HHS Public Access

Author manuscript

Clin Pharmacol Ther. Author manuscript; available in PMC 2024 July 11.

Published in final edited form as:

Clin Pharmacol Ther. 2023 March ; 113(3): 624–633. doi:10.1002/cpt.2820.

Genomewide Association Study Identifies Copy Number Variants Associated With Warfarin Dose Response and Risk of Venous Thromboembolism in African Americans

Honghong Zhang[1], **Cristina Alarcon**[1], **Larisa H. Cavallari**[2], **Edith Nutescu**[3], **Gemma L. Carvill**[1,4], **Minoli A. Perera**[1,*,†], **Wenndy Hernandez**[5,†]

[1]Department of Pharmacology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

[2]Department of Pharmacotherapy and Translational Research, Center for Pharmacogenomics and Precision Medicine, College of Pharmacy, University of Florida, Gainesville, Florida, USA

[3]Department of Pharmacy Practice, College of Pharmacy, University of Illinois Chicago, Chicago, Illinois, USA

[4]Department of Neurology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA

[5]Section of Cardiology, University of Chicago, Chicago, Illinois, USA.

Abstract

The anticoagulant warfarin is commonly used to control and prevent thrombotic disorders, such as venous thromboembolism (VTE), which disproportionately afflicts African Americans. Despite the importance of copy number variants (CNVs), few studies have focused on characterizing and understanding their role in drug response and disease risk among African Americans. In this study, we conduct the first genome-wide analysis of CNVs to more comprehensively account for the contribution of genetic variation in warfarin dose requirement and VTE risk among African Americans. We used hidden Markov models to detect CNVs from high-throughput single-nucleotide polymorphism arrays for 340 African American participants in the International Warfarin Pharmacogenetics Consortium. We identified 11,570 CNVs resulting in 2,038 copy number variable regions (CNVRs) and found 3 CNVRs associated with warfarin dose requirement and 3 CNVRs associated with VTE risk in African Americans. CNVRs 1q31.2del and 6q14.1del were associated with increased warfarin dose requirement ($\beta = 11.18$ and 4.94, respectively; $P_{emp} = < 0.002$); CNVR 19p13.31del was associated with decreased warfarin dose requirement ($\beta = -1.41$, $P_{emp} = 0.0004$); CNVRs (2p22.1del and 5q35.1-q35.2del) were found to be associated with

*Correspondence: Minoli A. Perera (minoli.perera@northwestern.edu).
†These authors contributed equally to this work.

AUTHOR CONTRIBUTIONS

M.A.P., W.H., and H.Z. wrote the manuscript. M.A.P. and G.L.C. designed the research. W.H., H.Z., C.A., L.H.C., G.L.C., and E.N. performed the research. H.Z. and W.H. analyzed the data.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

increased risk of VTE (odds ratios (ORs) = 1.88 and 14.9, respectively; $P_{emp} \leq 0.02$); and CNVR 10q26.12del was associated with a decreased risk of VTE (OR = 0.6; $P_{emp} = 0.05$). Modeling of the 10q26.12del in HepG2 cells revealed that this deletion results in decreased fibrinogen gene expression, decreased fibrinogen and WDR11 protein levels, and decreased secretion of fibrinogen into the extracellular matrix. We found robust evidence that CNVRs could contribute to warfarin dose requirement and risk of VTE in African Americans and for 10q26.3del describe a possible pathogenic mechanism.

Warfarin continues to remain a widely used anticoagulant with more than 15 million prescriptions dispensed in 2017, ranking warfarin as the 51st most commonly prescribed medication in the United States alone.[1] Unfortunately, warfarin can have serious adverse effects ranging from occult bleeding to death. Population differences in warfarin dosing requirements have been observed with African Americans having a higher rate of major bleeding compared to White people.[2] Clinical factors, demographics, and single nucleotide polymorphisms (SNPs) in genes involved in warfarin dose variability (*CYP2C9*, *VKORC1*, *CYP4F2*, *CALU*, and *GGCX*) have been well studied.[3] Yet combined, these factors are estimated to explain ~ 50% of warfarin dose requirements in populations of European ancestry and even <30% among populations of African ancestry[4–8] — highlighting the need to improve our understanding of the genetic architecture, beyond SNPs, of warfarin dose requirement especially in non-European ancestry populations.

Venous thromboembolism (VTE) is a significant health problem in the United States, resulting in up to 600,000 new cases and up to 100,000 deaths annually.[9] VTE is comprised of deep vein thrombosis (DVT) and its potentially fatal complication, pulmonary embolism (PE). VTE susceptibility is complex and involves both genetic and acquired risk factors, which include increasing age, surgery, immobilization, active cancer, prolonged hospitalization, and obesity.[10] The risk of VTE also differs by ancestry where individuals of African descent have a higher incidence compared to individuals of European ancestry. African Americans exhibit the highest incidence of DVT and mortality rates of PE.[11,12] The mechanisms causing idiopathic VTE remain, in great part, unknown. Previous genome-wide association studies (GWAS), conducted primarily on populations of European ancestry, have confirmed several VTE-associated loci either near or in *F5*, *FGA-FGG*, *F2*, *F11*, *PROCR*, *SCLC44A2*, and *ABO* genes.[13–19] Although the findings from two GWAS for VTE risk among African Americans did not replicate those found in Europeans, they both suggest that the variants found to be strongly associated with VTE risk in populations of European ancestry are not associated in African Americans.[15,16]

Copy number variants (CNVs) are deletions and duplications of the human genome that span at least 1 kb in length and can be polymorphic.[20] Several complex disorders have been associated with CNVs.[21–23] Additionally, several CNVs in genes associated with drug response, including *CYP2D6* and *CYP2A6*, have been identified.[24] The continued development of high-throughput, high-density genotyping technologies used in GWAS combined with statistical calling algorithms offer a unique advantage of size resolution and sensitivity for CNV detection and provide an opportunity to further characterize genomic architecture.[25–27]

Very few genome-wide studies have focused on identifying and characterizing CNVs in African Americans.[28,29] Furthermore, the impact of CNVs on the myriad of phenotypic traits, genetic conditions, and drug response in African Americans remains to be elucidated. Warfarin is an effective anticoagulant to reduce morbidity and mortality in patients suffering from VTE. Here, we report the first to date genome-wide study of CNVs for warfarin dose requirement and VTE risk in African Americans in order to help better understand the etiology of VTE and mechanisms regulating the warfarin dose requirement.

METHODS

Study population

Participants were part of a previously conducted GWAS, in collaboration with the International Warfarin Pharmacogenetics Consortium (IWPC), on warfarin dose response in African Americans.[30] In the present study, we only considered patients locally enrolled at The University of Chicago and The University of Illinois in Chicago ($n = 340$). Briefly, all participants were at least 18 years old, self-reported African American, and on a stable maintenance dose of warfarin, defined as the same dose for at least 3 consecutive clinic visits that produced and international normalized ratio (INR) within the therapeutic range of 2 to 3. Data collected related to warfarin dose requirement included age; height; weight; ethnicity; sex; liver disease; stable warfarin dose; INR at stable warfarin dose; indication for warfarin therapy; and concomitant medications, including potential interacting drugs (e.g., amiodarone, phenytoin, and carbamazepine). For VTE, warfarin patients with a documented history of VTE, defined as DVT or PE, and without strong known risk factors, including prolonged hospitalization, surgery, active cancer or history of malignancy <5 years, pregnancy or puerperium, hormone replacement therapy, or protein C or S deficiency, were classified as cases. VTE (DVT and/or PE) was diagnosed by physicians using different methods, including venous examination of the lower extremities using duplex ultrasound, spiral computed tomography, computed tomography pulmonary angiogram, or ventilation-perfusion scan. Warfarin patients free of VTE were classified as control subjects. VTE controls were on warfarin therapy due to atrial fibrillation/flutter, valve replacement, cardiomyopathy, or stroke. Patients provided a DNA sample (whole-blood, saliva, or mouthwash sample). The research protocol was approved by the local institutional review boards (The University of Chicago and The University of Illinois), and study participants gave written informed consent.

Genotyping and identification of CNVs

Genome-wide intensity data were retrieved from the existing GWAS dataset genotyped with the Illumina 610 Quad BeadChip (Illumina, San Diego, CA) at the RIKEN Center for Genomic Medicine (Yokohama, Japan).[30] Probe information for each sample, including genomic location, signal intensity (Norm R), allele frequency (Norm theta), log R ratios, and B allele frequencies (BAFs), were calculated and exported using GenomeStudio. A copy number state of 2 represents the normal diploid state, a lower value (0 and 1) represents a deletion, and a higher value[3,4] represents a duplication or triplication, respectively. The start and end coordinates of each CNV in our dataset were based on the human genome build GRCh37/hg19. For CNV calling, we chose two different segmentation algorithms that

implement the hidden Markov model, PennCNV[27] (June 2016) and QuantiSNP 2.04.[25] To reduce false positive results, only CNVs that were independently called by each algorithm were included in the analysis. CNV calls from each algorithm were classified as either deletion or duplication without accounting for hetero or homozygosity.

Stage 1 quality control

SNPs were excluded based on a genotyping rate <95%, a minor allele frequency of < 0.03, and failed Hardy–Weinberg equilibrium tests at $P < 0.00001$. We also removed all SNPs that did not get called in any individual because these loci may be CNV markers, particularly for the Quad 610 Bead chip. G/C and A/T SNPs were removed due to ambiguity. Genomewide genotype data were used to validate sex, as well as identity by descent. No sample had a call rate of <95%, missingness > 0.10, sex misspecification, or identity by descent >0.125.

Stage 2 quality control

CNV detection was performed for each chromosome separately. Individuals with a total number of CNVs exceeding the mean by more than three SDs were excluded, which helps reduce the number of false positive results and improves the reliability of CNV calls. We also removed samples based on previously used criteria for quality control: waviness factor ≥0.05, SD for autosomal log R ratio >0.28, a median BAF >0.55 or < 0.45, or a BAF drift of >0.002. A total of 326 samples passed quality control criteria and were included in the analysis.

We excluded immunoglobulin regions, telomeric, and centromeric regions due to the high frequency or repeats in these regions and the unreliability of calling CNVs in these regions. Adjacent calls with a gap of < 20% of total length were merged together into one single call. CNVs <10,000 bps were kept in the analysis if all of the following criteria were met: present in at least 2 individuals, found in at least 2 plates, had the same start and end positions, and the same SNPs on the genotyping platform were used to call the CNV. A total of 11,570 CNVs were identified in our African American cohort.

Copy number variable region determination

Due to technical fluctuations, the same CNV is frequently called with slightly different breakpoint locations in different individuals. Therefore, we used BedTools Intersect intervals from the Galaxy Project to merge CNVs overlapping at least one base pair into single CNV regions (CNVRs).[31] The breakpoints of each of these CNVRs were defined by the outer boundaries of the cumulative individual CNVs. Only CNVRs called by both PennCNV and QuantiSNP were included. The 11,570 CNVs identified resulted in 2,038 non-overlapping CNVRs, which were used for association analyses of warfarin dose requirement and VTE risk. For comparison to previously reported CNV loci, we obtained the March 29, 2012 release of the Database of Genomic Variants (DGVs).[32] A table summarizing DGV build 37 (hg19) was downloaded, and the start and end points of each DGV were used to identify any overlap with a CNVR. If a DGV overlapped by at least one base pair, then the CNVR was considered to be previously identified. The UCSC Genome Browser LiftOver tool (https://genome.ucsc.edu) was used to convert genomic coordinates from the original assembly to

GRCh37/hg19 to identify phenotypes previously associated with CNVs within our CNVRs and to build GRCh38/hg38 to identify genes more recently discovered.

Validation and determination of breakpoints for associated CNVRs

CNVRs associated with warfarin dose requirement and VTE risk were validated by high-resolution custom Agilent 8 X 15 K microarray-based Comparative Genomic Hybridization targeting each region specifically to confirm the breakpoints of each CNVR and processed and analyzed according to the manufacturer's instructions. We used a subset of eight subjects shown to carry the CNVRs of interest. We were able to validate the breakpoints of three CNVRs we found to be associated with either phenotypec (Tables 2 and 4).

CRISPR-Cas9 deletion of 10q26.12 CNVR

We chose CNVR at 10q26.12 due to the strong prior evidence for association with plasma fibrinogen and cardiovascular phenotypes. Guide RNAs (gRNAs) were designed using CHOPCHOP. The gRNA sets were tested for efficacy and accuracy of cutting using Guide-it sgRNA Screening Kit (Takara, 632639), resulting in 2 sets of gRNAs, dubbed gRNA1 + 2 and gRNA1 + 3 which cut at the base positions 121,009,657–121,024,970 (Genomic coordinate in build GRCh38/hg38) on Chromosome 10 and covered 83% of CNVR (Table S1). The gRNAs were cloned into eSpCas9–2A-Puro (PX459) or eSpCas9–2A-GFP (PX458) from GenScript. Cas9 and the gRNA expression plasmids were introduced into HepG2 cells by transient transfection using Lipofectamine 3,000 Reagent (ThermoFisher Scientific) according to the manufacturer's instructions. After 72 hours of puromycin selection and expansion, genomic DNA was isolated from pooled transfected cells, and polymerase chain reaction (PCR) was conducted to confirm the deletion of the selected CNVR. We designed primer pairs (Forward: 5'-GTAAG ATG TCT GCC ACC TACC-3′, Reverse: 5'-GTGAC TGT GGT TTT CAG TTGC-3') which flank the 15,623 base pair regions on chromosome 10 to confirm the deletion after the transfection of gRNA1 + 2 and gRNA1 + 3.

Assessment of protein concentration via Western blot

Transfected cells were lysed in RIPA lysis buffer (Thermo Scientific, 11,965,092) supplemented with protease inhibitor cocktail (Roche, 11,836,170,001). Whole cell lysates were centrifuged, and protein concentration was determined using Pierce Coomassie Protein Assay kit (Thermo Scientific, 23,200). Equal amounts of proteins were separated by standard SDS-PAGE and transferred to PVDF membrane. After blocking with 5% non-fat milk, the membranes were incubated with primary and secondary antibodies. The immune complexes were detected using Amersham ECL Western Blotting Detection Reagent (GE Healthcare). The protein quantification was done by Gels Analyze tool of Image J. WDR11 polyclonal antibody (A302–632A; Thermo Scientific), Fibrinogen alpha polyclonal antibody (PA5–35319; Invitrogen), β-Actin monoclonal antibody (3,700; Cell Signaling) were used in the Western blot experiment. Results represent Western blot obtained from at least three independent experiments.

Quantitative reverse-transcriptase polymerase chain reaction of WDR11 and fibrinogen in HepG2 cells

To determine the effect of the 10q26.12del CNVR on *WDR11* and fibrinogen gene expression, we conducted quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR). Briefly, RNA was isolated from HepG2 cells using the miRNeasy Mini Kit (QIAGEN; 217004). RNA concentration was determined using Nanodrop. Reverse transcription was performed using GoScript Reverse Trancriptase Kit (Promega; A5001) with random primers. The qRT-PCR was performed using GoTaq qPCR Master Mix (Promega; A6001) and QuantStudio 7 Fast Real Time PCR system. The relative gene expression was calculated by the double delta Ct method, using the Ct values from the genes of interest (*WDR11* and fibrinogen) and the housekeeping gene (TATA-box binding protein (*TBP*)) for both control and 10q26.12del samples. The sequence of primers are as follows: *WDR11* (Forward: 5′-AGACA GTG CTC GGA TTC CACCA-3′, Reverse: 5′-CTGGA TAC TCT GCC TTT CAAGTC-3′); fibrinogen (Forward: 5′-CAACG GCA TGT TCT TCA GCACG-3′, Reverse: 5′-GTATC TGC CGT TTG GAT TGGCTG-3′); *TBP* (Forward: 5′-TGTAT CCA CAG TGA ATC TTGGTTG-3′, Reverse: 5′-GGTTC GTG GCT CTC TTA TCCTC-3′).

Fibrinogen enzyme-linked immunosorbent assay from HepG2 supernatant

To determine the supernatant concentration of fibrinogen, enzyme-linked immunosorbent assay (ELISA) was performed in HepG2 cells using the Human Fibrinogen ELISA Kit (BIOMATIK; EKU04090) according to the manufacturer's instructions. Briefly, 100 µl of standards or samples were added to each well of a 96-well plate coated with fibrinogen. Following an hour incubation, 100 µL of Detection Reagent A working solution was added to each well and incubation was carried out for an hour. After washing with 350 µL of 1 × wash solution for 3 times, 100 µL of Detection Reagent B working solution was subsequently added to each well. After 30 minutes of incubation, the wells were washed 5 times. Next, 90 µL of substrate solution was then added to each well. After incubation for 30 minutes, 50 µL of stop solution was added to each well. All measurements were conducted at 450 nm immediately following the addition of stop solution. Calibration curves were constructed with the following concentration of fibrinogen: 8,000, 4,000, 2,000, 1,000, 500, 250, 125, and 0 pg/mL and were used to determine sample concentration.

Statistical analysis

Principal component (PC) analysis was conducted using genome-wide complex trait analysis,[33] a linkage disequilibrium-pruned ($r^2 > 0.2$) set of 138,409 markers and reference populations from the 1,000 Genomes Project phase I integrated variant set release in National Center for Biotechnology Information (NCBI) build 37 (hg19) coordinates (Figure S4). Percentage West African Ancestry was determined for each individual using ancestry-informative markers for European and West African ancestry.[34] The demographic and clinical characteristics were analyzed by the chi-square analysis or Fisher exact test using SPSS, and statistical significance for each covariate was determined as $P < 0.05$. Association between CNVRs and VTE risk was conducted using a logistic regression model adjusting for age, sex, and the first three PCs using PLINK.[35] For the association between CNVRs and

warfarin dose requirement, weekly stable warfarin dose was found to be normally distributed (Shapiro Wilk $P > 0.05$) and a linear regression model was used adjusted for age, weight, and VTE using PLINK.[35] One-sided empirical P values were calculated based on 1,000,000 permutations, statistical significance for each CNVR was determined as $P_{emp} < 0.05$. T-test and analysis of variance were used to compare protein and mRNA differences between control vs. gRNA1 + 2 and gRNA1 + 3 CNVR deletions.

RESULTS

Sample characteristics

The cohort's clinical and demographic characteristics are presented in Table 1. The mean weekly warfarin dose was 45.5 ± 18.8 mg. Increasing age was associated with a lower warfarin dose requirement whereas height and weight were associated with higher warfarin dose requirement ($\beta = -0.45$, $P < 0.001$; $\beta = 0.23$ and 0.16, $P \leq 0.01$, respectively; Table 1). Mean INR was 2.45 ± 0.35 (Table 1). Neither mean West African ancestry nor sex was associated with warfarin dose response ($\beta = 6.26$ and − 0.84, respectively; $P > 0.5$; Table 1). Nearly half of the cohort (47.9%) was taking warfarin for the indication of VTE, and this was significantly associated with an increased warfarin dose requirement ($\beta = 1.86$, $P < 0.001$; Table 1). The second most common indication for warfarin was atrial fibrillation (24.8%); cerebrovascular accident, mitral valve replacement, and sigmoid sinus thrombus comprised the remaining indications for warfarin therapy (Table 1).

CNV discovery

Figure S1 shows the flowchart for the CNV discovery and association analyses. All samples that passed quality control (QC; $n = 326$) were included in the CNV discovery phase regardless of warfarin dose or VTE status. The African American cohort that participated in IWPC first underwent stage 1 QC, where SNPs that did not pass QC thresholds were removed. After generating the CNV calls, samples underwent stage 2 QC, and the consensus of the CNV calls from the 2 algorithms was generated. After stringent QC filtering measures (Figure S1), a total of 11,570 CNVs were predicted in our African American cohort ($n = 326$) by both PennCNV and QuantiSNP (Table S2). Approximately 84% of CNVs detected were deletions, with heterozygous deletions (CNV event 1) accounting for 75.9% of all CNV events (Table S2). The median CNV length was 15,626 bp, with a minimum of 161 bp and maximum of nearly 2 Mbp (Table S2). The median number of SNPs used to detect CNVs was 17, with a minimum of 3 and a maximum of 1,264 SNPs (Table S2). An example of CNVs identified can be found in Figure S2. The breakpoints of each of these CNVRs were defined by the outer boundaries of the individual CNVs. The 11,570 CNVs resulted in a total of 2,038 CNVRs, ranging from 177 bp – 2,7 Mbp (Table S2). The distribution and density of CNVRs varied between chromosomes (Figure S3). Chromosomes 1, 2, and 4 contained the most CNVRs (177, 164, and 161, respectively) whereas chromosomes 18, 13, and 3 had the higher number of CNVRs per Mbp (26.65, 23.45, and 23.04, respectively; Figure S3). No unique CNVRs were detected as all were observed previously in the DGV.

Identification of candidate warfarin dose requirement CNVRs

Given the lack of studies exploring the role of CNVs in pharmacogenomics among African Americans, we tested for association between CNVRs and warfarin dose requirement in our cohort and found three CNVRs to be significantly associated with warfarin dose requirement (Table 2). All three CNVRs were deletions, ranged between ~ 51 Kbp and 555 Kbp, and with frequencies between 4% and 24% in our cohort (Table 2). Both 1q31.2del and 6q14.1del were associated with increased warfarin dose requirement ($\beta = 11.18 \pm 1.14$–21.23 and 4.94 ± 1.72–8.60, respectively; $P_{emp} = < 0.002$; Table 2) whereas 19p13.31del was associated with a decreased warfarin dose requirement in African Americans ($\beta = -1.41 \pm -3.23$ to −0.70, $P_{emp} = 0.0004$; Table 2). Using a custom comparative Agilent 8 X 15 K microarray-based Comparative Genomic Hybridization array, we were able to validate CNV 1q31.2del. Figure S5 shows the change in weekly warfarin dose of 57.09 mg vs. 45.9 mg with the presence of the 1q31.2del CNV vs. no CNV.

Identification of candidate VTE risk associated CNVRs

Our next step was to examine if CNVRs were associated with VTE risk in African Americans. The warfarin patient samples were divided into cases (patients with VTE) and controls (all other patients) status, and clinical and demographic characteristics are given in Table 3. Of the 172 VTE cases, 51% had DVT; 24% had DVT/PE, and 25% had isolated PE (Table 3). VTE cases were significantly younger compared with controls (54.5 ± 16.1 and 59.5 ± 12.9, respectively; $P = 0.001$; Table 3). Height, weight, and West African ancestry were not associated with risk of VTE ($P > 0.20$; Table 3). Men were more likely to be VTE cases than women (68.0% and 58.8%, respectively; $P = 0.01$; Table 3).

We identified 3 CNVRs associated with VTE risk in African Americans with frequencies ranging from 5.8 to 29.4% and lengths from ~ 18 Kbp–2 30 Kbp (Table 4). All CNVRs were deletions, and none have been previously associated with VTE risk (Table 4). Both 2p22.1del and 5q35.1-q35.2del were associated with increased risk of VTE (odds ratios (ORs) = 1.88 and 14.9, respectively; $P_{emp} \leq 0.02$; Table 4). In addition, we found 10q26.12del associated with a decreased risk of VTE (OR 0.6; $P = 0.05$; Table 4). We compared genomic coordinates (GRCh38/hg38) of the CNVRs from this study with those previously reported in the DECIPHER database and found that chr2:41000001–42,000,000 duplications and deletions have been associated with malformation of the heart and great vessels. Chr5:172500001–1 73,500,000 contain duplications and deletions, which span across CNVR 5q35.1-q35.2del, have been identified in patients with tachycardia, ventricular septal defect, tetralogy of Fallot, and atrial septal defect.[36] CNVR 10q26.12del is located within chr10:121000001–1 22,000,000, and duplications or deletions have been identified in patients with several conditions related to heart defects, including abnormal heart morphology, patent ductus arteriosus, and ventricular septal defect. These findings are particularly interesting as heart defects can increase the risk of thrombotic events, including VTE.[37]

Modeling of the 10q26.12del in VTE risk in HepG2 cells

Among the CNVRs found to be associated with VTE risk, 10q26.12del (validated) was of particular interest because local ancestry within this region has been previously associated to

myocardial infarction risk in African Americans.[38] Although among individuals of European ancestry, SNP rs2420915 located near the 10q26.12del region was found to be associated with lower fibrinogen plasma concentration in a GWAS.[39] Interestingly, SNP rs2420915 is also an expression quantitative trait loci for the gene *WDR11* in whole blood where the minor allele, A, was associated with decreased expression of *WDR11*,[39] a gene that is upstream of the discovered CNVR. Myocardial infarction is most commonly caused by a thrombotic event and fibrinogen, or clotting Factor I, is essential for proper blood clotting[40] and support our findings that the 10q26.12del may contribute to VTE risk in African Americans. To study the impact of 10q26.12del on gene expression in the region, we used two paired gRNAs (gRNA1/gRNA2 and gRNA1/ gRNA3) and CRISPR/Cas9-mediated genome engineering to delete this region (Table S1, Figure 1a). Using PCR amplification, as expected, no PCR band was detectable in control HepG2 cells (Figure 1b), as the PCR product would be too large to be amplified. However, we readily detected a PCR product in the sample in which gRNAs had been combined (Figure 1b), indicating the successful deletion of this region and the rejoining of the DNA ends after double-stranded breaks. Using semiquantitative analysis of PCR products spanning the internal breakpoints of the deletion, we determined that 30.6% and 80.4% of cells carried the 10q26.12del with gRNA1/gRNA2 and gRNA1/gRNA3, respectively. Therefore, we used gRNA1 + 2, gRNA1 + 3 as gRNA pair for all subsequent experiments.

To determine if the 10q26.12del affects the expression/abundance of the nearby gene/protein, *WDR11*, and fibrinogen plasma concentration, we performed qRT-PCR and immunoblotting and measured fibrinogen and *WDR11* mRNA expression and protein levels (Figure 2). The qRT-PCR analysis of 10q26.12del HepG2 cells (HepG2-del) compared with HepG2 control cells showed fibrinogen expression levels decreased by 42% in gRNA1 + 2 HepG2-del cells ($P = 0.002$) and by 55% in gRNA1 + 3 HepG2-del cells ($P = 0.0001$; Figure 2a), and *WDR11* mRNA expression was not statistically significant ($P = 0.25$; Figure 2b). We examined protein abundance using immunoblotting and found that the 10q26.12del also decreased fibrinogen and WDR11 protein levels (Figure 2c–e). We found that fibrinogen protein levels decreased by 50% in gRNA1 + 2 HepG2-del cells ($P = 0.003$) and by 71% in gRNA1 + 3 HepG2-del cells ($P = 0.0003$; Figure 2d), and WDR11 protein abundance decreased by 35% in gRNA1 + 2 HepG2-del cells ($P = 0.04$) and by 56% in gRNA1 + 3 HepG2-del cells ($P = 0.003$; Figure 2e). To further evaluate the effect of 10q26.12del CNVR deletion on fibrinogen, we examined the extracellular level of fibrinogen produced by HepG2 cells via ELISA. This experiment showed that the 10q26.12del decreased the secretion of fibrinogen into extracellular matrix by 63% in gRNA1 + 2 HepG2-del cells ($P = 0.001$) and by 55% in gRNA1 + 3 HepG2-del cells ($P = 0.03$; Figure 3). The observed decrease in fibrinogen expression and protein levels among cells with the 10q26.12del CNVR deletion strongly support our finding with decreased risk of VTE in African Americans.

DISCUSSION

CNVs contribute substantially to phenotypic diversity among individuals, but despite the importance of CNVs in disease risk and pharmacogenomics, coupled with the rich genetic diversity found in individuals of African ancestry, very few studies have focused on

characterizing and understanding the role of CNVs in African Americans.[28,29] In this study, we present the results of the first GWAS of CNVs in warfarin dose requirement and VTE risk among African Americans.

We identified three CNVRs significantly associated with warfarin dose requirement (Table 2). The CNVR 1q31.2del was associated with increased warfarin dose requirement (β = 11.18 ± 1.14–21.23, P_{emp} = 0.0023; Table 2). The wide confidence interval is due to the small cohort size and hence is less precise in estimating effect size. An approximate 11 mg/week increase in warfarin dose is similar to the effect seen with the African specific GWAS variant rs12777823. Well known coding SNPs such as *CYP2C9*2* and **3* have been associated with changes of between 14–20% and 21–49% of warfarin dose respectively.[30,41] These effects vary by population group. This locus harbors a cluster of lncRNAs. The role of lncRNAs in pharmacogenomics is only beginning to emerge and so far has been shown to be critical in cancer drug response.[42] Nearly 24% of our African American warfarin cohort were carriers of the CNVR 6q14.1del, which was also associated with a higher warfarin dose requirement (β = 4.94, P_{emp} = 0.0002; Table 2). According to the ENCODE Registry of candidate cis-Regulatory Elements, the region encompassing CNVR 6q14.1del contains four regulatory elements for genes PHIP, HMGN3, and IRAK1BP1 (Table 2). Although these genes have not been implicated in warfarin drug response, genetic variation of regulatory elements' sequence can be a significant contributor to human disease and drug response. One CNVR, 19p13.31del, was associated with a lower warfarin dose requirement (β = −0.41, P_{emp} = 0.0004; Table 2) and encompasses the *PSG* genes. In contrast to our findings, *PSG1* and *PSG9* have been shown to be antithrombotic by inhibiting platelet-f ibrinogen interaction during pregnancy.[43] Although CNVRs 6q14.1del and 19p13.31del were not validated in our study, these CNVRs have been previously reported in the DGV. In accordance with previous studies among African Americans, we did not find an association between warfarin dose requirement and CNVs among genes involved in warfarin dose variability (*CYP2C9*, *VKORC1*, *CYP4F2*, *CALU*, and *GGCX*) suggesting genetic variation outside of these well-studied genes may influence warfarin dose requirement in African Americans.[44]

We found 2 CNVRs, 2p22.1del and 5q35.1-q35.2del, significantly associated with increased risk of VTE (OR = 1.88 and 14.9, respectively, $P_{emp} \leq 0.02$; Table 4) and one CNVR, 10q26.12del, associated with decreased risk of VTE (OR = 0.60, P_{emp} = 0.05; Table 4) among African Americans. These three CNVRs contained clusters of miRNAs and/or lncRNAs. Although there is a significant gap in knowledge of the potential role of non-coding RNA molecules, studies have linked miRNAs to VTE risk and have suggested the use of miRNAs as therapeutic targets.[45] In addition to harboring a lncRNA cluster, CNVR 5q35.1-q35.2del also encompasses genes *DUSPI* and *ERGIC1* for which deregulated gene expression has been associated with risk of myocardial infarction.[46] We decided to pursue functional validation of the 10q26.12 deletion because of the previous evidence in both GWAS and gene expression studies linking this region to changes in fibrinogen plasma levels. We demonstrate that in 2 sets of HepG2 cells harboring the CNVR 10q26.12del, fibrinogen gene expression was decreased compared with controls (P = 0.02 and P = 0.0001; Figure 2a). *WDR11* gene expression showed no significant change in HepG2 cells carrying the 10q26.12del (P = 0.25; Figure 2b). We further examined the effect of 10q26.12del on

fibrinogen and WDR11 protein levels (Figure 2e). This result is particularly striking in light of the fact that we estimate that 30.6% and 80.4% of cells carried the 10q26.12del with gRNA1/gRNA2 and gRNA1/gRNA3, respectively, thus we anticipate that even stronger effects may be observed in clonal cell cultures. However, clonal lines cannot be established in HepG2 cells as they do not grow in single-cell cultures. It should be noted that the 10q26.12 CNVR is reported in DGV as occurring in Bantu Africans and one Bedouin (Accession Number: dgv1405n54).[47] In the gnomAD database, this CNV has an allele count of 1,039, in which 1,002 were found in African or African American subjects. Although none of our CNVRs harbor variants previously associated with risk of VTE,[14–16,48] taken together, our results provide strong evidence that CNVRs are associated with VTE risk in African Americans. Deficiency of fibrinogen expression of 10q26.12del CNVR carriers might be an underlying mechanism for their lower risk to VTE.

Our study has several limitations that must be taken into consideration. Approximately 84% of CNVs detected were deletions (9,737; Table S2). This is due to the genome-wide signal intensity data used, which is more sensitive to the detection of deletions as opposed to duplications.[27] CNVs that were < 10 Kbs were excluded if they were not present in at least 2 individuals even if the CNV was called by both algorithms. The stringent QC measures implemented in our study may result in CNVs missing but also served to minimize false-positive predictions. We were only able to functional validate one CNV using CRISPR/Cas9 genome editing. This is due to the limitations in the size of CNV that can be created using CRISPR methods. Nonetheless, our quality control measures increased the accuracy of CNVs detected with high confidence. None of the CNVs identified in our African American cohort were novel. However, this may be due to the SNP genotyping platform that was used to call the CNVs (Illumina 610 Quad BeadChip) which offers less coverage of genetic variants for non-European populations since these arrays have been designed based primarily on genetic variation found in European populations. A recent study, utilizing sequencing data, detected a much greater number of CNVRs among African populations and identified novel CNVs.[49] Nonetheless, our study builds upon the resources established from a GWAS on warfarin dose response among African Americans conducted in collaboration with the IWPC.[30] In addition, we were able to validate three out of the six CNVRs (50%) significantly associated with warfarin dose requirement or VTE risk.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

FUNDING

This research was supported by NIH/National Heart, Lung, and Blood Institute grant R21HL106097 (M.A.P.); NHLBI Research Supplements to Promote Diversity in Health-Related Research (R01 HL136962, W.H.); and KL2 TR 002387 (W.H.).

References

1. Kane S. Warfarin, ClinCalc DrugStats database, version 20.0. ClinCalc. In. <https://clincalc.com/DrugStats/Drugs/Warfarin> Accessed December 23, 2019. In: ClinCalc.

2. Limdi NA, Brown TM, Shendre A, Liu N, Hill CE & Beasley TM Quality of anticoagulation control and hemorrhage risk among African American and European American warfarin users. Pharmacogenet. Genomics 27, 347–355 (2017). [PubMed: 28806200]

3. Johnson JA & Cavallari LH Warfarin pharmacogenetics. Trends Cardiovasc. Med 25, 33–41 (2015). [PubMed: 25282448]

4. Gage BF et al. Use of pharmacogenetic and clinical factors to predict the therapeutic dose of warfarin. Clin. Pharmacol. Ther 84, 326–331 (2008). [PubMed: 18305455]

5. Limdi NA et al. VKORC1 polymorphisms, haplotypes and haplotype groups on warfarin dose among African-Americans and European-A mericans. Pharmacogenomics 9, 1445–1458 (2008). [PubMed: 18855533]

6. Rieder MJ et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N. Engl. J. Med 352, 2285–2293 (2005). [PubMed: 15930419]

7. Wadelius M. et al. The largest prospective warfarin-treated cohort supports genetic forecasting. Blood 113, 784–792 (2009). [PubMed: 18574025]

8. Limdi NA et al. Race influences warfarin dose changes associated with genetic factors. Blood 126, 539–545 (2015). [PubMed: 26024874]

9. Office of the Surgeion General (US); National Heart, Lung, and Blood Institute (US). The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism (Office of the Surgeon General (US), Rockville, MD, 2008). <https://www.ncbi.nlm.nih.gov/books/NBK44178/>

10. Heit JA, Spencer FA & White RH The epidemiology of venous thromboembolism. J. Thromb. Thrombolysis 41, 3–14 (2016). [PubMed: 26780736]

11. Schneider D, Lilienfeld DE & Im W. The epidemiology of pulmonary embolism: racial contrasts in incidence and in-hospital case fatality. J. Natl. Med. Assoc 98, 1967–1972 (2006). [PubMed: 17225843]

12. White RH & Keenan CR Effects of race and ethnicity on the incidence of venous thromboembolism. Thromb. Res 123(Suppl 4), S11–S17 (2009). [PubMed: 19303496]

13. Germain M. et al. Genetics of venous thrombosis: insights from a new genome wide association study. PLoS One 6, e25581 (2011).

14. Heit JA et al. A genome-wide association study of venous thromboembolism identifies risk variants in chromosomes 1q24.2 and 9q. J. Thromb. Haemost 10, 1521–1 531 (2012). [PubMed: 22672568]

15. Heit JA et al. Identification of unique venous thromboembolism-susceptibility variants in African-Americans. Thromb. Haemost 117, 758–768 (2017). [PubMed: 28203683]

16. Hernandez W. et al. Novel genetic predictors of venous thromboembolism risk in African Americans. Blood 127, 1923–1929 (2016). [PubMed: 26888256]

17. Klarin D. et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. Nat. Genet 51, 1574–1579 (2019). [PubMed: 31676865]

18. Klarin D, Emdin CA, Natarajan P, Conrad MF, Consortium I. & Kathiresan S. Genetic analysis of venous thromboembolism in UK biobank identifies the ZFPM2 locus and implicates obesity as a causal risk factor. Circ. Cardiovasc. Genet 10, e001643 (2017).

19. Tang W. et al. A genome-wide association study for venous thromboembolism: the extended cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium. Genet. Epidemiol 37, 512–521 (2013). [PubMed: 23650146]

20. Beckmann JS, Estivill X. & Antonarakis SE Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. Nat. Rev. Genet 8, 639–646 (2007). [PubMed: 17637735]

21. Fanciulli M. et al. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. Nat. Genet 39, 721–723 (2007). [PubMed: 17529978]

22. Fellermann K. et al. A chromosome 8 gene-c luster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. Am. J. Hum. Genet 79, 439–448 (2006). [PubMed: 16909382]

23. McGrath LM et al. Copy number variation in obsessive-compulsive disorder and tourette syndrome: a cross-disorder study. J. Am. Acad. Child Adolesc. Psychiatry 53, 910–919 (2014). [PubMed: 25062598]

24. Johansson I. & Ingelman-Sundberg M. CNVs of human genes and their implication in pharmacogenetics. Cytogenet. Genome Res 123, 195–204 (2008). [PubMed: 19287156]
25. Colella S. et al. QuantiSNP: an objective Bayes hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. 35, 2013–2025 (2007). [PubMed: 17341461]
26. Lin CF, Naj AC & Wang LS Analyzing copy number variation using SNP array data: protocols for calling CNV and association tests. Curr Protoc Hum Genet 79, Unit 1(27) (2013).
27. Wang K. et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17, 1665–1674 (2007). [PubMed: 17921354]
28. McElroy JP, Nelson MR, Caillier SJ & Oksenberg JR Copy number variation in African Americans. BMC Genet. 10, 15 (2009). [PubMed: 19317893]
29. Wineinger NE et al. Characterization of autosomal copy-n umber variation in African Americans: the HyperGEN study. Eur. J. Hum. Genet 19, 1271–1275 (2011). [PubMed: 21673747]
30. Perera MA et al. Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study. Lancet 382, 790–796 (2013). [PubMed: 23755828]
31. Goecks J, Eberhard C, Too T, Galaxy T, Nekrutenko A. & Taylor J. Web-based visual analysis for high-throughput genomics. BMC Genomics 14, 397 (2013). [PubMed: 23758618]
32. MacDonald JR., Ziman R., Yuen RK., Feuk L. & Scherer SW. The database of genomic variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 42, D986–D992 (2014). [PubMed: 24174537]
33. Yang J, Lee SH, Goddard ME & Visscher PM GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet 88, 76–82 (2011). [PubMed: 21167468]
34. Rosenberg NA, Li LM, Ward R. & Pritchard JK Informativeness of genetic markers for inference of ancestry. Am. J. Hum. Genet 73, 1402–1422 (2003). [PubMed: 14631557]
35. Purcell S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet 81, 559–575 (2007). [PubMed: 17701901]
36. Firth HV et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. Am. J. Hum. Genet 84, 524–533 (2009). [PubMed: 19344873]
37. Karsenty C, Zhao A, Marijon E. & Ladouceur M. Risk of thromboembolic complications in adult congenital heart disease: a literature review. Arch. Cardiovasc. Dis 111, 613–620 (2018). [PubMed: 29859704]
38. Shendre A. et al. Local ancestry and clinical cardiovascular events among African Americans from the atherosclerosis risk in communities study. J. Am. Heart Assoc 6, e004739 (2017).
39. Ide Vries PS et al. A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. Hum. Mol. Genet 25, 358–370 (2016). [PubMed: 26561523]
40. Kattula S, Byrnes JR & Wolberg AS Fibrinogen and fibrin in hemostasis and thrombosis. Arterioscler. Thromb. Vasc. Biol 37, e13–e21 (2017). [PubMed: 28228446]
41. Fung E. et al. Effect of genetic variants, especially CYP2C9 and VKORC1, on the pharmacology of warfarin. Semin. Thromb. Hemost 38, 893–904 (2012). [PubMed: 23041981]
42. Wang Y. et al. Systematic identification of non-coding pharmacogenomic landscape in cancer. Nat. Commun 9, 3192 (2018). [PubMed: 30093685]
43. Shanley DK et al. Pregnancy-specific glycoproteins bind integrin alphaIIbbeta3 and inhibit the platelet-fibrinogen interaction. PLoS One 8, e57491 (2013).
44. Scott SA et al. Copy number variation and warfarin dosing: evaluation of CYP2C9, VKORC1, CYP4F2, GGCX and CALU. Pharmacogenomics 13, 297–307 (2012). [PubMed: 22188360]
45. Morelli VM, Braekkan SK & Hansen JB Role of microRNAs in venous thromboembolism. Int. J. Mol. Sci 21, 2602 (2020). [PubMed: 32283653]
46. Qiu L. & Liu X. Identification of key genes involved in myocardial infarction. Eur. J. Med. Res 24, 22 (2019). [PubMed: 31269974]
47. Cooper GM et al. A copy number variation morbidity map of developmental delay. Nat. Genet 43, 838–846 (2011). [PubMed: 21841781]

48. Daneshjou R. et al. Population-specific single-nucleotide polymorphism confers increased risk of venous thromboembolism in African Americans. Mol Genet Genomic Med 4, 513–520 (2016). [PubMed: 27652279]

49. Nyangiri OA et al. Copy number variation in human genomes from three major ethno-linguistic groups in Africa. BMC Genomics 21, 289 (2020). [PubMed: 32272904]

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

- Studies suggest that combined genetic, demographic, and clinical factors explain < 30% of warfarin dose requirement in populations with African ancestry, and variants strongly associated with venous thromboembolism (VTE) risk in populations with European ancestry are not associated in African Americans. Copy number variants (CNVs) have been implicated in several disease risk and drug response studies.

WHAT QUESTION DID THIS STUDY ADDRESS?

- Are CNVs associated with warfarin dose requirement and VTE risk in African Americans?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

- Our results implicate CNVs in warfarin dose requirement and strongly support copy number variable region 10q26.12del as a novel protective locus for VTE in African Americans.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

- Understanding the biological mechanisms of associated genomic variants contributes to better understanding of drug response and disease predisposition, identification of novel therapeutically relevant targets for drug development, provides potential biomarkers that can be tested in larger clinical cohorts, and may help prioritize patients for therapy and tailor drug regimens.

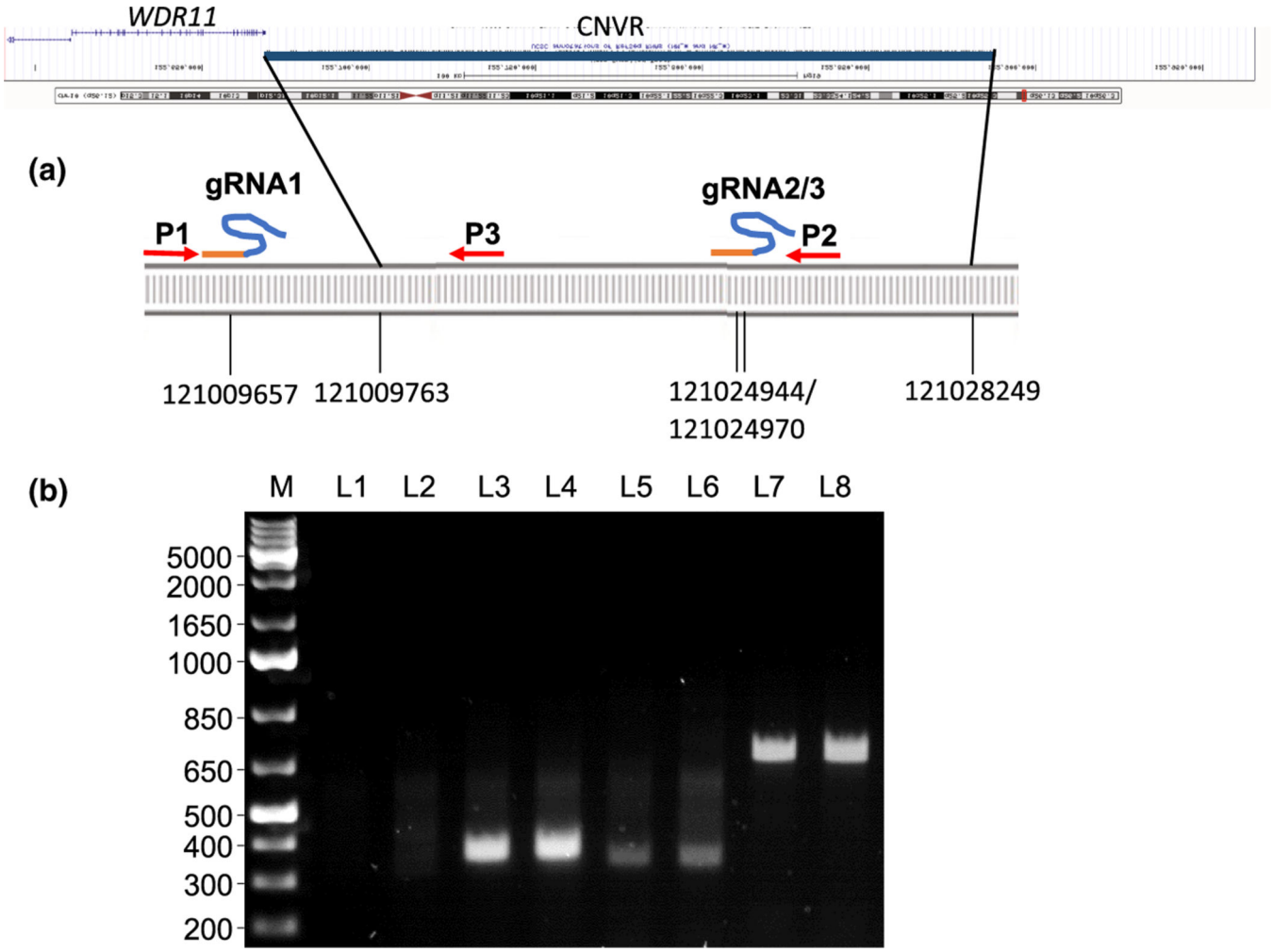

Figure 1.

Validation of 10q26.13 CNVR deletion in HepG2 cells (HepG2-del). (**a**) The diagram showing the genomic location of primers relative to the CNVR. The line in the top panel represents the location of the CNVR just downstream of the gene WDR11. Genomics positions on chromosome 10 are shown for the gRNA guides and CNVR. P1 5′-GTAAG ATG TCT GCC ACC TACC-3′, P2 5′-GTGAC TGT GGT TTT CAG TTGC-3′, P3 5′-GAGTG CCA TCT CAT GTAGG-3′. (**b**) Agarose gel electrophoresis (1.5% agarose) of PCR amplified products. Lane M, 1 kb plus DNA size marker. Lanes 1 and 2 (L1 and L2) showed HepG2 cells without gRNAs. Lanes 3 and 4 (L3 and L4) showed HepG2-del cells transfected by gRNA1 + 2. Lanes 5 and 6 (L5 and L6) showed HepG2-del cells transfected by gRNA1 + 3. Lanes 7 and 8 are HepG2 control cells. Lanes 1–6 are using primer set P1 and P2 flanking the 15,623 base pair regions of the CNVR. The expected PCR product is the gRNA1 + 2 and rRNA 1 + 3 lanes is 335 bp and 305 bp, respectively. Lanes 7 and 8 are using primer set P1 and P3 flanking the 709 base pair regions including the CNVR. CNVR, copy number variable region; PCR, polymerase chain reaction.

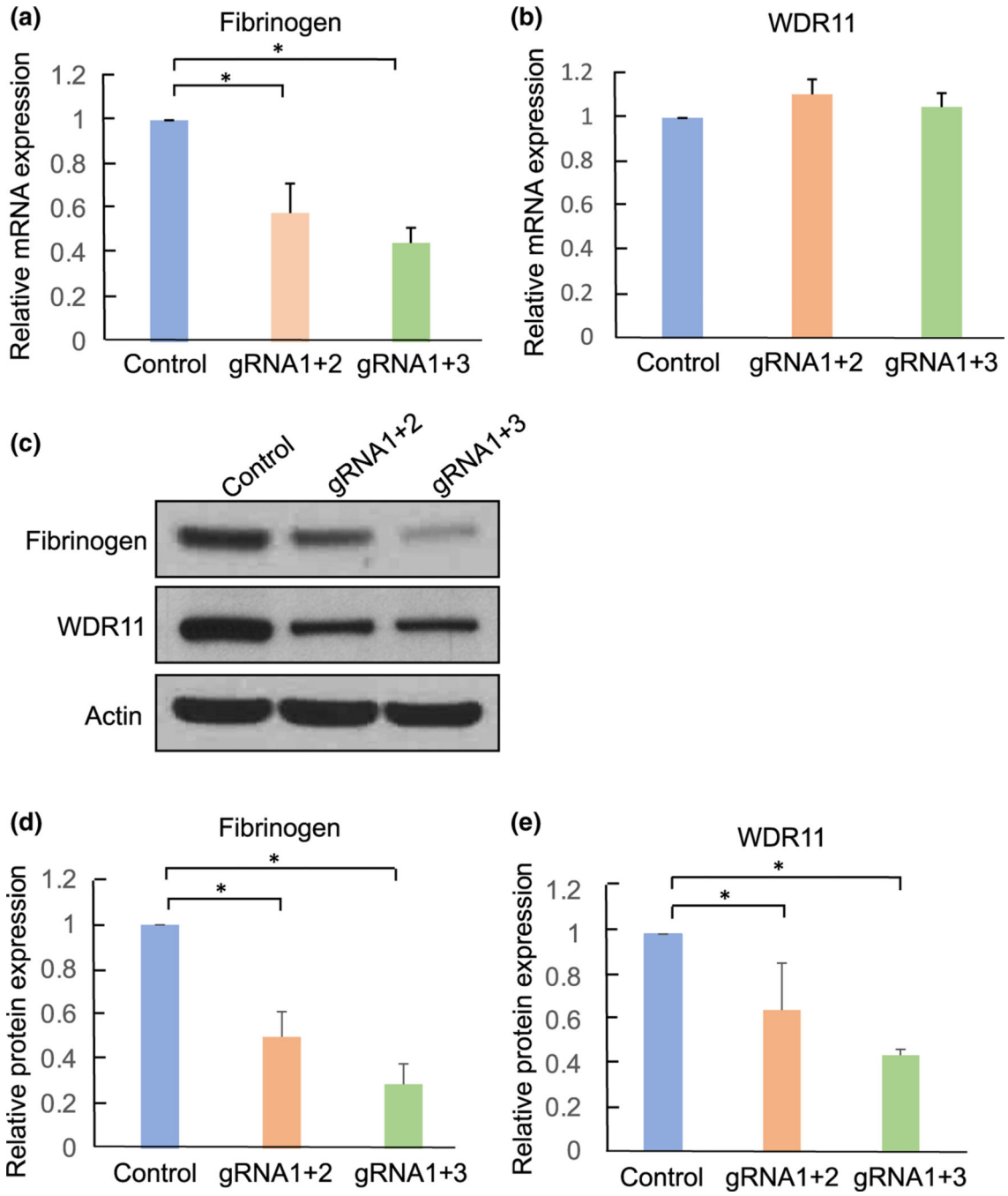

Figure 2.

Effect of 10q26.12del CNVR on fibrinogen and WDR11 mRNA and protein expression levels. Cells were transiently transfected by Cas9 and the gRNA expression plasmids for 72 hours. (**a**) Quantitative RT-PCR analysis showed that fibrinogen mRNA level were significantly decreased in 10q26.12 CNVR deleted HepG2 (HepG2-del) cells compared with HepG2 control cells (**b**) but not WDR11 mRNA level. (**c**) However, immunoblot analysis demonstrated a reduction of fibrinogen and WDR11 protein levels in HepG2-del cells compared with HepG2 control cells. (**d, e**) Protein quantification analysis showed a

statistically significant reduction of fibrinogen and WDR11 protein levels in HepG2-del cells compared with HepG2 control cells. Error bars indicate SEM. Representative images are shown from three biological replicates ($n = 3$). CNVR, copy number variable region; RT-PCR, reverse-transcriptase polymerase chain reaction.

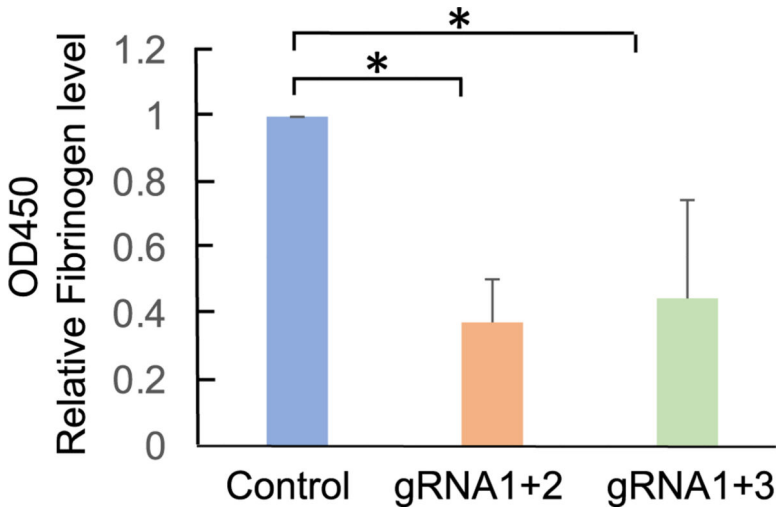

Figure 3.

ELISA analysis of fibrinogen in 10q26.12 CNVR deleted HepG2 cells (HepG2-del). Cells were transiently transfected by Cas9 and the gRNA expression plasmids for 72 hours. Cell culture supernatant was harvested and analyzed using ELISA for fibrinogen quantification. The quantitative analysis showed that fibrinogen production was significantly decreased in HepG2-del cells compared with HepG2 control cells. Error bars indicate SEM. Representative images are shown from three biological replicates ($n = 3$). ELISA, enzyme-linked immunosorbent assay.

Table 1.

Clinical and demographic characteristics of African American warfarin patients.

Variable	N = 327	β (95% CI)	P-value
Dose, mg/week	45.47 ±18.80		
Age	57 ±14.87	−0.45 (−0.58 – 0.32)	<0.001
Height, cm	169.08 ±11.58	0.23 (0.05 – 0.41)	0.01
Weight, kg	94.31 ±27.56	0.16 (0.09 – 0.24)	<0.001
West African Ancestry	0.82 ±0.11	6.26 (−16.48 – 29.00)	0.59
INR	2.45 ±0.35		
Females	62.4%	−0.84 (−5.08 – 0.39)	0.69
Males	37.6%		
Indication for Warfarin			
VTE	47.9%	1.86 (1.10 – 2.63)	<0.001
Atrial Fibrillation	24.8%		
Other	27.3%		

Values for continuous variables are given in mean ± standard deviation

CI, confidence interval; INR, international normalized ratio; VTE, venous thromboembolism

Table 2.

List of most significantly associated CNVRs with warfarin dose requirement in African Americans.

Cytoband	CN state	Validated	Intersecting Gene(s)/RE	CNVR	Size (bps)	Frequency	β (95% CI)	P_{emp}
1q31.2	Deletion (1)	Yes	lncRNA cluster	chr1: 191854738-191906622	51,884	4.2%	11.18 (1.14–21.23)	0.0023
6q14.1	Deletion (1)	No	PHIP, HMGN3, IRAK1BP1	chr6: 78255808-78323761	67,954	23.9%	4.94 (1.72–8.60)	0.0002
19p13.3	Deletion (1)	No	PSG cluster	chr19: 42779968-43335936	555,968	15.8%	−1.41 (−3.23 – −0.70)	0.0004

Adjusted for age, weight, and VTE.

Genomic coordinates in build GRCh38/hg38.

β represent mg/week change per copy of CNV.

CNVR, copy number variable regions; CN, copy number;1 = heterozygous deletion; RE, regulatory element; lncRNA, long non-coding RNA; PHIP, Pleckstrin Homology Domain Interacting Protein; HMGN3, High Mobility Group Nucleosomal Binding Domain 3; IRAK1BP1,(Interleukin 1 Receptor Associated Kinase 1 Binding Protein 1; PSG, Pregnancy Specific.

Table 3.

Association between demographic and clinical characteristics to VTE risk.

Variable	Cases (n=172)	Controls (n=155)	P-value
Age, years (mean ±SD)	54.47 ±16.1	59.54 ±12.9	0.001
Height, cm (mean ±SD)	168.69 ±10.3	170.18 ±10.7	0.29
Weight, kg (mean ±SD)	93.37 ±28.3	91.64 ±26.2	0.23
West African ancestry	0.81 ±0.02	0.80 ±0.01	0.54
Sex			
Female	32.0%	41.2%	0.01
Male	68.0%	58.8%	
VTE Location (Cases)			
DVT	51%		
PE	25%		
DVT/PE	24%		

SD, standard deviation; VTE, venous thromboembolism; DVT, deep vein thrombosis; PE, pulmonary embolism.

Table 4.

List of most significantly associated CNVRs with VTE risk.

Cytoband	CN State	Validated	Intersecting Gene(s)	CNVR	Length (bps)	Frequency	OR (95%CI)	P_{emp}
2p22.1	Deletion (0, 1)	Yes	miRNA cluster	chr2:41002158-41021328	18,170	29.4%	1.88 (1.21–3.10)	0.02
5q35.2	Deletion (1)	No	lncRNA cluster; DUSP1; ERGIC1	chr5:172657047-172887792	230,746	5.8%	14.9 (1.37–94.66)	0.01
10q26.13	Deletion (0, 1)	Yes	miRNA cluster	chr10:121009763-121028249	18,486	18.0%	0.6 (0.59–0.99)	0.05

Adjusted for age and sex.

Genomic coordinates in build GRCh38/hg38.

β represent mg/week change per copy of CNV.

CNVR, copy number variable regions; CN, copy number; 0 = homozygous deletion; 1 = heterozygous deletion; OR, odds ratio; miRNA, microRNA; lncRNA, long non-coding RNA; DUSP1, Dual Specificity Phosphatase 1; ERGIC1, Endoplasmic Reticulum-Golgi Intermediate Compartment 1.