Supplementary file

A Synthesis Approach of Mouse Studies to Identify Genes and Proteins in Arterial Thrombosis and Bleeding

Constance C. F. M. J. Baaten*, Stuart Meacham*, Susanne M. de Witt*, Marion A. H. Feijge, David J. Adams, Jan-Willem N. Akkerman, Judith M. E. M. Cosemans, Luigi Grassi, Steve Jupe, Myrto Kostadima, Nadine J. A. Mattheij, Martin H. Prins, Ramiro Ramirez-Solis, Oliver Soehnlein, Frauke Swieringa, Christian Weber, Jacqueline K. White, Willem H. Ouwehand*, Johan W. M. Heemskerk*

Supplementary methods

Selection of papers and study definition

A PubMed search over the years 1980-2018 was performed to identify papers on mouse studies and arterial thrombus formation. Search keywords were mouse, arterial, thrombus and thromboembolism in different combinations (Suppl. Table 1). We searched for primary papers: *(i)* in English language; *(ii)* describing effects of modification of a specific mouse gene or protein on arterial thrombus formation or thromboembolism; and *(iii)* providing quantitative data with statistics. Excluded were papers: *(i)* not indicating a well-defined genetic or pharmacological perturbation; *(ii)* using venous thrombosis models only; *(iii)* or containing no quantitative data or statistics. A second PubMed search was performed to identify papers describing quantitative effects of genetic modification on tail bleeding time. Keywords were: mouse, (tail) bleeding, and tail clipping in different combinations. Non-tail bleeding times were not included. Papers of the described sets as well as published reviews were further

screened for additional primary papers. For numbers of papers identified per search term combination, see Suppl. Table 1.

From the included papers, individual studies were defined as measurements in a defined mouse cohort of thrombus formation or thromboembolism *in vivo* or *in vitro,* using a particular experimental method (vascular bed, injury trigger) and a particular animal perturbation (modified gene or target protein inhibited) in comparison to the unperturbed condition (corresponding wildtype or pharmacological control), such as defined by the authors. Papers describing multiple models to measure thrombus formation*, e.g.,* in different vessels or *in vivo* and *in vitro*, hence, provided more than one study.

Registry of perturbed genes, experimental variables and outcome data per study

In Data File 1, all studies including references were categorized per target gene or corresponding protein, according to Mouse Genome Informatics (MGI). Basal genetic information was taken from open access databases (www.uniprot.org and www.informatics.jax.org): *i.e.* gene symbols and names; mouse genome index; chromosome location; Uniprot identification number. Per study are listed, according to the original papers: the genetic background; type of genetic modification; pharmacological intervention if applicable; mouse phenotype regarding altered activity of platelets, coagulation, fibrinolysis or vessel wall.

As indicated in Data File 1, the collected experimental details on the prothrombotic intervention differ per study class. Registered for class I (arterial thrombosis *in vivo*) are: vessel and type of injury, and dosing plus duration of the injury.

Recorded for class II (systemic thromboembolism *in vivo*) were: type and dosing of the intervention. Recorded for class III (thrombus formation *in vitro*) are: thrombogenic surface, and the blood flow conditions. With respect to tail bleeding (class IV), only times of (continued) bleeding are listed.

Quantitative information regarding the thrombotic or thromboembolic process is recorded at both the perturbed (genetic modification or pharmacological intervention) and unperturbed (wildtype or control) conditions (Data File 1). Original papers were mined for the following data (in units, numbers and means/medians, as published by the authors): For **Class I**: blood flow (BF); time to thrombus formation (TTF), time to occlusion (TO), percentage of occluded vessels (%Oc), platelet adhesion (PA), thrombus volume or size (ThS), number of thrombi (#Th), number of emboli (#Emb). For **Class II:** percentage of deaths (%Dth), time to death (TDth), platelet count (PlC). For **Class III:** time to thrombus formation (TTF), percentage surface area coverage (SAC); platelet adhesion (PA), procoagulant activity (ProA), thrombus size (ThS), number of thrombi or platelet aggregates (#Th), number of emboli (#Emb). For **Class IV**: per default tail bleeding time (TBT); if indicated re-bleeding events (RB).

Also listed are the experimental variables, such as type of measurement, detection method, and analysis procedure. Differences in mean or median values between mice with(out) perturbation were considered to be statistically significant, if indicated so by the authors.

Meta-analysis

Conventional meta-analysis was performed for six of the best studied genes*, i.e., Cd36*, *Clec1b, F12*, *Fcer1g*, *Gp6* and *Vwf* (Data File 2). For meta-analysis those studies were included in which: *(i)* arterial thrombosis was studied in the carotid, mesenteric or femoral vascular beds triggered by $FeCl_3$; *(ii)* outcomes were presented as time to vascular occlusion or time to (full) thrombus formation; *(iii)* data were presented as means with SE or SD (if *n* was given as a range, lowest values were taken); *(iv)* equal variance of groups was assumed with a correction based on effect size, if the SD of one group was missing. The meta-analyses were applied assuming a random effects model, based on the DerSimonian and Laird's methodology, according to the Nordic Cochrane Group Review Manager 5.1.[1,2] Calculated per gene were: standardized mean difference plus 95% confidence interval, and heterogeneity index I^2. Full meta-analysis data are given in Data File 2.

Combining of studies and outcome parameters for synthesis approach

Studies of class I (arterial thrombosis *in vivo*) were grouped according to similarity in the protocol of thrombosis induction (Suppl. Table 2). For class I, groups 1-12 differed in type of vascular bed and trigger of thrombosis induction. The severity of intervention (dosing and duration) was not evaluated separately, since this is often incompletely described.[3] For class II, groups 20-23 were formed based on trigger of (provoked) systemic thromboembolism or on observation of unprovoked thromboembolism. For class III, groups 30-34 differed depending on the thrombogenic surface for *in vitro* blood flow measurement. Studies of class IV (bleeding) were not sub-divided.

Regarding *in vivo* and *in vitro* thrombus formation (classes I and III), data were registered as time-dependent parameters (**a**), mass-dependent parameters (**b***i*: extent of platelet adhesion; **b***ii:* extent of occlusion or surface-area-coverage; **b***iii:* number, size or volume of thrombi), or stability parameters (**c**, number or frequency of embolic events). Regarding *in vivo* thromboembolism (class II), data were categorized as provoked (**d**, percentage of survival, time to death, residual platelet count) or unprovoked (**e**, spontaneous thrombosis or bleeding).

Scaling comparison of thrombosis studies in synthesis approach

Per study and parameter, original papers were mined for the published quantitative data on (sub)parameters of arterial thrombus formation, thromboembolism or thrombus formation *in vitro,* which were in accordance with the authors' conclusions. The reported effects of the perturbation (genetic modification or pharmacological intervention) were linearly normalized on both a 3-point and 5-point scale. For the 3-point scale, reported values indicating a *significant reduction or delay* in thrombus formation or thromboembolism after perturbation were scaled as -1; no effects as 0; and values indicating a *significant increase or acceleration* of thrombus formation as +1. For the 5-point scale, a further distinction was made between significantly reduced/delayed thrombus formation of < 50% (scale -2) or ≥ 50% (scale -1); and between an increased/accelerated thrombus formation of ≤ 200% (scale +1) or > 200% (scale +2). Perturbation effects were considered to be statistically significant, if indicated so by the authors. For convenience, these normalized, scaled values are indicated below as 'scores'. A complete overview of the mined data and the processing of the data are

provided in Data File 1. In addition, the legend part of this file gives examples of the way of data processing towards scores.

To assess effects on hemostasis, tail bleeding times were also normalized on 3-point and 5-point scale: For the 3-point scale, bleeding was scored as prolonged (-1), unchanged (0), or shortened (+1). For the 5-point scale, bleeding scores were further subdivided into significantly prolonged bleeding times of < 50% (scale -2) or ≥ 50% (scale -1); and between significantly shortened bleeding times of ≤ 200% (scale +1) or > 200% (scale +2). Again, effects were considered to be statistically significant, if indicated so by the authors. Given that the majority of mouse studies were carried out at similar, small sample size (*n* = 6-12 animals), no weighting was performed when averaging study scores.

Mean scores per study and gene were rounded off on both 3-point and 5-point scales. Data File 1 provides all scores per study and gene, as well as the rounded off values both on 3-point and 5-point scales. As an overview, the mean scores per gene for modification in arterial thrombus formation (classes I-III) and tail bleeding (class IV) are listed in Data File 3, which also includes a shortlist of reported functions of the corresponding proteins in platelet, coagulant or vascular activity, such as described in the original papers.

Reactome pathway analysis

Reactome (www.reactome.org/) is an open source database for the exploration and analysis of (human) biological pathways, based on curated literature analysis. It comprises a listing of protein-protein interactions based on biochemical reactions that

are mediated by specific proteins or classes of proteins (*e.g.*, protein tyrosine kinases or GTP-binding proteins). Using the Analyze Data tool,[4] Reactome also clusters cascades of protein-protein interactions into biological pathways. In addition, it performs overrepresentation analysis of protein lists to determine the coverage of matches per pathway; and allows the inclusion of attributes (*i.c.* scores) to produce a mean value of the attribute per pathway (*i.c.* mean scores).

Reactome analysis was performed for the following datasets: *(i)* all 409 scored mouse genes; *(ii)* all 334 genes or *(iii)* the 278 genes with modifying scores for thrombus formation (class I-III); *(iv)* all 311 genes or *(v)* the 192 genes with modifying scores for bleeding (class IV). Scored genes were submitted as Uniprot numbers for the orthologues human proteins.

For each of the five datasets, numbers of entities in Reactome were listed per biological pathway, as well as the mean scores of the involved proteins per pathway. Pathway lists were restricted to a minimal number of 4 entities for dataset *(i)*. For this dataset, also *P* values were given regarding the overrepresentation of the Uniprot lists in the concerning pathways (Data File 5).

Pathway analysis for the complete dataset of 409 scored genes/proteins indicated that a substantial part of the Reactome-registered hemostatic system was covered (148 out of 806 proteins). Well-covered pathways included *'Platelet activation, signaling and aggregation'* (93/288), *'Neutrophil degranulation'* (44/480), *'Platelet degranulation'* (32/137), *'Formation of fibrin clot'* (28/52), and various cell-signaling pathways, all of which with high significance of over-representation. Reactome analysis also revealed over-representation in several unexpected pathways, linked to cancer

(*e.g., 'Oncogenic MAPK signaling'* 18/87; *'Signaling by BRAF and RAF fusions'* 17/68) or to the vessel wall (*e.g., 'ECM proteoglycans'* 9/79, *'Nitric oxide stimulates guanylate cyclase'* 7/36).

The Reactome Analysis tool also allows calculation of the mean scores per pathway of identified proteins/genes. This analysis revealed several pathways, where the mean scores for (modifying) proteins in class I-III were less negative than for those in class IV (*i.e.*, *'Synthesis of 12/15-eicosatetraenoic acid derivatives', 'Activation of matrix metalloproteinases', 'Common pathway of fibrin clot formation'*), suggesting a more prominent role in hemostasis. Conversely, mean scores for proteins in Toll-like receptor pathways (*'IRAK4 deficiency TLR2/4'* and *'MyD88 deficiency TLR2/4*') in class I-III were more negative than for those in class IV, suggesting a more prominent role in thrombosis. (Data File 5)

Network of protein interactions in thrombosis and hemostasis (PITH)

For network analysis in Cytoscape, human orthologous genes were identified for 397 murine genes, of which 267 were included as modifying in class I-III. For these genes, corresponding human Uniprot protein identifiers were obtained, and used as primary bait to search for first order interactions using the public databases Reactome (78%), IntAct (22%) (www.ebi.ac.uk.intact/) and knowledge obtained from the literature (0.02%). Networks were constructed in Cytoscape, and network structures were analyzed using the CentiScape tool.[5]

Lists of core nodes (baits) and novel nodes (interactors) of all human orthologues in the network are provided in Data File 6. Indicated are for the core nodes the mean

scores of mouse genes in class I-III (thrombosis) and class IV (bleeding). Indicated are for all nodes the following network parameters: *(i)* Numbers of direct edges; *(ii)* Connectivity ('NeighborhoodConnectivity', *i.e.* steps to reach all other nodes); *(iii)* Proximity ('Radiality', *i.e*. average tendency to proximity to other nodes); *(iv*) Closeness ('ClosenessCentrality', *i.e.* central location to all other nodes); *(v)* Sub-clustering ('ClusteringCoefficient', *i.e*. tendency to organization in discrete clusters).

Indicated in Data File 6 are genes/proteins derived from the following GWAS and human (patho)physiological databases: *(i)* Thrombogenomics platform of genes linked to rare human Mendelian disorders of platelet or coagulation function from the BPD cohort (Bleeding and Platelet Diseases, NHS, UK);[6] *(ii)* GWAS of platelet count and volume [7]; *(iii)* variants of genomic sites associated with quantitative platelet traits;[8] *(iv)* genetic loci of the MEGASTROKE GWAS associated with stroke subtypes,[9] and *(v)* genetic loci associated with coronary artery disease.[10]

Furthermore, for relevant nodes, gene expression levels are indicated from the published Blueprint database,[11] such as recently updated regarding lineages of hematopoietic stem cells, including megakaryocytes (platelet precursor cells) and erythroblasts.[12]

mRNA Expression analysis

mRNA expression analysis of indicated endothelial and blood cells was as described. In brief, Trim Galore (v0.3.7) (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with parameters "-q 15 -s 3 --length 30 -e 0.05" was used to trim PCR and sequencing adapters. Trimmed reads were aligned to the Ensembl v75[13] human

transcriptome with Bowtie 1.0.1[14] using the parameters "-a --best --strata -S -m 100 -X 500 --chunkmbs 256 --nofw --fr". Subsequently, MMSEQ (v1.0.8a)[15,16] was used with default parameters to quantify and normalize gene expression.

Genetically modified mice to confirm constructed network

Animal experiments were approved by the local Animal Experimental Committees in Munich and Hinxton/Cambridge. Five strains of mice came from an unbiased selection of animals available from the Wellcome Sanger Institute Mouse Genetics Programme (Cambridge, United Kingdom), generated using targeted embryonic stem cells from the European Conditional Mouse Mutagenesis (EUCOMM) program or the Knockout Mouse Project (KOMP);[17] *i.e.* mice with homozygous deficiency in *Bnip2*, *Dlg4*, *Grm8*, *Ifnar1* or *Vps13a* with corresponding wildtype controls (tm1a/b allele on a C57BL/6N genetic background).

Three strains came from the breeding programme of the Institute for Cardiovascular Prevention in Munich (Germany) and were included because of established roles in atherosclerosis,[18,19] *i.e.* C57BL/6 mice with homozygous deficiency in *Anxa1*, *Apoe* or *Fpr2* with corresponding wildtypes.[20] Mice from all strains had normal platelet counts; however, platelets from only *Ifnar1*-deficient mice showed a 9.4% reduction in mean volume (P<0.01, n=5). No other relevant information on blood phenotype was available from Mouse Genome Informatics (www.informatics.jax.org).

Flow perfusion tests with 350 µL whole blood samples (collected on thrombin inhibitors) were performed as before.[21] In brief, freshly obtained samples were perfused over microspots coated with collagen type I at a wall shear rate of 1000 s^{-1} during 4 min

through a parallel-plate perfusion chamber. Brightfield microscopic images were taken using an EVOS microscope,[22] and analyzed for surface area coverage of multilayered platelet thrombi using the free program ImageJ.[21,23] All image analysis was blinded for the experimental condition.

Statistics and comparative analyses

Statistics on effects of gene perturbation on thrombus formation in individual studies were taken from the original papers. For all studies, correlation analysis was performed to compare, per study and class, scores of time-dependent, mass-dependent and thrombus stability parameters. A non-parametric Kendall's Tau-c test was used to determine association or rank correlation. Herein, the Kendall's τ value is computed as the excess of tabled concordant over discordant pairs, divided by a term representing the geometric mean of untied pairs. It reaches from 0.0 to 1.0 for tables where the entries are on one diagonal.

Calculation of consistency parameter (CP)

A standard method to assess distribution profiles of interval variables does not exist. We therefore defined a procedure to come to such a parameter. First, for all modifying genes (mean scores -2, -1, +1, +2) with $n \geq 2$, a one-sample t-test was performed for the scores of all individual studies in classes I-III. The obtained *P* value was used to calculate a consistency parameter (CP), defined as 1 – log (*P*), reflecting the extent of cohesion of pooled study scores (with 95% confidence intervals). If all studies scores were equal, the nearest *P* value was assigned. For genes with $n = 1$ studies, CP was

set at 1 per default. For non-modifying genes (mean score of 0, SD=0), the null hypothesis was considered to be true, and a P = 1.00 was assumed. Note that a CP value >2.30 corresponds to a P<0.05, and to 95% CI values not including zero. The statistical package for the social sciences was used for all analyses (SPSS version 23, IBM Statistics).

References

1. The Cochrane Collaboration. The Nordic Cochrane Centre Review Manager. *RevMan.* 2011:5.1.
2. McKenzie JE, Salanti G, Lewis SC, Altman DG. Meta-analysis and the Cochrane collaboration: 20 years of the Cochrane statistical methods group. *Systematic Rev.* 2013;2:80.
3. Denis CV, Dubois C, Brass LF, Heemskerk JW, Lenting PJ. Towards standardization of in vivo thrombosis studies in mice. *J Thromb Haemost.* 2011;9:1641-4.
4. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. *Nucleid Acids Res.* 2014;42:D472-7.
5. Scardoni G, Petterlini M, Laudanna C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics.* 2009;25:2857-9.
6. Thrombogenomics. *https://haemgenhaemcamacuk/thrombogenomics.* 2018.
7. Gieger C, Radhakrisknan A, Cvejic A, Tang W, Porcu E, et_al. New gene functions in megakaryopoiesis and platelet formation. *Nature.* 2011;480:201-8.
8. Petersen R, Lambourne JJ, Javierre BM, Grassi L, Kreuzhuber R, Ruklisa D, et al. Platelet function is modified by common sequence variation in megakaryocyte super enhancer. *Nat Commun.* 2017;8:16058.
9. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* 2018;50:524-37.

10. Van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res.* 2018;122:433-43.
11. Chen L, Kostadima M, Martens JH, Canu G, Garcia SP, Turro E, et al. Transcriptional diversity during lineage commitment of human blood progenitors. *Science.* 2014;345:125-33.
12. Blueprint. *blueprinthaemcamacuk/bloodatlas/.* 2018.
13. Cunninham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al. Ensembl 2015. *Nucleic Acids Res.* 2015;43:D662-9.
14. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the humane genome. *Genome Biol.* 2009;10:R25.
15. Turro E, Astle WJ, Tavaré S. Flexible analysis of RNA-seq data using miced effects models. *Bioinformatics.* 2014;30:180-8.
16. Turro E, Su SY, Gonçalves A, Coin LJ, Richardson S, Lewin A. Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. *Genome Biol.* 2011;12:R13.
17. Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, et al. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature.* 2011;474:337-42.
18. Koenen RR, von Hundelshausen P, Nesmelova IV, Zernecke A, Liehn EA, Sarabi A, et al. Disrupting functional interactions between platelet chemokines inhibits atherosclerosis in hyperlipidemic mice. *Nat Med.* 2009;15:97-103.
19. Weber C, Soehnlein O. ApoE controls the interface linking lipids and inflammation in atherosclerosis. *J Clin Invest.* 2011;121:3825-7.
20. Drechsler M, de Jong R, Rossaint J, Viola JR, Leoni G, Wang JM, et al. Annexin A1 counteracts chemokine-induced arterial myeloid cell recruitment. *Circ Res* 2015;116:827-35.
21. Kuijpers MJ, Mattheij NJ, Cipolla L, van Geffen JP, Lawrence T, Donners MM, et al. Platelet CD40L modulates thrombus growth via phosphatidylinositol 3-kinase beta,

and not via CD40 and IkB kinase alpha. *Arterioscler Thromb Vasc Biol.* 2015;35:1374-81.

22. De Witt SM, Swieringa F, Cavill R, Lamers MM, van Kruchten R, Mastenbroek T, et al. Identification of platelet function defects by multi-parameter assessment of thrombus formation. *Nat Commun.* 2014;5:4257.
23. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open source platform for biological-image analysis. *Nat Methods.* 2012;9:676-82.

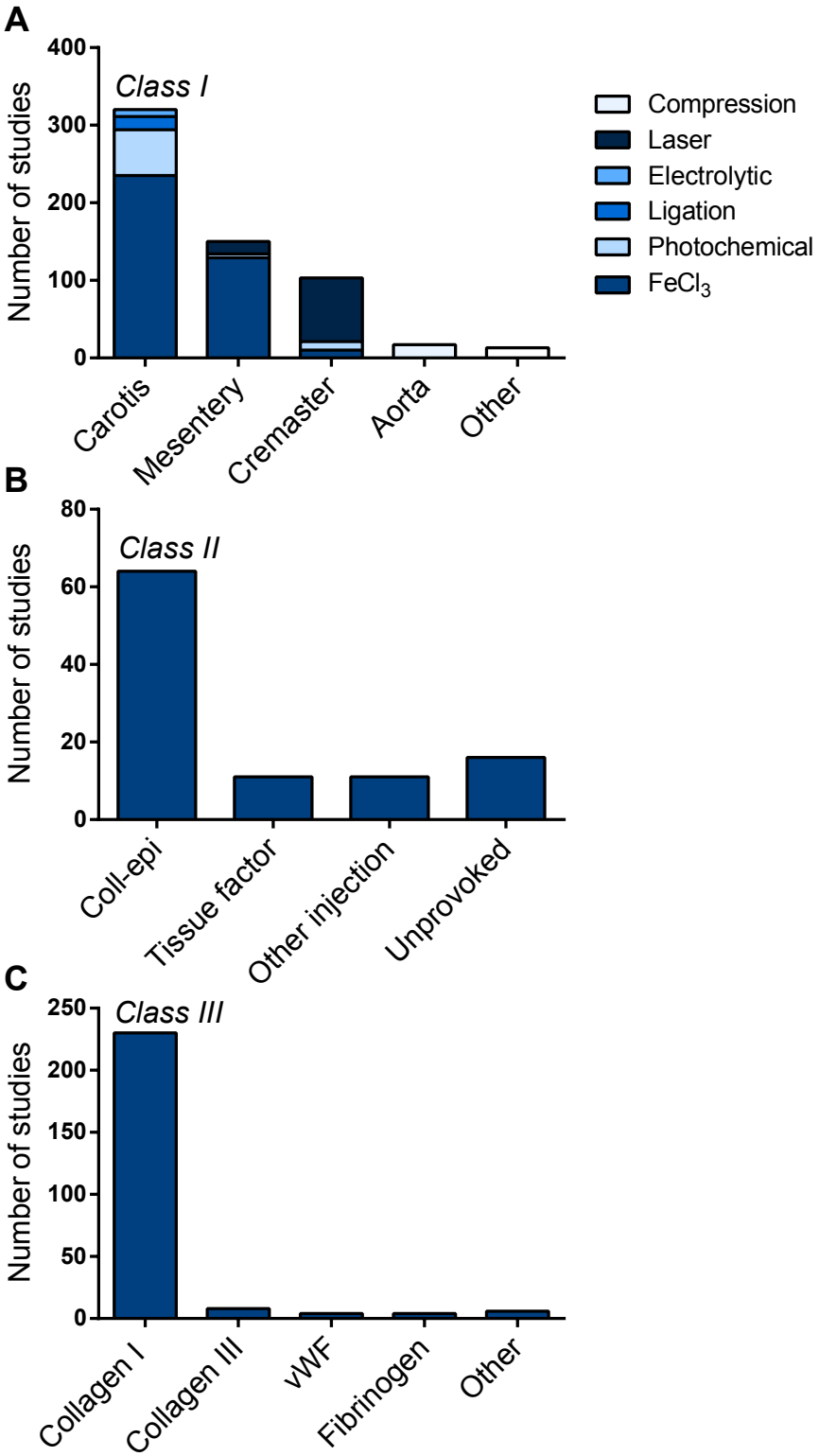
List of Supplementary Figures, Tables and Data Files

Suppl. Figure 1: Numbers of studies in class I-III with main experimental variables.
Suppl. Figure 2: Overview per mouse gene and study of scores in class I, II, III and IV.
Suppl. Figure 3: Network of protein interactions in thrombosis and hemostasis (PITH).
Suppl. Figure 4: Buildup of protein-protein interaction network.
Suppl. Figure 5: Relative expression levels of core and novel nodes in blood cells

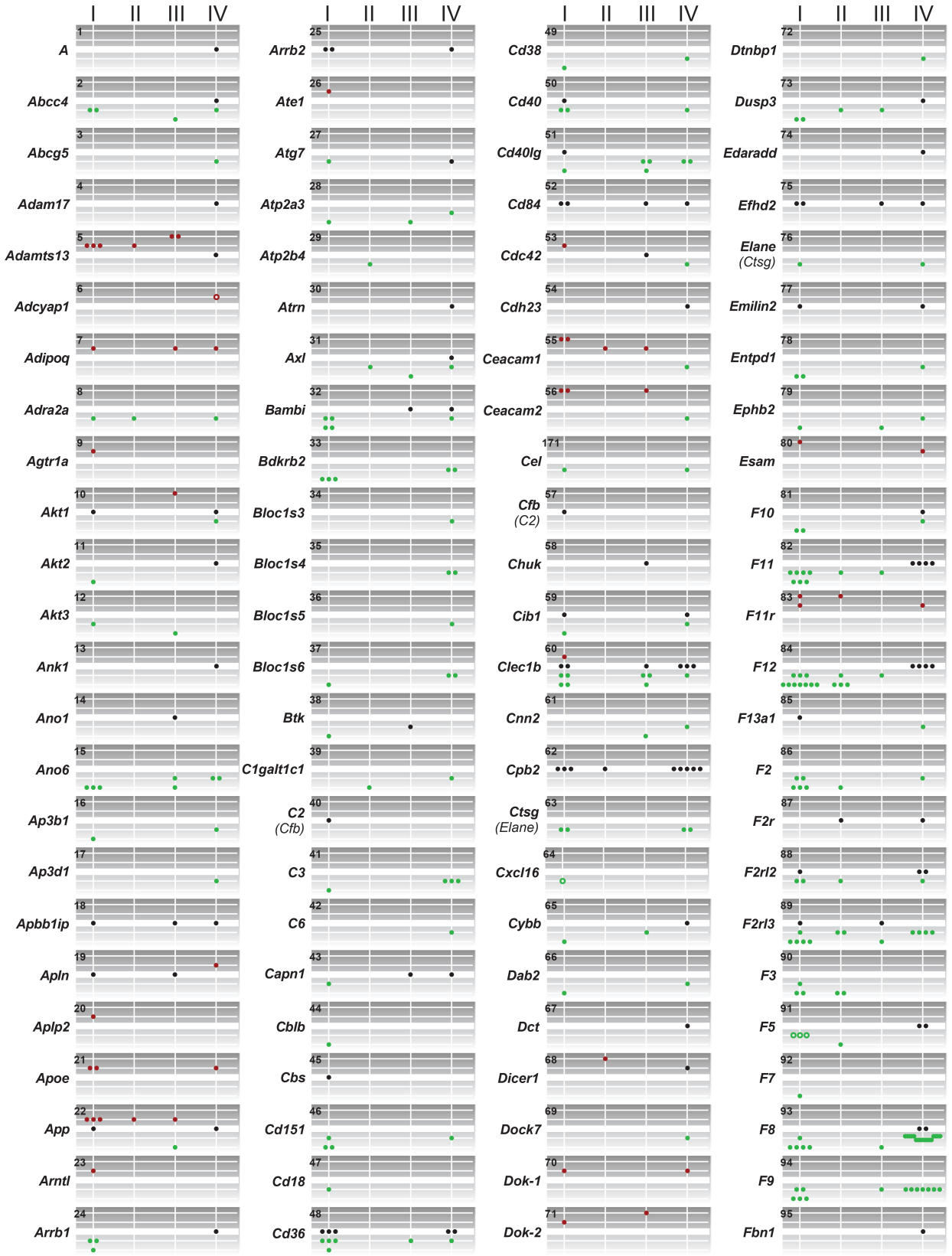
Suppl. Table 1: Keywords and numbers from search outcomes in PubMed.
Suppl. Table 2: Classification and grouping of analyzed mouse studies.
Suppl. Table 3: Correlations of intra-study parameters scored on 3- or 5-point scale.
Suppl. Table 4: Correlations of mean study scores per class on 3- or 5-point scale.
Suppl. Table 5: Numbers of scored mouse genes per class and group.
Suppl. Table 6: Node connections in PITH network for most connected proteins.
Suppl. Table 7: Mouse/human homology of relevant orthologous genes and proteins.

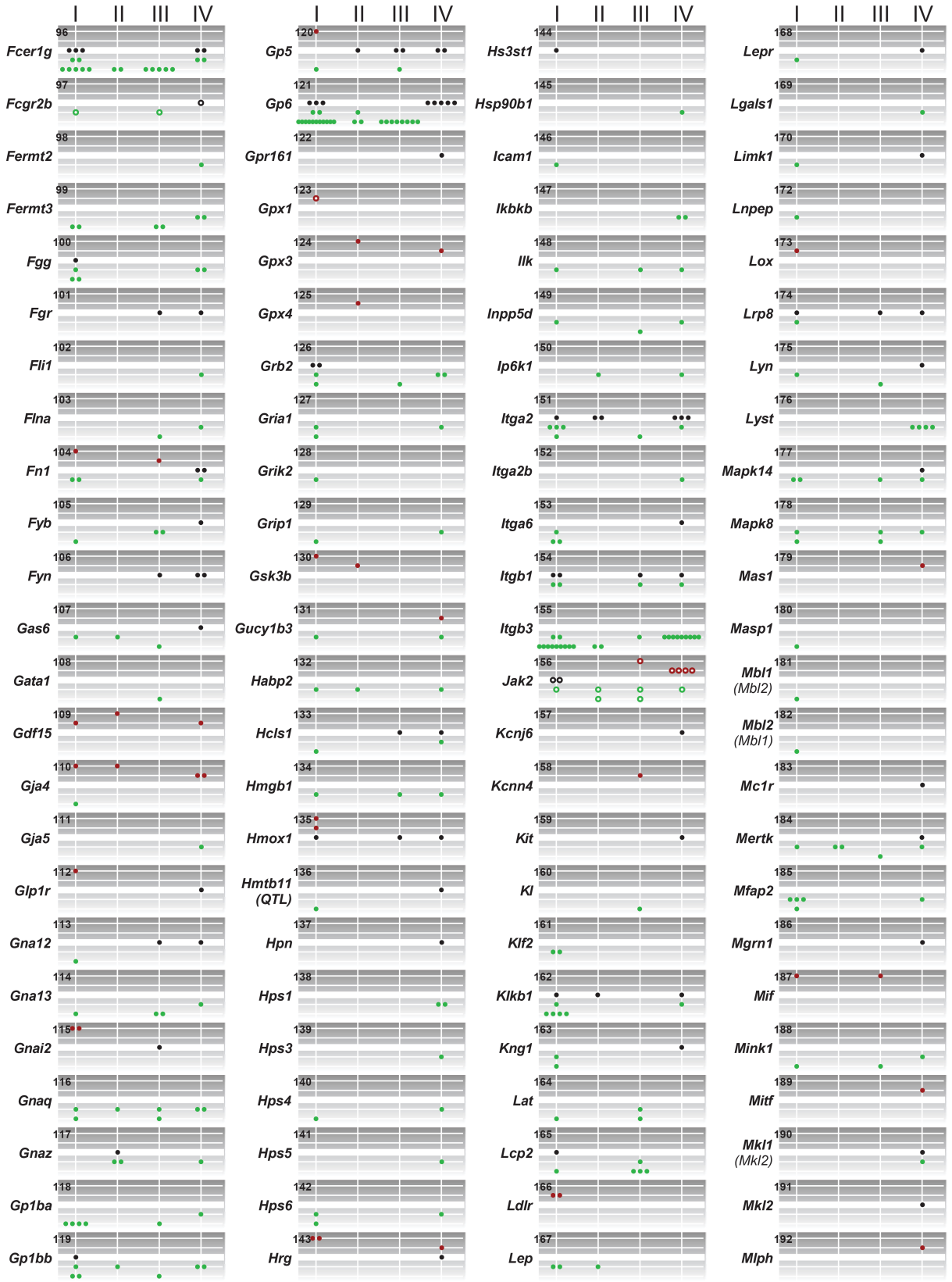
Supplementary Data File

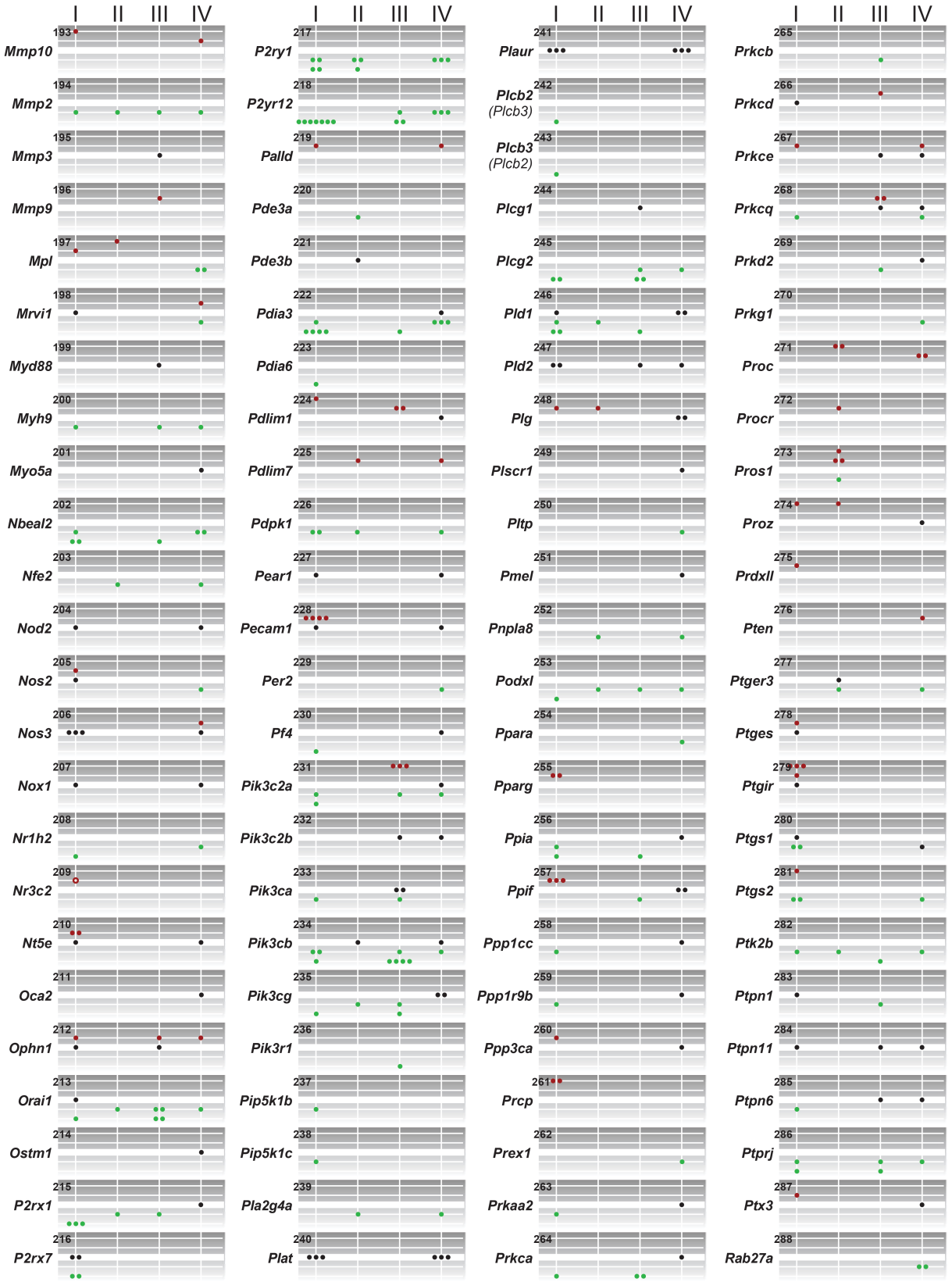
Legend. Includes examples of scoring/scaling procedure.
Data File 1. Primary dataset of included mouse studies, genetic modification, experimental variables, mined measurement outcomes, (sub)scores and references.
Data File 2. Input data of formal meta-analysis using RevMan 5.1 for six most studied genes.
Summary Data File 3. Mean scores and consistency parameter (CP) of alterations in arterial thrombus formation (classes I-III) and tail bleeding (class IV) calculated per gene.
Data File 4. Mouse/human homology, mRNA and protein expression levels. Listings for all genes included in the synthesis.
Data File 5. Reactome pathway analysis of over-representation for scored genes/ proteins with(out) modifying effect in class I-III (thrombosis) or class IV (hemostasis).
Data File 6. Core and novel proteins (nodes) identified in protein interaction network of thrombosis and hemostasis (PITH).



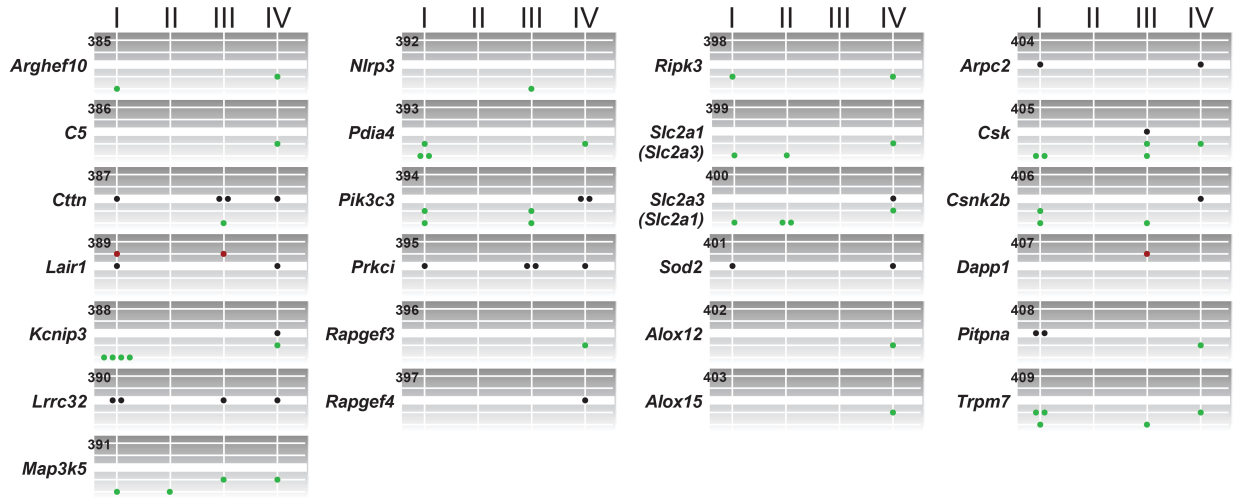
Suppl. Figure 1. **Numbers of studies analyzed in classes I-III with main experimental variables. A**, Class I studies grouped with separation according to vascular bed or type of injury. **B**, Class II studies grouped with separation according to type of systemic intervention. **C**, Class III studies with separation for type of thrombogenic surface.


I II III IV
1 A
2 Abcc4
3 Abcg5
4 Adam17
5 Adamts13
6 Adcyap1
7 Adipoq
8 Adra2a
9 Agtr1a
10 Akt1
11 Akt2
12 Akt3
13 Ank1
14 Ano1
15 Ano6
16 Ap3b1
17 Ap3d1
18 Apbb1ip
19 Apln
20 Aplp2
21 Apoe
22 App
23 Arntl
24 Arrb1
25 Arrb2
26 Ate1
27 Atg7
28 Atp2a3
29 Atp2b4
30 Atrn
31 Axl
32 Bambi
33 Bdkrb2
34 Bloc1s3
35 Bloc1s4
36 Bloc1s5
37 Bloc1s6
38 Btk
39 C1galt1c1
40 C2 (Cfb)
41 C3
42 C6
43 Capn1
44 Cblb
45 Cbs
46 Cd151
47 Cd18
48 Cd36
49 Cd38
50 Cd40
51 Cd40lg
52 Cd84
53 Cdc42
54 Cdh23
55 Ceacam1
56 Ceacam2
171 Cel
57 Cfb (C2)
58 Chuk
59 Cib1
60 Clec1b
61 Cnn2
62 Cpb2
63 Ctsg (Elane)
64 Cxcl16
65 Cybb
66 Dab2
67 Dct
68 Dicer1
69 Dock7
70 Dok-1
71 Dok-2
72 Dtnbp1
73 Dusp3
74 Edaradd
75 Efhd2
76 Elane (Ctsg)
77 Emilin2
78 Entpd1
79 Ephb2
80 Esam
81 F10
82 F11
83 F11r
84 F12
85 F13a1
86 F2
87 F2r
88 F2rl2
89 F2rl3
90 F3
91 F5
92 F7
93 F8
94 F9
95 Fbn1

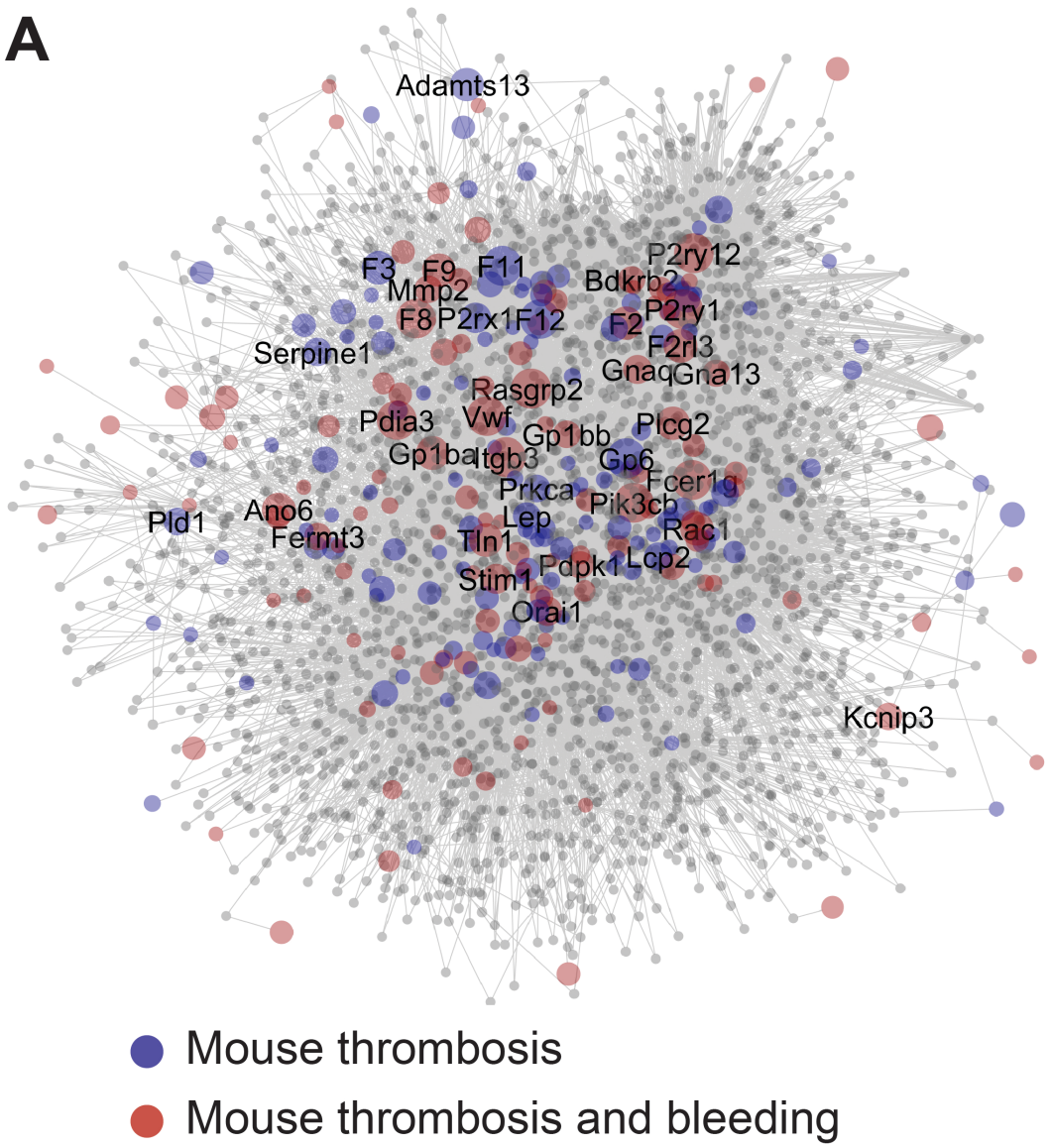

I II III IV
96 Fcer1g
97 Fcgr2b
98 Fermt2
99 Fermt3
100 Fgg
101 Fgr
102 Fli1
103 Flna
104 Fn1
105 Fyb
106 Fyn
107 Gas6
108 Gata1
109 Gdf15
110 Gja4
111 Gja5
112 Glp1r
113 Gna12
114 Gna13
115 Gnai2
116 Gnaq
117 Gnaz
118 Gp1ba
119 Gp1bb
120 Gp5
121 Gp6
122 Gpr161
123 Gpx1
124 Gpx3
125 Gpx4
126 Grb2
127 Gria1
128 Grik2
129 Grip1
130 Gsk3b
131 Gucy1b3
132 Habp2
133 Hcls1
134 Hmgb1
135 Hmox1
136 Hmtb11 (QTL)
137 Hpn
138 Hps1
139 Hps3
140 Hps4
141 Hps5
142 Hps6
143 Hrg
144 Hs3st1
145 Hsp90b1
146 Icam1
147 Ikbkb
148 Ilk
149 Inpp5d
150 Ip6k1
151 Itga2
152 Itga2b
153 Itga6
154 Itgb1
155 Itgb3
156 Jak2
157 Kcnj6
158 Kcnn4
159 Kit
160 Kl
161 Klf2
162 Klkb1
163 Kng1
164 Lat
165 Lcp2
166 Ldlr
167 Lep
168 Lepr
169 Lgals1
170 Limk1
172 Lnpep
173 Lox
174 Lrp8
175 Lyn
176 Lyst
177 Mapk14
178 Mapk8
179 Mas1
180 Masp1
181 Mbl1 (Mbl2)
182 Mbl2 (Mbl1)
183 Mc1r
184 Mertk
185 Mfap2
186 Mgrn1
187 Mif
188 Mink1
189 Mitf
190 Mkl1 (Mkl2)
191 Mkl2
192 Mlph

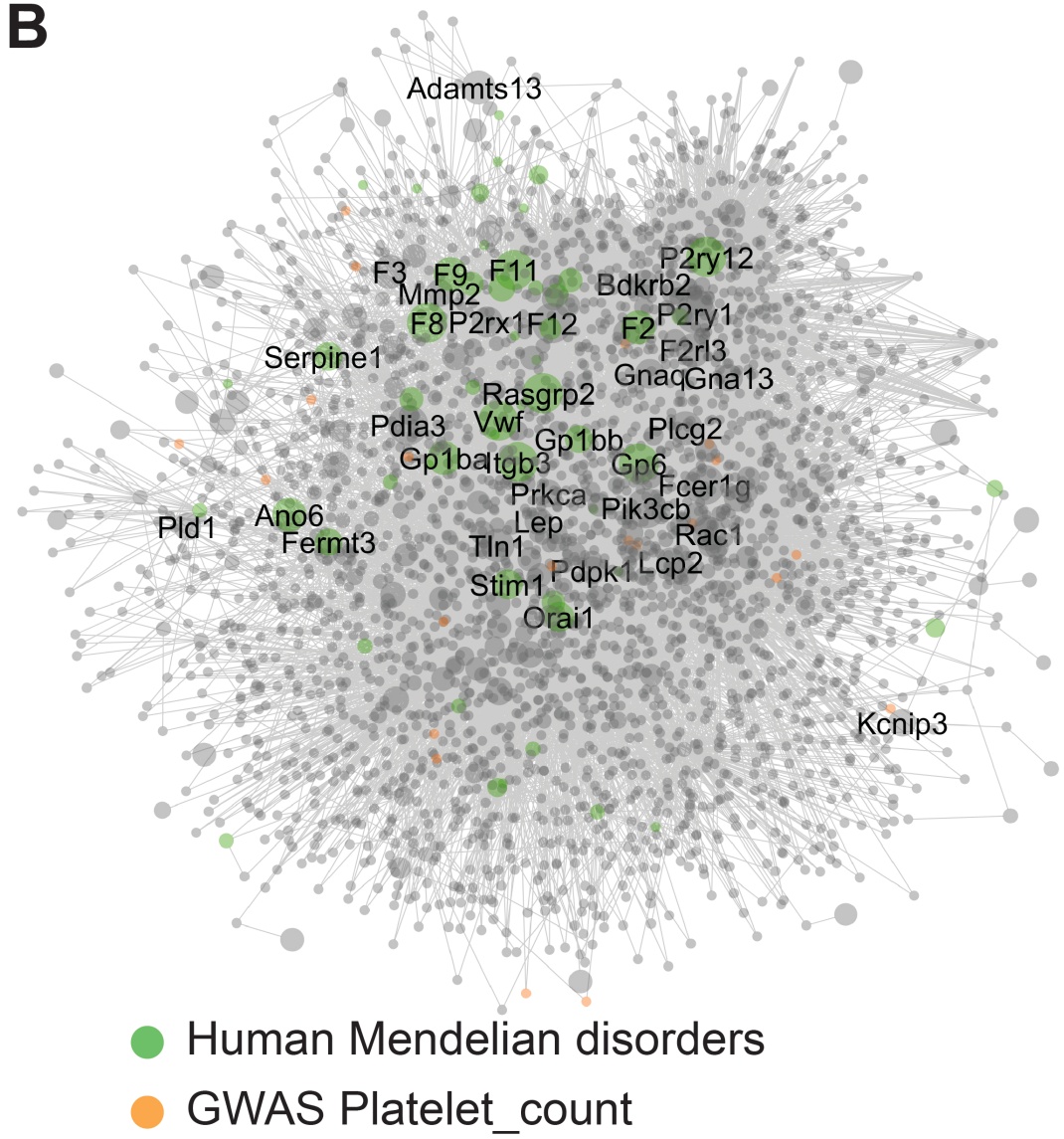

I II III IV
193 Mmp10
194 Mmp2
195 Mmp3
196 Mmp9
197 Mpl
198 Mrvi1
199 Myd88
200 Myh9
201 Myo5a
202 Nbeal2
203 Nfe2
204 Nod2
205 Nos2
206 Nos3
207 Nox1
208 Nr1h2
209 Nr3c2
210 Nt5e
211 Oca2
212 Ophn1
213 Orai1
214 Ostm1
215 P2rx1
216 P2rx7
217 P2ry1
218 P2yr12
219 Palld
220 Pde3a
221 Pde3b
222 Pdia3
223 Pdia6
224 Pdlim1
225 Pdlim7
226 Pdpk1
227 Pear1
228 Pecam1
229 Per2
230 Pf4
231 Pik3c2a
232 Pik3c2b
233 Pik3ca
234 Pik3cb
235 Pik3cg
236 Pik3r1
237 Pip5k1b
238 Pip5k1c
239 Pla2g4a
240 Plat
241 Plaur
242 Plcb2 (Plcb3)
243 Plcb3 (Plcb2)
244 Plcg1
245 Plcg2
246 Pld1
247 Pld2
248 Plg
249 Plscr1
250 Pltp
251 Pmel
252 Pnpla8
253 Podxl
254 Ppara
255 Pparg
256 Ppia
257 Ppif
258 Ppp1cc
259 Ppp1r9b
260 Ppp3ca
261 Prcp
262 Prex1
263 Prkaa2
264 Prkca
265 Prkcb
266 Prkcd
267 Prkce
268 Prkcq
269 Prkd2
270 Prkg1
271 Proc
272 Procr
273 Pros1
274 Proz
275 Prdxll
276 Pten
277 Ptger3
278 Ptges
279 Ptgir
280 Ptgs1
281 Ptgs2
282 Ptk2b
283 Ptpn1
284 Ptpn11
285 Ptpn6
286 Ptprj
287 Ptx3
288 Rab27a

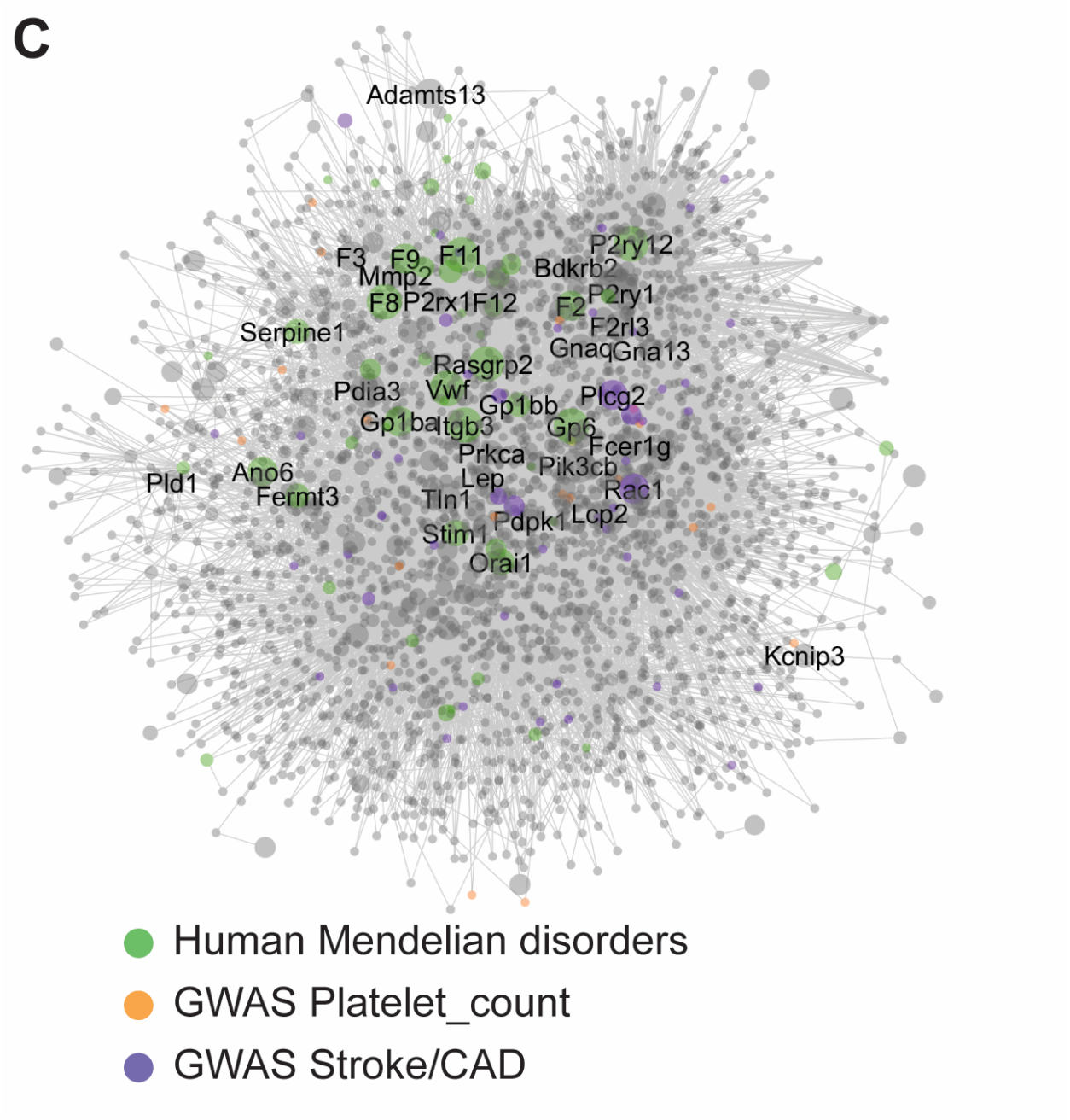

I II III IV
289 Rab27b
290 Rabggta
291 Rac1
292 Rac2
293 Ranbp10
294 Rap1b
295 Rasa3
296 Rasgrp2
297 Reln
298 Rgs10
299 Rgs18
300 Rgs2
301 Rhoa
302 Rhog
303 Robo1
304 Rock1
305 S100a9
306 Scarb1
307 Scube1
308 Selp
309 Selplg
310 Sema4d
311 Serpina10
312 Serpinc1
313 Serpind1
314 Serpine1
315 Serpine2
316 Serpinf2
317 Sgk1
318 Sh2b3
319 Sh2d1a
320 Sla (Sla2)
321 Sla2 (Sla)
322 Slamf1
323 Slc30a4
324 Slc45a2
325 Slc4a1
326 Slc7a11
327 Smpd1
328 Snap29
329 Sphk1
330 Src
331 Srf
332 Srgn
333 Stat3
334 Stat5a (Stat5b)
335 Stat5b (Stat5a)
336 Stim1
337 Stim2
338 Stx8
339 Stxbp5
340 Svil
341 Syk
342 Tac1
343 Tacr1
344 Tbxa2r
345 Tbxas1
346 Tfpi
347 Tgfb1
348 Tgfb1i1
349 Tgfbi
350 Thbd
351 Thbs1
352 Thbs2
353 Tln1
354 Tlr4
355 Tnfrsf1a (Tnfrsf1b)
356 Tnfrsf1b (Tnfrsf1a)
357 Tph1
358 Treml1
359 Trp53
360 Trpc1
361 Trpc6
362 Tslp
363 Tspan32
364 Tubb1
365 Tymp
366 Tyr
367 Tyro3
368 Tyrp1
369 Ubash3a (Ubash3b)
370 Ubash3b (Ubash3a)
371 Unc13d
372 Vamp3
373 Vamp7
374 Vamp8
375 Vav1 (Vav3)
376 Vav3 (Vav1)
377 Vcl
378 Vim
379 Vps33a
380 Vps33a
381 Vtn
382 Vwf
383 Wdr1
384 Ywhaz

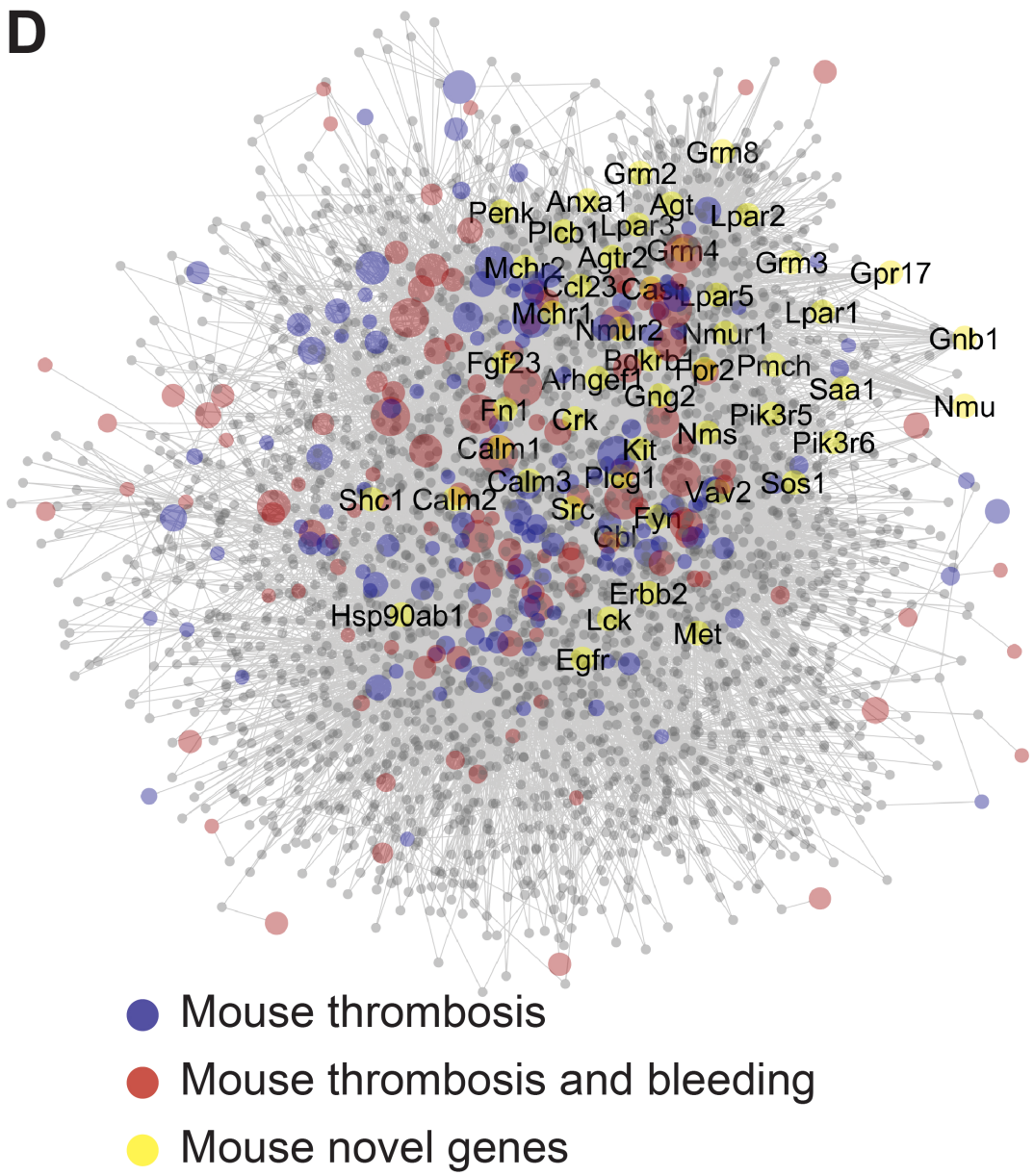


Suppl. Figure 2. Overview per mouse gene and study of scores in classes I-IV. Dots representing all individual studies, as scored in Classes I, II, III or IV, as indicated (for complete overview, see Data File 1). Open dots refer to a gain-of-function study, scored as if it were a loss-of-function. Green dots refer to negative scores, red dots to positive scores.


A
Adamts13
F3
F9
F11
Mmp2
Bdkrb2
P2ry12
F8
P2rx1
F12
F2
P2ry1
Serpine1
F2rl3
Gnaq
Gna13
Rasgrp2
Pdia3
Vwf
Gp1bb
Plcg2
Gp1ba
Itgb3
Gp6
Fcer1g
Prkca
Pik3cb
Pld1
Ano6
Lep
Fermt3
Tln1
Rac1
Pdpk1
Lcp2
Stim1
Orai1
Kcnip3
Mouse thrombosis
Mouse thrombosis and bleeding

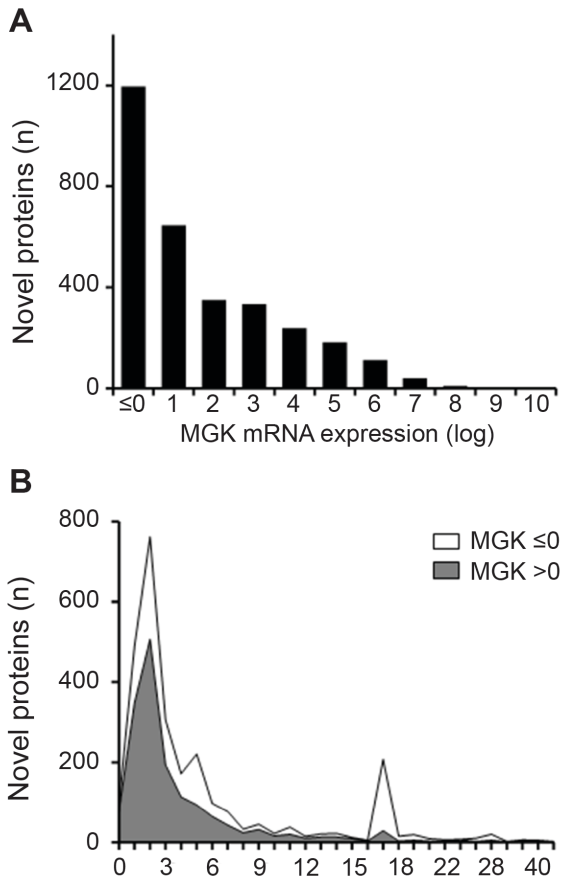

B
Adamts13
F3
F9
F11
Mmp2
F8
P2rx1
F12
Serpine1
Rasgrp2
Pdia3
Vwf
Gp1bb
Gp1ba
Itgb3
Prkca
Lep
Pld1
Ano6
Fermt3
Tln1
Stim1
Pdpk1
Orai1
Bdkrb2
P2ry12
P2ry1
F2
F2rl3
Gnaq
Gna13
Plcg2
Gp6
Fcer1g
Pik3cb
Rac1
Lcp2
Kcnip3
Human Mendelian disorders
GWAS Platelet_count


C
Adamts13
F3
F9
F11
Mmp2
Bdkrb2
P2ry12
F8
P2rx1
F12
F2
P2ry1
F2rl3
Serpine1
GnaqGna13
Rasgrp2
Pdia3
Vwf
Gp1bb
Plcg2
Gp1ba
Itgb3
Gp6
Fcer1g
Prkca
Pik3cb
Pld1
Ano6
Lep
Rac1
Fermt3
Tln1
Pdpk1
Lcp2
Stim1
Orai1
Kcnip3
Human Mendelian disorders
GWAS Platelet_count
GWAS Stroke/CAD

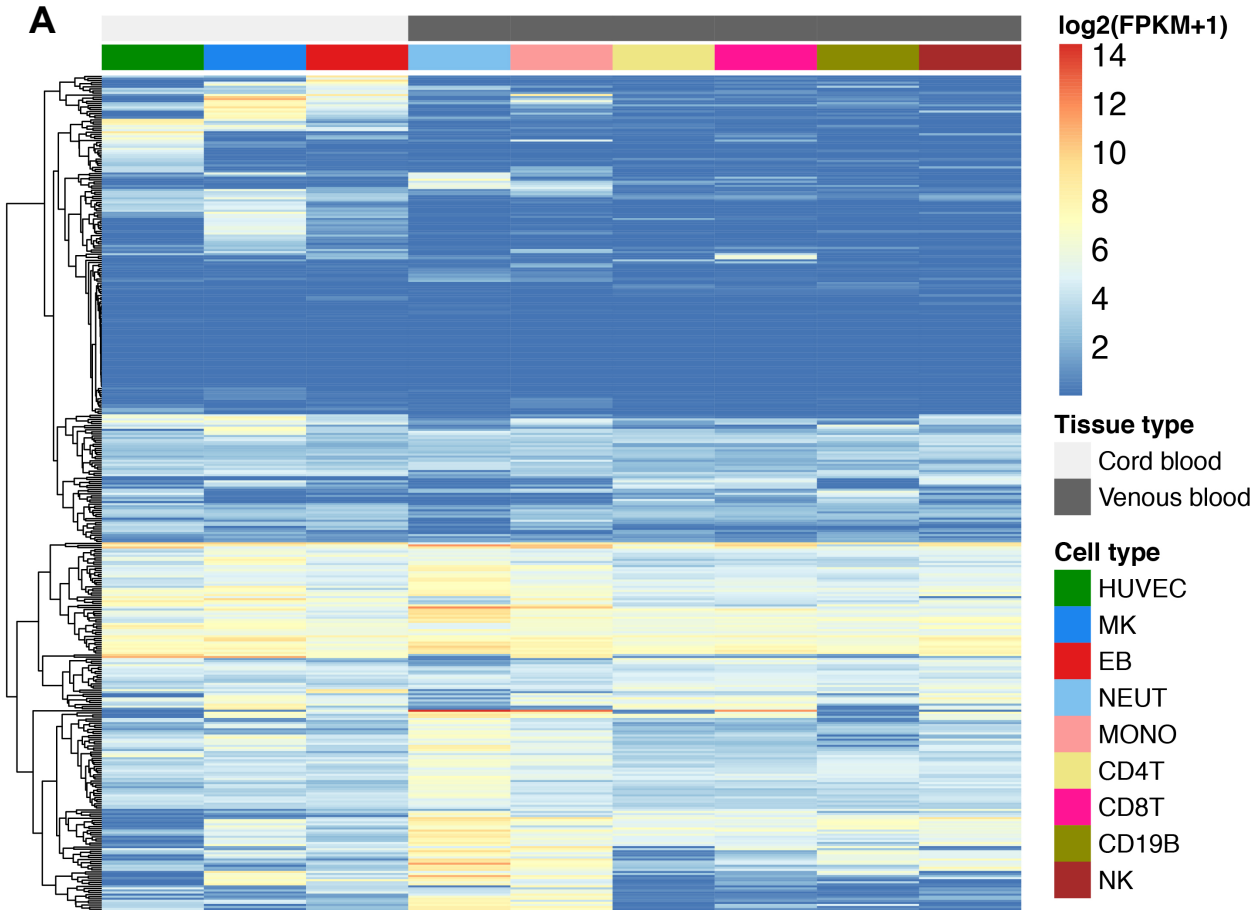


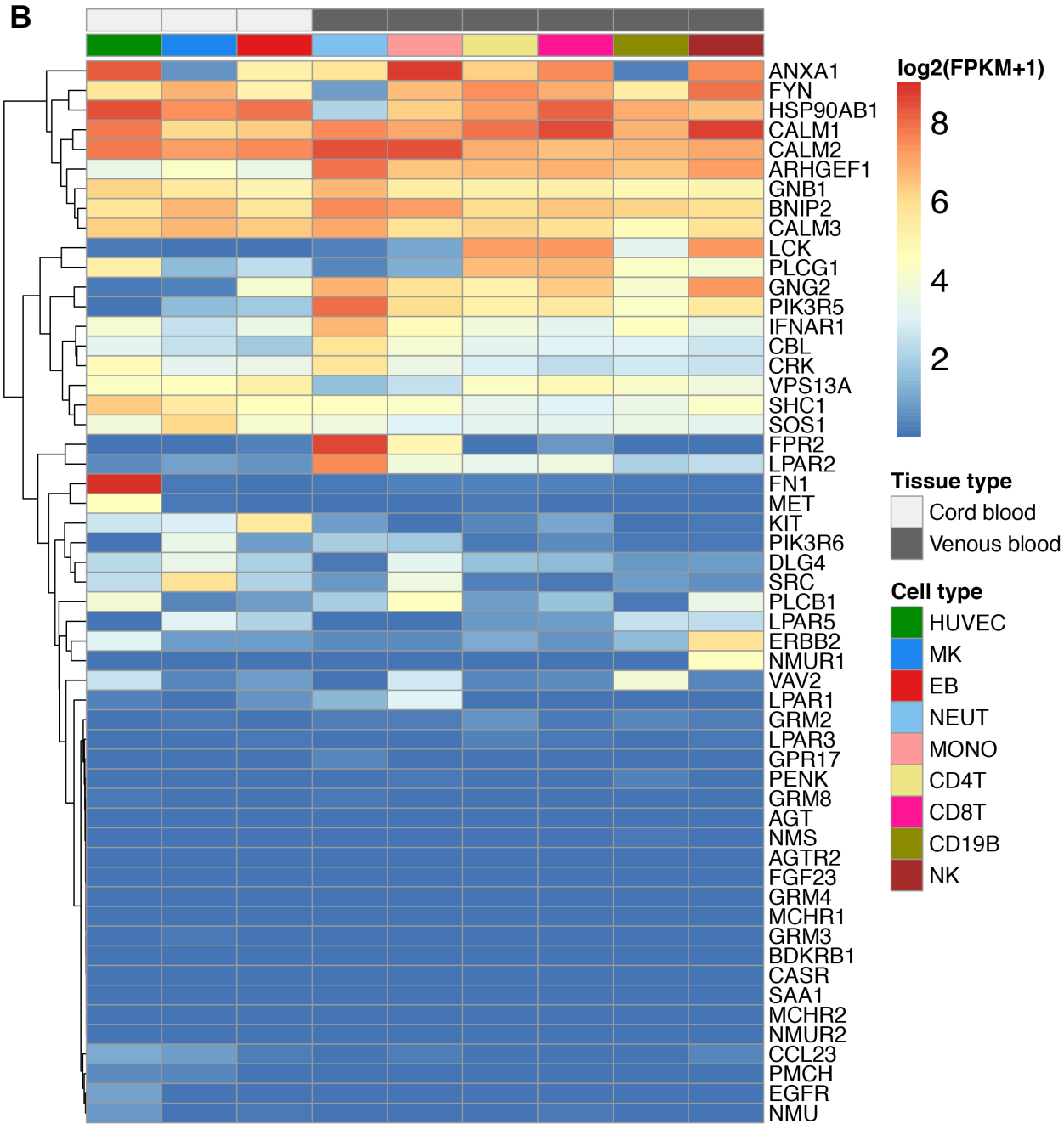
Suppl. Figure 3. Network of protein interactions in thrombosis and hemostasis (PITH). Network built in Cytoscape for 275 bait proteins, corresponding to mouse genes with modifying effect on arterial thrombosis of hemostasis. The final network has 267 core genes (bait nodes), 2679 new nodes, connected by 19721 interactions (edges). Of the edges, 15419 are derived from Reactome, 4299 from Intact, and 3 from manual assignment. Size of nodes is set as 1 per default, or as CP value. **A**, Network visualization with color-coded bait nodes: proteins modifying thrombosis (blue) or thrombosis + bleeding (red). Names are listed of 40 mouse genes with highest CP values. **B**, Network visualization with color-coded proteins linked to rare Mendelian disorders of human platelet or coagulant function (green, *n*=54); proteins linked to high association genes in GWAS of platelet count and volume (orange, *n*=21). **C**, Network visualization with additional color-coded proteins linked to Megastroke GWAS (purple,

n=11) or linked to coronary artery disease (*n*=40). **D**, Network visualization emphasizing 50 novel nodes with largest number of edges (≥25, yellow). Attribute lists are given in Data File 6.



Suppl. Figure 4. Buildup of protein-protein interaction network. **A**, Distribution of 2,741 novel proteins in network with respect to mRNA expression in megakaryocytes (MGK). Values are given as log2 (FPKM+1) (FPKM = fragments per kilobase of transcript per million mapped reads; values <1 are considered as a lack of transcription). **B**, Distribution of 2,741 novel proteins with respect to number of connections (edges) with modifying mouse proteins (thrombosis and/or bleeding).


A
log2(FPKM+1)
14
12
10
8
6
4
2
Tissue type
Cord blood
Venous blood
Cell type
HUVEC
MK
EB
NEUT
MONO
CD4T
CD8T
CD19B
NK



Suppl. Figure 5. Clustered heatmap of relative mRNA expression levels in relevant blood cells. Relative mRNA expression levels are shown for the 409 scored genes **(A)** and the top 50 most connected novel nodes **(B),** obtained from the Blueprint database,[12] of different cell types (see color code upper bar). Values are given as log2 (FPKM+1), as in Suppl. Figure 3. Abbreviations: resting endothelial cells of umbical vein (HUVEC), CD34-negative, CD41-positive, CD42-positive megakaryocytes (MK), erythroblasts (EB), mature neutrophils (NEUT), CD14-positive, CD16-negative classical monocytes (MONO), CD4-positive, alpha-beta T cells (CD4T), CD8-positive, alpha-beta T cells (CD8T), CD38-negative naive B cells (CD19B), and cytotoxic CD56-dim natural killer cells (NK).

Suppl. Table 1. Key words and numbers of publications from search outcomes in PubMed.

Keywords used for search	# Publications 1980-2017	2018*
Class I-III		
mouse & thrombus	5740	267
mouse & thrombus & arterial	1440	58
mouse & thromboembolism	771	24
mouse & thromboembolism & arterial	211	4
Class IV		
mouse & bleeding	8245	355
mouse & tail bleeding	594	32
mouse & tail clipping	45	0

*Studies added in 2018 (until August 1st) are included in gene lists, but not used for statistical analysis and calculations.

Suppl. Table 2. Classification and grouping of analyzed mouse studies. Indicated are the assignment of studies to classes and groups, as well as the evaluated output parameters of thrombus formation, thromboembolism or tail bleeding of indicated classes. For an overview, see Figure 1.

Class I: in vivo arterial thrombosis	
*Groups compared (**1-12**)*	
Carotid artery injury	$FeCl_3$ (**1**); photochemical (**2**); ligation (**3**); electrolytic (**4**)
Aorta injury	compression (**5**)
Mesenteric artery injury	$FeCl_3$ (**6**); photochemical (**7**); laser (**8**)
Cremaster artery injury	$FeCl_3$ (**9**); photochemical (**10**); laser (**11**)
Other artery/injury	any (**12**)
*Parameters compared (**a-c**)*	
Time-dependent parameters	time to platelet adhesion, time to occlusion or thrombus (**a**)
Mass-dependent parameters	extent of platelet adhesion (**b*i***); extent of occlusion (**b*ii***); number, size or volume of thrombi (**b*iii***)
Stability parameters	number or frequency of emboli (**c**)
Class II: in vivo thromboembolism	
*Groups compared (**20-23**)*	
Provoked ***(d)***	injection of collagen/epinephrine (**20**); tissue factor (**21**); or other agent (**22**)
Unprovoked ***(e)***	pre/after birth (**23**)
Parameters compared	
Thromboembolism parameters	percentage death, time to death, residual platelet count (**d**); unprovoked thrombosis or bleeding (**e**)
Class III: in vitro thrombus formation	
*Groups compared (**30-34**)*	
Adhesive surface	collagen I (**30**); collagen III (**31**); VWF (**32**), fibrinogen (**33**); other (**34**)
*Parameters compared (**f-h**)*	
Time-dependent parameters	time to platelet adhesion, time to thrombus formation (**f**)
Mass-dependent parameters	extent of platelet adhesion (**g*i***); surface area coverage (**g*ii***); number, size or volume of thrombi (**g*iii***)
Stability parameters	number or frequency of emboli (**h**)
Class IV: bleeding time	tail bleeding time

Suppl. Table 3. Correlation analysis of intra-study sub-parameters scored on 3-point or 5-point scale. Indicated are numbers of studies per parameter, Kendall's Tau correlation coefficients, and *P* values regarding sub-parameters **b**i-*iii*, **g***i-iii* and regarding parameters ***a-c*** and ***f-h***. **A**, Comparison of mass-dependent sub-parameters in class I studies (***bi***, platelet adhesion; ***bii***, extent of occlusion; ***biii***, thrombus size, volume or number). **B**, Comparison of mass-dependent sub-parameters in class III studies (***gi***, platelet adhesion; ***gii***, surface-area-coverage; ***giii***, thrombus size, volume or number). **C**, Comparison of time-dependent (***a***), mass-dependent (***b***), and stability (***c***) parameters in class I studies. Time-dependent parameters (***a/f***) included time to occlusion and time to thrombus formation; stability parameters (***c/h***) included number of emboli. **D**, Comparison of time-dependent (***f***), mass-dependent (***g***), and stability (***h***) parameters for all class III studies. Green cells indicate statistical significance.

A

Number of studies

	bi	bii	biii
bi	75		
bii	9	87	
biii	20	10	153

Kendall's τ (3-point scale)

	bi	bii	biii
bi			
bii	0.81		
biii	0.84	1.00	

P-value (3-point scale)

	bi	bii	biii
bi			
bii	0.015		
biii	<0.001	<0.01	

Kendall's τ (5-point scale)

	bi	bii	biii
bi			
bii	0.50		
biii	0.70	0.79	

P-value (5-point scale)

	bi	bii	biii
bi			
bii	0.102		
biii	<0.001	0.006	

B

Number of studies

	gi	gii	giii
gi	42		
gii	24	198	
giii	12	74	105

Kendall's τ (3-point scale)

	gi	gii	giii
gi			
gii	0.64		
giii	1.00	0.74	

P-value (3-point scale)

	gi	gii	giii
gi			
gii	0.001		
giii	<0.01	<0.001	

Kendall's τ (5-point scale)

	gi	gii	giii
gi			
gii	0.54		
giii	0.93	0.66	

P-value (5-point scale)

	gi	gii	giii
gi			
gii	0.003		
giii	<0.001	<0.001	

C

Number of studies

	a	b	c
a	402		
b	95	278	
c	48	46	80

Kendall's τ (3-point scale)

	a	b	c
a			
b	0.80		
c	0.43	0.65	

P-value (3-point scale)

	a	b	c
a			
b	<0.001		
c	0.002	<0.001	

Kendall's τ (5-point scale)

	a	b	c
a			
b	0.73		
c	0.16	0.61	

P-value (5-point scale)

	a	b	c
a			
b	<0.001		
c	0.226	<0.001	

D

Number of studies

	f	g	h
f	13		
g	10	239	
h	0	17	17

Kendall's τ (3-point scale)

	f	g	h
f			
g	0.67		
h	-	0.81	

P-value (3-point scale)

	f	g	h
f			
g	0.046		
h	-	0.001	

Kendall's τ (5-point scale)

	f	g	h
f			
g	0.56		
h	-	0.75	

P-value (5-point scale)

	f	g	h
f			
g	0.072		
h	-	0.001	

Suppl. Table 4. Correlation analysis of mean study scores per class on 3-point or 5-point scale. Comparisons of study scores on 3-point or 5-point scales separated according to classes I, II and III. **A**, Comparison between classes per studies published in the same paper (identical mouse strain). **B**, Comparison between classes per gene (loss-of function mutations only). Panels show numbers of studies or genes compared, Kendall's Tau correlation coefficients, and *P* values (3-point or 5-point scores, as indicated). Green cells indicate statistical significance with $P<0.05$.

A

Number of studies

	I	II	III
I	602		
II	54	102	
III	151	14	286

Kendall's τ (3-point scale)

	I	II	III
I			
II	0.79		
III	0.76	0.71	

P-value (3-point scale)

	I	II	III
I			
II	<0.001		
III	<0.001	0.006	

Kendall's τ (5-point scale)

	I	II	III
I			
II	0.68		
III	0.64	0.25	

P-value (5-point scale)

	I	II	III
I			
II	<0.001		
III	<0.001	0.286	

B

Number of genes

	I	II	III
I	278		
II	51	79	
III	129	32	161

Kendall's τ (3-point scale)

	I	II	III
I			
II	0.80		
III	0.70	0.52	

P-value (3-point scale)

	I	II	III
I			
II	<0.001		
III	<0.001	0.003	

Kendall's τ (5-point scale)

	I	II	III
I			
II	0.61		
III	0.58	0.09	

P-value (5-point scale)

	I	II	III
I			
II	<0.001		
III	<0.001	0.556	

Suppl. Table 5. Numbers of scored mouse genes per class and group. For description of classes and groups, see Suppl. Table 1.

Class I (number of genes)

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	Carotis $FeCl_3$	161											
2	Carotis photochemical	19	38										
3	Carotis ligation	7	3	13									
4	Carotis electrolytical	5	0	3	8								
5	Aorta compression	6	1	0	0	16							
6	Mesentery $FeCl_3$	43	9	7	2	12	97						
7	Mesentery photochemical	1	1	0	0	0	1	3					
8	Mesentery laser	8	0	1	1	1	6	0	14				
9	Cremaster $FeCl_3$	2	0	1	1	0	4	0	2	5			
10	Cremaster photochemical	6	0	0	0	0	3	0	1	0	12		
11	Cremaster laser	19	4	4	3	2	19	0	5	5	2	60	
12	Other	4	4	2	1	2	6	0	3	1	0	1	12

Class II (number of genes)

		20	**21**	**22**	**23**
20	Injection collagen-epinephrine	49			
21	Injection tissue factor	2	10		
22	Injection other	1	0	11	
23	Systemic other	0	1	0	12

Class III (number of genes)

		30	**31**	**32**	**33**	**34**
30	Collagen I	156				
31	Collagen III	4	8			
32	von Willebrand factor	4	0	4		
33	Fibrinogen	2	1	0	3	
34	Other	3	0	0	0	4

Class IV (number of genes)

		40
40	Tail bleeding	314

Suppl. Table 6. Node connections in PITH network for top 50 most connected proteins. Also indicated are: numbers of interactors in network, mRNA expression levels in megakaryocytes (MK)[12], human UniProt identifiers, protein names, OMIM identifiers regarding disease associations, and GWAS data (1: Mendelian disorders[6], 2: Platelet_count[7], 3: Platelet_traits[8], 4: Stroke[9] and 5: CAD[10])

Gene	Number of interactors in PITH network	mRNA expres-sion	UniProt identifier (human)	Protein name	OMIM identifier	GWAS
EGFR	51	0.02	P00533	Epidermal growth factor receptor	*131550	
FYN	43	6.79	P06241	Tyrosine-protein kinase Fyn	*137025	
GNB1	39	5.56	P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	*139380	
GNG2	37	0.35	P59768	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-2	*606981	
PIK3R5	37	1.51	Q8WYR1	Phosphoinositide 3-kinase regulatory subunit 5	*611317	
PIK3R6	37	3.43	Q5UE93	Phosphoinositide 3-kinase regulatory subunit 6	*611462	
SRC	37	5.80	P12931	Proto-oncogene tyrosine-protein kinase Src	*190090	
FN1	35	0.12	P02751	Fibronectin	*135600	
SHC1	35	5.47	P29353	SHC-transforming protein 1	*600560	
PLCB1	33	0.40	Q9NQ66	1-Phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-1	*607120	
ANXA1	31	0.64	P04083	Annexin A1	*151690	
ERBB2	31	0.89	P04626	Receptor tyrosine-protein kinase erbB2	*164870	
LCK	31	0.02	P06239	Tyrosine-protein kinase Lck	*153390	
PLCG1	31	1.50	P19174	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1	*172420	5
SOS1	31	6.11	Q07889	Son of sevenless homolog 1	*182530	
AGT	29	0.05	P01019	Angiotensinogen	+106150	5
HSP90AB1	29	7.41	P08238	Heat shock protein HSP 90-beta	*140572	
MET	29	0.09	P08581	Hepatocyte growth factor receptor	*164860	
CRK	28	3.31	P46108	Adapter molecule crk	*164762	
SAA1	28	0.00	P0DJI8	Serum amyloid A-1 protein	*104750	
BDKRB1	27	0.00	P46663	B1 bradykinin receptor	*600337	
CALM1	27	6.11	P0DP23	Calmodulin-1	*114180	

CALM2	27	7.16	P0DP24	Calmodulin-2	*114182	
CALM3	27	6.75	P0DP25	Calmodulin-3	*114183	
CASR	27	0.00	P41180	Extracellular calcium-sensing receptor	+601199	
CCL23	27	0.84	P55773	C-C motif chemokine 23	*602494	
FPR2	27	0.06	P25090	N-formyl peptide receptor 2	*136538	
GPR17	27	0.00	Q13304	Uracil nucleotide/cysteinyl leukotriene receptor	*603071	
KIT	27	2.91	P10721	Mast/stem cell growth factor receptor Kit	*164920	
LPAR1	27	0.06	Q92633	Lysophosphatidic acid receptor 1	*602282	
LPAR2	27	0.94	Q9HBW0	Lysophosphatidic acid receptor 2	*605110	
LPAR3	27	0.01	Q9UBY5	Lysophosphatidic acid receptor 3	*605106	
LPAR5	27	3.11	Q9H1C0	Lysophosphatidic acid receptor 5	*606926	
MCHR1	27	0.02	Q99705	Melanin-concentrating hormone receptor 1	*601751	
MCHR2	27	0.00	Q969V1	Melanin-concentrating hormone receptor 2	*606111	
NMS	27	0.01	Q5H8A3	Neuromedin-S	-	
NMU	27	0.01	P48645	Neuromedin-U	*605103	
NMUR1	27	0.00	Q9HB89	Neuromedin-U receptor 1	*604153	
NMUR2	27	0.00	Q9GZQ4	Neuromedin-U receptor 2	*605108	
PENK	27	0.02	P01210	Proenkephalin-A	*131330	
PMCH	27	0.40	P20382	Pro-MCH	*176795	
VAV2	27	0.38	P52735	Guanine nucleotide exchange factor VAV2	*600428	
CBL	26	2.46	P22681	E3 ubiquitin-protein ligase CBL	*165360	2/3
FGF23	26	0.00	Q9GZV9	Fibroblast growth factor 23	*605380	
AGTR2	25	0.00	P50052	Type-2 angiotensin II receptor	*300034	
ARHGEF1	25	4.33	Q92888	Rho guanine nucleotide exchange factor 1	*601855	
GRM2	25	0.14	Q14416	Metabotropic glutamate receptor 2	*604099	
GRM3	25	0.10	Q14832	Metabotropic glutamate receptor 3	*601115	
GRM4	25	0.01	Q14833	Metabotropic glutamate receptor 4	*604100	3
GRM8	25	0.04	O00222	Metabotropic glutamate receptor 8	*601116	

Suppl. Table 7: Mouse/human homology of 60 orthologous genes and proteins with role in arterial thrombosis with or without bleeding (see Table 2). Median homology on nucleotide level: 89.47 (IQR: 77.11-94.69), median pairwise alignment score on protein level: 91.80 (IQR: 81.35-95.40)

Gene	Protein	Score Class I-III	Score class IV	Homology nucleotide (mouse - human)	Pairwise alignment score protein (mouse - human)
Antithrombotic, no prolonged bleeding					
F11	Coagulation factor XI	-1	0	78.04	78.30
F12	Coagulation factor XII	-2	0	71.69	72.60
SERPINE1	Plasminogen activator inhibitor 1	-1	0	79.10	79.10
DUSP3	Dual specificity protein phosphatase 3	-2	0	92.97	93.50
GP6	Platelet glycoprotein VI	-2	0	60.06	69.60
ITGA6	Integrin alpha-6	-2	0	93.13	93.10
P2RX1	P2X purinoceptor 1	-2	0	89.47	89.50
SELP	P-selectin	-1	0	72.79	68.90
CSNK2B	Casein kinase II subunit beta	-2	0	54.09	100.00
PIK3C3	Phosphatidylinositol 3-kinase catalytic subunit type 3	-2	0	97.97	98.00
PIK3CG	Phosphatidylinositol 4,5-bisphoshpate 3-kinase catalytic subunit gamma isoform	-2	0	94.10	94.10
PLD1	Phospholipase D1	-1	0	91.31	91.30
PPIA	Peptidyl-prolyl cis-trans isomerase A	-2	0	n.a.	96.30
PRKCA	Protein kinase C alpha type	-2	0	99.40	99.60
RHOG	Rho-related GTP-binding protein RhoG	-1	0	100.00	100.00
S100A9	Protein S100-A9	-2	0	56.64	60.40
SGK1	Serine/threonine-protein kinase Sgk1	-2	0	94.47	94.50
Antithrombotic, prolonged bleeding					
F2	Prothrombin	-2	-1	81.72	82.00
F8	Coagulation factor VIII	-2	-2	73.65	74.20
F9	Coagulation factor IX	-2	-2	80.25	82.00
KLKB1	Plasma kallikrein	-1	-1	76.18	76.20
VWF	Von Willebrand factor	-2	-2	83.36	n.d.
ANO6	Anoctamin-6	-2	-1	90.34	89.90
CLEC1B	C-type lectin domain family 1 member B	-1	-1	63.76	64.60
F2RL3	Proteinase-activated receptor 4	-1	-2	74.49	78.30
FCER1G	High affinity immunoglobulin epsilon receptor subunit gamma	-2	-1	88.37	89.60
FERMT3	Fermitin family homolog 3	-2	-2	93.53	93.70

GNA13	Guanine nucleotide-binding protein subunit alpha-13	-2	-2	96.82	96.80
GNAQ	Guanine nucleotide-binding protein G(q) subunit alpha	-1	-2	99.72	99.70
GPIBA	Platelet glycoprotein Ib alpha chain	-2	-2	52.45	63.70
GPIBB	Platelet glycoprotein Ib beta chain	-1	-2	84.11	93.20
ITGA2	Integrin alpha-2	-1	-1	82.51	82.50
ITGB3	Integrin beta-3	-2	-2	90.72	90.70
MERTK	Tyrosine-protein kinase Mer	-1	-1	81.39	81.90
ORAI1	Calcium release-activated calcium channel protein 1	-1	-1	71.05	91.70
P2RY1	P2Y purinoceptor 1	-1	-2	94.91	94.90
P2RY12	P2Y purinoceptor 12	-2	-2	85.88	89.00
PDIA3	Protein disulfide-isomerase A3	-2	-1	93.07	93.10
PDIA4	Protein disulfide-isomerase A4	-2	-1	88.46	88.30
PTPRJ	Receptor-type tyrosine-protein phosphatase eta	-1	-1	66.44	72.70
STIM1	Stromal interaction molecule 1	-1	-1	96.79	96.80
TRPM7	Transient receptor potential cation channel subfamily M member 7	-2	-1	94.36	94.40
TSPAN32	Tetraspanin-32	-2	-1	63.28	64.00
CSK	Tyrosine-protein kinase CSK	-1	-2	99.11	99.10
MAP3K5	Mitogen-activated protein kinase kinase kinase 5	-2	-2	94.42	94.90
MAPK8	Mitogen-activated protein kinase 8	-2	-1	99.06	97.60
MAPK14	Mitogen-activated protein kinase 14	-1	-1	99.44	99.40
PDPK1	3-phosphoinositide-dependent protein kinase 1	-1	-1	94.28	95.10
PIK3CB	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit beta isoform	-1	-1	95.86	95.90
PLCG2	1-Phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2	-2	-2	94.15	94.20
RAC1	Ras-related C3 botulinum toxin substrate 1	-2	-2	98.10	100.00
RASGRP2	RAS guanyl-releasing protein 2	-2	-2	96.38	96.40
TLN1	Talin-1	-2	-2	98.74	98.70
NBEAL2	Neurobeachin-like protein 2	-2	-2	88.22	88.40
KCNIP3	Calsenilin	-2	-1	72.54	91.80
MMP2	72 kDa type IV collagenase	-1	-2	95.62	95.90
BAMBI	BMP and activin membrane-bound inhibitor homolog	-1	-1	93.08	93.10
BDKRB2	B2 bradykinin receptor	-2	-2	79.85	80.80
MFAP2	Microfibrillar-associated protein 2	-1	-1	87.57	90.60
Not antithrombotic, prolonged bleeding					
CEACAM1	Carcinoembryonic antigen-related cell adhesion molecule 1	2	-1	57.39	58.00