Next-generation re-sequencing of genes involved in increased platelet reactivity in diabetic patients on acetylsalicylic acid

Marek Postula, Piotr K. Janicki, Ceren Eyleten, Marek Rosiak, Agnieszka Kaplon-Cieslicka, Shigekazu Sugino, Radoslaw Wiliński, Dariusz A. Kosior, Grzegorz Opolski, Krzysztof J. Filipiak, & Dagmara Mirowska-Guzel

Department of Experimental and Clinical Pharmacology, Medical University of Warsaw, Center for Preclinical Research and Technology CEPT, Warsaw, Poland, Perioperative Genomics Laboratory, Penn State College of Medicine, Hershey, PA, USA, Department of Cardiology and Hypertension, Central Clinical Hospital, The Ministry of the Interior, Warsaw, Poland, Department of Cardiology, Medical University of Warsaw, Warsaw, Poland, Department of Cardiac Surgery, Medical University of Warsaw, Warsaw, Poland, and Department of Applied Physiology, Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland

Abstract

The objective of this study was to investigate whether rare missense genetic variants in several genes related to platelet functions and acetylsalicylic acid (ASA) response are associated with the platelet reactivity in patients with diabetes type 2 (T2D) on ASA therapy. Fifty eight exons and corresponding introns of eight selected genes, including PTGS1, PTGS2, TXBAS1, PTGIS, ADRA2A, ADRA2B, TXB2A2R, and P2RY1 were re-sequenced in 230 DNA samples from T2D patients by using a pooled PCR amplification and next-generation sequencing by Illumina HiSeq2000. The observed non-synonymous variants were confirmed by individual genotyping of 384 DNA samples comprising of the individuals from the original discovery pools and additional verification cohort of 154 ASA-treated T2DM patients. The association between investigated phenotypes (ASA induced changes in platelets reactivity by PFA-100, VerifyNow and serum thromboxane B2 level [sTxB2]), and accumulation of rare missense variants (genetic burden) in investigated genes was tested using statistical collapsing tests. We identified a total of 35 exonic variants, including 3 common missense variants, 15 rare missense variants, and 17 synonymous variants in 8 investigated genes. The rare missense variants exhibited statistically significant difference in the accumulation pattern between a group of patients with increased and normal platelet reactivity based on PFA-100 assay. Our study suggests that genetic burden of the rare functional variants in eight genes may contribute to differences in the platelet reactivity measured with the PFA-100 assay in the T2DM patients treated with ASA.

Keywords

Acetylsalicylic acid, diabetes mellitus, genetic polymorphism, next-generation sequencing, platelet reactivity

Introduction

Platelet phenotype on ASA therapy has been characterized by the inherited differences in platelet reactivity and function of platelet surface receptors, as well as enzymes related to platelet activation [1–3]. Until recently, most strategies aiming at elucidating the impact of genetic variation on platelet reactivity and function relied on common genetic variants (i.e. minor allele frequency or minor allele frequency (MAF) >5%) in candidate genes [4–7]. The results obtained so far suggest that the common variants in several genes may contribute to platelet reactivity in patients with type 2 diabetes (T2DM) [8–10]. They account, however, for relatively small effects on platelet function, leading to the hypothesis that rare variants with a potentially stronger damaging effect might also play a functional role in determining platelet reactivity phenotypes. It has been suggested that the combined effects of rare deleterious mutations (so-called genetic burden) in genes within specific functional pathways could explain a substantial fraction of the differences in platelet reactivity [11–18]. Identification of these rare variants requires a next generation sequencing (NGS) approach, which provides information on every base pair across the region of interest. The primary goal of this study was therefore to re-sequence selected genes with previously demonstrated functional common polymorphisms altering platelet response to ASA in diabetic patients.

Several single-nucleotide polymorphisms (SNPs) in both coding and non-coding regions of human COX-1 have been previously identified and were postulated to change its activity [6, 7, 19, 20]. Although low-dose ASA inhibits the COX-2 pathway only marginally, studies have also investigated the influence of the PTGS2 gene polymorphisms, however the results are inconsistent [21, 22].

Thromboxane A2 is a potent vasoconstrictor and stimulator of platelet aggregation with high levels produced in atherosclerosis. The enzyme that catalyzes its synthesis, thromboxane A synthase 1 (TXBAS1), is a key component for TXBA2 function along with the receptor that mediates its action, thromboxane A2 receptor (TXB2A2R). Of particular interest, specific TXBAS1 gene
polymorphisms may be a useful marker for development of cerebral infarction despite ASA therapy [23, 24]. Moreover, several common genetic variants within TXBA2R gene may also reduce the ability of ASA to inhibit platelet function in different populations of patients [8, 25, 26]. As TXA2 and prostacyclin have opposite effects on platelets, we decided to sequence PTGIS gene that was described to be associated with myocardial infarction (MI) and stroke risk [27, 28].

It was suggested previously that the platelet reactivity on ASA therapy may be partially dependent on neither COX-1 nor COX-2 pathways, thus limiting the possible impact of polymorphic variants in the COX genes [29–31]. We decided therefore to evaluate additional genomic targets not directly related to COX-1 platelet aggregation, namely alpha-2A adrenoreceptor (ADRA2A), alpha-2B adrenoreceptor (ADRA2B), and P2Y1 purinergic receptor (P2RY1). P2Y1 is a purinergic receptor mediating platelet activation upon stimulation by ADP, and its common polymorphic variants were reported to be associated with ASA efficacy in platelet aggregation studies [31–33]. In addition, ADRA2A polymorphism affects epinephrine-induced platelet aggregation and platelet reactivity under high shear stress conditions [34]. Both genome-wide meta-analyses and targeted genotyping identified and confirmed its association with platelet aggregation in response to epinephrine [8, 35]. Moreover, we decided to evaluate polymorphisms in the ADRA2B receptor gene, as it was recently postulated that the ADRA2B subtype does exist in platelets and is an important regulator of aggregation [36].

The primary goal of this study was to re-sequence genes with previously demonstrated common polymorphisms affecting platelet response to ASA in diabetic patients. In order to identify rare variants associated with platelet reactivity on ASA therapy in the Polish population, we performed pooled-DNA deep-sequencing of eight, described above genes, which have been associated with platelet activation and/or platelet response to ASA [2, 4, 6, 7, 19, 25, 28, 33–35, 37–43]. The discovery step was followed by the genotyping in individual patients with known platelet reactivity status on ASA therapy in order to verify and compare mutation frequencies.

Methods

Demographic and clinical data

DNA samples (N = 230) were obtained from the subjects participating in the AVOCADO (Aspirin Vs/Or Clopidogrel in Aspirin-resistant Diabetics inflammation Outcomes Study), a longitudinal, multi-center, prospective, randomized, open-label study in Polish population. The verification cohort consisted of patients from the original discovery phase (N = 154), remaining 70 subjects participating in the AVOCADO study and presenting to the outpatient clinics of the Central Teaching Hospital of the Medical University of Warsaw, as well as additional 84 patients from Department of Cardiac Surgery that were assigned to by-bass surgery due to coronary artery disease (CAD). The characterization of the study population, including the inclusion and exclusion criteria, was published previously (Table I) [8–10]. Briefly, the Caucasian subjects with T2DM were recruited who, at the time of enrollment, had been taking ASA tablets at the dose of 75 mg per day for at least 3 months for primary or secondary prevention of MI. Neither clopidogrel nor antiplatelet drugs other than ASA were used in any of the investigated patients. All patients had been taking oral antidiabetic agents and/or insulin for at least 6 months; diet-controlled diabetic patients were not included. Compliance to ASA therapy at the study entry was determined based upon the patient’s own statement and serum thromboxane B2 (sTxB2) level measurement by immunoassay kit (Cayman Chemicals, Ann Arbor, MI, USA). The compliance with ASA treatment was defined by the sTxB2 levels below 7.2 ng/mL [8–10].

Platelet function analysis

The VerifyNow Aspirin Assay (Accumetrics, San Diego, CA, USA) uses turbidimetric-based optical detection system, which measures platelet-induced agglutination as the increase in light transmittance in response to AA. The PFA-100 assay (Dade-Behring International, Inc., Newark, DE, USA) uses whole blood, simulates the high shear stress conditions, and records the time necessary for occlusion of the aperture, defined as closure time (CT), which is indicative of platelet function in the whole blood sample [8–10]. Cut-off values for normal (i.e. low) platelet reactivity for all assays are given in Figure 1.

Next-generation re-sequencing of DNA samples

DNA was extracted from a venous EDTA-whole blood sample by the membrane ultrafiltration method (FujiFilm Life Sciences distributed by Autogene, Holliston, MA USA), according to the manufacturer recommendations. The DNA samples in the pools were selected based on the contrasting values of platelet function assays (Figure 1) and pooled together (23–30 in each pool) in the equimolar fashion, followed by submission to Otogenetics Corporation (Norcross, GA, USA) for target capture and sequencing. The quality of the gDNA samples was assessed by agarose gel electrophoresis and the DNA concentration and purity were investigated using a spectrophotometer and optical density at 230/260/280 nm. The final DNA concentrations in pools were confirmed using the PicoGreen method (PicoGreen dsDNA Quantitation Reagent Kit, Molecular Probes Inc., Eugene, OR, USA). The gDNA fragmentation was performed by Covaris (Covaris, Inc., Woburn, MA, USA) and fragments were tested for size distribution and concentration using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) and Nanodrop (Thermo Fisher Scientific, Wilmington DE, USA). Illumina libraries were made from qualified fragmented gDNA using NEBNext reagents (New England Biolabs, Ipswich, MA, USA) and tested for enrichment of selected genes by qPCR and for size distribution and concentration by an Agilent Bioanalyzer 2100. The samples were then sequenced on an Illumina HiSeq2000 (Illumina, San Diego, CA, USA) which generated paired-end reads of 100 nucleotides.

Sequencing data analysis and identification of variants

The original source sequencing data files (from Illumina 1.8+ pipeline in fastq format) were analyzed by combined Web based Galaxy (main.g2.bx.psu.edu) and DNAnexus (www.dnanexus.com) platforms [44]. The reads in the fastq files were aligned (i.e. mapped) to a sequence of the human reference genome (Human Feb. 2009–GCRCh37/hg19) with Galaxy (s Burrows-Wheeler transformation routine which resulted in generation of the sorted Sequence Alignment/Map (SAM) format files for each pooled sample, and then were transformed into the indexed BAM files [45, 46]. Variant detection in each BAM file was performed using Galaxy’s FreeBayes (Bayesian genetic variant detector, github.com/ekg/freebayes#readme) [47].

In addition, data were independently analyzed using DNAnexus platform in the Nucleotide-Level Variation analysis mode. This platform, although similar to Burrows-Wheeler transform routine used in Galaxy, uses different algorithm for variant detection – GATK. Our approach applied GATK for base quality score recalibration, indel realignment, duplicate removal, and
Rare genetics and platelet reactivity in T2D

Table I. Demographic and clinical characteristics of the study population (N = 384).

Demographics

| Age (years) | 67.3 ± 8.78 |
| Female     | 140 (46.9%) |
| BMI (kg/m²) | 30.6 ± 10.88 |
| DM         | 311 (85.9%) |
| Dyslipidemia | 286 (79.0%) |
| Hypertension | 3290 (90.9%) |
| CAD        | 238 (65.7%) |
| Prior MI   | 111 (30.6%) |
| Prior stroke | 29 (8.0%) |
| Heart failure | 136 (37.5%) |
| History of smoking | 206 (56.9%) |
| Current smoking | 28 (7.7%) |

Concurrent medications

| Oral hypoglycemic       | 246 (67.9%) |
| Insulin                 | 93 (25.7%) |
| Beta-blockers           | 262 (72.6%) |
| ACE inhibitors          | 263 (72.6%) |
| Statins                 | 278 (77%) |

Biochemical and hematological parameters

| HGB (g/l)                        | 13.8 ± 1.3 |
| HCT (%)                          | 41.1 ± 4.4 |
| WBC (10³/µl)                     | 7.1 ± 2.2 |
| PLT (10³/µl)                     | 227.7 ± 58.2 |
| MPV (fl)                         | 9.9 ± 1.2 |
| eGFR (ml/min/1.73 m²)            | 70.8 ± 20.8 |
| HbA1c (%)                        | 6.9 ± 1.3 |
| hsCRP (mg/l)                     | 4.1 ± 5.6 |
| vWF%                             | 147.9 ± 62.8 |
| sTxB2 (ng/ml)                    | 1.0 ± 1.6 |
| VerifyNow (ARU)                  | 466.9 ± 67.2 |
| CEPI-CT (sec)                    | 232.8 ± 73.2 |
| CADP-CT (sec)                    | 124.07 ± 74.9 |

This table shows the characteristics of the overall population. Data are expressed as mean (SD) unless otherwise indicated. Abbreviations: BMI, body mass index (kg/m²); SBP, systolic blood pressure; DBP, diastolic blood pressure; CAD, coronary artery disease; MI, myocardial infarction; ACE, angiotensin-converting enzyme; HGB, hemoglobin (g/dL); HCT, hematocrit (%); WBC, white blood cells (10³/µl); PLT, platelet count (10³/µl); MPV, mean platelet volume (fl); eGFR, estimated glomerular filtration rate (ml/min/1.73); HbA1c, glycated hemoglobin; hsCRP, high sensitivity C-reactive protein; vWF, von Willebrand factor; sTxB2, serum thromboxane B2; ARU, aspirin reaction units; CEPI-CT, collagen/epinephrine closure time; CADP-CT, collagen/adenosine diphosphate closure time; sec, seconds.

Figure 1. Cut-off points and values for laboratory test used for platelet reactivity assessment. CEPI-CT, collagen/epinephrine closure time; ARU, aspirin reaction units; sTxB2, serum thromboxane B2; ASA, acetylsalicylic acid; LPR-ASA, low platelet reactivity on ASA; HPR-ASA, high platelet reactivity on ASA.

Individual SNP genotyping

Genotyping for selected markers in individual DNA samples was performed at Children’s Hospital Boston, using a custom Sequenom iPLEX assay in conjunction with the Mass ARRAY platform (Sequenom Inc., La Jolla, CA, USA) [8–10, 53]. Panels of SNP markers were designed using Sequenom Assay Design 3.2 software (Sequenom Inc., La Jolla, CA, USA).

Analysis of Burden Association Signal Due to Rare Variants

Standard statistical methods (e.g. Chi-square tests) used to test for association with single common genetic variants is underpowered for rare variants unless sample sizes or effect sizes are very large. A logical alternative approach is to employ burden tests that assess the cumulative effects of multiple variants in a genomic region. Burden tests proposed to date are based on collapsing or summarizing the rare variants within a region by a single value, which is then tested for association with the trait of interest.

For the testing of differences in the variant burden between study patients we employed different burden tests including score test (from logistic regression model) (SCORE) [54, 55], replication based test (RBT) [56], adaptation score test (ASCORE) [57, 58], sequential score sum test (SEQSUM) [59] and TTTEST = Hotelling T2 test (TTEST) [60]. The differences between the genetic burdens tests used in the study are related to the different algorithms of calculations (e.g. taking into account directionality of the changes, assumption of damaging or protective effects of minor alleles etc.). The detailed explanation of these statistical tests were performed using the R software package AssotesteR written by Gaston Sanchez (www.gastonsanchez.com) and available from cran.r-project.org/web/packages/AssotesteR/. For our primary analysis, we restricted the burden tests to rare missense variants both described and not described in dbSNP or the 1000 Genomes Project (release June 2011) under the hypothesis that truly rare mutations are more likely to be pathogenic. All tested variants were verified by individual genotyping by Sequenom and the numbers of analyzed alleles were obtained from the verification part of the study. In this analysis, we only included rare missense variants with MAF < 5% variants. We assessed statistical significance by 100 000 case-control permutations, with exception of t-test, which uses the asymptotic p value.

Power analysis

In 115 case-control pairs, for a investigated phenotype (e.g. ASA induced changes in platelet reactivity in diabetic population) of 25% population risk, there is a moderate power (70%) to detect a low-frequency (1 in 100) allele with a dominant 2.5 times increased risk significance p < 0.00625 (i.e. nominal significance 0.05 divided by the number of sequenced genes) by gene-test which is based on the aggregated allele frequency and averaged effect size for all rare mutations in each gene [61, 62].
Table II. Summary of variant observed in all re-sequenced genes in the study.

| Chr | (Position) and dbSNP | Base switch | Gene | Type | MAF for N = 384 | AA change | SIFT prediction | Mutation Taster prediction |
|-----|----------------------|-------------|------|------|----------------|-----------|----------------|---------------------------|
| Chr1 | (186643636) New C/T | PTGS1 SNP | 0.0013 | C/Y | D | D |
| Chr2 | (96780763) rs29000569 C/T | ADRA2B SNP | 0.015 | V/I | D | N |
| Chr2 | (96781257) rs93335368 C/G | ADRA2B SNP | 0.016 | G/A | T | N |
| Chr2 | (96781464) New A/T | ADRA2B SNP | 0.005 | V/D | D | D |
| Chr3 | (152553627) New C/T | PTGS1 SNP | 0.002 | A/V | T | N |
| Chr3 | (152554482) New C/T | PTGS1 SNP | 0.0013 | A/V | T | D |
| Chr7 | 139715641 rs1892868 C/A | TXBAS1 SNP | 0.012 | E/K | D | D |
| Chr7 | 139706915 rs3735354 A/G | TXBAS1 SNP | 0.0013 | E/K | T | D |
| Chr7 | 139652440 rs6952940 C/T | TXBAS1 SNP | 0.017 | P/S | T | N |
| Chr9 | (125133507) rs3842787 C/T | PTGS1 SNP | 0.045 | P/L | T | N |
| Chr9 | (125133479) rs1236913 T/A | PTGS1 SNP | 0.07* | W/R | T | N |
| Chr9 | (125143973) rs5789 C/A | PTGS1 SNP | 0.015 | L/M | D | D |
| Chr9 | (125152621) rs794 G/A | PTGS1 SNP | 0.007 | V/I | D | N |
| Chr10 | (112839001) New G/T | ADRA2A SNP | 0.0013 | C/E | D | D |
| Chr19 | (3595006) rs34486470 A/C | TXBA2R SNP | 0.07 | L/R | T | N |
| Chr20 | (48129726) New G/A | PTGIS SNP | 0.006 | A/V | T | N |

To further increase sensitivity and power of the study we analyzed combined pools (N = 115 patients) with the contrasting platelet comprehensive function data (i.e. PFA-100 CEPI-CT, VerifyNow, sTxB2) indicating high on ASA platelets reactivity (HPR) in one group and pools with low platelet reactivity on ASA (LPR) (N = 115 patients) in order to find possible association with the different observed platelet function phenotypes. Table III presents MAF data for all observed non-synonymous, and Table IV presents the differences in MAF for synonymous variants.

Results

Next generation re-sequencing and analysis of pooled genomic regions

The total output from Illumina sequencer for all pooled samples included 230,516,360 reads (100 bases each) for the total of 20.64 GB which yielded large amounts of high-quality data for each pool, and captured approximately 99% of our amplified target region of approximately 360 kb (Table I Suppl.). For the variants discovery in the pooled samples we have only accepted bases with FRED quality scale of at least Q30 for a given position. The median final read depth for the analyzed coding sequences was at least x230 per combined pool (23–30 diploid genomes in each pool), which means that each base position was interrogated at least 4 times for each allele. This level of coverage is enough to provide 97% sensitivity to detect a single alternate allele in pools [63].

The total number of observed single variant polymorphisms was 1037 over 8 genes loci span (Table II). In summary, 1002 variants were located outside of coding sequences (CDSs) of the interrogated genes (i.e. in introns and other untranslated sequences, Tables III and IV Supplement) and 35 were located within the CDSs. CDS variants included 17 synonymous variants (Table IV) and 18 non-synonymous variants (Table III). From 18 detected non-synonymous variants, 3 had MAF ≥5% and 15 had MAF <5%. When compared with available databases (dbSNP ver. 141 and 1000 Genome Project), 11 non-synonymous variants were reported before, and 7 were new (i.e. unreported in available databases).

Verification of the rare missense variants by individual genotyping

Individual genotyping was performed in total 384 of ASA treated DM patients, including 230 patients from the initial discovery group (i.e. used for NGS of pooled samples), and additional 154 verification patients. The results of genotyping for 11 rare variants from the discovery phase are presented in Table III.

Table III. Exonic missense variants (N = 18) observed in the re-sequenced genes.

| Chr | (Position) and dbSNP | Base switch | Gene | Type | MAF for N = 384 | AA change | SIFT prediction | Mutation Taster prediction |
|-----|----------------------|-------------|------|------|----------------|-----------|----------------|---------------------------|
| Chr1 | (186643636) New C/T | PTGS1 SNP | 0.0013 | C/Y | D | D |
| Chr2 | (96780763) rs29000569 C/T | ADRA2B SNP | 0.015 | V/I | D | N |
| Chr2 | (96781257) rs93335368 C/G | ADRA2B SNP | 0.016 | G/A | T | N |
| Chr2 | (96781464) New A/T | ADRA2B SNP | 0.005 | V/D | D | D |
| Chr3 | (152553627) New C/T | PTGS1 SNP | 0.002 | A/V | T | N |
| Chr3 | (152554482) New C/T | PTGS1 SNP | 0.0013 | A/V | T | D |
| Chr7 | 139715641 rs1892868 C/A | TXBAS1 SNP | 0.012 | E/K | D | D |
| Chr7 | 139706915 rs3735354 A/G | TXBAS1 SNP | 0.0013 | E/K | T | D |
| Chr7 | 139652440 rs6952940 C/T | TXBAS1 SNP | 0.017 | P/S | T | N |
| Chr9 | (125133507) rs3842787 C/T | PTGS1 SNP | 0.045 | P/L | T | N |
| Chr9 | (125133479) rs1236913 T/A | PTGS1 SNP | 0.07* | W/R | T | N |
| Chr9 | (125143973) rs5789 C/A | PTGS1 SNP | 0.015 | L/M | D | D |
| Chr9 | (125152621) rs794 G/A | PTGS1 SNP | 0.007 | V/I | D | N |
| Chr10 | (112839001) New G/T | ADRA2A SNP | 0.0013 | C/E | D | D |
| Chr19 | (3595006) rs34486470 A/C | TXBA2R SNP | 0.07 | L/R | T | N |
| Chr20 | (48129726) New G/A | PTGIS SNP | 0.006 | A/V | T | N |

Variants were initially identified in discovery pools and subsequently confirmed by individual genotyping. Shaded polymorphisms indicate rare variants (defined as MAF [minor allele frequency] < 5%). For newly discovered polymorphism, instead of dbSNP (ver.141) number, the chromosomal position (Hg19) in brackets is provided. * p = 0.003, Chi-square. Functional predictions, D = damaging, T = tolerated, N = neutral.
Accumulation of non-synonymous rare variants in diabetic individuals with different platelet reactivity on ASA

The group with increased on-ASA platelet activity demonstrated higher number of minor polymorphic alleles (32 vs 21 respectively). It should be noted however, that the difference in the absolute number of minor alleles for rare polymorphic variants between investigated groups is meaningless because in majority of cases we observed more than one (i.e. singletons) for investigated polymorphic alleles. In that case the absolute differences in the minor alleles do not reflect the differences in the minor allele distribution, which, in our case, was evaluated by the genetic burden tests. When combined together, the rare missense variants for eight investigated genes showed nominal burden signal of association (p<0.05) for the group differences based on CEPI-CT assay, but not for VerifyNow or sTxB2 level assays. All observed rare (MAF < 5%) non-synonymous variants were included in the platelet tests. The burden signal of association at these eight genes could not however be attributed to a bias in the sequencing coverage between case and control pools given that all variants included in the burden tests showed coverage in each pool.

Discussion

The relative extent to which common and rare variants in the protein-coding sequence of genes associated with platelet functions could contribute to the risk of abnormal platelet reactivity in T2DM patients remains unknown. The focus of this study was to directly identify all potentially causal DNA sequence variants implicated in the pathogenesis of high platelet activity in response to ASA treatment in T2DM patients by searching for low-frequency functional genetic variation in the exons of eight selected genes involved in the platelet response to ASA. The genes investigated in this study were reported previously to harbor common genetic variants associated with altered response of platelets to ASA in diabetic patients [8–10, 20]. We used medium-scale pooled sequencing of phenotypically similar individuals to identify rare SNPs in the target ASA related platelet signaling coding sequences, and applied this information to a case-control SNP association analysis of normal (low) and abnormal (high) reactivity of platelets in response to ASA treatment. Our results identified, and subsequently verified by individual genotyping, a cluster of variants within eight analyzed genes that were differently distributed in patients with normal and high on ASA platelet activity.

A major advantage of deep re-sequencing is its potential to identify rare gene variants typically missed by linkage analysis with marker SNPs (e.g. genome-wide association studies) in which any specific rare SNP affects only a small proportion of individuals and therefore tends to be discounted. As a class however, rare SNPs are collectively common as evidenced by our observations that, of all observed non-synonymous variants, 15 (83%) had a prevalence of less than 5%, including 8 with MAF < 1% (45%). Although previous studies have identified many common variants associated with abnormal platelet response to ASA, no studies to date have searched specifically the low-frequency allele spectrum (i.e. MAF < 5%) for abnormal platelet response to ASA - associated alleles [4–10, 43]. In this study, we addressed this issue by deep re-sequencing of 8 well characterized risk genes suggested by previous studies [2, 4–7, 19, 25, 28, 33–35, 37–43]. First, we documented the occurrence of rare coding variants in a pooled sequencing study of 115 cases with normal and 115 cases with abnormal response to ASA, as measured by the composite end-point consisting of VerifyNow ARU and sTxB2 level (COX-1 dependent pathway) and PFA-100 CEPI-CT (COX-independent pathway). Secondly, we verified the presence and frequency of genotypes of the non-synonymous variants from the previous step in the wider cohort consisting of 230 patients used in the initial discovery step and 154 additional verification patients (total of 384 patients).

The main finding of the present study is that there is a statistically significant difference in the accumulation burden of rare missense variants observed in all 8 investigated genes between patients with normal and increased platelet reactivity on ASA therapy (Table V) measured by PFA-100 CEPI-CT. On the other hand, no statistically significant differences in the genetic variant burden was observed between groups defined by alternative methods (i.e. VerifyNow ARU and sTxB2 level) what is in agreement with our previous observation about the sensitivity of the PFA-100 CEPI-CT method for the detection of functional significance of the platelet genes variability [8].
Table V. Summary of results of testing by several statistical methods for association between different aggregation phenotypes and genotype data of rare missense variants from individual genotyping.

| Statistical tests | CEPI-CT > 193 sec | VerifyNow ARU | sTxB2 <1 ng/ml |
|------------------|-------------------|---------------|----------------|
| SCORE            | 0.014             | 0.11          | 0.65           |
| RBT              | 0.015             | 0.12          | 0.61           |
| ASCORE           | 0.013             | 0.20          | 0.72           |
| SEQSUM           | 0.011             | 0.07          | 0.65           |
| TTEST            | 0.016             | 0.09          | 0.64           |

CEPI-CT, collagen/epinephrine closure time; ARU, aspirin reaction units; sTxB2, serum thromboxane B2; ASA, acetylsalicylic acid. All \( p \) values were obtained by 10000 permutations (with the exception of \( p \) value for TTEST which is asymptotic). Abbreviations for statistical tests names: \( \text{SCORE} \) = score test (from logistic regression model), \( \text{RBT} \) = replication based test, \( \text{ASCORE} \) = adaptation score test, \( \text{SEQSUM} \) = sequential score sum test, \( \text{TTEST} \) = Hotelling T2 test.

Our study cannot definitively confirm which coding variants are truly causative, however it suggests that at least some SNPs will ultimately contribute to risk of increased platelet reactivity on ASA therapy in T2DM patients.

In addition to investigated functional variants, we observed 972 variants in the introns and 30 in other untranslated parts of the investigated genes. Because the original study aim has been to compare the distribution of rare missense variants we did not verify the untranslated variants by individual genotyping, so our results can only be considered as preliminary, in particular in the part describing new untranslated variants, and approximate comparison of the pooled allele frequencies in the group with normal and high platelet reactivity on ASA therapy.

Limitations

The first potential limitation of the study is related to the coverage obtained in this study which was related to rather wide genomic target size (entire genetic loci were re-sequenced, containing both translated and untranslated parts for each gene). We believe however that the density of 230 should be sufficient to detect 97% of rare variants [63]. It is possible that some rare variants were missed with this level of coverage. Next limitation is related to the choice of tests for measurement of platelet activation. Light transmission aggregometry (LTA) is considered to be the gold standard platelet function test but is poorly standardized, requires a specialist laboratory and is unlikely to be used widely in routine clinical practice [65]. Thus, in our study we assessed platelet reactivity using two point-of-care tests (CEPI-CT by PFA-100, and ARU by VerifyNow Aspirin Assay) that are being used in multiple centers instead of the “gold standard” LTA. Another potential weakness is that our set of gene sequences did not include any small indels except of the common insertion rs4066772 in the ADRA2 gene. Indel calling in NGS is generally more challenging than SNP, and could be expected to be even more difficult in the context of pooled sequencing. In addition, the detection of larger copy number variants (CNVs) would require different methods.

Finally, the significance of scattered polymorphisms in eight genes as it relates to on-ASA platelet reactivity is unclear. In our study, we focused only on few targets that may potentially affect platelet function, there are many more potential regulators of platelet reactivity that were not studied for polymorphisms, and thus, further analysis would be essential to evaluate some other genes previously associated with abnormal ASA response [3–7]. Moreover, further studies should be undertaken in order to fully elucidate observed association between genetic variation and platelet reactivity using more specific methods for platelet function assessment.

Conclusion

We provide evidence that both rare, low-frequency, and common variants with a small to moderate effect size participate in the abnormal response of platelets to ASA. Further studies in larger diabetic populations are needed to confirm and extend the results we observed in this relatively small, but ethnically homogeneous, cohort.

Acknowledgments

Marek Postula and Piotr K. Janicki contributed equally to this work. The authors would like to thank Dr Agnieszka Serafin for help in preparing the database and Dr Agnieszka Kondrakca for technical support. Research subject implemented with CEPT infrastructure financed by the European Union—the European Regional Development Fund within the Operational Program “Innovative economy” for 2007–2013.”

Disclosure of conflict of interest

The authors state that they have no conflicts of interest.

Supplemental Material

Supplemental data for this article can be accessed on the publisher’s website.

ORCID

Marek Postula http://orcid.org/0000-0002-7826-4458
Piotr K. Janicki http://orcid.org/0000-0002-1733-4192
Ceren Eyileten http://orcid.org/0000-0002-3324-9625
Marek Rosiak http://orcid.org/0000-0002-8760-0812
Agnieszka Kaplon-Cieslicka http://orcid.org/0000-0003-2020-3027
Radoslaw Wilinski http://orcid.org/0000-0002-3408-3997
Grzegorz Opolski http://orcid.org/0000-0003-4744-2554
Krzysztof J. Filipiak http://orcid.org/0000-0002-6563-0877
Dagmara Mirowska-Guzel http://orcid.org/0000-0001-6294-3256

References

1. Davi G, Patrono C. Platelet activation and atherothrombosis. N Engl J Med 2007;357:2482–2494.
2. Marec AO, Curtin RJ, Chubb A, Dolan C, Cox D, O’Brien J, Crean P, Shields DC, Fitzgerald DJ. Cyclooxygenase-1 haplotype modulates platelet response to aspirin. J Thromb Haemost 2005;3:2340–2345.
3. Wurtz M, Lordkipanidze M, Grove EL. Pharmacogenomics in cardiovascular disease: Focus on aspirin and ADP receptor antagonists. J Thromb Haemost 2013;11:1627–1639.
4. Feher G, Feher A, Pusch G, Lupkovics G, Szapary L, Papp E. The genetics of antiplatelet drug resistance. Clin Genet 2009;75:1–18.
5. Johnson AD. The genetics of common variation affecting platelet development, function and pharmaceutical targeting. J Thromb Haemost 2011;9(Suppl 1):246–257.
6. Wurtz M, Kristensen SD, Hvas AM, Grove EL. Pharmacogenetics of the antiplatelet effect of aspirin. Curr Pharm Des 2012;18:5294–308.
7. Goodman T, Ferro A, Sharma P. Pharmacogenetics of aspirin resistance: a comprehensive systematic review. Br J Clin Pharmacol 2008;66:222–32.
Rarity genetics and platelet reactivity in T2D

10. Postula M, Janicki PK, Rosiak M, Kaplon-Cieslicka A, Kondracka A, Marth GT, Yu F, Indap AR, Garimella K, Gravel S, Leong WF, Kryukov GV, Pennacchio LA, Sunyaev SR. Most rare missense genes. J Thromb Haemost 2011;9:2291–2301.

23. Kimouli M, Gourvas V, Konstantoudaki X, Basta M, Miyakis S, Takahashi S, Ushida M, Komine R, Shimodaira A, Uchida T, Halushka MK, Walker LP, Chakravarti A, Filipiak KJ, Kosior DA, Czlonkowski A, Opolski G. Genetic determinants of platelet reactivity during acetylsalicylic acid therapy in diabetic patients: evaluation of 27 polymorphisms within candidate genes. J Thromb Haemost 2011;9:2291–2301.

26. Fujiwara T, Ikeda M, Esumi K, Fujita TD, Kono M, Tokushige H, Hatoyama T, Maeda T, Asai T, Ogawa T, et al. Exploratory aspirin resistance trial in healthy Japanese volunteers (J-ART) using platelet aggregation as a measure of thrombogenicity. Pharm J 2007;7:395–403.

29. Meen O, Brostad F, Hibianni H, Gjertsen E, Lauritsen ME, Pedersen TM, Bjornsen S, Schjelderup NM, Ameln L, Ge C, et al. No case of COX-1-related aspirin resistance found in 289 patients with symptoms of stable CHD remitted for coronary artery disease. Scand J Ther Pract 2008;68:185–191.

32. Jefferson BK, Foster JH, McCarthy JJ, Ginsburg G, Parker A, Michelson AD. Residual arachidonic acid-induced platelet activation via an adenosine diphosphate-dependent but cyclooxygenase-1 and cyclooxygenase-2-independent pathway: A 700-patient study of aspirin resistance. Circulation 2006;113:2888–2896.

35. Johnson AD, Yanek LR, Chen MH, Faraday N, Larson MG, Tofler GH. Cyclooxygenase polymorphisms and risk of cardiovascular events: A 700-patient study of aspirin resistance via an adenosine diphosphate-dependent but cyclooxygenase-1 and cyclooxygenase-2-independent pathway: A 700-patient study of aspirin resistance. Circulation 2006;113:2888–2896.

38. Lee CR, North KE, Bray MS, Couper DJ, Heiss G, Zeldin DC. Coronary artery disease patients. Thromb Haemost 2006;95:253–259.

40. Chevalier D, Lo-Guidice JM, Seralle E, Allorge D, Debusse D, Ferrari N, Libersa C, Lhermitte M, Broly F. Identification of functional variants in the human thromboxane synthase (CYP1A1) gene in healthy volunteers. Thromb Haemost 2006;96:356–360.

43. Faraday N, Becker DM, Becker LC. Pharmacogenomics of platelet responsiveness to aspirin. Pharmacogenomics 2007;8:1413–1425.
44. Blankenberg D, Gordon A, Von Kuster G, Coraor N, Taylor J, Nekrutenko A. Manipulation of FASTQ data with galaxy. Bioinformatics 2010;26:1783–1785.

45. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–1760.

46. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–2079.

47. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012. Preprint at arXiv:1207.3907v2 [qbio. GN]

48. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: A map reduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–1303.

49. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491–498.

50. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Rouzen D, Thibault J, et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinf 2013;11(1110): 11.10.1–11.10.33.

51. Chang X, Wang K. wANNOVAR: Annotating genetic variants for personal genomes via the web. J Med Genet 2012;49:433–436.

52. Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.

53. Janicki PK, Vealey R, Liu J, Escajeda J, Postula M, Welker K. Genome-wide association study using pooled DNA to identify candidate markers mediating susceptibility to postoperative nausea and vomiting. Anesthesiology 2011;115:54–64.

54. Clayton D, Chapman J, Cooper J. Use of unphased multilocus genotype data in indirect association studies. Genet Epidemiol 2004;27:415–428.

55. Chapman J, Whittaker J. Analysis of multiple SNPs in a candidate gene or region. Genet Epidemiol 2008;32:560–566.

56. Liu X, Jian X, Boerwinkle E. dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat 2011;32:894–899.

57. Han F, Pan W. A data-adaptive sum test for disease association with multiple common or rare variants. Hum Hered 2010;70:42–54.

58. Pan W, Kim J, Zhang Y, Shen X, Wei P. A powerful and adaptive association test for rare variants. Genetics 2014;197:1081–1095.

59. Basu S, Pan W. Comparison of statistical tests for disease association with rare variants. Genet Epidemiol 2011;35:606–619.

60. Xiong M, Zhao J, Boerwinkle E. Generalized T2 test for genome association studies. Am J Hum Genet 2002;70:1257–1268.

61. Purcell S, Cherny SS, Sham PC. Genetic power calculator: Design of linkage and association genetic mapping studies of complex traits. Bioinformatics 2003;19:149–150.

62. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet 2014;15:335–346.

63. Pihlstrom L, Rengmark A, Bjornara KA, Toft M. Effective variant detection by targeted deep sequencing of DNA pools: An example from Parkinson’s disease. Ann Hum Genet 2014;78:243–252.

64. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: A database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat 2013;34:E2393–E2402.

65. Michelson AD. Methods for the measurement of platelet function. Am J Cardiol 2009;103:20A–26A.