Original article

Association of ZFHX3 gene variation with atrial fibrillation, cerebral infarction, and lung thromboembolism: An autopsy study

Khin Thet Thet Zaw (MB, BS)\textsuperscript{a}, Noriko Sato (MD, PhD)\textsuperscript{a}, Shinobu Ikeda (DDS, PhD)\textsuperscript{b}, Kaung Si Thu (MB, BS)\textsuperscript{a}, Makiko Naka Mieno (PhD)\textsuperscript{c}, Tomio Arai (MD, PhD)\textsuperscript{d,e}, Seijiro Mori (MD, PhD)\textsuperscript{f}, Tetsushi Furukawa (MD, PhD)\textsuperscript{g}, Tetsuo Sasano (MD, PhD)\textsuperscript{h}, Motoji Sawabe (MD, PhD)\textsuperscript{i}, Masashi Tanaka (MD, PhD)\textsuperscript{j}, Masaaki Muramatsu (MD, PhD)\textsuperscript{a,*}

\textsuperscript{a} Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan
\textsuperscript{b} Data Science Department, JCRAC Data Center, Clinical Research Center, National Medical Research Center Hospital, Tokyo, Japan
\textsuperscript{c} Department of Medical Informatics, Center of Information, Jichi Medical University, Tochigi, Japan
\textsuperscript{d} Department of Pathology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan
\textsuperscript{e} Research Team for Geriatric Pathology, Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan
\textsuperscript{f} Center for Promotion of Clinical Investigation, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan
\textsuperscript{g} Department of Bioinformational Pharmacology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan
\textsuperscript{h} Department of Biofunctional Informatics, Tokyo Medical and Dental University, Tokyo, Japan
\textsuperscript{i} Department of Molecular-genetic Sciences, Division of Biomedical Laboratory Sciences, Molecular Pathophysiology, Graduate School of Health Care Sciences, Tokyo Medical and Dental University, Tokyo, Japan
\textsuperscript{j} Department of Clinical Laboratory, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

\textsuperscript{*} Corresponding author at: Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan. Fax: +81 3 5803 0247.

\textit{E-mail address: mura.matsu.epi@mri.tmd.ac.jp} (M. Muramatsu).

\textbf{A R T I C L E   I N F O}

Article history:
Received 14 September 2016
Received in revised form 28 October 2016
Accepted 11 November 2016
Available online 19 December 2016

Keywords:
ATBF1
Atrial fibrillation
Cerebral infarction
Lung thromboembolism
Single nucleotide polymorphism

\textbf{A B S T R A C T}

\textbf{Aim:} We aimed to study a single nucleotide polymorphism (SNP), rs2106261, in the transcription factor gene, \textit{ZFHX3}, in atrial fibrillation (AF) and other related phenotypes by phenome scanning in a Japanese population.

\textbf{Method:} We retrieved consecutive autopsy data \((n = 2433, \text{mean age} = 80 \text{ years})\) from the Japanese SNP database for geriatric diseases (JG-SNP). Clinical data, including an AF diagnosis, were collected from medical charts. Genotyping was performed with the DNA chip method. We also analyzed 42 pathological and 26 clinical phenotypes, including cerebral infarctions (CIs) and lung thromboembolisms (LTs), diagnosed by macroscopic inspection during the autopsy.

\textbf{Result:} Among the 2433 patients with available data, 18.6\% had AF, 29.4\% had CI, and 4.9\% had LT phenotypes. The A allele of the rs2106261 SNP was significantly associated with AF, after adjusting for age, sex, diabetes, hypertension, and smoking \((\text{AA} + \text{AG}/\text{GG}, \text{OR} = 1.51, 95\%\text{CI}: 1.16–1.97, p = 0.002)\). In the entire cohort, CI was not associated with rs2106261 \((p = 0.14)\). However, among patients under 80 years old, rs2106261 was significantly associated with CI \((\text{AA} + \text{AG}/\text{GG}, \text{OR} = 1.57, 95\%\text{CI}: 1.09–2.26, p = 0.01)\). LT was also associated with rs2106261 \((\text{AA} + \text{AG}/\text{GG}, \text{OR} = 1.99, 95\%\text{CI}: 1.31–3.01, p = 0.001)\).

\textbf{Conclusion:} We showed that the \textit{ZFHX3} polymorphism, rs2106261 \((\text{A allele})\), was a risk marker for AF and AF-related phenotypes. The roles of this variant in the development of AF and its related phenotypes warrant further investigation.

\textcopyright 2016 Japanese College of Cardiology. Published by Elsevier Ltd. All rights reserved.

\textbf{Introduction}

Atrial fibrillation (AF) is the most common cardiac arrhythmia; it affects more than 30 million people worldwide. The overall prevalence of AF is 0.56\%, but it increases with aging. The prevalence is <0.1\% in adults under 55 years old and 8\% in adults 80 years and above [1]. Additionally, among men and women over...
80 years old, the prevalence is 4.4% and 2.2%, respectively [2]. The prevalence of AF has increased in Japan, as the population has aged, similar to the phenomenon in Western countries. The accumulating number of patients with AF is projected to be 1.03 million in 2050, with an overall prevalence of 1.09%. A Danish twin study estimated that the heritability of AF was 62%. That finding indicated that genetic factors should play a substantial role in the risk of AF [3]. A number of genes that confer AF susceptibility were identified in genome wide association studies (GWAS) [4]. One of the major AF susceptibility-conferring genes was a transcription factor, ZFHX3, which was initially found by scanning with the single nucleotide polymorphism (SNP), rs2106261, in a Caucasian population. This finding was subsequently replicated in other populations [5–7]. The current study aimed to replicate the finding that rs2106261 had an effect on AF in a Japanese population. We took advantage of autopsy samples and employed a phenome scan with the rs2106261 probe. In addition to AF, we found signs that rs2106261 was positively associated with both cerebral infarctions (CIs) and lung thromboembolisms (LTs).

Materials and methods

Study population

We studied consecutive autopsy cases performed at Tokyo Metropolitan Geriatric Hospital from 1995 to 2006, which were recorded in the Japanese SNP database for geriatric research (JG-SNP). Initially 2433 elderly subjects were enrolled in this study. Due to unavailable pathological clinical history, missing DNA samples, and unsuccessful genotyping, we had to exclude 257 subjects for AF analysis, 921 subjects for CI, and 595 subjects for LT, and thus 2176 subjects for AF, 1512 subjects for CI, and 1838 subjects for LT were included in this study. The JG-SNP database was constructed to facilitate studies on the roles of genetic polymorphisms in geriatric diseases. This SNP database was created by storing the clinical data and pathological information for each autopsy together with genetic polymorphism data. There were 42 entities of pathological phenotypes and 26 entities of clinical phenotypes. All subjects were aged Japanese individuals, and the details of the autopsy findings were described elsewhere [8]. Data on AF, other clinical parameters, and patient drinking and smoking habits were obtained retrospectively from medical charts, which were recorded either upon hospital admission or in the outpatient clinic. The diagnoses of CI and LT were determined in macroscopic pathology inspections performed during the autopsy. Written informed consent was obtained from a family member of each subject involved in the study before performing the autopsy. The study was approved by the ethics committees of the Tokyo Medical and Dental University (2009-19-8) and the Tokyo Metropolitan Institute of Gerontology (2009-482).

Genotyping

Genomic DNA was extracted from the renal cortex with a standard procedure. Variations in exomes were analyzed with an Infinium Human Exome Bead Chip array, version 1.1 (Illumina Inc., San Diego, CA, USA), which was scanned with iScan, in accordance with protocols provided by the manufacturer. The genotype calling rate was checked for all samples with Gencall, the standard method provided by Illumina, which was the genotyping module used by Genome Studio data analysis software, version 1.9. The pathological assessments and the genotyping were performed in different institutions in a double-blinded fashion to minimize bias.

Statistical analysis

We conducted a phenome scan, which included 26 clinical phenotypes and 42 pathological phenotypes, based on the rs2106261 SNP. A nominal p-value of <0.05 was considered statistically significant. We examined Hardy–Weinberg equilibrium with the Chi-square test to check the distribution of genotyping results. Continuous variables were analyzed with the student t-test, and results are expressed as the mean ± SD. Categorical variables were analyzed with the Pearson Chi-square test, and results are expressed as the mean and percentage of cases. We performed multivariate logistic regression analyses in a dominant disease penetrance model for patients with AF and control subjects, and we adjusted the model for associated risk factors, such as age at death, gender, and smoking. We performed statistical analyses with IBM SPSS, version 19.0 (IBM, Chicago, IL, USA). We used PLINK software, version 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) for allelic association testing [9].

Results

The patient demographics are shown in Table 1. The phenome scan of all samples with rs2106261 revealed positive associations with CI and LT, by genotype association (data not shown). Therefore, we included CI and LT in Table 1, together with AF. There were 452 cases and 1981 controls for AF, 716 cases and 796 controls for CI, and 119 cases and 1719 controls for LT. Age was a prominent risk factor for all three phenotypes. In addition, CI was associated with diabetes mellitus and hypertension, both well known risk factors for CI.

The distribution of three rs2106261 genotypes was consistent with Hardy–Weinberg equilibrium (Table 2). We found that the A allele of rs2106261 was present at a higher frequency in the AF group than in the control population (A = 0.35% vs. G = 0.29%, p = 0.0006). We found that the AA genotype of rs2106261 significantly increased the risk of AF (p = 0.003). Phenome scanning with rs2106261 followed by genotype-phenotype association tests (not shown) identified positive associations between

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Atrial Fibrillation</th>
<th>Cerebral Infarction</th>
<th>Lung Thromboembolism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, n (M/F)</td>
<td>255/197</td>
<td>391/325</td>
<td>58/61</td>
</tr>
<tr>
<td>Age, years (Mean ± SD)</td>
<td>82.8 ± 8.2</td>
<td>81.5 ± 8.7</td>
<td>82.1 ± 9.5</td>
</tr>
<tr>
<td>DM, n (%)</td>
<td>43 (9.51)</td>
<td>121 (16.9)</td>
<td>15 (12.6)</td>
</tr>
<tr>
<td>HT, n (%)</td>
<td>103 (22.8)</td>
<td>251 (35.0)</td>
<td>28 (23.5)</td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td>126 (27.8)</td>
<td>324 (45.2)</td>
<td>38 (47.2)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.42</td>
<td>0.93</td>
<td>0.018</td>
</tr>
</tbody>
</table>

DM, diabetes mellitus; HT, hypertension; M, male; F, female; SD, standard deviation.
Among the 1319 patients, AF was found in 236 patients (Table 4). This group comprised 1319 patients. We created a table that included only subjects without any missing data for all three phenotypes. This group comprised 1319 patients. The association remained positive (\( p = 0.03 \) and \( p = 0.02 \), respectively) after adjusting for the presence of AF.

### Table 2

<table>
<thead>
<tr>
<th>Genotype</th>
<th>All subjects, n (%)</th>
<th>AF, n (%)</th>
<th>Control, n (%)</th>
<th>p-value</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GG</td>
<td>1064 (48.9)</td>
<td>175 (42.6)</td>
<td>889 (50.4)</td>
<td>0.003†</td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>907 (41.7)</td>
<td>182 (44.3)</td>
<td>725 (41.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>205 (9.4)</td>
<td>54 (13.1)</td>
<td>151 (8.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G (%)</td>
<td>0.70</td>
<td>0.65</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (%)</td>
<td>0.30</td>
<td>0.35</td>
<td>0.29</td>
<td>0.0006</td>
<td>1.329</td>
</tr>
<tr>
<td>HWE (p-value)</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; OR, odds ratio. * p-value was calculated with a Chi-square analysis. † p-value was calculated with allelic association testing, performed with PLINK software.

### Table 3

<table>
<thead>
<tr>
<th>Genotype</th>
<th>AF</th>
<th>CI</th>
<th>CI (age &lt; 80 years)</th>
<th>CI (age ≥ 80 years)</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Cont</td>
<td>Case</td>
<td>Cont</td>
<td>Case</td>
</tr>
<tr>
<td>GA + AA</td>
<td>236 (74.6)</td>
<td>876 (49.6)</td>
<td>328 (52.2)</td>
<td>336 (48.3)</td>
<td>144 (54.5)</td>
</tr>
<tr>
<td>GG</td>
<td>175 (74.6)</td>
<td>889 (50.4)</td>
<td>300 (47.8)</td>
<td>359 (51.7)</td>
<td>120 (45.5)</td>
</tr>
<tr>
<td>p-value</td>
<td>0.002</td>
<td>0.14</td>
<td>0.01</td>
<td>0.08</td>
<td>0.001</td>
</tr>
<tr>
<td>OR (95%CI)</td>
<td>1.51 (1.16–1.97)</td>
<td>1.17 (0.94–1.46)</td>
<td>1.57 (1.09–2.26)</td>
<td>1.37 (0.96–1.96)</td>
<td>1.99 (1.31–3.01)</td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; CI, cerebral infarction; LT, lung thromboembolism; OR, odds ratio. | Case | Cont | Case | Cont | Case | Cont | Case | Cont | Case | Cont |

p-values were calculated with a logistic regression analysis, adjusted for sex, age, hypertension, diabetes, and smoking.

### Table 4

<table>
<thead>
<tr>
<th>Other disease</th>
<th>AF (−) (n = 1083)</th>
<th>AF (+) (n = 236)</th>
<th>OR (95%CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CI (−)</td>
<td>600 (55.4)</td>
<td>92 (39.0)</td>
<td>1.71 (1.24–2.35)</td>
<td>0.001</td>
</tr>
<tr>
<td>CI (+)</td>
<td>483 (44.6)</td>
<td>204 (86.4)</td>
<td>1.87 (1.05–3.34)</td>
<td>0.034</td>
</tr>
<tr>
<td>LT (−)</td>
<td>1030 (95.1)</td>
<td>32 (13.6)</td>
<td>1.329</td>
<td></td>
</tr>
<tr>
<td>LT (+)</td>
<td>53 (4.9)</td>
<td>181 (50.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AF, atrial fibrillation; CI, cerebral infarction; LT, lung thromboembolism; OR, odds ratio. | Case | Cont | Case | Cont | Case | Cont | Case | Cont | Case | Cont |

p-values were calculated with a logistic regression analysis, adjusted for sex, age, hypertension, diabetes, and smoking.

rs2106261 and both CI (\( p = 0.03 \)) and LT (\( p = 0.005 \)). Table 3 shows the associations between rs2106261 genotypes and AF, CI, and LT in a minor allele, dominant disease penetrance model. After adjusting the model for covariates of age, gender, hypertension, diabetes, and smoking, we found that rs2106261 remained significantly associated with both AF (OR = 1.51, 95% CI = 1.16–1.97, adjusted \( p = 0.002 \)) and LT (OR = 1.99, 95% CI = 1.31–3.01, adjusted \( p = 0.001 \)) (Table 3).

When we tested the association between rs2106261 and CI, we found no significant association, after adjusting for the confounding factors. However, when we divided the subjects into ages below and above the mean age (<80 and ≥80 years), we found that subjects <80 years showed a significant association between rs2106261 and CI (OR = 1.57, 95% CI = 1.09–2.26, adjusted \( p = 0.01 \)).

To determine the relationship between AF, CI, and LT, we created a table that included only subjects without any missing data for all three phenotypes. This group comprised 1319 patients. Among the 1319 patients, AF was found in 236 patients (Table 4). We studied the association between AF and CI, AF and LT and found that the presence of AF was associated with both the presence of CI (OR = 1.71, \( p = 0.001 \)) and the presence of LT (OR = 1.87, \( p = 0.034 \)). We also performed genetic association and causality analyses with a table that excluded missing phenotype data (Supplementary Table S1). For the CI (age <80 years) and LT phenotypes, the association remained positive (\( p = 0.03 \) and \( p = 0.02 \), respectively) after adjusting for the presence of AF.

### Discussion

ZFHX3 was initially identified as ATBF1, a transcription factor that encoded four homeobox domains and 23 zinc-fingers. ZFHX3 was shown to be involved in suppressing alpha-fetoprotein transcription [10]. It was identified as a candidate tumor suppressor gene for prostate, breast, and gastric cancers, which acted by inducing cell cycle arrest [11,12]. ZFHX3 was also found to function in neuronal differentiation in the developing brain [11], and it was involved in cardiac development by regulating myogenic differentiation [13,14]. In cell signaling, ZFHX3 interacts with a protein that specifically inhibits signal transducer and activator of transcription 3 (STAT3). This protein inhibitor of activated STAT 3 (PIAS3) is activated by ZFHX3 [15]. Interestingly, in pacing-induced AF models, tachypacing induces a decrease in ZFHX3 expression, which then activates STAT3 signaling via the reduction in PIAS3 activity. Thus, inflammatory processes in atrial tissues that arise with the down-regulation of ZFHX3 may affect atrial arrhythmogenesis [16,17]. Indeed, elevations in the inflammation marker, C-reactive protein, predict the presence and the future development of AF [16]. These lines of evidence may explain, at least in part, the contribution of ZFHX3 to the occurrence of AF in humans [18].

We have replicated the finding that the ZFHX3 polymorphism, rs2106261, affected AF in a Japanese population. We showed that the risk allele was the same as that shown in previous reports [4,19]. However, because the ZFHX3 polymorphism, rs2106261, is
In the intron, the mechanism underlying these SNP effects on AF has not been well established. One previous study examined the effect of ZFHX3 SNPs on ZFHX3 expression, but they found that rs2106261 did not affect the expression. In contrast, other SNPs did affect ZFHX3 expression. Those authors suggested that the effect of an intronic SNP might be specific to cardiac tissue [20].

AF is one of the most common types of sustained cardiac arrhythmias. AF is well known to cause CI, a type of ischemic stroke [21]. AF increases the risk of stroke by twofold over the stroke risk without AF [22]. AF also causes cardio-embolic stroke, due to irregular cardiac muscle contractions, which lead to blood pooling in the atria, mainly the left atrial appendage. Blood stasis is associated with increased concentrations of fibrinogen, u-dimer, and von Willebrand factor. These factors contribute to a prothrombotic state, which in turn, predisposes to thrombus formation with the consequential increase in the rate of cerebral embolization [23]. On the other hand, the relationship between LT and AF appears to be highly complex. It has been suggested that the link between AF and LT is pathophysiological [24]. Acute LT increases pulmonary vascular resistance and right ventricular afterload, by obstructing the pulmonary arteries and triggering the release of vasoconstrictive mediators. Consequently, the increases in right ventricular and atrial pressures elicit stretch injuries that, in turn, can trigger AF [25,26].

In the present study, our phenome scanning showed that the rs2106261SNP was associated with CI (for patients aged <80 years) and, even more strongly, with LT. A logistic regression analysis adjusted for the presence of AF revealed that the rs2106261 variant was associated with CI (for patients aged <80 years) and LT, independently. Thus, the rs2106261 may have substantial pleiotropic effects on AF and other AF-related traits through a common mechanism. We speculated that this mechanism could involve inflammatory acceleration through the ZFHX3-PJAS3-STAT3 pathway.

Although acute LT is a well-known phenotype, which typically follows deep vein thrombosis, the phenotype of chronic LT is not well known. We know that AF causes blood stasis in the atrium, which enhances platelet aggregation and coagulation. Therefore, it logically follows that direct embolization in the right atrium precedes LT [27–29]. In fact, LT has been shown to be a rare complication of AF [30]. Nevertheless, our observations suggested that a common ZFHX3 genotype is a risk factor for AF and LT independently, which suggested that LT might be a risk factor for AF. However, this study was limited by the small sample size; therefore, Grethe results may not be widely generalizable; they should only be considered useful for generating hypotheses.

Conclusion

We found that the ZFHX3 SNP, rs2106261, was associated with AF, with CI in patients aged <80 years, and with LT. The relationship between AF and LT warrants further investigation.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This study was supported by grants from the following: the Smoking Research Foundation (TA and MS); the Mitsui Sumitomo Insurance Welfare Foundation (SM); the Pfizer Academic Contribution Fund (SM); the Joint Usage/Research Program of the Medical Research Institute, Tokyo Medical and Dental University (MM); the Ministry of Education, Culture, Sports, Science, and Technology of Japan; GMEXT/JSPS KAKENHI, grant numbers A-25242062, A-22240072, B-21390459, C-26670481, C-21590411, and C-26540514 (MT); the Ministry of Health, Labor, and Welfare of Japan; Grants-in-Aid for Research on Intractable Diseases (Mito-chondrial Disorders), grants 23-016, 23-116, and 24-005 (MT); and the Takeda Science Foundation (MT). We would like to thank Ms. Yasuko Hasegawa and the staff at the Department of Pathology, Tokyo Metropolitan Geriatric Hospital, for preparing DNA samples, and Dr. Makiko Naka-Mieno, for data management and collection.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jcc.2016.11.005.

References

[8] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a toolset for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:599–75.


