Original Paper

Cerebrovascular Diseases

Cerebrovasc Dis
DOI: 10.1159/000347066

Received: October 11, 2012
Accepted: January 1, 2013
Published online: ■■■

Public Awareness of Early Symptoms of Stroke and Information Sources about Stroke among the General Japanese Population: The Acquisition of Stroke Knowledge Study

Naomi Miyamatsu[a] Tomonori Okamura[b] Hirofumi Nakayama[c] Kazunori Toyoda[d] Kazuo Suzuki[f] Akihiro Toyota[g] Takashi Hata[h] Atsushi Hozawa[i] Tomofumi Nishikawa[j] Akiko Morimoto[e] Mihoko Ogita[j] Ayumi Morino[a] Takenori Yamaguchi[c, d]

[a]Department of Clinical Nursing, Shiga University of Medical Science, Otsu, [b]Keio University, Tokyo, [c]Japan Stroke Association, Osaka, [d]National Cerebral and Cardiovascular Center and [e]Osaka University, Suita, [f]Research Institute for Brain and Blood Vessels, Akita, [g]Chugoku Rosai Hospital, Kure, [h]Shizuoka City Shimizu Hospital, Shizuoka, [i]Tohoku University, Sendai, and [j]Kyoto Koka Women's University, Kyoto, Japan

Key Words
Stroke symptoms · Warning signs · Knowledge · Education campaign · Prehospital delay

Abstract
Background: It is important that the general population be aware of the early symptoms, since it has been shown that early arrival to hospitals leads better prognosis of stroke patients. However, the general population is not well informed about the early symptoms of stroke. This study was conducted to clarify which stroke symptoms are less well known and which information sources are related to awareness of stroke symptoms. ***Methods:*** A multiple-choice, mail-in survey involving 5,540 randomly selected residents, aged 40–74 years, of 3 cities in Japan was conducted. Their knowledge about stroke symptoms and their information sources were surveyed; information sources were classified as mass media (television/newspaper/radio) and personal communication sources (posters/leaflets/internet/health professionals/family and/or friends). 'Awareness' was defined as selecting all 5 of the correct stroke symptoms from among 10 listed symptoms with decoy choices. The estimated fraction of the possible impact due to each source on the whole population was also calculated by odds ratios (ORs) and the proportion of respondents who selected each source (Pe). The combined effects of mass media and personal communication sources on awareness were also assessed. ***Results:*** Of the 5,540 residents, only 23% selected all 5 correct symptoms. Visual disturbance was the least known of the 5 symptoms (35%). All sources were positively related to awareness, with ORs (Pe) of: television, 1.58 (72.5%); newspaper, 1.79 (48.0%); radio, 1.74 (13.3%); posters, 1.73 (7.6%); leaflets, 1.50 (24.7%); Internet, 1.66 (5.6%); health professionals, 1.33 (34.8%), and family/friends, 1.21 (44.6%). The estimated fraction of the possible impact due to each source was higher for mass media (television, 0.31 and newspaper, 0.28) than personal communication sources (Internet, 0.04 and leaflets, 0.12). Mass me-

KARGER
E-Mail karger@karger.com
www.karger.com/ced

1015–9770/13/0000–0000$38.00/0

Naomi Miyamatsu
Department of Clinical Nursing, Shiga University of Medical Science
Seta Tsukinowa-cho
Otsu 520-2192 (Japan)
E-Mail miyan@belle.shiga-med.ac.jp

dia only and mass media/personal communication sources were significantly associated (ORs: 1.66, 2.75, respectively). ***Conclusions:*** As a single method of public education, television could be the most effective strategy. Moreover, the combined approach involving mass media and personal communication sources might have a synergistic effect. Less well-known symptoms, such as visual disturbances, should be noted in public education campaigns.

Introduction

Stroke is the leading cause of permanent disability in adults and one of the major leading causes of death and disability worldwide [1–5]. Some previous studies examining time limit for initiation of intravenous thrombolysis therapy using alteplase demonstrated that alteplase treatment within 4.5 h of onset improved functional outcome [6, 7]. However, it was also reported that delay in hospital presentation of patients with acute stroke still remained substantial [8–11]. The reasons for this delay were attributed to both lack of awareness of stroke attacks and nonuse of ambulances in previous studies [8, 12]. Thus, it is important for improving stroke outcomes to educate citizens to react immediately and appropriately at their stroke onset. Such community education should be focused on the recognition of early symptoms of stroke onset and appropriate reaction at stroke onset, i.e., to call an ambulance as soon as possible. Despite the importance of recognizing early symptoms, the population remains poorly informed about the early symptoms of stroke [13–17].

The purpose of the present study was to clarify which early symptoms of stroke are less well known, and which information sources are used to obtain knowledge about stroke by the general Japanese population. The study aimed furthermore to identify which information sources were related to awareness of stroke symptoms.

Methods

Acquisition of Stroke Knowledge (ASK) Study

This study was conducted as a baseline survey of a nonrandomized community intervention trial, the Acquisition of Stroke Knowledge Study, to improve awareness about early symptoms of stroke and the appropriate response to stroke onset. Three cities in Japan, i.e., Akita, Shizuoka and Kure, were selected. Akita is located in the northern part of Japan. In the present study, two districts of Akita city, i.e., Kawabe and Yuwa, were selected in order to prepare for future intensive intervention targeting improvement of knowledge about stroke. Kure is a city in Hiroshima prefecture in the western part of Japan, and Shizuoka is located in the central part of Japan. A preintervention survey was followed by community intervention. After the community intervention, postintervention surveys have been performed to assess a long-term effect of our intervention, which are now ongoing and under analysis. This article was based on the preintervention survey in 2006.

From these three areas, 11,306 community dwellers, aged 40–74 years, were randomly selected by an age-stratified random sampling method from the Basic Resident Register. To align the age distribution among participants in the three areas, the number of listed participants in each age bracket, i.e., 40–49, 50–59, 60–69 and 70–74, was almost equal.

Measures and Procedure

A mail-in survey was conducted from April to July in 2006, and a closed-end questionnaire was mailed to each participant. The questionnaire consisted of the following: general knowledge of stroke, early symptoms of stroke, what to do at the time of stroke onset, information sources for knowledge about stroke, and sociodemographic factors (see 'Appendix'). The questions regarding 'early symptoms of stroke' consisted of 5 correct answers and 5 decoy answers as multiple-choice items. The questions regarding 'information sources about stroke' consisted of 11 multiple-choice items. For these questions, multiple answers were allowed. The questions regarding 'Response to stroke attack' consisted of 7 multiple-choice items, for which a single answer was required.

This study was approved by the ethics committee of Shiga University of Medical Science (17–97).

Analysis Methods

Differences in responses among age groups were assessed with the χ^2 test. In this study, respondents with 'awareness' of early symptoms were defined as those who selected all 5 correct early symptoms of stroke from the 10 symptoms listed, except for those who selected all items. The odds ratio (OR) of having 'awareness' of early symptoms from an information source was calculated with logistic regression analysis after adjusting for age, sex, survey site, with/without risk factors for stroke, with/without stroke patients close to respondents, and living alone or not, among respondents who did not have a past history of stroke. In addition, the estimated fraction of possible impact due to each information source on the whole population was calculated according to the ORs mentioned above and the proportion of respondents who chose each information source using this formula in analogy to Levin's [18] formula for population-attributable risk: [Pe × (multivariate adjusted OR -- 1)]/[Pe × (multivariate adjusted OR -- 1) + 1], where Pe is the proportion of respondents who chose one information source as their own information source, and OR is the odds ratio of having awareness of early symptoms in the same respondents. We believe that both of them are important to improve public awareness. Further logistic regression analysis was performed to assess the combined effect of mass media and personal communications on awareness of stroke symptoms.

All significance tests were two-tailed, and $p < 0.05$ was considered significant in all analyses. Data were analyzed with SPSS version 15.0 for Windows (SPSS Inc.).

Table 1. Response rate and characteristics of respondents by sex and age

	Overall	Sex		Age, years			
		men	women	40–49	50–59	60–69	70–74
Responses, n	5,540	2,618	2,922	1,390	1,524	1,779	847
Subjects, n	11,306	5,672	5,634	3,184	3,311	3,207	1,604
Response rate, %	49.0	46.2	51.9	43.7	46.0	55.5	52.8
	(n = 5,540)	(n = 2,618)	(n = 2,922)	(n = 1,390)	(n = 1,524)	(n = 1,779)	(n = 847)
Past history of stroke	128 (2.3)	81 (3.1)	47 (1.6)	7 (0.5)	14 (0.9)	60 (3.4)	47 (5.5)
Self-reported underlying diseases of stroke							
High blood pressure	1,211 (21.9)	656 (25.1)	555 (19.0)	107 (7.7)	270 (17.7)	518 (29.1)	316 (37.3)
High cholesterol	729 (13.2)	325 (12.4)	404 (13.8)	94 (6.8)	217 (14.2)	306 (17.2)	112 (13.2)
Diabetes	429 (7.7)	270 (10.3)	159 (5.4)	43 (3.1)	85 (5.6)	193 (10.8)	108 (12.8)
Heart diseases	213 (3.8)	127 (4.9)	86 (2.9)	10 (0.7)	26 (1.7)	103 (5.8)	74 (8.7)
Arrhythmia	423 (7.6)	241 (9.2)	182 (6.2)	50 (3.6)	102 (6.7)	172 (9.7)	99 (11.7)
Transient ischemic attack	31 (0.6)	19 (0.7)	12 (0.4)	1 (0.1)	4 (0.3)	11 (0.6)	15 (1.8)
At least one of diseases mentioned above	2,143 (38.7)	1,131 (43.2)	1,012 (34.6)	253 (18.2)	521 (34.2)	888 (49.9)	481 (56.8)
With stroke patients close to respondents	3,031 (54.7)	1,426 (54.5)	1,605 (54.9)	703 (50.6)	883 (57.9)	1,010 (56.8)	435 (51.4)
Living alone	471 (8.5)	230 (8.8)	241 (8.2)	92 (6.6)	115 (7.5)	166 (9.3)	98 (11.6)

Data shown as number of subjects with percentage in parentheses.

Results

Response Rate and Age Distribution of the Survey Population

The respondents of this study were 5,540 individuals (response rate 49.0%) with a mean age of 58.1 ± 9.8 years, 52.7% were women (table 1). Of the respondents, 2.3% reported a history of stroke, 54.7% reported that a close acquaintance had suffered a stroke, and 38.7% had at least one of the risk factors for stroke (table 1).

Knowledge of Stroke

As shown in table 2, approximately 80% of all respondents reported knowing what stroke was. Concerning the early symptoms of stroke, approximately 90% of respondents appropriately selected 'sudden hemiplegia' and 'sudden speech problem' as early symptoms, followed by 'sudden, severe headache' (72.3%) and 'sudden dizziness or loss of balance' (62.7%). However, only 35.0% of respondents selected 'sudden vision problem'. Of all respondents, 1,288 (23.2%) were classified as having 'awareness' of early symptoms, which was defined as complete selection of the 5 correct early symptoms. The proportion of respondents who had 'awareness' was higher in those aged less than 60 years (26.8, 19.3%, respectively, $p < 0.001$ for the χ^2 test), and was higher in those who had someone close who was a stroke patient than those who did not (26.1, 19.8%, respectively, $p < 0.001$ for the χ^2 test). Even among 4,285 respondents who reported that 'I generally know what a stroke is', only 1,076 (25.1%) were defined as those who had awareness of early symptoms of stroke (data not shown). Over 80% of respondents answered that they would call an ambulance at stroke onset (table 2).

Associations between Information Sources and Awareness of Early Symptoms of Stroke

In the present study, eight information sources were grouped into two major types of information sources. Television, newspaper, and radio were categorized as 'mass media', and Internet, family/friends, leaflets, posters, and health professionals were categorized as 'personal communication sources'. As shown in table 3, the proportion of respondents who chose television (72.4%) was higher than other information sources, especially 'personal communication sources'. Few respondents selected the Internet and posters as the information source about stroke. Substantial gender difference was not observed. Health professionals were chosen by elders, whereas Internet was chosen by young people.

The results of multivariate-adjusted logistic regression analyses, which were performed to assess associations between each information source and 'awareness' of early symptoms, are shown in table 4. These results demonstrated that respondents who obtained information for

Table 2. General knowledge, knowledge of risk factors and early symptoms of stroke, and reaction to stroke onset by sex and age

	Overall (n = 5,540)	Age, years				
		40–49 (n = 1,390)	50–59 (n = 1,524)	60–69 (n = 1,779)	70–74 (n = 847)	p value
Do you know what is 'stroke'?[a]						<0.001
Generally know	4,285 (77.3)	1,017 (73.2)	1,194 (78.3)	1,401 (78.8)	673 (79.5)	
Only name of disease	1,063 (19.2)	339 (24.4)	284 (18.6)	310 (17.4)	130 (15.3)	
Do not know	96 (1.7)	9 (0.6)	21 (1.4)	40 (2.2)	26 (3.1)	
Nonresponding	96 (1.7)	25 (1.8)	25 (1.6)	28 (1.6)	18 (2.1)	
What are the early symptoms of stroke onset?[b]						
Sudden one-sided numbness or weakness of the face, arm, or leg	4,797 (86.6)	1,242 (89.4)	1,346 (88.3)	1,512 (85.0)	697 (82.3)	<0.001
Sudden confusion or trouble speaking or understanding others	4,796 (86.6)	1,255 (90.3)	1,371 (90.0)	1,495 (84.0)	675 (79.7)	<0.001
Sudden severe headache with no known cause	4,008 (72.3)	1,077 (77.5)	1,158 (76.0)	1,222 (68.7)	551 (65.1)	<0.001
Sudden dizziness, trouble walking, or loss of balance or coordination	3,471 (62.7)	882 (63.5)	997 (65.4)	1,111 (62.5)	481 (56.8)	<0.001
Sudden trouble seeing in one or both eyes	1,938 (35.0)	571 (41.1)	579 (38.0)	544 (30.6)	244 (28.8)	<0.001
Palsy of both hands and/or fingers[c]	2,461 (44.4)	675 (48.6)	712 (46.7)	729 (41.0)	345 (40.7)	<0.001
Selected all 5 correct warning signs of stroke	1,288 (23.2)	373 (26.8)	408 (26.8)	344 (19.3)	163 (19.2)	<0.001
Selected all 5 correct warning signs of stroke[d]	1,267 (23.0)	367 (26.4)	399 (26.2)	339 (19.1)	162 (19.1)	<0.001
How do you respond if you are having a stroke attack?[a]						<0.001
Immediately call an ambulance	4,500 (81.2)	1,176 (84.6)	1,265 (83.0)	1,401 (78.8)	658 (77.7)	
Immediately call a primary physician at clinic or hospital	401 (7.2)	44 (3.2)	89 (5.8)	171 (9.6)	97 (11.5)	
Immediately call a large and/or special hospital	150 (2.7)	40 (2.9)	38 (2.5)	48 (2.7)	24 (2.8)	
Immediately see a primary physician at clinic or hospital	124 (2.2)	24 (1.7)	29 (1.9)	54 (3.0)	17 (2.0)	
Immediately see a doctor in large and/or special hospital	261 (4.7)	78 (5.6)	82 (5.4)	72 (4.0)	29 (3.4)	
See a primary physician at clinic or hospital during office hours	24 (0.4)	6 (0.4)	2 (0.1)	9 (0.5)	7 (0.8)	
See a doctor in large and/or special hospital during office hours	11 (0.2)	4 (0.3)	2 (0.1)	3 (0.2)	2 (0.2)	
Wait and observe symptoms for several days	11 (0.2)	3 (0.2)	3 (0.2)	4 (0.2)	1 (0.1)	
Unknown	58 (1.0)	15 (1.1)	14 (0.9)	17 (1.0)	12 (1.4)	

Data shown as number of subjects with percentage in parentheses. [a] Single answer was required. [b] Multiple answers allowed. [c] Decoy answer. [d] Of the 5,519 respondents, excluding those who chose all 10 items. p value for χ^2 test.

Table 3. Proportion of respondents who chose each information source among 5,391 members[a] of the general population (%, by sex and age)

Type of information source	Overall (n = 5,391)	Sex		Age, years			
		men (n = 2,527)	women (n = 2,864)	40–49 (n = 1,377)	50–59 (n = 1,501)	60–69 (n = 1,714)	70–74 (n = 799)
Mass media							
Television	72.4	68.5	75.9	74.2	71.8	73.2	69.0
Newspaper	48.0	48.1	47.8	48.0	50.2	49.5	40.3
Radio	13.2	14.4	12.1	11.1	14.9	14.3	11.1
Personal information							
Internet	5.5	7.1	4.1	10.7	6.9	2.1	1.3
Family/friends	44.5	40.5	48.0	39.7	48.7	47.0	39.5
Leaflets	24.6	25.6	23.7	20.0	29.2	25.9	20.9
Posters	7.5	8.2	6.9	7.8	9.3	7.0	4.6
Health professionals[b]	34.7	36.6	33.0	21.9	30.5	41.4	50.2

[a] 5,391 respondents excluding respondents who had a past history of stroke (n = 128) or selected all 10 as early symptoms (n = 21). [b] Medical doctors and/or nurses.

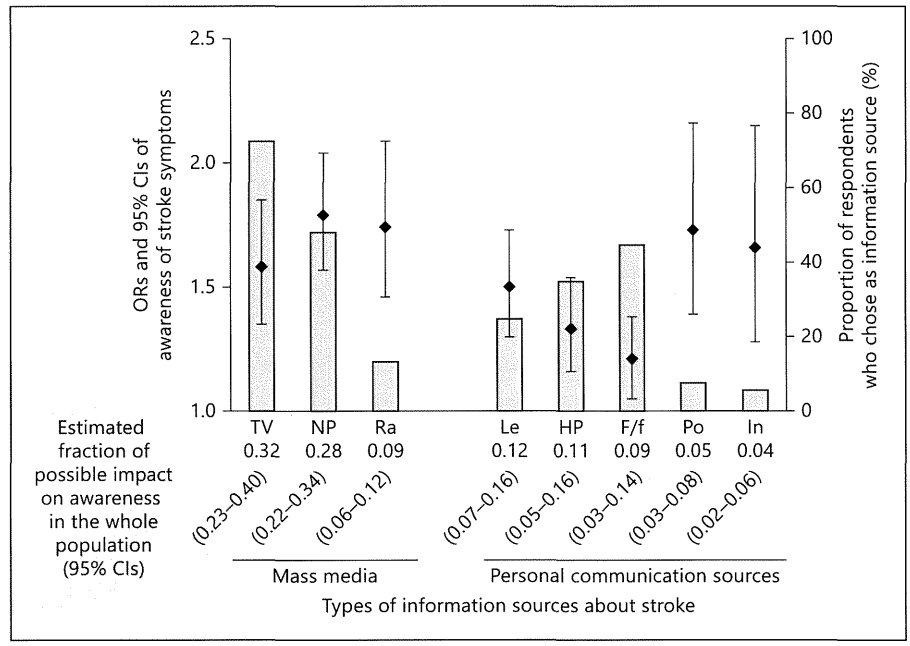


Fig. 1. Estimated fraction of possible impact of each information source on awareness about stroke symptoms on the whole population (for formula, see text). ORs were calculated adjusting for age, sex, survey site, with/without close stroke patients, living alone or not, with/without risk factors of stroke such as high blood pressure, high cholesterol, diabetes, heart disease, arrhythmia, and transient ischemic attack, along with each information source. Rhombus markers and whiskers indicate ORs and 95% CIs. Gray bars show the proportion of respondents who chose each information source. TV = Television; NP = newspaper; Ra = radio; Le = leaflet; HP = health professionals; F/f = family/friends; Po = poster; In = internet.

Table 4. Awareness of stroke symptoms by each information source among 5,391 members of the general population

Type of information source	Univariate		Multivariate adjusted[b]	
	ORs	95% CIs	ORs	95% CIs
Mass media				
Television	1.68	1.44–1.96	1.65	1.41–1.93
Newspaper	1.83	1.61–2.08	1.81	1.59–2.06
Radio	1.67	1.41–1.98	1.75	1.47–2.08
Personal information				
Internet	1.94	1.51–2.48	1.68	1.30–2.16
Family/friends	1.27	1.12–1.44	1.21	1.06–1.38
Leaflets	1.53	1.33–1.76	1.53	1.32–1.76
Posters	1.78	1.44–2.22	1.74	1.40–2.17
Health professionals[a]	1.23	1.08–1.40	1.34	1.16–1.54

'Awareness of stroke symptoms' was defined as to select all of 5 correct early symptoms of stroke (n = 1,241, 23.3%). ORs and 95% CIs were calculated among 5,391 respondents, excluding respondents who had a past history of stroke (n = 128) or selected all 10 as early symptoms (n = 21).

[a] Medical doctors and/or nurses.

[b] Multivariate-adjusted ORs and 95% CIs were calculated adjusting for age, sex, survey site, with/without closed person with stroke, living alone, with/without risk factors of stroke such as high blood pressure, high cholesterol, diabetes, heart disease, arrhythmia, and transient ischemic attack for each information source.

stroke from any type of sources were more likely to be aware of early symptoms of stroke than their counterparts: newspaper (OR: 1.81), radio (OR: 1.75), and television (OR: 1.65) as 'mass media', and posters (OR: 1.74), Internet (OR: 1.68), leaflets (OR: 1.53), health professionals (OR: 1.34), and family/friends (OR: 1.21) as 'personal communication sources'.

Estimated Fraction of Impact due to Each Information Source on the Whole Population

Figure 1 shows the estimated fraction of the impact due to each information source about awareness of stroke symptoms on the whole population. Mass media, especially television and newspapers, were found to have a large estimated fraction, i.e., 0.32 (95% confidence interval, CI, 0.23–0.40) for television and 0.28 (95% CI, 0.22–0.34) for newspaper. On the other hand, despite high ORs, posters and the Internet as personal communication sources had a relatively small estimated fraction due to the small proportion of respondents who chose them as information sources, 0.05 (95% CI, 0.03–0.08) for posters and 0.04 (95% CI, 0.02–0.06) for the Internet.

Combination of Mass Media and Personal Communications

To assess the combined effect of mass media and personal communications on 'awareness' of early symptoms, respondents were divided into four groups according to the

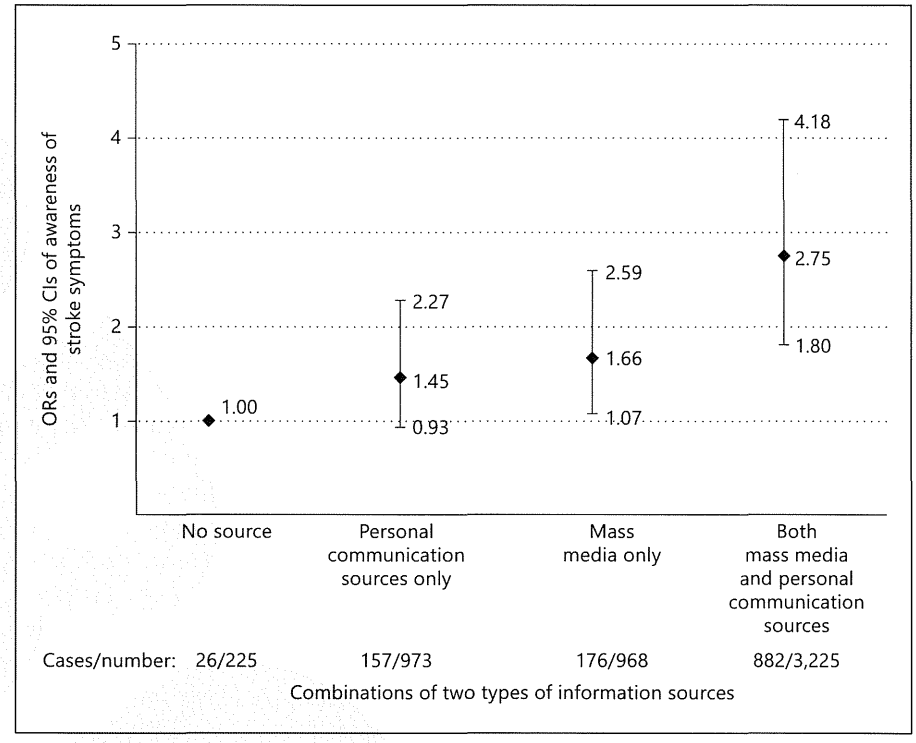


Fig. 2. ORs and 95% CIs of awareness of stroke symptoms by types of communication source among 5,391 participants of the general population. 'Awareness of stroke symptoms' was defined as selecting all 5 correct early symptoms of stroke. ORs and 95% CIs were calculated among the 5,391 respondents, except for respondents who had a past history of stroke (n = 128) or selected all 10 as early symptoms (n = 21). ORs were calculated adjusting for age, sex, survey site, with/without close stroke patients, living alone or not, with/without risk factors of stroke such as high blood pressure, high cholesterol, diabetes, heart disease, arrhythmia, and transient ischemic attack, along with each information source.

type of information source (no source, personal communication sources only, mass media only, both mass media and personal communication sources). On multivariate-adjusted logistic regression analysis using 'no information source' as the reference group, as shown in figure 2, obtaining knowledge about stroke from 'mass media only' was positively related to 'awareness' of early symptoms of stroke (OR: 1.66). The combination of mass media and personal communication sources was significantly related to better knowledge of early symptoms of stroke (OR: 2.75). 'Personal communication sources only' did not have a significant association with awareness of stroke symptoms.

Discussion

The present study is the first community-based study focusing on information sources for 'awareness' of early symptoms of stroke among the general population in Japan. In the present study, the proportion of the general population reported that they would call an ambulance if attacked by a stroke was higher than in the United States [19]. However, only about one fourth of respondents (23.2%) selected all of the correct early symptoms of stroke. Furthermore, despite the low prevalence of respondents who selected all 5 correct answers, nearly 80% of respondents reported that 'I generally know what stroke is'.

In previous studies that focused on knowledge about stroke, 35–70% of respondents had awareness of the early symptoms of stroke [20–22]. The discrepancy in the proportion of people with awareness between previous studies and the present study may be due to the difference in the definitions of 'to be aware' of stroke symptoms. In the present study, persons with awareness were defined as those who selected all 5 correct early symptoms of stroke from among 10 symptoms. On the other hand, knowledge of only 1–2 of correct symptoms was defined as awareness in the analysis of previous studies, i.e. Fogle et al. [22] defined a person reporting two or more symptoms of stroke as one who had awareness of stroke symptoms, and they reported that the proportion of people with awareness was 68–73% before the campaign. The reason why such a strict definition was used for the assessment of awareness about stroke symptoms in the present study is that people cannot choose their stroke symptom when it occurs. Robinson et al. [23] reported that symptoms not included in the FAST campaign (leg weakness and visual loss) were poorly recognized by British people and indicated that this lack of knowledge might lead to delays in hospital presentation. Actually, in the Greater

Cincinnati/Northern Kentucky Stroke Study, it was found that stroke/transient ischemic attack patients with numbness or visual changes were less likely to contact emergency medical services [24]. To call an ambulance and go as an emergency to the hospital, people who have stroke attacks and bystanders must recognize that it is a stroke, no matter which early symptom is observed. As to knowledge of which early symptom is lacking, less severe symptoms such as 'sudden vision problem' and 'sudden dizziness or loss of balance' tended to be recognized less as early symptoms than sudden hemiparesis/speech problem/severe headache in the present study, which should be emphasized in public education or a campaign.

This study also clarified how general citizens obtained knowledge about stroke, for example, from mass media, such as television, newspaper, and radio, or personal communication sources, such as the Internet, family/friends, leaflets, posters, and health professionals. Apparently, all information sources were positively associated with higher awareness; the OR for having awareness was highest in the persons who selected newspapers as information source. However, in the present study, the estimated fraction of impact due to each information source on the whole population was the largest for television, followed by newspapers. Silver et al. [20] performed a telephone survey before and after mass media campaigns, and they reported that television advertisements changed the ability to name the early symptoms of stroke in a Western population. We also reported that a 1-year intervention with a television campaign improved public awareness of stroke symptoms only in the intervention area in a comparative study design [25]. It seems that consensus has been obtained about the strong effect of television on improvement of stroke knowledge in community-based interventions. However, Silver et al. [20] also reported in the previously cited study that no significant change by receiving print (newspaper) advertising was observed in their study. On the other hand, the present study demonstrated that respondents who reported newspapers as their information source tended to be aware about the early symptoms of stroke, and the estimated impact of newspapers on awareness of stroke symptoms was the second largest after television. A possible reason for explaining this discrepancy is the difference in the daily newspaper subscription rate between study populations. Since about 49,063,000 households in Japan in 2006 paid for a total of 69,100,000 subscriptions [26], it was considered that most Japanese read newspapers. Campaigns involving newspapers could be a particularly effective intervention strategy for populations with a high newspaper subscription rate.

To assess a certain information source, its effect on personal education and the proportion of citizens who chose it as their information source are both important in the view of public education. Accordingly, we calculated the estimated fraction of impact due to each information source on the whole population. The present study indicated that mass media might have a huge advantage in educational campaigns to improve knowledge about stroke. Our findings also suggest that personal communication sources, including leaflets and the Internet, show a potential effect to increase individual knowledge; however, it is important to improve their distribution systems to ensure that many citizens can access the above-mentioned personal communication sources. Additionally, in the present study, it is very interesting that acquisition of knowledge via a combination of mass media, such as television and newspaper, and personal communication sources, such as leaflets and/or posters, was strongly related to awareness of early symptoms of stroke. Our findings are valuable for the development of an efficient strategy for education of the public at large.

There are several limitations to the present study. Firstly, there might have been a response bias in the survey because of the low response rate. Unfortunately, information to compare responders and nonresponders could not be obtained. Respondents may have been citizens who were relatively interested in stroke. Thus, the prevalence of citizens who have awareness of stroke symptoms might be overestimated. Secondly, the intensity of exposure, such as the duration of watching television or reading a newspaper daily concerning stroke, was not evaluated because of the limited space of the mail-in survey questionnaire. Thirdly, socioeconomic factors, such as educational level or annual income, which might be associated with information sources [27], were not considered. However, representatives of local characteristics such as household composition ratio, educational levels and unemployment rate, did not substantially differ between average of three study areas and the whole of Japan. Finally, the cross-sectional design cannot prove causality. In conclusion, this community-based survey involving a randomly selected sample of the general population showed that most respondents knew that they should immediately seek healthcare, including use of an ambulance, when they suspected or are suspected to have a stroke attack. However, they did not completely understand the early symptoms of stroke onset, especially mild symptoms. Less well-known symptoms, such as visual problems, should be noted in the campaign. In addition, the present study showed that any type of information,

especially a combination of mass media such as television and newspapers and personal communication sources such as the Internet and leaflets, was associated with awareness of stroke symptoms. A campaign including a mixture of mass media and personal communications, therefore, should be an effective strategy for public education about stroke. Further prospective or intervention studies are needed to assess the synergistic effects of complex intervention strategies.

Acknowledgments

This study was partly supported by a grant from the Japan Cardiovascular Foundation and Grants-in-Aid from the Ministry of Health, Labor and Welfare, Comprehensive Research on Lifestyle-Related Diseases including Cardiovascular Diseases and Diabetes Mellitus [H23-Junkankitou (Seishuu)-Ippan-009], in association with project expenses from the Japan Stroke Association, Osaka, Japan.

Disclosure Statement

None declared.

Appendix

Q1 Age _____ years

Q2 Sex
- 1 Man
- 2 Woman

Q3 Do you have any paid work/job now? (A part-time job is also included.)
- 1 Yes
- 2 No

Q4 Do you live alone?
- 1 Yes
- 2 No

Q5 Do you know what is stroke?
- 1 Yes: generally know
- 2 Yes: only name of the disease
- 3 No

Q6 Do you know that stroke is comprised of three types, i.e., cerebral hemorrhage, cerebral infarction, and subarachnoid hemorrhage?
- 1 Yes
- 2 No

Q7 Have you suffered any disease? Please check all of the following
- 1 No disease
- 2 Cancer
- 3 Heart disease
- 4 Arrhythmia
- 5 Stroke
- 6 Transient ischemic attack
- 7 Hypertension
- 8 Gastroduodenal ulcer
- 9 Liver disease
- 10 Diabetes mellitus
- 11 Hyperlipidemia/hypercholesterolemia
- 12 Kidney disease
- 13 Hyperuricemia/gout
- 14 Rheumatoid arthritis
- 15 Asthma
- 16 Cataract
- 17 Mental disorder
- 18 Dementia
- 19 Other disease

Q8 Do you have any family members, colleagues, or friends with a history of stroke?
- 1 Yes
- 2 No

Q9 What are the early symptoms of a stroke attack? Please choose all fitting symptoms from the following
- 1 Sudden one-sided numbness or weakness of the face, arm, or leg
- 2 Sudden nasal bleeding
- 3 Sudden fever
- 4 Sudden confusion or trouble speaking or understanding others
- 5 Sudden pain on left shoulder
- 6 Sudden trouble seeing in one or both eyes
- 7 Sudden dizziness, trouble walking, or loss of balance or coordination
- 8 Numbness of both hands and/or fingers
- 9 Sudden severe headache with no known cause
- 10 Sudden difficulty in breathing

Q10 How do you respond to a stroke attack? Please choose only one from the following
- 1 Immediately call an ambulance
- 2 Immediately call a family physician at clinic or hospital
- 3 Immediately call a large and/or specialized hospital
- 4 Immediately see a family physician at clinic or hospital
- 5 Immediately see a doctor in large and/or specialized hospital
- 6 See a family physician at clinic or hospital during office hours
- 7 See a doctor in large and/or specialized hospital during office hours
- 8 Wait and observe the symptoms for several days

Q11 What are your information sources of knowledge about stroke? Please choose all the sources
- 1 Medical doctor
- 2 Nurse and/or public health nurse
- 3 Leaflet
- 4 Poster
- 5 Television
- 6 Radio
- 7 Newspaper
- 8 Internet
- 9 Family and/or friend
- 10 Other (_____)
- 11 No information source

References

1 Hayakawa T, Okayama A, Ueshima H: Prevalence of impaired activities of daily living and impact of stroke and lower limb fracture on it in Japanese elderly people. CVD Prevention 2000;3:187–194.
2 Health and Welfare Statistics Association: Annual report of National Vital statistics 2007 (in Japanese). J Health Welfare Stat 2008;55: 38–41.
3 Health and Welfare Statistics Association: Lifestyle related diseases and health promotion strategies (in Japanese). J Health Welfare Stat 2009;56:81–96.
4 Heron M, Hoyert DL, Murphy SL, Xu J, Kochanek KD, Tejada-Vera B: Deaths: final data for 2006. Natl Vital Stat Rep 2009;57:1–15.
5 Centers for Disease Control and Prevention (CDC): Prevalence of disabilities and associated health conditions among adults – United States, 1999. MMWR Morb Mortal Wkly Rep 2001;50:120–125.
6 Wahlgren N, Ahmed N, Davalos A, et al: Thrombolysis with alteplase 3–4.5 h after acute ischaemic stroke (SITS-ISTR): an observational study. Lancet 2008;372:1303–1309.
7 Hacke W, Kaste M, Bluhmki E, et al: Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med 2008;359: 1317–1329.
8 Chang KC, Tseng MC, Tan TY: Prehospital delay after acute stroke in Kaohsiung, Taiwan. Stroke 2004;35:700–704.
9 Morris DL, Rosamond W, Madden K, Schultz C, Hamilton S: Prehospital and emergency department delays after acute stroke: the Genentech Stroke Presentation Survey. Stroke 2000;31:2585–2590.
10 Evenson KR, Rosamond WD, Morris DL: Prehospital and in-hospital delays in acute stroke care. Neuroepidemiology 2001;20:65–76.
11 Kimura K, Kazui S, Minematsu K, Yamaguchi T: Analysis of 16,922 patients with acute ischemic stroke and transient ischemic attack in Japan: a hospital-based prospective registration study. Cerebrovasc Dis 2004;18:47–56.
12 Williams LS, Bruno A, Rouch D, Marriott DJ: Stroke patients' knowledge of stroke: influence on time to presentation. Stroke 1997;28: 912–915.
13 Pancioli AM, Broderick J, Kothari R, Brott T, Tuchfarber A, Miller R, Khoury J, Jauch E: Public perception of stroke warning signs and knowledge of potential risk factors. JAMA 1998;279:1288–1292.
14 Pandian JD, Jaison A, Deepak SS, Kalra G, Shamsher S, Lincoln DJ, Abraham G: Public awareness of warning symptoms, risk factors, and treatment of stroke in northwest India. Stroke 2005;36:644–648.
15 Montaner J, Vidal C, Molina C, Alvarez-Sabin J: Selecting the target and the message for a stroke public education campaign: a local survey conducted by neurologists. Eur J Epidemiol 2001;17:581–586.
16 Sug Yoon S, Heller RF, Levi C, Wiggers J, Fitzgerald PE: Knowledge of stroke risk factors, warning symptoms, and treatment among an Australian urban population. Stroke 2001;32:1926–1930.
17 Becker K, Fruin M, Gooding T, Tirschwell D, Love P, Mankowski T: Community-based education improves stroke knowledge. Cerebrovasc Dis 2001;11:34–43.
18 Levin ML: The occurrence of lung cancer in man. Acta Unio Int Contra Cancrum 1953;19: 531–541.
19 Kamel H, Navi BB, Fahimi J: National trends in ambulance use by patients with stroke, 1997–2008. JAMA 2012;307:1026–1028.
20 Silver FL, Rubini F, Black D, Hodgson CS: Advertising strategies to increase public knowledge of the warning signs of stroke. Stroke 2003;34:1965–1968.
21 Hodgson C, Lindsay P, Rubini F: Can mass media influence emergency department visits for stroke? Stroke 2007;38:2115–2122.
22 Fogle CC, Oser CS, McNamara MJ, Helgerson SD, Gohdes D, Harwell TS: Impact of media on community awareness of stroke warning signs: a comparison study. J Stroke Cerebrovasc Dis 2010;19:370–375.
23 Robinson TG, Reid A, Haunton VJ, Wilson A, Naylor AR: The face arm speech test: does it encourage rapid recognition of important stroke warning symptoms? Emerg Med J 2012, E-pub ahead of print.
24 Kleindorfer D, Lindsell CJ, Moomaw CJ, Alwell K, Woo D, Flaherty ML, Adeoye O, Zakaria T, Broderick JP, Kissela BM: Which stroke symptoms prompt a 911 call? A population-based study. Am J Emerg Med 2010;28: 607–612.
25 Miyamatsu N, Kimura K, Okamura T, Iguchi Y, Nakayama H, Toyota A, Watanabe M, Morimoto A, Morinaga M, Yamaguchi T: Effects of public education by television on knowledge of early stroke symptoms among a Japanese population aged 40 to 74 years: a controlled study. Stroke 2012;43:545–549.
26 Chapter II: Country Report 102: Japan. World Press Trends 2007. Paris, World Association of Newspapers, 2007, pp 411–412.
27 Lundelin K, Graciani A, García-Puig J, Guallar-Castillón P, Taboada JM, Rodríguez-Artalejo F, Banegas JR: Knowledge of stroke warning symptoms and intended action in response to stroke in Spain: a nationwide population-based study. Cerebrovasc Dis 2012; 34:161–168.

栃木県脳卒中啓発プロジェクト

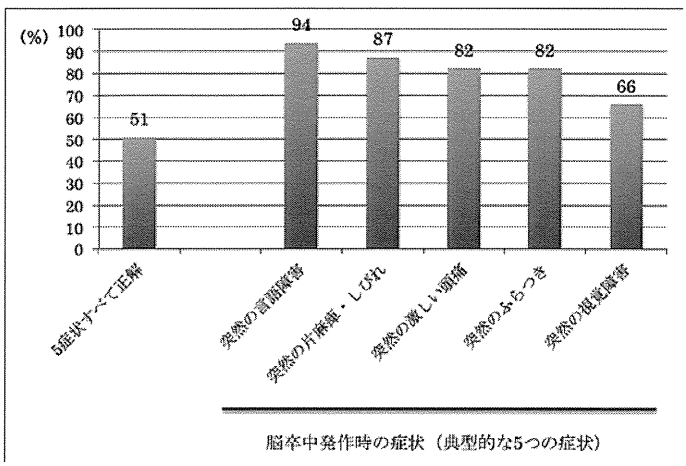
― 栃木県における包括的脳卒中市民啓発とその評価

滋賀医科大学　臨床看護学講座　宮松 直美

脳卒中はわが国の要介護状態の最大の原因であり、要介護度も高くなることが知られています。脳卒中による要介護数を減少させるためには、脳卒中の発症を予防すること、発症後の早期リハビリテーションなどで社会復帰を目指すことに加えて、脳卒中発作時にできるだけ早く専門医療機関を受診することが重要です。発症早期の専門医療機関受診のためには、市民の皆さんひとりひとりが脳卒中の発作を見逃さず、直ちに救急車を呼ぶことが大切になります。そのため、(公社) 日本脳卒中協会は、脳卒中になられた患者さんとご家族の支援や脳卒中予防の推進に加えて、脳卒中発症時の症状と適切な対処に関する知識の普及に長年にわたって取り組んできました。

市民の皆さんに脳卒中の症状について理解を深めていただく方法として、(公社) 日本脳卒中協会ではこれまでに厚生労働科学研究補助金事業「超急性期脳梗塞患者の救急搬送及び急性期病院受け入れ体制に関する実態調査研究」班（代表者：木村和美）およびNHK岡山放送局、ACジャパンなど多くの組織や団体の協力を得て、リーフレット・小冊子などの頻回な配布やテレビによる1年間の継続した啓発活動が脳卒中症状に関する知識を向上させること、ACジャパンによる脳卒中啓発のCMを見ている人は脳卒中症状に関する知識が高いことなどを示してきました。またこれらの活動はいずれも科学的な手法で検証して来ました。

しかしながら、このような脳卒中に関する市民啓発をどの地域でも継続的に進めるためには、自治体や地域の医療・福祉機関、企業、教育機関等が中心となった総合的な取り組みが不可欠です。そのため (公社) 日本脳卒中協会では平成24年度に、厚生労働科学研究補助金事業「慢性期ハイリスク者、脳卒中および心疾患患者に適切な早期受診を促すための地域啓発研究」班（代表者：宮本恵宏、以下地域啓発班）と共同で、脳卒中啓発を実施しようとする団体が無料で利用できる啓発動画を制作しました。そして現在は、これまでに制作した多くの啓発ツール (DVDやリーフレットなど) やACジャパンが制作した動画やポスター等を用いた地域啓発と、循環器病研究開発費「新しい脳卒中医療の開拓と均てん化のためのシステム構築に関する研究」班 (代表者：峰松一夫) から提供を受けたアニメやマンガ小冊子による中学生への脳卒中授業を中心とした、包括的な栃木県脳卒中啓発プロジェクトを実施しています。



図：栃木県民の脳卒中発作時症状に関する知識
（平成24年9月調査）

また啓発活動の実施に先立ち、脳卒中発作時の症状の理解と対応に関する知識調査を行いました。調査はRandom Digit Dialing法（ランダムに発生させた番号に電話して調査へのご協力をお願いする、社会調査でよく用いられる方法、以下RDD）により、同意が得られた40～74歳の市民の皆さんに対して実施されました。この調査の結果、脳卒中発作時の5つの典型的な症状のうち、「突然生じる言語障害」「突然生じる片側の麻痺やしびれ」「突然生じる激しい頭痛」「突然生じるふらつき・脱力感」はいずれも8割～9割強の方々が理解していることが示されました。しかしながら「突然生じる視覚

障害」を脳卒中発作の症状と知っている方はやや少なく、5つの典型的な症状すべてを正しく選択した方は全体の約半数のみでした。また、「脳卒中を疑ったとき、しばらく様子を見る」とした方が5%おられました。

今回の栃木県の脳卒中知識調査では、過去の数回の調査と同様に、脳卒中発作時の典型的な5つの症状すべてを理解している方は約半数にとどまること、比較的軽度の症状についての認識が低いことが示され、啓発によりこれらが向上するかどうかを検討することが必要だと考えられました。平成25年5月に栃木県下各地で開催される脳卒中市民講座が終了した後に同様の調査を行い、この8ヶ月にわたる包括的啓発活動の効果を評価する予定です。栃木県の取り組みで啓発効果が検証されれば、多くの都道府県で実施可能な脳卒中市民啓発のモデルとなると考えています。

Genetic Polymorphisms of the Human PNPLA3 Gene Are Strongly Associated with Severity of Non-Alcoholic Fatty Liver Disease in Japanese

Takahisa Kawaguchi[1,2], Yoshio Sumida[3], Atsushi Umemura[4], Keitaro Matsuo[5], Meiko Takahashi[1], Toshinari Takamura[6], Kohichiroh Yasui[7], Toshiji Saibara[8], Etsuko Hashimoto[9], Miwa Kawanaka[10], Sumio Watanabe[11], Sumio Kawata[12], Yasuharu Imai[13], Miki Kokubo[1], Toshihide Shima[4], Hyohun Park[4], Hideo Tanaka[5], Kazuo Tajima[5], Ryo Yamada[1], Fumihiko Matsuda[1,2]*, Takeshi Okanoue[4] for the Japan Study Group of Nonalcoholic Fatty Liver Disease (JSG-NAFLD)

1 Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, **2** Institut National de la Sante et de la Recherche Medicale (INSERM) Unite U852, Kyoto University Graduate School of Medicine, Kyoto, Japan, **3** Center for Digestive and Liver Diseases, Nara City Hospital, Nara, Japan, **4** Center of Gastroenterology and Hepatology, Saiseikai Suita Hospital, Suita, Japan, **5** Division of Epidemiology and Prevention, Aichi Cancer Center, Nagoya, Japan, **6** Department of Disease Control and Homeostasis, Kanazawa University, Graduate School of Medical Science, Kanazawa, Japan, **7** Department of Molecular Gastroenterology and Hepatology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan, **8** Department of Gastroenterology and Hepatology, Kochi Medical School, Kochi, Japan, **9** Department of Internal Medicine and Gastroenterology, Tokyo Women's Medical University, Tokyo, Japan, **10** Center of Liver Diseases, Kawasaki Hospital, Kawasaki Medical School, Okayama, Japan, **11** Department of Gastroenterology, Juntendo University School of Medicine, Tokyo, Japan, **12** Department of Gastroenterology, Yamagata University School of Medicine, Yamagata, Japan, **13** Department of Internal Medicine, Ikeda Municipal Hospital, Ikeda, Japan

Abstract

Background: Nonalcoholic fatty liver disease (NAFLD) includes a broad range of liver pathologies from simple steatosis to cirrhosis and fibrosis, in which a subtype accompanying hepatocyte degeneration and fibrosis is classified as nonalcoholic steatohepatitis (NASH). NASH accounts for approximately 10–30% of NAFLD and causes a higher frequency of liver-related death, and its progression of NASH has been considered to be complex involving multiple genetic factors interacting with the environment and lifestyle.

Principal Findings: To identify genetic factors related to NAFLD in the Japanese, we performed a genome-wide association study recruiting 529 histologically diagnosed NAFLD patients and 932 population controls. A significant association was observed for a cluster of SNPs in *PNPLA3* on chromosome 22q13 with the strongest *p*-value of 1.4×10^{-10} (OR = 1.66, 95%CI: 1.43–1.94) for rs738409. Rs738409 also showed the strongest association ($p=3.6\times10^{-6}$) with the histological classifications proposed by Matteoni and colleagues based on the degree of inflammation, ballooning degeneration, fibrosis and Mallory-Denk body. In addition, there were marked differences in rs738409 genotype distributions between type4 subgroup corresponding to NASH and the other three subgroups ($p=4.8\times10^{-6}$, OR = 1.96, 95%CI: 1.47–2.62). Moreover, a subgroup analysis of NAFLD patients against controls showed a significant association of rs738409 with type4 ($p=1.7\times10^{-16}$, OR = 2.18, 95%CI: 1.81–2.63) whereas no association was obtained for type1 to type3 ($p=0.41$). Rs738409 also showed strong associations with three clinical traits related to the prognosis of NAFLD, namely, levels of hyaluronic acid ($p=4.6\times10^{-4}$), HbA1c ($p=0.0011$) and iron deposition in the liver ($p=5.6\times10^{-4}$).

Conclusions: With these results we clearly demonstrated that Matteoni type4 NAFLD is both a genetically and clinically different subset from the other spectrums of the disease and that the *PNPLA3* gene is strongly associated with the progression of NASH in Japanese population.

Citation: Kawaguchi T, Sumida Y, Umemura A, Matsuo K, Takahashi M, et al. (2012) Genetic Polymorphisms of the Human PNPLA3 Gene Are Strongly Associated with Severity of Non-Alcoholic Fatty Liver Disease in Japanese. PLoS ONE 7(6): e38322. doi:10.1371/journal.pone.0038322

Editor: Takeshi Okanoue, Wageningen University, The Netherlands

Received March 8, 2012; **Accepted** May 3, 2012; **Published** June 14, 2012

Funding: This work was supported by the grant from Ministry of Labor and Welfare Japan [T.O., H20-Hepatitis-general-008], Core Research of Evolutional Science & Technology (CREST). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fumi@genome.med.kyoto-u.ac.jp

Introduction

Nonalcoholic fatty liver disease (NAFLD) includes a broad range of pathologies from fatty liver (simple steatosis), steatonecrosis, and steatohepatitis to cirrhosis [1–3]. NAFLD often accompanies other lifestyle-related pathologies of metabolic syndrome such as diabetes mellitus, hypertension and dyslipidemia, and the number of NAFLD patients is increasing worldwide along with the escalation in the incidence of metabolic syndrome [4]. Prevalence of NAFLD is considered as approximately 8% in Japanese and 6–35% in Europeans [4,5]. The majority of NAFLD

shows simple steatosis with a good prognosis, but approximately 10–30% of NAFLD histologically diagnosed as nonalcoholic steatohepatitis (NASH) shows hepatocyte degeneration (ballooning hepatocyte), necrosis, inflammation and fibrosis, with a higher frequency of liver-related death both in Japanese and European populations [6,7]. Insulin resistance and oxidative stress are considered to be key players in the progression of NASH [8,9]. However, the progression of NASH has been considered to be complex involving multiple genetic factors interacting with the environment and lifestyle, because only a portion of NAFLD patients develops NASH.

The first Genome-wide association (GWA) study searching for such genetic factors identified the *PNPLA3* gene as a major genetic determinant for the predisposition to NAFLD in Hispanic, African American and European American populations according to liver fat contents [10], which was subsequently confirmed in Europeans and Asians according to liver biopsy. Association of *PNPLA3* with not only fatty liver and TG content, but also inflammation and fibrosis were shown in the subsequent studies, so *PNPLA3* may be widely associated with the development of NAFLD [11–13]. More recently, another GWA study reported the association of four additional genes with NAFLD in Europeans [14]. Also, a candidate gene-based approach revealed the association between NAFLD and the apolipoprotein C3 gene in Indians [15]. However, the precise role of such genes in the development of NASH still remains to be elucidated. In addition, no GWA study has been reported for Asian populations to date although the genetic components and their relative contribution may be different between ethnicities.

The Japan NASH Study Group was founded in 2008 aiming at the identification of genetic determinants predisposing to NASH in the Japanese population. Here we report the first GWA study of NAFLD in the Japanese using DNA samples of patients with liver histology-based diagnoses recruited through this multi-institutional research network.

Results

Genome-wide Association Analysis of NAFLD in Japanese

We conducted a GWA study using DNA samples of 543 patients with NAFLD and 942 controls. After quality controls of genotyping results (see materials and methods for details), a total of 529 patients consisting of four NAFLD subgroups according to Matteoni's classification [2] (type1; 100, type2; 73, type3; 29, type4; 327) and 932 controls were subjected to statistical analyses (Table 1). This index pathologically classifies NAFLD according to the degree of inflammation, hepatocyte degeneration, and the existence of fibrosis and Mallory-Denk body in the liver. Genome scan results of 932 DNA samples collected for other genetic studies were used as general Japanese population controls [16]. After standard quality control procedure as described in materials and methods, genotype distributions of 484,751 autosomal SNP markers were compared between the NAFLD cases and control subjects by exact trend test. A slight inflation of p-values was observed by genomic control method ($\lambda = 1.04$) (Figure S1).

We identified six SNP markers located at chromosome 22q13 showing genome-wide significance ($p<1.04\times10^{-7}$) (Figure 1). Among them, four SNPs, namely, rs2896019, rs926633, rs2076211 and rs1010023, located in the *PNPLA3* gene and in strong linkage disequilibrium (LD) ($r^2>0.93$), returned p-values smaller than 1×10^{-9} ($p = 1.5\times10^{-10}$, 7.5×10^{-10}, 1.4×10^{-9} and 1.5×10^{-9}, respectively) (Table 2). Rs738407 and rs3810662 also located in *PNPLA3* showed significant but weaker associations ($p = 1.0\times10^{-7}$ and 1.0×10^{-7}, respectively) than the above four SNP markers. Rs738491, rs2073082, rs3761472, rs2235776, rs2143571 and rs6006473 were in the neighboring *SAMM50* gene which is outside of the linkage disequilibrium (LD) block where the top SNP markers were distributed (Figure 2). These markers were in moderate LD with each other ($r^2>0.42$) and showed p-values between 3.9×10^{-6} and 6.4×10^{-7} but did not reach genome-wide significance (Table S1). Rs738409, the SNP which showed the strongest association with NAFLD in the first GWA study [10], was not included in the SNP array used in our study. This SNP was therefore genotyped using Taqman technology in the same case and control samples that were used for genome scan. Rs738409 showed the strongest association with the disease ($p = 1.4\times10^{-10}$, OR = 1.66, 95%CI: 1.43–1.94) among all the SNP markers examined in this study. The association remained after the correction for population stratification with EIGENSTRAT [17] ($p = 2.3\times10^{-11}$). Although a peak consisting of a cluster of SNPs was observed at the *HLA* locus on chromosome 6 (minimal p-value of 4.10×10^{-7} for rs9262639 located at the 3′ of *C6orf15* gene), the association disappeared when EIGENSTRAT was applied ($p>1.6\times10^{-3}$). We consider this as a result of population stratification between the cases and controls.

Impact of *PNPLA3* Polymorphisms to the Pathogenicity of NAFLD

We next examined whether or not the seven SNPs in the *PNPLA3* gene were associated with the pathogenic status of NAFLD. The genotype distributions of these SNPs were compared by Jonckheere-Terpstra test among the four subgroups of NAFLD patients categorized by Matteoni's classification (type1 to type4). There was a significant increase in the frequency of the risk allele from Matteoni type1 to type4 for all of the seven SNPs (p-values ranging from 3.6×10^{-6} to 0.0017) (Table 2). Among them, rs738409 again showed the strongest association ($p = 3.6\times10^{-6}$) as seen in the simple case/control analysis. On the other hand, there was no significant association between control and Matteoni type1 ($p = 0.76$).

In order to clarify how rs738409 influences the pathogenicity of NAFLD, we performed pairwise comparisons of genotype distributions in the four subgroups of NAFLD patients. There were marked differences in genotype distributions between type4 subgroup and the other three subgroups by multivariable logistic regression adjusted for age, sex and body mass index (BMI) ($p = 2.0\times10^{-5}$, OR = 2.18, 95%CI: 1.52–3.18 between type1 and type4; $p = 1.4\times10^{-3}$, OR = 1.81, 95%CI: 1.26–2.62 between type2 and type4; $p = 0.027$, OR = 1.85, 95%CI: 1.07–3.19 between type3 and type4) (Figure 3). On the other hand, no significant associations were obtained for type1 to type3 in any combinations. When we performed the same analysis between type4 and the pooled genotypes of type1 to type3, we again obtained a significant difference ($p = 4.8\times10^{-6}$, OR = 1.96, 95%CI: 1.47–2.62).

We further examined the specific association of rs738409 with type4 subgroup by using the case/control association results of the initial genome scan. 529 NAFLD patients were divided into 202 patients with type1 to type3 and 327 patients with type4, and genotype distributions of rs738409 in each subgroup were compared with those of 932 control subjects. Exact trend test returned an extremely strong association of rs738409 with type4 subgroup ($p = 1.7\times10^{-16}$, OR = 2.18, 95%CI: 1.81–2.63) whereas no association was obtained for type1 to type3 subgroups ($p = 0.41$).

Table 1. Clinical characteristics according to the histological classification.

Phenotype	Matteoni classification of NAFLD				Control	*p*-value
	Type 1	Type 2	Type 3	Type 4		
Number of samples	100	73	29	327	932	
Sex (Male/Female)	59/41	47/26	13/16	130/197	471/461	0.0023‡
Age (year)	49.7±15.3	51.5±15.3	49.4±14.0	57.6±14.8	48.8±16.3	<0.001
Physical measurement						
BMI	26.2±4.3	27.7±4.8	27.6±3.5	27.7±5.2	–	0.054
Amount of visceral fat (cm^2)	146.8±65.3	154.3±47.7	136.8±53.8	151.7±57.4	–	0.46
Abdominal circumscript (cm)	90.9±9.9	94.1±10.0	88.5±10.2	94.1±11.8	–	0.10
Biochemical trait						
AST (IU/L)	31.1±14.6	36.4±18.5	52.4±35.1	57.7±48.4	–	<0.001
ALT (IU/L)	48.6±30.8	62.8±47.6	81.5±46.9	74.9±48.4	–	<0.001
GGT (IU/L)	71.0±62.5	67.1±66.9	96.1±91.3	76.6±73.9	–	0.25
Albumin (g/dL)	4.5±0.4	4.4±0.3	4.5±0.3	4.3±0.4	–	<0.001
Total bilirubin (mg/dL)	0.9±0.5	0.9±0.5	0.9±0.6	0.8±0.4	–	0.063
Cholinesterase (unit)	389.1±97.0	354.3±97.2	371.1±109.9	348.9±93.2	–	<0.001
Type IV collagen 7S (ng/dL)	3.8±0.7	3.9±0.9	3.9±0.8	5.1±1.7	–	<0.001
Hyaluronic acid (ng/dL)	25.6±22.5	33.6±29.5	31.5±24.0	80.9±84.3	–	<0.001
Triglycerides (mg/dL)	151.9±73.8	154.0±92.1	166.1±86.5	161.2±85.7	–	0.23
Total cholesterol (mg/dL)	209.1±32.8	194.0±38.0	203.0±39.9	200.3±39.0	–	0.093
HbA1c (%)	6.1±1.1	5.9±1.2	6.5±1.8	6.2±1.3	–	0.13
IRI (μg/dL)	9.1±5.4	11.4±9.0	10.4±6.3	14.9±9.9	–	<0.001
FPG (mg/dL)	112.9±33.7	107.3±27.4	109.9±27.7	114.8±33.8	–	0.14
HOMA-IR	2.4±1.5	2.9±2.4	3.0±2.1	4.2±3.0	–	<0.001
hs-CRP (mg/dL)	1078.9±1407	1048.3±1185.0	865.8±658.4	1579.2±2377.9	–	0.027
Adiponectin (μg/mL)	7.4±4.4	8.5±6.6	6.6±2.6	6.9±4.3	–	0.24
Leptin (ng/mL)	9.9±7.4	9.1±6.2	11.3±9.4	12.4±7.9	–	<0.001
Ferritin (ng/mL)	145.8±101.1	176.5±134.0	271.2±307.0	208.3±180.3	–	0.027
Uric acid (mg/dL)	5.9±1.5	5.7±1.2	5.4±1.9	5.7±1.6	–	0.77
PLT ($\times 10^4/\mu L$)	23.0±5.9	22.9±4.9	21.9±6.7	20.2±6.4	–	<0.001
ANA (0/1/2/3/4)	42/17/4/0/0	31/8/4/1/2	15/6/2/0/0	147/76/31/8/12	–	0.015
Clinical history						
Diabetes (NGT/IGT/DM)	36/11/34	24/7/27	12/8/7	103/35/119	–	0.45*
Hyperlipidemia (+/−)	31/68	31/42	9/20	120/206	–	0.60‡
Hypertension (+/−)	64/35	33/40	19/10	155/172	–	0.013‡
Liver biopsy feature						
Brunt grade (1/2/3)	–	–	19/3/2	149/133/44	–	<0.001‡
Brunt stage (1/2/3/4)	–	–	–	123/74/105/24	–	–
Fat droplet (1/2/3/4)	38/32/19/11	14/29/18/7	7/3/10/4	51/99/104/52	–	<0.001
Iron deposition (0/1/2/3/4)	30/14/21/10/1	24/9/12/2/1	10/5/2/2/0	132/56/29/29/11	–	0.16

Measurements are shown as mean ± standard deviation. Categorical values are shown by the count number. *P*-values are calculated by Jonckheere-Terpstra test unless otherwise stated;
‡Chochran-Armitage trend test,
*Kruskal-Wallis test. Abbreviations used for each trait are summarized in materials and methods.
doi:10.1371/journal.pone.0038322.t001

Association of rs738409 Genotypes with Clinical Traits

The quantitative effects of rs738409 genotypes to clinical traits were examined by multivariable regression adjusted for age, sex and BMI (statistical calculation 1, Table 3). Five categorical ordinals, namely, anti-nuclear antibody (ANA), Brunt grade, Brunt stage, fat deposition and iron deposition, were also tested by an ordinal logistic regression analysis. Potential associations ($p<0.05$) were obtained for 11 traits, namely, aspartate transaminase (AST), alanine aminotransferase (ALT), type IV collagen 7S, hyaluronic acid, hemoglobin A1c (HbA1c), fasting immunoreactive insulin (IRI), fasting plasma glucose (FPG), platelet count (PLT), Brunt grade, fat deposition and iron deposition (Table 3). When the results were further adjusted for Matteoni type (statistical calculation 2), AST, hyaluronic acid, HbA1c, FPG,

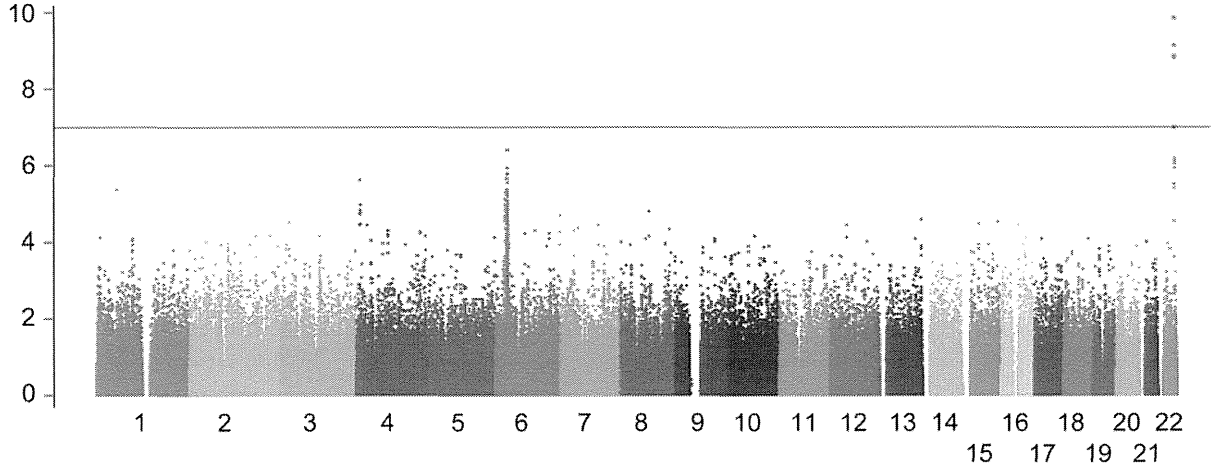


Figure 1. Manhattan plot of the GWA study. Association *p*-values are calculated by exact trend test and plotted along the chromosome in $-\log_{10}$ scale. The horizontal line indicates Bonferroni-adjusted significance threshold ($p = 1.03 \times 10^{-7}$).
doi:10.1371/journal.pone.0038322.g001

PLT, Brunt grade and iron deposition showed *p*-values smaller than 0.05. The level of serum triglyceride was not significant in the initial analysis, but became significant after being adjusted for Matteoni's type ($p = 0.013$). Among them, only three traits, namely, hyaluronic acid, HbA1c and iron deposition, remained significant ($p < 0.0021$) after Bonferroni's correction for multiple testing (Table 3).

Associations of Previously Reported SNPs with NAFLD

Previous genetic studies identified four chromosomal loci, namely, *LYPLAL1* at 1q41, *GCKR* at 2p23, *NCAN* at 19p12 and *PPP1R3B* at 8p23.1, associated with NAFLD in populations of European descent [14]. We examined whether or not the associations were reproduced in the Japanese population by extracting genotype information of SNP markers corresponding to these four loci. As shown in Table 4, the association of rs780094 in *GCKR* with NAFLD was at the border of significance ($p = 0.011$, OR = 0.82, 95%CI: 0.70–0.91) in the case/control analysis. However, the association was lost when examined between rs780094 genotypes and Matteoni types. There were no associations of rs2228603 in *NCAN* and rs12137855 in *LYPLAL1* with either NAFLD or Matteoni types. Rs4240624 in *PPP1R3B* was not in the SNP array used for this study, and this marker was not polymorphic or at a very low frequency in the Japanese (0 in 90

Table 2. List of the SNP markers in the *PNPLA3* locus at chromosome 22q showing genome wide significance.

		Genotyping Result and Allele Frequency of A2						Statistics		
			NAFLD					NAFLD vs. Control		Matteoni
dbSNPID	A1/A2	Control	Total	Type 1	Type 2	Type 3	Type 4	*p*-value†	OR (95%CI)	*p*-value‡
rs738407	T/C	124/447/361	46/200/283	12/51/37	10/28/35	4/14/11	20/107/200	1.0×10^{-7}	1.56(1.32–1.83)	3.4×10^{-5}
		(0.627)	(0.724)	(0.625)	(0.671)	(0.621)	(0.775)			
rs738409	C/G*	247/468/217	88/236/203	20/59/21	21/30/22	8/11/9	39/136/151	1.4×10^{-10}	1.66(1.43–1.94)	3.6×10^{-6}
		(0.484)	(0.609)	(0.505)	(0.507)	(0.518)	(0.672)			
rs2076211	C/T*	248/473/211	92/242/195	21/58/21	21/30/22	8/11/10	42/143/142	1.4×10^{-9}	1.61(1.38–1.87)	3.2×10^{-5}
		(0.480)	(0.597)	(0.500)	(0.507)	(0.534)	(0.653)			
rs2896019	T/G*	246/473/213	91/234/204	20/57/23	22/29/22	7/12/10	42/136/149	1.5×10^{-10}	1.66(1.42–1.93)	2.6×10^{-5}
		(0.482)	(0.607)	(0.515)	(0.500)	(0.552)	(0.664)			
rs1010023	T/C*	249/473/210	94/239/196	21/57/22	22/29/22	7/12/10	44/141/142	1.5×10^{-9}	1.61(1.38–1.87)	6.5×10^{-5}
		(0.479)	(0.596)	(0.505)	(0.500)	(0.552)	(0.650)			
rs926633	G/A*	247/474/211	93/237/199	21/56/23	22/29/22	7/12/10	43/140/144	7.5×10^{-10}	1.62(1.39–1.89)	5.8×10^{-5}
		(0.481)	(0.600)	(0.510)	(0.500)	(0.552)	(0.654)			
rs3810622	T*/C	330/445/157	263/208/58	40/48/12	28/29/16	14/12/3	181/119/27	1.0×10^{-7}	0.64(0.55–0.75)	0.0017
		(0.407)	(0.306)	(0.360)	(0.418)	(0.310)	(0.265)			

Reference (A1) and non-reference (A2) alleles refer to NCBI Reference Sequence Build 36.3 with the effective allele marked by an asterisk. Genotyping results are shown by genotype count of A1A1/A1A2/A2A2 with allele frequency of A2 in parenthesis.
†*P*-values are calculated by exact trend test with odds ratios (OR) calculated for A2 with 95% confidence interval (CI).
‡*P*-values are calculated by Jonckheere-Terpstra test in NAFLD patients for Matteoni type and additive model of genotype. SNPs are ordered by chromosomal location.
doi:10.1371/journal.pone.0038322.t002

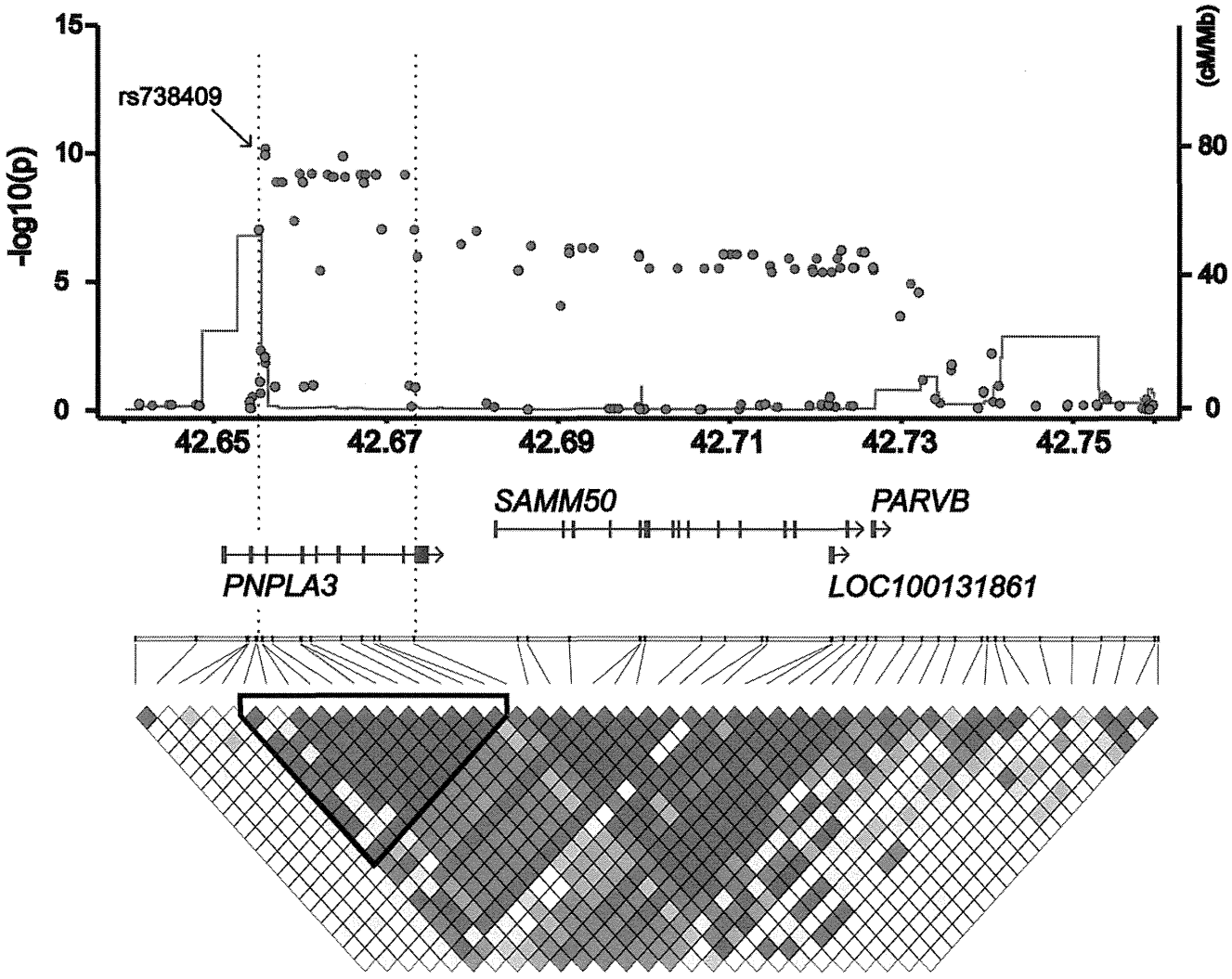


Figure 2. A schematic organization of the human *PNPLA3* locus at 22q13.31 with the genome scan results. *P*-values calculated by the exact trend test were plotted in $-\log_{10}$ scale. Red and blue dots indicate the *p*-values of genotyped and imputed SNPs, respectively. Local recombination rate obtained from HAPMAP release 22 is indicated by a red line plotted in cm/Mb scale. The structure and orientation of four genes in the region are shown below the plots with their transcriptional orientations according to NCBI Reference Sequence Build 36.3. LD blocks were generated according to pairwise LD estimates of the SNPs located within the region using the genome scan results. The LD block showing the strongest association is highlighted with the triangle, and the corresponding chromosomal region is represented by the dotted lines.
doi:10.1371/journal.pone.0038322.g002

chromosomes in the Japanese result of the International HapMap Project).

Discussion

NASH is a type of hepatic steatosis in NAFLD with poor prognosis accompanying liver fibrosis, and subsequent liver cirrhosis and hepatocellular carcinoma [18]. Despite the extensive biochemical and histological investigation of NAFLD, whether or not NASH forms a distinct disease entity in NAFLD still remains unclear. The principle aim of this study was to identify the genetic factors related to the pathogenic status of NAFLD by collecting DNA samples of Japanese NAFLD patients with critically diagnosed disease status by liver biopsy. To our knowledge, this is the first GWA study of NAFLD using patients with known histology-based Matteoni type. In the initial association study using pooled genotyping results of all the cases, we found a significant association of the *PNPLA3* gene at chromosome 22q13.31 with NAFLD in the Japanese. Rs738409 which showed the strongest association with NAFLD in the GWA study of Caucasians was also genotyped and its strongest association with NAFLD was confirmed. These results were in agreement with the former GWA analyses in populations of European descent and in Hispanics, giving strong evidence of the involvement of *PNPLA3* in NAFLD beyond ethnicities. Rs738409 is located in exon3 of the *PNPLA3* gene which is expressed in the liver and adipose tissue. This SNP introduces an amino acid substitution from isoleucine to methionine (I148M), and biological studies demonstrated that its risk allele (G) abolishes the triglyceride hydrolysis activity of *PNPLA3* [19]. These observations strongly suggest rs738409 to be a causative genetic variation for NAFLD. However, future genomic analyses by fine mapping or extensive sequencing may identify additional genetic determinants within the *PNPLA3* locus.

In the current study we did not find other genetic loci showing genome-wide significance ($p<1.0\times10^{-7}$). However, two additional chromosomal loci with *p*-values being smaller than 1×10^{-5} were

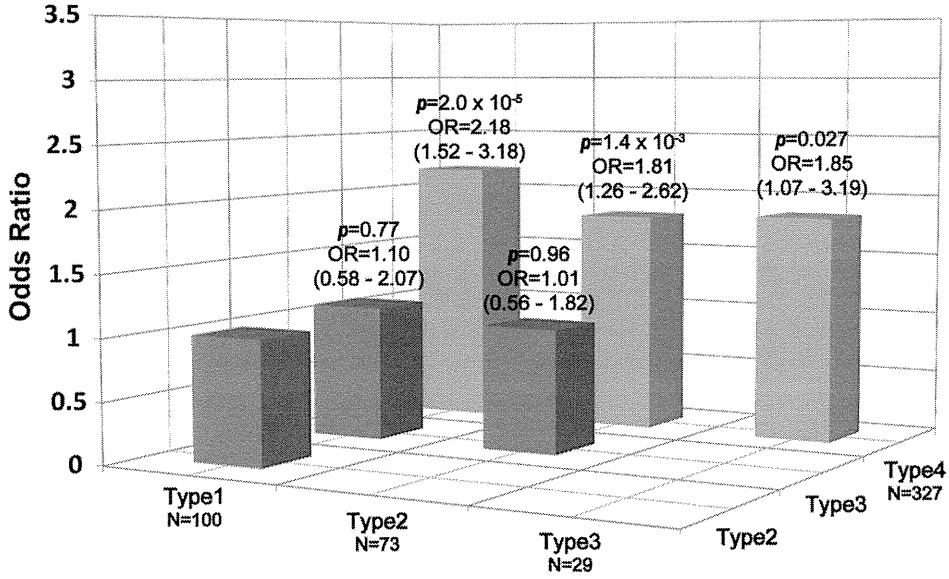


Figure 3. Histogram of odds ratios for genotype distribution of rs738409 between Matteoni types. Each box denotes the odds ratio (OR) comparing the corresponding Matteoni types on the horizontal axes. N represents the number of samples. Odds ratios and *p*-values are calculated for the higher Matteoni type per risk allele (G) on additive model by multivariable logistic regression adjusted for age, sex and BMI, and are shown with 95% CI above each box.
doi:10.1371/journal.pone.0038322.g003

identified on chromosome 1p (rs11206226) and chromosome 4p (rs1390096) neither of which has been reported as being associated with NAFLD in Caucasians (Table S1). Statistical calculation by taking their allele frequencies and effect sizes into account showed that approximately three times as many case and control samples are required to obtain sufficient statistical power (>0.8) for genome wide significance. Hence, further confirmation is required using a larger collection of patients and controls although they may be potential candidates of low-penetrance genes for susceptibility to NAFLD in Japanese.

Subsequent analyses through comparison of genotype distribution among four subgroups of NAFLD (type1 to type4) categorized by Matteoni's classification revealed that the seven NAFLD-associated SNPs in the *PNPLA3* gene were also significantly associated with the pathogenic status of NAFLD. There were also marked differences in genotype distribution of rs738409 between type4 subgroup and the other three groups ($p = 4.8 \times 10^{-6}$, OR = 1.96, 95%CI: 1.47–2.62 between type4 and pooled genotypes of type1 to type3). Moreover, a case/control analysis of rs738409 between Matteoni type4 and controls returned a surprisingly strong association ($p = 1.7 \times 10^{-16}$) which was much stronger than the initial analysis using all NAFLD cases ($p = 1.4 \times 10^{-10}$), whereas the analysis using Matteoni type1 to type3 as cases didn't show significance ($p = 0.41$). There were differences in the score of HOMA-IR and hs-CRP, indicators of insulin resistance and inflammation, respectively, between Matteoni type1 to type3 and type4 subgroups (Table 1). Our results provide compelling evidence that NASH corresponding to Matteoni type4 is both a clinically and genetically different disease subset from other spectrums of NAFLD. Previous studies showed association between *PNPLA3* and fatty liver, inflammation, fibrosis grade and NASH [13]. In our result, strong association between rs738409 and fatty liver was not observed by comparing control and Matteoni type1. In addition, strong association between rs738409 and lobular inflammation was not observed by comparing Matteoni Type1 and Type2. In contrast, a strong association between rs738409 and NASH was observed. Although we could not observe the strong association between rs738409 and fibrosis stage, strong association between rs738409 and Hyaluronic acid suggests that an association exists between *PNPLA3* and fibrosis.

We have also undertaken association analyses of rs738409 and clinical traits in the patients. The multivariable regression analysis adjusted for age, sex, BMI and Matteoni type followed by the correction for multiple testing revealed hyaluronic acid and HbA1c as being significantly associated with rs738409. Hyaluronic acid is one of the principle components of the extracellular matrix and its involvement in fibrosis has been previously suggested [20]. This may indicate another possible functional involvement of *PNPLA3* in the progression of liver fibrosis by influencing the circulating hyaluronic acid levels. A weak association of rs738409 and HbA1c levels was observed in our study population. However, there are no reports to date indicating such an association, and confirmation with different sample sets is needed for definitive conclusion. Also, the association between rs738409 and iron deposition was demonstrated by an ordinal logistic regression analysis. Since the association still remained after the results were adjusted with Matteoni type, rs738409 may play a functional role in the oxidative stress through iron absorption in the liver.

Recently, a genetic analysis of Japanese NAFLD patients was reported demonstrating a significant association in the increase of AST, ALT, ferritin levels and fibrosis stage (Brunt stage) and in the decrease of serum triglyceride with the risk allele (G) of rs738409 [12]. In our study, the association of rs738409 with AST ($p = 1.2 \times 10^{-4}$) and ALT ($p = 0.0016$) was reproduced and that of AST still remained after the results were adjusted for Matteoni type ($p = 0.038$). No association was observed for ferritin level. Brunt stage was available for Matteoni type4 patients only in our study. Although the odds ratio was slightly high (OR = 1.28, 95%CI: 0.95–1.72), it was not possible to examine the association. In addition, the inverse association of the risk allele of rs738409 with decrease of serum triglyceride was confirmed in our study ($p = 0.013$ after being adjusted for Matteoni type). For all of these

Table 3. Association of rs738409 with clinical traits.

Biochemical traits	Statistical calculation1		Statistical calculation 2	
Phenotype	**Coef. (S.E.)**	***p*-value**	**Coef. (S.E.)**	***p*-value**
Biological traits				
AST (IU/L)	0.22 (0.056)	**1.2×10^{-4}**	0.11 (0.052)	0.038
ALT (IU/L)	0.19 (0.058)	**0.0016**	0.093 (0.056)	0.098
GGT (IU/L)	−0.056 (0.061)	0.37	−0.088 (0.062)	0.16
Albumin (g/dL) *	0.015 (0.051)	0.77	−0.012 (0.052)	0.81
Total bilirubin (mg/dL)	−0.011 (0.063)	0.86	0.0059 (0.064)	0.93
Cholinesterase (unit) *	0.062 (0.040)	0.12	0.069 (0.041)	0.092
Type IV collagen 7S (ng/dL) *	−0.19 (0.064)	0.0025	−0.11 (0.062)	0.069
Hyaluronic acid (ng/dL)	0.30 (0.065)	**4.9×10^{-6}**	0.22 (0.063)	**4.6×10^{-4}**
Triglycerides (mg/dL)	−0.10 (0.058)	0.072	−0.15 (0.059)	0.013
Total cholesterol (mg/dL)	−0.066 (0.060)	0.27	−0.057 (0.061)	0.34
HbA1c (%)	−0.17 (0.053)	**0.0012**	−0.18 (0.054)	**0.0011**
IRI (μg/dL)	0.16 (0.063)	0.012	0.086 (0.061)	0.16
FPG (mg/dL)	−0.14 (0.049)	0.0047	−0.15 (0.05)	0.0035
HOMA-IR	0.084 (0.064)	0.19	0.0092 (0.062)	0.88
Hs-CRP (mg/dL)	−0.013 (0.048)	0.79	−0.031 (0.049)	0.52
Adiponectin (μg/mL)	0.048 (0.066)	0.47	0.12 (0.066)	0.072
Leptin (ng/mL)	0.11 (0.068)	0.11	0.10 (0.069)	0.15
Ferritin (ng/mL)	0.031 (0.047)	0.51	−0.0042 (0.048)	0.93
Uric acid (mg/dL)	−0.097 (0.061)	0.11	−0.11 (0.062)	0.067
PLT (x10^4/μL)	−0.056 (0.020)	0.0052	−0.045 (0.020)	0.028
Immunological/histological traits				
ANA (0/1/2/3/4)	0.92 (0.70–1.21)	0.56	0.86 (0.65–1.15)	0.31
Brunt grade (1/2/3)	1.42 (1.06–1.92)	0.021	1.38 (1.02–1.87)	0.036
Brunt stage (1/2/3/4)	1.28 (0.95–1.72)	0.11		
Fat deposition (1/2/3/4)	1.44 (1.15–1.81)	0.0019	1.24 (0.98–1.56)	0.76
Iron deposition (0/1/2/3/4)	0.61 (0.47–0.80)	**3.0×10^{-4}**	0.62 (0.47–0.81)	**5.6×10^{-4}**

Associations between distribution of rs738409 genotypes and clinical traits are calculated by multivariable regression. Statistical calculation1 is adjusted for age, sex and BMI, while the Matteoni types are additionally included as covariate in statistical calculation 2. Statistics are calculated by multivariable linear regression for biochemical traits and by multivariable ordinal logistic regression for immunological and histological traits.

Coefficients and odds ratios are calculated for the increase of each trait per risk allele (G). The *p*-values showing significance after Bonferroni's correction for multiple testing ($p = 0.0021$) was shown in bold.

*Reciprocal numbers are used for normalization and a negative coefficient implicates an increase in value according to the increase of the risk allele.

doi:10.1371/journal.pone.0038322.t003

biomarkers, however, the significance was lost after the correction for multiple testing.

A replication analysis of other genetic loci that had been reported for their association with NAFLD in East coast white Americans [14] was performed in our sample collection. We confirmed the association of rs780094 in *GCKR* with NAFLD in a case/control analysis but at a much weaker level ($p = 0.011$, OR = 0.82, 95%CI: 0.70–0.95) than that shown for the populations of European-descent. No associations were found for *LYPLAL1* and *NCAN* loci in our study. There are several reasons to explain such differences, such as the insufficient statistical power with a limited number of study subjects in our study due to the difficulty in the collection of a larger number of histologically diagnosed NAFLD patients. The difference in genetic background between the Japanese and Europeans is also conceivable. Indeed, the risk allele frequency of rs12137855 in *LYPLAL1* was 0.944 in our control subjects but approximately 0.79 in the European populations [14]. Similarly, there was a difference in the risk allele frequency of rs2228603 in *NCAN* (0.049 in Japanese and 0.08 in Europeans). Rs4240624 in *PPP1R3B* was not polymorphic in the Japanese while its risk allele frequency was 0.91 in Europeans.

Materials and Methods

Ethics Statement

In compliance with the Declaration of Helsinki, ethical approval for this study was given by the respective Institutional Review Board and subject written informed consent were obtained for all subjects (Ethical committee of Nara City Hospital; Ethical committee of Saiseikai Suita Hospital; Medical Ethics Committee of Kanazawa University; Ethics committee of Kyoto Prefectural University of Medicine; Ethical Committee of Aichi Cancer Center; Ethical Committee of Kochi Medical School, Kochi University; Ethics Committee of Tokyo Women's Medical University; Ethical Committee on Kawasaki Medical School and Kawasaki Medical School Hospital; Ethical Committee of

Juntendo University; Ethics Committee of Yamagata University School of Medicine; Ethical Committee of the Ikeda Municipal Hospital; Institutional Review Board and Ethics Committee of Kyoto University School of Medicine).

Study Population

A total of 543 patients histologically diagnosed for NAFLD in 2007–2009 were recruited through the Japan study of Nonalcoholic Fatty Liver Disease. Biopsy specimens were stained with H&E and Masson's trichrome for morphological review and assessment of fibrosis. Perl's Prussian blue was performed to evaluate iron load. Biopsy specimens were reviewed by a hepatopathologist (T.O). NAFLD patients were classified into four categories by liver histology according to the classification by Matteoni *et al* [2] as follows; type1: fatty liver alone, type2: fat accumulation and lobular inflammation, type3: fat accumulation and ballooning degeneration, type4: fat accumulation, ballooning degeneration, and either Mallory-Denk body or fibrosis. With these criteria, the 543 patients were classified as type1; 102, type2; 75, type3; 31 and type4; 335. The histological grade and fibrosis stage were also evaluated by the classification of Brunt *et al* [21] for advanced NAFLD cases (type3 and type4) as follows; grade 1: steatosis involving up to 66% of biopsy, occasional ballooned zone 3 hepatocytes and absence or mild portal chronic inflammation, grade2: steatosis, ballooning hepatocytes mild to moderate chronic inflammation, grade3: panacinar steatosis, ballooning and disarray obvious and mild or portal mild to moderate inflammation, stage1: perivenular and/or perisinusoidal fibrosis in zone3, stage2: combined pericellular portal fibrosis, stage3: septal/bridging fibrosis, stage4: cirrhosis. The degree of fat deposition was evaluated by amount of fat droplets as observed under the microscope as follows; 0: <5%, 1: 5–<10%, 2: 10–<34%, 3: 34–<67%, 4: >67%. The degree of iron deposition was categorized by the presence of granules of free iron observed under the microscope as follows; 0: absence by x400, 1: easily identifiable by x400 and rarely identifiable by x250, 2: identifiable by x100, 3: identifiable by x25, 4: identifiable at lower than x25.

Inclusion criteria for NAFLD patients were as follows; (i) no history of alcoholism, (ii) no history for HBV/HCV/HIV infection, (iii) diagnosed by liver biopsy, (iv) information regarding age and BMI available. The sex of two samples was unknown, and was imputed from the results of the genome scan. As general Japanese population controls, the genome scan results of 942 healthy Japanese volunteers from Aichi Cancer Center Hospital and Research Institute were used [22].

Anthropometric and Laboratory Evaluation

We employed conventional methods for the measurement of anthropometry (height, weight, amount of visceral fat and abdominal circumscript). BMI was calculated from the measurements. The following biochemical/hematological/immunological traits were also measured by conventional methods; aspartate aminotransferase (AST), alanine aminotransferase (ALT), γ-glutamyl transpeptidase (GGT), albumin, total bilirubin, cholinesterase, type IV collagen 7S, hyaluronic acid, triglyceride, total cholesterol, hemoglobin A1c (HbA1c), fasting immunoreactive insulin (IRI), fasting plasma glucose (FBS), high sensitive CRP (hs-CRP), adiponectin, leptin, ferritin, uric acid, and platelet (PLT) count. Anti nuclear antibody (ANA) was measured by ELISA and categorized by the detection limit in a serial dilution as follows; 0: <40x, 1: 40–80x, 2: 81–160x, 3: 160x, 4: >320x. Homeostasis model assessment-insulin resistance (HOMA-IR) was calculated from the measurements. Patients were assigned a diagnosis of diabetes mellitus (DM) when they had documented use of oral hypoglycemic medication, a random glucose level >200 mg/dl, or FPG >126 mg/dl. Hyperlipidemia was diagnosed with the cholesterol level being >200 mg/dl and/or triglyceride level being >160 mg/dl. Hypertension was diagnosed when the patient was taking antihypertensive medication and/or had a resting recumbent blood pressure ≥140/90 mmHg on at least two occasions.

DNA Preparation

Genomic DNA was extracted from peripheral blood mononuclear cells by standard phenol-chloroform extraction and resuspended in TE buffer. DNA concentration and purity were measured with Nanodrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). The samples were stored at −20°C until use.

Genome-wide Genotyping and Quality Control

Genome scan was conducted for 543 patients with NAFLD and 942 healthy subjects using Illumina Human 610-Quad Bead Chip on a Bead Station 500G Genotyping System (Illumina, Inc., San Diego, CA, USA) and subjected to the following quality controls. Initially, ten patients and six control subjects were removed due to low call rates (<0.99). Regarding the SNP markers, 85,472 SNPs with minor allele frequency of smaller than 0.01 in either case or control group, 6,479 SNPs with lower success rates (<0.98) and 35 SNPs with distorted Hardy-Weinberg equilibrium ($p<10^{-7}$) were removed, resulting in 484,751 SNP markers being used for analysis. Principal component analysis by EIGENSOFT [17] including phase II HapMap (http://hapmap.ncbi.nlm.nih.gov/) samples identified no samples that were deviated from the Japanese population. Subsequently, the degree of kinship between individuals was examined by pi-hat in PLINK 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) [23]. Of the eight pairs of samples (four case pairs and four control pairs) showing high degrees of kinship (PI-HAT>0.4), the sample with the lower call rate in each pair was removed. After these steps, 529 case and 932 controls were used for the analysis.

Statistical Analysis

A case/control association analysis was performed by exact trend test between NAFLD patients and control subjects [24]. The correction of obtained *p*-values for population stratification was performed using EIGENSTRAT [17]. In addition, an association between Matteoni classification (type1 to type4) and additive model of genotype for each SNP was examined using Jonckheere-Terpstra test for NAFLD patients. Assessment of population stratification of inflation of *p*-value was carried out by the genomic control method for asymptotic trend test [25]. Association between each quantitative trait and the genotype of significant SNPs in NAFLD patients were calculated by multivariable linear regression or multivariable ordinal regression adjusted for age, sex and BMI. Each quantitative trait was transformed as follows; natural log for ALT, AST, HOMA-IR, HbA1c, IRI, triglyceride, total bilirubin, adiponectin, hs-CRP, hyaluronic acid, leptin, reciprocal number for albumin, cholinesterase, type IV collagen 7S and square root for uric acid and ferritin. The values of FPG, PLT, total cholesterol, amount of visceral fat, and abdominal circumscript were not transformed. For each trait, values that were within only 4 S.D. were included for analysis. LD indices were calculated by default setting of Haploview [26] and the LD block was defined manually.

Table 4. Replication study of previously reported SNPs.

			Genotyping Result and Allele Frequency of A2					Statistics		
			NAFLD					NAFLD vs. Control		Matteoni
dbSNPID	A1/A2	Gene	Control	Type 1	Type 2	Type 3	Type 4	*p*-value†	OR (95%CI)	*p*-value‡
rs12137855	C*/T	*LYPLAL1*	828/102/2	90/10/0	67/6/0	24/5/0	294/33/0	0.55	0.89	0.98
			(0.056)	(0.050)	(0.041)	(0.086)	(0.050)		(0.64–1.25)	
rs780094	T*/C	*GCKR*	321/433/178	34/54/12	28/34/11	17/11/1	133/139/55	0.011	0.82	0.92
			(0.423)	(0.390)	(0.383)	(0.224)	(0.381)		(0.70–0.95)	
rs4240624	G/A	*PPP1R3B*	–	–	–	–	–	–	–	–
rs2228603	C/T*	*NCAN*	842/88/2	93/7/0	65/8/0	28/1/0	292/31/4	0.80	1.05	0.58
			(0.049)	(0.035)	(0.054)	(0.017)	(0.059)		(0.75–1.48)	

Reference (A1) and non-reference (A2) alleles refer to NCBI Reference Sequence Build 36.3 with the effective allele marked by an asterisk. Genotyping results are shown by genotype count of A1A1/A1A2/A2A2 with allele frequency of A2 in parenthesis. †*P*-values are calculated by exact trend test with odds ratios (OR) calculated for A2 with 95% confidence interval (CI). ‡*P*-values are calculated by Jonckheere-Terpstra test in NAFLD patients for Matteoni type and additive model of genotype.
doi:10.1371/journal.pone.0038322.t004

Supporting Information

Figure S1 QQ plot of the GWA study comparing distribution of the observed and expected *p*-values. Upper box is expressed in antilog scale and the lower box is expressed in $-\log_{10}$ scale. The X- and Y-axis correspond to expected and observed *p*-values. Blue and red dots denote before and after correction by genomic control method ($\lambda = 1.04$), respectively.
(DOC)

Table S1 List of the SNPs showing $p<1.0\times10^{-5}$ in the GWA study. Reference (A1) and non-reference (A2) alleles refer to NCBI Reference Sequence Build 36.3 with the effective allele marked by an asterisk. Genotyping results are shown by genotype count of A1A1/A1A2/A2A2 with allele frequency of A2 in parenthesis. †*P*-values are calculated by exact trend test with odds ratios (OR) calculated for A2 with 95% confidence interval (CI). ‡*P*-values are calculated by Jonckheere-Terpstra test in NAFLD patients for Matteoni type and additive model of genotype. SNPs are ordered by chromosomal location.
(DOC)

Acknowledgments

The authors would like to thank Yutaka Kohgo, Hirofumi Uto and Tetsuo Takehara for sample collection and Hisako Imamura and Hiroyuki Uneme for data management.

Author Contributions

Conceived and designed the experiments: FM TO. Performed the experiments: MT M. Kokubo. Analyzed the data: TK RY FM. Contributed reagents/materials/analysis tools: TK YS AU KM MT TT KY T. Saibara EH M. Kokubo SW SK YI M. Kawanaka T. Shima HP HT KT RY. Wrote the paper: TK MT RY FM TO.

References

1. Ludwig J, Viggiano TR, McGill DB, Oh BJ (1980) Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. Mayo Clin Proc 55: 434–438.
2. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, et al. (1999) Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. Gastroenterology 116: 1413–1419.
3. Cohen JC, Horton JD, Hobbs HH (2011) Human fatty liver disease: old questions and new insights. Science 332: 1519–1523. doi:10.1126/science.1204265.
4. Vernon G, Baranova A, Younossi ZM (2011) Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther 34: 274–285. doi:10.1111/j.1365–2036.2011.04724.x.
5. Okanoue T, Umemura A, Yasui K, Itoh Y (2011) Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis in Japan. J Gastroenterol Hepatol 26 Suppl 1: 153–162. doi:10.1111/j.1440–1746.2010.06547.x.
6. Williams CD, Stengel J, Asike MI, Torres DM, Shaw J, et al. (2011) Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. Gastroenterology 140: 124–131. doi:10.1053/j.gastro.2010.09.038.
7. Okanoue T (2011) Recent progress in the research of NASH/NAFLD in Japan. Nihon Shokakibyo Gakkai Zasshi 108: 1161–1169.
8. Berson A, De Beco V, Lettéron P, Robin MA, Moreau C, et al. (1998) Steatohepatitis-inducing drugs cause mitochondrial dysfunction and lipid peroxidation in rat hepatocytes. Gastroenterology 114: 764–774.
9. Day CP (2006) From fat to inflammation. Gastroenterology 130: 207–210. doi:10.1053/j.gastro.2005.11.017.
10. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, et al. (2008) Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 40: 1461–1465. doi:10.1038/ng.257.
11. Sookoian S, Castaño GO, Burgueño AL, Gianotti TF, Rosselli MS, et al. (2009) A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res 50: 2111–2116. doi:10.1194/jlr.P900013-JLR200.
12. Hotta K, Yoneda M, Hyogo H, Ochi H, Mizusawa S, et al. (2010) Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease. BMC Med Genet 11: 172. doi:10.1186/1471–2350–11–172.
13. Sookoian S, Pirola CJ (2011) Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. Hepatology 53: 1883–1894. doi:10.1002/hep.24283.
14. Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, et al. (2011) Genome-wide association analysis identifies variants associated with nonalcoholic Fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 7: e1001324. doi:10.1371/journal.pgen.1001324.
15. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, et al. (2010) Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease. N Engl J Med 362: 1082–1089. doi:10.1056/NEJMoa0907295.
16. Terao C, Yamada R, Ohmura K, Takahashi M, Kawaguchi T, et al. (2011) The human AIRE gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. Human Molecular Genetics 20: 2680–2685. doi:10.1093/hmg/ddr161.
17. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38: 904–909. doi:10.1038/ng1847.
18. Yasui K, Hashimoto E, Komorizono Y, Koike K, Arii S, et al. (2011) Characteristics of patients with nonalcoholic steatohepatitis who develop hepatocellular carcinoma. Clin Gastroenterol Hepatol 9: 428–433; quiz e50. doi:10.1016/j.cgh.2011.01.023.