Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Pharmacogenomics

Influence of common and rare genetic variation on warfarin dose among African–Americans and European–Americans using the exome array

Aim: We conducted a genome-wide association study using the Illumina Exome Array to identify coding SNPs that may explain additional warfarin dose variability. **Patients & methods:** Analysis was performed after adjustment for clinical variables and genetic factors known to influence warfarin dose among 1680 warfarin users (838 European–Americans and 842 African–Americans). Replication was performed in an independent sample. **Results:** We confirmed the influence of known genetic variants on warfarin dose variability. Our study is the first to show the association between rs12772169 and warfarin dose in African–Americans. In addition, genes *COX15* and *FGF5* showed significant association in European–Americans. **Conclusion:** We identified some novel genes/SNPs that underpin warfarin dose response. Further replication is needed to confirm our findings.

First draft submitted: 7 March 2017; Accepted for publication: 19 May 2017; Published online: 7 July 2017

Keywords: *CYP2C9* • *CYP4F2* • exome array • genome-wide association study • prediction models • *VKORC1* • warfarin • warfarin dose

Nianjun Liu[1], Marguerite R Irvin[2], Degui Zhi[3], Amit Patki[4], T Mark Beasley[4], Deborah A Nickerson[5], Charles E Hill[6], Jinbo Chen[7], Stephen E Kimmel[7,8] & Nita A Limdi*[,9]

[1]Department of Epidemiology & Biostatistics, School of Public Health – Bloomington, Indiana University, Bloomington, IN 47405, USA

[2]Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[3]School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA

[4]Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[5]Department of Genome Sciences, School of Medicine, University of Washington, Seattle, WA 98195, USA

[6]Department of Pathology & Laboratory Medicine, Emory University School of Medicine, Atlanta, GA 30322, USA

[7]Department of Biostatistics & Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[8]Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[9]Department of Neurology, School of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA

*Author for correspondence: nlimdi@uabmc.edu

Warfarin remains the mainstay of oral anticoagulant therapy [1–3], commanding more than 70% of the total market share despite the introduction of newer agents [3–6]. However, warfarin therapy is challenging due to the great interpatient variability in dose requirements. Further, its narrow therapeutic index contributes to a high rate of adverse effects [7–13], earning warfarin a consistent ranking among the top ten drugs associated with serious adverse events [14,15]. Achieving therapeutic anticoagulation rapidly and predictably is critical for safe and effective therapy.

To this end, multiple investigations have identified the influence of clinical (e.g., comorbidity, concurrent medications), demographic (e.g., age, weight) and genetic factors on warfarin dose response [16–27]. To date, candidate gene and genome-wide associations studies (GWAS) have confirmed that the majority of the genetic influence is accounted for by SNPs in two genes: *CYP2C9*, which codes for the enzyme CYP450 2C9 that metabolizes *S*-warfarin [28,29], and *VKORC1*, which codes for warfarin's target, vitamin K epoxide reductase [30,31]. In addition, *CYP4F2* accounts for a small but significant proportion of the variability in warfarin dose in European–Americans (EAs) [16,32] but not in African–Americans (AAs) [17,23,27].

In both EAs and AAs, clinical factors account for approximately 15–20% of the variability in warfarin dose. SNPs in *CYP2C9*, *CYP4F2* and *VKORC1* account for a greater proportion of dose variability among EAs (35–40%) than in AAs (7–10%) [23]. To dentify the influence of common SNPs on warfarin dose in AAs, the

Future Medicine part of fsg

International Warfarin Pharmacogenetics Consortium has published a GWAS analysis of warfarin dose representing data on approximately 1000 AA adults (aged ≥18 years) who were taking a stable maintenance dose of warfarin [25]. After conditioning their models on *VKORC1* –1639G>A, *CYP2C9**2 and *CYP2C9**3, they identified rs12777823 within the *CYP2C* gene cluster on chromosome 10 as being significantly associated with warfarin dose.

To date all genomic studies of warfarin dose have considered common variants from GWAS or candidate gene variants. GWAS variants are selected to capture a snapshot of common variation across the genome largely focusing on tagging SNPs [33]. Overall, potentially functional protein coding variations have not been represented in previous publications. We hypothesized that further examination of coding SNPs (both common and rare) may harbor additional clues that explain the variability in warfarin dose. With this aim, we assessed 247,870 SNPs using the Illumina Exome Array (Illumina, CA, USA) [34] in a prospective cohort of 1680 warfarin users (838 EAs and 842 AAs).

Patients & methods

Our study population includes participants (≥20 years old) initiating warfarin treatment with a target international normalized ratio (INR) of 2 to 3 were enrolled at the beginning of treatment in an inception cohort. This study was approved by the institutional review boards of the University of Alabama at Birmingham, Emory University and the University of Pennsylvania. A total of 1680 patients were included in this study. Among them, 1240 patients (701 EAs and 539 AAs) from the University of Alabama at Birmingham and Emory were included in the discovery cohort. The validation cohort included 440 patients (137 EAs and 303 AAs) from the INR Adherence and Genetics 2 study [35], a multicenter, prospective cohort of patients initiating warfarin therapy between October 2009 and August 2013 at three urban anticoagulation clinics: the Hospital of the University of Pennsylvania, the Corporal Michael J Crescenz Veterans Affairs Medical Center and the Johns Hopkins Medical Institutions under the approval of their respective IRBs. Exclusion criteria were age <21 years, inability to give consent or abnormal INR before initiating therapy.

Patient demographics, including self-identified race, indication for therapy, comorbidities and medications were collected as previously reported [23]. We examined 247,870 SNPs using the Illumina Exome Array [34] among participants in the discovery and validation cohorts. Analysis was performed using Illumina GenomeStudio software (v2011.1). A project was created using the beadchip data, sample statistics were calculated and samples were clustered using Illumina GenomeStudio. A subset of SNPs was manually checked for clustering performance. After clustering, each sample was assigned a call rate (indicating the percentage of SNPs that the sample was able to be clustered for) to determine how well it clustered with the other samples. Samples were then sorted by call rate, and those with a call rate of <98% were considered a failed sample and excluded from the project. Warfarin dose was defined as the average maintenance dose after the attainment of three consecutive INRs in target range measured at least 2 weeks apart. Log-transformation on warfarin dose was used to attain normality [36,37].

Statistical analysis

All SNPs were tested for Hardy–Weinberg equilibrium (HWE) at a threshold of $p > 10^{-5}$. Singletons (i.e., SNPs with only one copy of alleles), SNPs violating the HWE assumption and SNPs with a missing rate larger than 5% were excluded from the analyses. There were 85,723 SNPs left for AAs, with a Bonferroni threshold of 5.83×10^{-7}. To ensure that important signals approaching (but not reaching) genome-wide significance are not missed, in addition to Bonferroni thresholds, GWAS commonly have thresholds for 'suggestive significance' [25,38–40]. A threshold of 1.0×10^{-5} was used as 'suggestive significance' for AAs. For EAs, a total of 62,461 SNPs were included in the analysis, with a Bonferroni threshold of 8.01×10^{-7} (a threshold of 1.0×10^{-5} was used as 'suggestive significance'). Both Bonferroni and false discovery rate (FDR) corrections were applied for multiple testing. The same quality control criteria were used for the validation cohort.

In our single-marker analysis, we assessed the effect of SNPs on dose using the additive model in PLINK (v. 1.07) [37] after adjusting for the first two principal components (i.e., PC1 and PC2) [41], clinical factors (e.g., age, BMI, gender, kidney function, amiodarone therapy) and SNPs known to be associated with warfarin dose. The latter included *CYP2C9* (*2 [rs1799853], *3 [rs1057910]), *CYP24F2* (rs2108622) and *VKORC1* (rs9923231). Principal components were included in the model to control for population structure/admixture within each self-reported race group and in the pooled analysis across race groups [41,42], calculated using software EIGENSTRAT (v.3.0) [42,43] on the 29,543 tagged SNPs for EAs and 32,695 tagged SNPs for AAs with linkage disequilibrium (LD; using r^2) <0.05 among those SNPs. The first ten PCs were calculated and were included in the initial analysis. As the first two adequately controlled population structure/admixture, the final models adjusted for these two PCs. A pooled

Table 1. Summary of SNPs influencing warfarin dose in single marker analysis adjusted for clinical variables and top two principal components.

SNP	CHR	Gene	BP	Minor allele	MAF	BETA	p-value	BONF	FDR	R^2
European–Americans										
exm-rs10871454	16	*STX4*	31,048,079	A	37.44%	-0.3275	6.68E-46	4.17E-41	4.17E-41	0.2189
rs9923231†	16	*VKORC1*	31,107,689	T	36.41%	-0.3281	9.28E-45	5.80E-40	2.90E-40	0.2164
exm1235282	16	*ZNF646*	31,088,625	G	35.01%	-0.336	2.68E-44	1.68E-39	5.58E-40	0.2188
exm1235743	16	*MYST1*	31,142,271	G	36.58%	0.2223	7.45E-20	4.66E-15	1.16E-15	0.0997
exm-rs4086116	10	*CYP2C9*	96,707,202	A	20.57%	-0.2635	1.24E-19	7.76E-15	1.55E-15	0.0947
exm844046†	10	*CYP2C9*	96,741,053	A	6.67%	-0.3571	1.01E-13	6.31E-09	9.01E-10	0.0637
exm-rs10509680	10	*CYP2C9*	96,734,339	C	6.67%	-0.3571	1.01E-13	6.31E-09	9.01E-10	0.0645
exm1235538	16	*PRSS53*	31,096,164	G	38.09%	0.1774	3.59E-13	2.24E-08	2.80E-09	0.0613
exm1235457	16	*ZNF646*	31,092,075	A	36.57%	0.1772	6.85E-13	4.28E-08	4.75E-09	0.0814
exm-rs7197475	16		30,642,867	A	36.42%	0.1716	1.76E-12	1.10E-07	1.10E-08	0.0611
exm1235050	16	*HSD3B7*	30,999,142	A	36.11%	0.1697	2.41E-12	1.51E-07	1.37E-08	0.0579
exm-rs7186852	16		30,635,659	G	36.03%	0.1706	3.19E-12	1.99E-07	1.66E-08	0.0592
exm-rs10782001	16	*FBXL19*	30,942,625	G	35.40%	0.1675	5.09E-12	3.18E-07	2.45E-08	0.0576
exm-rs11150610	16	*ITGAM*	31,334,236	A	39.25%	-0.137	1.36E-08	8.49E-04	6.07E-05	0.0371
exm1236199	16	*ITGAM*	31,289,396	G	15.78%	0.1729	3.32E-08	2.07E-03	1.38E-04	0.0378
exm1236395	16	*ITGAX*	31,374,535	C	33.99%	-0.1303	1.27E-07	7.93E-03	4.96E-04	0.0323
exm-rs2185570	10	*CYP2C9*	96,751,270	G	13.66%	-0.1789	1.79E-07	1.12E-02	6.58E-04	0.0314
exm-rs1799853†	10	*CYP2C9*	96,702,047	A	13.66%	-0.1775	2.34E-07	1.46E-02	8.12E-04	0.0315
exm844097	10	*CYP2C8*	96,798,749	G	12.72%	-0.1755	4.93E-07	3.08E-02	1.62E-03	0.0292
exm2272433	16		30,672,719	G	26.39%	0.1259	2.68E-06	1.67E-01	8.37E-03	0.0279
exm1236311	16	*ITGAX*	31,367,318	G	14.55%	0.1497	3.95E-06	2.47E-01	0.0117	0.0267
exm1233627	16	*PRR14*	30,666,367	A	26.06%	0.1209	5.93E-06	3.70E-01	0.0168	0.0244
exm848216	10	*ENTPD7*	101,421,367	A	0.21%	1.786	7.25E-06	4.53E-01	0.0197	0.0236
African–Americans										
rs9923231†	16	*VKORC1*	31,107,689	T	10.16%	-0.2252	1.80E-08	1.54E-03	9.61E-04	0.0506
exm-rs10871454	16	*STX4*	31,048,079	A	10.47%	-0.2243	2.24E-08	1.92E-03	9.61E-04	0.0471
exm1279100	17	*P2RX1*	3,807,286	A	0.19%	-1.255	4.56E-06	3.90E-01	0.1224	0.0355
exm1235612	16	*VKORC1*	31,104,720	A	0.28%	1.001	6.91E-06	5.92E-01	0.1224	0.0293
exm1205118	16	*TSC2*	2,115,529	A	0.19%	-1.226	8.36E-06	7.16E-01	0.1224	0.0286
exm582506	6	*TXLNB*	139,564,007	A	2.09%	-0.4073	1.01E-05	8.66E-01	0.1224	0.0283

†rs9923231 (*VKORC1* SNP –1639G>A); exm844046 is commonly referred to as *CYP2C9*3* (rs1057910), and exm-rs1799853 is commonly referred to as *CYP2C9*2 (rs1799853)*.
All SNPs met the Hardy–Weinberg equilibrium assumption; p-values (not corrected for multiple testing) are displayed; FDR is FDR adjusted q-value.
BETA: Beta coefficient from regression analysis; BONF: Bonferroni adjusted p-value; BP: Base-pair position; CHR: Chromosome; FDR: False discovery rate; MAF: Minor allele frequency; R^2: Semipartial R^2 from regression.

analysis was performed for single markers across race groups in the discovery cohort using the meta-analysis tool in PLINK (v. 1.07).

After exclusion of singletons and SNPs violating the HWE assumption, the gene-based analysis included 14,089 genes for EAs and 14,755 genes for AAs. For each participant, rare variations (minor allele frequency [MAF] <5%) within each gene were represented by a rare variant score (RVS) equal to the proportion of rare variants at which the individual possessed the minor allele [44]. For each gene, single common variants (MAF ≥5%) and the RVS were included

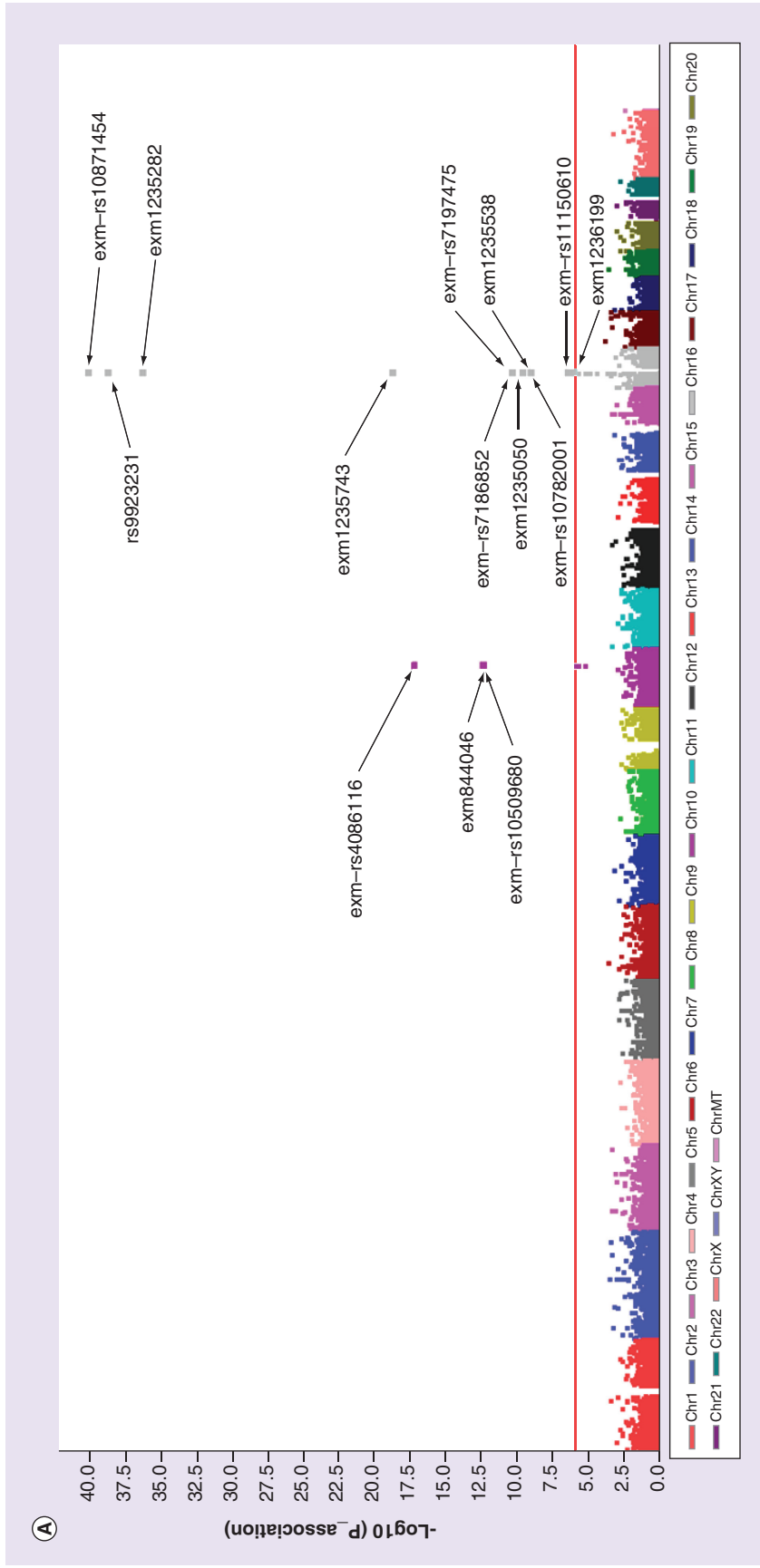

Figure 1. Manhattan plot for single marker association analysis with warfarin dose adjusted for clinical variables and top two principal components among European–Americans (A) and African–Americans (B; for B see overleaf).
For color figures please see http://www.futuremedicine.com/doi/full/10.2217/pgs-2017-0046

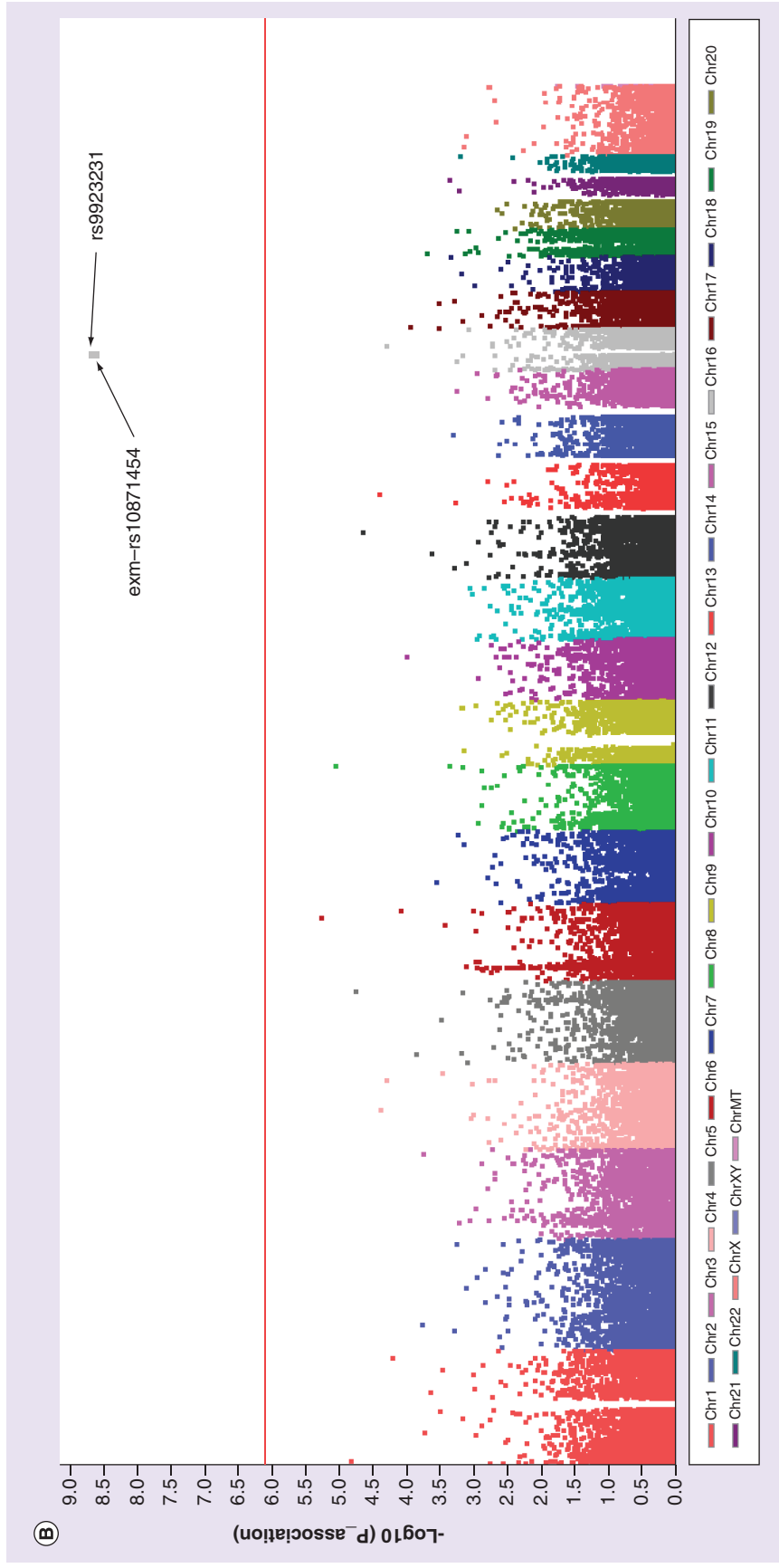

Figure 1. Manhattan plot for single marker association analysis with warfarin dose adjusted for clinical variables and top two principal components among European–Americans (A) and African–Americans (B) (cont.).
For color figures please see http://www.futuremedicine.com/doi/full/10.2217/pgs-2017-0046

Table 2. Single marker analysis after adjustment for clinical factors, *VKORC1* (rs9923231), *CYP2C9* (rs1057910 and rs1799853), *CYP4F2* (rs2108622) and top two principal components.

SNP	Gene	CHR	BP	Minor allele	MAF	BETA	p-value	BONF	FDR	R^2
European–Americans										
exm847442	*C10orf28*	10	99,995,247	A	0.143%	-1.174	1.20E-6	0.075	0.0110	0.0165
exm409265	*FGF5*	4	81,188,221	A	0.499%	0.584	6.96E-6	0.435	0.0111	0.0157
exm2262194	Intergenic *TSG1* (dist = 9331), *MANEA* (dist = 1529883)	6	94,495,530	A	1.141%	-0.381	1.13E-05	0.706	0.0176	0.0134
exm156678	*OBSCN*	1	228,559,994	G	49.070%	-0.078	1.6E-05	0.999	0.0244	0.0131
exm1619580	*BRD1*	22	50,169,762	A	0.143%	-1.039	1.8E-05	1	0.0267	0.0127
exm1020921	*CCT2*	12	69,983,450	G	0.143%	-1.027	2.17E-05	1	0.0311	0.0125
exm638477	*ZNF655*	7	99,170,533	G	1.284%	0.347	2.23E-05	1	0.0311	0.0124
exm508543	*COL23A1*	5	177,674,821	G	0.143%	1.032	2.24E-05	1	0.0311	0.0115
exm67802	*MYOM2*	1	70,501,827	C	0.143%	1.020	2.63E-05	1	0.0358	0.0121
exm1188445	*PLIN1*	15	90,214,777	G	0.143%	-1.020	2.86E-05	1	0.0379	0.0119
exm674510	*MLL3*	7	151,949,698	A	0.143%	-1.000	3.7E-05	1	0.0453	0.0117
exm251401	*ZC3H15*	2	187,360,053	G	0.143%	-0.980	5.23E-05	1	0.0453	0.0111
exm1197593	*WDR90*	16	703,648	G	0.143%	-1.382	5.3E-05	1	0.0453	0.0113
exm1274615	*SCARF1*	17	1,538,499	A	0.143%	-1.382	5.3E-05	1	0.0453	0.0113
exm425409	*MAML3*	4	140,640,527	A	0.499%	0.526	5.6E-05	1	0.046	0.0112
exm1163794	*DYX1C1*	15	55,731,682	T	0.143%	0.976	6.15E-05	1	0.0499	0.0109
exm1201118	*TPSD1*	16	1,308,333	A	1.000%	0.367	7.84E-05	1	0.0627	0.0109
exm1432416	*RAD23A*	19	13,059,147	G	3.067%	0.213	8.31E-05	1	0.0657	0.0106
exm413801	*MMRN1*	4	90,856,377	A	0.785%	0.407	9.22E-05	1	0.0714	0.0104
exm1013500	*TIMELESS*	12	56,815,233	A	0.357%	-0.601	9.26E-05	1	0.0714	0.0103
exm-rs7197475		10	101,421,367	A	0.214%	1.200	9.61E-05	1	0.0732	0.0000
exm357145	*ZIC4*	3	147,113,821	C	0.927%	0.349	9.75E-05	1	0.0734	0.0103
African–Americans										
exm-rs12777823	Intergenic *HELLS* (dist = 43646), *CYP2C18* (dist = 37749)	10	96,405,502	A	24.910%	-0.134	1.66E-07	0.0142	0.0142	0.0356
exm1279100	*P2RX1*	17	3,807,286	A	0.191%	-1.298	7.10E-07	0.0609	0.0186	0.0407
exm1028932	*UHRF1BP1L*	12	100,444,984	A	8.442%	-0.190	5.72E-06	0.4903	0.0186	0.0231
exm1034271	*SELPLG*	12	109,017,758	A	0.186%	-1.641	8.13E-06	0.6969	0.0186	0.0274
exm800598	*LCN12*	9	139,848,395	A	0.186%	-1.158	9.35E-06	0.8015	0.0195	0.0267
exm-rs12772169		10	96,405,329	A	41.360%	-0.102	9.64E-06	0.8264	0.0197	0.0226
exm177799	*ASXL2*	2	25,967,276	A	0.186%	-1.146	1.13E-05	0.9687	0.0220	0.0263
exm1235612	*VKORC1*	16	31,104,720	A	0.278%	0.938	1.23E-05	1	0.0235	0.0260
exm384884	*OTOP1*	4	4,198,893	A	0.278%	-0.922	1.49E-05	1	0.0278	0.0259

All SNPs satisfied the Hardy–Weinberg equilibrium assumption test.
BETA: β coefficient from regression analysis; BP: Base-pair position; BONF: Bonferroni adjusted p-value; CHR: Chromosome; FDR: False discovery rate; MAF: Minor allele frequency; R^2: Semipartial R^2 from regression.

Table 2. Single marker analysis after adjustment for clinical factors, *VKORC1* (rs9923231), *CYP2C9* (rs1057910 and rs1799853), *CYP4F2* (rs2108622) and top two principal components (cont.).

SNP	Gene	CHR	BP	Minor allele	MAF	BETA	p-value	BONF	FDR	R^2
African–Americans (cont.)										
exm1608547	*DMC1*	22	38,917,703	G	0.186%	-1.124	1.77E-05	1	0.0324	0.0249
exm1205118	*TSC2*	16	2,115,529	G	0.186%	-1.149	2.00E-05	1	0.0358	0.0230
exm748191	*PIGO*	9	35,092,607	A	0.186%	-1.114	2.19E-05	1	0.0383	0.0249
exm199426	*WDR35*	2	70,033,584	A	32.470%	-0.106	2.28E-05	1	0.0391	0.0219
exm249675	*CCDC141*	2	179,732,845	A	15.860%	0.133	2.41E-05	1	0.0406	0.0303
exm28590	*HSPG2*	1	22,154,900	A	0.278%	-0.893	2.71E-05	1	0.0447	0.0243
exm608896	*DNAH11*	7	21,934,511	A	0.464%	-0.692	3.08E-05	1	0.0499	0.0237
exm327045	*C3orf67*	3	58,853,578	G	0.186%	-1.080	3.57E-05	1	0.0557	0.0236
exm711934	*NIPAL2*	8	99,217,391	G	0.186%	-1.080	3.57E-05	1	0.0557	0.0236
exm296431	*ZCWPW2*	3	28,476,661	A	0.464%	-0.679	4.74E-05	1	0.0725	0.0222
exm826588	*ANK3*	10	61,802,477	A	0.464%	-0.741	6.26E-05	1	0.0821	0.0224
exm2261915		5	3,311,493	G	45.820%	0.092	6.74E-05	1	0.0821	0.0209
exm1454975	*GRAMD1A*	19	35,506,824	A	0.278%	-0.847	6.92E-05	1	0.0821	0.0218
exm1484329	*BBC3*	19	47,735,796	T	0.286%	-0.846	7.53E-05	1	0.0821	0.0280
exm619489	*TNS3*	7	47,408,443	G	0.742%	-0.563	7.62E-05	1	0.0821	0.0106
exm533622	*DOM3Z*	6	31,939,416	G	0.280%	0.854	8.38E-05	1	0.0821	0.0243
exm1134736	*PACS2*	14	105,833,623	A	0.186%	-1.454	8.76E-05	1	0.0821	0.0215
exm218846	*ST6GAL2*	2	107,460,345	A	0.289%	0.865	8.81E-05	1	0.0821	0.0220
exm-rs7495052	*SLCO3A1*	15	92,552,029	G	28.070%	-0.094	9.12E-05	1	0.0823	0.0180

All SNPs satisfied the Hardy–Weinberg equilibrium assumption test.
BETA: β coefficient from regression analysis; BP: Base-pair position; BONF: Bonferroni adjusted p-value; CHR: Chromosome; FDR: False discovery rate; MAF: Minor allele frequency; R^2: Semipartial R^2 from regression.

in the model together with the aforementioned clinical and known genetic factors using the R version 3.1.1 software [45]. In addition to the test of RVS (p-value denoted as Prare), joint test of common variants (p-value denoted as Pcommon) and joint test of common variants and the RVS were reported (p-value denoted as Pcombined). ANNOVAR software (version 2014Nov12) was used to annotate all SNPs [46,47]. After consideration of the number of genes tested, the Bonferroni thresholds for gene-based tests were 3.55×10^{-6} for EAs and 3.39×10^{-6} for AAs. The suggestive threshold was 5×10^{-4}. SNPs showing significant or suggestive association in the discovery cohort were all tested in the validation cohort. The same statistical models and genetic and clinical factors were used in the validation cohort as in the discovery cohort. For SNPs tested in the validation cohort, a meta-analysis was also performed in the discovery and validation cohorts using the meta-analysis tool in PLINK (v. 1.07). To calculate the unique contribution of individual SNPs and genes toward warfarin dose variation we computed semipartial R^2, adjusting for the clinical factors and known genetic variants.

Results

Table 1 presents the results from analysis adjusted for only clinical factors and top two principal components. The genomic inflation factor λ was 0.984 for EA and 0.981 for AA, indicating that population stratification was well controlled for in the analysis. *VKORC1* (rs9923231 and exm-rs10871454) was the single most important genetic influence on warfarin dose (Figure 1 & Table 1) in both EAs ($p = 9.28 \times 10^{-45}$ and p-value $= 6.68 \times 10^{-46}$) and AAs (p-value $= 1.80 \times 10^{-8}$ and $p = 2.24 \times 10^{-8}$). As previously reported rs9923231 is in tight LD with rs10871454 (located in the Syntaxin [*STX4*] gene 60 kb 5' of *VKORC1*) among EAs ($r^2 = 0.950$) and AAs ($r^2 = 0.975$). This association explains about 22% of the variability in warfarin dose among EAs and 5% in AAs.

Among EAs, after adjusting for clinical covariates, the single most significant *CYP2C9* SNP was rs4086116 (MAF 20.6%; $p = 1.24 \times 10^{-19}$) as reported by Teichert *et al.* [48]. *CYP2C9**3 ($p = 1.01 \times 10^{-13}$) and *CYP2C9**2 ($p = 2.34 \times 10^{-7}$) demonstrated significant influence on warfarin dose. As the LD between rs4086116 and *CYP2C9**3 ($r^2 = 0.269$) and with *CYP2C9**2 ($r^2 = 0.611$) is not high, we assessed dose variability explained by rs4086116. Inclusion of this SNP explains an additional (0.6%) variability in dose. *CYP4F2* (rs2108622) was also associated with dose in EAs (p = 0.009) and explained 0.5% of the dose variability. Among AAs, no *CYP2C9* SNP demonstrated statistically significant association with warfarin dose. Inclusion of rs4086116 did not explain additional variability in dose for AAs.

Table 2 shows results after adjusting for clinical factors, top two principal components and the aforementioned known genetic variants. The genomic inflation factor λ was 0.930 for EAs and 0.976 for AAs in this analysis. The SNP rs12777823 identified in the previous GWAS of AAs [25], intergenic to *HELLS* and *CYP2C18*, showed significant association with dose (Bonferroni adjusted p = 0.014 and FDR adjusted q = 0.014) and explained an additional 3.6% of the variability in warfarin dose. *CYP4F2* did not demonstrate significant influence on warfarin dose among AAs (Table 2).

In addition to the confirmation of some known findings, we identified novel significant association after adjustment of clinical covariates, known SNPs in *VKORC1*, *CYP2C9* and *CYP4F2* and top two principal components. Specifically, in AAs, SNP exm1279100 (Bonferroni adjusted p = 0.061; FDR q = 0.019) showed significant association explaining 4.71% of the variability in warfarin dose. Among EAs, SNP exm847442 showed marginally significant evidence (Bonferroni corrected p = 0.075; FDR q = 0.011) and explained an additional 1.7% warfarin dose variability. In addition to these top hits, several genetic markers reaching suggestive threshold in both races (Table 2; FDR $q < 0.1$) in the discovery cohort were tested in the validation cohort.

In addition to the separate analysis in EA and AA samples, we also conducted analysis in the pooled samples in the discovery cohort. Several SNPs in *VKORC1* or *CYP2C9*, or in genes flanking them reach genome-wide significance. However, no novel SNPs identified (data not shown). The genomic inflation factor λ was 1.015 for the pooled analysis.

For the gene-based analysis, genes demonstrating significant (or near significant) influence on warfarin dose are presented in Table 3. Specifically, among EAs, two genes *COX15* and *FGF5* showed significant association with warfarin dose (Bonferroni adjusted p-values are 0.0038 and 0.0338; FDR q-values are 0.0038 and 0.0169, respectively), and explained additional 2.4 and 1.6% of warfarin dose variability, respectively. Gene *ZNF776* showed marginally significant effect among AAs (Bonferroni adjusted p-value is 0.0995 and FDR q-values is 0.0859) and explained 3.3% variability in warfarin dose. In addition, multiple genes (Table 3) showed suggestive significance in both EAs and AAs.

Validation efforts included the 22 SNPs in EAs and 28 SNPs in AAs, which demonstrated significant or suggestive association in the discovery cohort. Nine of the 22 SNPs in EAs and 22 of the 28 SNPs in AAs were observed in the validation cohort (the others were either not observed or were singletons in the validation cohort). After accounting for the clinical factors and known SNPs in *CYP2C9*, *VKORC1* and *CYP4F2*, none of the SNPs showed significant evidence in EAs. Among AAs, three SNPs showed significant influence on warfarin dose after Bonferroni correction (Table 4). SNP exm1235612 (Bonferroni corrected p = 0.00022 and FDR q = 0.00016), a rare variant on chromosome 16 located in *VKORC1*, explains 2.6% dose variability. Both SNP rs12772169 (Bonferroni corrected p = 0.00044 and FDR q = 0.00018) and rs12777823 (Bonferroni corrected p = 0.00066 and FDR q = 0.00021) are common variants on chromosome 10, upstream from *CYP2C18*, with medium LD ($r^2 = 0.47$) between them. The allele frequencies and coefficient estimates are very similar in the discovery and validation cohorts, indicating the consistency in the distributions of those two SNPs and their effects. Their inclusion explains 2.3 and 3.6% variability in dose, respectively. For 18 genes (8 genes in EAs and 10 genes in AAs) showing significant or suggestive evidence in discovery cohort, almost none of them have the same set of markers in the validation cohort. Therefore, we could not perform validation for those genes.

We performed meta-analysis on the SNPs that were replicated in the validation cohort. The results shown in Table 4 are only for the 31 SNPs that are also observed in the validation cohort. Results are similar to those from validation study. The three SNPs (exm1235612, rs12772169 and rs12777823) that were significant in the validation study for AAs showed stronger signals in the meta-analysis. None of the other SNPs showed significant association.

In this study, clinical factors accounted for 22.4 and 16.4% of the variability in warfarin dose in AAs and EAs, respectively. The known genetic factors that we adjusted for in our analyses (i.e., *CYP2C9* [*2 (rs1799853), *3 (rs1057910)], *CYP24F2* [rs2108622] and *VKORC1*

Table 3. Gene-based analysis testing common and rare variation separately and together using the rare variant score after adjusting for clinical factors, *VKORC1, CYP2C9, CYP4F2* variants and top two principal components.

Gene	Common variants		Rare variants		P combined[§]	BONF	FDR	R^2
	n (SNPs)	P common[†]	n (SNPs)	P rare[‡]				
European–Americans								
COX15	2	0.8519	2	1.06E-08	2.71E-07	0.0038	0.0038	0.024
FGF5	0	NA	4	2.40E-06	2.40E-06	0.0338	0.0169	0.016
NDUFB3	2	0.012239	10	2.08E-04	1.67E-05	0.235	0.061	0.012
MAML3	2 (exm2269854, exm425416)	0.012239	10 (exm425409, exm425417, exm425423, exm425432, exm425468, exm425490, exm425520, exm425527, exm425529, exm425540)	2.08E-04	1.67E-05	0.235	0.061	0.018
ZC3H15	0	NA	2	1.73E-05	1.73E-05	0.244	0.061	0.013
GAK	3 (exm379302, exm-rs11248051, exm-rs1564282)	0.000182	13 (exm2080762, exm2256768, exm379300, exm379319, exm379354, exm379355, exm379369, exm379374, exm379392, exm379442, exm379538, exm379552, exm379570)	2.25E-03	2.77E-05	0.39	0.0651	0.019
C6orf120	0	NA	2	6.60E-05	6.60E-05	0.929	0.0929	0.011
SLC10A7	1	0.056912	2	4.86E-05	9.26E-05	1	0.1186	0.013
African–Americans								
ZNF776	1	0.05769	1	3.93E-06	6.74E-06	0.0995	0.0859	0.033
PLIN5	1	0.235	1	6.17E-06	2.06E-05	0.304	0.1014	0.03
BBC3	0	NA	2	4.86E-05	4.86E-05	0.718	0.1794	0.023
C2orf28	1	0.072	1	8.35E-05	6.62E-05	0.976	0.1953	0.027
SLCO3A1	4	7.90E-05	2	0.125731	9.71E-05	1	0.2215	0.037
PDGFRL	0	NA	14	0.000205	2.05E-04	1	0.2811	0.02
ANXA4	1	4.64E-05	2	0.654896	2.47E-04	1	0.2811	0.023
CCRL2	5	0.000173	9	0.154093	2.48E-04	1	0.2811	0.037
C17orf87	1	0.46	1	6.10E-05	2.32E-04	1	0.2811	0.023
CCDC141	4	0.000144	14	0.530749	2.83E-04	1	0.2978	0.033
CYP2C9	2 (exm844029, exm-rs4086116)	0.000303	3 (exm844046, exm-rs10509680, exm-rs1799853)	NA	3.03E-04	1	0.2982	0.023

FDR indicates FDR adjusted q-value.
[†]Indicates p-value for all common variants.
[‡]Indicates p-value for rare variants represented by a RVS.
[§]Indicates p-value for testing all common and rare variants (summarized by RVS) together.
BONF: Bonferroni adjusted p-value; FDR: False discovery rate; NA: Not calculated because some of the SNPs were adjusted for in the regression model; R^2: Semipartial R^2 from regression; RVS: Rare variant score.

[rs9923231]) explained 7.5% of the dose variability in AAs and 34.6% in EAs. Those are all consistent with the findings in the literature. The newly identified SNP rs12772169 explained an additional 2.3% dose variation in AAs. Together with the marginal significant SNP (exm1279100) and gene (*ZNF776*), those novel genetic factors explained an extra 9.9% dose variation in AAs. The newly identified genes (*COX15* and *FGF5*) accounted 3.3% of warfarin dose variation in EAs. Inclusion of the marginally significant SNP (exm847442) can explain 4.9% of new dose variation in EAs. In our data, all genetic variants (known and newly identified) accounted for 17.4 and 39.5% warfarin dose variation in AAs and EAs, respectively.

Table 4. Results from validation study and meta-analysis of discovery and validation cohorts after adjustment of clinical, genetic factors and top two principal components.

SNP	Gene	CHR	BP	Minor allele	Discovery cohort			Validation cohort			Meta p-value
					MAF	BETA	p-value	MAF	BETA	p-value	
European–Americans											
exm2262194	Intergenic *TSG1*- (dist = 9331), *MANEA* (dist = 1529883)	6	94,495,530	A	1.14%	-0.381	1.13E-05	1.83%	-0.1074	6.13E-01	5.36E-04
exm156678	*OBSCN*	1	2.29E+08	G	49.07%	-0.078	1.60E-05	48.54%	-0.0077	8.75E-01	1.57E-03
exm1619580	*BRD1*	22	50,169,762	A	0.14%	-1.039	1.80E-05	0.37%	0.599	1.05E-01	2.94E-05
exm1163794	*DYX1C1*	15	55,731,682	T	0.14%	0.976	6.15E-05	0.37%	0.07785	8.32E-01	2.85E-03
exm1201118	*TPSD1*	16	1,308,333	A	1.00%	0.367	7.84E-05	1.83%	-0.2156	2.26E-01	2.63E-04
exm1432416	*RAD23A*	19	13,059,147	G	3.07%	0.213	8.31E-05	2.56%	0.05228	7.20E-01	2.40E-03
exm413801	*MMRN1*	4	90,856,377	A	0.79%	0.407	9.22E-05	1.10%	-0.0232	9.14E-01	4.50E-03
exm1013500	*TIMELESS*	12	56,815,233	A	0.36%	-0.601	9.26E-05	0.37%	-0.1249	7.33E-01	2.65E-03
exm357145	*ZIC4*	3	1.47E+08	C	0.93%	0.349	9.75E-05	1.10%	0.1592	4.62E-01	1.06E-03
African–Americans											
exm-rs12777823	Intergenic *HELLS* (dist = 43646), *CYP2C18* (dist = 37749)	10	96,405,502	A	24.91%	-0.134	1.66E-07	24.83%	-0.17	3.00E-05	2.89E-11
exm1028932	*UHRF1BP1L*	12	1E+08	A	8.44%	-0.19	5.72E-06	6.60%	-0.06	3.51E-01	1.10E-04
exm1034271	*SELPLG*	12	1.09E+08	A	0.19%	-1.641	8.13E-06	0.33%	0.28	2.85E-01	9.22E-05
exm800598	*LCN12*	9	1.4E+08	A	0.19%	-1.158	9.35E-06	0.50%	0.16	4.48E-01	2.43E-04
exm-rs12772169		10	96,405,329	A	41.36%	-0.102	9.64E-06	38.74%	-0.15	2.00E-05	8.01E-10
exm1235612	*VKORC1*	16	31,104,720	A	0.28%	0.938	1.23E-05	0.17%	1.64	1.00E-05	5.13E-10
exm1608547	*DMC1*	22	38,917,703	G	0.19%	-1.124	1.77E-05	0.33%	-0.44	9.78E-02	2.60E-05
exm748191	*PIGO*	9	35,092,607	A	0.19%	-1.114	2.19E-05	0.17%	-0.34	3.58E-01	2.61E-04
exm199426	*WDR35*	2	70,033,584	A	32.47%	-0.106	2.28E-05	32.01%	0.00	9.72E-01	2.53E-03
exm249675	*CCDC141*	2	1.8E+08	A	15.86%	0.133	2.41E-05	16.67%	-0.04	2.73E-01	1.69E-04

Meta p-value is for p-values from meta-analysis of discovery and validation cohorts.
BETA: β coefficient from regression analysis; BP: Base-pair position; CHR: Chromosome; MAF: Minor allele frequency.

Table 4. Results from validation study and meta-analysis of discovery and validation cohorts after adjustment of clinical, genetic factors and top two principal components (cont.).

SNP	Gene	CHR	BP	Minor allele	Discovery cohort			Validation cohort			Meta p-value
					MAF	BETA	p-value	MAF	BETA	p-value	
African–Americans (cont.)											
exm608896	*DNAH11*	7	21,934,511	A	0.46%	-0.692	3.08E-05	0.99%	-0.10	5.31E-01	6.98E-04
exm327045	*C3orf67*	3	58,853,578	G	0.19%	-1.08	3.57E-05	0.17%	0.36	3.27E-01	2.99E-04
exm711934	*NIPAL2*	8	99,217,391	G	0.19%	-1.08	3.57E-05	0.33%	0.53	4.56E-02	1.45E-05
exm296431	*ZCWPW2*	3	28,476,661	A	0.46%	-0.679	4.74E-05	0.66%	-0.19	3.20E-01	3.43E-04
exm826588	*ANK3*	10	61,802,477	A	0.46%	-0.741	6.26E-05	0.50%	-0.06	7.73E-01	2.41E-03
exm2261915		5	3,311,493	G	45.82%	0.092	6.74E-05	42.50%	0.02	5.61E-01	1.24E-03
exm1484329	*BBC3*	19	47,735,796	T	0.29%	-0.846	7.53E-05	0.66%	-0.15	4.05E-01	7.03E-04
exm619489	*TNS3*	7	47,408,443	G	0.74%	-0.563	7.62E-05	0.99%	0.10	5.67E-01	1.36E-03
exm533622	*DOM3Z*	6	31,939,416	G	0.28%	0.854	8.38E-05	0.83%	0.20	2.87E-01	4.09E-04
exm1134736	*PACS2*	14	1.06E+08	A	0.19%	-1.454	8.76E-05	0.83%	-0.21	2.76E-01	3.94E-04
exm218846	*ST6GAL2*	2	1.07E+08	A	0.29%	0.865	8.81E-05	0.50%	0.02	9.35E-01	4.64E-03
exm-rs7495052	*SLCO3A1*	15	92,552,029	G	28.07%	-0.094	9.12E-05	32.89%	-0.02	5.78E-01	1.58E-03

Meta p-value is for p-values from meta-analysis of discovery and validation cohorts.
BETA: β coefficient from regression analysis; BP: Base-pair position; CHR: Chromosome; MAF: Minor allele frequency.

Discussion

Identification of novel genetic variants influencing warfarin response can provide additional insight with regard to genetic influences on warfarin pharmacokinetics and pharmacodynamics and enable more precise dosing [21,25,49]. In this study, we employed the Illumina Exome Array to identify novel coding SNPs that may explain additional variability in warfarin dose in both EAs and AAs. To our knowledge, this is the first study utilizing exome chip to search for genetic factors associated with warfarin dose response.

Concordant with prior studies, *VKORC1* demonstrated important influence on warfarin dose in both EAs and AAs. Among EAs, *CYP2C9*2* and **3* and *CYP4F2* also demonstrated significant influence on warfarin dose. In our study, among AAs, neither *CYP2C9* nor *CYP4F2* SNPs was significantly associated with warfarin dose. As previously reported, SNP rs12777823 was significantly associated with warfarin dose in AAs [25]. We also identified and validated in an independent sample that rs12772169 has significant influence on warfarin dose among AAs. Previous studies [48] showed that rs12772169 was significantly associated with coumarin dose but did not include AAs. Both rs12772169 and rs12777823 are on chromosome 10q23, upstream from *CYP2C18*. They reside within the *CYP2C* gene cluster that includes the *CYP2C9*, *CYP2C8*, *CYP2C18* and *CYP2C19* genes. They are 230 base pairs away from each other with medium LD (r^2 = 0.47). Perera *et al.* [25] stated that the association between rs12777823 and warfarin dose in AAs "might not be due to rs12777823 itself, but a causal SNP in LD with rs12777823 in AAs but not in other populations." When we included both SNPs in the model in AAs in the discovery cohort, adjusted for all aforementioned known clinical and genetic factors, SNP rs12777823 was still significant, but rs12772169 was not, indicating that the effect of rs12772169 may

be due to rs12777823. Another possibility is that there is a causal SNP in LD with both rs12777823 and rs12772169. However, rs12777823 did not show an association with warfarin dose in previous studies in different racial groups such as European, Japanese and Egyptian [18–19,32,50]. On the other hand, rs12772169 showed an association with warfarin dose in Caucasians in a previous study [48]. Taken together with the low LD between those two SNPs, a third possibility is that rs12772169 is a causative variant, but there exists another causal variant in LD with rs12777823. Further study is needed to determine the relationship of these two SNPs with warfarin dose response.

It should be noted that in the replication study only 9 out of the 22 top SNPs in EAs and 22 out of the 28 top SNPs in AAs were observed and tested in the validation cohort. The remaining 19 SNPs that were not observed are all rare variants. This is not unexpected, given the small sample size of the validation cohort. However, these SNPs may hold clues to variability in warfarin dosing and are worthy of further validation. Among them, SNPs exm1279100 (Bonferroni adjusted p = 0.061 and FDR q = 0.019) and exm847442 (Bonferroni corrected p = 0.075 and FDR q = 0.011) showed marginally significant association in the discovery cohort in AAs and EAs, respectively, even after Bonferroni correction. SNP exm1279100 locates on chromosome 17 in gene *P2RX1*, a member of the purigenic receptor family, associated with bleeding disorders, platelet activation congestive heart failure and neurogenic bladder [51–53]. Given that gene *P2RX1* is in the platelet activation, signaling and aggregation pathway and is associated with bleeding disorders, it is likely that *SNP* exm1279100 is associated with warfarin dose response. SNP exm409265 showed suggestive association in EAs (Bonferroni corrected p = 0.435 and FDR q = 0.011). It located in gene *FGF5* that showed significant association in our gene-based analysis in EAs.

In gene-based analysis, among EAs, genes *COX15* and *FGF5* showed significant association with warfarin dose after multiple testing corrections. Gene *ZNF776* showed marginally significant effect among AAs. *COX15*, a protein coding gene, located on the chromosome 10q24, catalyzes the electron transfer from reduced cytochrome c to oxygen and is associated with Alzheimer's disease, cardioencephalomyopathy and Leigh syndrome [54,55]. Porphyrin and chlorophyll metabolism is among its related pathways and also includes two CYP450 genes *CYP2A6* and *CYP3A4*. *CYP2A6* is associated with metabolism of coumarin that is used as a precursor reagent in the synthesis of several synthetic anticoagulants including warfarin [56]. Warfarin is administered as a racemic mixture of the *R* and *S* stereoisomers where (*R*)-warfarin is mainly metabolized via *CYP3A4*. Gene *FGF5*, located on the chromosome 4q21.2, encodes the FGF5, which is involved in a variety of biological processes, including tissue repair, tumor growth and invasion [57,58]. Studies have showed that *FGF5* is associated with blood pressure and hypertension [59] and may have an important

Summary points

Background

- Anticoagulation therapy with warfarin, the most widely used oral anticoagulant, remains challenging. Despite efforts, significant portion of the dose variability remains unexplained. We hypothesized that coding SNPs may explain additional warfarin dose variability.

Patients & methods

- A genome-wide association study was performed using the Illumina Exome Array to identify coding SNP variations associated with warfarin dose response.
- A total of 1680 warfarin users (838 European–Americans [EAs] and 842 African–Americans [AAs]) were included in the study. Among them 1240 (701 EAs and 539 AAs) were used in the discovery cohort and 440 patients (137 EAs and 303 AAs) were included in the validation cohort.

Results

- We confirmed the influence of *VKORC1* (rs9923231) on warfarin dose variability in both EAs and AAs, and the influence of *CYP2C9* (**2*, **3* and rs4086116) among EAs and rs12777823 among EAs.
- Our study is the first to show the association between rs12772169 ($p = 9.64 \times 10^{-6}$) and warfarin dose in AAs.
- In our study, genes *COX15* ($p = 2.71 \times 10^{-7}$) and *FGF5* ($p = 2.40 \times 10^{-6}$) showed significant association with warfarin dose in EAs.

Conclusion

- We identified some novel genes and SNPs associated with warfarin maintenance dose in EAs and AAs.
- The information provided by the newly identified SNPs and genes may be limited for the prediction of a patient's dose. However, with more and more such genetic variants being identified, collectively they may make a difference in clinical practice.
- The genetic architecture of warfarin dose response may be very different for EAs and AAs.

role in therapeutic angiogenesis [60]. Gene-based analysis in AAs yielded a different list of potential candidates. Perilipin 5 (*PLIN5*), located on chromosome 19p13.3, coats lipid droplets protecting them from degradation and is expressed in highly oxidative tissues such as the heart, kidney and liver, the primary site of warfarin metabolism [61,62]. A very recent study shows that *PLIN5* alters cardiac lipid metabolism and is protective in the ischemic heart [63]. Zinc Finger Protein (*ZNF776*), also locates on chromosome 19, has been implicated in the control of blood, bone and neural progenitor cells. Further validation and confirmation of these associations with the warfarin dose phenotype is needed.

Although our sample size is considered large for warfarin related study, it is still small for a genome-wide study such as our work. Therefore, we recognize that this study may be under-powered. Genotype imputation is commonly used in genetic studies [64–66]. Research has shown that study samples need to have enough genotype data in order to have good imputation results [67–69]. We acknowledge this limitation as our analysis was restricted to exome-array data, which provides a less than optimal scaffold to achieve good imputation results. Therefore, we did not perform imputation.

Future perspective

The importance of genetic and environmental factors in determining warfarin dose response is widely recognized. Our study evaluated the effect of rare and common variants confirming the effect of known variants and identifying novel variants on warfarin dose. The novel genetic variants accounted for an additional warfarin dose variation in AAs and EAs, respectively, compared with our recent report [23]. Whether, incorporation of these additional gene variants can improve anticoagulation control and reduce the risk of hemorrhagic or thromboembolic complications remain to be determined.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Web resources

The URLs for data presented herein are as follows:

ANNOVAR, http://annovar.openbioinformatics.org/en/latest/ [70]

PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/res.shtml [71]

HapMap project, www.genome.gov/10001688/international-hapmap-project/ [72]

EIGENSTRAT, http://genetics.med.harvard.edu/reich/Reich_Lab/Software.html [73]

References

1 Ageno W, Gallus AS, Wittkowsky A, Crowther M, Hylek EM, Palareti G. Oral anticoagulant therapy: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 141(2), e44S–e88S (2012).

2 Chan PS, Maddox TM, Tang FM, Spinler S, Spertus JA. Practice-level variation in warfarin use among outpatients with atrial fibrillation (from the NCDR PINNACLE program). *Am. J. Cardiol.* 108(8), 1136–1140 (2011).

3 Kirley K, Qato DM, Kornfield R, Stafford RS, Alexander GC. National trends in oral anticoagulant use in the United States, 2007 to 2011. *Circ. Cardiovasc. Qual.* 5(5), 615–621 (2012).

4 Hsu JC, Maddox TM, Kennedy KF *et al.* Oral anticoagulant therapy prescription in patients with atrial fibrillation across the spectrum of stroke risk: insights from the NCDR PINNACLE registry. *JAMA Cardiol.* 1(1), 55–62 (2016).

5 Raji MA, Lowery M, Lin YL, Kuo YF, Baillargeon J, Goodwin JS. National utilization patterns of warfarin use in older patients with atrial fibrillation: a population-based study of Medicare Part D beneficiaries. *Ann. Pharmacother.* 47(1), 35–42 (2013).

6 Khan AS, Chaudhry S, Qureshi AI. Antithrombotic utilization trends after noncardioembolic ischemic stroke or TIA in the setting of large antithrombotic trials (2002–2009). *J. Vasc. Interv. Neurol.* 8(1), 20–26 (2015).

7 Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation* 115(21), 2689–2696 (2007).

8 Hylek EM, Go AS, Chang Y *et al.* Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N. Engl. J. Med.* 349(11), 1019–1026 (2003).

9 Hylek EM, Singer DE. Risk factors for intracranial hemorrhage in outpatients taking warfarin. *Ann. Intern. Med.* 120(11), 897–902 (1994).

10 Landefeld CS, Goldman L. Major bleeding in outpatients treated with warfarin: incidence and prediction by factors known at the start of outpatient therapy. *Am. J. Med.* 87(2), 144–152 (1989).

11 Poli D, Antonucci E, Marcucci R *et al.* Risk of bleeding in very old atrial fibrillation patients on warfarin: relationship with ageing and CHADS(2) score. *Thromb. Res.* 121(3), 347–352 (2007).

12 Shehab N, Sperling LS, Kegler SR, Budnitz DS. National estimates of emergency department visits for hemorrhage-related adverse events from clopidogrel plus aspirin and from warfarin. *Arch. Intern. Med.* 170(21), 1926–1933 (2010).

13 Wittkowsky AK, Whitely KS, Devine EB, Nutescu E. Effect of age on international normalized ratio at the time of major bleeding in patients treated with warfarin. *Pharmacotherapy* 24(5), 600–605 (2004).

14 Budnitz DS, Lovegrove MC, Shehab N, Richards CL. Emergency hospitalizations for adverse drug events in older Americans. *N. Engl. J. Med.* 365(21), 2002–2012 (2011).

15 Budnitz DS, Shehab N, Kegler SR, Richards CL. Medication use leading to emergency department visits for adverse drug events in older adults. *Ann. Intern. Med.* 147(11), U755–U726 (2007).

16 Caldwell MD, Awad T, Johnson JA *et al. CYP4F2* genetic variant alters required warfarin dose. *Blood* 111(8), 4106–4112 (2008).

17 Cavallari LH, Langaee TY, Momary KM *et al.* Genetic and clinical predictors of warfarin dose requirements in African–Americans. *Clin. Pharmacol. Ther.* 87(4), 459–464 (2010).

18 Cha PC, Mushiroda T, Takahashi A *et al.* Genome-wide association study identifies genetic determinants of warfarin responsiveness for Japanese. *Hum. Mol. Genet.* 19(23), 4735–4744 (2010).

19 Cooper GM, Johnson JA, Langaee TY *et al.* A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood* 112(4), 1022–1027 (2008).

20 Eriksson N, Wallentin L, Berglund L *et al.* Genetic determinants of warfarin maintenance dose and time in therapeutic treatment range: a RE-LY genomics substudy. *Pharmacogenomics* 17(13), 1425–1439 (2016).

21 Klein TE, Altman RB, Eriksson N *et al.* Estimation of the warfarin dose with clinical and pharmacogenetic data. *N. Engl. J. Med.* 360(8), 753–764 (2009).

22 Limdi NA, Beasley TM, Crowley MR *et al. VKORC1* polymorphisms, haplotypes and haplotype groups on warfarin dose among African–Americans and European–Americans. *Pharmacogenomics* 9(10), 1445–1458 (2008).

23 Limdi NA, Brown TM, Yan Q *et al.* Race influences warfarin dose changes associated with genetic factors. *Blood* 126(4), 539–545 (2015).

24 Limdi NA, Wadelius M, Cavallari L *et al.* Warfarin pharmacogenetics: a single *VKORC1* polymorphism is predictive of dose across 3 racial groups. *Blood* 115(18), 3827–3834 (2010).

25 Perera MA, Cavallari LH, Limdi NA *et al.* Genetic variants associated with warfarin dose in African–American individuals: a genome-wide association study. *Lancet* 382(9894), 790–796 (2013).

26 Rieder MJ, Reiner AP, Gage BF *et al.* Effect of *VKORC1* haplotypes on transcriptional regulation and warfarin dose. *N. Engl. J. Med.* 352(22), 2285–2293 (2005).

27 Shendre A, Brown TM, Liu NJ *et al.* Race-specific influence of CYP4F2 on dose and risk of hemorrhage among warfarin users. *Pharmacotherapy* 36(3), 263–272 (2016).

28 Kaminsky LS, Zhang ZY. Human P450 metabolism of warfarin. *Pharmacol. Ther.* 73(1), 67–74 (1997).

29 Rettie AE, Korzekwa KR, Kunze KL *et al.* Hydroxylation of warfarin by human cDNA-expressed cytochrome P-450: a role for P-4502C9 in the etiology of (S)-warfarin-drug interactions. *Chem. Res. Toxicol.* 5(1), 54–59 (1992).

30 Li T, Chang CY, Jin DY, Lin PJ, Khvorova A, Stafford DW. Identification of the gene for vitamin K epoxide reductase. *Nature* 427(6974), 541–544 (2004).

31 Rost S, Fregin A, Ivaskevicius V *et al.* Mutations in *VKORC1* cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature* 427(6974), 537–541 (2004).

32 Takeuchi F, McGinnis R, Bourgeois S *et al.* A genome-wide association study confirms *VKORC1*, *CYP2C9*, and *CYP4F2* as principal genetic determinants of warfarin dose. *PLoS Genet.* 5(3), e1000433 (2009).

33 The International Hapmap Consortium. The International HapMap Project. *Nature* 426(6968), 789–796 (2003).

34 Exome Chip Design. http://genome.sph.umich.edu/wiki/Exome_Chip_Design

35 Kimmel SE, Chen Z, Price M *et al.* The influence of patient adherence on anticoagulation control with warfarin – results from the International Normalized Ratio Adherence and Genetics (IN-RANGE) Study. *Arch. Intern. Med.* 167(3), 229–235 (2007).

36 PLINK (version 1.07). http://pngu.mgh.harvard.edu/~purcell/plink/res.shtml

37 Purcell S, Neale B, Todd-Brown K *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81(3), 559–575 (2007).

38 Zeng Z, Shaffer JR, Wang X *et al.* Genome-wide association studies of pit-and-fissure- and smooth-surface caries in permanent dentition. *J. Dent. Res.* 92(5), 432–437 (2013).

39 Mangold E, Ludwig KU, Birnbaum S *et al.* Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. *Nat. Genet.* 42(1), 24–26 (2010).

40 Garg P, Ludwig KU, Bohmer AC *et al.* Genome-wide analysis of parent-of-origin effects in non-syndromic orofacial clefts. *Eur. J. Hum. Genet.* 22(6), 822–830 (2014).

41 Liu NJ, Zhao HY, Patki A, Limdi NA, Allison DB. Controlling population structure in human genetic association studies with samples of unrelated individuals. *Stat. Interface* 4(3), 317–326 (2011).

42 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38(8), 904–909 (2006).

43 Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2(12), e190 (2006).

44 Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. *Genet. Epidemiol.* 34(2), 188–193 (2010).

45 The R project for statistical computing. www.r-project.org/

46 Wang K, Li MY, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38(16), e164 (2010).

47 Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* 10(10), 1556–1566 (2015).

48 Teichert M, Eijgelsheim M, Rivadeneira F *et al.* A genome-wide association study of acenocoumarol maintenance dosage. *Hum. Mol. Genet.* 18(19), 3758–3768 (2009).

49 Johnson JA, Caudle KE, Gong L *et al.* Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for pharmacogenetics-guided warfarin dosing: 2017 update. *Clin. Pharmacol. Ther.* doi:10.1002/cpt.668 (2017) (Epub ahead of print).

50 Takeuchi F, Kashida M, Okazaki O *et al.* Evaluation of pharmacogenetic algorithm for warfarin dose requirements in Japanese patients. *Circ. J.* 74(5), 977–982 (2010).

51 Malmsjo M, Bergdahl A, Moller S *et al.* Congestive heart failure induces downregulation of P2X1-receptors in resistance arteries. *Cardiovasc. Res.* 43(1), 219–227 (1999).

52 Cattaneo M. The P2 receptors and congenital platelet function defects. *Semin. Thromb. Hemost.* 31(2), 168–173 (2005).

53 Cattaneo M. Molecular defects of the platelet P2 receptors. *Purinerg Signal* 7(3), 333–339 (2011).

54 Papadopoulou LC, Sue CM, Davidson MM *et al.* Fatal infantile cardioencephalomyopathy with COX deficiency and mutations in *SCO2*, a *COX* assembly gene. *Nat. Genet.* 23(3), 333–337 (1999).

55 Oquendo CE, Antonicka H, Shoubridge EA, Reardon W, Brown GK. Functional and genetic studies demonstrate that mutation in the *COX15* gene can cause Leigh syndrome. *J. Med. Genet.* 41(7), 540–544 (2004).

56 Pearce R, Greenway D, Parkinson A. Species-differences and interindividual variation in liver microsomal cytochrome P450 2a enzymes: effects on coumarin, dicumarol, and testosterone oxidation. *Arch. Biochem. Biophys.* 298(1), 211–225 (1992).

57 Zhan XI, Bates B, Hu XG, Goldfarb M. The human FGF-5 oncogene encodes a novel protein related to fibroblast growth-factors. *Mol. Cell. Biol.* 8(8), 3487–3495 (1988).

58 Allerstorfer S, Sonvilla G, Fischer H *et al.* FGF5 as an oncogenic factor in human glioblastoma multiforme: autocrine and paracrine activities. *Oncogene* 27(30), 4180–4190 (2008).

59 Liu C, Li HX, Qi QB *et al.* Common variants in or near *FGF5*, *CYP17A1* and *MTHFR* genes are associated with blood pressure and hypertension in Chinese Hans. *J. Hypertens.* 29(1), 70–75 (2011).

60 Rissanen TT, Markkanen JE, Arve K *et al.* Fibroblast growth factor 4 induces vascular permeability, angiogenesis, and arteriogenesis in a rabbit hind limb ischemia model. *FASEB J.* 16(13), 100–102 (2002).

61 Kimmel AR, Sztalryd C. Perilipin 5, a lipid droplet protein adapted to mitochondrial energy utilization. *Curr. Opin. Lipidol.* 25(2), 110–117 (2014).

62 Kuramoto K, Okamura T, Yamaguchi T *et al.* Perilipin 5, a lipid droplet-binding protein, protects heart from oxidative burden by sequestering fatty acid from excessive oxidation. *J. Biol. Chem.* 287(28), 23852–23863 (2012).

63 Drevinge C, Dalen KT, Mannila MN *et al.* Perilipin 5 is protective in the ischemic heart. *Int. J. Cardiol.* 219, 446–454 (2016).

64 Li L, Li Y, Browning SR *et al.* Performance of genotype imputation for rare variants identified in exons and flanking regions of genes. *PLoS ONE* 6(9), e24945 (2011).

65 Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu. Rev. Genomics Hum. Genet.* 10, 387–406 (2009).

66 Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* 11(7), 499–511 (2010).

67 Pei YF, Li J, Zhang L, Papasian CJ, Deng HW. Analyses and comparison of accuracy of different genotype imputation methods. *PLoS ONE* 3(10), e3551 (2008).

68 Pei YF, Zhang L, Li J, Deng HW. Analyses and comparison of imputation-based association methods. *PLoS ONE* 5(5), e10827 (2010).

69 Zhang B, Zhi D, Zhang K, Gao G, Limdi NA, Liu N. Practical consideration of genotype imputation: sample size, window size, reference choice, and untyped rate. *Stat. Interface* 4(3), 339–352 (2011).

70 ANNOVAR. ANNOVAR Documentation. http://annovar.openbioinformatics.org/en/latest/

71 Partners healtcare. PLINK. http://pngu.mgh.harvard.edu/~purcell/plink/res.shtml

72 NIH National Human Genome Research Institute. International HapMap project. www.genome.gov/10001688/international-hapmap-project/

73 Reich Laboratory Medical and popular genetics. Software. http://genetics.med.harvard.edu/reich/Reich_Lab/Software.