A Platelet Function Modulator of Thrombin Activation Is Causally Linked to Cardiovascular Disease and Affects PAR4 Receptor Signaling

Benjamin A.T. Rodriguez,[1,14] Arunoday Bhan,[2,14] Andrew Beswick,[3] Peter C. Elwood,[4] Teemu J. Niiranen,[5,6] Veikko Salomaa,[5] FinnGen Study, David-Alexandre Trégouët,[7] Pierre-Emmanuel Morange,[8,9,10] Mete Civelek,[11,12] Yoav Ben-Shlomo,[13] Thorsten Schlaeger,[2] Ming-Huei Chen,[1] and Andrew D. Johnson[1,*]

Dual antiplatelet therapy reduces ischemic events in cardiovascular disease, but it increases bleeding risk. Thrombin receptors PAR1 and PAR4 are drug targets, but the role of thrombin in platelet aggregation remains largely unexplored in large populations. We performed a genome-wide association study (GWAS) of platelet aggregation in response to full-length thrombin, followed by clinical association analyses, Mendelian randomization, and functional characterization including iPSC-derived megakaryocyte and platelet experiments. We identified a single sentinel variant in the *GRK5* locus (rs10886430-G, $p = 3.0 \times 10^{-42}$) associated with increased thrombin-induced platelet aggregation ($\beta = 0.70$, SE = 0.05). We show that disruption of platelet *GRK5* expression by rs10886430-G is associated with enhanced platelet reactivity. The proposed mechanism of a GATA1-driven megakaryocyte enhancer is confirmed in allele-specific experiments. Utilizing further data, we demonstrate that the allelic effect is highly platelet- and thrombin-specific and not likely due to effects on thrombin levels. The variant is associated with increased risk of cardiovascular disease outcomes in UK BioBank, most strongly with pulmonary embolism. The variant associates with increased risk of stroke in the MEGASTROKE, UK BioBank, and FinnGen studies. Mendelian randomization analyses in independent samples support a causal role for rs10886430-G in increasing risk for stroke, pulmonary embolism, and venous thromboembolism through its effect on thrombin-induced platelet reactivity. We demonstrate that G protein-coupled receptor kinase 5 (GRK5) promotes platelet activation specifically via PAR4 receptor signaling. GRK5 inhibitors in development for the treatment of heart failure and cancer could have platelet off-target deleterious effects. Common variants in *GRK5* may modify clinical outcomes with PAR4 inhibitors, and upregulation of GRK5 activity or signaling in platelets may have therapeutic benefits.

Introduction

Activated platelets provide the link between inflammation, thrombosis, and atherosclerotic cardiovascular disease.[1] Platelet reactivity is highly heritable,[2,3] but the limited number of previously identified loci explain only a small portion of the estimated heritability.[4] Despite thrombin being the principal enzyme of hemostasis and viewed as the strongest platelet agonist,[5] the genetics of thrombin-induced platelet reactivity is not well understood and heretofore has not been investigated on a genome-wide scale. Many of thrombin's cellular effects are initiated by protease-activated receptors (PARs) which are G protein-coupled receptors (GPCRs).[6] PAR1 (MIM: 187930) and PAR4 (MIM: 602779) are the receptors primarily responsible for mediating the effects of thrombin in human platelets.[5]

Dual antiplatelet therapy (DAPT) reduces the occurrence of both stent-related and spontaneous myocardial infarction (MI) after acute coronary syndrome (ACS), but with concomitant increase in bleeding risk.[7] Thus, there is a need for milder DAPT targets in order to maintain or increase efficacy while reducing bleeding, given the narrow therapeutic window of most antiplatelet treatments.[8] Development of more effective strategies could potentially expand anti-platelet therapy into primary prevention where, due to bleeding risks, it generally is not recommended. Given their role in platelet biology, PAR1 and PAR4 have both been the focus of antithrombotic drug development. Targeting PAR1, the high-affinity thrombin receptor, led to vorapaxar, approved for preventing thrombotic events in patients with MI when used in combination with standard-of-care DAPT.[9] Of limited clinical utility, vorapaxar is associated with increased risk of major bleeding events,[9] where the mortality risk due to bleeding can be comparable to or greater than that due to MI.[10] Compared to targeting PAR1, there is evidence that targeting PAR4 is associated with a lower bleeding risk and can achieve an effective antithrombotic

[1]National Heart, Lung, and Blood Institute, Division of Intramural Research, Population Sciences Branch, The Framingham Heart Study, Framingham, MA 01702, USA; [2]Boston Children's Hospital, Boston, MA 02644, USA; [3]School of Clinical Sciences, University of Bristol, Bristol, BS8 1TH UK; [4]Division of Population Medicine, Cardiff University, Cardiff, CF14 4YS UK; [5]Finnish Institute for Health and Welfare, Helsinki, FI-00271 Finland; [6]Department of Medicine, Turku University Hospital and University of Turku, Turku, 20521 Finland; [7]INSERM UMR_S 1219, Bordeaux Population Health Research Center, University of Bordeaux, 333076 Bordeaux, France; [8]Laboratory of Haematology, La Timone Hospital, 13885 Marseille, France; [9]Centre for Cardiovascular and Nutrition Research, Aix-Marseille Université , INSERM, INRA, 13885 Marseille, France; [10]Centre de Ressources Biologiques Assistance Publique-Hôpitaux de Marseille, HemoVasc, 13885 Marseille, France; [11]Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA; [12]Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908, USA; [13]School of Social and Community Medicine, University of Bristol, Bristol, BS8 1TH UK
[14]These authors contributed equally to this work
*Correspondence: johnsonad2@nhlbi.nih.gov
https://doi.org/10.1016/j.ajhg.2020.06.008.

profile, though large trials of PAR4 inhibition are still lacking.[11–13]

While prospective studies have demonstrated the association of platelet function with cardiovascular disease (CVD) events in patients with established CVD, there is less existing evidence that platelet function predicts CVD or CVD outcomes in the healthy population.[14] Clinical trials have demonstrated a relationship between high on-treatment platelet reactivity and adverse clinical ischemic events, but tailoring therapy based on platelet reactivity remains uncertain.[15] Platelet function traits remain relatively unexplored in large populations, in particular for thrombin and PAR1/PAR4 platelet activation. In order to address this broad knowledge gap—to discover genes that may mediate CVD or bleeding risk, classify treatment subpopulations or suggest new therapeutic targets—we performed the first genome-wide association study (GWAS) of thrombin-induced platelet aggregation.

Subjects and Methods

Participants and Genome-wide Analyses

We conducted and present analyses from the following methodological approaches: (1) a GWAS of platelet thrombin activation, (2) platelet and other cell and tissue expression quantitative trait loci (eQTL) analyses, including Mendelian randomization (MR), for our lead locus, and signal co-localization analyses, (3) two-sample MR for CVD outcomes from multiple consortia and biobanks, (4) integration of megakaryocyte and other epigenetic data at our lead locus, (5) site-directed mutagenesis and regulatory enhancer assays in three cell backgrounds, and (6) induced pluripotent stem cell (iPSC)-derived megakaryocyte, and platelet, small interfering RNA (siRNA) and chemical inhibitor experiments to dissect functional effects on platelets. The samples and resources utilized are described in Table S1. The demographic characteristics of the GWAS sample are shown in Table S2.

The Caerphilly Prospective Study assessed platelet aggregation induced by full-length thrombin (0.056 units/mL, Sigma Aldrich) in middle-aged males through the use of light transmission aggregometry (LTA).[16] All participants provided written informed consent. Genotyping of 1,248 samples was performed with the Affymetrix UK BioBank Axiom array. Following sample and genotyping quality control, imputation was done on 1,184 samples through the use of the Haplotype Reference Consortium (HRC) 1.1 panel. We conducted a GWAS using a linear mixed model adjusting for age and medication usage. A significance threshold of $p < 7 \times 10^{-9}$ was adopted to account for all variants tested. Conditional analyses adjusting for the strongest peak SNP in *GRK5* (MIM: 600870), rs10886430, were conducted by adding SNP dosage as a covariate to the base model. Our methods for platelet aggregation, genotyping, quality control procedures, imputation, and GWAS are further described in the Supplemental Information.

Causal Analysis and Multi-trait Colocalization for GRK5 Locus

Testing for causal association with eQTL was conducted with the use of the platelet RNA and expression 1 (PRAX1) platelet eQTL dataset.[17] We investigated whether thrombin-induced aggregation and platelet cell traits share a common association signal at the *GRK5* locus by performing a co-localization analysis with quantitative traits derived from multiple blood cell lineages: platelets (mean platelet volume [MPV], platelet count [PLT], and platelet distribution width [PDW]), red blood cell count (RBC), and white blood cell count (WBC).[18] Furthermore, we conducted similar genetic colocalization analyses for the *GRK5* locus through the use of genome-wide study data for thrombin generation potential;[19] platelet aggregation to ADP, collagen, and epinephrine;[4,20] eQTLs from 44 cells and tissues from the Genotype-Tissue Expression (GTEx) Project version 7; five white blood cell types and platelets from the CEDAR project;[21] and aortic endothelial cells.[22] Two-sample MR analysis was conducted using the rs10886430-G allele as the genetic instrument and thrombin-induced platelet aggregation as exposure in separate analyses for nine pulmonary, stroke, or heart disease outcomes from the UK BioBank,[23] four stroke outcomes from the MEGASTROKE consortium,[24] and CVD codes in the FinnGen Biobank (version 4).[25] Only this SNP was used in MR analyses because in conditional analyses of chromosome 10, this was the only independently significant SNP associated with thrombin reactivity or platelet *GRK5* expression levels. Details regarding these analyses are given in the Supplemental Information.

Regulatory Function

We integrated epigenetic regulatory maps of chromatin accessibility, enhancer RNA (eRNA), histone marks, enhancer elements, and DNA-binding proteins assessed in megakaryocyte-erythroid lineage models (ENCODE)[26,27] and cultured primary megakaryocytes (BLUEPRINT) to annotate potential functional impacts of the intronic *GRK5* rs10886430 variant.[28,29] Protein network analysis of transcriptional regulators that bind the rs10886430 variant was performed with STRING 10.5. We used site-directed mutagenesis to investigate the impact of the rs10886430 variant on enhancer activity *in vitro* in *GATA1*- (MIM: 305371) and *GATA2*- (MIM: 137295) overexpressing HEK293 cells, as well as HUVEC and K562 cells. Details regarding the data integration, network analysis, enhancer reporter luciferase assays, and conditional overexpression of *GATA1* and *GATA2* are provided in the Supplemental Information.

Platelet Function siRNA and Inhibitor Experiments

The role of G protein-coupled receptor kinase 5 (GRK5) in platelet function was assessed through the use of siRNA as well as pharmacologic inhibition of G protein-coupled receptor kinase (GRK) activity. Immortalized megakaryocyte progenitor cell lines (imMKCLs) were generated from human induced pluripotent stem cells and maintained as previously described.[30] For siRNA experiments, imMKCLs were transfected with *GRK5* or control EGFP siRNA for 48 h, then analyzed using qRT-PCR and *in vitro* flow cytometric analysis of platelet function (via P-selectin and PAC-1) following stimulation with either 20 μM ADP plus 20 μM thrombin receptor-activating peptide 6 (TRAP-6 that activates platelets via PAR1) or control. Pharmacologic inhibition of GRK activity was investigated with the pan-GRK inhibitor CCG215022 (MedChemExpress) which exhibits nanomolar IC_{50} values against both GRK2 as well as GRK5 and selectivity against PKA.[31] Platelet-rich plasma (PRP) samples obtained from healthy donors ($n = 3$) were treated with 0.78 μM CCG215022 or DMSO vehicle for 45 min prior to stimulation with increasing concentrations of one of three platelet agonists: PAR4 Activating Peptide

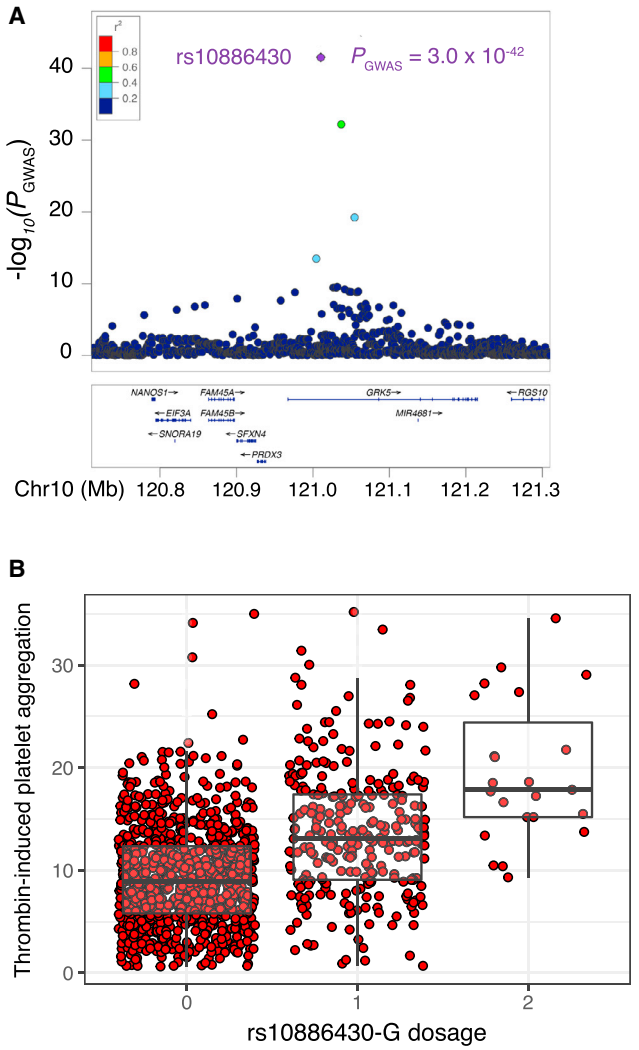

Figure 1. Regional and SNP Associations of *GRK5* with Thrombin-Induced Platelet Aggregation

(A) Locus Zoom plot of the lead SNP (rs10886430) from genome-wide association analysis of thrombin-induced platelet aggregation in Caerphilly Prospective study (n = 1,184). The plot depicts the 600-kb window flanking the rs10886430 (purple) variant which is located in the first intron of *GRK5*.

(B) Boxplot of the effect of rs10886430-G variant dosage on thrombin-induced platelet aggregation (0.056 U/mL). Data points are plotted as red circles. The bold horizontal line represents the median. The upper and lower hinges indicate the 25th and 75th percentiles, respectively. The whiskers extending from the hinges represent the values no further than 1.5 * interquartile range.

(PAR4-AP, 1, 20, 50 μM), TRAP-6 (1, 10, 20 μM), ADP (1, 10, 20 μM) or vehicle control followed by flow cytometric analysis of platelet function (via P-selectin and PAC-1). Details regarding the siRNA experiments in imMKCL, GRK inhibition in PRP samples, and platelet function assessment by flow cytometry are provided in the Supplemental Information.

Results

We performed a GWAS of thrombin-induced platelet aggregation in the Caerphilly Prospective Study including >7.75 million common and low-frequency (minor allele frequency [MAF] > 0.01) SNPs imputed via the HRC panel. There was no evidence for inflation of test statistics (λ = 1.005) (Figure S1A). We observed 17 variants that surpassed the genome-wide significance threshold ($p < 7 \times 10^{-9}$) for association with thrombin-induced aggregation, all localized to 10q26.11 (Figure 1A). Conditional analysis identified no additional signals independent of the sentinel variant located in the first intron of the *GRK5* locus (*GRK5*, rs10886430, $p = 3.0 \times 10^{-42}$) ~43-kb downstream of the transcription start site (TSS) within consensus intron 1 (Figure S1B, Table S3). The minor G allele (MAF 0.136) of the *GRK5* SNP was associated with increased platelet reactivity to thrombin ($\beta = 0.70$, SE = 0.05; with other covariates fixed, this is ~3.9% per allele increase in thrombin reactivity) (Figure 1B). Variance component analysis indicates that this single variant explained 18.3% of variation in the thrombin phenotype. There was no significant population structure in Caerphilly, which consists overwhelmingly of participants of European ancestry, based on principal components clustering with multi-ethnic samples in the UK BioBank (Figure S2). Likewise, we found little relatedness in Caerphilly: only n = 303 had any first-, second-, or third-degree relations. Inclusion of the genetic relatedness matrix in the GWAS accounted for this. Nonetheless, we conducted a sensitivity analysis removing the n = 303 individuals and analyzing chromosome 10, and rs10886430 remained highly significant ($p = 2.01 \times 10^{-31}$), indicating that the results are population associations rather than strong family effects.

We next asked whether thrombin-induced aggregation and platelet, red cell, and white cell count traits share a common association signal at the *GRK5* locus. We performed a Bayesian test for co-localization between the Caerphilly thrombin GWAS and multiple blood cell lineage traits from the UK BioBank/INTERVAL study meta-analysis,[18] interrogating shared variants in an ~1.8 Mb independent linkage disequilibrium block containing the *GRK5* lead SNP. We observed strong evidence for co-localization (posterior probability >0.99) between thrombin reactivity and platelet cell traits (MPV, PLT, and PDW) but not WBC or RBC; this result supports the hypothesis that a single variant affects these traits in platelets (Figure 2A). The *GRK5* lead SNP (rs10886430) was identified as the shared, potentially causal variant in each case (Table S4). With the association limited to platelets, we next asked whether the rs10886430 variant affects platelet reactivity mediated by other agonists. We conducted a GWAS on aggregation to ADP and collagen in Caerphilly participants, and we performed colocalization analyses with the thrombin GWAS. No evidence of colocalization was observed between thrombin reactivity and either agonist (Table S5). Further, no colocalization was observed with aggregation to ADP, collagen, or epinephrine in the largest such GWAS published to date in independent cohorts[4] (Table S6). We next asked whether the effect on

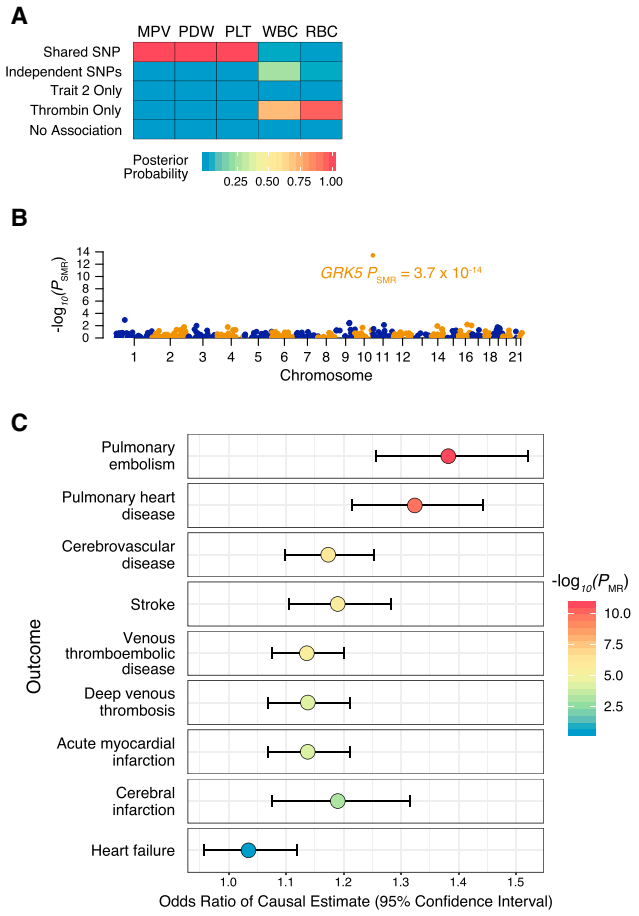

Figure 2. *GRK5* rs10886430-G Is Linked to Platelet Traits and Causally Effects Platelet *GRK5* Expression and Multiple CVD Outcomes

The rs10886430 variant regulates platelet cell traits and *GRK5* platelet gene expression, as well as both cardiovascular and cerebrovascular disease pathologies.

(A) Heatmap of posterior probabilities from Bayesian colocalization analyses of thrombin reactivity and five blood cell traits in the 1.8 Mb LD block containing the rs10886430 variant (10q26.11). Shared SNP—probability of one shared SNP associated with both traits; Independent SNPs—probability of two independent SNPs associated with each trait; Trait 2 Only—probability of association with the blood cell trait and not with thrombin-induced aggregation; Thrombin Only—probability of association with thrombin-induced aggregation and not with the blood cell trait; No Association—probability of no association with either trait.

(B) Manhattan plot depicting summary data-based Mendelian Randomization (SMR) analysis of association between platelet gene expression and platelet reactivity to thrombin (0.056 U/mL).

(C) Mendelian Randomization analysis of thrombin reactivity (rs10886430G instrument) and cardiopulmonary phenotypes in UK BioBank (Outcomes). Plotted are the OR of the causal estimates (circles) and associated 95% confidence intervals (error bars), color of circle indicates -log transformed p value of estimate. Further description of medical outcomes codes and statistics is found in the Supplemental Information and Table S12.

thrombin reactivity was limited to platelet activation or was also observed in thrombin generation potential traits, as this could indicate an effect mediated via thrombin levels. No colocalization was observed between platelet reactivity to thrombin and three phenotypic markers of thrombin generation[19] (Table S7).

To determine potential regulatory impacts of the rs10886430 variant, we first examined its association with functional expression changes cataloged in the platelet-specific PRAX1 study eQTL dataset.[17,32] We utilized summary-data-based Mendelian Randomization (SMR) analysis with the PRAX1 eQTL dataset to test for association between platelet gene expression and platelet reactivity to thrombin. We identified *GRK5* as the only gene at a genome-wide significance level ($\beta_{SMR} = -1.54$, $SE_{SMR} = 0.20$, $P_{SMR} = 3.67 \times 10^{-14}$) (Figure 2B, Table S8). Finding no evidence to suggest that the SMR association could be due to genetic linkage ($P_{HEIDI} = 0.16$), we concluded that expression of *GRK5* is associated with platelet reactivity driven by the rs10886430 variant. As predicted by our SMR analysis, the minor rs10886430-G allele is a strong *cis*-eQTL for decreased platelet *GRK5* expression ($\beta_{eQTL} = -0.456$, $P_{eQTL} = 8.27 \times 10^{-20}$) in the PRAX1 study.[17] We further replicated this strong *GRK5* eQTL in the independent CEDAR platelet dataset[21] ($\beta_{eQTL} = -0.429$, $P_{eQTL} = 1.11 \times 10^{-20}$). To investigate the tissue specificity of the genetic effect of rs10886430-G on *GRK5* expression, we conducted co-localization analyses between the Caerphilly thrombin GWAS in the *GRK5* locus and 51 tissue or cell types. We observed no evidence of co-localization between thrombin reactivity and *GRK5* expression among 44 tissues profiled by the GTEx Project (Table S9), five other white blood cell types[21] (Table S10), or vascular endothelial cells[22] (Table S11).

Having observed the strong association of rs10886430-G with lower *GRK5* expression exclusively in platelets, we asked whether the genetic effect of the variant on platelet reactivity was further associated with relevant cardiovascular or cerebrovascular disease pathology. We utilized GWAS statistics for nine pulmonary, stroke, or heart disease diagnoses in the UK BioBank cohort. There was evidence for an association of rs10886430-G with several diagnoses, most strongly with pulmonary embolism (PE) (odds ratio [OR] = 1.25, SE = 1.03, p = 8.1×10^{-13}) (Table S12). We performed two-sample MR to test for a causal relationship between thrombin reactivity (exposure) and disease outcome for the nine UK BioBank diagnoses. We observed strong, positive association between thrombin-induced platelet reactivity at the *GRK5* locus and multiple disease outcomes, most significantly with diagnosis of PE (OR_{MR} = 1.38 [1.26 – 1.52], $P_{MR} = 2.40 \times 10^{-11}$), also with diagnoses of cerebral infarction (OR_{MR} = 1.19 [1.08 – 1.32], $P_{MR} = 7.43 \times 10^{-4}$) and acute MI ($OR_{MR}$ = 1.14 [1.07 – 1.21], $P_{MR} = 6.67 \times 10^{-5}$), but not with heart failure (OR_{MR} = 1.04 [0.96 – 1.12], $P_{MR} = 0.39$) (Figure 2C, Table S12). Next, we investigated the rs10886430-G variant in stroke subtypes from MEGASTROKE,[24] the cohorts of which are independent of UK BioBank. There was evidence for a stronger association in cardioembolic stroke (OR = 18.42, SE = 2.96, p = 6.16×10^{-4}) than ischemic stroke (OR = 1.61, SE = 1.15, p = 2.29×10^{-4}), and no association for

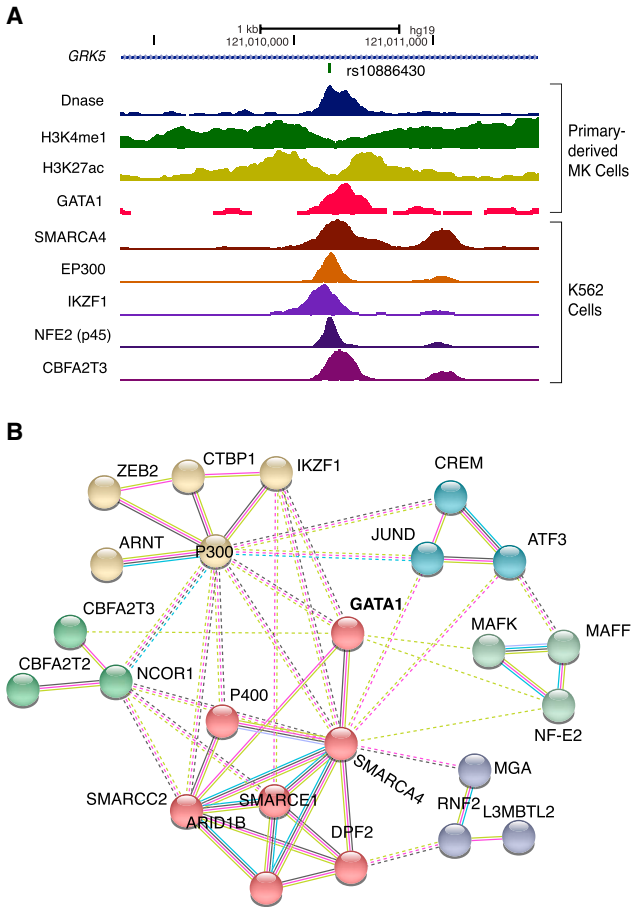

Figure 3. *GRK5* rs10886430 Overlaps an Active Meta-erythroid Lineage Enhancer Bound by Interconnected Factors

The rs10886430 variant resides in a megakaryocyte enhancer element.

(A) Epigenetic regulatory maps of primary-derived MK and K562 cells.

(B) STRING Protein Network Analysis of DNA binding factors recruited to the rs10886430 variant in ENCODE mega-erythroid cell models. Proteins are represented as circles. Colors indicate network cluster membership. Solid lines indicate interactions within a network cluster. Dotted lines indicate interactions between proteins in different clusters. Line colors indicate type of evidence: cyan—known interaction from curated database; magenta—known interaction experimentally determined; yellow-green—text mining; black—co-expression.

large-artery stroke (OR = 4.32, SE = 5.09, p = 0.39) (Table S13). In MR analyses, we observed strong, positive association between thrombin-induced platelet reactivity at the rs10886430-G variant and cardioembolic stroke (OR_{MR} = 64.36 [3.09 – 1340.44], $P_{MR} = 7.18 \times 10^{-3}$), all stroke ($OR_{MR}$ = 1.79 [1.19 – 1.27], $P_{MR} = 7.69 \times 10^{-4}$), and ischemic stroke ($OR_{MR}$ = 1.97 [1.22 – 1.33], $P_{MR} = 7.17 \times 10^{-4}$), but not large artery stroke (OR_{MR} = 10.24 [0.08 – 773.88], P_{MR} = 0.37) (Table S13). Finally, the G allele was again associated with increased risk of multiple CVD outcomes in the FinnGen study—deep venous thrombosis (DVT) of lower extremities (OR_{MR} = 1.25 [1.11– 1.41], $P_{MR} = 2.28 \times 10^{-4}$), ischemic stroke ($OR_{MR}$ = 1.15 [1.06 – 1.25], $P_{MR} = 5.75 \times 10^{-4}$), portal vein thrombosis (OR_{MR} = 2.51 [1.30-4.85], $P_{MR} = 6.31 \times 10^{-3}$), right bundle branch block (OR_{MR} = 1.68 [1.14 – 2.48], $P_{MR} = 8.51 \times 10^{-3}$), MI ($OR_{MR}$ = 1.12 [1.03 – 1.22], $P_{MR} = 1.09 \times 10^{-2}$), atrioventricular block (OR_{MR} = 1.24 [1.04– 1.47], $P_{MR} = 1.64 \times 10^{-2}$)—and with reduced risk for hypertension (OR_{MR} = 0.94 [0.89 – 0.98], $P_{MR} = 7.49 \times 10^{-3}$) and cardiomyopathies ($OR_{MR}$ = 0.79 [0.68 – 0.91], $P_{MR} = 1.85 \times 10^{-3}$) (Table S14).

A GRK5 gain-of-function coding variant (p.Gln41Leu) leading to enhanced β-adrenergic receptor (βAR) desensitization of excessive catecholamine signaling has been proposed to provide a "genetic β-blockade" that improves survival in African Americans with heart failure.[33] Functional studies have also shown β2-ARs can inhibit platelet aggregation and adhesion.[34] In our study, this variant was not associated with thrombin reactivity (p = 0.51, data not shown). Having observed associations with multiple pathologic cardiopulmonary traits, we asked whether the *GRK5* variant's effect on platelet reactivity was modified by pharmacological β-blockade. A small subset of Caerphilly participants were taking cardiac-specific, β1-AR selective (5.1%) or non-selective drugs (3.1%) at the time platelet reactivity was measured. Fitting linear models with specific and non-specific beta-blockers as additional covariates, we observed a negative interaction effect on platelet reactivity between the *GRK5* variant and β1-AR selective drugs (β = −0.60, ^{SE} = 0.24, p = 0.01) (Figure S3).

To study the regulatory function of the *GRK5* variant, we next integrated cell-type-specific epigenome maps derived from primary megakaryocyte cells.[28,29] The variant localizes to a region of open chromatin (DNase hypersensitivity peak site) in a predicted enhancer region characterized by broad enrichment of active marks H3K27Ac and H3K4me1 (Figure 3A). Active enhancers are often characterized by short, unstable bi-directional transcripts termed "enhancer RNAs (eRNAs)." Integrating nascent transcription maps in K562 cells, we observed that the *GRK5* variant localizes to the predicted TSS region of an eRNA[27] (Figure S4). Having observed multiple lines of evidence that the variant is in an enhancer element, we asked whether the position was occupied by transcriptional regulators *in vivo*, as this could provide a clear testable mechanism of action in a non-coding DNA context. To this end, we scanned transcription factor (TF) binding datasets from mega-erythroid cell models generated by the ENCODE consortium.[26] The variant position was bound by 27 factors, including the master hematopoietic TF GATA1 in peripheral blood erythroblasts and the histone acetyltransferase p300 (MIM: 602700) in K562 cells (Table S15).

We verified that binding of the *GRK5* variant by GATA1 is also observed in primary megakaryocyte (MK) cells[35] and in hematopoietic stem and/or progenitor-derived erythroid precursors.[36] The GATA1 binding motif is highly enriched at active eRNA in K562 cells.[37] We next considered whether there was functional connectivity or association between the factors binding the *GRK5* variant locus. Using the STRING algorithm, we constructed a highly

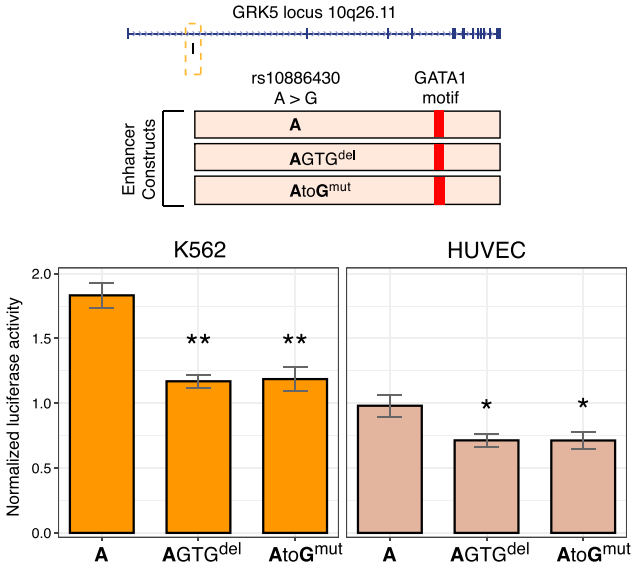

Figure 4. Mutagenesis Causing Deletion or rs10886430 A to G Transition Disrupts Enhancer Activity in Multiple Cell Backgrounds

The rs10886430 variant modulates megakaryocyte enhancer activity. Top, diagram of GRK5 enhancer constructs: row 1, WT allele; row 2, four-base deletion including allele position; row 3, substitution with "G" effect allele. Bottom, normalized luciferase activity in mega-erythroid K562 cells and endothelial HUVEC. Results are the mean of three independent experiments performed in quadruplicate (n total = 12), error bars represent SEM. *$p < 1 \times 10^{-4}$, **$p < 5 \times 10^{-5}$.

connected network model incorporating 24 of 26 mappable transcriptional regulators (protein-protein interaction enrichment p value: $< 1.0 \times 10^{-16}$) (Figure 3B). Clustering of the network revealed several properties: SWI/SNF chromatin remodeling complex clustering with GATA1 (red cluster), transcriptional repressors (including one cluster of polycomb-related proteins) (green cluster), cAMP responsive factors (blue cluster), p300 and lineage factor IKZF1 (MIM: 603023) (khaki cluster), and NF-E2 (MIM: 601490) complex (light green cluster). These included p45 and MAFK (MIM: 600197), subunits of the heterodimeric NF-E2 complex required for megakaryocyte maturation and platelet production *in vivo*.[38] Collectively, these results suggest that the *GRK5* non-coding variant affects platelet reactivity by modulating a functional megakaryocyte lineage enhancer, leading to platelet populations with altered *GRK5* expression and function.

Using luciferase reporter assay, we confirmed that the locus drives enhancer activity *in vitro* through binding of GATA1 and not through binding of GATA2 (Figure S5A–S5C). Enhancer activity in endogenous GATA1-expressing cells is largely abolished upon deletion of the GATA1 core binding motif ($GATA1^{del}$; Figure S5B). Given its close proximity to the GATA1 motif, we hypothesized that the rs10886430-G could reduce the regulatory potential of the enhancer by interfering with binding of GATA1. Thus, we investigated the impact of the GRK5 variant on enhancer activity, both by targeted four-base deletion as well as by single-base substitution of the minor "G" allele (Figure 4). In K562 cells, introduction of the GRK5 variant effect allele ($AtoG^{mut}$) repressed enhancer activity 1.5-fold ($p < 0.001$). Targeted deletion of the base position ($AGTG^{del}$) produced nearly identical results. Endothelial HUVEC cells, which express lower levels of *GATA1* (confirmed using qRT-PCR, data not shown), exhibited weak enhancer activity and a diminished capacity for the GRK5 variant effect allele to repress enhancer activity (1.3-fold, $p < 0.05$) (Figure 4).

Having established a potential mechanism by which the DNA variant regulates *GRK5* expression, we next investigated the role of GRK5 in platelet physiology. To this end, we utilized iPSC-derived imMKCLs which produce functional platelets expressing cell markers CD42b (MIM: 606672) and VWF (MIM: 613160).[30] We first assessed the impact of reducing *GRK5* expression in imMKCL through the use of siRNA testing, achieving a knockdown efficiency of ~75% (Figure S6A). We performed platelet function testing of GRK5-depleted and control platelet progeny cells stimulated with agonists ADP/TRAP-6 via flow cytometry, measuring platelet surface activated GPIIb-IIIa (PAC1 antibody) and platelet surface P-Selectin (CD62P). Transient knockdown of GRK5 increased the percentages of platelets that were positive for P-Selectin 1.6-fold ($p < 0.05$) and activated GPIIb-IIIa 1.2-fold ($p < 0.05$) (Figure 5A). The amounts of P-Selectin and activated GPIIb-IIIa exposed on the surface of each platelet, as judged by the geometric mean fluorescence intensity (MFI), were largely unchanged (Figure 5A).

We next investigated the effect of repressing GRK activity on platelet function in the specific context of PAR1- or PAR4-activated signaling *ex vivo* via CCG215022, a pan-specific small-molecule inhibitor exhibiting nanomolar IC_{50} values against both GRK2 (MIM: 109635) and GRK5. To this end, we performed platelet function testing on treated and control PRP stimulated with either TRAP-6 (PAR1 activator) or PAR4-AP via flow cytometry. Treatment with the GRK inhibitor increased the percentages of platelets positive for P-Selectin up to 2.3-fold in the presence of PAR4-AP but not in the presence of TRAP-6 (Figure S6B). Given that we observed activation in siRNA experiments upon ADP/TRAP-6 co-stimulation, the absence of an effect by PAR1 signaling alone in healthy donor PRP treated with GRK inhibitor led us to consider whether stimulation with ADP alone would have an effect. We performed additional platelet function testing with a range of concentrations of TRAP-6, PAR4-AP, or ADP following GRK inhibition (Figure 5B, Figure S7). Significant increases in P-Selectin-positive cells were again observed for PAR4-AP (20uM: 2.7-fold, $p < 0.01$) and to a lesser extent for ADP (1 uM: 1.2-fold, $p < 0.05$) but not TRAP-6 (Figure 5B, Figure S7). Because TRAP-6 activation may have been saturated at doses between 1 uM to 20 uM, we ran independent experiments at lower doses of TRAP-6 in donor PRP (0.001 uM, 0.01 uM, and 0.10 uM), and again we found no effect of GRK inhibition on platelet activation (Figure S8). Together,

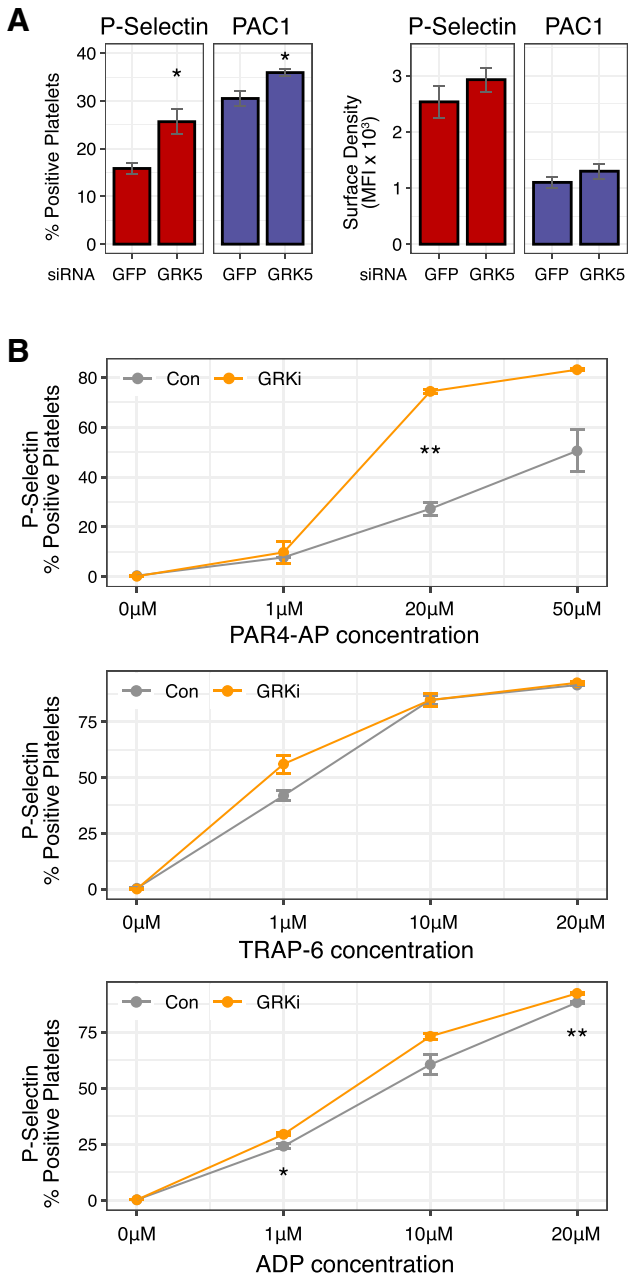

Figure 5. Disrupting Platelet GRK5 via siRNA or Chemical Inhibitor Causes Increased Thrombin Activation via a PAR4 Receptor Signaling Process

GRK5 perturbation promotes platelet activation.

(A) Change in platelet activation markers P-Selectin and activated GPIIb-IIIa (PAC1) in response to 20 μM ADP/TRAP6 stimulation (compared to untreated) following siRNA knockdown of *GRK5* expression (or Green Fluorescent Protein [GFP] siRNA negative control) in imMKCL cells. Data represent mean ± SEM of four independent experiments. Differences assessed via Student's T-Test. * $p < 0.05$, ** $p < 0.01$.

(B) Change in P-Selectin in response to specific activation of either ADP, PAR1 (TRAP-6), or PAR4 (PAR4-AP) signaling (compared to untreated) following pharmacologic inhibition of GRK activity with 0.78 μM CCG215022 (GRKi) or vehicle control (Con) in PRP samples from healthy donors. Plotted data represent mean ± SEM (n = 3) difference in percentages of marker-positive platelets following treatment with the indicated concentrations of agonist. Differences between PRP treated with GRK inhibitor versus control assessed via Student's T-Test. *$p < 0.05$, **$p < 0.01$.

these results suggest that inhibition of platelet GRK5 promotes PAR4-mediated platelet activation and to a lesser extent ADP-mediated activation, but not PAR1-mediated platelet activation.

Discussion

In a GWAS of thrombin-induced platelet aggregation, we identify a *GRK5* non-coding variant (rs10886430-G) strongly associated with increased reactivity to thrombin. We observed that thrombin-induced aggregation shares a common association signal at the *GRK5* locus with three platelet cell indices (MPV, PLT, and PDW), the rs10886430 SNP being the shared causal variant in each case. Interestingly, the SNP had previously been identified as a sentinel variant positively associated with two indices: MPV and PDW,[18] which are suggested by some as partial surrogates for platelet activation. Our study suggests a direct role for GRK5 in platelet activation, with an overall mechanism outlined in Figure 6.

To determine the regulatory potential of the rs10886430 variant, we first applied SMR to test for association between platelet gene expression and reactivity to thrombin. The analysis supported a model whereby decreased *GRK5* expression is associated with increased platelet reactivity through disruption by the rs10886430-G variant. The strong rs10886430-G eQTL for *GRK5* replicated in two independent platelet datasets[17,21] but none of the other 44 tissues in GTEx, five white blood cell subtypes, or aortic endothelial cells; these results indicate significant platelet specificity. Utilizing megakaryocyte and erythroid epigenetic datasets, we found evidence for an active cell lineage enhancer at the SNP site bound by master hematopoietic TF GATA1. GATA1 plays a critical role in megakaryocyte maturation and platelet formation *in vivo*.[39] We supported the disruption of *GRK5* expression through multiple experiments. We further investigated the role of GRK5 in platelet function through the use of an iPSC-derived megakaryocyte and platelet production model. Knockdown of GRK5 by siRNA increased functional markers of platelet activation in agonist-stimulated cells, including α-granule release (surface P-Selectin) and inside-out activation of the platelet integrin complex GPIIb-IIIa. These results indicate that GRK5 plays an important functional role in negatively regulating thrombin-induced platelet reactivity.

Thrombin is the most potent endogenous platelet activator, and it plays an important role in clot promotion and inhibition, and cell signaling, as well as additional processes that influence fibrinolysis and inflammation. The role of GRK5 could be mediated through several mechanisms: first, through canonical GRK GPCR desensitization of signal initiators PAR1 or PAR4,[6] and second, GPCR desensitization of numerous downstream effectors of PAR signaling such as β2-AR (MIM: 107941),[34,40] Akt (MIM: 164730), or the SRC family kinases.[41] To further determine whether GRK5 affects PAR1 or PAR4 signaling, we

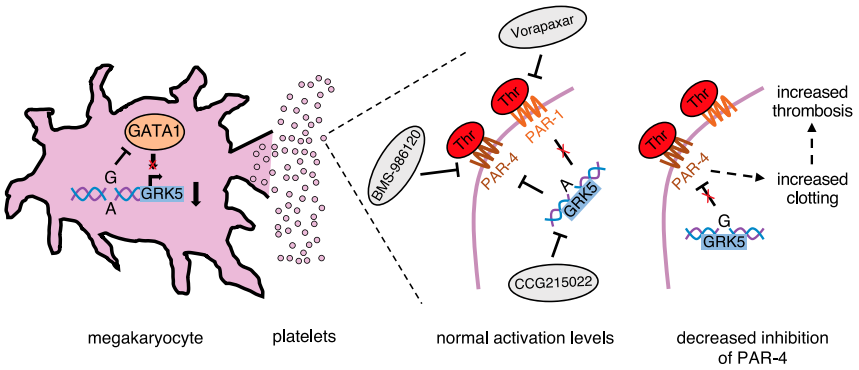

Figure 6. Suggested Mechanism of *GRK5* Regulatory Variant in Influencing PAR4 Platelet Activation and Thrombosis Risk
Schematic depicting the platelet-specific effects of *GRK5* variant rs10886430 via a GATA1 enhancer and modified suppression of platelet PAR4 signaling on thrombin activation. PAR4 has been studied as a drug target (BMS-986120). The model is supported by thrombin platelet reactivity association in PRP, a strong platelet eQTL in two independent samples, a lack of eQTL co-localization in other cells and tissues, mutagenesis and enhancer assays, siRNA and inhibitor (CCG215022) experiments in iPSC-derived megakaryocytes and platelets, and multiple population genetic studies for CVD outcomes (UK BioBank, MEGASTROKE, FinnGen, INVENT, MVP).

conducted experiments that showed GRK5 inhibitory effects are mediated via inhibition of PAR4-driven platelet activation.

PAR4 is involved in sustained platelet activation, and it invokes sustained intracellular calcium response in platelets, phosphyltidylserine exposure, thrombin generation, and fibrin deposition.[13] Thus, PAR4 has been suggested as a novel anti-platelet therapeutic target, with primate models and other studies indicating that PAR4 inhibition could provide superior inhibition with reduced bleeding diatheses.[11,12] Our results suggest that a significant genetically determinable fraction of the population could potentially receive greater benefit from PAR4 inhibition. Interestingly, a missense variant, p.Ala120Thr, in PAR4 has been described as affecting activation.[42] However, this variant was not associated with full thrombin reactivity in our study ($p = 0.07$, data not shown), and responses to a monoclonal antibody directed at PAR4 did not vary by genotype in another study.[43] Taken together, this suggests that *GRK5* variant rs10886430 may have stronger implications for thrombin-based platelet activation than does the previously reported *PAR4* coding variant. Our study does have some limitations. Given the modest sample size, false negatives are expected. Because there are no other reported thrombin, PAR1, or PAR4 platelet reactivity GWASs, we cannot yet assess replication of those findings.

Given the importance of thrombin in clot formation, we used two-sample MR to determine whether the genetic effect of the variant on platelet reactivity may play a putatively causal role in cardiovascular or cerebrovascular disease pathology. Whereas the role of platelets in arterial thrombosis is well established, our MR analysis utilizing the strong GRK5 instrumental variable also suggests that thrombin-driven platelet reactivity contributes to the trajectory of venous thromboembolism (VTE), both in DVT and in PE. Recently, the variant was independently tied to increased VTE risk in the INVENT Consortium and this was replicated in the Million Veteran's Project.[44] Critical for DVT propagation *in vivo*, platelets are recruited to developing venous thrombi where they support leukocyte accumulation and promote formation of procoagulant neutrophil extracellular traps.[45] Markers of platelet activation are elevated in acute PE, correlate with the severity of right ventricular dysfunction, and can persist for several months.[46] The contributing role of platelets in VTE is further supported by the observation that aspirin therapy reduces the risk of DVT and PE in patients undergoing orthopedic surgery.[47] Thrombosis and excess platelet activation are common pathological features of pulmonary arterial hypertension,[48] another cause of pulmonary heart disease. While the etiologic heterogeneity characterizing ischemic stroke makes it difficult to assign a causative role for platelet reactivity to any given subtype, the platelet content of embolized thrombi is twice that of *in situ* thrombi,[49] suggesting that the platelet aggregate increases propensity for embolization. The importance of platelets in pathogenesis of acute MI is supported by both clinical and *in vivo* animal studies which show that the initial thrombus (following endothelial injury) is primarily composed of activated platelets.[50]

In our investigation of the rs10886430-G variant among stroke subtypes available in MEGASTROKE, we observed a markedly strong effect in the prevalence of cardioembolic stroke, though not of large vessel or small vessel disease. Compared to cardioembolic stroke, the large artery atherosclerosis and small vessel occlusion subtypes have vastly different etiologies.[24] MR analysis supported a causal role for thrombin-driven platelet reactivity specifically in the cardioembolic subtype. Causal associations with PE and cardioembolic stroke suggest that the rs10886430-G variant is particularly enriched in emboli forming distally to the site of vascular occlusion. More broadly, our results underline the importance of thrombin-driven platelet reactivity in both venous and arterial disease. It remains to be seen whether rs10886430-G is an important variant to segment populations at risk relative to treatment for either venous or arterial disease.

Beta-blockers are a common preventative therapy following MI and a mainstay for the management of heart failure.[33] Functional studies have shown β2-ARs to inhibit

platelet aggregation and adhesion through activation of platelet nitric oxide synthase.[34] Given the role of GRK5 in desensitizing β-AR signaling, we investigated whether the *GRK5* variant's effect on platelet reactivity was modified by beta-blocker therapy in a subset of Caerphilly participants. We observed a negative interaction effect on platelet reactivity between the *GRK5* variant and β1-AR selective drugs, and we saw no effect with non-selective drugs. The absence of an appreciable effect with the latter may in part be explained by low sample sizes in the model (selective n = 56, non-selective n = 34) as well as by differences in chemistry which may affect platelet uptake.[51] Future work could potentially examine the genotypic effect of the *GRK5* variant on multiple drug classes in larger samples including direct thrombin inhibitors, beta-adrenergic blockers, and other anticoagulant and anti-platelet therapies. Overexpression of cardiac *GRK5* leads to early heart failure after pressure overload in mouse models.[52] Also upregulated in human heart failure, GRK5 is being investigated as a therapeutic target with selective small-molecule inhibitors under development.[53] Our work highlights the potential for significant platelet-driven off-target effects with this or other strategies seeking to inhibit GRK5. Notably, in platelet RNA-sequencing data, *GRK5* is by far the most expressed member of the GRK family. The next highest-expressed GRK family members are *GRK6* (MIM: 600869) (~18% expression level of *GRK5*) and *GRK4* (MIM: 137026) (~2% expression level of *GRK5*), suggesting that GRK5 is likely to be the critical protein family member active in platelets.[54] Given the role of GRK5 in controlling PAR4-mediated platelet activation and the association of the rs10886430-G genetic effect on platelet reactivity with cardiovascular and cerebrovascular embolic events, we suggest that finding a mechanism to maintain GRK5 activity in platelets could prove beneficial in preventing venous and arterial CVD.

Data and Code Availability

Caerphilly genotypes and phenotypes are available by request to the study steward at the University of Bristol: https://www.bristol.ac.uk/population-health-sciences/projects/caerphilly/. Other code and data are either freely available at the websites listed in the Web Resources or in the Supplemental Information, or by request to the corresponding author.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2020.06.008.

Acknowledgments

This work was supported by National Heart, Lung, and Blood Institute (NHLBI) Intramural Research Program funding (to B.A.T., M-H.C., and A.D.J.). Additional support came from the National Blood Foundation/American Association of Blood Banks (FP01021164), the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; U54DK110805) and the National Research Service Award (NRSA)'s Joint Program in Transfusion Medicine (T32 4T32HL066987-15 to Ar.B.). D.A.T. was supported by the «EPIDEMIOM-VTE» Senior Chair from the Initiative of Excellence of the University of Bordeaux. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the NHLBI, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Declarations

The authors declare no competing interests.

Received: January 24, 2020
Accepted: June 3, 2020
Published: July 9, 2020

Web Resources

BLUEPRINT, http://dcc.blueprint-epigenome.eu
COLOC, https://cran.r-project.org/web/packages/coloc/index.html
EMMAX, https://genome.sph.umich.edu/wiki/EMMAX
ENCODE, https://www.encodeproject.org/summary/?type=Experiment
Gene ATLAS, http://geneatlas.roslin.ed.ac.uk/
Genotype-Tissue Expression (GTEx) Project, https://www.gtexportal.org/home/
GRASP Full GWAS Results (for Astle et al.), https://grasp.nhlbi.nih.gov/downloads/FullResults/2016/2016_Astle/AstleREADME.txt
Haplotype Reference Consortium, http://www.haplotype-reference-consortium.org/
LD Blocks, http://bitbucket.org/nygcresearch/ldetect-data/src
Michigan Imputation Server, https://imputationserver.sph.umich.edu
Online Mendelian Inheritance in Man (OMIM), https://omim.org/
Summary-data-based Mendelian Randomization (SMR), https://cnsgenomics.com/software/smr/#Overview
STRING, https://string-db.org/

References

1. Fuentes Q, E., Fuentes Q, F., Andrés, V., Pello, O.M., Font de Mora, J., and Palomo G, I. (2013). Role of platelets as mediators that link inflammation and thrombosis in atherosclerosis. Platelets *24*, 255–262.
2. O'Donnell, C.J., Larson, M.G., Feng, D., Sutherland, P.A., Lindpaintner, K., Myers, R.H., D'Agostino, R.A., Levy, D., Tofler, G.H.; and Framingham Heart Study (2001). Genetic and environmental contributions to platelet aggregation: the Framingham heart study. Circulation *103*, 3051–3056.
3. Faraday, N., Yanek, L.R., Mathias, R., Herrera-Galeano, J.E., Vaidya, D., Moy, T.F., Fallin, M.D., Wilson, A.F., Bray, P.F., Becker, L.C., and Becker, D.M. (2007). Heritability of platelet responsiveness to aspirin in activation pathways directly and indirectly related to cyclooxygenase-1. Circulation *115*, 2490–2496.
4. Johnson, A.D., Yanek, L.R., Chen, M.-H., Faraday, N., Larson, M.G., Tofler, G., Lin, S.J., Kraja, A.T., Province, M.A., Yang,

Q., et al. (2010). Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat. Genet. *42*, 608–613.
5. Crawley, J.T.B., Zanardelli, S., Chion, C.K.N.K., and Lane, D.A. (2007). The central role of thrombin in hemostasis. J. Thromb. Haemost. *5* (*Suppl 1*), 95–101.
6. Tiruppathi, C., Yan, W., Sandoval, R., Naqvi, T., Pronin, A.N., Benovic, J.L., and Malik, A.B. (2000). G protein-coupled receptor kinase-5 regulates thrombin-activated signaling in endothelial cells. Proc. Natl. Acad. Sci. USA *97*, 7440–7445.
7. Wallentin, L., Becker, R.C., Budaj, A., Cannon, C.P., Emanuelsson, H., Held, C., Horrow, J., Husted, S., James, S., Katus, H., et al.; PLATO Investigators (2009). Ticagrelor versus clopidogrel in patients with acute coronary syndromes. N. Engl. J. Med. *361*, 1045–1057.
8. Scaglione, F. (2013). New oral anticoagulants: comparative pharmacology with vitamin K antagonists. Clin. Pharmacokinet. *52*, 69–82.
9. Tricoci, P., Huang, Z., Held, C., Moliterno, D.J., Armstrong, P.W., Van de Werf, F., White, H.D., Aylward, P.E., Wallentin, L., Chen, E., et al.; TRACER Investigators (2012). Thrombin-receptor antagonist vorapaxar in acute coronary syndromes. N. Engl. J. Med. *366*, 20–33.
10. Valgimigli, M., Costa, F., Lokhnygina, Y., Clare, R.M., Wallentin, L., Moliterno, D.J., Armstrong, P.W., White, H.D., Held, C., Aylward, P.E., et al. (2017). Trade-off of myocardial infarction vs. bleeding types on mortality after acute coronary syndrome: lessons from the Thrombin Receptor Antagonist for Clinical Event Reduction in Acute Coronary Syndrome (TRACER) randomized trial. Eur. Heart J. *38*, 804–810.
11. Wong, P.C., Seiffert, D., Bird, J.E., Watson, C.A., Bostwick, J.S., Giancarli, M., Allegretto, N., Hua, J., Harden, D., Guay, J., et al. (2017). Blockade of protease-activated receptor-4 (PAR4) provides robust antithrombotic activity with low bleeding. Science Translational Medicine *9*, eaaf5294.
12. Wilson, S.J., Ismat, F.A., Wang, Z., Cerra, M., Narayan, H., Raftis, J., Gray, T.J., Connell, S., Garonzik, S., Ma, X., et al. (2018). PAR4 (Protease-Activated Receptor 4) Antagonism With BMS-986120 Inhibits Human Ex Vivo Thrombus Formation. Arterioscler. Thromb. Vasc. Biol. *38*, 448–456.
13. French, S.L., Arthur, J.F., Lee, H., Nesbitt, W.S., Andrews, R.K., Gardiner, E.E., and Hamilton, J.R. (2016). Inhibition of protease-activated receptor 4 impairs platelet procoagulant activity during thrombus formation in human blood. J. Thromb. Haemost. *14*, 1642–1654.
14. Puurunen, M.K., Hwang, S.J., Larson, M.G., Vasan, R.S., O'Donnell, C.J., Tofler, G., and Johnson, A.D. (2018). ADP Platelet Hyperreactivity Predicts Cardiovascular Disease in the FHS (Framingham Heart Study). J. Am. Heart Assoc. *7*, e008522.
15. Sibbing, D., Aradi, D., Alexopoulos, D., Ten Berg, J., Bhatt, D.L., Bonello, L., Collet, J.P., Cuisset, T., Franchi, F., Gross, L., et al. (2019). Updated Expert Consensus Statement on Platelet Function and Genetic Testing for Guiding $P2Y_{12}$ Receptor Inhibitor Treatment in Percutaneous Coronary Intervention. JACC Cardiovasc. Interv. *12*, 1521–1537.
16. Elwood, P.C., Renaud, S., Sharp, D.S., Beswick, A.D., O'Brien, J.R., and Yarnell, J.W. (1991). Ischemic heart disease and platelet aggregation. The Caerphilly Collaborative Heart Disease Study. Circulation *83*, 38–44.
17. Simon, L.M., Chen, E.S., Edelstein, L.C., Kong, X., Bhatlekar, S., Rigoutsos, I., Bray, P.F., and Shaw, C.A. (2016). Integrative Multi-omic Analysis of Human Platelet eQTLs Reveals Alternative Start Site in Mitofusin 2. Am. J. Hum. Genet. *98*, 883–897.
18. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell *167*, 1415–1429.e19.
19. Rocanin-Arjo, A., Cohen, W., Carcaillon, L., Frère, C., Saut, N., Letenneur, L., Alhenc-Gelas, M., Dupuy, A.-M., Bertrand, M., Alessi, M.-C., et al.; CardioGenics Consortium (2014). A meta-analysis of genome-wide association studies identifies ORM1 as a novel gene controlling thrombin generation potential. Blood *123*, 777–785.
20. Eicher, J.D., Xue, L., Ben-Shlomo, Y., Beswick, A.D., and Johnson, A.D. (2016). Replication and hematological characterization of human platelet reactivity genetic associations in men from the Caerphilly Prospective Study (CaPS). J. Thromb. Thrombolysis *41*, 343–350.
21. Momozawa, Y., Dmitrieva, J., Théâtre, E., Deffontaine, V., Rahmouni, S., Charloteaux, B., Crins, F., Docampo, E., Elansary, M., Gori, A.-S., et al.; International IBD Genetics Consortium (2018). IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. Nat. Commun. *9*, 2427.
22. Erbilgin, A., Civelek, M., Romanoski, C.E., Pan, C., Hagopian, R., Berliner, J.A., and Lusis, A.J. (2013). Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. J. Lipid Res. *54*, 1894–1905.
23. Canela-Xandri, O., Rawlik, K., and Tenesa, A. (2018). An atlas of genetic associations in UK Biobank. Nat. Genet. *50*, 1593–1599.
24. Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.-K., van der Laan, S.W., Gretarsdottir, S., et al.; AFGen Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; International Genomics of Blood Pressure (iGEN-BP) Consortium; INVENT Consortium; STARNET; BioBank Japan Cooperative Hospital Group; COMPASS Consortium; EPIC-CVD Consortium; EPIC-InterAct Consortium; International Stroke Genetics Consortium (ISGC); METASTROKE Consortium; Neurology Working Group of the CHARGE Consortium; NINDS Stroke Genetics Network (SiGN); UK Young Lacunar DNA Study; and MEGASTROKE Consortium (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat. Genet. *50*, 524–537.
25. Mars, N., Koskela, J.T., Ripatti, P., Kiiskinen, T.T.J., Havulinna, A.S., Lindbohm, J.V., Ahola-Olli, A., Kurki, M., Karjalainen, J., Palta, P., et al.; FinnGen (2020). Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat. Med. *26*, 549–557.
26. Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Res. *22*, 1798–1812.
27. Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nat. Genet. *46*, 1311–1320.

28. Stunnenberg, H.G., Hirst, M.; and International Human Epigenome Consortium (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell *167*, 1145–1149.
29. Petersen, R., Lambourne, J.J., Javierre, B.M., Grassi, L., Kreuzhuber, R., Ruklisa, D., Rosa, I.M., Tomé, A.R., Elding, H., van Geffen, J.P., et al. (2017). Platelet function is modified by common sequence variation in megakaryocyte super enhancers. Nat. Commun. *8*, 16058.
30. Nakamura, S., Takayama, N., Hirata, S., Seo, H., Endo, H., Ochi, K., Fujita, K., Koike, T., Harimoto, K., Dohda, T., et al. (2014). Expandable megakaryocyte cell lines enable clinically applicable generation of platelets from human induced pluripotent stem cells. Cell Stem Cell *14*, 535–548.
31. Homan, K.T., Waldschmidt, H.V., Glukhova, A., Cannavo, A., Song, J., Cheung, J.Y., Koch, W.J., Larsen, S.D., and Tesmer, J.J.G. (2015). Crystal Structure of G Protein-coupled Receptor Kinase 5 in Complex with a Rationally Designed Inhibitor. J. Biol. Chem. *290*, 20649–20659.
32. Simon, L.M., Edelstein, L.C., Nagalla, S., Woodley, A.B., Chen, E.S., Kong, X., Ma, L., Fortina, P., Kunapuli, S., Holinstat, M., et al. (2014). Human platelet microRNA-mRNA networks associated with age and gender revealed by integrated plateletomics. Blood *123*, e37–e45.
33. Liggett, S.B., Cresci, S., Kelly, R.J., Syed, F.M., Matkovich, S.J., Hahn, H.S., Diwan, A., Martini, J.S., Sparks, L., Parekh, R.R., et al. (2008). A GRK5 polymorphism that inhibits β-adrenergic receptor signaling is protective in heart failure. Nat. Med. *14*, 510–517.
34. Queen, L.R., Xu, B., Horinouchi, K., Fisher, I., and Ferro, A. (2000). beta(2)-adrenoceptors activate nitric oxide synthase in human platelets. Circ. Res. *87*, 39–44.
35. Tijssen, M.R., Cvejic, A., Joshi, A., Hannah, R.L., Ferreira, R., Forrai, A., Bellissimo, D.C., Oram, S.H., Smethurst, P.A., Wilson, N.K., et al. (2011). Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. Dev. Cell *20*, 597–609.
36. Pinello, L., Xu, J., Orkin, S.H., and Yuan, G.-C. (2014). Analysis of chromatin-state plasticity identifies cell-type-specific regulators of H3K27me3 patterns. Proc. Natl. Acad. Sci. USA *111*, E344–E353.
37. Azofeifa, J.G., Allen, M.A., Hendrix, J.R., Read, T., Rubin, J.D., and Dowell, R.D. (2018). Enhancer RNA profiling predicts transcription factor activity. Genome Res. *28*, 334–344.
38. Shivdasani, R.A., Rosenblatt, M.F., Zucker-Franklin, D., Jackson, C.W., Hunt, P., Saris, C.J.M., and Orkin, S.H. (1995). Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. Cell *81*, 695–704.
39. Vyas, P., Ault, K., Jackson, C.W., Orkin, S.H., and Shivdasani, R.A. (1999). Consequences of GATA-1 deficiency in megakaryocytes and platelets. Blood *93*, 2867–2875.
40. Tran, T.M., Jorgensen, R., and Clark, R.B. (2007). Phosphorylation of the β2-adrenergic receptor in plasma membranes by intrinsic GRK5. Biochemistry *46*, 14438–14449.
41. Kim, S., Jin, J., and Kunapuli, S.P. (2006). Relative contribution of G-protein-coupled pathways to protease-activated receptor-mediated Akt phosphorylation in platelets. Blood *107*, 947–954.
42. Edelstein, L.C., Simon, L.M., Lindsay, C.R., Kong, X., Teruel-Montoya, R., Tourdot, B.E., Chen, E.S., Ma, L., Coughlin, S., Nieman, M., et al. (2014). Common variants in the human platelet PAR4 thrombin receptor alter platelet function and differ by race. Blood *124*, 3450–3458.
43. French, S.L., Thalmann, C., Bray, P.F., Macdonald, L.E., Murphy, A.J., Sleeman, M.W., and Hamilton, J.R. (2018). A function-blocking PAR4 antibody is markedly antithrombotic in the face of a hyperreactive PAR4 variant. Blood Adv. *2*, 1283–1293.
44. Lindstrom, S., Wang, L., Smith, E.N., Gordon, W., van Hylckama Vlieg, A., de Andrade, M., Brody, J.A., Pattee, J.W., Haessler, J., Brumpton, B.M., et al. (2019). Genomic and Transcriptomic Association Studies Identify 16 Novel Susceptibility Loci for Venous Thromboembolism. Blood, blood.2019000435.
45. von Brühl, M.-L., Stark, K., Steinhart, A., Chandraratne, S., Konrad, I., Lorenz, M., Khandoga, A., Tirniceriu, A., Coletti, R., Köllnberger, M., et al. (2012). Monocytes, neutrophils, and platelets cooperate to initiate and propagate venous thrombosis in mice in vivo. J. Exp. Med. *209*, 819–835.
46. Chung, T., Connor, D., Joseph, J., Emmett, L., Mansberg, R., Peters, M., Ma, D., and Kritharides, L. (2007). Platelet activation in acute pulmonary embolism. J. Thromb. Haemost. *5*, 918–924.
47. 2000). Prevention of pulmonary embolism and deep vein thrombosis with low dose aspirin: Pulmonary Embolism Prevention (PEP) trial. Lancet 355, 1295–1302.
48. Lannan, K.L., Phipps, R.P., and White, R.J. (2014). Thrombosis, platelets, microparticles and PAH: more than a clot. Drug Discov. Today *19*, 1230–1235.
49. Wysokinski, W.E., Owen, W.G., Fass, D.N., Patrzalek, D.D., Murphy, L., and McBane, R.D., 2nd. (2004). Atrial fibrillation and thrombosis: immunohistochemical differences between in situ and embolized thrombi. J. Thromb. Haemost. *2*, 1637–1644.
50. Silvain, J., Collet, J.-P., Nagaswami, C., Beygui, F., Edmondson, K.E., Bellemain-Appaix, A., Cayla, G., Pena, A., Brugier, D., Barthelemy, O., et al. (2011). Composition of coronary thrombus in acute myocardial infarction. J. Am. Coll. Cardiol. *57*, 1359–1367.
51. Walle, T., Webb, J.G., Walle, U.K., and Bagwell, E.E. (1991). Stereoselective accumulation of the β-receptor blocking drug atenolol by human platelets. Chirality *3*, 451–453.
52. Martini, J.S., Raake, P., Vinge, L.E., DeGeorge, B.R., Jr., Chuprun, J.K., Harris, D.M., Gao, E., Eckhart, A.D., Pitcher, J.A., and Koch, W.J. (2008). Uncovering G protein-coupled receptor kinase-5 as a histone deacetylase kinase in the nucleus of cardiomyocytes. Proc. Natl. Acad. Sci. USA *105*, 12457–12462.
53. Traynham, C.J., Hullmann, J., and Koch, W.J. (2016). "Canonical and non-canonical actions of GRK5 in the heart". J. Mol. Cell. Cardiol. *92*, 196–202.
54. Eicher, J.D., Wakabayashi, Y., Vitseva, O., Esa, N., Yang, Y., Zhu, J., Freedman, J.E., McManus, D.D., and Johnson, A.D. (2016). Characterization of the platelet transcriptome by RNA sequencing in patients with acute myocardial infarction. Platelets *27*, 230–239.

The American Journal of Human Genetics, Volume *107*

Supplemental Data

A Platelet Function Modulator of Thrombin Activation Is Causally Linked to Cardiovascular Disease and Affects PAR4 Receptor Signaling

Benjamin A.T. Rodriguez, Arunoday Bhan, Andrew Beswick, Peter C. Elwood, Teemu J. Niiranen, Veikko Salomaa, FinnGen Study, David-Alexandre Trégouët, Pierre-Emmanuel Morange, Mete Civelek, Yoav Ben-Shlomo, Thorsten Schlaeger, Ming-Huei Chen, and Andrew D. Johnson

Supplementary Figures and Legends

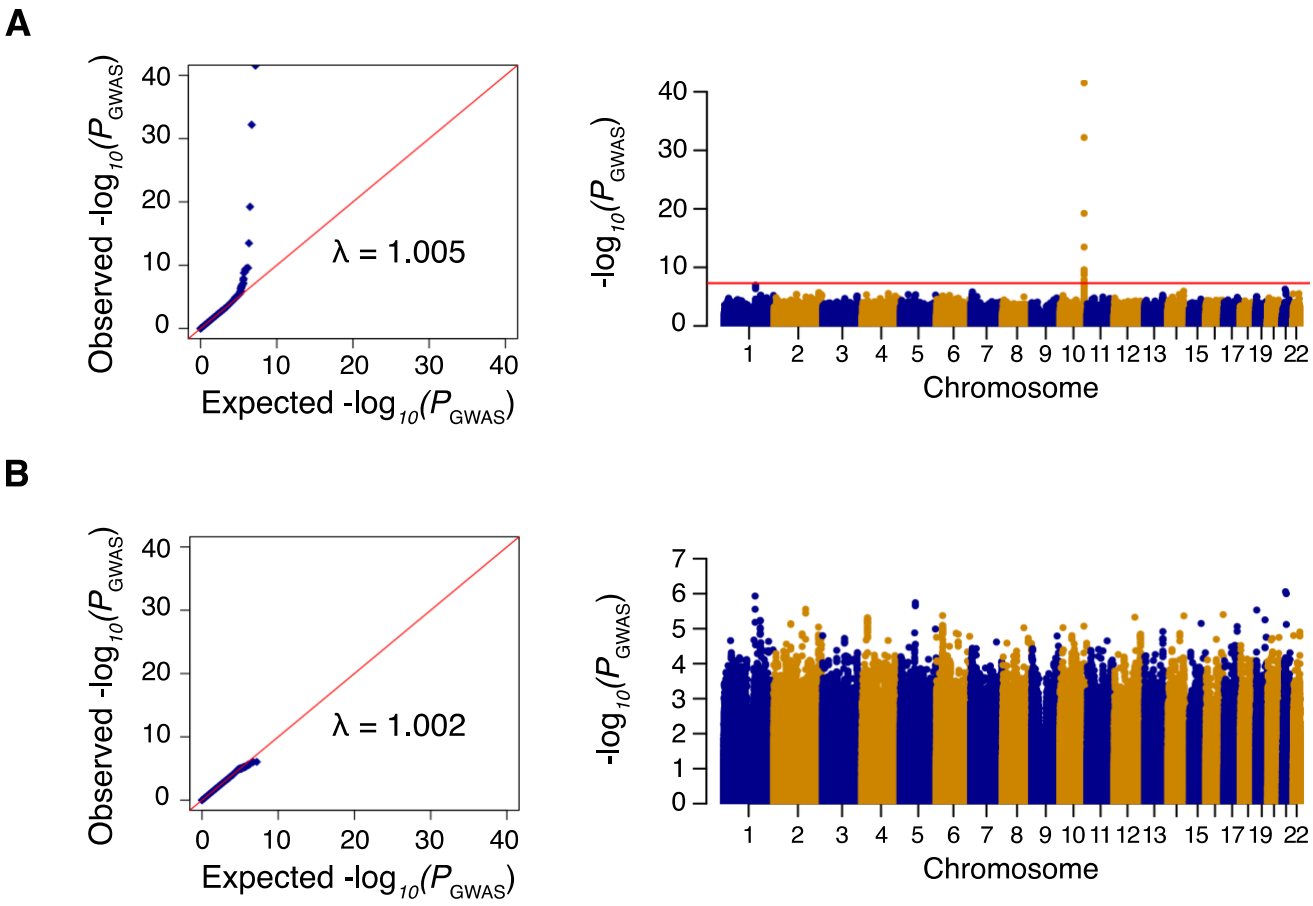

Figure S1. Manhattan and Quantile-Quantile (QQ) plots from genome-wide association analysis of Thrombin-induced platelet aggregation in the Caerphilly Prospective Study. A, initial analysis of Caerphilly participants (n = 1184). B, analysis conditioning on GRK5 peak SNP rs10886430. On Manhattan plots, negative log10-transformed P-values are plotted according to physical genomic locations by chromosome. The red horizontal line indicates P-value threshold of 7×10^{-9}, corresponding to genome-wide significance.

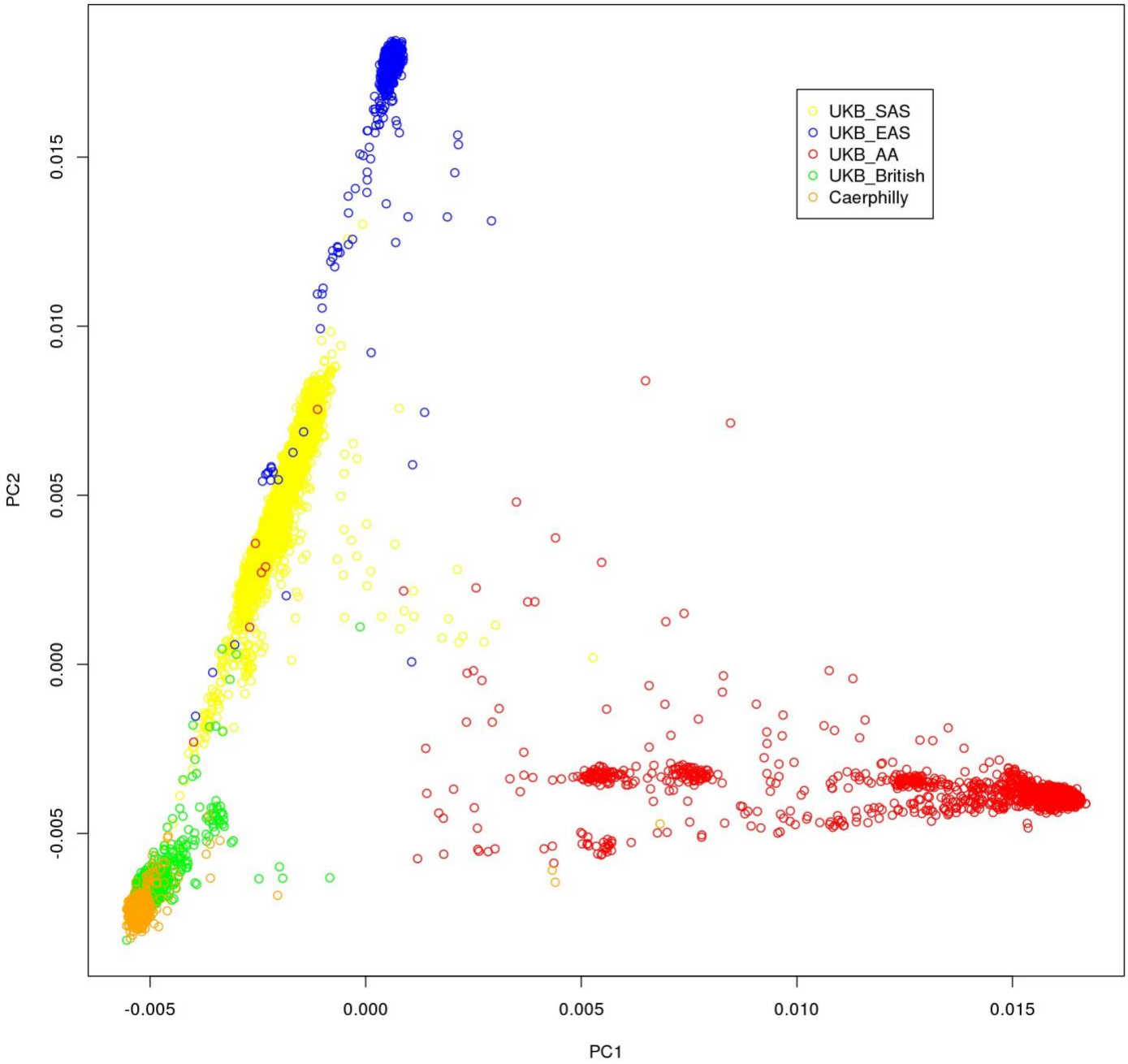

Figure S2. Principal components analysis of Caerphilly with UK BioBank samples reveals that the Caerphilly sample is overwhelmingly European ancestry. Caerphilly samples are indicated by orange circles and UK BioBank British/European ancestry samples by green circles.

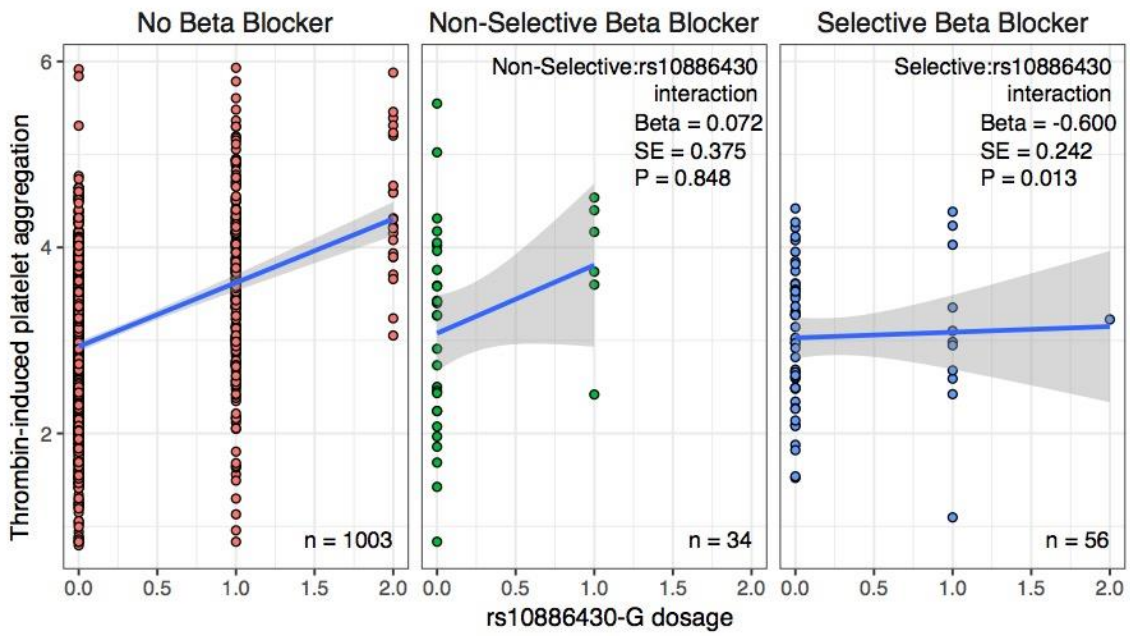

Figure S3. Scatterplot of the effect of rs10886430-G variant dosage on thrombin-induced platelet aggregation in Caerphilly participants not taking Beta Blockers (left), taking Non-Selective Beta Blockers (middle), and taking Selective Beta Blockers (right) with linear fit and 95% confidence interval. Platelet aggregation values were square root transformed prior to modeling and plotting. Beta, standard error (SE), and P value for interaction effects with rs10886430-G dosage are given for Non-Selective (B) and Selective (C) Beta Blockers.

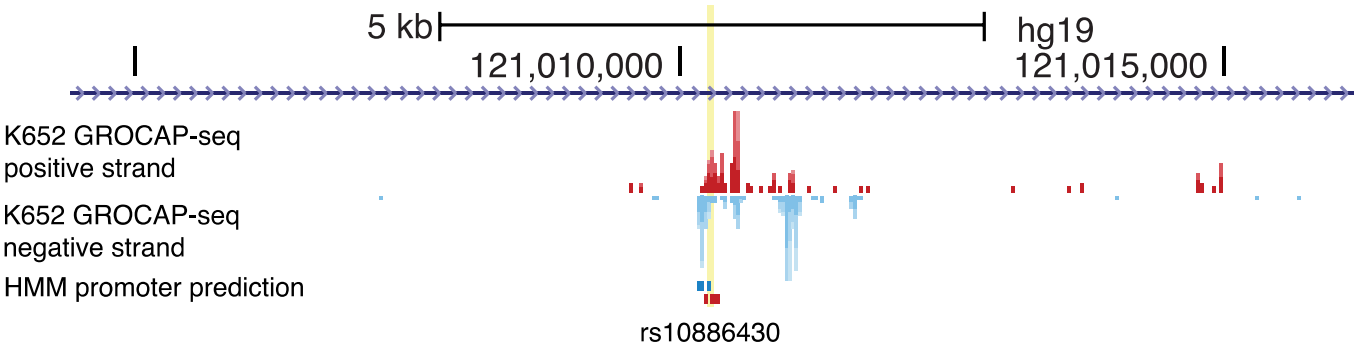

Figure S4. **Genome browser plot of K562 cell GROCAP-seq (+ strand red, - strand blue) signal density and Hidden Markov Model bidirectional promoter prediction.** The yellow line indicates the position of the rs10886430 variant in the bidirectional eRNA promoter.

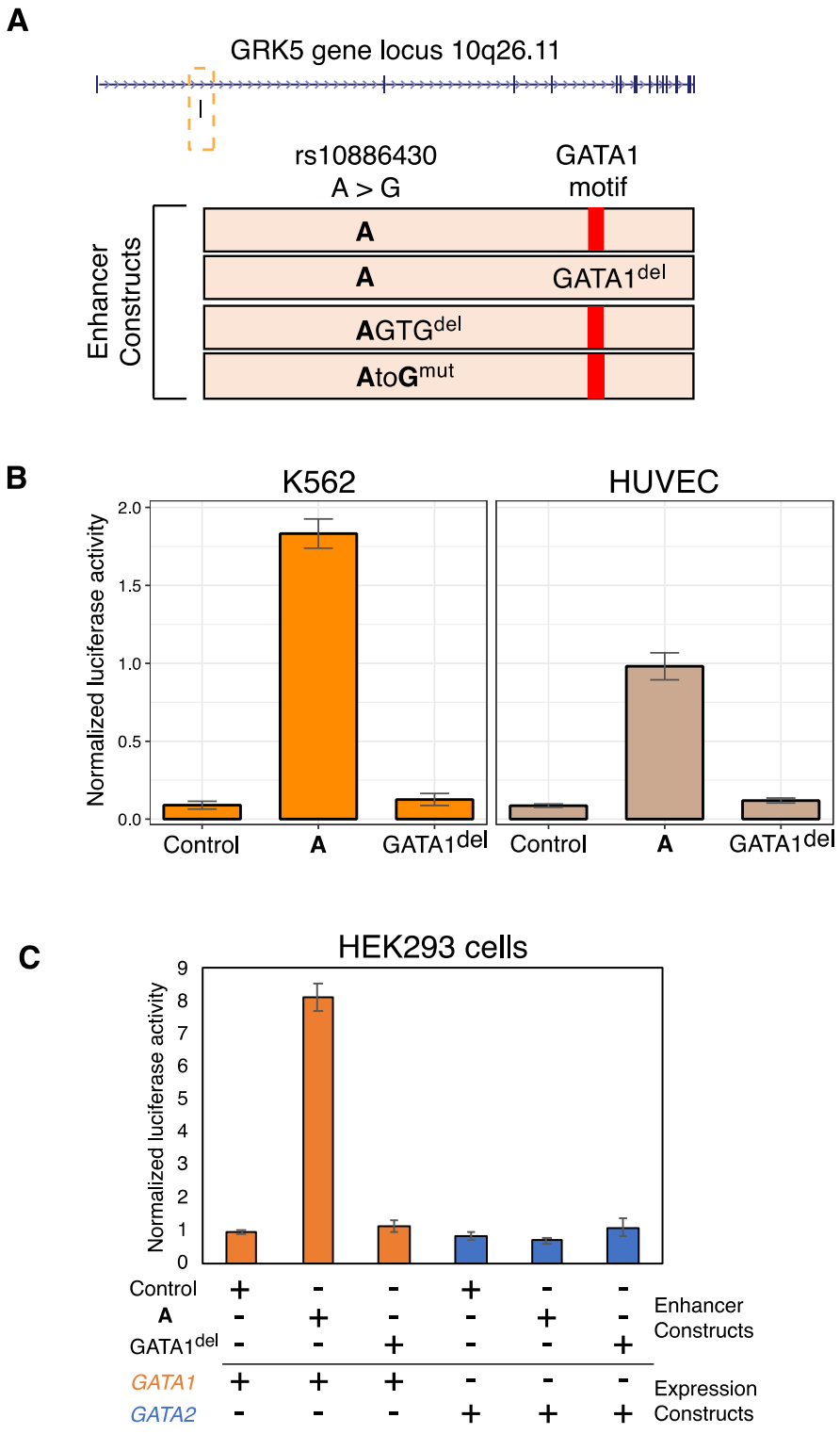

Figure S5. **The GRK5 SNP locus is a functional enhancer driven by GATA1 but not GATA2**. A, diagram of GRK5 enhancer constructs. B, normalized luciferase activity in K562 and HUVEC cells. C, normalized luciferase activity in GATA1- or GATA2-transduced HEK293 cells. B – C, Results are the mean of 3 independent experiments performed in quadruplicate (n total=12), error bars represent SEM.

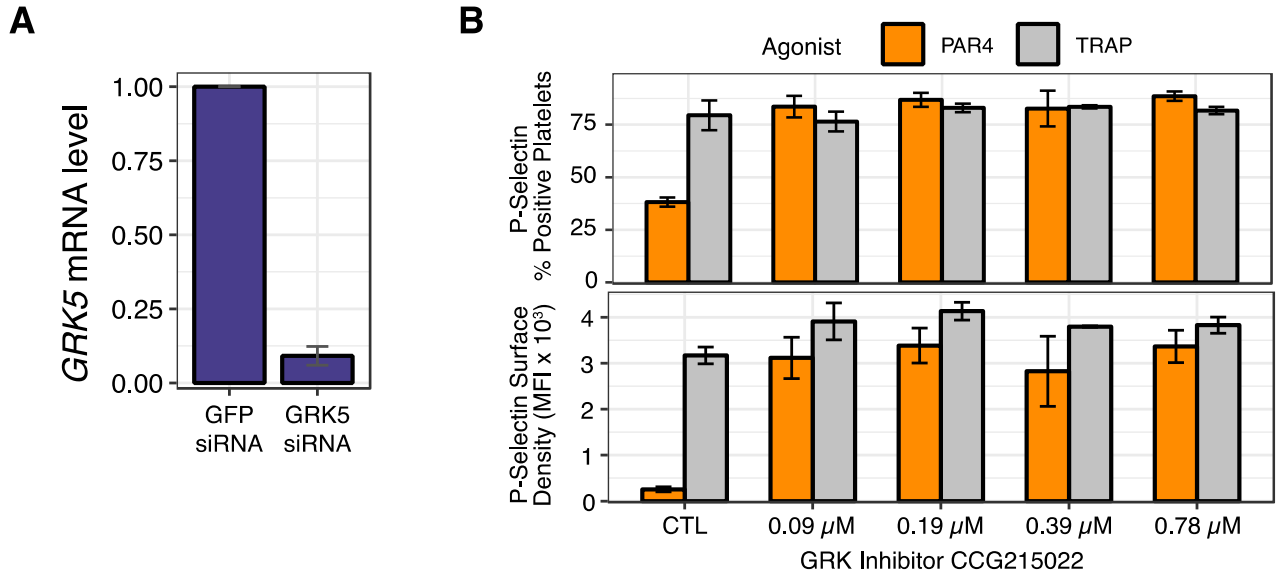

Figure S6. **siRNA treatment in megakaryocytic cells (imMKCLs) significantly reduces *GRK5* expression, and GRK Inhibition by CCG215022 significantly increased platelet activation via PAR4 but not PAR1 (TRAP-6), supporting Figure 5**. A, qRT-PCR analyses of GRK5 siRNA knockdown in Day 6 imMKCLs (4 biological replicates assayed in triplicate). Fold change in *GRK5* expression calculated by the $\Delta\Delta Ct$ method using *ACTB* as reference gene. Error bars represent SEM (n=4). B, Change in P-Selectin in response to specific activation of either PAR1 (20 μM TRAP6) or PAR4 (50 μM PAR4-AP) signaling (compared to untreated) following pharmacologic inhibition of GRK activity with increasing concentrations of CCG215022 in PRP samples from healthy donors (n=3). Top, difference in percentages of marker-positive platelets (agonist-treated compared to no agonist). Bottom, difference in marker platelet surface density (as measured by geometric MFI). Error bars represent SEM (n=3).

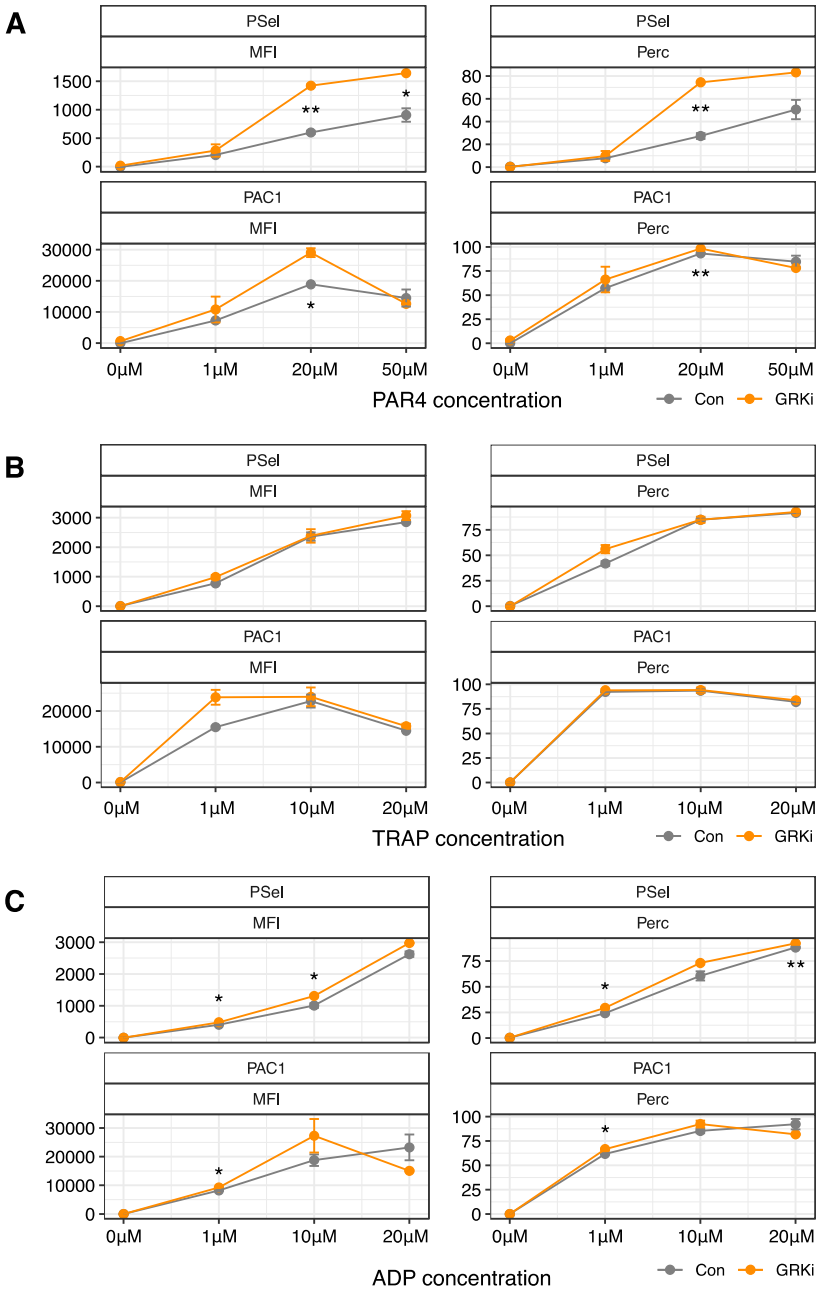

Figure S7. **Utilizing different doses of ADP, PAR4AP and TRAP-6 in platelet rich plasma from healthy donors, GRK inhibition by CCG215022 significantly increases platelet activation via moderate doses of PAR4 and to a lesser extent ADP, but not TRAP-6, supporting Figure 5.** Changes in P-Selectin or PAC1 in response to specific activation by increasing concentrations of either PAR4 (A, PAR4-AP), PAR1 (B, TRAP6) or ADP (C) signaling (compared to untreated) following pharmacologic inhibition of GRK activity with 0.78 μM CCG215022 (GRKi, orange points on plot) in PRP samples from healthy donors (n=3). Top row of each facet indicates the platelet function marker, P-Selectin (PSel) or PAC1. Second row of each facet represents the FACS analysis metric, Percent Positive Platelets (Perc) or Cell Surface Mean Fluorescent Intensity (MFI). Differences between PRP treated with GRK inhibitor versus control assessed by Student's T Test **, $p < 0.01$; *, $p < 0.05$.

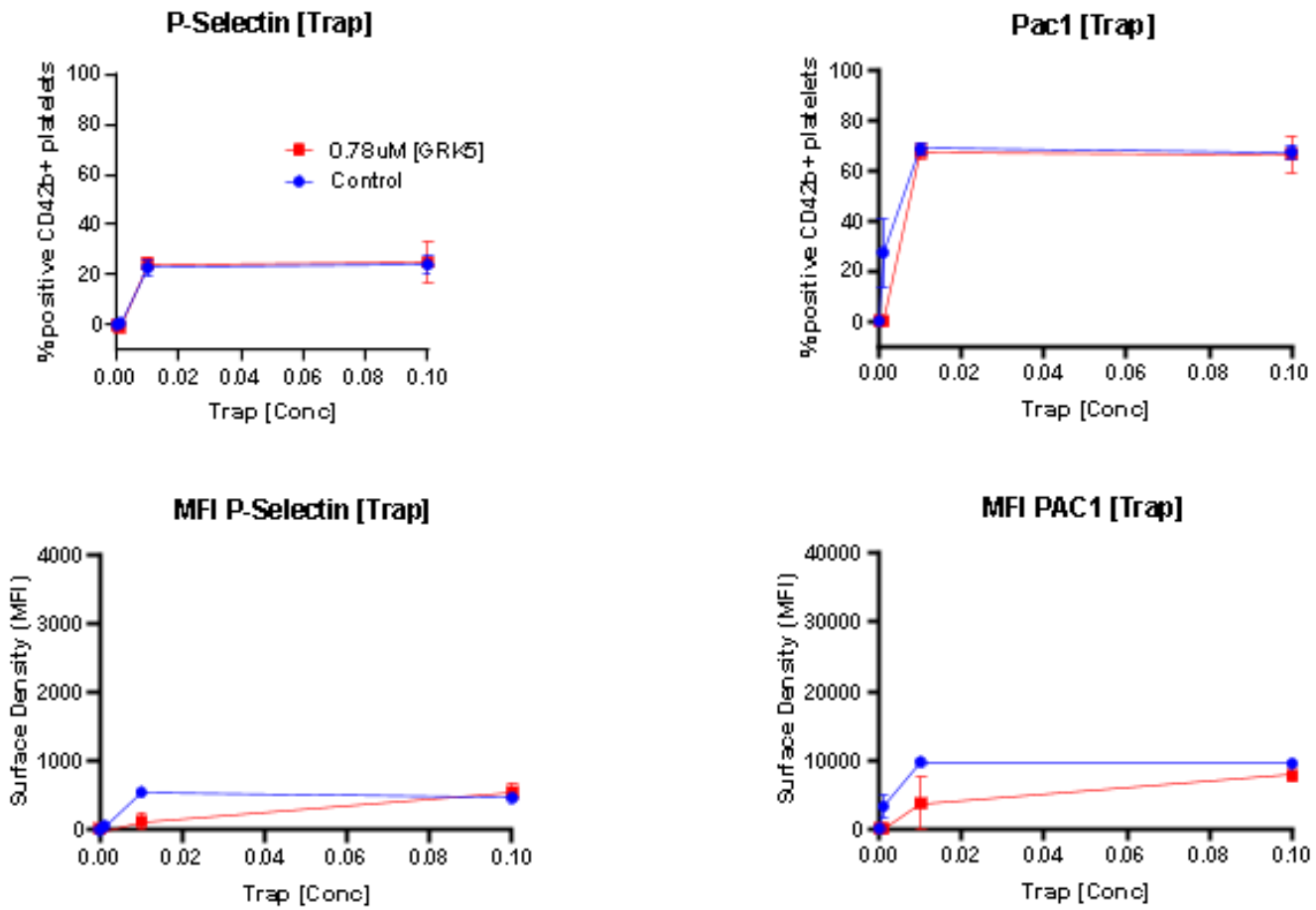

Figure S8. **Inhibition of platelet GRK by CCG215022 does not affect PAR1 (TRAP-6) mediated platelet activation at submaximal TRAP-6 doses (0.001 uM, 0.01 uM, 0.10 uM).** This is an extension of Figures 5 and S6 utilizing independent PRP donors (n=3). Changes in P-Selectin or PAC1 in response to specific activation by increasing concentrations of PAR1 (TRAP6: 0.001 uM, 0.01 uM, 0.10 uM) following pharmacologic inhibition of GRK activity with 0.78 μM CCG215022 (GRKi, red points on plot), or vehicle control (blue points), in PRP samples from healthy donors (n=3). The FACS analyses shown are: top left (% double-positive platelets for P-selectin (CD62P+)/CD42b+), bottom left (Cell Surface Mean Fluorescent Intensity (MFI) for CD62P+/CD42b+ platelets), top right (% double-positive platelets for fibrinogen binding (PAC1+)/CD42b+), bottom right (Cell Surface MFI for PAC1+/CD42b+ platelets).

Supplementary Tables.

Dataset	Phenotype, Sample Size / Ancestry	Figure/Table	PubMed ID
Caerphilly GWAS sample	LTA in PRP for Thrombin (n = 1184) / EA	Figures 1, S1, S2, S3, Tables S2, S3	this study
UK BioBank+Interval	PLT (n=166066), MPV (n=164454), PDW (n=164433), RBC (n=172952), WBC (n=172435) / EA	Figure 2A, S2, Table S4	27863252
Caerphilly	LTA in PRP: ADP (n=1184), Collagen (n=811) / EA	Table S5	this study
FHS / GeneStar	LTA in PRP: ADP (n = 4 trait concentration meta-analyses, n=3024-3857), Epinephrine (n = 4 trait concentration meta-analyses, n=2452-3596), Collagen lag time (3467) / EA	Table S6	20526338
MARTHA / 3C	Thrombin generation potential: ETP (n=1941), Peak (n=1950), Lag (n=1967) / EA	Table S7	24357727
PRAX1	Platelet gene expression QTLs (Affymetrix Human ST Gene 1.0 microarray) and 1000 Genomes Project imputation / n=80 EA / n=74 AA	Figure 2B, Table S8	27132591
GTEx eQTL version 7	Gene expression QTLs in 44 tissues, sample size range (n=70-361)	Table S9	29022597
CEDAR	Gene expression QTLs (Illumina HT12) in blood cell subtypes: CD4+ (n=303), CD8+ (n=294), CD19+ (n=282), CD14+ (n=286), CD15+ (n=289), Platelets (n=251) / EA	Table S10	29930244
Aortic endothelial cells	Gene expression QTLs (n = 147) / Trans-ethnic, primarily EA > AA	Table S11	23667179
Platelet RNA-seq	Expression levels of GRK family members from RNA-seq data (n=32) / EA	Discussion	26367242
iPSC-derived megakaryocytes and platelets	iPS line derived from fibroblast male donor with healthy karyotype, with derived imMKCLs used in experiments all n ≥ 3	Figure S6, S8	24529595
Donor PRP	Healthy donor PRP at Boston Children's Hospital (all n ≥ 3 independent donors)	Figure S7	this study
UK BioBank (GeneAtlas)	9 disease outcomes, cases (n=3502-11636), total study (n=452264) / EA	Figure 2C, Table S12	30349118
MEGASTROKE	4 disease outcomes, cases (n=688-67162), controls (n=454450) / Trans-ethnic	Table S13	29531354
FinnGen	16 disease outcomes, cases (n=3271-43576), controls (n=97214- 170880) / EA	Table S14	32273609

Associated Table refers to the table where we report results using the indicated dataset(s). LTA, light transmission aggregometry; PRP, platelet rich plasma; EA, European ancestry; MPV, mean platelet volume; PLT, platelet count; PDW, platelet distribution width; WBC, white blood cell; RBC, red blood cell; FHS, Framingham Heart Study; 3C, 3 City Study; ETP, endogenous thrombin potential; PRAX1, Platelet RNA and Expression; AA, African ancestry; TBD, to be determined. NA, not applicable.

Table S1. Sample sizes, ancestry and sources of datasets used in this publication.

Phenotype	Mean (SD), or Number Yes (%)
Age (years)	56.8 (4.5)
Body mass index (kg/m2)	26.7 (3.6)
Fasting glucose (mg/dL)	96.3 (22.3)
HDL cholesterol (mg/dL)	39.5 (9.6)
LDL cholesterol (mg/dL)	143.5 (35.7)
Triglycerides (mg/dL)	168.4 (97.3)
Systolic blood pressure (mmHg)	151.1 (23.3)
Diastolic blood pressure (mmHg)	89.2 (12.3)
Smoker (current; yes/no)	484 (40.9%)
Thrombin LTA in PRP (percent aggregation)	10.5 (5.7)
ADP LTA in PRP (percent aggregation)	22.3 (17.9)
Collagen LTA in PRP (percent aggregation)	9.8 (8.4)
Taking antiplatelet medication (current; yes/no)	126 (10.6%)

HDL, high density lipoprotein; LDL, low density lipoprotein; LTA, light transmission aggregometry; PRP, platelet rich plasma.

Table S2. Summary demographics and phenotype values for the Caerphilly GWAS study sample (n=1184).

rsID	Beta	SE	P	Ref	Alt	Alt Freq	MAF	Rsq	Chr	Pos	Cond	Beta	SE	P val
rs10886430	0.700	0.051	2.97E-42	A	G	0.136	0.136	0.908	10	121010256		n/a	n/a	n/a
rs10886437	0.532	0.045	6.57E-33	C	T	0.181	0.181	0.982	10	121037154	>	0.111	0.080	0.1649
rs10886443	0.354	0.039	5.94E-20	A	G	0.294	0.294	0.968	10	121054613	>	0.075	0.047	0.1058
rs56095167	0.301	0.040	3.39E-14	G	A	0.324	0.324	0.923	10	121004561	>	0.022	0.045	0.6337
rs928670	-0.229	0.036	2.58E-10	C	T	0.531	0.469	0.997	10	121031659	>	0.056	0.037	0.1325
rs10886436	-0.228	0.036	2.84E-10	T	G	0.533	0.467	0.997	10	121032130	>	0.055	0.037	0.1379
rs7076992	-0.227	0.036	3.22E-10	C	T	0.530	0.470	1.000	10	121028387	>	0.053	0.037	0.1480
rs7077224	-0.227	0.036	3.23E-10	G	T	0.530	0.470	1.000	10	121028390	>	0.053	0.037	0.1482
rs7071131	-0.227	0.036	3.52E-10	A	G	0.469	0.469	0.995	10	121026712	>	-0.052	0.037	0.1607
rs7090417	-0.226	0.036	3.83E-10	C	G	0.529	0.471	0.997	10	121030668	>	0.054	0.037	0.1388
rs10510056	-0.224	0.036	5.95E-10	C	T	0.533	0.467	0.994	10	121041733	>	0.052	0.037	0.1591
rs1889743	-0.218	0.036	1.24E-09	T	C	0.517	0.483	0.997	10	121059347	>	0.059	0.036	0.1055
rs4752274	-0.220	0.036	1.32E-09	G	T	0.534	0.466	0.994	10	121047585	>	0.050	0.037	0.1735
rs11198861	-0.221	0.036	1.42E-09	G	A	0.534	0.466	0.990	10	121049819	>	0.049	0.037	0.1815
rs10886444	-0.217	0.036	1.52E-09	G	A	0.517	0.483	0.999	10	121058459	>	0.057	0.036	0.1134
rs77984972	0.670	0.111	1.60E-09	G	C	0.032	0.032	0.887	10	120976563	>	0.080	0.115	0.4865
rs4752282	-0.217	0.036	1.64E-09	A	G	0.517	0.483	0.994	10	121057333	>	0.057	0.036	0.1196

GWAS, genome-wide association study; SE, standard error; Ref, reference allele; Alt, alternative allele; Freq, frequency; MAF, minor allele frequency; Rsq, imputation score; Chr, chromosome; Pos, position; Cond, GWAS results conditioned on rs10886430.

Table S3. Genome-wide significant (P < 7E-9) variants from GWAS of thrombin-induced platelet aggregation in Caerphilly participants and conditioned on lead SNP rs10886430.

	H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant					
Number of SNPs tested	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**	**Trait 2**	**Lead SNP**	**log Approx Bayes Factor Trait 1**	**log Approx Bayes Factor Trait 2**	**SNP Posterior Probability H4**
4548	5.18E-45	2.13E-14	2.44E-34	1.47E-11	1	MPV	rs10886430	79.700	33.783	1
4548	2.21E-34	9.08E-04	2.56E-34	5.31E-05	0.999	PLT	rs10886430	79.700	9.306	1
4548	4.26E-41	1.75E-10	2.44E-34	1.59E-11	1	PDW	rs10886430	79.700	24.770	1
4548	1.74E-31	0.715	5.95E-32	0.244	0.040	WBC	rs10886430	79.700	-0.573	1
4548	2.32E-31	0.951	1.08E-32	0.044	0.004	RBC	rs10886430	79.700	-3.085	1

GWAS, genome-wide association study; LD, linkage disequilibrium; H, hypothesis; Approx, approximate; MPV, mean platelet volume; PLT, platelet count; PDW, platelet distribution width; WBC, white blood cell count; RBC, red blood cell count.

Table S4. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the rs10886430-G containing LD haplotype block with quantitative blood cell traits (WBC, PLT, PDW, MPV, RBC).

		H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
ADP (0.725uM)	5427	1.54E-31	0.633	7.88E-32	0.323	0.044
Collagen (42.7ug/mL)	5427	1.64E-31	0.675	6.96E-32	0.286	0.040

Table S5. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the rs10886430-G containing LD haplotype block with Caerphilly participants ADP- and Collagen-induced aggregation signals.

		H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
ADP 3um 2um	1469	2.19E-31	0.901	1.41E-32	0.058	0.041
ADP 5um 10um	1469	2.19E-31	0.899	1.46E-32	0.060	0.041
ADP EC50 10um	1469	2.25E-31	0.923	1.27E-32	0.052	0.025
ADP EC50 2um	1469	2.24E-31	0.918	1.44E-32	0.059	0.023
Collagen lag time	1469	2.17E-31	0.893	2.23E-32	0.092	0.016
Epi 1uM 2uM	1469	2.24E-31	0.920	1.54E-32	0.063	0.017
Epi 3uM 10uM	1469	2.16E-31	0.885	1.95E-32	0.080	0.035
Epi EC50 10uM	1469	2.22E-31	0.910	1.83E-32	0.075	0.015
Epi EC50 2uM	1469	2.21E-31	0.906	1.96E-32	0.080	0.014

Trait 2 platelet agonist GWAS were published in Nature Genetics 42: 608. Trait names represent best matched platelet phenotypes between two cohorts used for meta-analyses.

Table S6. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the rs10886430-G containing LD haplotype block with published aggregation GWAS signals for ADP, Collagen, and Epinephrine.

		H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
Peak	1659	2.03E-31	0.835	2.74E-32	0.112	0.053
Lag time	1649	2.13E-31	0.873	2.43E-32	0.100	0.027
ETP	1575	2.11E-31	0.868	2.49E-32	0.102	0.030

Trait 2 thrombin generation phenotypes were published in Blood 123(5): 777-785. ETP, endogenous thrombin potential.

Table S7. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the rs10886430-G containing LD haplotype block with GWAS signals for thrombin generation potential (peak, lag time, endogenous thrombin potential).

Gene Chr	Gene	Gene TSS	Top SNP	SNP Chr	SNP Position	Effect Allele	Ref Allele	Beta	SE	SMR P value	HEIDI P value
10	GRK5	120967196	rs10886430	10	121010256	G	A	-1.536	0.203	3.67E-14	1.61E-01
1	PGM1	64058946	rs855300	1	64078224	A	G	-0.536	0.165	1.19E-03	6.32E-01
9	HEMGN	100700675	rs7861658	9	100679954	G	A	-0.322	0.110	3.35E-03	6.13E-01
9	HABP4	99212436	rs10990651	9	99105420	A	G	-0.160	0.056	4.25E-03	5.47E-01
16	NUDT21	56485261	rs28793026	16	56415786	T	C	0.563	0.207	6.58E-03	1.18E-02
11	LAMTOR1	71814433	rs58417942	11	71840047	G	A	-1.690	0.632	7.54E-03	NA[1]
16	ZDHHC7	85045141	rs59109628	16	85059628	C	A	-0.541	0.208	9.23E-03	NA[1]

SMR, summary-based mendelian randomization analysis; GWAS, genome-wide association study; PRAX1, platelet RNA and expression 1; Chr, chromosome; TSS, transcriptional start site; SE, standard error; HEIDI, heterogeneity in dependent instruments. 1, insufficient number of SNPs to perform HEIDI test.

Table S8. SMR analysis of GWAS of thrombin-induced platelet aggregation of Caerphilly participants and PRAX1 study platelet eQTL.

		H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
Adipose Subcutaneous	5837	1.39E-31	0.569	9.67E-32	0.397	0.034
Adipose Visceral Omentum	5837	1.58E-31	0.649	7.67E-32	0.315	0.036
Adrenal Gland	5836	1.52E-31	0.625	7.83E-32	0.321	0.054
Artery Aorta	5836	1.72E-31	0.706	6.31E-32	0.259	0.035
Artery Coronary	5833	1.51E-31	0.619	8.05E-32	0.330	0.051
Artery Tibial	5837	1.53E-32	0.063	2.28E-31	0.935	0.003
Brain Anterior cingulate cortex BA24	5814	1.48E-31	0.607	8.50E-32	0.349	0.044
Brain Caudate basal ganglia	5833	1.21E-31	0.495	1.14E-31	0.466	0.038
Brain Cerebellar Hemisphere	5825	1.45E-31	0.594	8.75E-32	0.359	0.046
Brain Cerebellum	5830	1.50E-31	0.616	8.29E-32	0.340	0.044
Brain Cortex	5830	1.07E-31	0.438	1.30E-31	0.535	0.027
Brain Frontal Cortex BA9	5826	1.41E-31	0.581	9.07E-32	0.372	0.047
Brain Hippocampus	5823	1.29E-31	0.528	1.05E-31	0.433	0.039
Brain Hypothalamus	5826	1.46E-31	0.598	8.41E-32	0.345	0.057
Brain Nucleus accumbens basal ganglia	5831	1.15E-31	0.473	7.26E-32	0.298	0.229
Brain Putamen basal ganglia	5829	1.45E-31	0.595	8.40E-32	0.345	0.060
Breast Mammary Tissue	5837	1.69E-31	0.692	6.75E-32	0.277	0.030
Cells EBV-transformed lymphocytes	5823	1.38E-31	0.567	9.54E-32	0.392	0.041
Cells Transformed fibroblasts	5837	1.29E-31	0.530	1.09E-31	0.447	0.023
Colon Sigmoid	5837	1.52E-31	0.623	8.23E-32	0.338	0.039
Colon Transverse	5837	1.60E-31	0.656	7.76E-32	0.318	0.025
Esophagus Gastroesophageal Junction	5837	1.47E-31	0.603	8.23E-32	0.338	0.059
Esophagus Mucosa	5837	1.54E-31	0.630	8.24E-32	0.338	0.031
Esophagus Muscularis	5837	1.54E-31	0.632	8.23E-32	0.338	0.030
Heart Atrial Appendage	5837	1.30E-31	0.536	1.07E-31	0.439	0.025
Heart Left Ventricle	5837	1.31E-31	0.539	1.01E-31	0.417	0.045

Liver	5833	1.43E-31	0.588	8.92E-32	0.366	0.045
Lung	5837	2.65E-33	0.011	2.41E-31	0.989	0.001
Muscle Skeletal	5837	1.41E-31	0.578	9.75E-32	0.400	0.022
Nerve Tibial	5837	1.33E-31	0.544	1.06E-31	0.434	0.021
Ovary	5821	1.44E-31	0.590	8.37E-32	0.343	0.066
Pancreas	5836	1.48E-31	0.608	8.47E-32	0.348	0.044
Pituitary	5829	1.50E-31	0.617	8.46E-32	0.347	0.035
Prostate	5830	1.50E-31	0.615	8.49E-32	0.349	0.037
Skin Not Sun Exposed Suprapubic	5837	1.73E-31	0.711	6.51E-32	0.267	0.022
Skin Sun Exposed Lower leg	5837	1.17E-31	0.482	9.25E-32	0.380	0.138
Small Intestine Terminal Ileum	5825	1.25E-31	0.513	1.09E-31	0.448	0.039
Spleen	5828	1.46E-31	0.600	8.51E-32	0.349	0.051
Stomach	5836	1.35E-31	0.552	1.01E-31	0.416	0.032
Testis	5837	1.43E-31	0.586	8.16E-32	0.335	0.079
Thyroid	5837	4.15E-41	1.70E-10	2.44E-31	1.000	7.41E-12
Uterus	5807	1.42E-31	0.582	9.33E-32	0.383	0.035
Vagina	5823	1.50E-31	0.617	8.32E-32	0.342	0.042
Whole Blood	5837	7.43E-32	0.305	1.55E-31	0.634	0.061

Variants tested for tissue GRK5 cis eQTL in GTEx version 7. All variants in the region spanning up to 1-MB upstream and downstream of the GRK5 TSS and TTS, respectively (hg19 chr10:119967083-122219257). TSS, transcriptional start site; TTS, transcriptional termination site.

Table S9. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the *GRK5* locus with 44 GTEx tissue eQTL signals.

	H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
CD4+ T cells	4099	3.00E-34	0.001	2.43E-31	0.997	0.002
CD8+ T cells	4099	2.63E-33	0.011	2.40E-31	0.987	0.002
CD19+ B cells	4099	1.60E-31	0.657	6.54E-32	0.268	0.074
CD14+ monocytes	4099	1.64E-31	0.672	6.91E-32	0.284	0.044
CD15+ granulocytes	4099	1.96E-32	0.080	2.23E-31	0.914	0.005
Platelets	4099	4.41E-44	0.000	2.44E-34	0.000	1.000

All variants in the region spanning up to 1-MB upstream and downstream of the GRK5 TSS and TTS, respectively (hg19 chr10:119967083-122219257).

Table S10. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the *GRK5* locus with 6 blood cell type eQTL from the CEDAR cohort.

		H0: No causal variant	H1: Causal variant for trait 1 only	H2: Causal variant for trait 2 only	H3: Two distinct causal variants	H4: One common causal variant
Trait 2	**Number of SNPs tested**	**Posterior Probability H0**	**Posterior Probability H1**	**Posterior Probability H2**	**Posterior Probability H3**	**Posterior Probability H4**
Aortic endothelial cells	6294	1.39E-31	0.572	9.74E-32	0.400	0.029

All variants in the region spanning up to 1-MB upstream and downstream of the GRK5 TSS and TTS, respectively (hg19 chr10:119967083-122219257).

Table S11. Colocalization analysis of thrombin-induced platelet aggregation GWAS signals of Caerphilly participants in the *GRK5* locus with endothelial cell eQTL.

Outcome[1]	GWAS OR	GWAS SE	GWAS P value	Cases	OR of MR Estimate	MR SE	MR CI Lower	MR CI Upper	MR P value
I26 Pulmonary embolism http://geneatlas.roslin.ed.ac.uk/trait/?trait=401	1.25	1.03	8.05E-13	3952	1.382	1.050	1.257	1.520	2.40E-11
I26-I28 Pulmonary heart disease and diseases of pulmonary circulation http://geneatlas.roslin.ed.ac.uk/trait/?trait=297	1.22	1.03	1.42E-11	4800	1.324	1.045	1.214	1.443	1.97E-10
1083 Cerebrovascular disease self-reported http://geneatlas.roslin.ed.ac.uk/trait/?trait=88	1.12	1.02	8.57E-07	8385	1.173	1.034	1.099	1.253	1.85E-06
1098 Stroke, self-reported http://geneatlas.roslin.ed.ac.uk/trait/?trait=308	1.13	1.03	2.35E-06	6465	1.190	1.039	1.104	1.282	5.05E-06
1085 Venous thromboembolic disease, self-reported http://geneatlas.roslin.ed.ac.uk/trait/?trait=497	1.09	1.02	3.90E-06	11636	1.136	1.029	1.075	1.201	6.63E-06
1112 Deep venous thrombosis (DVT), self-reported http://geneatlas.roslin.ed.ac.uk/trait/?trait=262	1.09	1.02	3.73E-05	9059	1.138	1.033	1.068	1.212	5.69E-05
I21 Acute myocardial infarction, clinical http://geneatlas.roslin.ed.ac.uk/trait/?trait=397	1.09	1.02	4.41E-05	8764	1.138	1.033	1.068	1.212	6.67E-05
I63 Cerebral infarction, clinical http://geneatlas.roslin.ed.ac.uk/trait/?trait=526	1.13	1.04	5.01E-04	3502	1.190	1.053	1.076	1.316	7.43E-04
I50 Heart Failure http://geneatlas.roslin.ed.ac.uk/trait/?trait=56	1.02	1.03	3.92E-01	5901	1.035	1.040	0.957	1.118	0.392

MR, mendelian randomization; GWAS, genome-wide association study; UK, United Kingdom; OR, odds ratio; SE, standard error; CI, 95% confidence interval. 1, given are the full labels of outcomes as cataloged in the Gene ATLAS database and the corresponding URL. Outcomes with prefix "I" are derived from ICD9 and / or ICD10 codes.

Table S12. MR analysis of rs10886430-G thrombin-induced platelet aggregation GWAS signal in Caerphilly participants and cardiovascular, cardiopulmonary outcomes in the UK BioBank cohort, in support of Figure 2C.

Outcome[1]	GWAS OR	GWAS SE	GWAS P value	Cases	OR of MR Estimate	MR SE	MR CI Lower	MR CI Upper	MR P value
Cardioembolic Stroke	18.42	2.96	6.16E-04	9006	64.36	4.71	3.09	1340.44	7.18E-03
Any Stroke	1.50	1.13	2.83E-04	67162	1.79	1.19	1.27	2.51	7.69E-04
Any Ischemic Stroke	1.61	1.15	2.29E-04	60341	1.97	1.22	1.33	2.92	7.17E-04
Large Artery Stroke	4.32	5.09	3.85E-01	6688	8.10	10.24	0.08	773.88	3.69E-01

MR, mendelian randomization; GWAS, genome-wide association study; OR, odds ratio; SE, standard error; CI, 95% confidence interval. 1, GWAS of stroke outcomes correspond to the trans-ethnic analyses published in Nature Genetics 50, 524-537 (2018). The rs10886430 variant did not pass QC in Small Vessel Stroke subtype trans-ethnic analysis, thus this subtype was not included in our analysis; however, there was no evidence of association in the European Only analysis (P = 0.82)

Table S13. **MR analysis of rs10886430-G thrombin-induced platelet aggregation GWAS signal in Caerphilly participants and stroke outcomes in the MEGASTROKE consortium.**

Outcome	pheWAS OR	pheWAS SE	pheWAS P value	Cases	Controls	OR of MR Estimate	MR SE	MR CI Lower	MR CI Upper	MR P value
Deep vein thrombosis (DVT) of lower extremities	1.17	1.04	2.32E-04	3592	153951	1.25	1.06	1.11	1.41	2.28E-04
Ischemic Stroke, excluding all hemorrhages	1.11	1.03	5.78E-04	8046	164286	1.15	1.04	1.06	1.25	5.75E-04
DVT of lower extremities and pulmonary embolism	1.11	1.03	1.22E-03	6019	170880	1.17	1.05	1.06	1.28	1.22E-03
Cardiomyopathy	0.84	1.06	1.84E-03	2342	129908	0.79	1.08	0.68	0.91	1.85E-03
Stroke, excluding subarachnoid hemorrhage (SAH)	1.08	1.03	3.50E-03	8877	163535	1.12	1.04	1.04	1.21	3.57E-03
Portal vein thrombosis	1.90	1.27	6.30E-03	110	153951	2.51	1.40	1.30	4.85	6.31E-03
Cardiomyopathies, Primary/intrinsic	0.84	1.07	6.52E-03	1678	129908	0.78	1.10	0.65	0.93	6.53E-03
Hypertension	0.96	1.02	7.65E-03	43545	133323	0.94	1.02	0.89	0.98	7.49E-03
Hypertensive diseases	0.96	1.02	7.82E-03	43576	133323	0.94	1.02	0.89	0.98	7.80E-03
Stroke, including SAH	1.07	1.03	8.32E-03	9592	162861	1.11	1.04	1.03	1.19	8.25E-03
Right bundle-branch block	1.44	1.15	8.54E-03	314	129908	1.68	1.22	1.14	2.48	8.51E-03
Venous thromboembolism (VTE)	1.08	1.03	1.00E-02	6913	169986	1.12	1.05	1.03	1.22	9.97E-03
Myocardial infarction	1.08	1.03	1.09E-02	10003	152621	1.12	1.04	1.03	1.22	1.09E-02
Myocardial infarction, strict	1.08	1.03	1.25E-02	9145	152621	1.12	1.05	1.02	1.22	1.24E-02
Cerebrovascular diseases	1.06	1.02	1.27E-02	11859	165040	1.09	1.04	1.02	1.17	1.26E-02
Other arterial embolism and thrombosis	1.45	1.17	1.61E-02	256	167843	1.70	1.25	1.10	2.61	1.62E-02
Atrioventricular (AV) block	1.16	1.06	1.63E-02	1727	129908	1.24	1.09	1.04	1.47	1.64E-02
Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	1.04	1.02	3.64E-02	22948	153951	1.06	1.03	1.00	1.12	3.69E-02
Cerebrovascular diseases (FINNGEN)	1.07	1.03	3.65E-02	10367	97214	1.09	1.04	1.01	1.19	3.63E-02
Conduction disorders	1.10	1.05	4.12E-02	3271	129908	1.14	1.07	1.01	1.29	4.13E-02
Sequelae of cerebrovascular disease	1.09	1.04	4.42E-02	3520	165040	1.13	1.06	1.00	1.28	4.45E-02

1MR, mendelian randomization; OR, odds ratio; SE, standard error; CI, 95% confidence interval. DVT, deep venous thrombosis; SAH, Subarachnoid hemorrhage.

Table S14. MR analysis of rs10886430-G thrombin-induced platelet aggregation GWAS signal in Caerphilly participants and circulatory system disease outcomes in the FinnGen study

Protein	Cell model	Chrom	Start	Stop	Signal value	Q value	Experiment Name	File Name
GATA1	PBDE	chr10	121010043	121010647	31.7	4.90	ENCSR000EXP	ENCFF957CWW
JUND	K562	chr10	121010130	121010410	56.5	4.92	ENCSR000EGN	ENCFF337DKJ
CBFA2T3	K562	chr10	121010119	121010523	110.9	4.77	ENCSR697YLJ	ENCFF993GXU
NFE2	K562	chr10	121010182	121010343	108.3	4.75	ENCSR000FCC	ENCFF495MHZ
C11orf30	K562	chr10	121010068	121010419	331.9	4.74	ENCSR350XWY	ENCFF996ZGL
MAFF	K562	chr10	121010123	121010399	56.9	4.44	ENCSR000EGI	ENCFF308IXJ
ZNF316	K562	chr10	121010091	121010400	370.8	4.37	ENCSR167KBO	ENCFF788EBS
IKZF1	K562	chr10	121010059	121010362	178.8	4.32	ENCSR948VFL	ENCFF886VSU
L3MBTL2	K562	chr10	121010154	121010754	56.8	4.26	ENCSR530XQI	ENCFF314ULQ
DPF2	K562	chr10	121010087	121010385	110.3	4.22	ENCSR715CCR	ENCFF823SYE
SMARCE1	K562	chr10	121009987	121010503	58.9	4.05	ENCSR157TCS	ENCFF622RBW
ARID1B	K562	chr10	121010142	121010437	234.3	4.03	ENCSR822CCM	ENCFF225MPC
CREM	K562	chr10	121010129	121010429	19.7	4.00	ENCSR077DKV	ENCFF948TXN
CTBP1	K562	chr10	121010073	121010613	56.6	3.88	ENCSR201NQZ	ENCFF152VMJ
ATF3	K562	chr10	121010038	121010538	99.4	3.81	ENCSR028UIU	ENCFF958KNK
SMARCA4	K562	chr10	121010175	121010403	106.3	3.71	ENCSR587OQL	ENCFF883TOD
EP300	K562	chr10	121010096	121010412	71.5	3.68	ENCSR000EGE	ENCFF549TYR
MAFK	K562	chr10	121010091	121010371	95.7	3.21	ENCSR000EGX	ENCFF812QPN
ZNF592	K562	chr10	121010196	121010556	109.2	3.06	ENCSR249BHQ	ENCFF273TYA
ZEB2	K562	chr10	121010140	121010590	41.9	2.81	ENCSR322CFO	ENCFF407AOX
SMARCC2	K562	chr10	121010093	121010583	50.9	2.70	ENCSR519WMW	ENCFF114PTZ
MGA	K562	chr10	121010171	121010735	140.5	2.64	ENCSR710WLO	ENCFF670ZCR
RNF2	K562	chr10	121009997	121010597	112.3	2.61	ENCSR608XTF	ENCFF644WLI
ARNT	K562	chr10	121010132	121010608	40.5	2.53	ENCSR613NUC	ENCFF471ZTZ
NCOR1	K562	chr10	121010073	121010483	101.0	2.40	ENCSR910JAI	ENCFF165BCW
CBFA2T2	K562	chr10	121010085	121010455	41.7	1.88	ENCSR699PVC	ENCFF642BNC
EP400	K562	chr10	121010101	121010811	51.3	1.67	ENCSR817QKV	ENCFF925ANU

ENCODE, Encyclopedia of DNA Elements; PBDE, peripheral blood derived erythroblasts; Chr, chromosome; Overlapping peak signal and Q value are extracted from file indicated by ENCODE File Name. Experiment Name is the ENCODE record for the experiment which produced the given file corresponding to File Name.

Table S15. **Transcription factors binding the rs10886430 variant in ENCODE mega-erythroid cell models.**

MAF	beta	power	MAF	beta	power	MAF	beta	power	MAF	beta	power	MAF	beta	power
0.01	0.1	0	0.02	0.1	0	0.03	0.1	0	0.04	0.1	0	0.05	0.1	0
	0.2	0		0.2	0		0.2	0		0.2	0.0001		0.2	0.0003
	0.3	0		0.3	0.0002		0.3	0.0015		0.3	0.0053		0.3	0.0143
	0.4	0.0002		0.4	0.0034		0.4	0.0211		0.4	0.0709		0.4	0.1636
	0.5	0.0011		0.5	0.0265		0.5	0.1391		0.5	0.3533		0.5	0.5948
	0.6	0.0061		0.6	0.1237		0.6	0.4477		0.6	0.7656		0.6	0.9294
	0.7	0.0254		0.7	0.3537		0.7	0.7962		0.7	0.9668		0.7	0.9967
	0.8	0.0805		0.8	0.6592		0.8	0.9637		0.8	0.9985		0.8	0.9999
	0.9	0.1984		0.9	0.8851		0.9	0.9972		0.9	0.9999		0.9	0.9999
	1	0.3854		1	0.9771		1	0.9999		1	0.9999		1	0.9999
0.1	0.1	0	0.15	0.1	0.0001	0.2	0.1	0.0002	0.25	0.1	0.0003	0.3	0.1	0.0005
	0.2	0.0064		0.2	0.0317		0.2	0.0839		0.2	0.1558		0.2	0.2327
	0.3	0.2055		0.3	0.5541		0.3	0.8055		0.3	0.9225		0.3	0.968
	0.4	0.8055		0.4	0.9845		0.4	0.9992		0.4	0.9999		0.4	0.9999
	0.5	0.9949		0.5	0.9999		0.5	0.9999		0.5	0.9999		0.5	0.9999
	0.6	0.9999		0.6	0.9999		0.6	0.9999		0.6	0.9999		0.6	0.9999
	0.7	0.9999		0.7	0.9999		0.7	0.9999		0.7	0.9999		0.7	0.9999
	0.8	0.9999		0.8	0.9999		0.8	0.9999		0.8	0.9999		0.8	0.9999
	0.9	0.9999		0.9	0.9999		0.9	0.9999		0.9	0.9999		0.9	0.9999
	1	0.9999		1	0.9999		1	0.9999		1	0.9999		1	0.9999

Table S16. Power analysis results.

FinnGen

Steering Committee

Aarno Palotie	University of Helsinki / FIMM
Mark Daly	University of Helsinki / FIMM

Pharma

Howard Jacob	Abbvie
Athena Matakidou	Astra Zeneca
Heiko Runz	Biogen
Sally John	Biogen
Robert Plenge	Celgene
Julie Hunkapiller	Genentech
Meg Ehm	GSK
Dawn Waterworth	GSK
Caroline Fox	Merck
Anders Malarstig	Pfizer
Kathy Klinger	Sanofi
Kathy Call	Sanofi

UH & Biobanks

Tomi Mäkelä	University of Helsinki / FIMM
Jaakko Kaprio	University of Helsinki / FIMM
Petri Virolainen	Auria BB / Univ. of Turku /VSSHP
Kari Pulkki	Auria BB / Univ. of Turku /VSSHP
Terhi Kilpi	THL Biobank (BB) / THL
Markus Perola	THL Biobank (BB) / THL
Jukka Partanen	Finnish Red Cross Blood Service/FHRB
Anne Pitkäranta	HUS/Univ Hosp Districts
Riitta Kaarteenaho	Borealis BB/Univ. of Oulu/PPSHP
Seppo Vainio	Borealis BB/Univ. of Oulu/PPSHP
Kimmo Savinainen	Tampere BB/Univ Tampere/PSHP
Veli-Matti Kosma	Eastern Finland BB/UEF/PSSHP
Urho Kujala	Central Finland BB /University of Jyväskylä

Other Experts/ Non-Voting Members

Outi Tuovila	Business Finland
Minna Hendolin	Business Finland
Raimo Pakkanen	Business Finland

Scientific Committee

Pharma

Jeff Waring	Abbvie
Bridget Riley-Gillis	AbbVie
Athena Matakidou	Astra Zeneca
Heiko Runz	Biogen
Jimmy Liu	Biogen
Shameek Biswas	Celgene
Julie Hunkapiller	Genentech
Dawn Waterworth	GSK
Meg Ehm	GSK
Josh Hoffman	GSK
Dorothee Diogo	Merck
Caroline Fox	Merck
Anders Malarstig	Pfizer
Catherine Marshall	Pfizer
Xinli Hu	Pfizer
Kathy Call	Sanofi
Kathy Klinger	Sanofi

UH & Biobanks

Samuli Ripatti	University of Helsinki / FIMM
Johanna Schleutker	Auria BB / Univ. of Turku /VSSHP
Markus Perola	THL Biobank (BB) / THL
Tiina Wahlfors	Finnish Red Cross Blood Service/FHRB
Olli Carpen	HUS/Univ Hosp Districts
Johanna Myllyharju	Borealis BB/Univ. of Oulu/PPSHP
Johannes Kettunen	Borealis BB/Univ. of Oulu/PPSHP Reijo
Laaksonen	Tampere BB/Univ Tampere/PSHP
Arto Mannermaa	Eastern Finland BB/UEF/PSSHP
Juha Paloneva	Central Finland BB /University of Jyväskylä/KSSHP
Urho Kujala	Central Finland BB /University of Jyväskylä

Other Experts/ Non-Voting Members

Outi Tuovila	Business Finland
Minna Hendolin	Business Finland
Raimo Pakkanen	Business Finland

Clinical Groups

Neurology Group

Hilkka Soininen LEAD Kuopio
Valtteri Julkunen Kuopio
Anne Remes Oulu
Reetta Kälviäinen Kuopio
Mikko Hiltunen Kuopio
Jukka Peltola Tampere
Pentti Tienari Helsinki
Juha Rinne Turku
Adam Ziemann AbbVie
Jeffrey Waring AbbVie
Sahar Esmaeeli AbbVie
Nizar Smaoui AbbVie
Anne Lehtonen AbbVie
Susan Eaton Biogen
Heiko Runz Biogen
Sanni Lahdenperä Biogen
Janet van Adelsberg Celgene
Shameek Biswas Celgene
John Michon Genentech
Geoff Kerchner Genentech
Julie Hunkapiller Genentech
Natalie Bowers Genentech
Edmond Teng Genentech
John Eicher Merck
Vinay Mehta Merck
Padhraig Gormley Merck Kari Linden Pfizer
Christopher Whelan Pfizer
Fanli Xu GSK
David Pulford GSK

Gastroenterology Group Martti Färkkilä LEAD Helsinki
Sampsa Pikkarainen HUS
Airi Jussila Tampere
Timo Blomster Oulu
Mikko Kiviniemi Kuopio
Markku Voutilainen Turku
Bob Georgantas AbbVie
Graham Heap AbbVie
Jeffrey Waring AbbVie
Nizar Smaoui AbbVie
Fedik Rahimov AbbVie
Anne Lehtonen AbbVie
Keith Usiskin Celgene
Tim Lu Genentech
Natalie Bowers Genentech
Danny Oh Genentech
John Michon Genentech
Vinay Mehta Merck
Dermot Reilly Merck
Kirsi Kalpala Pfizer
Melissa Miller Pfizer
Xinli Hu Pfizer
Linda McCarthy GSK

Rheumatology Group

Kari Eklund LEAD Helsinki
Antti Palomäki Turku
Pia Isomaki Tampere
Laura Pirilä Turku
Oili Kaipiainen-Seppänen Kuopio
Tuulikki Sokka-Isler KSSHP
Markku Kauppi Päijät-Häme Central Hospital/University of Tampere
Johanna Huhtakangas Oulu Eleanore Wigmore AstraZeneca
Bob Georgantas AbbVie
Jeffrey Waring AbbVie
Fedik Rahimov AbbVie
Apinya Lertratanakul AbbVie
Nizar Smaoui AbbVie
Anne Lehtonen AbbVie
Marla Hochfeld Celgene
Natalie Bowers Genentech
John Michon Genentech
Dorothee Diogo Merck
Vinay Mehta Merck
Kirsi Kalpala Pfizer
Nan Bing Pfizer
Xinli Hu Pfizer
Jorge Esparza Gordillo GSK
Nina Mars University of Helsinki / FIMM

Pulmonology Group

Tarja Laitine LEAD Tampere
Margit Pelkonen Kuopio
Paula Kauppi Helsinki
Hannu Kankaanranta Tampere
Terttu Harju Oulu
Nizar Smaoui AbbVie
Eleanore Wigmore AstraZeneca
Susan Eaton Biogen
Steven Greenberg Celgene
Hubert Chen Genentech
Natalie Bowers Genentech
John Michon Genentech
Vinay Mehta Merck
Jo Betts GSK
Soumitra Ghosh GSK

Cardiometabolic Diseases Group

Veikko Salomaa LEAD THL Teemu Niiranen THL
Markus Juonala Turku
Kaj Metsärinne Turku
Mika Kähönen Tampere
Juhani Junttila Oulu
Markku Laakso Kuopio

Jussi Pihlajamäki	Kuopio
Juha Sinisalo	Helsinki
Marja-Riitta Taskinen	Helsinki
Tiinamaija Tuomi	Helsinki
Jari Laukkanen	Keski-Suomen Keskussairaala/ University of Jyväskylä
Ben Challis	AstraZeneca
Keith Usiskin	Celgene
Andrew Peterson	Genentech
Julie Hunkapiller	Genentech
Natalie Bowers	Genentech
John Michon	Genentech
Dorothee Diogo	Merck
Dermot Reilly	Merck
Audrey Chu	Merck
Vinay Mehta	Merck
Jaakko Parkkinen	Pfizer
Melissa Miller	Pfizer
Anthony Muslin	Sanofi
Dawn Waterworth	GSK

Oncology Group

Heikki Joensuu LEAD	Helsinki
Tuomo Meretoja	Helsinki
Olli Carpen	Helsinki
Lauri Aaltonen	Helsinki
Annika Auranen	Tampere
Peeter Karihtala	Oulu
Saila Kauppila	Oulu
Päivi Auvinen	Kuopio
Klaus Elenius	Turku
Relja Popovic	AbbVie
Jeffrey Waring	AbbVie
Bridget Riley-Gillis	AbbVie
Anne Lehtonen	AbbVie
Athena Matakidou	AstraZeneca
Jennifer Schutzman	Genentech
Julie Hunkapiller	Genentech
Natalie Bowers	Genentech
John Michon	Genentech
Vinay Mehta	Merck
Andrey Loboda	Merck
Aparna Chhibber	Merck
Heli Lehtonen	Pfizer
Stefan McDonough	Pfizer
Marika Crohns	Sanofi
Diptee Kulkarni	GSK

Opthalmology Group

Kai Kaarniranta LEAD	Kuopio
Joni Turunen	HUS/ Secretary
Terhi Ollila	HUS
Sanna Seitsonen	HUS
Hannu Uusitalo	Tampere
Vesa Aaltonen	Turku
Hannele Uusitalo-Järvinen	PSHP
Marja Luodonpää	Oulu
Nina Hautala	Oulu
Heiko Runz	Biogen
Stephanie Loomis	Biogen
Erich Strauss	Genentech
Natalie Bowers	Genentech
Hao Chen	Genentech
John Michon	Genentech
Anna Podgornaia	Merck
Vinay Mehta	Merck
Dorothee Diogo	Merck
Joshua Hoffman	GSK

Dermatology Group

Kaisa Tasanen LEAD	Oulu
Laura Huilaja	Oulu
Katariina Hannula-Jouppi	HUS
Teea Salmi	Tampere
Sirkku Peltonen	Turku
Leena Koulu	Turku
Ilkka Harvima	Kuopio
Kirsi Kalpala	Pfizer
Ying Wu	Pfizer
David Choy	Genentech
John Michon	Genentech
Nizar Smaoui	AbbVie
Fedik Rahimov	AbbVie
Anne Lehtonen	AbbVie
Dawn Waterworth	GSK

FinnGen Teams

Administration Team

Anu Jalanko	University of Helsinki / FIMM
Risto Kajanne	University of Helsinki / FIMM
Ulrike Lyhs	University of Helsinki / FIMM

Communication

Mari Kaunisto	University of Helsinki / FIMM

Analysis Team

Justin Wade Davis	Abbvie
Bridget Riley-Gillis	Abbvie
Danjuma Quarless	Abbvie
Slavé Petrovski	Astra Zeneca
Jimmy Liu	Biogen
Stephanie Loomis	Biogen
Paola Bronson	Biogen
Robert Yang	Celgene
Joseph Maranville	Celgene
Shameek Biswas	Celgene

Diana Chang	Genentech
Julie Hunkapiller	Genentech
Tushar Bhangale	Genentech
Natalie Bowers	Genentech
Dorothee Diogo	Merck
Emily Holzinger	Merck
Padhraig Gormley	Merck
Xulong Wang	Merck
Xing Chen	Pfizer
Åsa Hedman	Pfizer
Joshua Hoffman	GSK
Clarence Wang	Sanofi
Ethan Xu	Sanofi
Franck Auge	Sanofi
Clement Chatelain	Sanofi
Mitja Kurki	University of Helsinki / FIMM/ Broad Institute
Samuli Ripatti	University of Helsinki / FIMM
Mark Daly	University of Helsinki / FIMM
Juha Karjalainen	University of Helsinki / FIMM/ Broad Institute
Aki Havulinna	University of Helsinki / FIMM
Anu Jalanko	University of Helsinki / FIMM
Kimmo Palin	University of Helsinki
Priit Palta	University of Helsinki / FIMM
Pietro della Briotta Parolo	University of Helsinki / FIMM
Wei Zhou	Broad Institute Susanna
Lemmelä	University of Helsinki / FIMM
Manuel Rivas	University of Stanford
Jarmo Harju	University of Helsinki / FIMM
Aarno Palotie	University of Helsinki / FIMM
Arto Lehisto	University of Helsinki / FIMM
Andrea Ganna	University of Helsinki / FIMM
Vincent Llorens	University of Helsinki / FIMM
Antti Karlsson	Auria BB / Univ. of Turku /VSSHP
Kati Kristiansson	THL BB / THL
Mikko Arvas	Finnish Red Cross Blood Service BB /FHRB
Kati Hyvärinen	Finnish Red Cross Blood Service BB /FHRB
Jarmo Ritari	Finnish Red Cross Blood Service BB /FHRB
Tiina Wahlfors	Finnish Red Cross Blood Service BB /FHRB
Miika Koskinen	Helsinki BB/HUS/Univ Hosp Districts
Olli Carpen	Helsinki BB/HUS/Univ Hosp Districts
Johannes Kettunen	Borealis BB/Univ. of Oulu/PPSHP
Katri Pylkäs	Borealis BB/Univ. of Oulu/PPSHP
Marita Kalaoja	Borealis BB/Univ. of Oulu/PPSHP
Minna Karjalainen	Borealis BB/Univ. of Oulu/PPSHP Tuomo
Mantere	Borealis BB/Univ. of Oulu/PPSHP Eeva
Kangasniemi	Tampere BB/Univ Tampere/PSHP
Sami Heikkinen	Eastern Finland BB/UEFPSSHP
Arto Mannermaa	Eastern Finland BB/UEF/PSSHP
Eija Laakkonen	Central Finland BB /University of Jyväskylä/KSSHP
Juha Kononen	Central Finland BB /University of Jyväskylä/KSSHP

Sample Collection Coordination

Anu Loukola	Helsinki BB/HUS/Univ Hosp Districts

Sample Logistics

Päivi Laiho	THL BB /THL
Tuuli Sistonen	THL BB /THL
Essi Kaiharju	THL BB /THL
Markku Laukkanen	THL BB /THL
Elina Järvensivu	THL BB /THL
Sini Lähteenmäki	THL BB /THL
Lotta Männikkö	THL BB /THL
Regis Wong	THL BB /THL

Registry Data Operations

Kati Kristiansson	THL BB / THL
Hannele Mattsson	THL BB / THL
Susanna Lemmelä	University of Helsinki / FIMM
Tero Hiekkalinna	THL BB / THL
Manuel González Jiménez	THL BB /THL

Genotyping

Kati Donner	University of Helsinki / FIMM

Sequencing Informatics

Priit Palta	University of Helsinki / FIMM

Kalle Pärn	University of Helsinki / FIMM
Javier Nunez-Fontarnau	University of Helsinki / FIMM

Data Management and IT Infrastructure

Jarmo Harju	University of Helsinki / FIMM
Elina Kilpeläinen	University of Helsinki / FIMM
Timo P. Sipilä	University of Helsinki / FIMM
Georg Brein	University of Helsinki / FIMM
Alexander Dada	University of Helsinki / FIMM
Ghazal Awaisa	University of Helsinki / FIMM
Anastasia Shcherban	University of Helsinki / FIMM
Tuomas Sipilä	University of Helsinki / FIMM

Clinical Endpoint Development

Hannele Laivuori	University of Helsinki / FIMM
Aki Havulinna	University of Helsinki / FIMM
Susanna Lemmelä	University of Helsinki / FIMM
Tuomo Kiiskinen	University of Helsinki / FIMM

Trajectory Team

Tarja Laitinen	Tampere University Hospital
Harri Siirtola	University of Tampere
Javier Gracia Tabuenca	University of Tampere

Biobank Directors

Lila Kallio	Auria Biobank
Sirpa Soini	THL Biobank
Jukka Partanen	Blood Service Biobank
Kimmo Pitkänen	Helsinki Biobank Seppo Vainio	Northern Finland Biobank Borealis
Kimmo Savinainen	Tampere Biobank
Veli-Matti Kosma	Biobank of Eastern Finland
Teijo Kuopio	Central Finland Bioban

Supplementary Methods

PARTICIPANTS AND GENOMEWIDE ANALYSES

Participants and thrombin-induced platelet aggregation:

The Caerphilly Prospective Study assessed platelet aggregation induced by full-length thrombin (0.056 unit/mL, Sigma Aldrich), ADP (0.725 µM), and Collagen (42.7 µg/mL) in middle aged males using light transmission aggregometry (LTA) [1]. All participants provided written informed consent. The maximum optical density increase due to platelet aggregation was measured and expressed as a proportion of the difference in optical density between platelet-rich plasma (PRP) and autologous platelet-poor plasma (PPP) using a Rubel-Renaud coaguloaggregometer (Mean=10.54, SD=5.66).

Genotyping, imputation and genetic analyses:

Genotyping was performed with the Affymetrix UK BioBank Axiom array (n=830115 variants) using the Affymetrix Axiom Analysis Suite (AAS) software (https://www.thermofisher.com/us/en/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/axiom-analysis-suite.html) according to the manufacturer's recommended instructions. The starting sample size was 1248 (including 26 duplicates and reference samples). After sample quality control procedures (excluding duplicate/reference samples, samples with genome-wide heterozygosity rate > 6 standard deviations from the mean, sample call rate < 0.97, and possible sex mismatches) and genotyping quality control procedures (excluding variants with Hardy Weinberg Equilibrium $P < 1x10^{-6}$, minor allele frequency (MAF) < 0.01, SNP call rate < 0.97), the subsequent dataset for genotype imputation consisted of 1184 samples and 646137 variants. Imputation of the 22 autosomes was performed using the Haplotype Reference Consortium (HRC) release 1.1, 2016 reference panel (European ancestry) with the University of Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html)[2]. Imputation resulted in 7751469 autosomal single-nucleotide polymorphisms (SNPs) with a MAF > 0.01 and imputation score > 0.3.

To assess any potential population structure, we merged Caerphilly sample data with UK BioBank data (n=6,000 European ancestry samples; n=1,500 East Asian ancestry samples; 7,700 South Asian ancestry samples; 3,300 African ancestry samples) and computed principal components in PLINK. The plot of PC1 and PC2 (**Figure S2**) indicates that all but n=3 of the Caerphilly samples clustered with the European ancestry samples in UK BioBank, indicating little population structure.

Platelet aggregation values for thrombin were square root transformed, collagen were cube-root transformed, and ADP were log-transformed. We performed a genome wide association study (GWAS) using a linear mixed model adjusting for age and medication usage (anticoagulant, antiplatelet, antilipid, hypoglycemics), accounting for genetic relatedness with a genotype relationship matrix. To further assess whether genetic relatedness could contribute to results, we used KING to estimate pairwise kinship coefficients for the Caerphilly samples. In total we identified 63, 19 and 126 pairs of 1^{st}, 2^{nd}, and 3^{rd} order relationships, respectively, involving a total of n=303 unique samples. While our GWAS analysis accounted for relatedness, to further provide a sensitivity analysis of possible family effects, we conducted additional analyses of chromosome 10 after removing the n=303 unique samples who had 1^{st}, 2^{nd}, and/or 3^{rd} order relationships.

We used the Quanto package to estimate statistical power at various combinations of MAF and genotypic effect sizes [3]. We had 80% power to detect alleles with frequency > 10% conferring an effect size of Beta = 0.4. Similarly, we had 99% power to detect > 10% alleles with an effect size of Beta = 0.7. A more extended range of power calculations is given in **Table S16**. There were no missing data in the model covariates thus the complete sample size for GWAS was n=1184. The GWAS was performed with EMMAX[4]. A significance threshold of $P < 7x10^{-9}$ was adopted to account for all variants tested. Conditional analyses adjusting for the strongest peak SNP in *GRK5*, rs10886430, were conducted by adding the SNP dosage as a covariate to the base model. Conditional analyses were performed with EMMAX[4]. EMMAX reports a pseudo-heritability estimate: the proportion of phenotype variation explained by the marker-based kinship matrix[4]. By comparing the main GWAS heritability estimate (54.36%) with the estimate obtained in the conditional analysis

(35.98%), where the only difference was adjusting for rs10886430, we find this single SNP accounts for roughly 18.38% in thrombin phenotype variation.

CAUSAL ANALYSIS OF GRK5

We integrated our GWAS with the platelet RNA and expression 1 (PRAX1) platelet eQTL dataset (**Table S1** in [5]) utilizing the SMR package [6] which implements a Mendelian randomization approach to test for a joint association in GWAS and eQTL data. SMR further measures association profiles of nearby co-inherited DNA variants of the input datasets for a heterogeneity in dependent instruments (HEIDI) test. The 1000 Genomes Project Phase 3 version 5 reference panel [7] was used to calculate linkage disequilibrium (LD) for this analysis. In the case where profiles are dissimilar (a significant HEIDI test), identified GWAS and eQTL signals are less likely to be driven by a common variant, and the co-association can be explained by genetic linkage where the top *cis*-eQTL is likely in LD with two distinct causal variants, one affecting the outcome and the other affecting gene expression. A significant SMR test cannot directly prove causality, but does support a mechanism of pleiotropy [6]. We further conducted eQTL lookups for the rs10886430 variant in an independent platelet study imputed with UK10K + 1000G Phase 3 haplotype panels [8] as well as 44 tissue types profiled in the Genotype-Tissue Expression (GTEx) project version 7 portal (https://gtexportal.org).

We performed a series of Bayesian co-localization analyses of our thrombin-induced platelet aggregation GWAS with different quantitative traits: hematologic, platelet function, thrombin generation potential, and eQTL. The datasets comprising the different quantitative traits are listed in **Table S1.** First, co-localization analyses were conducted with multiple blood cell types in the UK BioBank/INTERVAL study meta-analysis [9]: mean platelet volume (MPV), platelet cell count (PLT), platelet distribution width (PDW), red blood cell count (RBC), and white blood cell count (WBC). Second, analyses were conducted with ADP and collagen-stimulated platelet aggregation phenotypes from our Caerphilly cohort. Third, analyses were conducted with additional published platelet aggregation traits for agonists ADP, collagen and epinephrine [10].

Fourth, analyses were conducted with three traits for thrombin generation potential [11]. Fifth, we conducted co-localization analyses of our thrombin GWAS with three additional eQTL datasets: 44 GTEx tissues (project version 7), 6 blood cell types [8], and vascular endothelial cells [12]. The co-localization analyses was performed with the *coloc* package version 3.1 in R (https://github.com/chr1swallace/coloc). For each pair of thrombin and blood cell traits, thrombin generation traits, or platelet aggregation traits we tested all shared bi-allelic SNPs (MAF > 0.01) in the ~1.8 Mb approximately independent LD block[13] containing the *GRK5* rs10886430 variant. The LD block file used was 1000G phase 1 European hg19 (https://bitbucket.org/nygcresearch/ldetect-data/src/master/EUR). For colocalization analyses involving eQTL traits, we tested all shared bi-allelic SNPs (MAF > 0.01) up to 1-MB upstream or downstream from the *GRK5* gene locus.

The vascular endothelial cell eQTL described above were derived from a new analysis of previously published data[12]. Briefly, human aortic endothelial cells (HAEC) were isolated from aortic explants of 147 heart transplant donors in the UCLA transplant program. Gene expression was measured using the Affymetrix HT HG-U133A microarray, SNP genotyping was performed using the Affymetrix SNP 6.0 microarray and genotypes were imputed using the 1000 genomes reference panel. We calculated the association of SNPs with *GRK5* expression using FaST-LMM[14]. Genotype dosages from all autosomal chromosomes and expression data were used. For FaST-LMM implementation, to improve power when testing all the variants on chromosome N for association, we constructed the kinship matrix by using the variants from all other chromosomes besides N. This procedure allowed us to include the variant being tested for association in the regression equation only once.

Mendelian Randomization (MR) analysis was conducted using the rs10886430 G allele as the genetic instrument and thrombin-induced platelet aggregation as exposure in separate analyses for: (A) 9 pulmonary, stroke, or heart disease outcomes from the UK BioBank, (B) 4 stroke outcomes from MEGASTROKE, and (C) 13 circulatory system disease outcomes in FinnGen using the *MendelianRandomization* package version 0.3.0 (https://cran.r-project.org/web/packages/MendelianRandomization/index.html) in R. We implemented a fixed-

effect model using the ratio method for a single genetic variant. From Gene ATLAS release version 2 (http://geneatlas.roslin.ed.ac.uk) we downloaded UK BioBank GWAS summary statistics for the following nine outcomes: I26 Pulmonary embolism (clinical), I26-I28 Pulmonary heart disease (clinical), 1083 cerebrovascular disease (self-reported), 1098 stroke (self-reported), 1085 venous thromboembolic disease (self-reported), 1112 deep venous thrombosis (dvt) (self-reported), I21 acute myocardial infarction (clinical), I63 Cerebral infarction (clinical), I50 Heart Failure (clinical) [15]. Outcomes with prefix "I" are derived from ICD9 and / or ICD10 codes. The number of cases for each outcome is reported in **Table S12**. The Gene ATLAS release version 2 GWAS were conducted on 452264 individuals of European ancestry in the UK Biobank in a Linear Mixed Model (LMM) framework including as fixed effects sex, array batch, UK Biobank Assessment Center, age, age$_2$, and the leading 20 genomic principal components as computed by UK Biobank. Polygenic effect that captures the population structure was fitted as a random effect and corrected for using a leave-one-chromosome-out approach [15]. Additional MR analyses were performed as described using four stroke outcomes from the MEGASTROKE consortium. From MEGASTROKE (http://megastroke.org/index.html) we downloaded fixed-effects trans-ethnic meta-analyses for any stroke and three stroke subtypes: any ischemic stroke, large artery stroke, and cardioembolic stroke [16]. The number of cases per stroke outcome is reported in **Table S13** and the number of controls in the meta-analyses was 454450 [16]. The number of cases and controls for each outcome in FinnGen are reported in **Table S14**. Prior to conducting MR we transformed the linear mixed-model beta estimates from Gene ATLAS and MEGASTROKE to odds ratios (OR) [17]. With FinnGen having utilized SAIGE for binary outcomes[18] , the transformation of beta estimates was exponentiation. The transformed OR and the corresponding standard errors from the Gene ATLAS, MEGASTROKE, FinnGen GWAS are reported in **Table S12, Table S13** and **Table S14,** respectively. For the MR analyses we obtained the OR of the causal estimates, associated 95% confidence intervals, standard errors, and p-values for each outcome.

REGULATORY FUNCTION

Epigenetic Regulatory Data Integration:

We obtained relevant epigenetic regulatory maps from multiple sources. We downloaded available ChIP-seq datasets produced by the ENCODE processing pipeline [19] for cell models K562 and primary human peripheral blood-derived erythroblasts. For inclusion criteria, we considered only those datasets that were officially released and audit-compliant (that is, no entries in ENCODE's "Audit NOT_COMPLIANT" metadata field). Specifically, we downloaded the hg19-derived optimal irreproducible discovery rate (IDR) thresholded peaks (based on two isogenic replicates) for each dataset meeting the above criteria. We downloaded cultured, primary megakaryocyte datasets for chromatin accessibility (Dnase hypersensitivity), histone marks H3K4me1, H3K27ac from the BLUEPRINT Epigenome project [20]. Files on genome build hg38 were lifted over to hg19. We downloaded published megakaryocyte enhancer regions predicted by chromatin segmentation of BLUEPRINT ChIP-seq datasets[21]. To investigate enhancer RNAs (eRNA) we downloaded transcription start site (TSS) maps of K562 cells identified by a Hidden Markov Model from global nuclear run-on sequencing enriched for 5'-capped (m^7G) RNAs (GRO-cap) data[22]. We intersected this set of regulatory maps with a VCF file containing the GRK5 rs10886430 variant and retained overlapping elements for curation and further analyses.

For our Protein Network Analysis we constructed a network of the transcriptional regulators in ENCODE K562 and primary human peripheral blood-derived erythroblasts that bind the rs10886430 variant site using STRING version 10.5 (http://string-db.org/). We applied the Markov Chain Clustering (MCL) algorithm to identify potential functional relationships between network members.

Lentivirus production:

The following vectors were used in this protocol: pInducer-21 lentiviral vector (Addgene), pMD2.G envelope plasmid (Addgene), psPAX2 packaging plasmid (Addgene). 293T-17 cells (ATCC) were maintained in a DMEM media supplemented with 5% FBS (Sigma). Cultures were kept in 5% CO2 in a humidified environment at 37^0C and passaged every 1 to 2 days. Lentiviral plasmids GATA1 and GATA2 were cloned into

pInducer-21 lentiviral vector. For lentivirus production, 293T-17 cells (ATCC) were transfected with third generation packaging plasmids pMD2.G and psPAX2 (Addgene) and lentiviral plasmids (GATA1 and GATA2). Viruses were harvested 48 h post transfections and concentrated by ultracentrifugation at 24,000 rpm for 2 h at 4 °C. Viruses were titrated by serial dilution on 293T cells using GFP as an indicator.

Conditional overexpression of GATA1 and GATA2 in HEK293 cells:

Lentiviral transductions were carried out in a total volume of 200 μL. The multiplicity of infection (MOI) for GATA1 and GATA2 was 5. Virus was concentrated onto the HEK293 cells by centrifugation at 2500 rpm for 30 min at RT. Infections were carried out for 24 hours. After gene transfer, GATA1 and GATA2 cells were cultured in DMEM and 5% FBS (Gemini/Benchmark) for 1 day, followed by addition of Dox (μg/mL) (Sigma). Cultures were kept at a density of $<1 \times 10_6$ cells/mL in 5% CO2 in a humidified environment at 37_0C and passaged every 1 to 2 days.

Enhancer function reporter assay:

The following vectors were used in this protocol: pGL4.23[*luc2*/minP] luciferase reporter (Promega), pGL4.74[*hRluc*/TK] control vector (Promega). A 284-bp non-coding putative *GRK5* enhancer region (10q26.11) containing a GATA1 transcription factor binding site and the major "A" allele of the rs10886430 variant was cloned into the pGL4.23 luciferase vector. We created three modified constructs to assess functionality of the locus: deletion of the GATA1 site, deletion of four bases flanking the variant (AGTG), and knock-in of the rs10886430 minor "G" allele. The modified constructs were generated with the QuikChange™ Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer's instructions using the mutagenesis primers listed in methods section ***Oligonucleotides***. Constructs were sequenced to confirm the expected genotype and to ensure no off-target mutations were introduced. Dual luciferase reporter assays were performed as described previously with minor modifications [23]. Briefly, GATA1- or GATA2-overexpressing HEK293 cells, mega-erythroid K562 cells, or Human Umbical Vein Endothelial Cells (HUVEC) were co-transfected with one of four pGL4.23 luciferase vectors described above as well pGL4.74 control according to the manufacturer's instructions. Firefly

and Renilla luciferase reporter activity of cell extracts were measured using the Dual-Glo Luciferase Assay System (Promega) on a microplate reader according to the manufacturer's instructions. Each treatment was performed in duplicate and the experiment was repeated three times.

PLATELET FUNCTION siRNA AND INHIBITOR EXPERIMENTS

imMKCL culture:

Immortalized megakaryocyte progenitor cell lines (imMKCLs) were maintained in presence of 5 µg/mL doxycycline (DOX) as previously described [24]. Removal of DOX results in imMKCL maturation and the generation of platelets after six days [24].

siRNA transfection assay:

The imMKCLs growing in the differentiation medium (without DOX) on day 4 were seeded in a 24-well plate, 24 hours prior to transfection. The cells were transfected with 500nM GRK5 or control eGFP siRNA oligonucleotides (see methods section ***Oligonucleotides***) using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. 48 hours post-transfection, the cells were harvested for RT-qPCR and *in vitro* flow cytometric analysis of platelets.

RNA extraction, reverse transcription, and RT-qPCR:

RNA extraction was performed using an RNAeasy kit (Qiagen). Reverse transcription was performed using Superscript III (Invitrogen), using Oligo (dT) 15 primer. Quantitative PCR was performed in triplicate with SYBR Green and CFX96 real-time PCR detection system (Bio-rad). Target transcript abundance was calculated relative to *ACTB* (reference gene) using the $2^{-\Delta\Delta CT}$ method. Gene specific primer pairs are present in methods section ***Oligonucleotides***.

Human blood collection and platelet isolation:

A written informed consent (in accordance with the Declaration of Helsinki prior to participation in this Boston Children's Hospital IRB-approved study) was obtained from healthy donor volunteers prior to blood draw.

Blood was collected by venipuncture with a 21-gauge butterfly needle into evacuated tubes containing 3.2% sodium citrate. Blood was drawn from healthy volunteers who were free from antiplatelet agents and non-steroidal anti-inflammatory drugs for 10 days prior to the donation. The same phlebotomist performed all the blood draws. Complete blood cell counts were determined with a Sysmex XN-1000 Hematology Analyzer. Platelet-rich plasma (PRP) was collected after centrifugation at 200xg (15 min, RT).

Flow cytometric analysis of platelet cell markers:

Flow cytometric analysis of Day 6 platelets from GRK5 siRNA knockdown experiments was performed on Day 6 cultures growing in differentiation medium without DOX. Briefly, cell aliquots were incubated for 20 min with fluorescently labeled monoclonal antibodies and either 20 μM adenosine 5'-diphosphate (ADP) plus 20 μM thrombin receptor-activating peptide 6 (TRAP6) or control HEPES-buffered Tyrode's solution. The antibodies used were: phycoerythrin (PE)-mouse anti-human anti–P-selectin (CD62P) antibody (BD Pharmingen; Cat# 555524); fluorescein isothiocyanate (FITC)- mouse anti-human antibody PAC1 (BD Pharmingen; Cat# 340507), which only binds to the glycoprotein IIb/IIIa (GPIIb/IIIa) complex of the activated platelets or near the platelet fibrinogen receptor; and PE-Cy5–mouse anti-human anti-CD42b (GPIb) monoclonal antibody (BD Pharmingen; Cat# 551141). Staining was fixed in 0.5% paraformaldehyde solution. Data were collected on a FACSCalibur flow cytometer (BD Biosciences) and analyzed with FlowJo software (FlowJo LLC). Mean fluorescence intensity (geometric mean) of PE, FITC and PE-%positive, FITC-%positive events were determined.

For PRP pharmacological inhibition experiments, PRP samples were incubated with varying concentrations (0.0098 – 0.781 μM) of pan-GRK inhibitor CCG215022 (MedChemExpress) or vehicle for 45 min. To treated PRP were then added two fluorescently labeled antibodies described above (PE-conjugated CD62P and APC-conjugated CD42b) along with either 50 μM PAR4 Activating Peptide (PAR4-AP) (amino acid sequence AYPGKF, TOCRIS Bioscience), 20 μM TRAP-6 (Thrombin Receptor Agonist Peptide, amino acid sequence SFLLRN which binds PAR1), or vehicle for 20 minutes. Subsequent PRP pharmacological inhibition

experiments incubated PRP samples with 0.781 μM CCG215022 or vehicle for 45 min followed by varying concentrations of one of three platelet agonists: PAR4-AP (0, 1, 20, 50 μM), TRAP-6 (0, 1, 10, 20 μM), ADP (0, 1, 10, 20 μM), or vehicle for 20 min. Samples were fixed with 1% paraformaldehyde. Data were collected on a LSRII flow cytometer (BD Biosciences) and analyzed with FlowJo software (FlowJo LLC). Mean fluorescence intensity (geometric mean) of PE and PE-%positive events were determined.

Oligonucleotides:

	Forward Primer (5′-3′)	**Reverse Primer (5′-3′)**	**Amplicon size**
qRT-PCR			
GAPDH	ACCCACTCCTCCACCTTTGA	CTGTTGCTGTAGCCAAATTCGT	101
GATA1	CTACACCAGGTGAACCGGC	CTTTTCCAGATGCCTTGCGG	76
GATA2	GCAGAACCGACCACTCATCA	AATTTGCACAACAGGTGCCG	74
GRK5	GAGCTGGAAAACATCGTGGC	CTTGCTTTTCCCTTTGCGCT	81
siRNA			
GRK5	GGAAATTATGACCAAGTACCT		
GRK5	GCAGATCCTCGAGAACGTCAA		
eGFP	GCCACAACGTCTATATCAT		
Mutatagenesis primers			
A90G	GTGAACGTTGGAGAAGGTGG CTTAGTCATG	CATGACTAAGCCACCTTCTCCA ACGTTCAC	
AGTG deletion	GAACGTTGGAGAAGCTTAGTC ATGAC	GTCATGACTAAGCTTCTCCAAC GTTC	
GATA1 motif deletion	GAACAATGCAGTTCTCATCCGT TGTTGGGTGATATG	CATATCACCCAACAACGGATGA GAACTGCATTGTTC	

Supplementary References

1. Elwood, P.C., Renaud, S., Sharp, D.S., Beswick, A.D., O'Brien, J.R., and Yarnell, J.W. (1991). Ischemic heart disease and platelet aggregation. The Caerphilly Collaborative Heart Disease Study. Circulation 83, 38-44.
2. Das, S., Forer, L., Schonherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation genotype imputation service and methods. Nat Genet 48, 1284-1287.
3. Gauderman, W.J. (2002). Sample Size Requirements for Association Studies of Gene-Gene Interaction. American Journal of Epidemiology 155, 478-484.
4. Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.-y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nature Genetics 42, 348.
5. Simon, Lukas M., Chen, Edward S., Edelstein, Leonard C., Kong, X., Bhatlekar, S., Rigoutsos, I., Bray, Paul F., and Shaw, Chad A. (2016). Integrative Multi-omic Analysis of Human Platelet eQTLs Reveals Alternative Start Site in Mitofusin 2. The American Journal of Human Genetics 98, 883-897.
6. Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., et al. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nature Genetics 48, 481.
7. The Genomes Project, C., Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., et al. (2015). A global reference for human genetic variation. Nature 526, 68.
8. Momozawa, Y., Dmitrieva, J., Théâtre, E., Deffontaine, V., Rahmouni, S., Charloteaux, B., Crins, F., Docampo, E., Elansary, M., Gori, A.-S., et al. (2018). IBD risk loci are enriched in multigenic regulatory modules encompassing putative causative genes. Nature Communications 9, 2427.
9. Astle, W.J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A.L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M.A., et al. (2016). The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 167, 1415-1429.e1419.
10. Johnson, A.D., Yanek, L.R., Chen, M.-H., Faraday, N., Larson, M.G., Tofler, G., Lin, S.J., Kraja, A.T., Province, M.A., Yang, Q., et al. (2010). Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nature Genetics 42, 608.
11. Rocanin-Arjo, A., Cohen, W., Carcaillon, L., Frère, C., Saut, N., Letenneur, L., Alhenc-Gelas, M., Dupuy, A.-M., Bertrand, M., Alessi, M.-C., et al. (2014). A meta-analysis of genome-wide association studies identifies ORM1 as a novel gene controlling thrombin generation potential. Blood 123, 777-785.
12. Erbilgin, A., Civelek, M., Romanoski, C.E., Pan, C., Hagopian, R., Berliner, J.A., and Lusis, A.J. (2013). Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. Journal of Lipid Research 54, 1894-1905.
13. Berisa, T., and Pickrell, J.K. (2016). Approximately independent linkage disequilibrium blocks in human populations. Bioinformatics 32, 283-285.
14. Lippert, C., Listgarten, J., Liu, Y., Kadie, C.M., Davidson, R.I., and Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. Nature Methods 8, 833-835.
15. Canela-Xandri, O., Rawlik, K., and Tenesa, A. (2018). An atlas of genetic associations in UK Biobank. Nature Genetics 50, 1593-1599.
16. Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.-K., van der Laan, S.W., Gretarsdottir, S., et al. (2018). Multiancestry genome-wide association study of

520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nature Genetics 50, 524-537.
17. Lloyd-Jones, L.R., Robinson, M.R., Yang, J., and Visscher, P.M. (2018). Transformation of Summary Statistics from Linear Mixed Model Association on All-or-None Traits to Odds Ratio. Genetics 208, 1397.
18. Mars, N., Koskela, J.T., Ripatti, P., Kiiskinen, T.T.J., Havulinna, A.S., Lindbohm, J.V., Ahola-Olli, A., Kurki, M., Karjalainen, J., Palta, P., et al. (2020). Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. Nat Med 26, 549-557.
19. Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Research 22, 1798-1812.
20. Stunnenberg, H.G., Abrignani, S., Adams, D., de Almeida, M., Altucci, L., Amin, V., Amit, I., Antonarakis, S.E., Aparicio, S., Arima, T., et al. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell 167, 1145-1149.
21. Petersen, R., Lambourne, J.J., Javierre, B.M., Grassi, L., Kreuzhuber, R., Ruklisa, D., Rosa, I.M., Tomé, A.R., Elding, H., van Geffen, J.P., et al. (2017). Platelet function is modified by common sequence variation in megakaryocyte super enhancers. Nature Communications 8, 16058.
22. Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. Nature Genetics 46, 1311.
23. Bhan, A., Hussain, I., Ansari, K.I., Bobzean, S.A.M., Perrotti, L.I., and Mandal, S.S. (2014). Histone Methyltransferase EZH2 Is Transcriptionally Induced by Estradiol as Well as Estrogenic Endocrine Disruptors Bisphenol-A and Diethylstilbestrol. Journal of Molecular Biology 426, 3426-3441.
24. Nakamura, S., Takayama, N., Hirata, S., Seo, H., Endo, H., Ochi, K., Fujita, K.-i., Koike, T., Harimoto, K.-i., Dohda, T., et al. (2014). Expandable Megakaryocyte Cell Lines Enable Clinically Applicable Generation of Platelets from Human Induced Pluripotent Stem Cells. Cell Stem Cell 14, 535-548.