Diversity and impact of rare variants in genes encoding the platelet G protein-coupled receptors

Matthew L. Jones1; Jane E. Norman1; Neil V. Morgan2; Stuart J. Mundell3; Marie Lordkipanidzé4; Gillian C. Lowe3; Martina E. Daly4; Michael A. Simpson5; Sian Drake2; Steve P. Watson2; Andrew D. Mumford1; on behalf of the UK GAPP study group

1School of Cellular and Molecular Medicine, University of Bristol, Level 7 Bristol Royal Infirmary, Bristol, UK; 2School of Physiology and Pharmacology, University of Bristol, Bristol, UK; 3Centre for Cardiovascular Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK; 4Centre de recherche de l’Institut de cardiologie de Montréal, and Faculté de pharmacie, Université de Montréal, Quebec, Canada; 5Department of Cardiovascular Science, University of Sheffield Medical School, Sheffield, UK; 6Division of Genetics and Molecular Medicine, King’s College, London, UK

Introduction

G protein-coupled receptors (GPCRs) are seven transmembrane domain proteins that mediate signal transduction from a wide range of extracellular stimuli. GPCRs are expressed widely in hematopoietic and vascular tissues, including platelets, in which they mediate activation signals from agonists such as thrombin (protease activated receptors [PAR] 1 and 4), thromboxane A2 (thromboxane A2 receptor [TP]), epinephrine (α2A-adrenoreceptor) and ADP (P2Y12 and P2Y1 receptors). Platelets also express Gαi-coupled GPCRs such as the prostacyclin (IP3), adenosine 2A (A2A) and prostaglandin D2 (DP1) receptors, which mediate inhibitory signals from prostacyclin, adenosine and PGD2 respectively, to suppress platelet activation.

Platelet GPCR activity varies between individuals within the population, in part because of common genetic sequence variants (minor allele frequency (MAF) ≥ 1.0%) near, or within GPCR genes. Examples include the variant rs1472122 (downstream of the P2Y12 gene P2RY12), which affects ADP-induced platelet fibrinogen binding and P-selectin exposure (1) and the variant rs4311994 (downstream of the α2A-adrenoreceptor gene ADRA2A), which affects epinephrine-induced platelet aggregation (2). Similar associations have been demonstrated between common variants in the PAR1 (P2R), PAR4 (F2RL3) and the TP receptor (TBXA2R) genes and function of the corresponding GPCRs (3–5). Some common variants also influence susceptibility to cardiovascular disease and responses to anti-platelet drugs (4–6). Since the common GPCR gene variants lie exclusively in non-coding regions, these effects are most likely caused by changes in receptor expression, and not altered receptor function (4, 7).

Although the evidence linking common variants near platelet GPCR genes and GPCR activity is compelling, the individual effect size of common variants is small (2). For other genes, rare (MAF <1%) single nucleotide variants (SNVs), with large individual
effect size, provide a greater source of inter-individual genetic variation than common variants (8–10). However, for platelet GPCR genes, descriptions of rare variants affecting platelet function are restricted to SNVs in P2RY12 and TBXA2R in isolated pedigrees with inherited platelet function disorders (IPFD) (11–17). It is likely that the impact of rare GPCR gene variants in the population is much greater than implied from these limited descriptions, but this has not been confirmed by systematic analysis. In order to assess the population diversity and impact of rare SNVs in platelet GPCR genes, we have surveyed and annotated coding and splice region SNVs in public databases of 7595 individuals and in 31 cases with IPFD of unknown genetic basis.

**Materials and methods**

**G protein-coupled receptors in human platelets**

Class A GPCRs that were listed in the International Union of Basic and Clinical Pharmacology (IUPHAR) GPCR Database (Suppl. Table 1, available online at www.thrombosis-online.com) were selected for analysis if present in the Proteomics Identifications Database (PRIDE), the PlateletWeb resource (Suppl. Table 1, available online at www.thrombosis-online.com) and in the human platelet transcriptome with >1.0 reads per kilobase of exon model per million mapped reads (18).

**GPCR gene variations in population datasets**

We identified coding sequence and splice region (from 3 exonic to 8 intronic nucleotides flanking the exon-intron boundaries) SNVs in the GPCR gene shortlist in the April 2012 Integrated Variant Set release of the 1,000 Genomes project and the NHLBI Exome Sequencing Project (ESP) dataset release number ESP6500, accessed through Ensembl Variation 74 (H. sapiens Short Variation GRCh37.p13 dataset) using the BioMart tool (Suppl. Table 1, available online at www.thrombosis-online.com). Nucleotide variations were annotated to the consensus coding sequence (CCDS) database transcript of each platelet GPCR.

**GPCR gene variations in inherited platelet function disorders**

Genomic DNA was isolated from peripheral venous blood from a representative sub-group of 31 unrelated cases with IPFD recruited at UK Haemophilia Comprehensive Care Centres to the Genotyping and Phenotyping of Platelets (GAPP) study (ISRCTN 77951167, UK REC 06/MRE07/36) according to previously reported eligibility criteria (19). For all cases, platelet function was evaluated using light transmission aggregation and ATP secretion assays using nine agonists at least two weeks after exposure to 19 drugs known to affect platelet function (19, 20). Genomic DNA was enriched for the target GPCR genes either using a custom made bait library for platelet genes (21) or the Agilent SureSelect All Exon 50Mb kit (Agilent Technologies, Wokingham, UK). Sequence data were captured using an Illumina HiSeq 2000 analyser (Illumina Inc San Diego, CA, USA). Sequence reads were mapped to the reference genome GRCh37.p11, Feb 2009 and SNVs were annotated to the consensus CCDS records using the ANNOVAR tool (Suppl. Table 1, available online at www.thrombosis-online.com). Since the IPFD cases all showed reduced platelet responses, we analysed stimulatory platelet GPCRs and excluded the Gα2βγ- coupled inhibitory GPCRs IP1, DP1 receptor and A2A receptor. All potentially damaging SNVs were confirmed by PCR amplification of individual exons and direct cycle sequencing.

**Functional annotation of GPCR gene variants using computational algorithms**

SNVs that were identified in population databases and in cases with IPFD were classified according to sequence ontology terminology used in Ensembl release 74 (Suppl. Table 1, available online at www.thrombosis-online.com). The likely pathogenicity of each SNV was determined using the MAP, PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT and SNAP prediction tools on the PredictSNP server (Suppl. Table 1, available online at www.thrombosis-online.com). SNVs were classified as potentially damaging if identified as ‘damaging’ by the PredictSNP meta-analysis tool with a consensus likelihood of >0.5 (22). Splice region variants were analysed using the Human SpliceFinder tool (Suppl. Table 1, available online at www.thrombosis-online.com) and were classified as potentially damaging if the difference between the splice site prediction scores of the wild type and variant sequences exceeded 30% that of the wild-type sequence (23).

**Manual functional annotation of GPCR gene variants**

Missense SNVs were also annotated using a manual strategy in which variants were classified as potentially damaging if any of the following criteria were met:

- The substituted amino-acid was within a functional GPCR sequence motif identified in UniProt (Suppl. Tables 1 and 2, available online at www.thrombosis-online.com).
- The substituted amino-acid, expressed in Ballesteros-Weinstein nomenclature, contributed to inter-helical interactions, the ligand binding pocket or to the G-protein binding sites in the consensus Class A GPCR structure (25) or in published crystal structures for the PAR1 (26), A2A (27) or P2Y12 receptors (28) (Suppl. Table 2, available online at www.thrombosis-online.com).
- There was published experimental evidence of a change in GPCR function from site-directed mutagenesis in a heterologous system, determined from the GPCRDB resource (Suppl. Table 1, available online at www.thrombosis-online.com).

**Analysis of the P2Y12, R122H and V207A variants in transfected cells**

R122H and V207A HA-tagged human P2Y12 constructs were generated by site-directed mutagenesis (Eurofins MWG Operon, Ebersberg, Germany) and were transfected into either HEK293 or...
1321N1 cells according to previously described methods (17). Cell surface P2Y12 expression in the transfected cells was determined by enzyme linked immunosorbent assay (ELISA) and by immunofluorescence microscopy using murine anti-HA antibody (HA-11) as described previously (17). P2Y12 receptor function was measured by incubating the transfected cells with 1 µM forskolin (Sigma-Aldrich, Gillingham, UK) to increase basal cAMP levels. The cells were then incubated with 50 µM-10 nM ADP before residual cAMP concentrations were determined in cell lysates by ELISA (Sigma-Aldrich cAMP Enzyme Immunoassay Kit, Gillingham, UK).

**Results**

**Identification of GPCRs in human platelets**

Using the IUPHAR database, we identified 18 Class A GPCRs with robust evidence of expression in human platelets at transcript and protein levels. The coding regions of the 18 GPCR genes had median length of 1121 kb (interquartile range [IQR] 1043–1248) and median GC content of 56.4 % (IQR 49.0–64.6; ▶ Table 1).

**GPCR gene variations in population datasets**

In order to assess the allelic diversity of the platelet GPCR genes, we surveyed the 1,000 genomes (1,092 subjects) and ESP (6,503 subjects) datasets for coding and splice region SNVs in the 18 selected GPCR genes. We identified 740 SNVs in the target regions (median 41.5 [IQR 30–48] SNVs per GPCR gene) of which 318 (43 %) were synonymous, 410 (55 %) were missense and 7 (1 %) were stop-gain (▶ Figure 1A). There were five intronic splice region SNVs and 1 exonic splice region SNP that was also a synonymous coding region SNP. Amongst the 740 SNVs in the target regions, 58 (8 %) had global MAF ≥1 %, 163 (22 %) had MAF 0.99–0.05 % and 519 (70 %) had global MAF <0.05 % or were singletons, indicating very low population frequency.

**Predicting the functional impact of GPCR gene variants**

We used both computational and manual annotation to assess whether missense SNVs in the GPCR gene target regions were potentially damaging to GPCR function. Computational annotation using the PredictSNP server enabled meta-analysis of predictions from six tools that utilise trained decision (PhD-SNP, Polyphen-2 and SNAP), evolutionary conservation (SIFT), physicochemical (MAPP) and expert rule (Polyphen-1) algorithms to generate a consensus likelihood of pathogenicity for each SNV (22). Using this strategy, 122 (30 %) of the 410 missense SNVs in the GPCR gene target regions were classified as potentially damaging (Suppl. Table 3, available online at www.thrombosis-online.com). None of the SNVs had a consensus likelihood of pathogenicity of >0.99, indicating that the remaining 288 (71 %) SNVs were not functionally damaging, or were benign.

**Table 1: Platelet G-protein coupled receptors.** The table lists the 18 G-protein-coupled receptors described in IUPHAR terminology, listed in order of decreasing transcript abundance in human megakaryocytes (18). †Principal coupled G protein. ‡Major CCDS annotated transcript with experimental evidence in megakaryocytes.
the six splice region SNVs were predicted by computation to disrupt transcript splicing.

Our manual annotation strategy classified missense SNVs as potentially damaging if the predicted amino-acid substitution affected a functional GPCR sequence motif or a critical residue in the consensus or specific GPCR crystal structures or if previous experimental mutagenesis of the residue caused loss of receptor function. This identified 60 (15%) of the 410 missense SNVs in the GPCR gene target regions as potentially damaging (Suppl. Table 3, available online at www.thrombosis-online.com). Seven stop-gain SNVs were also classified as potentially damaging by manual annotation since they predicted protein truncation.

The total number of all classes of SNV that were classified as potentially damaging by either computational or manual annotation was 156 (21% of all SNVs; Suppl. Table 3, available online at www.thrombosis-online.com). Forty missense SNVs were classi-
fied as potentially damaging by both computational and manual annotation (Suppl. Table 2, available online at www.thrombosis-online.com).

**Distribution of damaging missense GPCR gene variants**

The 149 potentially damaging missense SNVs were represented in all of the 18 selected GPCR genes (Figure 1B) and predicted amino-acid substitutions that were more common in the TM domains and C-terminal intracellular region (CT) than other regions (Figure 2). Twenty-five SNVs predicted amino-acid substitutions at sites shown in the consensus Class A GPCR structure to contribute to inter-helical interactions between the TM domains.

![Distribution of potentially damaging missense SNVs across GPCR domains](image)

**Characteristics of patients with inherited platelet function disorders**

The IPFD collection comprised 31 unrelated cases (11 males and 20 females; age range 6–82 years) with abnormal platelet function determined by light transmission aggregation and ATP release assays (19, 20). The collection comprised cases in which the main laboratory defect was within the Gi-pathway (n=15), secretion pathway (n=11) and thromboxane synthesis pathway (n=3) according to previous diagnostic criteria (19). Two cases showed complex defects that could not be classified. This collection was selected as a representative sub-group of a larger collection of 111 previously reported cases with inherited platelet function disorders enrolled into the UK GAPP study and showed a similar distribution of pathway defects to the group as a whole (19).

**GPCR gene variations in cases with inherited platelet disorders**

Among the 31 cases with IPFD, we identified 256 SNVs in the target regions of the genes encoding the stimulatory platelet GPCRs PAR1, P2Y12, TRP, LPA5, CXCR4, PAR4, P2Y1, α2A-adrenoceptor, CCR4, V1a receptor, PAF receptor, FP1, EP3 receptor, succinate, and 5-HT2A receptor. These comprised 38 individual SNVs of which 22 (58.9%) were synonymous and 16 (42.1%) were missense. There were no stop-gain or splice region SNVs. Thirty-four unique SNVs were present in the 1,000 genomes and ESP population datasets (12 with global MAF < 1% and 22 with global MAF ≥ 1%) and four were undocumented. Using an identical strategy to the analysis of the population datasets, we classified three heterozygous missense SNVs as potentially damaging in the IPFD cases, all within P2RY12 (Table 3). A wider analysis of variants identified in other platelet genes did not identify any single candidate variants that could completely account for the platelet phenotype of each case.

**Characteristics of cases with P2Y12 variants**

The P2Y12 R122H variant was identified in a female index case 1.1 (Figure 3A) with a history of prolonged bleeding from minor wounds and after a vaginal delivery. There was no abnormal bleeding after two other vaginal deliveries or after tonsillectomy. The P2Y12 P258T variant was identified in an unrelated male index case 2.1 (Figure 3D) who had experienced recurrent gastrointestinal bleeding throughout adulthood but had no other bleeding symptoms. Platelets from case 1.1 and from case 2.1 showed normal shape change but reduced aggregation responses to 10–100 μM ADP compared to healthy controls that was reversible with 10 μM ADP (Figure 3B and E), indicating selective loss of P2Y12 function. Compared to control subjects, platelets from both cases...
also showed reduced aggregation responses to 3–30 µM epinephrine and 1 µg/ml collagen, but not to 3 µg/ml collagen, which are consistent with loss of P2Y₁₁ function. However, there were also reduced aggregation responses to 0.5–1 mM arachidonic acid in case 1.1 and reduced responses to ristocetin 1.25–1.5 mg/ml and a markedly reduced response to high concentration (100 µM) epinephrine in case 2.1. The latter findings indicate that cases 1.1 and 2.1 have distinct and complex aggregation phenotypes, neither of which can be completely explained by loss of P2Y₁₁ function. Platelets from other pedigree members 1.2 and 2.2, analysed in parallel with the respective index cases, also showed reduced aggregation responses to ADP compared with controls (▶ Figure 3C and F). Cases 1.2 and 2.2 were subsequently shown to harbour the R122H and P258T variations respectively, but neither had abnormal bleeding symptoms.

The P2Y₁₁ V207A variant was identified in an asymptomatic female index case 3.1 (▶ Figure 3G) who also harboured a P2Y₁₂ SNV on the same allele that predicted a T223R substitution, that was classified as benign. Platelets from 3.1 showed normal platelet shape change but reduced aggregation responses to 5–20 µM ADP (▶ Figure 3H) and 5–10 µM epinephrine compared to controls. The maximum amplitude of responses to ADP and epinephrine

### Table 2: Potentially damaging missense variants in platelet GPCR genes in the 1,000 Genomes and ESP6500 databases

The table lists the subgroup missense SNV classified as potentially deleterious by both manual and computational annotation. For each missense SNV the results of six independent annotation tools: P-SNP (the consensus classifier), MAPP, PhD-SNP, PP-1, PP-2, SIFT and SNAP are indicated as + for ‘damaging’ and – for ‘not damaging’. / indicates the tool did not return a prediction for the SNV.

| GPCR | Variation | Substitution | Region | Manual annotation | P-SNP | MAPP | PhD-SNP | PP-1 | PP-2 | SIFT | SNAP |
|------|-----------|--------------|--------|-------------------|-------|------|---------|------|------|------|------|
| PAR1 | rs144474562 | Y266C       | EL2    | Tethered ligand binding region | +     | /    | +       | +    | +    | -    | +    |
| PAR1 | rs372280945 | L354F       | TM7    | Consensus ligand binding pocket | +     | /    | -       | +    | +    | -    | -    |
| P2Y₁₂ | rs372954515 | T283I       | TM7    | TM6-TM7 conserved interaction | +     | +    | +       | +    | +    | +    | -    |
| P2Y₁₂ | rs367926037 | M108L       | TM3    | TM3-TM6 conserved interaction; ligand binding pocket | +     | +    | +       | +    | +    | +    | -    |
| P2Y₁₂ | rs202099742 | P258T       | EL3    | Identified in subject with IPFD | +     | -    | +       | +    | +    | +    | +    |
| IP₁  | rs147448416 | R77C        | TM2    | Reduced surface expression (reduced Bmax) | +     | /    | +       | +    | +    | +    | +    |
| IP₁  | rs4987262  | R212C       | IL3    | Reduced surface expression (reduced Bmax) | +     | /    | +       | +    | +    | +    | +    |
| IP₁  | rs201340109 | R215C       | IL3    | Reduced surface expression (reduced Bmax) | +     | /    | +       | +    | +    | +    | +    |
| IP₁  | rs370909150 | H237R       | TM6    | Site of G protein interaction within helices | +     | /    | +       | +    | +    | +    | +    |
| TPα  | rs34377097 | R60L        | ICL1   | TM2-TM3 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| TPα  | rs201421330 | M126T       | TM3    | TM3-TM2 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| TPα  | rs372994525 | S218G       | TM5    | Site of G protein interaction within helices | +     | +    | -       | +    | +    | +    | +    |
| TPα  | rs61731124 | T286M       | TM7    | Consensus ligand binding pocket | +     | -    | +       | +    | +    | +    | +    |
| TPα  | rs374635591 | P305L       | TM7    | TM1-TM7 conserved interaction | +     | +    | +       | +    | +    | +    | +    |
| TPα  | rs370418735 | R326W       | CT     | Serine phosphorylation site | +     | -    | -       | +    | +    | +    | +    |
| LPAR₁ | rs149664830 | M103K       | TM3    | TM3-TM4 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| LPAR₁ | rs187536858 | P294L       | TM7    | TM1-TM7 conserved interaction | +     | +    | +       | +    | +    | +    | +    |
| LPAR₁ | rs374800227 | Y298H       | TM7    | TM7-TM2 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| A₂₃₃₅ | rs142560733 | R293C       | CT     | Minus 5 from Thr phosphorylation site | +     | +    | -       | +    | +    | +    | +    |
| CXC4 | rs368016542 | D84H        | TM2    | TM1-TM2 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| CXC4 | rs367718547 | D193Y       | EL2    | Mediates dimerization | +     | -    | -       | +    | +    | -    | -    |
| PAR4 | rs374965245 | G13E        | NT     | Signal peptide | +     | -    | +       | -    | +    | +    | +    |
| PAR4 | rs201697829 | R47H        | NT     | Site of cleavage of tethered ligand | +     | /    | -       | +    | +    | -    | -    |
| PAR4 | rs2227346  | F296V       | TM6    | TM3-TM6 conserved interaction; ligand binding pocket | +     | +    | -       | +    | +    | +    | +    |

© CC-BY This is an open access article under the terms of the Creative Commons License, which permits use and distribution in any medium, provided the original work is properly cited.
Table 2: Continued

| GPCR | Variation | Substitution | Region | Manual annotation | P-SNP | MAPP | PhD-SNP | PP-1 | PP-2 | SIFT | SNAP |
|------|-----------|--------------|--------|-----------------|-------|------|---------|------|------|------|------|
| PAR4 | rs111890288 | S329C        | TM7    | Consensus ligand binding pocket | +     | -    | -       | +    | +    | +    | -    |
| a1A  | rs149350078 | I234M        | TM5    | Site of G protein interaction within helices | +     | +    | +       | +    | +    | +    | +    |
| a1A  | rs370313798 | R383G        | TM6    | Site of G protein interaction within helices | +     | +    | +       | +    | +    | +    | +    |
| a1A  | rs375454021 | V390M        | TM6    | Site of G protein interaction within helices | +     | +    | -       | +    | +    | +    | -    |
| CCR4 | rs200003145 | A240V        | TM6    | Site of G protein interaction within helices | +     | -    | +       | +    | +    | +    | +    |
| DP1  | rs145604058 | T69M         | TM2    | TM1-TM2 conserved interaction | +     | +    | -       | -    | +    | +    | -    |
| DP1  | rs370262391 | A127T        | TM3    | TM3-TM5 conserved interaction | +     | +    | +       | -    | +    | +    | +    |
| V1A  | rs180760072 | F133C        | TM3    | TM3-TM4 conserved interaction | +     | +    | +       | -    | +    | +    | +    |
| V1A  | rs369710823 | M145T        | TM3    | TM3-TM2 conserved interaction | +     | -    | +       | -    | -    | -    | -    |
| V1A  | rs377107751 | V153L        | TM3    | Site of G protein interaction within helices | +     | +    | +       | -    | -    | +    | -    |
| PAF  | rs139524224 | R229Q        | TM6    | Site of G protein interaction within helices | +     |       |         |      |      |      |      |
| FPR1 | rs369354920 | D122H        | TM3    | E/DRY motif       | +     | +    | -       | +    | +    | +    | +    |
| FPR1 | rs149931707 | R123H        | TM3    | Site of G protein interaction within helices | +     | +    | -       | +    | +    | +    | +    |
| FPR1 | rs368075541 | R238H        | TM6    | Site of G protein interaction within helices | +     | -    | +       | -    | +    | +    | +    |
| succinate | rs371440145 | N41K         | TM1    | TM1-TM2 conserved interaction | +     | +    | +       | +    | +    | +    | +    |
| succinate | rs374436394 | I240N        | TM6    | TM3-TM6 conserved interaction | +     | +    | -       | +    | +    | +    | +    |
| succinate | rs142256005 | R252Q        | TM6    | Agonist dependent activation | +     | -    | +       | +    | +    | +    | +    |
| succinate | rs142852744 | R291Q        | TM7    | Agonist dependent activation; TM6-TM7 interaction | +     | -    | +       | +    | +    | +    | +    |
| 5-HT2A | rs375024989 | N38K         | NT     | N-linked glycosylation site | +     | -    | -       | +    | +    | +    | +    |

were within the reference interval of responses determined from a panel of 30 locally recruited healthy controls, but fell within the lowest 10th percentile of control responses, consistent with reduced P2Y12 function, that was less pronounced than index cases 1.1 and 2.1. Platelets from case 3.1 also showed reduced aggregation responses to 0.5–1 mM arachidonic acid suggesting an additional platelet defect. A pedigree member 3.2 with wild-type P2Y12 showed platelet aggregation responses to ADP that were similar to control subjects (►Figure 31).

**Analysis of the P2Y12 R122H and V207A in HEK-293 cells**

Since the R122H and V207A substitutions had not been previously associated with P2Y12 receptor deficiency in humans, we examined the phenotype of these substituted P2Y12 receptors in transfected cells. Expression of P2Y12 R122H and V207A was observed predominantly at the cell surface by immunofluorescence microscopy (data not shown). When cell-surface expression was quantified by ELISA, the normalised expression levels of the substituted receptors were almost identical to that of wild-type receptor (R122H mean 113% ± S.E.M. 14.5% and V207A 97.5% ± 7.9%; ►Figure 4A) indicating that neither substitution significantly affected P2Y12 receptor trafficking. When P2Y12 receptor function was tested by measuring the ability of ADP to reduce cellular cAMP, the substituted P2Y12 receptors showed less reduction in cAMP at ADP concentrations of 1 µM to 10nM compared to P2Y12 wild-type (p=0.013 for R122H and p=0.019 for V207A; 1 way ANOVA; ►Figure 4B). These data indicate that both substitutions reduce P2Y12 function, with a weaker effect from P2Y12 V207A substitution, consistent with the less marked platelet aggregation defect.

© CC-BY This is an open access article under the terms of the Creative Commons License, which permits use and distribution in any medium, provided the original work is properly cited.
Table 3: Potentially damaging missense SNV in P2RY12 in cases with IPFD. This table lists the missense SNV found in P2RY12 in cases with a demonstrable platelet aggregation defect. For each missense SNV the results of six independent annotation tools: P-SNP (the consensus classifier), MAPP, PhD-SNP, PP-1, PP-2, SIFT and SNAP are indicated as + for ‘damaging’ and – for ‘not damaging’.

| Substitution | Variation | Manual annotation | P-SNP | MAPP | PhD-SNP | PP-1 | PP-2 | SIFT | SNAP |
|--------------|-----------|-------------------|-------|------|---------|------|------|------|------|
| R122H        | novel     | DRY motif         | +     | +    | +       | +    | +    | +    | +    |
| P258T        | rs202099742 | TM6; adjacent to Y259 that contributes to ligand binding | +     | -    | -       | +    | +    | +    | +    |
| V207A        | rs370983746 | TM5; adjacent to C208 that has multiple binding contacts with TM3 | +     | +    | -       | -    | +    | +    | +    |

Discussion

We have reported the results of a unique survey of coding and splice region SNVs in 18 platelet GPCR genes from 7,595 exomes in the 1,000 genomes and ESP databases and from 31 cases with IPFD. Our main findings were that: (i) in the population databases, the GPCR gene target regions contained potentially damaging SNVs that were individually rare, but collectively numerous; ii) the potentially damaging SNVs were diverse and were predicted to alter GPCR activity through several mechanisms, and, iii) a representative collection of cases with IPFD also had SNVs in platelet GPCR genes, including potentially damaging variants affecting P2Y12 in three cases.

Our strategy for identifying potentially damaging SNVs was based on computational annotation using six bioinformatic tools with different methodologies (22), complemented by manual annotation using resources that are unique to GPCRs. These included the GPCRDB database that catalogues previous GPCR mutagenesis experiments, the high resolution structures for the PAR1, A{sub}2{sub}A and P2Y{sub}12 receptors (26–28) and the consensus structure for Class A GPCRs (25) that provides structural data for GPCRs with unsolved crystal structures. Combined computational and manual annotation has provided a valuable insight into the diversity and impact of rare variants in human GPCR genes. However, our analysis has focussed on missense rather than synonymous coding region SNVs. Since 6% of all synonymous SNVs in the ESP exome dataset were computed to be potentially damaging, primarily through codon usage effects (10, 29), our analysis is likely to have underestimated the overall burden of GPCR gene variation.

Within the ESP and 1,000 genomes databases, we found 740 SNVs in the GPCR gene target regions, of which 56% were missense and 70% had a global MAF <0.05% or were singleton records. These characteristics are similar to the entire ESP exome dataset comprising >500,000 SNVs, of which 58% are missense and 72% are present in only three alleles or less (10), indicating an exome-wide abundance of rare missense variants. Our prediction that 21% of SNVs in the GPCR gene target regions were potentially damaging, is also similar to exome-wide estimates of 17% determined by computation (10). One noteworthy finding from our survey is that the platelet GPCR genes contained a median of 41.5 SNVs per coding region, compared with 24 SNVs per coding region exome-wide (10). The high variation rate in GPCR genes cannot be explained by differences in the length of coding region because the GPCR genes had median coding length 1121 bp, similar to the exome median of 1,100 bp (30). However, this difference could be related to GC content (31) which was 56.4% in platelet GPCR genes compared to 51% exome-wide (32). Consistent with this, the GPCR genes F2RL3 and PTGIR with high GC content, had more SNVs than others with lower GC content, although this trend was inconsistent across all platelet GPCRs. We also showed that SNVs classified as potentially damaging were more common in gene regions encoding GPCR TM helices and CT intracellular regions compared to other areas. This reflects the essential roles of the TM helices in maintaining GPCR tertiary structure and defining the ligand binding and G-protein interaction sites (25) and the CT intracellular regions in regulating GPCR signalling and trafficking (33).

Consistent with the population databases, the 31 cases with IPFD also harbour rare missense SNVs in genes encoding the stimulatory platelet GPCRs. Although these were represented in all of the target GPCR genes, the three variants that were predicted to be potentially damaging were exclusively in P2RY12 and occurred as heterozygous traits.

These included the P2Y12 P258T substitution which occurs adjacent to Y259 in TM6 that is required for ligand binding (28). This substitution was identified and characterised previously in an unrelated IPFD pedigree who also displayed reduced platelet responses to ADP (34), identical to the phenotype in the P2Y12 P258T pedigree in our study. Since this independent data provides good evidence that the P258T variation is causally related to loss of P2Y12 receptor function, we performed no further characterisation.

The other observed P2Y12 variants had not been previously reported. These included an SNV predicting an R122H substitution within the P2Y12 DRY motif which has multiple postulated roles in regulating receptor conformation, G-protein interactions and receptor trafficking (35). This substitution also occurs at a residue affected by a different substitution (P2Y12 R122C) in a previously reported IPFD pedigree with P2Y12 dysfunction (17). The P2Y12 V207A substitution affects a residue not previously associated with an IPFD, but which is adjacent to C208 in TM5 which has multiple interactions with TM3 and is, thereby, required for receptor structural integrity (25). Consistent with these significant structural predictions, we confirmed that the P2Y12 V207A and R122H substitutions were responsible for loss of P2Y12 receptor function by demonstrating diminished ADP-mediated reduction in
Figure 3: Platelet phenotype in cases with potentially damaging SNV in P2RY12. A) The pedigree of index case 1.1, who was heterozygous for the P2Y12 R122H substitution (half back shading), showing pedigree member 1.2, also with the P2Y12 R122H substitution. B) Light transmission aggregation (LTA) responses to 10–100 µM ADP in platelet-rich plasma from 1.1 and a healthy control (HC). C) LTA responses to 30 µM ADP in 1.2. D) The pedigree of index case 2.1, who was heterozygous for the P2Y12 P258T substitution (half black shading) showing pedigree member 2.2, also with the P2Y12, P258T substitution. E) LTA response to 10–100 µM ADP in platelet-rich plasma from 2.1 and an HC. F) LTA responses to 30 µM ADP in 2.2 and a HC. G) The pedigree of the index case 3.1 who was heterozygous for the P2Y12 V207A and T223R substitutions (indicated by half black and half grey shading), showing pedigree member 3.2 with wild-type P2Y12. H) LTA response to 2–10 µM ADP in platelet-rich plasma from 1.1 and an HC. I) LTA responses to 2 µM ADP in 3.2 and a HC.
cytoplasmic cAMP levels in transfected cells, which is a sensitive and highly specific measure of P2Y12 function (36). Our demonstration that cell surface expression of the substituted P2Y12 receptors was the same as wild-type suggest that both V207A and R122H disrupt function by impairing ligand binding, receptor activation or signal transduction, rather than by affecting receptor trafficking.

It is noteworthy that in the IPFD cases in this series, the R122H and P258T substitutions, and to a less pronounced extent the V207A substitution, were associated with reduced platelet responses to ADP consistent with impaired P2Y12 function. Platelet responses to low concentrations of other activating agonists were also reduced, in keeping with impaired P2Y12-mediated positive feedback from ADP released via dense granules, and simi-
lar to previous IPFD cases with loss-of-function P2Y12 variants (11, 34, 37). Despite this platelet phenotype, there was an inconsistent relationship between the heterozygous P2Y12 variants and abnormal bleeding, suggesting that partial loss of P2Y12 function alone is insufficient to affect haemostasis. It is also noteworthy that in all three index cases with potentially damaging P2Y12 variants, there were abnormal responses to other activating agonists that may not be explained solely by loss of P2Y12 function because of the magnitude of the other defects. Our data do not allow us to exclude a dominant negative effect from the observed heterozygous P2Y12 variants, or a further non-coding P2Y12 variant in trans that reduces expression of the other allele. However, a more plausible explanation is that in addition to a variant affecting P2Y12, the IPFD cases also harboured loss-of-function variants in other platelet genes that contributed to bleeding and to the complex laboratory phenotypes. The concept that some IPFD have complex heritability is supported by previous descriptions of pedigrees with phenotypes that are the composite effect of independent variants affecting P2Y12 and PAR1 (17) and P2Y12 and von Willebrand factor (11). We speculate that the apparent over-representation of P2Y12 defects in our case series and in these previous reports, may reflect that clinical diagnostic LTA agonist panels have greater sensitivity for defective stimulatory GPCRs in non-redundant feedback pathways such as P2Y12, than other GPCRs, thereby introducing a selection effect.

Through a systematic analysis of platelet GPCR genes, we have highlighted the burden of rare SNVs in the general population and in selected patients with IPFDs. The variety and burden of potentially damaging SNVs in the healthy population recruited into the 1,000 genome and ESP databases highlight the incomplete penetration of these variants. This suggests the possibility that mild IPFDs are more frequent in the general population, but may commonly go unnoticed until a challenge, such as childbirth, surgery or initiation of antiplatelet therapy, is applied. The spread of bleeding manifestations even in patients diagnosed with IPFDs and harbouring identical SNVs also highlights the challenges of applying genetic screening with platelet function testing approaches to large populations. Rare variants affecting GPCR function may be solely responsible for the phenotypes of some isolated pedigrees with IPFD, but could also contribute to complex defects caused by variants in other platelet genes.

Acknowledgements
This study was funded through the British Heart Foundation Programme RG/09/007 and was supported by the NIHR Bristol Cardiovascular Biomedical Research Unit. JN was funded through a University Hospitals Bristol Clinical PhD studentship. GL was funded through the Wellcome Trust Clinical Training Fellowship 093994. SPW is a British Heart Foundation Chair. We acknowledge the technical assistance of Dr Jessica Harris and Dr Shaista Nisar, University of Bristol.

Author contributions
M.L. Jones contributed to study design, data acquisition, data analysis and co-wrote the manuscript. N.V. Morgan, M. Lordkipanidzé, G. C. Lowe, S. Drake and S. P. Watson recruited and phenotyped cases, coordinated exome sequencing and critically proofread the manuscript. M. Simpson carried out exome sequencing. J.E Norman and S.J Mundell performed the cell transfection and cAMP determination experiments. S.J. Mundell and M.E. Daly contributed to design of the study. A.D. Mumford recruited cases, designed the study, analysed data, and co-wrote the manuscript.

Conflicts of interest
None declared.

References
1. Jones CI, Bray S, Garner SF, et al. A functional genomics approach reveals novel inherited platelet function disorders. Thromb Haemost 2006; 96: 356–360.
2. Johnson AD, Yanek LR, Chen MH, et al. Genome-wide meta-analyses identifies seven loci associated with platelet aggregation in response to agonists. Nature Gen 2010; 42: 608–613.
3. Dupont A, Fontana P, Bachelot-Loza C, et al. An intronic polymorphism in the PAR-1 gene is associated with platelet receptor density and the response to SFLRN. Blood 2003; 101: 1833–1840.
4. Muehlbacher JD, Perry TE, Liu KY, et al. Polymorphism in the protease-activated receptor-4 gene region associates with platelet activation and perioperative myocardial injury. Am J Hematol 2012; 87: 161–166.
5. Fontana P, Gaudreille S, Remones V, et al. Identification of functional polymorphisms of the thromboxane A2 receptor gene in healthy volunteers. Thromb Haemost 2006; 96: 356–360.
6. Zee KY, Michaud SE, Diehl KA, et al. Purinergic receptor P2Y, G-protein coupled, 12 gene variants and risk of incident ischemic stroke, myocardial infarction, and venous thromboembolism. Atherosclerosis 2008; 197: 694–699.
7. Small KM, Brown KM, Seman CA, et al. Complex haplotypes derived from non-coding polymorphisms of the intronless alpha2A-adrenergic gene diversify receptor expression. Proc Natl Acad Sci USA 2006; 103: 5472–5477.
8. Genomes Project C, Abecasis GR, Altshuler D, et al. A map of human genome variation from population-scale sequencing. Nature 2010; 467: 1061–1073.
9. Li Y, Vinckenbosch N, Tian G, et al. Resequencing of 200 human exomes identifies an excess of low-frequency non-synonymous coding variants. Nature Gen 2010; 42: 969–972.
10. Tennesen JA, Bigham AW, O’Connor TD, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 2012; 337: 64–69.
11. Daly ME, Dawood BB, Lester WA, et al. Identification and characterization of a novel P2Y12 variant in a patient diagnosed with type 1 von Willebrand disease in the European MCMDDM-1VWD study. Blood 2009; 113: 4110–4113.
12. Fontana G, Ware J, Cattaneo M. Haploinsufficiency of the platelet P2Y12 gene in a family with congenital bleeding diathesis. Haematologica 2009; 94: 581–584.
13. Mumford AD, Dawood BB, Daly ME, et al. A novel thromboxane A2 receptor D304N variant that abrogates ligand binding in a patient with a bleeding diathesis. Blood 2010; 115: 363–369.
14. Mumford AD, Nisar S, Darnige L, et al. Platelet dysfunction associated with the novel Thr29Cys thromboxane A(2) receptor variant. J Thromb Haemost 2013; 11: 547–554.
15. Nisar S, Daly ME, Federici AB, et al. An intact PDZ motif is essential for correct P2Y12 purinoceptor traffic in human platelets. Blood 2011; 118: 5641–5651.
16. Nisar SP, Lordkipanidze M, Jones ML, et al. A novel thromboxane A2 receptor N42S variant results in reduced surface expression and platelet dysfunction. Thromb Haemost 2014; 111: 923–932.
17. Patel YM, Lordkipanidze M, Lowe GC, et al. A novel mutation in the p2y receptor and a function-reducing polymorphism in par-1 in a patient with chronic bleeding. J Thromb Haemost 2014; Epub ahead of print.
18. Rowley JW, Oler AJ, Tolley ND, et al. Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. Blood 2011; 118: e101–111.
19. Dawood BB, Lowe GC, Lordkipanidze M, et al. Evaluation of participants with suspected heritable platelet function disorders including recommendation and validation of a streamlined agonist panel. Blood 2012; 120: 5041–5049.
20. Dawood BB, Wilde J, Watson SP. Reference curves for aggregation and ATP secretion to aid diagnose of platelet-based bleeding disorders: effect of inhibition of ADP and thromboxane A(2) pathways. Platelets 2007; 18: 329–345.
21. Jones ML, Murden SL, Bern D, et al. Rapid genetic diagnosis of heritable platelet function disorders with next-generation sequencing: proof-of-principle with Hermansky-Pudlak syndrome. J Thromb Haemost 2012; 10: 306–309.
22. Bend J, Stourac J, Saldana O, et al. PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. PLoS Comp Biol 2014; 10: e1003440.
23. Colombo M, De Vecchi G, Calea L, et al. Comparative in vitro and in silico analyses of variants in splicing regions of BRCA1 and BRCA2 genes and characterization of novel pathogenic mutations. PloS One 2013; 8: e57173.
24. Ballesteros JA, Weinstein H. Analysis and refinement of criteria for predicting the structure and relative orientations of transmembranal helical domains. Bioinformatics 1992; 62: 107–109.
25. Venkatar Krishnan AJ, Deupi X, Lebon G, et al. Molecular signatures of G-protein-coupled receptors. Nature 2013; 494: 185–194.
26. Zhang C, Srinivasan Y, Arlow DH, et al. High-resolution crystal structure of human protease-activated receptor 1. Nature 2012; 492: 387–392.
27. Xu F, Wu H, Katritch V, et al. Structure of an agonist-bound human A2A adenosine receptor. Science 2011; 332: 322–327.
28. Zhang K, Zhang J, Gao ZG, et al. Structure of the human P2Y receptor in complex with an antithrombotic drug. Nature 2014; Epub ahead of print.
29. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nature Rev Gen 2011; 12: 683–691.
30. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860–921.
31. Fryxell KJ, Moon WJ. CpG mutation rates in the human genome are highly dependent on local GC content. Mol Biol Evol 2005; 22: 650–658.
32. Kozlowski P, de Mezer M, Krzyzosiak WJ. Trinucleotide repeats in human genome and exome. Nucl Acid Res 2010; 38: 4027–4039.
33. Marchese A, Paing MM, Temple BR, et al. G protein-coupled receptor sorting to endosomes and lysosomes. Annu Rev Pharmacol Toxicol 2008; 48: 601–629.
34. Remijn JJ, MJ, Struk AL, et al. Novel molecular defect in the platelet ADP receptor P2Y12 of a patient with haemorrhagic diathesis. Clin Chem Lab Med 2007; 45: 187–189.
35. Salon JA, Lodziowski DT, Palczewski K. The significance of G protein-coupled receptor crystallography for drug discovery. Pharmacol Rev 2011; 63: 901–937.
36. Cattaneo M, Zighetti ML, Lombardi R, et al. Molecular bases of defective signal transduction in the platelet P2Y12 receptor of a patient with congenital bleeding. Proc Natl Acad Sci USA 2003; 100: 1978–1983.
37. Cattaneo M, Lecchi A, Lombardi R, et al. Platelets from a patient heterozygous for the defect of P2Y12C receptors for ADP have a secretion defect despite normal thromboxane A2 production and normal granule stores: further evidence that some cases of platelet ‘primary secretion defect’ are heterozygous for a defect of P2Y12C receptors. Arterioscl Thromb Vasc Biol 2000; 20: E101–106.