Title

Platelet-related variants identified by exome chip meta-analysis in 157,293 individuals

Authors

John D. Eicher[1]*, Nathalie Chami[2,3]*, Tim Kacprowski[4,5]*, Akihiro Nomura[6,7,8,9,10]*, Ming-Huei Chen[1], Lisa R. Yanek[11], Salman M. Tajuddin[12], Ursula M. Schick[13,14], Andrew J. Slater[15,16], Nathan Pankratz[17], Linda Polfus[18], Claudia Schurmann[13], Ayush Giri[19], Jennifer A. Brody[20], Leslie A. Lange[21], Ani Manichaikul[22], W. David Hill[23,24], Raha Pazoki[25], Paul Elliot[26], Evangelos Evangelou[26,27], Ioanna Tzoulaki[26], He Gao[26], Anne-Claire Vergnaud[26], Rasika A. Mathias[28,29], Diane M. Becker[29], Lewis C. Becker[29,30], Amber Burt[31], David R. Crosslin[32], Leo-Pekka Lyytikäinen[33,34], Kjell Nikus[35,36], Jussi Hernesniemi[33,34], Mika Kähönen[37,38], Emma Raitoharju[33,34], Nina Mononen[33,34], Olli Raitakari[39,40], Terho Lehtimäki[33,34], Mary Cushman[41], Neil A. Zakai[41], Deborah A. Nickerson[42], Laura M. Raffield[21], Rakale Quarells[43], Cristen J. Willer[44,45,46], Gina M. Peloso[6,7,47], Goncalo R. Abecasis[48], Dajiang J. Liu[49], Global Lipids Genetics Consortium, Panos Deloukas[50,51], Nilesh J. Samani[52,53], Heribert Schunkert[54,55], Jeanette Erdmann[56,57], CARDIoGRAM Exome Consortium, Myocardial Infarction Genetics Consortium, Myriam Fornage[58], Melissa Richard[58], Jean-Claude Tardif[2,3], John D. Rioux[2,3], Marie-Pierre Dube[2,3], Simon de Denus[3,59], Yingchang Lu[13], Erwin P. Bottinger[13], Ruth J. F. Loos[13], Albert Vernon Smith[60,61], Tamara B. Harris[62], Lenore J. Launer[62], Vilmundur Gudnason[60,61], Digna R. Velez Edwards[63], Eric S. Torstenson[19], Yongmei Liu[64], Russell P. Tracy[65], Jerome I. Rotter[66,67], Stephen S. Rich[22], Heather M. Highland[68,69], Eric Boerwinkle[70,71], Jin Li[72], Ethan Lange[21,73], James G. Wilson[74], Evelin Mihailov[75], Reedik Mägi[75], Joel Hirschhorn[7,76], Andres Metspalu[75], Tõnu Esko[7,75], Caterina Vacchi-Suzzi[77], Mike A. Nalls[78], Alan B. Zonderman[12], Michele K. Evans[12], Gunnar Engström[79,80], Marju Orho-Melander[79,80], Olle Melander[79,80], Michelle L.

O'Donoghue[81], Dawn M. Waterworth[82], Lars Wallentin[83], Harvey D. White[84], James S. Floyd[20], Traci M. Bartz[85], Kenneth M. Rice[85], Bruce M. Psaty[86,87], J.M. Starr[23], David C. M. Liewald[23,24], Caroline Hayward[88], Ian J. Deary[23,24], Andreas Greinacher[89], Uwe Völker[4,5], Thomas Thiele[90], Henry Völzke[5,91], Frank J. A. van Rooij[25], André G. Uitterlinden[25,92,93], Oscar H. Franco[25], Abbas Dehghan[25], Todd L. Edwards[94], Santhi K. Ganesh[95], Sekar Kathiresan[6,7,8,9], Nauder Faraday[96]*, Paul L. Auer[97]*, Alex P. Reiner[98,99]*, Guillaume Lettre[2,3]*, Andrew D. Johnson[1]*

*These authors contributed equally to this study.

Affiliations

[1]Population Sciences Branch, National Heart Lung and Blood Institute, The Framingham Heart Study, Framingham, MA, USA, 01702
[2]Department of Medicine, Université de Montréal, Montréal, Québec, Canada, H3T 1J4
[3]Montreal Heart Institute, Montréal, Québec, Canada, H1T 1C8
[4]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald and Ernst-Mortiz-Arndt University, Greifswald, Germany, 17489
[5]DZHK German Centre for Cardiovascular Research, partner site Greifswald, Greifswald, Germany, 13347
[6]Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA, 02114
[7]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA, 02142
[8]Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA, 02114
[9]Department of Medicine, Harvard Medical School, Boston, MA, USA, 02115
[10]Division of Cardiovascular Medicine, Kanazawa University Graduate School of Medical Science, Kanazawa, Ishikawa, Japan, 9200942
[11]Department of Medicine, Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, 21205
[12]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA, 21224
[13]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA, 10029
[14]The Genetics of Obesity and Related Metabolic Traits Program, Icahn School of Medicine at Mount Sinai, New York, NY, USA, 10029
[15]Genetics, Target Sciences, GlaxoSmithKline, Research Triangle Park, NC, USA, 27709
[16]OmicSoft Corporation, Cary, NC, USA, 27513
[17]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN, USA, 55454
[18]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA, 77030
[19]Division of Epidemiology, Institute for Medicine and Public Health, Vanderbilt University Nashville, TN, USA, 37235
[20]Department of Medicine, University of Washington, Seattle, WA, USA, 98101
[21]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA, 27514
[22]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA, 22908
[23]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK, EH8 9JZ
[24]Department of Psychology, University of Edinburgh, Edinburgh, UK, EH8 9JZ
[25]Department of Epidemiology, Erasmus MC, Rotterdam, Netherlands, 3000
[26]Department of Epidemiology and Biostatistics, MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK, W2 1PG

[27]Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece, 45110
[28]Department of Medicine, Divisions of Allergy and Clinical Immunology, Johns Hopkins University School of Medicine, Baltimore, MD, USA, 21205
[29]Department of Medicine, Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, 21205
[30]Department of Medicine, Divisions of Cardiology, Johns Hopkins University School of Medicine, Baltimore, MD, USA, 21205
[31]Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA, USA, 98195
[32]Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA, USA, 98105
[33]Fimlab Laboratories, Tampere, Finland, 33521
[34]Department of Clinical Chemistry, University of Tampere, Tampere, Finland, 33520
[35]Department of Cardiology, Heart Hospital, Tampere University Hospital, Tampere, Finland, 33521
[36]Department of Cardiology, University of Tampere, Tampere, Finland, 33520
[37]Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland, 33521
[38]Department of Clinical Physiology, University of Tampere, Tampere, Finland, 33520
[39]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland, 20520
[40]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland, 20520
[41]Departments of Medicine and Pathology, University of Vermont College of Medicine, University of Vermont College of Medicine, Burlington, VT, USA, 05405
[42]Department of Genome Sciences, University of Washington, Seattle, WA, USA, 98105
[43]Morehouse School of Medicine, Social Epidemiology Research Center, Cardiovascular Research Institute, Atlanta, GA, USA, 30310
[44]Department of Internal Medicine, Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI, USA, 48108
[45]Department of Computational Medicine and Bioinformatics, Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA, 48108
[46]Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA, 48108
[47]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA, 02118
[48]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA, 48108
[49]Department of Public Health Sciences, College of Medicine, Pennsylvania State University, Hershey, PA, USA, 17033
[50]William Harvey Research Institute, Queen Mary University London, London, UK, E1 4NS
[51]Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia, 21589
[52]Department of Cardiovascular Sciences, University of Leicester, Leicester, UK, LE1 7RH
[53]NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, UK, LE3 9QP
[54]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, Munich, Germany, 80333
[55]Deutsches Herzzentrum München Technische Universität München, Munich, Germany, 80333
[56]Institute for Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany, 23562
[57]DZHK (German Research Centre for Cardiovascular Research), partner site Hamburg/Lübeck/Kiel, Lübeck, Germany, 23562
[58]Institute of Molecular Medicine, The University of Texas Health Science Center at Houston, Houston, TX, USA, 77030
[59]Faculty of Pharmacy, Université de Montréal, Montréal, Québec, Canada, H3T 1J4
[60]Icelandic Heart Association, Kopavogur, Iceland, IS-201
[61]Faculty of Medicine, University of Iceland, Reykjavik, Iceland, 101
[62]Laboratory of Epidemiology, Demography, and Biometry, National Institute on Aging, Intramural Research Program, National Institutes of Health, Bethesda, MD, USA, 21224
[63]Vanderbilt Epidemiology Center, Department of Obstetrics & Gynecology, Institute for Medicine and Public Health, Vanderbilt Genetics Institute, Vanderbilt University Nashville, TN, USA, 37235
[64]Center for Human Genetics, Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA, 27157
[65]Departments of Pathology and Laboratory Medicine and Biochemistry, University of Vermont College of Medicine, Colchester, VT, USA, 05446
[66]Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, Torrance, CA, USA, 90502
[67]Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA, USA, 90502
Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA, 22908
[68]The University of Texas School of Public Health, The University of Texas Graduate School of Biomedical Sciences at Houston, The University of Texas Health Science Center at Houston, Houston, TX, USA, 77030
[69]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 27514
[70]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA, 77030
[71]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA, 77030
[72]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Palo Alto, CA, USA, 94305

[73]Department of Biostatistics University of North Carolina at Chapel Hill, University of North Carolina, Chapel Hill, NC, USA, 27514
[74]Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA, 39216
[75]Estonian Genome Center, University of Tartu, Tartu, Estonia, 51010
[76]Department of Endocrinology, Boston Children's Hospital, Boston, MA, USA, 02115
[77]Department of Family, Population and Preventive Medicine, Stony Brook University, Stony Brook, NY, USA, 11794
[78]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA, 21224
[79]Department of Clinical Sciences Malmö, Lund University, Malmö, Sweden, 221 00
[80]Skåne University Hospital, Malmö, Sweden, 222 41
[81]Cardiovascular Division, Brigham and Women's Hospital, Boston, MA, USA, 02115
[82]Genetics, Target Sciences, GlaxoSmithKline, King of Prussia, PA, USA, 19406
[83]Department of Medical Sciences, Cardiology, and Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden, 751 05
[84]Green Lane Cardiovascular Service, Auckland City Hospital and University of Auckland, Auckland, New Zealand, 1142
[85]Department of Biostatistics, University of Washington, Seattle, WA, USA, 98195
[86]Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA, USA, 98105
[87]Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA, 98105
[88]Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK, EH4 2XU
[89]Institute for Immunology and Transfusion Medicine, University Medicine Greifswald, Greifswald, Germany, 17475
[90]Institute for Immunology and Transfusion Medicine, University Medicine Greifswald, Greifswald, Germany, 17475
[91]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany, 13347
[92]Department of Internal Medicine, Erasmus MC, Rotterdam, Netherlands, 3000
[93]Netherlands Consortium for Healthy Ageing (NCHA), Rotterdam, Netherlands, 3015
[94]Division of Epidemiology, Department of Medicine, Institute for Medicine and Public Health, Vanderbilt Genetics Institute, Vanderbilt University Nashville, TN, USA, 37235
[95]Departments of Internal and Human Genetics, University of Michigan, Ann Arbor, MI, USA, 48108
[96]Department of Anesthesiology & Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA, 21205
[97]Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI, USA, 53205
[98]Department of Epidemiology, University of Washington, Seattle, WA, USA, 98105
[99]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, 98109

Corresponding Author:

Andrew D. Johnson

Tenure Track Investigator

Population Sciences Branch

National Heart, Lung, and Blood Institute

The Framingham Heart Study

73 Mt. Wayte Ave. Suite #2

Framingham, MA 01702

Email: johnsonad2@nhlbi.nih.gov

Phone: (508)-663-4082

Abstract

Platelet production, maintenance, and clearance are tightly controlled processes indicative of platelets' important roles in hemostasis and thrombosis. Platelets are common targets for primary and secondary prevention of several conditions. They are monitored clinically by complete blood counts, specifically with measurements of platelet count (PLT) and mean platelet volume (MPV). Identifying genetic effects on PLT and MPV can provide mechanistic insights into platelet biology and their role in disease. Therefore, we formed the Blood Cell Consortium (BCX) to perform a large-scale meta-analysis of exome chip association results for PLT and MPV in 157,293 and 57,617 individuals, respectively. Using the low-frequency/rare coding variant enriched exome chip platform, we sought to identify genetic variants associated with PLT and MPV. In addition to confirming 47 known PLT and 20 known MPV associations, we identified 32 PLT and 18 MPV associations not previously observed in the literature across the allele frequency spectrum, including rare large effect (*FCER1A*), low-frequency (*IQGAP2*, *MAP1A*, *LY75*), and common (*ZMIZ2*, *SMG6*, *PEAR1*, *ARFGAP3*/*PACSIN2*) variants. Several variants associated with PLT/MPV (*PEAR1*, *MRVI1*, *PTGES3*) were also associated with platelet reactivity. In concurrent BCX analyses, there was overlap of platelet associated variants with red (*MAP1A*, *TMPRSS6*, *ZMIZ2*) and white blood cell (*PEAR1*, *ZMIZ2*, *LY75*) traits, suggesting common regulatory pathways with shared genetic architecture among these hematopoietic lineages. Our large-scale exome chip effort successfully identified numerous previously undocumented associations with platelet traits and further indicates that several complex quantitative hematological, lipid, and cardiovascular traits share genetic factors.

Abstract Word Count: 243

Introduction

The number and size of circulating blood cells are determined by multiple genetic and environmental factors, and abnormal values are a common manifestation of human disease. The three major cell types—red blood cells (RBCs), white blood cells (WBCs), and platelets—have distinct biological roles, with platelets serving as important mediators of hemostasis and wound healing. Platelet count (PLT) and mean platelet volume (MPV), a measure of platelet size, are clinical blood tests that are used to screen for and diagnose disease. A number of well-described rare genetic disorders, including Bernard-Soulier Syndrome (MIM: 231200), Glanzmann's Thrombasthenia (MIM: 273800), and Wiskott-Aldrich Syndrome (MIM: 301000), as well as common conditions such as acute infection are characterized by abnormalities in the number, size, and/or reactivity of circulating blood platelets. MPV has also been reported to be an independent risk factor for myocardial infarction (MI) in population-based studies[1]. Accordingly, anti-platelet medications including aspirin and ADP/$P2Y_{12}$ receptor blockers such as clopidogrel and GIIb/IIIa inhibitors that reduce platelet reactivity are commonly used in the primary and secondary prevention of several cardiovascular conditions, including stroke and MI[2, 3]. Investigating the biological mechanisms that govern platelet number (PLT) and size (MPV) can provide insights into platelet development and clearance, and has the potential to enhance our understanding of human diseases.

Genome-wide association studies (GWAS) have successfully identified numerous loci where variants are associated with PLT and MPV[4-13]. To date, the largest GWAS of PLT (n=66,867) and MPV (n=30,194) identified 68 distinct loci[8]. Subsequent functional experiments of several identified genes, including *ARHGEF3* (MIM: 612115), *DNM3* (MIM: 611445), *JMJD1C* (MIM: 604503), and *TPM1* (MIM: 191010), demonstrated their roles in hematopoiesis

and megakaryopoesis[8, 14], as well as the potential of human genetic association methods to identify genetic factors that functionally contribute to platelet biology and dysfunction in disease.

Despite these successes, much of the heritability of these traits remains unexplained[15]. GWAS of PLT and MPV have largely focused on more common (minor allele frequency [MAF] > 0.05) genetic variation, with many of the associated markers located in intronic or intergenic regions. The examination of rare (MAF < 0.01) and low-frequency (MAF: 0.01-0.05) variants, particularly those in protein coding regions, has the potential to identify previously unidentified causal variants. Indeed, recent studies reaching sample sizes of 31,340 individuals have identified rare to low-frequency coding variants associated with PLT in *MPL* (MIM: 159530), *CD36* (MIM: 173510), and *JAK2* (MIM: 147796), among others[16, 17]. Studies with larger sample size are needed to further characterize the contribution of rare and low-frequency genetic variation to PLT and MPV.

To conduct such a study of platelet related traits, we formed the Blood Cell Consortium (BCX) to perform a large scale meta-analysis of exome chip association results of blood cell traits. In this report, we describe results from a meta-analysis of exome chip association data in 157,293 and 57,617 participants for PLT and MPV, respectively. The exome chip is a customized genotyping platform enriched for rare to low-frequency coding variants, as well as common variants previously identified in GWAS of complex disorders and traits. With increased sample size and use of the exome chip array, our goal was to identify rare, low-frequency, and common variants associated with PLT and MPV.

Materials and Methods

Study participants

The Blood Cell Consortium (BCX) was formed to identify genetic variants associated with blood cell traits using the exome chip platform. As the BCX is interested in the genetics of common hematological measures, our collaborative group is divided into three main working groups: RBC, WBC, and platelet[18, 19]. For the platelet working group, our sample is comprised of 157,293 participants from 26 discovery and replication cohorts of five ancestries: European (EA), African-American (AA), Hispanic, East Asian, and South Asian. Detailed descriptions of the participating cohorts are provided in the Tables S1-S4. All participants provided informed consent, and all protocols were approved by the participating studies' respective institutional review boards. In the platelet working group, we analyzed two traits: PLT ($x10^9$/L of whole blood) and MPV (fL) (Table S3).

Genotyping and Quality Control

Each participating study used one of the following exome chip genotyping arrays: Illumina ExomeChip v1.0, Illumina ExomeChip v1.1_A, Illumina ExomeChip-12 v1.1, Illumina ExomeChip-12 v1.2, Affymetrix Axiom Biobank Plus GSKBB1, or Illumina HumanOmniExpressExome Chip (Table S2). Genotypes were called either 1) using a combination of the Illumina GenomeStudio and zCall software or 2) the exome chip joint calling plan developed by the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium (Table S2)[20]. Standard quality control criteria were applied by each study. Exclusion criteria included: 1) sample call rates, 2) excess heterozygosity rate, 3) Hardy-Weinberg equilibrium p-values < $1x10^{-6}$, and 4) sex mismatch. Additionally, ancestry was confirmed through principal components or multi-dimensional scaling analyses using linkage

disequilibrium (LD) pruned markers ($r^2 < 0.2$) with MAF > 1%. Scatter plots anchored using the 1000 Genomes Project populations were visually inspected, and ancestry outliers excluded. We only included autosomal and X chromosome variants. All remaining variants (including monomorphic variants) were aligned to the forward strand and alleles checked to ensure that the correct reference allele was specified. We performed study specific level quality control on each trait association results using EasyQC[21]. We plotted variant allele frequencies from each study against ethnicity specific reference population allele frequencies to identify allele frequency deviations and presence of flipped alleles. Following all quality control procedures, each study generated an indexed variant call file (VCF) for subsequent analyses that was checked for allele alignment using the checkVCF package.

Association analysis

To assess the association between the blood cell traits and exome chip variants in the BCX, we considered blood cell traits measured in standard peripheral complete blood counts. When possible, we excluded individuals with blood cancer, leukemia, lymphoma, bone marrow transplant, congenital or hereditary anemia, HIV, end-stage kidney disease, dialysis, splenectomy, and cirrhosis, and those with extreme measurements of platelet traits. We also excluded individuals on erythropoietin treatment as well as those on chemotherapy. Additionally, we excluded women who were pregnant and individuals with acute medical illness at the time of complete blood count.

For platelet traits, we used raw values of PLT ($x10^9$/L) and MPV (fL). In each participating study, residuals for PLT and MPV were first calculated from linear regression models that adjusted for age, age^2, sex, study center (where applicable), and principal components of genotype data. We then transformed these residuals using the rank-based inverse

normal transformation. To confirm proper implementation of this transformation in each cohort, a scatter plot of the median standard error versus study specific sample size was visually inspected for deviations using EasyQC[21]. Autosomal and X chromosome variants were then tested for association with each blood cell trait using either RvTests or RAREMETALWORKER. Within individual cohorts, we performed analyses in ancestry-stratified groups: EA, AA, Hispanic, East Asian, and South Asian. Both statistical packages generate single variant association score summary statistics, variance-covariance matrices containing LD relationships between variants within a 1MB window, and variant-specific parameters including MAF, chromosome, position, strand, genotype call rate, and Hardy-Weinberg equilibrium p-values.

Discovery association meta-analysis

We performed ancestry-stratified (EA and AA) and combined all ancestry (All) meta-analyses of single variant association results using the Cochran-Mantel-Haenszel approach implemented in RareMETALS[22]. In the multi-ancestry meta-analyses (All), we combined individuals of EA, AA, Hispanic, South Asian, and East Asian ancestries. We included variants in the meta-analysis if the genotype call rate was $\geq$ 95%. For palindromic variants (i.e., A/T and C/G variants), we compared allele frequencies taken across the entire consortium in order to detect flipped alleles. We kept variants with an allele frequency difference < 0.30 or < 0.60 for race-specific (EA and AA) or combined all ancestry analyses, respectively[21]. Heterogeneity metrics (I^2 and heterogeneity p-value) were calculated using METAL[23]. Using single variant score statistics and variance-covariance matrices of LD estimates, we performed two types of gene-based tests: (1) variable threshold (VT) burden test with greatest power when all rare variants in a gene are associated consistently with a trait[24] and (2) sequence kernel association

test (SKAT)[25] with better power than the burden approach when rare variants in a gene have heterogeneous effects. For all gene-based tests, we only considered missense, nonsense, and splice site SNVs with MAF ≤1%. Similar to the single variant meta-analyses, gene-based results were generated for each major ancestry group (EA and AA) and for the combined multi-ancestry (All) samples.

Conditional analysis

To identify independent signals around significant associations, we performed step-wise conditional analyses conditioning on the most significant single variant in a 1MB window in RareMETALS. This procedure was repeated until there was no new signal identified in each region, defined as a p-value that accounts for the number of markers tested in each ancestry group. For discovery and conditional single variant analyses, the threshold was: AA $p<3.03x10^{-7}$, EA $p<2.59x10^{-7}$, and All $p<2.20x10^{-7}$. For gene-based tests, the significance threshold accounted for the number of genes tested: AA $p<2.91x10^{-6}$, EA $p<2.90x10^{-6}$, and All $p<2.94x10^{-6}$. In regions like chromosome 12q24 with known extended LD structure spanning more than 1MB, we performed a step-wise conditional analysis in GCTA using the Montreal Heart Institute Biobank cohort to disentangle 7 independent PLT-associated SNVs (Table S9)[26], conditioning on the most significant variant in the region.

Replication meta-analysis

We attempted to replicate PLT and MPV associations with independent SNVs that reached significance levels in 6 independent cohorts (Figure 1, Table S4). Single variant association results of the 6 independent cohorts were combined in RareMETALS. Contributing replication cohorts adhered to identical quality control and association analysis procedures described previously for the discovery phase. We combined results in EA (PLT n=19,939, MPV

n=15,519) and All (PLT n=35,436, MPV n=16,088) ethnicity groupings (Table S4). The results of discovery and replication phases were further combined using fixed effects inverse variance weighted meta-analysis in METAL[23].

Platelet Function Exome Chip

Two BCX cohorts, GeneSTAR and the Framingham Heart Study (FHS), measured platelet aggregation in a subset of genotyped participants. Platelet aggregation measures are described in detail elsewhere and briefly below (Table S18)[27]. Both studies isolated platelet-rich plasma from fasting blood samples and measured platelet aggregation after addition of agonists using a four-channel light transmission aggregometer (Bio/Data Corporation). FHS (Offspring Exam 5) tested aggregation for periods of 4 minutes after administration of ADP (0.05, 0.1, 0.5, 1.0, 3.0, 5.0, 10.0, and 15.0 μM) and 5 minutes after administration of epinephrine (0.01, 0.03, 0.05, 0.1, 0.5, 1.0, 3.0, 5.0, and 10.0 μM), as well as lag time(s) to aggregation with 190 μg/mL calf skin–derived type I collagen (Bio/Data Corporation). Threshold concentrations (EC_{50}) were determined as the minimal concentration of agonist required to produce a >50% aggregation. The maximal aggregation response (% aggregation) was also determined for each participant at each concentration tested. GeneSTAR recorded maximal aggregation (% aggregation) for periods of 5 minutes after ADP (2.0 and 10.0 μM) and 5 minutes after epinephrine administration (2.0 and 10.0 μM), as well as lag time(s) to aggregation with equine tendon–derived type I collagen (1, 2, 5, and 10 μg/mL). Exome chip genotyping, quality control, and association analyses adhered to methods described previously for PLT and MPV analysis. We queried independent SNVs associated with PLT (n=79) and/or MPV (n=38) in these platelet aggregation association results and report platelet aggregation associations with $p<0.001$.

Further Variant Annotation

In addition to primary analyses completed in this investigation, we took advantage of several existing resources to annotate our associated SNVs. Associated variants were cross-referenced with Combined Annotation Dependent Depletion (CADD) scores for exome chip[28]. The Global Lipids Genetics Consortium (GLGC), the CARDIoGRAM Exome Consortium, and Myocardial Infarction Genetics Consortium have each performed independent exome chip analysis of lipids traits and coronary heart disease (CHD)[29, 30]. The CHD phenotype reflected a composite endpoint that included MI, CHD, coronary artery bypass graft, and hospitalized angina, among others[29]. Similar to the platelet aggregation lookups, we queried our list of PLT and/or MPV associated SNVs against their exome chip association results for lipids and CHD. We report lipid and CHD associations with $p<0.0001$. From a curated collection of over 100 separate expression quantitative trait loci (QTL) datasets, we conducted a more focused query of whether platelet loci were also associated with transcript expression in blood, arterial and adipose related tissues. A general overview of a subset of >50 eQTL studies has been published (Supplemental Note)[31]. Separately, we queried transcripts in loci corresponding to previously unreported associated variants and/or marginally associated variants showing further evidence of association in our replication analyses to assess their platelet expression levels using the largest platelet RNA-seq dataset to date (n=32 patients with MI)[32].

Results

Discovery Meta-Analysis

In our discovery phase, we performed a meta-analysis of the associations of 246,925 single nucleotide variants (SNVs) with PLT and MPV in 131,857 and 41,529 individuals, respectively (Figure 1, Figures S1-S2, Tables S1-S4). Following the initial meta-analyses, we ran conditional analyses to identify independent loci and found 79 independent PLT and 38 independent MPV SNVs (Tables 1-2, Tables S5-S8). One association, rs12692566 in *LY75-CD302*, with PLT in EA did not surpass the initial discovery statistical significance threshold but surpassed the threshold when conditioned on nearby rs78446341 ($p=2.48x10^{-7}$). There were no associations unique to the AA ancestry group, which had a limited sample size (Tables S10-S11). Single variant meta-analysis results for each ancestry grouping that met our significance thresholds are available in the Supplement (Tables S10-S11). Additionally, full discovery meta-analysis results are available online (Web Resources).

Of these independently associated single variants, 32 PLT and 18 MPV variants were in loci not previously reported (Tables 1-2). Four of these 32 PLT loci had previously been identified as MPV loci (Table 1), while ten of the 18 MPV loci had previously been identified with PLT (Table 2)[8, 9, 17]. Of the independent loci in our study, 23 SNVs showed association with both PLT and MPV (Table 3, Figure 2). All but one (rs6136489 intergenic to *SIRPA* (MIM: 602461) and *LOC727993*) had opposite directions of effect for PLT and MPV. Additionally, the observed effect sizes for PLT and MPV displayed strong negative correlations (Figure 2), indicative of the strong negative correlation between these traits.

Associated variants ranged in allele frequency and included rare, low-frequency, and common SNVs. Most of the previously unreported associations were with common variants

(PLT n=25, MPV n=15), although associations with low-frequency (PLT n=6, MPV n=2) and rare (PLT n=1, MPV n=1) variants were observed. Rare (PLT n=6, MPV n=1) SNVs associated with PLT and MPV had larger effects compared to common and low-frequency SNVs (Tables 1-2, Tables S5-S8). A large majority of associated SNVs did not exhibit heterogeneous effects; however, one previously unreported association with *MRVI1* and a few known associated loci (e.g., *MYL2*/*SH2B3*/*ATXN2*, *ARHGEF3*, *WDR66*/*HPD*, and *JAK2*) did show moderate to substantial heterogeneity across discovery studies (Table S23). Gene-based tests of missense, nonsense, and splice-site rare variants that found significant results largely reflected rare and low-frequency single variant results, with variants in *TUBB1* (MIM: 612901), *JAK2*, *LY75* (MIM: 604524), *IQGAP2* (MIM: 605401), and *FCER1A* (MIM: 147140) showing associations (Tables S12-S13).

Replication Meta-Analysis

We attempted to replicate our associations in 6 independent cohorts (PLT n=25,436, MPV n=16,088) (Figure 1, Table S4). Of the loci not previously associated, 20/32 PLT and 11/18 MPV variants showed evidence of replication with $p<0.05$ and the same direction of effect (Tables 1-2). In addition to the significant markers in our discovery analysis, we carried forward 13 PLT and 10 MPV sub-threshold markers that approached discovery significance thresholds with p-values ranging from 2.47×10^{-7} to 1.99×10^{-6} (Tables S14-S15). Of these, 7/13 PLT and 4/10 MPV showed associations in same direction of effect with $p<0.05$ and surpassed significance thresholds when discovery and replication results were combined (Tables S14-S15).

Intersection with Other Cardiovascular and Blood Traits

As the BCX also completed analyses of RBC and WBC traits, we cross-referenced our list of PLT and MPV associated SNVs with the results of the other blood cell traits[18, 19]. Of our

replicated platelet loci previously unreported in the literature, six SNVs in *TMPRSS6* (MIM: 609862), *MAP1A* (MIM: 600178), *PNPLA3* (MIM: 609567), *FADS2* (MIM: 606149), *TMEM50A* (MIM: 605348), and *ZMIZ2* (MIM: 611196) showed association with RBC-related traits ($p<0.0001$) (Table 4). Similarly, five replicated platelet SNVs previously unreported in the literature in *PEAR1* (MIM: 610278), *CD33* (MIM: 159590), *SIRPA*, *ZMIZ2*, and *LY75* showed association with WBC-related traits ($p<0.0001$) (Table 4). To explore possible shared genetic associations of platelet size/number with platelet reactivity, we examined the association of PLT/MPV associated SNVs with platelet reactivity to collagen, epinephrine, and ADP in GeneSTAR and FHS. Eight SNVs associated with PLT and/or MPV were also associated with platelet reactivity ($p<0.001$) (Table 5, Tables S18-S19). The most strongly associated SNVs were located in genes implicated with platelet reactivity in prior GWAS, including *PEAR1*, *MRVI1* (MIM: 604673), *JMJD1C*, and *PIK3CG* (MIM: 601232)[27]. However, we did observe new suggestive relationships between platelet reactivity and SNVs in *PTGES* (MIM: 607061), *LINC00523*, and *RASGRP4* (MIM: 607320) (Table 5).

In addition to examining possibly shared genetic associations with blood cell specific traits, we queried our list of associated platelet SNVs against independent exome chip efforts in lipids and CHD by the GLGC, CARDIoGRAM Exome Consortium, and Myocardial Infarction Genetics Consortium exome chip studies[29, 30]. Numerous platelet associated SNVs (n=37), including those in *GCKR* (MIM: 600842), *FADS1* (MIM: 606148), *FADS2*, *MAP1A*, *APOH* (MIM: 138700), and *JMJD1C*, showed association with one or more lipids traits ($p<0.0001$) (Table S20). Far fewer (n=4; *MYL2* (MIM: 160781), *SH2B3* (MIM: 605093), *BRAP* (MIM: 604986), *APOH*)) showed association with CHD ($p<0.0001$) (Table S20).

Annotation of Associated Variants

We used various resources to annotate our platelet associated variants. First, we used CADD to predict the severity of associated variants[28]. As expected, rare and low-frequency coding SNVs were predicted to be more severe than common, non-coding variation (Tables 1-2, Tables S5-S8). To assess potential impact on gene expression, we queried our list of platelet associated SNVs against a collection of results from existing eQTL datasets[31]. Many (n=67) platelet-associated SNVs were also associated with gene expression in blood, arterial, or adipose tissues (Table S21). These included the reported *trans*-eQTL rs12485738 in *ARHGEF3* with several platelet-related transcript targets (e.g., GP1BA, GP6, ITGA2B, MPL, TUBB1, and VWF)[33], as well as eQTLs in newly identified PLT/MPV loci (e.g., rs1018448 with ARFGAP3/PACSIN2, rs1050331 with ZMIZ2, and rs174546 with FADS1/FADS2/TMEM258 expression). Using platelet RNA-seq data from 32 subjects with MI, we found that almost all of the genes closest to our previously unreported associated SNVs or marginal SNVs with evidence of replication were expressed in platelets, indicating the feasibility of potential functional roles in the relevant target cell type (Table S22)[32].

Discussion

Here, we present a large-scale meta-analysis of exome chip association data with two clinical platelet measurements, PLT and MPV. By combining exome chip association results in 157,293 and 57,617 participants, respectively, we detected numerous associations with rare, low-frequency, and common variants. There was substantial overlap of our platelet associations with concurrent exome chip association findings for RBC and WBC traits, indicating shared genetic influence on regulatory and functional mechanisms among the three different blood cell lineages[18, 19]. More surprisingly, we observed shared associations of platelet and lipids loci. The identification of shared blood cell and lipids associations as well as identifying genes with entirely new associations reveal candidates for further examination in order to further elucidate the mechanisms underlying platelet development and function.

Using Exome Chip to Identify Previously Unreported Genetic Associations

Using the exome chip which has an emphasis on rarer and coding variation, we found associations of variants that ranged from common to rare in allele frequency. We attempted to replicate independent associations, although our replication cohorts were underpowered to associations of rare variants. To inform our replication criteria, we conducted a power analysis using a sample size of 20,000 and considering multiple combinations of allele frequencies and effect sizes. Based on allele frequency and effect size, our most difficult to replicate variant was rs56106611 (MAF=0.012, Beta=0.11). However, we still had approximately 80% power to detect this association in the replication stage. Despite this, replication of extremely rare variants remains a challenge. For example, there were associations with rare coding variants with large effect sizes in *FCER1A*, *MPL*, *JAK2*, *SH2B3*, *TUBB1*, and *IQGAP2*[16, 17]. The overall effect size of these rare variants must be validated in independent studies. The PLT associated and predicted deleterious variant rs200731779 in *FCER1A* (Leu114Val) had a large effect (β=-2.96) in

discovery analyses, but could not be replicated in available samples due to its extremely rare allele frequency (MAF=$1.48x10^{-5}$ in EA). The affected amino acid is extracellularly positioned near the interface of two Ig-like domains, an area of the protein critical for FC-IgE interaction as shown through its crystal structure, biochemical data, and mutagenesis studies[34-37]. Other variants in *FCER1A*, a subunit of the allergy response IgE receptor and basophil differentiation factor, have previously been associated with IgE levels and monocyte counts[38, 39]. Increased platelet activation has been postulated to contribute to or be a consequence of allergic and inflammatory responses[40]. Our association of rare deleterious variation in *FCER1A* to reduced PLT provides a further link between platelet biology and allergy response.

Although SNVs in *IQGAP2* have previously been associated with PLT, we detected independent *IQGAP2* low-frequency and rare missense variants associated with increased MPV (Table 2, Figures S3-S4)[8, 17]. Located proximal to thrombin receptor *F2R* (MIM: 187930), IQGAP2 functions in the cytoskeletal dynamics in response to thrombin-induced platelet aggregation[41]. We did not observe *IQGAP2* associations with platelet aggregation, which may be due to the rare/low-frequency nature of the SNVs and the absence of thrombin-induced aggregation data in the available cohorts. Nonetheless, the associations of rare and low-frequency variants in *IQGAP2* further strengthen its contribution to platelet biology. In addition to *IQGAP2*, we observed other low-frequency associations, including nonsynonymous coding variants in *ITGA2B* (MIM: 607759), *LY75*, *MAP1A*, and *APOH*. The SNV rs76066357 in *ITGA2B*, a gene implicated in Glanzmann's Thrombasthenia (MIM: 273800), was associated with decreased PLT (Table 1). Moreover, *ITGA2B* codes for the platelet glycoprotein alpha-IIb, which part of the target receptor of GIIb/IIIa inhibitors (e.g., eptifibatide and abciximab) used in the acute management of acute coronary syndromes. Although ClinVar lists rs76066357 as

pathogenic (ID: 216944) with limited evidence, rs76066357 is a non-rare, predicted benign variant that contributes to population variability in PLT in our study as opposed to a severe Mendelian disorder of platelet reactivity[42]. Previous studies do suggest a potential role for variants in *ITGA2B* and *ITGB3* (MIM: 173470) leading to thrombocytopenia as well as abnormalities in platelet reactivity[43].

In addition to rare and low-frequency variant associations, we detected previously unreported associations for PLT and MPV at 25 and 15 common loci, respectively. For example, a common missense SNV rs1018489 in *ARFGAP3* (MIM: 612439) showed association with decreased PLT and increased MPV. This variant is an eQTL for both *ARFGAP3* and neighboring gene *PACSIN2* (MIM: 604960) in blood tissues (Table S21, Figures S5-S6). Although the possible role of the androgen receptor (AR) gene target and cellular secretory factor *ARFGAP3* is unknown in platelets[44-46], *PACSIN2* functions in the formation of the megakaryocyte demarcation membrane system during platelet production through interactions with FlnA[47]. Genetic variation that influences *PACSIN2* expression may hinder the formation of the megakaryocyte demarcation membrane system and lead to the production of fewer, but larger and potentially more reactive platelets. We also observed several other novel associations with common variants, including those in *SMG6* (MIM: 610963), a mediator of embryonic stem cell differentiation through nonsense mediated decay, and *LY75* an endocytotic immunity-related receptor highly expressed on dendritic cells where it is involved in recognition of apoptotic and necrotic cells[48-50].

Overlap with other platelet and blood cell traits

There was substantial overlap of variants associated with both PLT and MPV (n=23) as well as a strong negative correlation in effect sizes, consistent with the documented negative

correlation between the two traits in population studies (Figure 2)[51]. Only rs6136489, a reported eQTL for SIRPA, showed the same direction of effect for both PLT and MPV. SIRPA directly interacts with CD47, and SIRPA/CD47 signaling plays an important role in platelet clearance and the etiology of immune thrombocytopenia purpura[52-54]. Knockout *Sirpa* mice exhibit thrombocytopenia phenotypes, although have similar MPV to control animals[54]. How genetic variation in *SIRPA* influences MPV in addition to its demonstrated contribution to PLT remains to be characterized. In addition to shared associations of PLT and MPV, there was overlap in the parallel exome chip analyses of platelet reactivity. Largely mirroring results from previous GWAS, markers within *PEAR1*, *JMJD1C*, *PIK3CG*, and *MRVI1* showed the strongest associations with PLT/MPV and platelet reactivity[27, 55-57]. Other PLT/MPV associated markers in *PTGES3*, *LINC00523*, and *RASGRP4*, showed marginal associations. Notably, PTGES3 is linked to prostaglandin synthesis and the RasGRP family has been shown to have functional roles in blood cells including in platelet adhesion[58]. The association of platelet reactivity genes, particularly *PEAR1* and *MRVI1*, with PLT/MPV further supports a biological relationship between processes that control platelet function, megakaryopoiesis, and clearance[51, 59, 60]. However, these large-scale association analyses are unable to demonstrate whether these shared associations indicate shared biological mechanisms or simply reflect the epidemiological correlations among these traits.

In addition to platelet traits, there was substantial overlap of genetic associations with RBC and WBC traits examined by the BCX[18, 19]. The shared genetic associations with the two other primary blood cell lineages further supports other studies proposing that mechanisms that govern platelet size and number also influence RBC and WBC traits[61]. In BCX analyses, rs1050331 in the 3' untranslated region (UTR) of *ZMIZ2* was associated with increased PLT,

mean corpuscular hemoglobin (MCH), and mean corpuscular volume (MCV), as well as with decreased WBC count[18, 19]. rs1050331 is also an eQTL for ZMIZ2 expression in whole blood (Table S21)[62]. There are known sex differences in cell counts, with females consistently having higher PLT and mixed results on MPV[63, 64]. Similar to well-established PLT and MPV associated transcriptional regulator JMJD1C, ZMIZ2 directly interacts with AR to modulate AR-mediated transcription and influences mesodermal development, and thus genetic variation in *ZMIZ2* could potentially contribute to hormonally mediate differences in PLT across genders[65-67]. Also associated with increased PLT and decreased RBC indices was rs55707100 in *MAP1A*[18]. Though typically examined in a neurological context, *MAP1A* is involved in microtubule assembly, a process important in blood cell development and function[68]. Our observed association of *MAP1A* and its expression in platelets and RBCs suggests that the known role of *MAP1A* in developmental and cytoskeletal processes in neural tissues may extend to blood cells (Table S22). How these shared genetic factors specifically influence the development, maintenance, or clearance of multiple blood cell types remains to be determined.

Overlap with non-blood cell traits

While the overlap with other blood cell traits may be intuitive, we also observed overlap with quantitative lipids traits. In cross-trait lookups, several known PLT/MPV loci confirmed in this study (e.g., *JMJD1C*, *GCKR*, and *SH2B3*) showed associations with lipids traits, and several known lipids loci showed association to PLT/MPV (e.g., *FADS1*, *FADS2*, *APOH*, and *TMEM50A*). Moreover, *SH2B3,* which is also expressed in human vascular endothelial cells where it modulates inflammation, has been associated with blood pressure and the risk of MI[69-71]. Our study further suggests that a regulation of platelets could also contribute to potential implication of *SH2B3* in the development of cardiovascular diseases. The associated SNVs in the

FADS1/*FADS2* locus (rs174546 and rs174583) are eQTLs for multiple lipids-related transcripts in blood-related tissues, including TMEM258, FADS1, FADS2, and LDLR (Table S21)[62]. Intriguingly, expression of TMEM258 has also been shown to be a transcriptional regulatory target of cardiovascular disease implicated *CDKN2B-AS1* (MIM: 613149), a region marginally associated with PLT (Discovery EA $p=1.00x10^{-6}$, Replication EA p=0.0577, Combined EA $p=1.56x10^{-7}$) (Table S14)[72, 73]. Our genetic association results link the underlying genetic architecture of platelet and lipids traits as suggested by previous epidemiological, genetic, and animal studies[63, 74-77]. However, these observed shared genetic associations do not demonstrate whether these reflect direct genetic pleiotropy or indirect relationships. Several variants previously implicated in lipids (e.g., *FADS1*, *FADS2*, *SH2B3*, *TMEM50A*, and *GCKR*) have stronger associations with lipids traits relative to our platelet associations, suggesting that their primary effects are on lipids pathways (Table S20). Determining the directionality and causality among genetic variants, lipids, and platelets remains an important future step in dissecting which genetic variants may reveal new insights into platelet biology.

Conclusions

By performing a large meta-analysis of exome chip association results, we identified rare, low-frequency, and common variants that influence PLT and MPV. Despite our ability to detect numerous associations with SNVs across a wide range of allele frequencies, the exome chip interrogated a limited fraction of genomic variation. Sequencing-based studies across the genome in large sample sizes will be necessary to fully assess the contribution of variants across the allele frequency spectrum, particularly of rare variants in intergenic regions. Nonetheless, our results identify several intriguing genes and genetic mechanisms of platelet biology. Many of these associations overlapped with related blood cell and lipids traits, pointing to common

mechanisms underlying their development and maintenance. As blood cells share developmental lineages and several of our platelet associated genes have known developmental or transcriptional regulatory functions, we hypothesize that the origins of these shared genetic associations are mainly in blood cell development in the bone marrow. How these genes function and interact in RBC, WBC, and platelet development will need to be tested in future experiments in both functional and human-based studies. Advances in these domains could provide key insights into genes that influence human blood disorders and reveal new mechanisms for the development of novel therapeutic applications.

Acknowledgements

We thank all participants and study coordinating centers. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services. The FHS authors acknowledge that the computational work reported on in this paper was performed on the Shared Computing Cluster, which is administered by Boston University's Research Computing Services (www.bu.edu/tech/support/research/). The MHI Biobank acknowledges the technical support of the Beaulieu-Saucier MHI Pharmacogenomic Center. We would like to thank Liling Warren for contributions to the genetic analysis of the SOLID-TIMI-52 and STABILITY datasets. The University Medicine Greifswald is a member of the Caché Campus program of the InterSystems GmbH. The SHIP and SHIP-TREND samples were genotyped at the Helmholtz Zentrum München. EGCUT would like to acknowledge Mr. V. Soo, Mr. S. Smith, and Dr. L. Milani. The Airwave Health Monitoring Study thanks Louisa Cavaliero who assisted in data collection and management as well as Peter McFarlane and the Glasgow CARE, Patricia Munroe at Queen Mary University of London, and Joanna Sarnecka and Ania Zawodniak at Northwick Park. FINCAVAS thanks the staff of the Department of Clinical Physiology for collecting the exercise test data. YFS acknowledges the expert technical assistance in the statistical analyses by Ville Aalto and Irina Lisinen.

Supplemental Data

Supplemental Data included a note on eQTL analyses and additional funding information, 6 figures, and 23 tables.

Web Resources

OMIM: http://www.omim.org/

ClinVar: http://www.ncbi.nlm.nih.gov/clinvar/

CheckVCF: https://github.com/zhanxw/checkVCF

RareMetalWorker: http://genome.sph.umich.edu/wiki/RAREMETALWORKER

RVTests: http://genome.sph.umich.edu/wiki/RvTests

RareMETALS: http://genome.sph.umich.edu/wiki/RareMETALS

1000 Genomes Project: http://www.1000genomes.org/

Summary Association Statistics: http://www.mhi-humangenetics.org/en/resources

References

1. Chu SG, Becker RC, Berger PB, Bhatt DL, Eikelboom JW, Konkle B, Mohler ER, Reilly MP, and Berger JS (2010) Mean platelet volume as a predictor of cardiovascular risk: a systematic review and meta-analysis. J Thromb Haemost 8 (1):148-156

2. Sutcliffe P, Connock M, Gurung T, Freeman K, Johnson S, Kandala NB, Grove A, Gurung B, Morrow S, and Clarke A (2013) Aspirin for prophylactic use in the primary prevention of cardiovascular disease and cancer: a systematic review and overview of reviews. Health Technol Assess 17 (43):1-253

3. Hennekens CH, Dyken ML, and Fuster V (1997) Aspirin as a therapeutic agent in cardiovascular disease: a statement for healthcare professionals from the American Heart Association. Circulation 96 (8):2751-2753

4. Schick UM, Jain D, Hodonsky CJ, Morrison JV, Davis JP, Brown L, Sofer T et al (2016) Genome-wide Association Study of Platelet Count Identifies Ancestry-Specific Loci in Hispanic/Latino Americans. Am J Hum Genet

5. Soranzo N, Rendon A, Gieger C, Jones CI, Watkins NA, Menzel S, Doring A et al (2009) A novel variant on chromosome 7q22.3 associated with mean platelet volume, counts, and function. Blood 113 (16):3831-3837

6. Shameer K, Denny JC, Ding K, Jouni H, Crosslin DR, de AM, Chute CG, Peissig P, Pacheco JA, Li R, Bastarache L, Kho AN, Ritchie MD, Masys DR, Chisholm RL, Larson EB, McCarty CA, Roden DM, Jarvik GP, and Kullo IJ (2014) A genome- and phenome-wide association study to identify genetic variants influencing platelet count and volume and their pleiotropic effects. Hum Genet 133 (1):95-109

7. Qayyum R, Snively BM, Ziv E, Nalls MA, Liu Y, Tang W, Yanek LR, Lange L, Evans MK, Ganesh S, Austin MA, Lettre G, Becker DM, Zonderman AB, Singleton AB, Harris TB, Mohler ER, Logsdon BA, Kooperberg C, Folsom AR, Wilson JG, Becker LC, and Reiner AP (2012) A meta-analysis and genome-wide association study of platelet count and mean platelet volume in african americans. PLoS Genet 8 (3):e1002491

8. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, Serbanovic-Canic J et al (2011) New gene functions in megakaryopoiesis and platelet formation. Nature 480 (7376):201-208

9. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, Willenborg C et al (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet 41 (11):1182-1190

10. Kim YK, Oh JH, Kim YJ, Hwang MY, Moon S, Low SK, Takahashi A, Matsuda K, Kubo M, Lee J, and Kim BJ (2015) Influence of Genetic Variants in EGF and Other

Genes on Hematological Traits in Korean Populations by a Genome-Wide Approach. Biomed Res Int 2015:914965

11. Oh JH, Kim YK, Moon S, Kim YJ, and Kim BJ (2014) Genome-wide association study identifies candidate Loci associated with platelet count in koreans. Genomics Inform 12 (4):225-230

12. Li J, Glessner JT, Zhang H, Hou C, Wei Z, Bradfield JP, Mentch FD, Guo Y, Kim C, Xia Q, Chiavacci RM, Thomas KA, Qiu H, Grant SF, Furth SL, Hakonarson H, and Sleiman PM (2013) GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. Hum Mol Genet 22 (7):1457-1464

13. Guerrero JA, Rivera J, Quiroga T, Martinez-Perez A, Anton AI, Martinez C, Panes O, Vicente V, Mezzano D, Soria JM, and Corral J (2011) Novel loci involved in platelet function and platelet count identified by a genome-wide study performed in children. Haematologica 96 (9):1335-1343

14. Nurnberg ST, Rendon A, Smethurst PA, Paul DS, Voss K, Thon JN, Lloyd-Jones H, Sambrook JG, Tijssen MR, Italiano JE, Jr., Deloukas P, Gottgens B, Soranzo N, and Ouwehand WH (2012) A GWAS sequence variant for platelet volume marks an alternative DNM3 promoter in megakaryocytes near a MEIS1 binding site. Blood 120 (24):4859-4868

15. Johnson AD (2011) The genetics of common variation affecting platelet development, function and pharmaceutical targeting. J Thromb Haemost 9 Suppl 1:246-257

16. Auer PL, Johnsen JM, Johnson AD, Logsdon BA, Lange LA, Nalls MA, Zhang G, Franceschini N, Fox K, Lange EM, Rich SS, O'Donnell CJ, Jackson RD, Wallace RB, Chen Z, Graubert TA, Wilson JG, Tang H, Lettre G, Reiner AP, Ganesh SK, and Li Y (2012) Imputation of exome sequence variants into population- based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. Am J Hum Genet 91 (5):794-808

17. Auer PL, Teumer A, Schick U, O'Shaughnessy A, Lo KS, Chami N, Carlson C et al (2014) Rare and low-frequency coding variants in CXCR2 and other genes are associated with hematological traits. Nat Genet 46 (6):629-634

18. Chami N, Chen MH, Slater A.J., et al. (2016) Several new pleiotropic variants associated with red blood cell traits identified by exome genotyping. Am J Hum Genet, in press.

19. Schick UM, Tajuddin S, et al. (2016) Large-scale exome-wide association analysis identifies loci for white blood cell traits and pleiotropy with immune-mediated diseases. Am J Hum Genet, in press.

20. Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, Hansen M et al (2013) Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. PLoS One 8 (7):e68095

21. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Magi R, Ferreira T, Fall T, Graff M, Justice AE, Luan J, Gustafsson S, Randall JC, Vedantam S, Workalemahu T, Kilpelainen TO, Scherag A, Esko T, Kutalik Z, Heid IM, and Loos RJ (2014) Quality control and conduct of genome-wide association meta-analyses. Nat Protoc 9 (5):1192-1212

22. Liu DJ, Peloso GM, Zhan X, Holmen OL, Zawistowski M, Feng S, Nikpay M, Auer PL, Goel A, Zhang H, Peters U, Farrall M, Orho-Melander M, Kooperberg C, McPherson R, Watkins H, Willer CJ, Hveem K, Melander O, Kathiresan S, and Abecasis GR (2014) Meta-analysis of gene-level tests for rare variant association. Nat Genet 46 (2):200-204

23. Willer CJ, Li Y, and Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26 (17):2190-2191

24. Price AL, Kryukov GV, de Bakker PI, Purcell SM, Staples J, Wei LJ, and Sunyaev SR (2010) Pooled association tests for rare variants in exon-resequencing studies. Am J Hum Genet 86 (6):832-838

25. Wu MC, Lee S, Cai T, Li Y, Boehnke M, and Lin X (2011) Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet 89 (1):82-93

26. Yang J, Lee SH, Goddard ME, and Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88 (1):76-82

27. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler G, Lin SJ, Kraja AT, Province MA, Yang Q, Becker DM, O'Donnell CJ, and Becker LC (2010) Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nat Genet 42 (7):608-613

28. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, and Shendure J (2014) A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46 (3):310-315

29. (2016) Coding Variation in ANGPTL4, LPL, and SVEP1 and the Risk of Coronary Disease. N Engl J Med

30. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, Ganna A et al (2013) Discovery and refinement of loci associated with lipid levels. Nat Genet 45 (11):1274-1283

31. Zhang X, Gierman HJ, Levy D, Plump A, Dobrin R, Goring HH, Curran JE, Johnson MP, Blangero J, Kim SK, O'Donnell CJ, Emilsson V, and Johnson AD (2014) Synthesis of 53 tissue and cell line expression QTL datasets reveals master eQTLs. BMC Genomics 15:532

32. Eicher JD, Wakabayashi Y, Vitseva O, Esa N, Yang Y, Zhu J, Freedman JE, McManus DD, and Johnson AD (2015) Characterization of the platelet transcriptome by RNA sequencing in patients with acute myocardial infarction. Platelets:1-10

33. Fehrmann RS, Jansen RC, Veldink JH, Westra HJ, Arends D, Bonder MJ, Fu J et al (2011) Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet 7 (8):e1002197

34. Sandomenico A, Monti SM, Marasco D, Dathan N, Palumbo R, Saviano M, and Ruvo M (2009) IgE-binding properties and selectivity of peptide mimics of the FcvarepsilonRI binding site. Mol Immunol 46 (16):3300-3309

35. Mackay GA, Hulett MD, Cook JP, Trist HM, Henry AJ, McDonnell JM, Beavil AJ, Beavil RL, Sutton BJ, Hogarth PM, and Gould HJ (2002) Mutagenesis within human FcepsilonRIalpha differentially affects human and murine IgE binding. J Immunol 168 (4):1787-1795

36. Cook JP, Henry AJ, McDonnell JM, Owens RJ, Sutton BJ, and Gould HJ (1997) Identification of contact residues in the IgE binding site of human FcepsilonRIalpha. Biochemistry 36 (50):15579-15588

37. Garman SC, Kinet JP, and Jardetzky TS (1999) The crystal structure of the human high-affinity IgE receptor (Fc epsilon RI alpha). Annu Rev Immunol 17:973-976

38. Granada M, Wilk JB, Tuzova M, Strachan DP, Weidinger S, Albrecht E, Gieger C, Heinrich J, Himes BE, Hunninghake GM, Celedon JC, Weiss ST, Cruikshank WW, Farrer LA, Center DM, and O'Connor GT (2012) A genome-wide association study of plasma total IgE concentrations in the Framingham Heart Study. J Allergy Clin Immunol 129 (3):840-845

39. Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, Dean E et al (2011) Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). PLoS Genet 7 (6):e1002108

40. Page C and Pitchford S (2014) Platelets and allergic inflammation. Clin Exp Allergy 44 (7):901-913

41. Schmidt VA, Scudder L, Devoe CE, Bernards A, Cupit LD, and Bahou WF (2003) IQGAP2 functions as a GTP-dependent effector protein in thrombin-induced platelet cytoskeletal reorganization. Blood 101 (8):3021-3028

42. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, Das K, Toy T, Harry B, Yourshaw M, Fox M, Fogel BL, Martinez-Agosto JA, Wong DA, Chang VY, Shieh PB, Palmer CG, Dipple KM, Grody WW, Vilain E, and Nelson SF (2014) Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA 312 (18):1880-1887

43. Nurden AT, Pillois X, Fiore M, Heilig R, and Nurden P (2011) Glanzmann thrombasthenia-like syndromes associated with Macrothrombocytopenias and mutations in the genes encoding the alphaIIbbeta3 integrin. Semin Thromb Hemost 37 (6):698-706

44. Obinata D, Takayama K, Urano T, Murata T, Ikeda K, Horie-Inoue K, Ouchi Y, Takahashi S, and Inoue S (2012) ARFGAP3, an androgen target gene, promotes prostate cancer cell proliferation and migration. Int J Cancer 130 (10):2240-2248

45. Kartberg F, Asp L, Dejgaard SY, Smedh M, Fernandez-Rodriguez J, Nilsson T, and Presley JF (2010) ARFGAP2 and ARFGAP3 are essential for COPI coat assembly on the Golgi membrane of living cells. J Biol Chem 285 (47):36709-36720

46. Weimer C, Beck R, Eckert P, Reckmann I, Moelleken J, Brugger B, and Wieland F (2008) Differential roles of ArfGAP1, ArfGAP2, and ArfGAP3 in COPI trafficking. J Cell Biol 183 (4):725-735

47. Begonja AJ, Pluthero FG, Suphamungmee W, Giannini S, Christensen H, Leung R, Lo RW, Nakamura F, Lehman W, Plomann M, Hoffmeister KM, Kahr WH, Hartwig JH, and Falet H (2015) FlnA binding to PACSIN2 F-BAR domain regulates membrane tubulation in megakaryocytes and platelets. Blood 126 (1):80-88

48. Li T, Shi Y, Wang P, Guachalla LM, Sun B, Joerss T, Chen YS, Groth M, Krueger A, Platzer M, Yang YG, Rudolph KL, and Wang ZQ (2015) Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. EMBO J 34 (12):1630-1647

49. Butler M, Morel AS, Jordan WJ, Eren E, Hue S, Shrimpton RE, and Ritter MA (2007) Altered expression and endocytic function of CD205 in human dendritic cells, and detection of a CD205-DCL-1 fusion protein upon dendritic cell maturation. Immunology 120 (3):362-371

50. Cao L, Shi X, Chang H, Zhang Q, and He Y (2015) pH-Dependent recognition of apoptotic and necrotic cells by the human dendritic cell receptor DEC205. Proc Natl Acad Sci U S A 112 (23):7237-7242

51. Karpatkin S (1978) Heterogeneity of human platelets. VI. Correlation of platelet function with platelet volume. Blood 51 (2):307-316

52. Catani L, Sollazzo D, Ricci F, Polverelli N, Palandri F, Baccarani M, Vianelli N, and Lemoli RM (2011) The CD47 pathway is deregulated in human immune thrombocytopenia. Exp Hematol 39 (4):486-494

53. Olsson M, Bruhns P, Frazier WA, Ravetch JV, and Oldenborg PA (2005) Platelet homeostasis is regulated by platelet expression of CD47 under normal conditions and in passive immune thrombocytopenia. Blood 105 (9):3577-3582

54. Yamao T, Noguchi T, Takeuchi O, Nishiyama U, Morita H, Hagiwara T, Akahori H, Kato T, Inagaki K, Okazawa H, Hayashi Y, Matozaki T, Takeda K, Akira S, and Kasuga

M (2002) Negative regulation of platelet clearance and of the macrophage phagocytic response by the transmembrane glycoprotein SHPS-1. J Biol Chem 277 (42):39833-39839

55. Qayyum R, Becker LC, Becker DM, Faraday N, Yanek LR, Leal SM, Shaw C, Mathias R, Suktitipat B, and Bray PF (2015) Genome-wide association study of platelet aggregation in African Americans. BMC Genet 16:58

56. Lewis JP, Ryan K, O'Connell JR, Horenstein RB, Damcott CM, Gibson Q, Pollin TI, Mitchell BD, Beitelshees AL, Pakzy R, Tanner K, Parsa A, Tantry US, Bliden KP, Post WS, Faraday N, Herzog W, Gong Y, Pepine CJ, Johnson JA, Gurbel PA, and Shuldiner AR (2013) Genetic variation in PEAR1 is associated with platelet aggregation and cardiovascular outcomes. Circ Cardiovasc Genet 6 (2):184-192

57. Eicher JD, Xue L, Ben-Shlomo Y, Beswick AD, and Johnson AD (2015) Replication and hematological characterization of human platelet reactivity genetic associations in men from the Caerphilly Prospective Study (CaPS). J Thromb Thrombolysis

58. Stone JC (2011) Regulation and Function of the RasGRP Family of Ras Activators in Blood Cells. Genes Cancer 2 (3):320-334

59. van der Loo B and Martin JF (1999) A role for changes in platelet production in the cause of acute coronary syndromes. Arterioscler Thromb Vasc Biol 19 (3):672-679

60. Kauskot A, Vandenbriele C, Louwette S, Gijsbers R, Tousseyn T, Freson K, Verhamme P, and Hoylaerts MF (2013) PEAR1 attenuates megakaryopoiesis via control of the PI3K/PTEN pathway. Blood 121 (26):5208-5217

61. Bertin A, Mahaney MC, Cox LA, Rogers J, VandeBerg JL, Brugnara C, and Platt OS (2007) Quantitative trait loci for peripheral blood cell counts: a study in baboons. Mamm Genome 18 (5):361-372

62. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW et al (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45 (10):1238-1243

63. Sloan A, Gona P, and Johnson AD (2015) Cardiovascular correlates of platelet count and volume in the Framingham Heart Study. Ann Epidemiol 25 (7):492-498

64. Panova-Noeva M, Schulz A, Hermanns MI, Grossmann V, Pefani E, Spronk HM, Laubert-Reh D, Binder H, Beutel M, Pfeiffer N, Blankenberg S, Zeller T, Munzel T, Lackner KJ, Ten CH, and Wild PS (2016) Sex-specific differences in genetic and nongenetic determinants of mean platelet volume: results from the Gutenberg Health Study. Blood 127 (2):251-259

65. Daly ME (2011) Determinants of platelet count in humans. Haematologica 96 (1):10-13

66. Moreno-Ayala R, Schnabel D, Salas-Vidal E, and Lomeli H (2015) PIAS-like protein Zimp7 is required for the restriction of the zebrafish organizer and mesoderm development. Dev Biol 403 (1):89-100

67. Peng Y, Lee J, Zhu C, and Sun Z (2010) A novel role for protein inhibitor of activated STAT (PIAS) proteins in modulating the activity of Zimp7, a novel PIAS-like protein, in androgen receptor-mediated transcription. J Biol Chem 285 (15):11465-11475

68. Liu Y, Lee JW, and Ackerman SL (2015) Mutations in the microtubule-associated protein 1A (Map1a) gene cause Purkinje cell degeneration. J Neurosci 35 (11):4587-4598

69. Ganesh SK, Tragante V, Guo W, Guo Y, Lanktree MB, Smith EN, Johnson T et al (2013) Loci influencing blood pressure identified using a cardiovascular gene-centric array. Hum Mol Genet 22 (8):1663-1678

70. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, Najjar SS et al (2009) Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet 41 (6):666-676

71. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, Thorleifsson G et al (2009) Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 41 (3):342-347

72. Bochenek G, Hasler R, El Mokhtari NE, Konig IR, Loos BG, Jepsen S, Rosenstiel P, Schreiber S, and Schaefer AS (2013) The large non-coding RNA ANRIL, which is associated with atherosclerosis, periodontitis and several forms of cancer, regulates ADIPOR1, VAMP3 and C11ORF10. Hum Mol Genet 22 (22):4516-4527

73. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, Ingelsson E et al (2013) Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet 45 (1):25-33

74. Gomes AL, Carvalho T, Serpa J, Torre C, and Dias S (2010) Hypercholesterolemia promotes bone marrow cell mobilization by perturbing the SDF-1:CXCR4 axis. Blood 115 (19):3886-3894

75. Su Y, Wang Z, Yang H, Cao L, Liu F, Bai X, and Ruan C (2006) Clinical and molecular genetic analysis of a family with sitosterolemia and co-existing erythrocyte and platelet abnormalities. Haematologica 91 (10):1392-1395

76. Wang Z, Cao L, Su Y, Wang G, Wang R, Yu Z, Bai X, and Ruan C (2014) Specific macrothrombocytopenia/hemolytic anemia associated with sitosterolemia. Am J Hematol 89 (3):320-324

77. Murphy AJ, Bijl N, Yvan-Charvet L, Welch CB, Bhagwat N, Reheman A, Wang Y, Shaw JA, Levine RL, Ni H, Tall AR, and Wang N (2013) Cholesterol efflux in megakaryocyte progenitors suppresses platelet production and thrombocytosis. Nat Med 19 (5):586-594

Figures Titles and Legends

Figure 1: Study Design and Flow. Individual study level association analyses were performed using RareMetalWorker or RVTests. To perform quality control of individual study association results, we used EasyQC v8.6 to ensure proper trait transformations, to assess allele frequency discrepancies, and to evaluate other metrics. We then combined results in meta-analysis with RareMETALS v5.9 in three groups: African ancestry (AA), European ancestry (EA), and combined all five (AA, EA, Hispanic-Latino, East Asian, South Asian) ancestries (All). Independent variants identified by conditional analysis in RareMETALS with a p-value less than the threshold corrected for multiple testing (All: $p<2.20x10^{-7}$, EA: $p<2.59x10^{-7}$, AA: $p<3.03x10^{-7}$) were carried forward for replication. Markers showed replication if they had $p<0.05$ in the same direction of effect in the replication analyses. Associated markers were further annotated using various resources: (1) concurrent BCX exome chip analyses of RBC and WBC traits, (2) on-going exome chip analyses of platelet aggregation, quantitative lipids, and coronary heart disease (CHD) traits, (3) severity prediction by CADD, (4) an internal database of reported eQTL results, and (5) platelet RNA-seq data.

Figure 2: Shared PLT and MPV genetic associations. A) Comparing PLT and MPV effects sizes (r=-0.84) in European ancestry (EA) analyses of all identified SNVs identified (n=124). Examined SNPs include all those from Tables 1-2, Table S5-S9, and Tables S14-S15. **B)** 56 independent SNVs showed association to PLT only, while 15 independent SNVs were associated with MPV only. 23 independent SNVs were associated with both PLT and MPV. Named genes indicate that the association was not previously reported in the literature.

Table 1: Previously unreported associations (n=32) with PLT. We show variants in previously unreported loci and retained after conditional analyses in European Ancestry (EA) (p<2.59E-7) and All Ancestry (All) (p<2.20E-7) analyses. Associations in African Ancestry (AA) had previously been reported in the literature (Table S10). **Bolded** variants (20/32) showed evidence of replication (p<0.05, same direction of effect). If multiple genes/transcripts were annotated to a variant, the transcript most expressed in Eicher et al. 2015 (Table S22) was selected. Full results and annotations are available in the supplement (Table S5).

				European Ancestry (EA)						Combined All Ancestry (All)					
				Discovery (n=108,598)			Replication (n=19,939)		Combined	Discovery (n=131,857)			Replication (n=25,436)		Combined
rsID	**Ref/Alt**	**Function**	**Gene**	**EAF**	**Beta**	**P-value**	**Beta**	**P-value**	**P-value**	**EAF**	**Beta**	**P-value**	**Beta**	**P-value**	**P-value**
rs3091242	C/T	intron	*TMEM50A*	0.54	-0.026	9.68E-8	-0.017	0.124	3.85E-8	0.50	-0.02	1.03E-5	-0.0084	0.390	1.24E-5
rs12566888	**G/T**	**intron**	***PEAR1***	**0.094**	**0.040**	**1.42E-7**	**0.061**	**1.26E-3**	**1.17E-9**	**0.16**	**0.034**	**2.09E-8**	**0.047**	**4.31E-4**	**5.71E-11**
rs200731779	C/G	missense	*FCER1A*	1.5E-5	-2.96	2.48E-7	NA	NA	2.48E-7	1.2E-5	-2.96	2.48E-7	NA	NA	2.48E-7
rs6734238	A/G	intergenic	*IL1F10/IL1RN*	0.41	0.022	9.55E-6	0.0075	0.487	1.64E-5	0.41	0.026	7.19E-9	0.015	0.117	3.77E-9
rs12692566[b]	**C/A**	**missense**	***LY75-CD302***	**0.82**	**-0.029**	**9.19E-7**	**-0.042**	**2.50E-3**	**1.23E-8**	**0.83**	**-0.026**	**2.27E-6**	**-0.05**	**7.84E-5**	**3.65E-9**
rs78446341	**G/A**	**missense**	***LY75-CD302***	**0.02**	**0.092**	**4.16E-9**	**0.14**	**5.01E-5**	**1.98E-12**	**0.018**	**0.094**	**3.06E-10**	**0.13**	**9.23E-5**	**1.97E-13**
rs56106611[a]	**T/G**	**missense**	***KALRN***	**0.012**	**0.11**	**3.51E-8**	**0.11**	**7.14E-3**	**8.51E-10**	**0.01**	**0.11**	**8.59E-8**	**0.11**	**7.37E-3**	**2.14E-9**
rs1470579	A/C	intron	*IGF2BP2*	0.32	-0.028	1.08E-7	-0.0073	0.562	2.82E-7	0.38	-0.023	6.07E-7	-0.012	0.272	5.15E-7
rs1126673	C/T	ncRNA	*LOC100507053*	0.69	0.026	6.38E-8	0.019	9.63E-2	1.81E-8	0.71	0.025	1.87E-8	0.014	0.168	1.12E-8
rs1473247[a]	**T/C**	**intron**	***RNF145***	**0.27**	**-0.029**	**3.01E-8**	**-0.022**	**8.32E-2**	**7.28E-9**	**0.32**	**-0.026**	**1.32E-8**	**-0.025**	**1.85E-2**	**7.66E-10**
rs2256183	A/G	intron	*MICA*	0.56	0.03	6.78E-7	-0.022	0.104	2.60E-6	0.59	0.028	2.13E-7	0.011	0.389	3.20E-7
rs1050331	**T/G**	**3'UTR**	***ZMIZ2***	**0.47**	**0.037**	**1.32E-15**	**0.036**	**5.80E-4**	**3.28E-18**	**0.48**	**0.035**	**3.09E-17**	**0.031**	**8.80E-4**	**1.26E-19**
rs755109	**T/C**	**intron**	***HEMGN***	**0.37**	**0.028**	**2.87E-9**	**0.039**	**6.84E-4**	**1.17E-11**	**0.34**	**0.028**	**9.03E-11**	**0.044**	**2.18E-5**	**2.59E-14**
rs2068888	G/A	nearGene-3	*EXOC6*	0.45	-0.023	2.81E-7	-0.012	0.266	2.47E-7	0.44	-0.022	1.19E-7	-0.012	0.212	8.61E-8
rs3794153	**C/G**	**missense**	***ST5***	**0.45**	**-0.027**	**7.28E-9**	**-0.026**	**1.53E-2**	**3.57E-10**	**0.40**	**-0.027**	**2.19E-9**	**-0.023**	**2.47E-2**	**1.74E-10**
rs174583	**C/T**	**intron**	***FADS2***	**0.34**	**0.031**	**8.79E-9**	**0.048**	**1.22E-4**	**1.03E-11**	**0.34**	**0.028**	**4.72E-9**	**0.042**	**1.10E-4**	**4.42E-12**
rs45535039	T/C	3'UTR	*CCDC153*	0.28	0.04	4.02E-10	0.071	5.31E-2	8.48E-11	0.28	0.04	2.5E-12	0.056	8.56E-2	6.25E-13
rs11616188	**G/A**	**nearGene3**	***LTBR***	**0.42**	**-0.025**	**1.26E-8**	**-0.031**	**3.59E-3**	**1.81E-10**	**0.37**	**-0.025**	**7.57E-9**	**-0.033**	**1.07E-3**	**4.20E-11**
rs10506328[a]	**A/C**	**intron**	***NFE2***	**0.64**	**0.033**	**5.63E-11**	**0.06**	**5.88E-8**	**2.01E-16**	**0.69**	**0.038**	**3.79E-15**	**0.059**	**2.33E-8**	**2.73E-21**
rs2279574	C/A	missense	*DUSP6*	0.54	-0.023	2.47E-7	-0.0082	0.442	4.28E-7	0.50	-0.021	1.57E-7	-0.006	0.531	4.04E-7
rs61745424	**G/A**	**missense**	***CUX2***	**0.025**	**-0.064**	**2.36E-6**	**-0.085**	**6.79E-3**	**6.49E-8**	**0.023**	**-0.068**	**1.37E-7**	**-0.073**	**1.43E-2**	**6.30E-9**
rs2784521	A/G	nearGene-5	*DDHD1*	0.83	0.025	1.62E-5	0.0096	0.486	2.24E-5	0.76	0.028	2.92E-8	0.01	0.363	5.56E-8
rs55707100	**C/T**	**missense**	***MAP1A***	**0.032**	**0.095**	**7.03E-14**	**0.073**	**3.87E-2**	**9.53E-15**	**0.028**	**0.092**	**6.85E-14**	**0.082**	**1.62E-2**	**3.77E-15**
rs10852932	**G/T**	**intron**	***SMG6***	**0.36**	**-0.024**	**1.82E-6**	**-0.042**	**8.93E-4**	**1.42E-8**	**0.39**	**-0.025**	**4.79E-8**	**-0.036**	**6.99E-4**	**2.15E-10**
rs76066357	**G/C**	**missense**	***ITGA2B***	**0.014**	**-0.17**	**6.92E-16**	**-0.19**	**2.88E-5**	**1.05E-19**	**0.013**	**-0.16**	**1.92E-15**	**-0.18**	**6.00E-5**	**5.78E-19**
rs1801689	**A/C**	**missense**	***APOH***	**0.036**	**0.083**	**6.34E-12**	**0.13**	**2.44E-5**	**1.82E-15**	**0.032**	**0.090**	**8.64E-15**	**0.12**	**2.03E-5**	**1.57E-18**
rs892055	A/G	missense	*RASGRP4*	0.34	0.029	5.30E-10	0.018	9.87E-2	2.01E-10	0.38	0.025	3.49E-9	0.017	8.13E-2	9.96E-10
rs3865444	**C/A**	**5'UTR**	***CD33***	**0.32**	**-0.026**	**1.11E-6**	**-0.034**	**2.52E-3**	**1.27E-8**	**0.29**	**-0.026**	**2.10E-7**	**-0.032**	**3.03E-3**	**2.59E-9**
rs6136489[a]	**T/G**	**intergenic**	***SIRPA***	**0.34**	**-0.033**	**8.69E-13**	**-0.028**	**1.24E-2**	**4.00E-14**	**0.39**	**-0.030**	**1.8E-12**	**-0.024**	**1.30E-2**	**8.78E-14**
rs855791	**A/G**	**missense**	***TMPRSS6***	**0.56**	**-0.031**	**3.96E-11**	**-0.017**	**0.130**	**2.34E-11**	**0.60**	**-0.029**	**2.34E-11**	**-0.022**	**3.52E-2**	**2.97E-12**
rs1018448	A/C	missense	*ARFGAP3*	0.54	-0.028	4.02E-10	-0.0053	0.618	2.62E-9	0.59	-0.025	1.55E-9	-0.0065	0.515	6.13E-9
rs738409	**C/G**	**missense**	***PNPLA3***	**0.23**	**-0.042**	**1.49E-14**	**-0.042**	**1.75E-3**	**1.03E-16**	**0.22**	**-0.044**	**1.33E-18**	**-0.038**	**1.61E-3**	**9.73E-21**

[a]Previous association with MPV, [b]Surpasses significance threshold after conditioning on rs78446341 ($p=2.48E^{-7}$)

Abbreviations: PLT, platelet count; MPV, mean platelet volume; REF, reference allele; ALT, alternate allele; EAF, effect allele frequency

Table 2: Previously unreported associations (n=18) with MPV. We show variants in previously unreported MPV loci and retained after conditional analyses in European Ancestry (EA) (p<2.59E-7) and All Ancestry (All) (p<2.20E-7) analyses. Associations in African Ancestry (AA) had previously been reported in the literature (Table S11). **Bolded** variants (11/18) showed evidence of replication (p<0.05, same direction of effect). If multiple genes/transcripts were annotated to a variant, the transcript more expressed in Eicher et al. 2015 (Table S22) was selected. Full results and annotations are available in the supplement (Table S7).

				European Ancestry (EA)						Combined All Ancestry (All)					
				Discovery (n=34,021)			Replication (n=15,519)		Combined	Discovery (n=41,529)			Replication (n=16,088)		Combined
rsID	Ref/Alt	Function	Gene	EAF	Beta	P-value	Beta	P-value	P-value	EAF	Beta	P-value	Beta	P-value	P-value
rs6687605	**T/C**	**missense**	***LDLRAP1***	**0.53**	**0.046**	**8.27E-12**	**0.025**	**3.74E-2**	**1.80E-9**	**0.51**	**0.046**	**9.92E-11**	**0.024**	**3.58E-2**	**3.80E-11**
rs56043070[a]	**G/A**	**splice**	***GCSAML***	**0.069**	**0.094**	**1.30E-9**	**0.19**	**4.48E-16**	**1.12E-21**	**0.064**	**0.092**	**2.25E-10**	**0.19**	**3.66E-16**	**2.42E-22**
rs1339847[a]	G/A	missense	*TRIM58*	0.10	-0.10	1.47E-13	-0.037	5.44E-2	9.31E-13	0.10	-0.11	2.18E-17	-0.032	9.77E-2	1.06E-15
rs34968964[a]	G/C	missense	*IQGAP2*	0.0049	0.32	7.65E-9	0.12	9.18E-2	1.99E-8	0.004	0.32	2.11E-9	0.11	0.106	8.18E-9
rs34950321[a]	**C/T**	**missense**	***IQGAP2***	**0.018**	**0.18**	**7.80E-10**	**0.14**	**1.49E-3**	**6.03E-12**	**0.016**	**0.17**	**2.61E-9**	**0.14**	**1.59E-3**	**1.86E-11**
rs34592828[a]	**G/A**	**missense**	***IQGAP2***	**0.037**	**0.22**	**1.72E-27**	**0.16**	**2.73E-9**	**1.61E-34**	**0.032**	**0.23**	**1.68E-31**	**0.16**	**2.95E-9**	**2.98E-38**
rs1012899[a]	G/A	missense	*LRRC16A*	0.77	0.051	1.40E-7	0.012	0.417	1.24E-6	0.77	0.042	1.32E-6	0.016	0.273	2.50E-6
rs664370	**A/G**	**missense**	***PXT1***	**0.30**	**-0.034**	**8.03E-5**	**-0.025**	**5.61E-2**	**1.39E-5**	**0.35**	**-0.042**	**5.77E-8**	**-0.028**	**2.78E-2**	**7.23E-9**
rs2343596[a]	C/A	intron	*ZFPM2*	0.31	0.062	2.02E-13	0.012	0.357	3.32E-11	0.38	0.052	1.59E-11	0.012	0.339	4.35E-10
rs55895668[a]	T/C	missense	*PLEC*	0.43	-0.042	5.94E-7	-0.013	0.350	2.19E-6	0.47	-0.041	1.23E-7	-0.011	0.409	5.97E-7
rs4909945	**T/C**	**missense**	***MRVI1***	**0.68**	**-0.048**	**1.25E-8**	**-0.035**	**8.41E-3**	**5.19E-10**	**0.71**	**-0.041**	**3.96E-7**	**-0.035**	**7.42E-3**	**1.06E-8**
rs11125	A/T	missense	*LGALS3*	0.078	-0.091	1.55E-8	-0.037	0.117	2.76E-8	0.07	-0.09	4.22E-9	-0.037	0.117	7.21E-9
rs2010875[a]	**C/T**	**missense**	***PLEKHO2***	**0.14**	**-0.076**	**1.33E-7**	**-0.042**	**1.62E-2**	**2.10E-8**	**0.15**	**-0.063**	**3.01E-7**	**-0.042**	**1.62E-2**	**2.43E-8**
rs10512472[a]	**T/C**	**missense**	***SLFN14***	**0.18**	**-0.059**	**1.37E-8**	**-0.059**	**1.96E-4**	**1.12E-11**	**0.18**	**-0.058**	**3.15E-10**	**-0.059**	**1.20E-4**	**1.67E-13**
rs35385129	**C/A**	**missense**	***PVR***	**0.16**	**-0.058**	**6.24E-8**	**-0.044**	**7.36E-3**	**2.01E-9**	**0.15**	**-0.055**	**3.00E-8**	**-0.043**	**7.13E-3**	**8.79E-10**
rs2243603	C/G	missense	*SIRPB1*	0.77	0.044	5.89E-6	0.077	0.167	2.62E-6	0.79	0.049	4.58E-8	0.088	7.78E-2	1.25E-8
rs1018448	**A/C**	**missense**	***ARFGAP3***	**0.55**	**0.056**	**1.13E-12**	**0.051**	**1.78E-5**	**1.04E-16**	**0.60**	**0.055**	**1.52E-13**	**0.05**	**2.16E-5**	**1.68E-17**
rs1997715	**G/A**	**3'UTR**	***ZXDB***	**0.26**	**0.048**	**1.93E-9**	**0.084**	**5.83E-2**	**4.26E-10**	**0.35**	**0.04**	**4.58E-8**	**0.08**	**3.99E-2**	**8.88E-9**

[a]Previous association with PLT

Abbreviations: MPV, mean platelet volume; PLT, platelet count; REF, reference allele; ALT, alternate allele; EAF, effect allele frequency

Table 3: Variants associated with both PLT and MPV. All variants listed here showed association with both PLT and MPV in the opposite direction of effect as indicated by the arrows, except for rs6136489 (denoted by asterisk) which showed association with decreased PLT and decreased MPV.

rsID	Gene	PLT	MPV
rs12566888	*PEAR1*	↑	↓
rs1668873	*TMCC2*	↑	↓
rs56043070	*GCSAML*	↓	↑
rs12485738	*ARHGEF3*	↑	↓
rs56106611	*KALRN*	↑	↓
rs34592828	*IQGAP2*	↓	↑
rs1012899	*LRRC16A*	↓	↑
rs342293	*PIK3CG*	↓	↑
rs2343596	*ZFPM2*	↓	↑
rs10761731	*JMJD1C*	↑	↓
rs11602954	*BET1L*	↑	↓
rs10506328	*NFE2*	↑	↓
rs2958154	*PTGES3*	↓	↑
rs7961894	*WDR66*	↓	↑
rs1465788	*ZFP36L1*	↑	↓
rs2297067	*EXOC3L4*	↑	↓
rs2138852	*TAOK1*	↓	↑
rs10512472	*SLFN14*	↑	↓
rs11082304	*CABLES1*	↓	↑
rs6136489*	*SIRPA/LOC727993*	↓	↓
rs41303899	*TUBB1*	↓	↑
rs6070697	*TUBB1*	↑	↓
rs1018448	*ARFGAP3*	↓	↑

Abbreviations: PLT, platelet count; MPV, mean platelet volume

Table 4: Intersection of platelet associated variants with red blood cell (RBC) and white blood cell (WBC) traits (p<0.0001). We cross-referenced novel variants associated with platelet count (PLT) and/or mean platelet volume (MPV) in RBC and WBC association analyses in the Blood Cell Consortium (BCX). Here, we show RBC/WBC associated platelet variants with p<0.0001. Full details of RBC/WBC associations are shown in Table S16 and Table S17. Arrows denote direction of effect for the platelet and other blood cell trait(s).

SNP	MarkerName	Gene	PLT	Trait	Other Blood Cell
rs855791	22:37462936	*TMPRSS6*	↓	MCH, MCV, HGB MCHC, HCT	↑
rs855791	22:37462936	*TMPRSS6*	↓	RDW	↓
rs55707100	15:43820717	*MAP1A*	↑	HGB, MCH, HCT, MCHC	↓
rs174583	11:61609750	*FADS2*	↑	RDW	↓
rs174583	11:61609750	*FADS2*	↑	HGB, RBC, HCT, MCHC	↑
rs738409	22:44324727	*PNPLA3*	↓	HCT, HGB	↑
rs3091242	1:25674785	*TMEM50A*	↓	RDW	↑
rs1050331	7:44808091	*ZMIZ2*	↑	MCH, MCV	↓
rs1050331	7:44808091	*ZMIZ2*	↑	WBC	↑
rs6734238[a]	2:113841030	*IL1F10/IL1RN*	↑	MCH	↓
rs6734238[a]	2:113841030	*IL1F10/IL1RN*	↑	WBC, NEU	↑
rs12566888	1:156869047	*PEAR1*	↑	WBC, NEU, MON	↓
rs3865444	19:51727962	*CD33*	↓	WBC	↓
rs6136489	20:1923734	*SIRPA/LOC727993*	↓	WBC, LYM	↓
rs2256183[a]	6:31380529	*MICA*	↑	BAS	↑
rs12692566	2:160676427	*LY75-CD302*	↓	WBC	↓

[a]Marker not replicated in platelet analyses

Abbreviations: BCX, Blood Cell Consortium; RBC, red blood cell; WBC, white blood cell; PLT, platelet count; MCH, mean corpuscular hemoglobin; MCV, mean corpuscular volume; HGB, hemoglobin; MCHC, mean corpuscular hemoglobin concentration; HCT, hematocrit; RDW, red blood cell distribution width; PLT, platelet count; NEU, neutrophil; MON, monocyte; LYM, lymphocyte; BAS, basophil

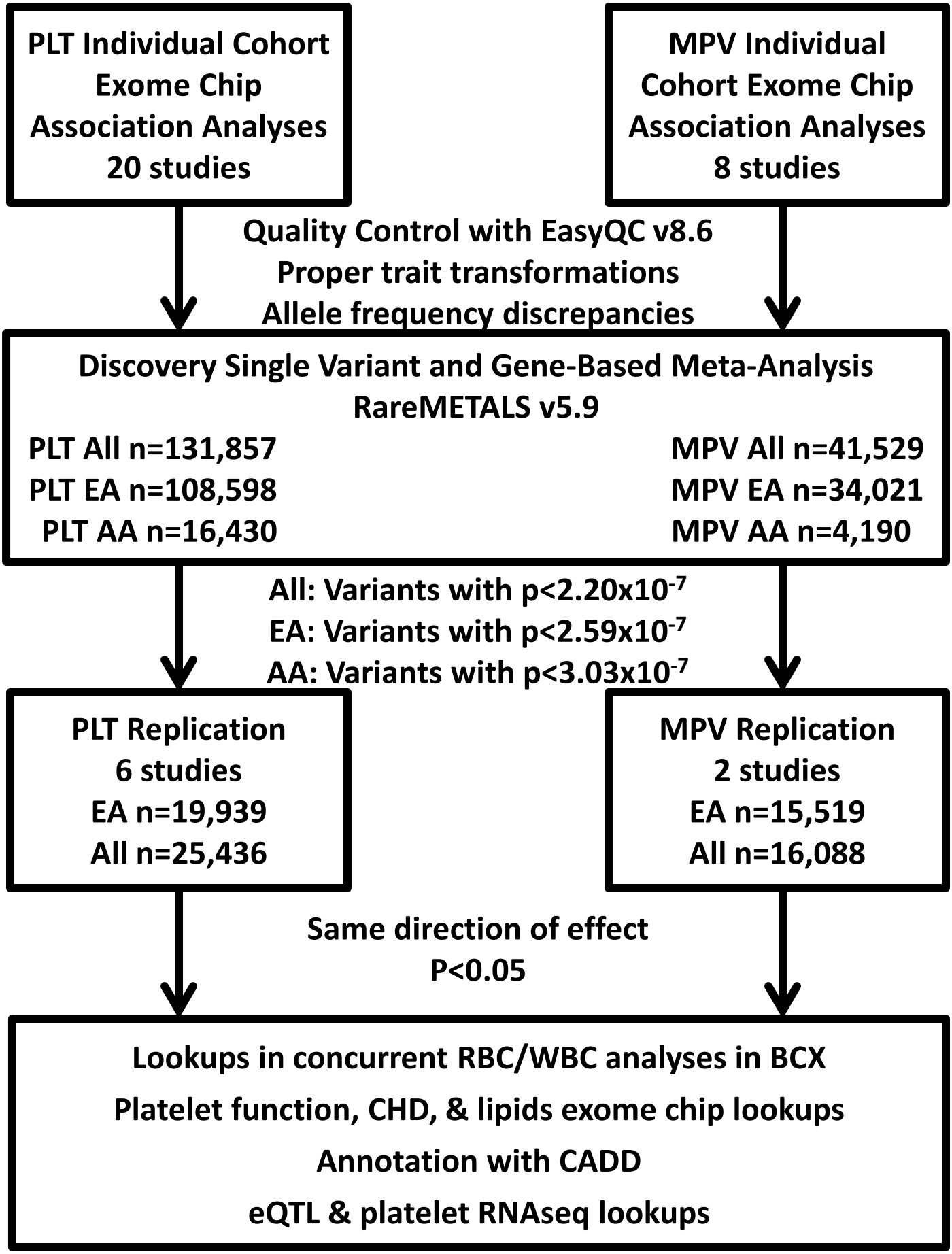
Table 5: Overlap of associations of platelet count (PLT) and mean platelet volume (MPV) variants with platelet reactivity (p<0.001). Variants were examined using platelet reactivity phenotypes (Table S18) in GeneSTAR and the Framingham Heart Study (FHS). Arrows denote direction of effect for PLT, MPV, and platelet reactivity. Multiple arrows refer to direction for respective agonist for platelet reactivity. Detailed association results for platelet reactivity are given in Table S19.

rsID	Gene	PLT	MPV	Agonist(s)[a]	Direction of Effects[b]
rs12566886	*PEAR1*	↑	↓	Epi, ADP, Collagen	↓↓↓
rs10761731	*JMJD1C*	↑	↓	Epi, ADP	↑↑
rs12355784	*JMJD1C*	↑	ns	Epi	↑
rs342293	*PIK3CG*	↓	↑	Epi	↓
rs4909945	*MRVI1*	ns	↓	Epi, ADP	↓↓
rs2958154	*PTGES3*	↓	↑	Collagen	↑
rs12883126	*LINC00523*	↑	ns	Epi	↑
rs892055	*RASGRP4*	↑	ns	Epi	↓

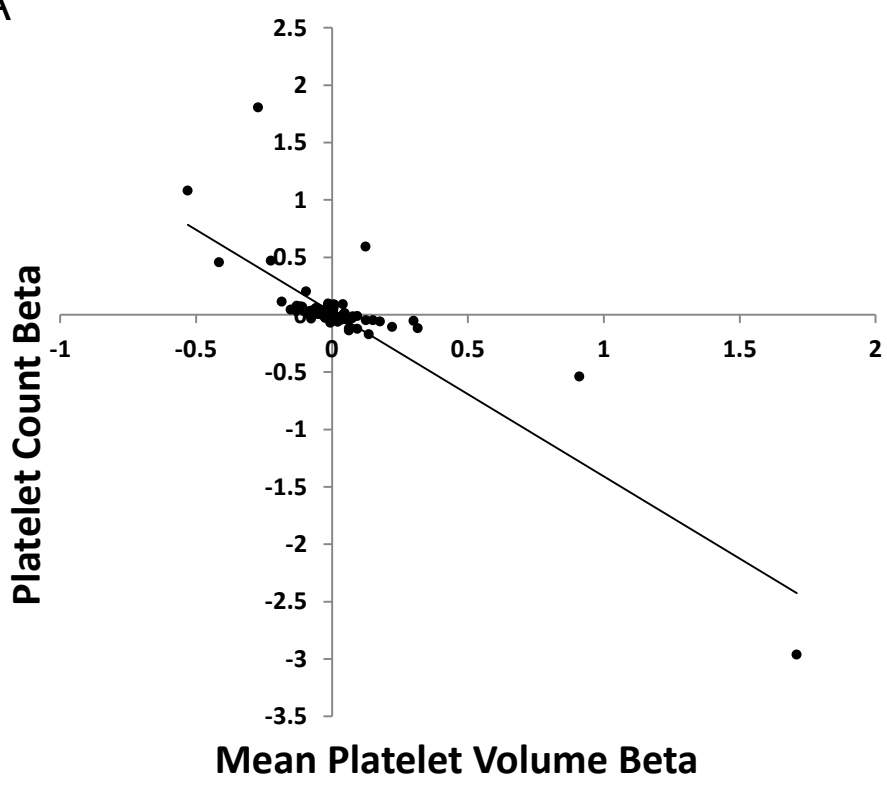
[a]Platelet reactivity associations with p<0.001

[b]As collagen measurements reflect lag time to aggregation, direction of effect has been flipped to denote a negative direction of effect as less reactive and positive direction of effect as more reactive

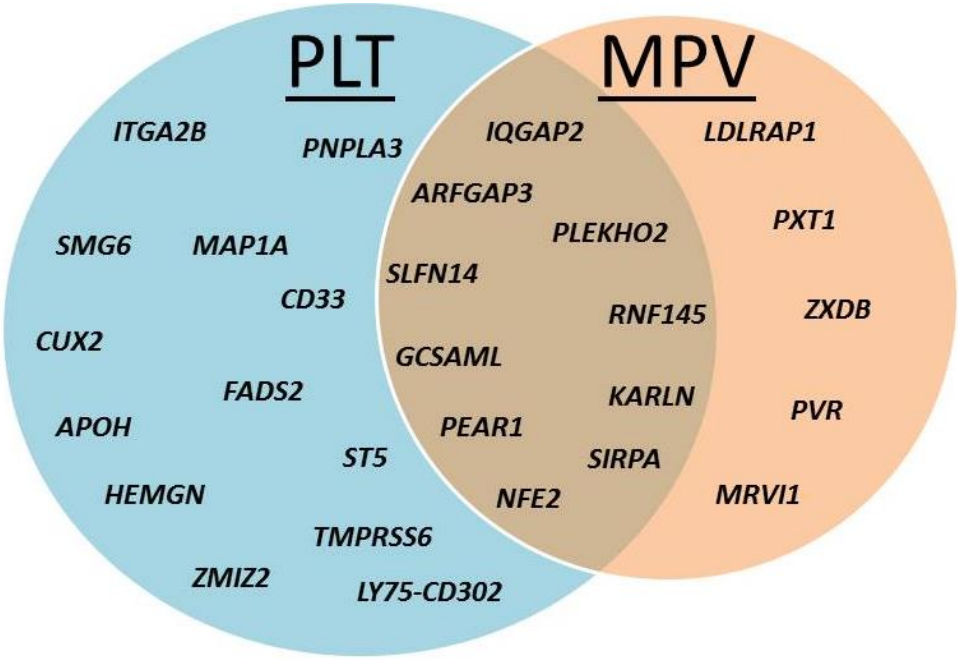
Abbreviations: PLT, platelet count; MPV, mean platelet volume; ns, not significant (p>0.05), Epi, epinephrine


PLT Individual Cohort Exome Chip Association Analyses 20 studies
MPV Individual Cohort Exome Chip Association Analyses 8 studies
Quality Control with EasyQC v8.6
Proper trait transformations
Allele frequency discrepancies
Discovery Single Variant and Gene-Based Meta-Analysis
RareMETALS v5.9
PLT All n=131,857
PLT EA n=108,598
PLT AA n=16,430
MPV All n=41,529
MPV EA n=34,021
MPV AA n=4,190
All: Variants with p<2.20x10-7
EA: Variants with p<2.59x10-7
AA: Variants with p<3.03x10-7
PLT Replication
6 studies
EA n=19,939
All n=25,436
MPV Replication
2 studies
EA n=15,519
All n=16,088
Same direction of effect
P<0.05
Lookups in concurrent RBC/WBC analyses in BCX
Platelet function, CHD, & lipids exome chip lookups
Annotation with CADD
eQTL & platelet RNAseq lookups

A


2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-1
-0.5
0
0.5
1
1.5
2
Platelet Count Beta
Mean Platelet Volume Beta

B


PLT
MPV
ITGA2B
PNPLA3
IQGAP2
LDLRAP1
ARFGAP3
PXT1
PLEKHO2
SMG6
MAP1A
SLFN14
CD33
RNF145
ZXDB
CUX2
GCSAML
FADS2
KARLN
PVR
APOH
PEAR1
ST5
SIRPA
HEMGN
NFE2
MRVI1
TMPRSS6
ZMIZ2
LY75-CD302