis and physiological regulation*, vol. 11. New York: Academic Press; 1976, p. 203.
3 Revak SD, Cochrane CG, Johnston AR, Hugli TE. Structural changes accompanying enzymatic activation of human Hageman factor. *J Clin Invest* 1974; **54**:619–627.
4 Pixley RA, Colman RW. Factor XII: Hageman factor. *Methods Enzymol* 1993; **222**:51–65.
5 Cool DE, MacGillivray RT. Characterization of the human blood coagulation factor XII gene, Intron/exon gene organization and analysis of the 5′-franking region. *J Biol Chem* 1987; **262**:13662–13673.
6 Royle NJ, Nigli M, Cool D, MacGillivray RT, Hamerton JL. Structural gene encoding human factor XII is located at 5q33-qter. *Somat Cell Mol Genet* 1988; **14**:217–221.
7 Cool DE, Edgell CJ, Louie GV, Zoller MJ, Brayer GD, MacGillivray RT. Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor XIIa. *J Biol Chem* 1985; **260**:13666–13676.
8 Tripodi M, Citarella F, Guida S, Galeffi P, Fantoni A, Cortese R. cDNA sequence coding for human coagulation factor XII (Hageman). *Nucl Acids Res* 1986; **14**: 3146
9 Ratnoff OD, Busse RJ, Sheron R. The demise of John Hageman. *N Engl J Med* 1968; **279**:760–761.
10 Goodnough LT, Saito H, Ratnoff OD. Thrombosis or myocardial infarction in congenital clotting factor abnormalities and chronic thrombocytopenias: a report of 21 patients and a review of 50 previously reported cases. *Medicine (Baltimore)* 1983; **62**:248–255.
11 Saito H. Contact factors in health and disease. *Semin Thromb Hemost* 1987; **13**:36–49.

12 Kondo S, Tokunaga F, Kawano S, Oono Y, Kumagai S, Koide T. Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a proteasome-mediated degradation. *Blood* 1999; **93**:4300–4308.
13 Kanaji T, Kanaji S, Osaki K, Kuroiwa M, Sakaguchi M, Mihara K, *et al.* Identification and characterization of two novel mutations (Q421K and R123P) in congenital factor XII deficiency. *Thromb Haemost* 2001; **86**:1409–1415.
14 Miyata T, Kawabata S, Iwanaga S, Takahashi I, Alving B, Saito H. Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor XIIa results from Cys-571-Ser substitution. *Proc Natl Acad Sci USA* 1989; **86**:8319–8322.
15 Hovinga JK, Schaller J, Stricker H, Wuillemin WA, Furlan M, Lammle B. Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg353 → Pro leading to loss of a kallikrein cleavage site. *Blood* 1994; **84**:1173–1181.
16 Kanaji T, Okamura T, Osaki K, Kuroiwa M, Shimoda K, Hamasaki N, *et al.* A common genetic polymorphism (46 C to T substitution) in the 5′-untranslated region of the coagulation factor XII gene is associated with low translation efficiency and decrease in plasma factor XII level. *Blood* 1998; **91**:2010–2014.
17 Nishimura N, Nakayama T, Tomoike H, Kojima K, Kato S. Direct polymerase chain reaction from whole blood without DNA isolation. *Ann Clin Biochem* 2000; **37**:674–680.
18 Kasai Y, Kato K, Hayashi T, Abe Y, Nishioka J, Morishita Y, *et al.* Molecular characterization of coagulation factor XII deficiency in a Japanese family. Abstracts–18th International Congress on Clinical Chemistry, 2002 Kyoto. *Clin Chem Lab Med* 2002; 40, Special Supplement: S195 (abstract).
19 Schloesser M, Zeerleder S, Lutze G, Halbmayer WM, Hofferbert S, Hinney B, *et al.* Mutations in the human factor XII gene. *Blood* 1997; **90**:3967–3977.
20 Helenius A, Aebi M. Intracellular functions of N-linked glycans. *Science* 2001; **291**:2364–2369.
21 Rutishauser F, Spiess M. Endoplasmic reticulum storage disease. *Swiss Med Wkly* 2002; **132**:211–222.
22 Hampton RY. ER-associated degradation in protein quality control and cellular regulation. *Curr Opin Cell Biol* 2002; **14**:476–482.

DOI: 10.1111/j.1464-5491.2004.01252.x

Paraoxonase 1 ^{192}Gln/Arg polymorphism is associated with the risk of microangiopathy in Type 2 diabetes mellitus

M. Murata, T. Maruyama*, Y. Suzuki*, T. Saruta and Y. Ikeda

Department of Medicine, Keio University, Tokyo, and *Department of Medicine, Saitama Social Insurance Hospital, Saitama, Japan

Accepted 15 October 2003

Abstract

Aims To investigate possible associations between diabetic microangiopathy and genetic polymorphisms in factors relevant to arterial thrombosis.

Methods We conducted a case-control study on a total of 280 patients with Type 2 diabetes, comparing those without retinopathy or nephropathy (n = 92) and those with microangiopathies (n = 188), for the association of polymorphisms in four candidate genes, paraoxonase 1 (PON1), plasminogen activator inhibitor-1, fibrinogen, and platelet glycoprotein Ibα.

Results There were no differences between the two study groups in gender distribution, age at diagnosis of diabetes (47.9 ± 8.4 and 49.0 ± 11.4 years, respectively), or duration of diabetes (14.9 ± 4.5 and 14.5 ± 8.4 years, respectively). Among the gene polymorphisms tested, the ^{192}Gln/Arg polymorphism of PON1 was associated with the prevalence of retinopathy [odds ratio (OR) = 3.13, 95% confidence interval (CI) = 1.42–6.89, P = 0.0046, Gln/Gln vs. Gln/Arg and Arg/Arg]. This polymorphism was also associated with nephropathy (OR = 3.01, 95% CI = 1.30–6.98, P = 0.0103). There were no differences between the three PON1 genotypes (Gln/Gln, Gln/Arg, and Arg/Arg) with regard to the present disease status. Logistic regression analysis for the adjustment of other risk factors revealed that genotypes with PON1 ^{192}Arg were an independent predictor of retinopathy. No associations were found between microangiopathies and the other polymorphisms evaluated (plasminogen activator inhibitor-1, fibrinogen, and platelet glycoprotein Ibα).

Conclusions This study suggests that the presence of the ^{192}Arg-allele in the PON1 gene is a genetic risk factor for microangiopathy in Type 2 diabetes mellitus.

Diabet. Med. 21, 837–844 (2004)

Keywords genetics, microangiopathy, paraoxonase 1, polymorphism, Type 2 diabetes

Introduction

The development of diabetic microangiopathy depends mostly on the duration of diabetes and glycaemic control [1]. It has been also noted, however, that diabetic microangiopathy develops in a fraction of diabetic patients, and some patients do not develop clinically evident microangiopathic complications. Familial clustering of microangiopathy is also reported [2–8]. These observations suggest the presence of non glycaemic genetic factors that predispose to or protect against the development of diabetic microangiopathy. Several genetic factors, including HLA typing, angiotensin converting enzyme, angiotensinogen, aldose reductase, and β_3-adrenoreceptor gene polymorphisms, have been reported in association with retinopathy or nephropathy in diabetes mellitus [6–9].

Correspondence to: Mitsuru Murata MD, Department of Medicine, School of Medicine, Keio University, 35 Shinanomachi, Shinjuku-ku, Tokyo 160–8582, Japan. E-mail: murata@sc.itc.keio.ac.jp

In patients with diabetes mellitus, abnormalities in blood flow, activation of blood coagulation/fibrinolysis systems, and enhanced platelet functions have been reported [10–12]. These abnormalities include increased plasma fibrinogen levels, elevated levels of plasminogen activator inhibitor-1 (PAI-1) that inhibit fibrinolysis, and increased platelet adhesiveness and aggregation in response to agonists. Under such thrombophilic circumstances, microthombi might be generated and local circulation impaired, which in turn might cause tissue ischaemia.

In the present study, we assessed the association between diabetic microangiopathy and genetic variations of four factors that have crucial roles in atherosclerosis and/or thrombosis, because atherosclerosis and microthrombosis might predispose patients to impaired microcirculation and cause ischaemic damage to vascular cells. The factors included were (i) paraoxonase 1 (PON1), a calcium-dependent esterase that catalyses reactive oxygen species produced under oxidative stress during atherogenesis and contributes to the prevention of LDL from oxidation [13–15], (ii) PAI-1, a key regulating factor of fibrinolysis, (iii) fibrinogen, the central blood coagulation factor, and (iv) platelet glycoprotein Ibα, which mediates shear-stress-dependent platelet aggregation [16]. Polymorphisms of these factors are associated with plasma levels or activities of these factors [17–29].

We conducted allelic association studies comparing two groups of patients with Type 2 diabetes mellitus; i.e. those without retinopathy or nephropathy vs. those with retinopathy and/or nephropathy. Although the patients were initially registered randomly, we selected patients based on the following entry criteria, as duration of diabetes is the major determinant for the development of microangiopathies; i.e. no minimum duration of diabetes was required for the microangiopathy-positive group, whereas for the microangiopathy-negative group, only those with a disease duration of at least 10 years were eligible for this study. These criteria would avoid the enrolment of microangiopathy-negative patients with short disease durations. As a result, however, the two groups matched each other in terms of age at diagnosis of diabetes, gender distribution, and mean duration of diabetes. The findings of the present study indicate that one polymorphism in the PON gene, ^{192}Gln/Arg, is associated with the prevalence of diabetic retinopathy or nephropathy.

Patients and methods

Study population and evaluation of diabetic complications

This was a cross-sectional study comparing diabetic patients with and without microangiopathies (retinopathy and nephropathy). The subjects in this study were 280 genetically unrelated Japanese patients who were diagnosed with Type 2 diabetes as defined by the World Health Organization criteria [30]. Those who were followed up at the outpatient clinic of Saitama Social Insurance Hospital on a regular basis were randomly entered into the study, if they were judged to be eligible (see below). Patients were divided into two groups, i.e. 188 patients

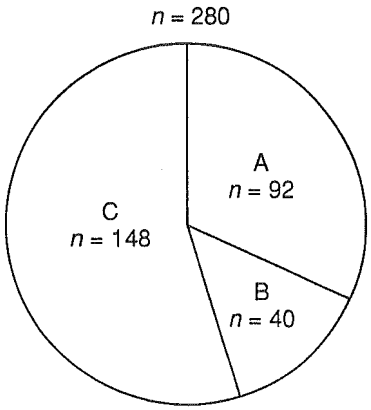


Figure 1 Classification of 280 patients with Type 2 diabetes. Patients were classified as (A) no retinopathy or nephropathy, (B) with retinopathy but no nephropathy or (C) both retinopathy and nephropathy.

with retinopathy or nephropathy (40 patients with retinopathy alone, 148 patients with retinopathy and nephropathy), and 92 patients without retinopathy or nephropathy, based on the definitions below (Fig. 1). For the latter group, only those patients whose disease duration exceeded 10 years were eligible for the study, while no minimum disease duration was required for the former group. As a result, however, the two groups matched each other in terms of gender distribution, age at diagnosis of diabetes, and the mean duration of diabetes. Mean (± SD) ages at diagnosis were 47.9 ± 8.4 years for those without and 49.0 ± 11.4 years for those with microangiopathy; mean disease duration was 14.9 ± 4.5 and 14.5 ± 8.4 years for those without and with microangiopathy, respectively (Table 1). Age at diagnosis was defined as follows; for those without symptoms before diagnosis, ages at which the attending physician made a diagnosis were recorded; for those who had symptoms before diagnosis, onset ages were predicted by the attending physicians.

Diabetic retinopathy was documented by diabetic ophthalmologists using standard fundus photos and was classified as simple, pre-proliferative, and proliferative retinopathy. Diagnosis of diabetic nephropathy was defined by the presence of continuous microalbuminuria or proteinuria with concomitant presence of diabetic retinopathy (group C in Fig. 1), and classified as microalbuminuria (urinary albumin index 30–300 mg/g creatinine), overt proteinuria (> 300 mg/g creatinine), and chronic renal failure. Patients with a microalbuminuria with urinary albumin index less than 20 mg/g creatinine were considered to have no diabetic nephropathy. Patients with a microalbuminuria with urinary albumin index of 20–30 mg/g creatinine were excluded from the study. Patients with urinary albumin index more than 30 mg/g creatinine but without retinopathy were not considered to have nephropathy, and were excluded from the study. Those patients who had had proteinuria before the diagnosis of diabetes were also excluded. Information on a variety of characteristics relevant to the assessment of diabetic microangiopathy, including age at diagnosis, body mass index (BMI) at first visit, duration of diabetes before genotyping, family history of diabetes, treatment with anti-hypertensive or anti-hyperlipidemic drugs, type of therapy, and haemoglobin A_{1c}

Table 1 Characteristics of 280 patients with Type 2 diabetes

	No retinopathy or nephropathy	With retinopathy (with or without nephropathy)	P
Number	92	188	
Age at diagnosis of diabetes (years)*	47.9 ± 8.4	49.0 ± 11.4	0.411†
Male/female	39/53	81/107	0.912‡
Family history of diabetes (%)	66.3	68.8	0.876‡
BMI at first visit (kg/m^2)*	23.3 ± 3.2	22.9 ± 3.5	0.293†
Duration of diabetes (years)*	14.9 ± 4.5	14.5 ± 8.4	0.655†

*Mean ± SD.
†t-test.
‡χ^2-test.

(HbA_{1c}) level were obtained from patient medical records. This study was approved by the ethics committee of Saitama Social Insurance Hospital, and all subjects enrolled in the study gave informed consent.

Genotyping

Blood was drawn from the peripheral veins and genomic DNA was isolated from leucocytes as described previously [31]. The assays for genotyping of PON1 (^{192}Gln/Arg, identical to A- and B-alleles, respectively, used by other investigators), PAI-1 (promoter 4G/5G), fibrinogen ($^{-455}$G/A), and platelet glycoprotein Ibα (^{145}Thr/Met) were based on amplification of genomic DNA with polymerase chain reaction (PCR) using a DNA thermal cycler (Perkin Elmer, Takara Biomedicals, Ohtsu, Japan) followed by restriction enzyme digestion, as described previously [26,28,29,32]. Oligonucleotide primers used were 5′-TATTGTTGCTGTGGGACCTGAG-3′ and 5′-CACGCTAAACCCAAATACATCTC-3′ for PON1, 5′-CACAGAGAGAGTCTGGCCACGT-3′ (with a −681C > A substitution to create a *Bsl*I cutting site) and 5′-CCAACAGAGGACTCTTGGTCT-3′ for PAI-1, 5′-AAGAATTTGGGAATGCAATCTCTGCTACCT-3′ and 5′-CTCCTCATTGTCGTTGACACCTTGGGAC-3′ for fibrinogen, and 5′-GGACGTCTCCTTCAACCGGC-3′ and 5′-GCTTTGGTGGGGAACTTGAC-3′ for glycoprotein Ibα. PCR products were digested with *Alw*I (Toyobo Biochemicals, Osaka, Japan), *Bsl*I (New England Biolabs, Beverly, MA, USA), *Hae*III (Takara Shuzo, Ohtsu, Japan), and *Bbi*II (Takara Shuzo) for genotyping of PON1, PAI-1, fibrinogen, and platelet glycoprotein Ibα, respectively.

Statistical analysis

Age, BMI at first visit, and duration of diabetes were compared between those with and without microangiopathies using Student's t-test. Statistical analyses of frequency counts were performed using the Chi-square (χ^2) test or Fisher's exact test for small samples. Comparison of continuous variables (age, duration of diabetes, BMI and HbA_{1c} levels) between the three PON genotypes was performed with the use of analysis of variance (ANOVA). A logistic regression analysis was performed to evaluate the interaction between the PON1^{192}Arg/Gln genotypes and other variables in relation to the prevalence of microangiopathy. In this analysis, the dependent variable was the presence or absence of retinopathy (A vs. B+C, Fig. 1). Independent variables included in this analysis were age at diagnosis of diabetes (quantitative), sex (male or female), family history of diabetes (yes or no), BMI at first visit (quantitative), duration of diabetes (quantitative), present HbA_{1c} levels (quantitative), treatment for hypertension (yes or no), treatment for hyperlipidaemia (yes or no), and PON1 genotypes (Gln/Gln vs. Gln/Arg+Arg/Arg). The analysis was performed using the SPSS statistical program version 6.12 for Macintosh (SAS Institute, Cary, NC, USA).

Results

Table 1 shows number of patients, age at diagnosis of diabetes, sex, family history of diabetes, BMI at first visit, and duration of diabetes. There were no statistically significant differences between the two groups on these parameters.

Genotype distributions of PON1^{192}Gln/Arg polymorphism are shown in Table 2. For those without retinopathy or nephropathy (group A, Fig. 1), genotypes for Gln/Gln, Gln/Arg and Arg/Arg were 15 (16.3%), 46 (50.0%), and 31 (33.7%), respectively, whereas for those with retinopathy (group B+C), they were 11 (5.9%), 98 (52.1%), and 79 (42.0%), respectively (P = 0.0149, when compared with a 2 × 3 contingency table). When those without and with the Arg-allele were compared (i.e. Gln/Gln vs. Gln/Arg+Arg/Arg), the odds ratio (OR) was 3.13 [95% confidence interval (CI), 1.42–6.89, P = 0.0046]. For those with nephropathy (group C), genotypes were 9 (6.1%), 77 (52.0%), and 62 (41.9%), respectively, which were significantly different from those for the microangiopathy-negative group (group A, P = 0.0308 by a 2 × 3 contingency table, and P = 0.0103 by a 2 × 2 contingency table, OR = 3.01, 95% CI, 1.30–6.98). The genotype distribution in the general Japanese population previously published from our laboratory [29] is also shown in Table 2. The effect of PON1 genotype was more obvious when analysis was confined to male patients; i.e. the OR was 3.92 (1.27–12.15) for retinopathy and 4.00 (1.19–13.39) for nephropathy. However, genotype effect was weaker for female patients, although in each analysis the OR exceeded 2.0.

Table 3 shows the genotype frequencies in relation to the stage of microangiopathy. There was an increased frequency