Supplementary Data

Table 1. Factors identified using Bayesian factor regression modeling of peripheral blood gene expression data and their characteristics.

Factor Number	Number of genes within Factor	GO terms	KEGG terms
1	521	protein modification, cellular physiological process, protein transport, establishment of protein localization, protein localization organismal physiological process ubiquitin cycle, G-protein coupled receptor protein signaling pathway neurophysiological process, cell surface receptor linked signal transduction, protein metabolism intracellular protein transport, response to external stimulus, cellular protein metabolism, intracellular transport, cellular macromolecule metabolism, macromolecule metabolism, response to stimulus	NS
2	142	cellular physiological process	NS
3	16	NS	NS
4	58	protein biosynthesis, macromolecule biosynthesis, cellular biosynthesis biosynthesis, cellular protein metabolism, protein metabolism macromolecule metabolism, cellular macromolecule metabolism	oxidative phosphorylation
5	29	response to biotic stimulus, immune response, defense response, deoxyribonucleotide metabolism	NS
6	91	response to biotic stimulus, immune response, defense response, response to stimulus, organismal physiological process, progesterone metabolism	NS
7	120	NS	NS
8	291	immune response, response to external biotic stimulus, response to pest, pathogen or parasite, response to biotic stimulus, defense response, inflammatory response, response to wounding	NS
9	153	NS	NS
10	68	NS	NS
11	11	NS	NS
12	248	cellular physiological process	NS
13	53	gas transport, oxygen transport	NS

14	62	blood coagulation, coagulation hemostasis, regulation of body fluids, platelet activation	NS
15	99	NS	NS
16	32	response to biotic stimulus	Nicotinate and nicotinamide metabolism
17	115	NS	NS
18	49	NS	NS
19	52	NS	NS
20	73	NS	NS
GO = Gene Ontology; KEGG = Kyoto Encyclopedia of Genes and Genomes; NS = no significantly associated terms/pathways (see Methods for significance level)			

Table 2. Validation of microarray based gene expression for Factor 14 probe sets with RT-PCR in HV2 cohort whole blood RNA.

Gene name	Taqman RTPCR assay name	Affymetrix microarray probe set name	Correlation coefficient	Correlation p-value
ITGA2B	Hs01116228_m1	206494_s_at	-0.94421	9.05E-26
TREML1	Hs00698316_m1	1555659_a_at	-0.90545	3.04E-20
ITGA2B	Hs01116228_m1	206493_at	-0.90253	6.28E-20
MPL	Hs00180489_m1	207550_at	-0.89971	1.24E-19
MYL9	Hs00697086_m1	201058_s_at	-0.89826	1.74E-19
PCSK6	Hs00159844_m1	207414_s_at	-0.893	5.76E-19
ITGA2B	Hs01116228_m1	216956_s_at	-0.88851	1.53E-18
CLU	Hs00971656_m1	208791_at	-0.88747	1.90E-18
CLU	Hs00971656_m1	208792_s_at	-0.88153	6.40E-18
SPARC	Hs00277762_m1	200665_s_at	-0.87198	3.95E-17
ALOX12	Hs00167524_m1	207206_s_at	-0.86138	2.53E-16
CMTM5	Hs00370784_m1	230942_at	-0.85545	6.71E-16
SH3BGRL2	Hs00230283_m1	225354_s_at	-0.85496	7.26E-16
CTDSPL	Hs00505109_m1	201906_s_at	-0.84553	3.12E-15
PPBP	Hs00234077_m1	214146_s_at	-0.84171	5.48E-15
PF4	Hs00236998_m1	206390_x_at	-0.83419	1.59E-14
PBX1	Hs00295499_s1	212151_at	-0.82951	3.01E-14
CTTN	Hs01124225_m1	201059_at	-0.82216	7.89E-14
SPARC	Hs00277762_m1	212667_at	-0.80984	3.60E-13
ITGB3	Hs01001469_m1	204628_s_at	-0.79402	2.17E-12
TUBB1	Hs00258236_m1	230690_at	-0.79401	2.17E-12
GNG11	Hs00914578_m1	204115_at	-0.78661	4.78E-12
TGFB1I1	Hs00210887_m1	209651_at	-0.76354	4.63E-11
PBX1	Hs00295499_s1	212148_at	-0.76222	5.23E-11
GP1BB	Hs00236857_m1	206655_s_at	-0.74866	1.75E-10
TUBB1	Hs00258236_m1	208601_s_at	-0.74411	2.59E-10
ELOVL7	Hs00405151_m1	227180_at	-0.72899	8.91E-10
THBS1	Hs00962914_m1	201108_s_at	-0.70496	5.42E-09
SELP	Hs00356351_m1	206049_at	-0.69728	9.30E-09
SLC24A3	Hs00221141_m1	57588_at	-0.65927	1.07E-07
CALD1	Hs00921982_m1	212077_at	-0.64826	2.04E-07
SLC24A3	Hs00221141_m1	219090_at	-0.63347	4.65E-07
MYLK	Hs00364926_m1	224823_at	-0.6135	1.33E-06
ARHGAP6	Hs00241801_m1	206167_s_at	-0.60772	1.77E-06
PRKAR2B	Hs00176966_m1	203680_at	-0.57731	7.45E-06
PDE5A	Hs00903251_m1	227088_at	-0.55683	1.81E-05
FSTL1	Hs00907496_m1	208782_at	-0.53538	4.32E-05
TTC7B	Hs00406077_m1	226152_at	-0.52431	6.62E-05

GUCY1B3	Hs00168336_m1	203817_at	-0.52258	7.07E-05
MMD	Hs00948031_m1	203414_at	-0.39744	0.003528
FRMD3	Hs00604157_m1	230645_at	-0.21862	0.119455
SDPR	Hs00190538_m1	222717_at	-0.18796	0.182072
CPNE5	Hs00326218_m1	227189_at	-0.07961	0.574802
CCDC90A	Hs00254417_m1	227451_s_at	-0.06118	0.666593
CDC14B	Hs00269351_m1	221556_at	-0.04119	0.771874
*coefficient and p-value correspond to Pearson correlation test between RMA normalized microarray expression for a given probe set and delta Ct (where higher values represent lower transcript abundance) for the corresponding RT-PCR assay.				

Table 3. Concordance of correlations with PFS between microarray and RT-PCR based gene expression for Factor 14 genes in HV2 cohort			
Gene Name	Affymetrix probe set ID	RT-PCR based Correlation coefficient*	Microarray based Correlation coefficient*
CTDSPL	201906_s_at	0.34	-0.38
FSTL1	208782_at	0.18	-0.38
ITGA2B	216956_s_at	0.22	-0.37
TREML1	1555659_a_at	0.35	-0.35
SPARC	212667_at	0.23	-0.31
ITGA2B	206494_s_at	0.22	-0.30
ITGA2B	206493_at	0.22	-0.30
MPL	207550_at	0.30	-0.29
CTTN	201059_at	0.24	-0.29
CMTM5	230942_at	0.24	-0.29
SELP	206049_at	0.19	-0.28
CLU	208791_at	0.23	-0.27
GP1BB	206655_s_at	0.27	-0.27
ITGB3	204628_s_at	0.30	-0.26
TGFB1I1	209651_at	0.20	-0.25
ALOX12	207206_s_at	0.19	-0.24
PBX1	212151_at	0.26	-0.24
CPNE5	227189_at	0.33	-0.24
MYL9	201058_s_at	0.31	-0.24
CLU	208792_s_at	0.23	-0.22
PPBP	214146_s_at	0.24	-0.22
SH3BGRL2	225354_s_at	0.22	-0.20
PF4	206390_x_at	0.30	-0.20
THBS1	201108_s_at	0.13	-0.20
SLC24A3	57588_at	0.06	-0.20
PBX1	212148_at	0.26	-0.19
SPARC	200665_s_at	0.23	-0.18
GNG11	204115_at	0.24	-0.16
TUBB1	208601_s_at	0.16	-0.16
PCSK6	207414_s_at	0.28	-0.15
SLC24A3	219090_at	0.06	-0.15
PDE5A	227088_at	0.20	-0.14
GUCY1B3	203817_at	0.17	-0.12
CALD1	212077_at	0.13	-0.11
PRKAR2B	203680_at	0.23	-0.10
TTC7B	226152_at	0.00	-0.10
ARHGAP6	206167_s_at	0.20	-0.10
TUBB1	230690_at	0.16	-0.09

ELOVL7	227180_at	0.20	-0.08
MYLK	224823_at	0.22	-0.05
MMD	203414_at	0.21	-0.05
SDPR	222717_at	0.21	-0.05
FRMD3	230645_at	-0.12	0.00
CDC14B	221556_at	0.13	0.18
CCDC90A	227451_s_at	-0.14	0.28

*Correlation coefficients represent Pearson correlations between gene expression based on microarray or RTPCR for Factor 14 transcripts and post-aspirin platelet function score (PFS) in HV2 cohort.

Table 4. Comparison of platelet and megakaryocyte genes with Factor 14

Gene set description*	Number of features	Number of overlapping genes with Factor 14***	GSEA NES	GSEA P-value	GSEA Q-value**	Reference
Factor 14 genes (reference)	62	62	2.23	<0.0001	<0.0001	N/A
Platelet genes (A)	248	25	2.18	<0.0001	<0.0001	[1]
Platelets specific genes (B)	36	12	2.12	<0.0001	0.001	[2]
Megakaryocytes genes (C)	261	20	2.06	<0.0001	0.001	[3]
Platelet genes (D)	196	14	1.97	<0.0001	0.008	[4]
Platelet proteins (E)	99	31	1.84	<0.0001	0.01	[5]
Platelet genes (F)	36	6	1.73	0.004	0.02	[6]

GSEA = Gene Set Enrichment Analysis (GSEA); NES = Normalized enrichment score; *The gene lists were as follows: A) the abundant platelet transcripts identified by RNA-sequencing[1]; B) platelet-specific transcripts identified in platelets of sickle cell disease patients or controls[2]; C) megakaryocyte-specific genes[3]; D) abundant platelet genes identified in platelets of patients with systemic lupus or controls[4]; or E) platelet proteins from healthy donors[5]; F) the top 50 genes identified in purified platelets from healthy volunteers[6]; ** refers to false discovery rate q-value;***p ≤ 0.001 for all overlaps.

Table 5. Comparison of Factor 14 genes with non-platelet derived peripheral blood cell-type genes*.

Gene set description*	Number of features	Number of overlapping genes with Factor 14*	GSEA NES	GSEA P-value**	GSEA Q-value***
CD4+ Th lymphocytes	36	0	-0.7	0.7	0.7
CD8+ Tc lymphocytes	4	0	0.7	0.7	0.8
CD14+ Monocytes	164	0	0.6	0.8	0.9
CD19+ B lymphocytes	53		0.5	1.0	0.9
CD56+ NK cells	605	0	1.4	0.2	0.2
CD66+ granulocytes	257	0	1.7	0.03	0.03
Erythroblasts	38	0	1.3	0.2	0.2

*Cell-type specific gene lists were obtained from Watkins et al[3]; Abbreviations and footnotes are per Table 2.

Table 6. Baseline characteristics in CATHGEN cohorts and their association with Factor 14 levels			
Variable	Case:Control Cohort (n = 397)	Observational cohort (n = 190)	P-value for association with Factor 14*
Age (years, median, [IQR])	65 [57 -73]	56 [47 -67]	0.15
Hypertension (%)	70.3%	70.5%	0.23
CAD	26%	78%	0.47
Diabetes (%)	34.3%	23.2%	0.14
RACE (%) White African-American Other	74.3% 21.7% 4.0%	66.8% 26.8% 6.3%	Reference 2.13e-06 0.53
Female (%)	32.2%	41.1%	7.37e-06
Smoking history (%)	47.9%	53.7%	0.06
Hyperlipidemia (%)	61.0%	61.6%	0.37
*The combined Case:Control and Observational cohorts were used to test for associations with Factor 14 levels, while controlling for cohort using linear regression.			

Table 7. Factor 14 individual probe sets are associated with death or myocardial infarction in two independent datasets.

Probe set ID	Gene name	Odds Ratio*	Odds Ratio p-value	Hazard Ratio*	Hazard Ratio p-value
Factor 14	n/a	1.21	0.028036541	1.47	0.001454
Factor 14 individual probe set associations					
201058_s_at	MYL9	1.30	0.00400748	1.75	7.65E-06
206494_s_at	ITGA2B	1.37	0.00351754	1.83	1.88E-05
216956_s_at	ITGA2B	1.52	0.000531982	1.82	3.09E-05
230942_at	CMTM5	1.31	0.066298986	2.52	3.50E-05
1555659_a_at	TREML1	1.36	0.010689841	1.89	5.63E-05
212573_at	ENDOD1	1.94	4.95E-05	2.10	0.000188
212148_at	PBX1	1.46	2.45E-06	1.47	0.0002
226188_at	LGALSL	1.38	0.071992059	2.62	0.000206
212151_at	PBX1	1.65	1.48E-05	1.73	0.000281
206493_at	ITGA2B	1.46	0.002887947	1.83	0.000351
210986_s_at	TPM1	1.87	1.45E-05	1.83	0.000422
57588_at	SLC24A3	1.15	0.440525089	2.19	0.000472
209651_at	TGFB1I1	1.32	0.015158949	1.63	0.000616
203819_s_at	IGF2BP3	1.45	0.033315085	1.86	0.000619
206049_at	SELP	1.46	0.021063226	2.01	0.000636
1552773_at	CLEC4D	1.48	0.00175683	1.74	0.000708
226152_at	TTC7B	1.43	0.029162426	2.06	0.000931
204628_s_at	ITGB3	1.41	0.059389555	1.81	0.00112
208791_at	CLU	1.56	0.002006075	1.84	0.001166
206655_s_at	GP1BB	1.32	0.043694557	1.69	0.001235
210387_at	HIST1H2BG	1.74	0.000124728	1.98	0.001269
201108_s_at	THBS1	1.39	0.02328534	1.87	0.001633
208792_s_at	CLU	1.51	0.003791609	1.80	0.001869
219090_at	SLC24A3	1.16	0.373804241	2.01	0.001907

225974_at	TMEM64	0.59	0.008235562	0.41	0.002877
207206_s_at	ALOX12	1.21	0.138154505	1.74	0.003564
229778_at	C12orf39	1.39	0.004259803	1.58	0.005073
215779_s_at	HIST1H2BG	1.68	0.002837678	1.81	0.005208
37966_at	PARVB	1.04	0.771921449	1.81	0.005396
225354_s_at	SH3BGRL2	1.34	0.049135136	1.79	0.00551
212667_at	SPARC	1.25	0.167170138	1.78	0.005604
207414_s_at	PCSK6	1.23	0.150039721	1.72	0.006252
235331_x_at	PCGF5	1.93	5.27E-06	1.42	0.006562
200665_s_at	SPARC	1.44	0.008859476	1.72	0.006808
207808_s_at	PROS1	1.45	0.018336331	1.58	0.006933
212077_at	CALD1	1.11	0.338890694	1.52	0.006972
201906_s_at	CTDSPL	1.30	0.11440241	1.90	0.007299
241133_at	TRBV27	1.28	0.019888406	1.52	0.011429
206390_x_at	PF4	1.13	0.398156787	1.85	0.0117
211026_s_at	MGLL	1.16	0.327461963	1.85	0.012801
230645_at	FRMD3	1.85	0.000131807	1.79	0.014483
206110_at	HIST1H3H	1.72	3.26E-05	1.57	0.018138
208601_s_at	TUBB1	1.16	0.201369942	1.50	0.01831
203680_at	PRKAR2B	1.43	0.005278287	1.54	0.025644
227180_at	ELOVL7	1.25	0.070697319	1.49	0.032551
214146_s_at	PPBP	1.55	0.000423437	1.83	0.037707
227189_at	CPNE5	0.71	0.079532372	1.63	0.03795
208782_at	FSTL1	1.28	0.086692126	1.53	0.039651
1560262_at	LRRC32	1.24	0.155422462	1.55	0.050918
201059_at	CTTN	1.16	0.299273772	1.53	0.057153
220496_at	CLEC1B	1.50	0.002955397	1.43	0.079647
212651_at	RHOBTB1	1.32	0.08218736	1.43	0.096564
204115_at	GNG11	1.41	0.004779687	1.45	0.118268
206167_s_at	ARHGAP6	1.14	0.422952376	1.36	0.151355

227088_at	*PDE5A*	1.35	0.046937627	1.37	0.151501
203817_at	*GUCY1B3*	1.28	0.098280545	1.35	0.160362
207550_at	*MPL*	1.05	0.717720328	1.28	0.210728
1553842_at	*BEND2*	1.42	0.079659139	1.33	0.233
212813_at	*JAM3*	0.97	0.872083306	1.34	0.240894
230690_at	*TUBB1*	1.09	0.589122948	1.31	0.265404
207156_at	*HIST1H2AG*	1.54	0.004128653	1.27	0.303256
225166_at	*ARHGAP18*	1.53	0.034773026	1.30	0.338505
224823_at	*MYLK*	1.24	0.099105304	1.21	0.357528
227451_s_at	*CCDC90A*	1.07	0.751398281	0.76	0.381367
228708_at	*RAB27B*	1.21	0.257379973	0.84	0.546707
207815_at	*PF4V1*	1.07	0.397575492	1.07	0.581265
221556_at	*CDC14B*	0.78	0.197214947	0.86	0.60266
222717_at	*SDPR*	1.08	0.545883997	1.10	0.679631
206272_at	*RAB4A*	1.54	0.020878373	0.91	0.803266
214974_x_at	*CXCL5*	0.88	0.201323204	0.97	0.867518
205442_at	*MFAP3L*	0.90	0.409405359	1.01	0.968416
203414_at	*MMD*	1.19	0.292053865	1.01	0.985414

*Individual Factor 14 probe sets were associated with death or myocardial infarction in the case control and observational cohorts, yielding odds ratios and hazards ratios, respectively. For each probe set analyses were adjusted for log(platelet count), race, and sex.

Table 8. Net Reclassification Table for combined CATHGEN cohorts

	10%	10-20%	>20%	% Reclassified
Factor 14				
10%	10	4	0	29
10-20%	4	55	6	15
>20%	0	14	474	3
			Net Reclassification Index	4.6% (p = 0 007)
ITGA2B				
	10%	10-20%	>20%	% Reclassified
10%	8	5	1	43
10-20%	13	42	10	35
>20%	1	24	463	5
			Net Reclassification Index	7.5% (p = 0.002)