Genome-wide association study identifies genetic determinants of warfarin responsiveness for Japanese

Pei-Chieng Cha1, Taisei Mushiroda2, Atsushi Takahashi3, Michiaki Kubo4, Shiro Minami5, Naoyuki Kamatani3 and Yusuke Nakamura1,*

1Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, 2Research Group for Pharmacogenomics, 3Laboratory for Statistical Analysis and 4Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, Yokohama, Japan and 5Department of Bioregulation, Nippon Medical School, Kawasaki, Japan

Received April 3, 2010; Revised and Accepted September 6, 2010

Warfarin is a commonly used anticoagulant, whose dose needs to be determined for each individual patient owing to large inter-individual variability in its therapeutic dose. Although several clinical and genetic variables influencing warfarin dose have been identified, uncovering additional factors are critically important for safer use of warfarin. Through a genome-wide association study, we identified single-nucleotide polymorphism (SNP) rs2108622 [cytochrome P450, family 4, subfamily F, polypeptide 2 (CYP4F2)] as a genetic determinant of warfarin responsiveness for Japanese. Stratifying subjects who have been pre-classified according to the genotypes of SNP rs10509680 [cytochrome P450, family 2, subfamily C, polypeptide 9 (CYP2C9)] and SNP rs9923231 [vitamin K epoxide reductase complex subunit 1 (VKORC1)], based on their genotypes of rs2108622 allowed identification of subjects who require higher dose of warfarin. Incorporating genotypes of rs2108622 into a warfarin dosing algorithm that considers age, body surface area, status of amiodarone co-administration and genotypes of SNPs in the CYP2C9 and VKORC1 genes improved the model’s predictability to 43.4%. In this study, the association of CYP4F2 with warfarin dose of the Japanese has been established for the first time. Besides, a warfarin dosing algorithm that incorporates genotypes of rs2108622 and amiodarone co-administration status was suggested for the Japanese. Our study also implied that common SNPs other than those in the CYP2C9, VKORC1 and CYP4F2 genes that show strong effect on the therapeutic warfarin dose might not exist.

INTRODUCTION

Warfarin is the most widely prescribed anticoagulant for thromboembolic therapy (1). However, owing to its narrow therapeutic index and the large inter-individual variability in its maintenance dose, warfarin is not easy to dose (2–4). To determine the effective therapeutic dose, frequent monitoring of the international normalized ratio (INR) followed by careful dose adjustments is required for each individual and it often takes 30–60 days to find out the maintenance dose (2).

A recent study has pointed out that warfarin was one of the 10 most common drugs that were responsible for emergency room visits by senior citizens with adverse reactions to medications (5). In addition, warfarin was reported to be under-prescribed for stroke prevention in patients with atrial fibrillation owing to doctors’ fear for causing hemorrhage in patients’ brain and digestive tracts (6). The evidence implied the importance and necessity of having a reliable warfarin dosing algorithm that helps to minimize the risk of adverse events and to provide more appropriate medical management.

It is known that inter-individual variability in the therapeutic warfarin dose could be attributable to demographical and clinical variables such as age, body size [represented by weight and body surface area (BSA)] (7), ethnicity (8–10), administration of concomitant drugs, dietary intake of...
vitamin K (11,12) and target INR (13). In addition, genetic variations in the CYP2C9 (cytochrome P450, family 2, subfamily C, polypeptide 9) and VKORC1 (vitamin K epoxide reductase complex subunit 1) genes could significantly influence warfarin responsiveness across various ethnic groups, because products of these genes play essential roles in the pharmacokinetics and pharmacodynamics of warfarin (9,14–18). However, a combined effect of currently known variables has been implicated to explain ~30–60% of total warfarin dose variation (9,13–15,19–23), suggesting that other unknown factors that determine inter-individual variability in warfarin dose may exist.

Most of the attempts to identify additional genetic factors associated with warfarin responsiveness, as reviewed by Wadelius and Pirmahomed (24) focused on genes involved in the pharmacokinetic and pharmacodynamic pathways of warfarin. Although several studies suggested minor effects of several genes on warfarin responsiveness (25,26), the results revealed significant inconsistencies across populations and often failed to be replicated (27–29). Furthermore, inclusion of these additional variants into the prediction model did not significantly improve the previously suggested model that considered only genetic influences of CYP2C9 and VKORC1 genes (30,31).

On the other hand, two genome-wide association studies (GWAS) aiming to detect additional genetic variants associated with warfarin responsiveness in the Caucasians have been conducted by far (32,33). Although neither of the studies found any common variants with strong effects on warfarin metabolism (GWAS) test, 25,043 SNPs that had a call rate of 99% and 106,878 SNPs that had minor allele frequency (MAF) of <0.01 were also excluded from the subsequent analysis. In total, we analyzed 485,227 SNPs for 1,508 subjects consisting of 807 Japanese patients with a low therapeutic warfarin dose and 701 Japanese patients with a high therapeutic warfarin dose.

Results of GWAS by univariate logistic regression analysis were shown in Figure 1A. A total of 28 SNPs revealed genome-wide significant P-values (<1.03 × 10⁻⁸), and all of these SNPs are located in or near the VKORC1 gene; the SNP showing the strongest association was rs9923231 (P = 8.65 × 10⁻³⁵). This SNP is strongly linked to the SNP rs10871454 (r² = 0.992, D' = 0.997) that, in the GWAS of Caucasians, achieved the strongest association (P = 6.2 × 10⁻¹³) (32). SNPs located in or near the CYP2C9 gene came next in the list: an SNP rs10509680 showed a P-value of 3.84 × 10⁻⁷. In total, 19 SNPs located in or near to the CYP2C9 gene showed P-values of <1 × 10⁻⁵. Eighteen additional SNPs in other chromosomal regions showed P-values of <1 × 10⁻⁵. These 18 SNPs were further genotyped in the replication samples and meta-analysis was performed. Results as in Supplementary Material, Table S1 showed that none of the 18 SNPs revealed significant association after the Bonferroni correction.

Figure 1B illustrates the results of GWAS by using multivariate logistic regression analysis after adjustment of the effects of age, BSA and genotypes of SNP rs10509680 (CYP2C9) and SNP rs9923231 (VKORC1) on warfarin dose. Although no SNP achieved a genome-wide significant association, 18 SNPs showed P-values of <1 × 10⁻⁵. These SNPs were further examined in the replication study and a meta-analysis was performed. Results as illustrated in Table 2 indicated that SNP, rs2108622, not only reached a Bonferroni-corrected P-value of <0.05 in the replication study, but also revealed a combined P-value of 2.57 × 10⁻⁸ that reached the genome-wide significance level.

In the subsequent analyses that evaluate the association of SNPs in each of the CYP2C9, VKORC1 and CYP4F2 genes with the therapeutic warfarin dose of subjects in the replication study, box and whisker plots (Fig. 2A) indicated that the GG (CYP2C9*1/*1) genotype of the CYP2C9’s SNP (rs10509680) as well as the TC and CC genotypes of the VKORC1’s SNP (rs9923231) both are strikingly associated with a higher therapeutic warfarin dose (P = 0.001626 and P < 2.2 × 10⁻¹⁶, respectively, by Mann–Whitney U-test). On the other hand, for the CYP4F2 SNP, rs2108622, we did not observe the distinctive trend of association. However, when we classified subjects based on their genotypes of SNPs in the CYP2C9 and the VKORC1 genes by using a scoring system [warfarin responsive index (WFRI); see Materials and Methods] previously described by Mushiroda et al. (17) (Fig. 2B) and then further stratify subjects based on their genotypes of rs2108622 (Fig. 2C), we observed that subjects carrying the CT/TT genotypes for rs2108622 in a group of subjects with a WFRI of two required a significantly higher dose of warfarin (P = 0.037).

To evaluate if other covariates such as sex, height, weight, INR and co-administration of amiodarone could also significantly influence warfarin dose of the Japanese subjects, we performed univariate linear regression analyses. Table 3 illustrates the results for 10 covariates (including age, BSA and SNPs in each of the CYP2C9, VKORC1 and CYP4F2 genes). We found that status of amiodarone co-administration could also significantly influence therapeutic warfarin dose of
Concomitant drugs
Genes, SNPs (Minor allele frequencies)

VKORC1
warfarin dose variation. SNP rs9923231 (four covariates [age, BSA and genotypes of rs10509680 replication samples while the model that considered only this model explains 43.4% of warfarin dose variation in the with warfarin dose in the univariate linear regression analysis. [(P = 8.39 × 10^{-4}). Although sex, height and weight also were considered to have significant influences on warfarin dose, effects of these covariates could be accounted by incorporating BSA into the multivariate regression analysis (data not shown).

Table 4 shows the results of multivariate linear regression analysis considering six covariates that include age, BSA, status of amiodarone co-administration and genotypes of rs10509680 (CYP2C9), rs9923231 (VKORC1) and rs2108622 (CYP4F2), each of which showed significant association with warfarin dose in the univariate linear regression analysis. This model explains 43.4% of warfarin dose variation in the replication samples while the model that considered only four covariates [age, BSA and genotypes of rs10509680 (CYP2C9) and rs9923231 (VKORC1)] explains 40.3% of warfarin dose variation. SNP rs9923231 (VKORC1) showed the greatest influence on warfarin dose variation (25.7%). About 6.7 and 6.4% of the dose variation could be attributed to age and BSA, respectively. However, SNP rs10509680 (CYP2C9) and SNP rs2108622 (CYP4F2) explained only 1–2% of the residual warfarin dose variation. The status of co-administration of amiodarone could explain 2.1% of the residual warfarin dose variation. Although the influence of increase in age (in year) and co-administration of amiodarone decreases the warfarin dose, increase in the number of alleles associated with high warfarin dose in the VKORC1, CYP2C9 and CYP4F2 genes (alleles C, G and T, respectively) increases warfarin dose for 1.39, 0.70 and 0.23 mg/day, respectively. A warfarin dosing algorithm developed for the Japanese population based on the model considering these six covariates was shown in Figure 3A and B, with status of amiodarone co-administration, and genotypes of each of the three SNPs were coded as indicated in Figure 3C.

Finally, considering that genotypes of rs2108622 as well as the status of amiodarone co-administration could also significantly influence warfarin dose, we performed another genome-wide association analysis to examine if other SNPs that show significant influence on warfarin dose could be identified if effects from all of the six covariates aforementioned have been accounted for. Although we found no SNP that showed the association with genome-wide significance, 15 SNPs showed a minimum P-value of <1 × 10^{-5}. We performed replication study and meta-analysis for these SNPs, but none of them showed significant association after the Bonferroni correction (Supplementary Material, Table S2).

**DISCUSSION**

In this GWAS that compared 807 Japanese patients with a low therapeutic warfarin dose to 701 Japanese patients with a high therapeutic warfarin dose, we report here for the first time that SNP rs2108622 (CYP4F2) is a genetic determinant of warfarin responsiveness for the Japanese population. We selected subjects with extreme phenotypes as cases and controls for the GWAS to increase the chances to detect variants associated with warfarin responsiveness as variants associated with high/low dose of warfarin were usually enriched in these subjects. To avoid possible effects from sampling bias, the finding from GWAS was evaluated by using a subsequent replication study that investigated 444 Japanese subjects, and all the subsequent analysis on contribution of SNP rs2108622 to the warfarin dose variation and potential clinical application of the...
SNP in prediction of the warfarin dose were based on the replication samples.

In the initial genome-wide analysis that used univariate logistic regression analysis, only SNPs in or near the VKORC1 gene, which was later found to contribute 25.7% to the warfarin dose variation in the Japanese population, showed a genome-wide significant level of association with therapeutic warfarin dose. Neither SNPs in the CYP2C9 gene nor the SNP in the CYP4F2 gene revealed genome-wide significant association. This is probably due to their relatively smaller contributions (~1–2%) to warfarin dose variation in the Japanese population. Similar to previous findings in Caucasians (33), we did not identify other common SNPs that have an influence as strong as that of SNPs in the VKORC1, CYP2C9 and CYP4F2 genes on the warfarin dose.

SNP, rs2108622 (C>T), which is located in exon 11 of the CYP4F2 gene, causes amino-acid substitution from valine to methionine. A recent study (34) found that CYP4F2 encodes a vitamin K<sub>1</sub> (VK1) oxidase and carriers of the CYP4F2 V433M allele (allele T carriers for rs2108622) may require a higher dose of warfarin due to the reduced capacity of CYP4F2 for metabolizing VK1. Since the association of rs2108622 with warfarin dose was first reported in 2008 (35), studies on Caucasians who were administered warfarin (33,36,37) and also the rs9923231 (VKORC1) that are known to influence warfarin dose of Japanese.

Our results indicated that incorporating genotypes of rs2108622 and addition of subjects’ status of amiodarone co-administration into the warfarin dosing algorithm, that considered only age, BSA and genotypes information of the rs10509680 (CYP2C9) and rs9923231 (VKORC1), improved the predictability of the warfarin requirement dose. Contribution from the CYP4F2 was estimated to be 1.4% for the Japanese and is similar to that reported for the Caucasians as 1–4% (33,35,36). To our knowledge, our study is the first that reports that the inclusion of the status of amiodarone co-administration could improve the predictability of the warfarin dosing algorithm for the Japanese. In our study, subjects who were co-administered with amiodarone required a significantly lower dose of warfarin ($P = 8.39 \times 10^{-4}$). This is in concordance with an enhanced anticoagulant effect resulting from inhibition of CYP2C9 by amiodarone (40). Although the total proportion of dose variation explained by our model is similar to that reported for the Han-Chinese population recently (21), it is much lower than the effect predicted by dosing algorithms developed for Caucasians that could explain up to 60% of the variance in the therapeutic warfarin dose (13–15).

A possible reason for this difference could be the weaker influence of the CYP2C9 gene due to lower allelic frequency (0.7% in our study) of the CYP2C9*3 variant and also the absence of the CYP2C9*2 variant in the Japanese population, compared with Caucasians and Africans (41,42). Although in
Table 2. Replication study and meta-analysis for the 18 SNPs with minimum $P$-values of multivariate logistic regression analysis [adjusted for the effects of 4 covariates known to influence warfarin dose of Japanese: age, BSA, and SNPs in the CYP2C9 (rs10509680) and the VKORC1 (rs9923231) genes, known to influence warfarin dose of Japanese] < 1 $\times$ 10^{-8}

<table>
<thead>
<tr>
<th>CHR</th>
<th>Chromosomal location</th>
<th>SNP</th>
<th>Allele 1</th>
<th>Allele 2</th>
<th>Study</th>
<th>$n$</th>
<th>$P^1$</th>
<th>Model</th>
<th>Bonferroni $P^2$ ($&lt; 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>15851431</td>
<td>rs2108622</td>
<td>T</td>
<td>C</td>
<td>GWAS</td>
<td>1508</td>
<td>8.67 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>5.47 $\times$ 10^{-3}</td>
</tr>
<tr>
<td>13</td>
<td>55670697</td>
<td>rs1512651</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>5.51 $\times$ 10^{-6}</td>
<td>REC</td>
<td>&gt;1</td>
</tr>
<tr>
<td>5</td>
<td>96451215</td>
<td>rs316222</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>3.56 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>20</td>
<td>4494579</td>
<td>rs2208303</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>6.11 $\times$ 10^{-7}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>20a</td>
<td>4495633</td>
<td>rs6084794</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>7.48 $\times$ 10^{-7}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>20a</td>
<td>4492779</td>
<td>rs911165</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>1.18 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>3</td>
<td>14459117</td>
<td>rs7628386</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>4.95 $\times$ 10^{-6}</td>
<td>DOM</td>
<td>&gt;1</td>
</tr>
<tr>
<td>16</td>
<td>13485305</td>
<td>rs12708772</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>9.25 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>20a</td>
<td>4456090</td>
<td>rs6116402</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>1.18 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>3</td>
<td>86283674</td>
<td>rs6549079</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>7.84 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>13a</td>
<td>55598251</td>
<td>rs281659</td>
<td>C</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>9.25 $\times$ 10^{-6}</td>
<td>REC</td>
<td>&gt;1</td>
</tr>
<tr>
<td>3</td>
<td>86300649</td>
<td>rs4688887</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>7.39 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>20a</td>
<td>4461311</td>
<td>rs6107476</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>2.84 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>15</td>
<td>49773127</td>
<td>rs2607117</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>2.92 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
<tr>
<td>16</td>
<td>79938647</td>
<td>rs6564871</td>
<td>A</td>
<td>C</td>
<td>GWAS</td>
<td>1508</td>
<td>4.74 $\times$ 10^{-6}</td>
<td>DOM</td>
<td>&gt;1</td>
</tr>
<tr>
<td>1</td>
<td>149787355</td>
<td>rs6587597</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>4.85 $\times$ 10^{-6}</td>
<td>REC</td>
<td>&gt;1</td>
</tr>
<tr>
<td>16</td>
<td>50558008</td>
<td>rs10163230</td>
<td>G</td>
<td>A</td>
<td>GWAS</td>
<td>1508</td>
<td>6.57 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>5.34 $\times$ 10^{-2}</td>
</tr>
<tr>
<td>20a</td>
<td>4483572</td>
<td>rs12626073b</td>
<td>A</td>
<td>G</td>
<td>GWAS</td>
<td>1508</td>
<td>1.18 $\times$ 10^{-6}</td>
<td>ADD</td>
<td>&gt;1</td>
</tr>
</tbody>
</table>

CHR, chromosome; SNP, single nucleotide polymorphism; $n$, sample size; ADD, additive model; DOM, Dominant model; REC, Recessive model; NA, not available; GWAS, genome-wide association study; Rep, replication study; Meta, combined study or meta-analysis. $^1P$-value for the GWAS refers to the minimum $P$-value among the additive, dominant and recessive models of logistic regression analysis; $P$-value for the replication study refers to the $P$-value of the linear regression analysis considering the same model; $P$-value for the combined study refers to the $P$-value of inverse variance meta-analysis. $^2P$-value of SNPs after the Bonferroni correction that considers number of independent tests. Only the Bonferroni-corrected $P$-value of < 1 are shown. $^3$SNPs that are highly linked with at least another SNP in the same table, and were therefore not considered in the Bonferroni correction. $^4$Failed to be genotyped in the replication study, but is completely linked with another genotyped SNP, rs911165.
most of the studies of Caucasians, variants in \textit{CYP2C9} could explain up to 10% of the variance in warfarin dose, in our study and another study of the Japanese (43), \textit{CYP2C9*3} explained only 1–5% of the warfarin dose variation.

In the present study, the contribution from rs9923231 (\textit{VKORC1}) (25.7%) was similar to 28% (13) and 28.5% (14) reported previously for the Caucasian populations, but much higher than the African population reported as 7.3%, probably due to the difference in the allelic frequency (9,10,16,22). In our study, the C-allele of rs9923231 is significantly enriched in the high-dose group (allele frequency = 0.219) compared with the low-dose group (allele frequency = 0.07; Table 1).

In conclusion, although the strongest genetic determinant of warfarin responsiveness for the Japanese population remains the \textit{VKORC1} gene, followed by the \textit{CYP2C9} gene, \textit{CYP4F2} has been identified for the first time to be a genetic determinant for warfarin responsiveness. Inclusion of its genotypes into the dosing algorithm improved warfarin dose predictability. Although our results implied that common SNPs other than those in the \textit{CYP2C9}, \textit{VKORC1} and \textit{CYP4F2} genes that show strong effect on therapeutic warfarin dose might not exist for the Japanese, the existence of other rare variants that influence on therapeutic warfarin dose should not be excluded. Considering the retrospective nature of our current study, the dose prediction algorithm derived from our data should be considered preliminary. Further prospective studies that involve a larger number of subjects with more well-monitored clinical information, especially on dietary intake of vitamin K and concomitant drugs, should be helpful not only in validating and improving the predictability of the current dosing algorithm, but also in allowing weight to be given to each variant for development of the more accurate WFRI.

**MATERIALS AND METHODS**

**Subjects**

The GWAS involved 1515 subjects requiring extreme dose of warfarin who were registered into ‘the Biobank Japan Project’ in the Ministry of Education, Culture, Sports, Science and Technology, Japan (44). Among them, 811 subjects required a daily warfarin dose of 1 mg or less, whereas the remaining 704 subjects required a daily warfarin dose of 4 mg or more. These 1515 subjects were selected from the 4309 subjects who received warfarin therapy that were registered into the BioBank Japan Project from June 2003 to December 2007. We selected subjects whose warfarin maintenance dose was two times lower or higher than the average daily warfarin dose required by the general Japanese population (2–2.5 mg) for the GWAS, after considering the statistical power of the study. For the replication study, a total of 444 subjects administrated with warfarin were independently collected in Nippon Medical School (Kawasaki, Japan) from July 2007 to July 2008. None of these patients have registered record of high dietary intake of vitamin K [such as natto (fermented soybean) or chlorella (a kind of algae) that had been known to contain high level of vitamin K]. All patients were
monitored from the first date of warfarin use until the end of follow-up at hospitals and their clinical information was collected from the patients’ medical records. The latest dates for the data collection were December 2008 and July 2008 for the BioBank Japan Project and Nippon Medical School, respectively. All subjects who participated in the study had clinical information for age, sex, BSA (calculated based on the DuBois formula) and status of amiodarone co-administration. Subjects had a target INR range of 1.5–3.0 and warfarin was prescribed either for prevention or treatment of thromboembolic diseases. The daily maintenance dosage was defined as the amount of warfarin (mg) required for achievement of a stable INR in the target range in two consecutive blood samplings. Demographical information of subjects was summarized in Table 1. All subjects had given written informed consent to participate in the study in accordance with the process approved by Ethical Committee at each of The Institute of Medical Science of the University of Tokyo, The Nippon Medical School and the Center for Genomic Medicine of RIKEN.

**Genome-wide association study**

A case–control GWAS was conducted by comparing 811 subjects requiring a daily warfarin dose of 1 mg or less, with the 704 subjects who required a daily warfarin dose of 4 mg or more. In the GWAS, the 1515 subjects were genotyped by using the Illumina HumanHap610 Genotyping BeadChip (San Diego, CA, USA) and quality control of genotyping data was performed. Subjects with a call rate of <95%, SNPs with a call rate of <99% or MAF of <0.01 as well as SNPs with a HWE test’s $P$-value of $<1 \times 10^{-6}$ were excluded from the subsequent statistical analysis. The genotyping data were analyzed by using univariate and multivariate logistic regression analyses. In the univariate analysis, the association of each SNP with warfarin dose was examined without being adjusted by any other factors; whereas in the multivariate logistic regression analysis, effects of four covariates known to significantly influence therapeutic warfarin dose (age, BSA, genotypes of SNP rs10509680 in the CYP2C9 and SNP rs9923231 in the VKORC1 genes) were used for adjustment of association. Each of the additive, dominant and recessive models of inheritance was used to evaluate the association between each SNP and warfarin dose. Minimum $P$-value was used to evaluate the significance of association. SNPs with minimum $P$-value of $<10^{-5}$ in the GWAS were selected for the replication study. The strategy of the study was summarized in Supplementary Material, Figure S1.

**Replication study and meta-analysis**

In the replication study, the 18 SNPs with minimum $P$-value of $<10^{-5}$ by univariate logistic regression analysis in the first GWAS were further genotyped in 444 subjects by a multiplex PCR-Invader assay method and analyzed by univariate linear regression analysis in the replication study. The 18 SNPs with a minimum $P$-value of $<1 \times 10^{-5}$ by multivariate logistic regression analysis in the GWAS were genotyped similarly and then analyzed by multivariate linear regression analysis. The Bonferroni correction considering the number of independent tests (after considering linkage disequilibrium between tested SNPs) was applied for the judgment of statistical significance of association between each SNP and warfarin dose. Inverse variance meta-analysis, which calculated an $z$-statistic summarizing the magnitude and the direction of effect relative to a reference allele in both GWAS and replication study, was used to calculate the combined $P$-value of SNP considering both the GWAS and the replication study.

**Validating association between rs2108622 and warfarin dose**

Subsequent analyses to validate the effects of SNPs on therapeutic warfarin dose of the Japanese subjects were performed only for subjects of the replication study because subjects who participated in the GWAS were selected from the two extreme-dose groups. Box and whisker plots were used to graphically illustrate association between genotypes of each SNP and therapeutic warfarin dose of subjects. The association between the newly identified SNP, rs2108622, and warfarin dose after adjusting for the genetic influences from SNP rs10509680 in the CYP2C9 and SNP rs9923231 in the

---

**Table 3. Univariate linear regression analysis for association of each covariate with warfarin dose for samples of the replication study**

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Replication (n = 440)</th>
<th>Beta coefficient</th>
<th>$P$-value</th>
<th>Adjusted $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$-0.032$</td>
<td>$4.61 \times 10^{-11}$</td>
<td>0.092</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.257</td>
<td>$3.05 \times 10^{-2}$</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Height</td>
<td>0.030</td>
<td>$1.58 \times 10^{-7}$</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td>0.026</td>
<td>$7.92 \times 10^{-11}$</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>1.759</td>
<td>$2.51 \times 10^{-11}$</td>
<td>0.095</td>
<td></td>
</tr>
<tr>
<td>INR</td>
<td>0.180</td>
<td>$1.97 \times 10^{-1}$</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Amiodarone</td>
<td>$-0.894$</td>
<td>$8.39 \times 10^{-4}$</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>CYP2C9</td>
<td>0.732</td>
<td>$3.54 \times 10^{-3}$</td>
<td>0.017</td>
<td></td>
</tr>
<tr>
<td>(rs10509680)</td>
<td>1.416</td>
<td>$&lt;2.00 \times 10^{-16}$</td>
<td>0.262</td>
<td></td>
</tr>
<tr>
<td>VKORC1</td>
<td>(rs9923231)</td>
<td>0.213</td>
<td>$1.71 \times 10^{-2}$</td>
<td>0.011</td>
</tr>
<tr>
<td>CYP4F2</td>
<td>(rs2108622)</td>
<td>$2.00 \times 10^{-1}$</td>
<td>0.059</td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 4. Multiple linear regression analysis considering the six covariates that are significantly influencing warfarin dose of samples of the replication study**

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Beta coefficient</th>
<th>$P$-value</th>
<th>Partial $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$-0.024$</td>
<td>$3.62 \times 10^{-5}$</td>
<td>0.067</td>
</tr>
<tr>
<td>BSA</td>
<td>1.118</td>
<td>$3.13 \times 10^{-6}$</td>
<td>0.064</td>
</tr>
<tr>
<td>CYP2C9</td>
<td>(rs10509680)</td>
<td>0.698</td>
<td>$2.61 \times 10^{-4}$</td>
</tr>
<tr>
<td>VKORC1</td>
<td>(rs9923231)</td>
<td>1.386</td>
<td>$&lt;2.00 \times 10^{-16}$</td>
</tr>
<tr>
<td>CYP4F2</td>
<td>(rs2108622)</td>
<td>0.227</td>
<td>$8.67 \times 10^{-4}$</td>
</tr>
<tr>
<td>Amiodarone</td>
<td>$-0.756$</td>
<td>$2.74 \times 10^{-4}$</td>
<td>0.021</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td></td>
<td></td>
<td>0.434</td>
</tr>
</tbody>
</table>

---

* $n$, sample size; BSA, body surface area; INR, international normalized ratio.
* Four among the 444 subjects who participated in the replication study failed to be genotyped at rs2108622 (CYP4F2), and were therefore not included in the subsequent analysis.
VKORC1 genes, was also examined; subjects were stratified based on their genotypes of rs2108622, only after they were pre-classified according to their genotypes of rs10509680 (CYP2C9) and rs9923231 (VKORC1) using a scoring system (WFRI) previously described by Mushiroda et al. (17). In the case of rs10509680 (CYP2C9), a score of 1 was assigned to individuals with GG genotype (CYP2C9∗1/*1) and 0 to individuals with TG genotype (CYP2C9∗1/*3). For rs9923231 (VKORC1), a score of 0 was given to individuals with the TT genotype, which is associated with warfarin sensitivity, whereas score 1 and score 2 were assigned to individuals with TC and CC genotypes, respectively, because of the additive effect of the C allele of the SNP. The scores of rs10509680 (CYP2C9) and rs9923231 (VKORC1) were then summed for each subject to represent the WFRI of the subject. The contribution of SNP rs2108622 to warfarin dose variation of Japanese subjects as a single covariate as well as in combination with other genetic and clinical variables was evaluated by univariate and multivariate linear regression analyses, respectively. Ten variables including age, BSA, sex, INR, height, weight, status of amiodarone co-administration, genotypes of SNPs in the CYP2C9 and VKORC1 genes and genotypes of rs2108622 were tested independently and simultaneously against the therapeutic warfarin dose of subjects. Covariates showing significant association with warfarin dose were incorporated into the subsequent multiple regression model. An algorithm that predicts the appropriate therapeutic warfarin dose for the Japanese population was derived on the basis of this model. The effect of each covariate on the difference in warfarin dose for subjects and the contribution of each covariate to explained warfarin dose variation were determined.
Statistical analysis
Quality control of genome-wide genotyping data were performed by using the PLINK statistical software (v1.06) (46,47). Manhattan plots were generated by using the WGA-Viewer software (48). Logistic and linear regression analysis were performed either by using the PLINK statistical software (v1.06) or the R statistical environment version 2.7.10 (49). Inverse variance meta-analysis was performed by using the METAL program (50). Box and whisker plots were generated by using the R statistical environment version 2.7.10. The Mann–Whitney U-test was performed by using the R statistical environment version 2.7.10. Scatter plot that illustrates the effect of equaling upon warfarin dose requirements was generated by using the Excel program.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We express our gratitude to patients who kindly participated in the current study. We also thank members of the genotyping team at the Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, for their excellent technical assistance.

Conflict of Interest statement. None declared.

FUNDING
This work was supported by Leading Project for Personalized Medicine in Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES


49. R statistical environment version 2.7.10. http://cran.r-project.org/