SUPPLEMENTARY INFORMATION

Method S1	**BeviMed genetic association analysis**	**Page 1**
Method S2	**Variant confirmation in cases and pedigree members**	**Page 2**
Method S3	**Expression of *THPO* variants in HEK293T cells**	**Page 3**
Method S4	**Confocal immunofluorescence microscopy**	**Page 4**
Results S1	**Hematological parameters and serum TPO levels for cases with *THPO* variants**	**Page 5**
Reference		**Page 5**

Method S1: BeviMed genetic association analysis

Using the BeviMed method for genetic association[1], we compared the genotypes at rare non-synonymous variants of 105 unrelated thrombocytopenia cases with those of 10,152 unrelated NIHR BioResource for Rare diseases (NBR-RD) participants who did not have unexplained thrombocytopenia. Thrombocytopenia cases were defined as patients with a platelet count (PLT) <130 $\times 10^9$/L or with the HPO term thrombocytopenia (HP:0001873).

We excluded SNVs and indels with an allele count in any GnomAD population higher than would be expected given a true population minor allele frequency of 1/10,000 (under a dominant model) or 1/1,000 (under a recessive model). We further excluded SNVs with a CADD Phred score of less than 10. Large deletions with an internal allele count in excess of what would be expected given a population allele frequency of 1/200 were also excluded. The BeviMed method compares the statistical support

for a baseline model in which disease risk is independent of the genotypes with various association models in which disease risk depends on the allele configuration at the given rare variant sites, a latent partition of variants into pathogenic and benign groups, and a Mendelian mode of inheritance. The association models are fitted to different subsets of variants at a given locus corresponding to different predicted consequences, which imposes a prior correlation structure on the pathogenicity of the variants that reflects competing potential disease mechanisms.

Method S2: Variant confirmation in cases and pedigree members

THPO variants were confirmed in index cases and tested in pedigree members: Genomic DNA was extracted from peripheral blood or saliva using a Maxwell® 16 Instrument as per the manufacturer's instructions. The relevant sections of genomic DNA were then amplified by polymerase chain reaction (PCR) using MyTaq HS Mix (BIOLINE BIO- 25045) and oligonucleotide primers designed with Primer3Plus (https://primer3plus.com/cgi- bin/dev/primer3plus.cgi) and used at a final concentration of 0.5μM in each 20μl PCR reaction.

Primer Sequences

Pedigree A	Forward: CAGGGAGCCTGTGTCTGATG (20) Reverse: AGATCTGGCCCTGGTGTTTG (20) Annealing temperature 60°C
Pedigree B	Forward: AGTCCTCACACTGAACGAGC (20) Reverse: GGAGCAGAAGGGTCAGGAAG (20) Annealing temperature 66°C
Pedigree C	Forward AGACCAGGCTGGACTAGAGG (20) Reverse: GACAGGATGCCAGTACCCAG (20) Annealing temperature 62°C

Thermal cycling was performed with a GeneAmp PCR System 9700 (Applied Biosystems, ThermoFisher) with an initial denaturation step of 3 minutes at 95°C, then 15 cycles of 95°C for 30sec, 60-66°C 1 min touchdown and 72°C for 30sec. The final cycling phase was 25 cycles of 95°C for 30 sec, 55°C for 1 min and 72°C for 30 sec, with a final extension of 72°C for 5 min. The PCR products were checked using an Agilent D1000 ScreenTape System (Agilent Technologies) as per the manufacturer's instructions. Purified, quantified amplicons were sequenced using Sanger methodology (DBS Genomics, Durham University) with the same primers that were used to amplify the target.

Method S3: Expression of THPO variants in HEK293T cells

HEK293T cells (kind gift from Dr Graciela Newby, University of Bristol) were cultured in Dulbecco's Modified Eagle medium supplemented with 10% fetal-bovine serum and 1% Penicillin-Streptomycin. Aliquots of approximately $4x10^5$ cells were seeded in six-well culture plates (in antibiotic-free medium) and transfected 24hours later with pcDNA3.1+ expression vectors containing N-terminal flag-tagged THPO c.DNAs for either wildtype, p.E204Gfs*123, p.L269Pfs*58 or p.R99W variants (purchased from Genscript) or a pcDNA3.1+ negative control. Lipofectamine 2000 (ThermoFisher) was used for transfections at a 3:1 (reagent : DNA) ratio. Medium was changed 6 hours post-transfection. THPO-conditioned supernatants were collected 24hours post-transfection and frozen at –20°C until use. Cells were lysed in 300μl RIPA Buffer containing protease inhibitor (cOmplete™, Mini, EDTA-free protease inhibitor cocktail, Roche).

THPO-conditioned cell supernatants and lysates were analysed using an enzyme-linked immunosorbent assay for the quantification of DYKDDDK (FLAG)-tagged proteins (Cayman Chemical). Results were normalised against the total protein from each cell lysate (determined using a Pierce™ BCA Protein Assay Kit, ThermoFisher). In order to confirm that FLAG-tagged protein levels accurately represented thrombopoietin expression, the thrombopoeitin concentration was measured using a thrombopoietin ELISA (R&D Biosystems) in parallel with the FLAG-ELISA in supernatants of each transfected cell line. There was strong correlation between FLAG-tagged protein and thrombopoietin concentrations.

Methods S4: Confocal immunofluorescence microscopy

For confocal immunofluorescence microscopy, HEK293T cells were transfected as above, then applied to Poly-L-lysine coated coverslips. Cells were fixed with 4% paraformaldehyde for 15mins, permeabilised with 0.2% Triton X-100 for 10mins and blocked with 5% Bovine Serum Albumin (BSA) for 30mins at room temperature. Cells were incubated for 2 hours at room temperature with a primary antibody against the FLAG epitope (Caymen Chemical) at a dilution of 1:200 in 1% BSA/0.1% PBS-Tween, and then incubated for 1 hour at room temperature with an AlexaFluor® plus 647 Goat anti-Mouse secondary antibody (ThermoFisher) at a dilution of 1:500 in 1%BSA/0.1% PBS Tween. Coverslips were mounted onto glass slides using Vectashield® mounting medium with DAPI (Vector Laboratories). Cells were visualised using a Leica SP5-II Confocal microscope with a 63x/1.4NA lens and analysed with ImageJ/Fiji

Results S1: Hematological parameters and serum thrombopoietin (TPO) levels for case with *THPO* variants.

Case	**A I.1**	**A II.2**	**A III.1**	**B I.2**	**BII.1**	**C II.1**
Age (years) / Gender	76 M	42 F	<1 F	70 F	33 F	39 F
Platelet count x 10^9/L	102	121	139	120	109	143
MPV, fl	14.2	12.2	n/a	12.5	n/a	12.1
PDW, fl	20.3	16.2	n/a	17.5	n/a	15.3
IPF, %	17.3	6.7	n/a	8.2	n/a	6.4
Haemoglobin, g/l	142	152	204	139	137	139
MCV, fl	86.9	88.4	104.3	95.1	94.8	90.6
Reticulocytes x 10^9/L	36.5	44.8	183	68.3	n/a	31.2
WBC count x 10^9/L	7.25	8.51	10.0	5.8	5.2	4.6
Neutrophil count x 10^9/L	4.44	4.74	5.5	3.6	3.5	2.6
Serum thrombopoietin (pg/mL)	44.7	17.0	n/a	56.9	n/a	47.9

IPF, immature platelet fraction; MCV, mean corpuscular volume; MPV, mean platelet volume; PDW, platelet distribution width; WBC, white blood cell

REFERENCES

1. Greene, D., BioResource, N., Richardson, S. & Turro, E. A Fast Association Test for Identifying Pathogenic Variants Involved in Rare Diseases. *Am J Hum Genet* **101**, 104-114 (2017).