

Downloaded from UvA-DARE, the institutional repository of the University of Amsterdam (UvA)
<http://dare.uva.nl/document/97933>

File ID 97933
Filename Chapter 7: Comparative gene expression profiling of in vitro differentiated megakaryocytes and erythroblasts identifies novel activatory and inhibitory platelet membrane proteins

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type Dissertation
Title From stem cell to platelet
Author M.R. Tijssen
Faculty Faculty of Medicine
Year 2008
Pages 151
ISBN 9789090226293

FULL BIBLIOGRAPHIC DETAILS:

<http://dare.uva.nl/record/261783>

Copyright

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use.

Chapter 7

Comparative gene expression profiling of *in vitro* differentiated megakaryocytes and erythroblasts identifies novel activatory and inhibitory platelet membrane proteins

I.C. Macaulay¹, M.R. Tijssen², D.C. Thijssen-Timmer², A. Gusnanto³, M. Steward⁴, P. Burns¹, C.F. Langford⁵, P. Ellis⁵, F. Dudbridge³, J.J. Zwaginga^{2,6}, N.A. Watkins¹, C.E. van der Schoot^{2,7}, and W.H. Ouwehand¹ on behalf of the Bloodomics consortium

ICM and MRT contributed equally to this work

¹Department of Haematology, University of Cambridge and National Blood Service, Cambridge, UK

²Department of Experimental Immunohematology, Sanquin Research and Landsteiner Laboratory, Academic Medical Centre, University of Amsterdam, The Netherlands

³MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK

⁴Domantis Limited, Cambridge, UK

⁵Microarray Facility, Wellcome Trust Sanger Institute, Cambridge, UK

⁶Department of Immunohematology-Blood transfusion, Leiden University Medical Centre, Leiden, The Netherlands

⁷Department of Hematology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands.

Abstract

To identify previously unknown platelet receptors we compared the transcriptomes of *in vitro* differentiated megakaryocytes (MKs) and erythroblasts (EBs). RNA was obtained from purified, biologically paired MK and EB cultures and compared using cDNA microarrays. Bioinformatical analysis of MK-upregulated-genes identified 151 transcripts encoding transmembrane domain-containing proteins. Although many of these were known platelet genes, a number of previously unidentified or poorly characterized transcripts were also detected. Many of these transcripts, including G6b, G6f, LRRC32, LAT2, and the G protein-coupled receptor SUCNR1, encode proteins with structural features or functions that suggest they may be involved in the modulation of platelet function. Immunoblotting on platelets confirmed the presence of the encoded proteins, and flow cytometric analysis confirmed the expression of G6b, G6f, and LRRC32 on the surface of platelets. Through comparative analysis of expression in platelets and other blood cells we demonstrated that G6b, G6f, and LRRC32 are restricted to the platelet lineage, whereas LAT2 and SUCNR1 were also detected in other blood cells. The identification of the succinate receptor SUCNR1 in platelets is of particular interest, because physiologically relevant concentrations of succinate were shown to potentiate the effect of low doses of a variety of platelet agonists.

Introduction

Megakaryocytes (MKs) and erythroblasts (EBs) share a common, bipotent progenitor cell when differentiating from hematopoietic progenitor cells (HPCs).¹ The close association between the differentiation pathways of the 2 cell types is evidenced by a number of observations. Firstly, there is significant overlap of transcription factor dependency (eg GATA1,² GATA2,³ NFE1,⁴ NFE2,⁵ and FOG1⁶) between the 2 cell types, suggesting the presence of common transcriptional networks and concomitant gene expression. The 2 cell types have also been shown to have a bipotent progenitor (BFU-E/MK) in methylcellulose and single cell cultures of CD34⁺CD38^{lo} hematopoietic stem cells⁷ and in the bone marrow of normal mice.⁸ The common origin of MKs and EBs suggests that restriction towards either lineage is a late event in HPC differentiation. Comparative transcriptome analysis of the MK and EB lineages thus gives an insight into the mechanisms of cellular differentiation as well as allowing the identification of novel genes that are expressed following terminal lineage commitment of the MK. The study of the MK transcriptome complements the numerous ongoing gene discovery studies of the platelet transcriptome and proteome⁹ and, furthermore, may allow the identification of transcripts and proteins that are not readily detectable in platelets using transcriptomic or proteomic approaches.

Here, we report the gene expression profiles of MKs and EBs differentiated *in vitro* from cord blood-derived HPCs. We compared 5 biologically paired samples (in which the MKs and EBs were differentiated from the same CD34⁺ HPC harvest) in an experimental design that included extensive technical replication and dye-swaps to minimize false discovery. Statistical analysis of this data identified both well-characterized and novel genes that are differentially expressed between the 2 lineages. The aim of this comparative study was primarily to identify genes encoding novel transmembrane proteins that may act as modulators of platelet function (eg,

immunoglobulin superfamily [Ig-SF] members, G protein-coupled receptors [GPCRs], tetraspanins, adaptor molecules, and leucine-rich repeat proteins [LRRs]). Using both purpose-made polyclonal antisera and commercially available antibodies, we confirmed the expression of these novel proteins in platelets. Analysis of transcript and protein expression in mature blood cells shows that some of these proteins are restricted to the MK and platelet lineage. Functional characterization of SUCNR1, a GPCR with succinate as its specific ligand,¹⁰ suggests that the expression of SUCNR1 mRNA and protein in MKs and platelets adds a novel and physiologically relevant agonist pathway in platelet activation.

Materials and methods

Umbilical cord blood was collected after informed consent according to the guidelines of Eurocord Nederland. Mononuclear cells were isolated by density gradient centrifugation over Ficoll (1.077 g/cm³; Pharmacia Biotech, Uppsala, Sweden). CD34⁺ cells were isolated by magnetic cell sorting (91-98% purity; VarioMACS system; Miltenyi Biotec, Gladbach, Germany) according to the manufacturer's instructions and cryopreserved until required. After thawing, cells were washed in Iscove Modified Dulbecco Medium (BioWhittaker, Berkshire, United Kingdom) with 0.2% (wt/vol) human serum albumin (Sanquin Reagents, Amsterdam, The Netherlands) and cultured in Stemspan expansion medium (StemCell Technologies, Vancouver, Canada). To generate MKs, CD34⁺ cells (1 x 10⁵/mL) were cultured for 7 days in serum-free medium supplemented with rhTpo (10 ng/mL, Strathmann Biotec AG, Hamburg, Germany) and interleukin-1 β (IL-1 β ; 10 ng/mL, Strathmann Biotec AG). To generate EBs, CD34⁺ cells (1 x 10³/mL) were cultured for 10 days in the presence of erythropoietin (3 U/ml; R&D Systems, Minneapolis, MN, USA), interleukin-3 (IL-3; 10 ng/mL; Strathmann Biotec AG) and stem cell factor (100 ng/mL, R&D Systems).

Flow cytometry on MKs and antibodies

The following primary antibodies were used for cellular phenotyping: FITC-conjugated monoclonal IgG1 isotype control, anti-CD41 and anti-CD42b (Sanquin Reagents); phycoerythrin (PE)-conjugated IgG1 isotype control (BD Biosciences, San Jose, CA, USA); anti-CD34 (Immunotech, Marseille, France); anti-CD41 and anti-CD235a (DakoCytomation, Cambridgeshire, United Kingdom).

In vitro differentiated MKs were fixed with 1% (wt/vol) paraformaldehyde for 10 minutes. Cells were centrifuged and resuspended in phosphate buffered saline (PBS) containing 0.2% (wt/vol) bovine serum albumin (BSA) and 5 mM ethylenediamine tetraacetic acid (EDTA). Cells were incubated with primary antibodies for 30 minutes at 4°C, with isotype-matched mouse IgG subtypes as controls. After 30 minutes, the cells were washed with PBS/BSA/EDTA and analyzed on a FACScan or LSRII (Becton Dickinson, San Jose, CA, USA).

Sorting of MKs and EBs

To ensure a high level of purity prior to transcriptome analysis, both MKs and EBs were flow-sorted following culturing. MKs were stained with anti-CD41-FITC, EBs with anti-CD41-PE and anti-CD235-FITC. Both cell types were sorted on a MoFlo cell sorter (DakoCytomation) to obtain highly purified populations of CD41⁺ MKs and CD235a⁺/CD41⁻ EBs.

Isolation of RNA and hybridization onto cDNA microarrays

Total RNA was extracted from purified cells using Trizol (Invitrogen, Paisley, United Kingdom), according to the manufacturer's instructions. Extracted RNA was resuspended in RNase free, diethylpyrocarbonate-treated water (Ambion, Huntingdon, United Kingdom) and quantified using a NanoDrop spectrophotometer (Nanodrop Technologies, Wilmington, DE, USA). RNA was amplified using an adaptation of the SMART cDNA synthesis kit (BD Clontech, Oxford, United Kingdom), which amplifies RNA by a template switching-polymerase chain reaction (TS-PCR) process.¹¹ Total RNA (100 ng) was mixed with 10 pmol each of SMART CDS 3' Primer IIA (5'-AAGCAGTGGTATCAACGCAGAGTACT₃₀VN-3') and TS primer (5'-dAAGCAGTGGTATCAACGCAGAGTACGC)r(GGG)-3'). The reaction was denatured at 72°C for 2 minutes and cooled immediately. RNA was reverse transcribed by the addition of 2 µL first-strand buffer (5 X), 1 µL DTT (20 mM), 1 µL dNTPs (10 mM), and 1 µL PowerScript Reverse Transcriptase, followed by incubation at 42°C for 1 hour. The RT reaction (5 µL) was used as template for global amplification of the cDNA by PCR using Advantage II PCR reagents and the SMART PCR primer (5'-AAGCAGTGGTATCAACGCAGAGT-3') for 17 cycles of PCR consisting of 95°C for 5 seconds, 65°C for 5 seconds, and 68°C for 6 minutes. Following amplification the product was purified using a PCR clean-up kit (Qiagen, Crawley, United Kingdom). Of the amplified TS-PCR product, 250 ng was labeled by incorporation of Cy3 or Cy5 dCTP using the Bioprime Labelling Kit (Invitrogen), and the labeled products were purified on Autoseq G-50 columns (GE Healthcare, Buckinghamshire, United Kingdom).

Differentially labeled biologically paired MK and EB samples were pooled and hybridized to the Hver 2.1.1 cDNA microarrays (Wellcome Trust Sanger Institute; WTSI). Hybridization and washing of the arrays were performed according to WTSI protocols (<http://www.sanger.ac.uk/Projects/Microarrays/>). Following hybridization, the slides were promptly scanned at 10-µm resolution on an Agilent Microarray Scanner (G2505B; Agilent Technologies, Stockport, United Kingdom). The images generated were exported into GenePix version 4.1 (Molecular Devices, Sunnyvale, CA, USA) for spotfinding and feature extraction.

In total, 20 cDNA microarrays were hybridized, representing 5 biological replicates each with 4 technical replicates. To control for labeling bias, 2 dye-swaps were performed per pairing.

Statistical analysis of cDNA microarray data

The Limma package¹² from Bioconductor (<http://www.bioconductor.org>) was used to process the expression data. Data were first background corrected and then, in addition to automated and manual flagging of features in GenePix, a further automated quality control step¹³ was applied to remove poor quality features. A Loess normalization was applied to remove systematic variability, including possible dye bias.^{14,15,16}

To identify differentially expressed transcripts, we used a linear model similar to those previously developed.^{17,12} Following the discussion of Repsilber *et al.*¹⁸ and Dobbin *et al.*,¹⁴ we introduced a new term to take into account the effect of 5 biological replicates in our data and regard it as sample effect. The model is fitted using the lme4 package¹⁹ in R (R Development Core Team, <http://cran.R-project.org>). Differential expression of each feature was assessed by testing the null hypothesis that the corresponding parameter in the model is zero. The *P* value obtained from this

test is then corrected for control of false discovery rate (fdr).²⁰ Differentially expressed genes were identified as those with an average fold difference of 2 or more and an fdr-corrected *P* value no greater than .05.

Confirmation of transcript expression by quantitative RT-PCR

The differential expression of selected transcripts observed in the microarray experiments was validated by quantitative reverse transcriptase (RT)-PCR. TaqMan mRNA specific assays (Applied Biosystems, Warrington, United Kingdom) were used to measure the relative transcript abundance in MKs and EBs. Transcript levels were also measured in RNA from highly purified populations of platelets, CD4⁺, CD8⁺, CD14⁺, CD16⁺, and CD19⁺ peripheral blood cells. Platelets were purified by 3 rounds of centrifugation followed by negative selection with anti-CD45 magnetic beads (Invitrogen) to remove contaminating leucocytes. All other blood cell populations were purified by positive selection from Ficoll-separated peripheral blood mononuclear cells using antibody-conjugated magnetic microbeads (Miltenyi Biotech).

The expression of the following transcripts was examined: leukocyte specific transcript (*LST1*, Assay ID HS00394683_m1), tetraspannin NET-7 (*TSPAN15*, Hs00202548_m1), G6b (Hs00228935_m1), linker for activation of T-cells 2 (*LAT2*, Hs00247916_m1), succinate receptor 1 (*SUCNR1*, HS00263701_m1), G6f (Hs00222013_m1), pan-hematopoietic expression (*PHEMX*, Hs00185077_m1), LRRC32 (Hs00194136_m1), CD61 (*ITGB3*, Hs00173978_m1), CD41 (*ITGA2B*, Hs00166246_m1), glycoprotein 6 (*GP6*, Hs00212574_m1), CD36 (Hs00169627_m1), CD42a (*GP9*, Hs00236828_m1), CD110 (*MPL*, Hs00180489_m1), CD235a (Hs00266777_m1), and GAPDH (Hs99999905_m1).

RNA (100 ng) was reverse transcribed using the TaqMan RT kit (Applied Biosystems), from which 0.5 ng of cDNA was then used as template for real-time PCR following the manufacturer's instructions (Applied Biosystems). Reactions were incubated at 50°C for 2 minutes then 95°C for 10 minutes and real-time PCR reactions were performed over 40 cycles (95°C for 15 seconds, 60°C for 1 minute) on an MX-4000 (Stratagene, La Jolla, CA, USA). Threshold values (Ct) were normalized to GAPDH to allow comparison between samples.

G6b splice variants in platelets and MKs

Splice variant specific primers were designed to identify the predominant isoforms of G6b in platelets and MKs. Briefly, 1 ng platelet or MK cDNA was amplified by 40 cycles of PCR (94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute). The forward primer ATGGCTGTGTTTCTGCAG was used in all reactions in combination with the following reverse primers: G6b-A/C/G (GACAGAGCAAATCTAGGGA, expected product size 548 bp [base pair], 457 bp, and 548 bp respectively), G6b-B (AAGTGGAGCAAATCTAGGG, 549-bp product), G6b-D (GACAGAGCATGGGTAGG, 416-bp product), G6-E (GTTTTCAAGTGGAGCATG, 425-bp product), G6b-F (GACAGAGCTGAAGGATGG, 666-bp product), and G6b-G (CCTTACCGGTTCTGG, 646-bp product). All primers were obtained from Sigma Genosys (Cambridge, United Kingdom).

Generation of polyclonal antisera against novel platelet proteins

Polyclonal antisera were generated against *Escherichia coli*-produced recombinant proteins in mice. cDNAs corresponding to the extracellular domains of G6b (NM_

025260, corresponding to amino acid residues 18-142), G6f (NM_001003693, amino acid residues 17-234), and LRRC32 (NM_005512, amino acids 29-296) were expressed as N-terminal 6 X His-tagged proteins, which were subsequently refolded and purified from inclusion bodies by Ni-NTA affinity chromatography.

Female Balb/C mice, 6 to 8 weeks old, were immunized to generate polyclonal immune sera. Protein (50 µg/mouse) was formulated with TiterMax Gold (Strattech Scientific, Soham, Cambridgeshire, United Kingdom), and mice were immunized with 200 µL (subcutaneous injection) on days 0 and 14. Subsequent immunizations (50 µg in 200 µL PBS, days 28 and 42) were delivered intraperitoneally. Test bleeds (day 49) were assessed for reactivity and specificity by enzyme-linked immunosorbent assay, FACS, and Western blot, after which mice were given an intravenous boost of 25 µg protein in 100 µL PBS, followed by sacrifice (+4 days) and terminal bleed. Serum from the terminal bleed was collected and used in assays without further purification.

Glycosidase treatment of platelet lysates

Washed platelets were prepared in 10 mM phosphate buffer, pH 6.5, 0.1% SDS, and 50 mM β-mercaptoethanol. Lysates from approximately 2×10^8 platelets were treated with *N*-glycosidase F or *O*-glycosidase as previously described.²¹ Reactions were then stopped by addition of an equal volume of 2 x Laemmli buffer containing 10% β-mercaptoethanol and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Protein detection by Western blot

Mouse monoclonal anti-LAT2 (Clone NAP-07) and rabbit polyclonal anti-SUCNR1 were obtained from Abcam (Cambridge, United Kingdom). The SUCNR1 antibody is a polyclonal rabbit antibody raised against a peptide from the second cytoplasmic loop of the SUCNR1 protein (KYPFREHLLQKKE). Washed platelets (1×10^7 /lane) or purified blood cells (0.5×10^6 cells/lane) were lysed in 2 x Laemmli buffer, and proteins separated on a 10% Bis-Tris Gel (Invitrogen). Blotting was performed using the Chemiluminescent WesternBreeze kit following the manufacturer's protocol (Invitrogen). Primary antibodies were diluted 1:1000 for LRRC32, G6b, and G6f mouse polyclonals and 1:5000 for the LAT2 mouse monoclonal and SUCNR1 rabbit polyclonal.

Flow cytometry on platelets

Platelets were fixed for 15 minutes at room temperature in 10% formaldehyde (DakoCytomation), washed in PBS/EDTA/BSA and incubated for 40 minutes with 1 µL antiserum or primary antibody, washed, and detected using the Alexa Fluor-488 signal amplification kit (Invitrogen) according to the manufacturer's instructions. Matched preimmune sera or unlabeled IgG1 were used as a negative control where appropriate. Where the primary antibody was raised against an intracellular antigen (LAT2 and SUCNR1), cells were fixed and permeabilized prior to staining using an Intrastain kit (DakoCytomation) according to the manufacturer's instructions. Samples were analyzed on a Coulter EPICS XL-MCL flow cytometer.

Platelet aggregometry

Whole blood was obtained from blood donors in 3.2% citrate Vacuette tubes (Grenier Bio-one, Gloucester, United Kingdom). Platelet-rich plasma was prepared by

centrifugation at 150g for 15 minutes, after which the platelet count was adjusted to $250 \times 10^9/\text{mL}$ using platelet-poor plasma. Platelet aggregation in response to a suboptimal dose of ADP (1 μM ; Sigma-Aldrich, Irvine, United Kingdom), thrombin receptor-activating peptide (TRAP-6; Bachem, Torrance, CA, USA) and cross-linked collagen-related peptide (CRP-XL) in the presence of a range of concentrations of succinate in HBS pH 7.4 (50-1000 μM , Sigma-Aldrich) was analyzed using a PAP-4D Aggregometer (Bio-data, Horsham, PA, USA).

Results

Differentiation of CD34⁺ HPCs

CD34⁺ cells were isolated from a total of 5 cord blood donations, and paired MK and EB cultures performed. In the MK cultures 53.1% (range, 31.0%-63.4%) of the cells expressed CD41 (further details of this culture system are summarized in van den Oudenrijn *et al.*²²). Morphologic analysis of the cells shows that the MKs are predominantly at the megakaryoblast and pro-MK stage of differentiation (Figure 1A), with some larger granular MKs present (Figure 1B). Ploidy and immunohistochemical analysis shows 37.5% of the cells are 4N or higher and CD42 is expressed on the cell surface (data not shown).²³ The EBs are at the pro-erythroblast stage based on the size of the cells (Figure 1C) and Glycophorin A (CD235a) expression.^{24,25} In the EB cultures 66.2% (range, 60.4%-70.2%) of the cells expressed CD235a. The differentiated cells were purified as described (Figure 1D-E), with cell purity prior to RNA isolation being $95.4\% \pm 2.4\%$ for CD41⁺ (92.4% – 99.3%) and $93.7\% \pm 7.5\%$ for CD235a⁺ (80.8% – 98.8%). TaqMan real-time PCR was used to confirm differential expression of known lineage associated transcripts in the purified cell populations following RNA isolation (Figure 1F).

cDNA microarray analysis

Comparative transcriptome analysis of the MK and EB samples was performed using the Sanger Hver 2.1.1 15K cDNA microarray. This microarray comprises 15 552 features representing approximately 12 000 unique transcripts. We used a biologically paired experimental design, with each MK/EB culture pair being compared directly. A total of 20 hybridizations were performed representing 4 technical replicates per biological pair. The data were analyzed as described to identify differentially expressed features. In total, 596 features (516 unique annotated transcripts) were up-regulated in MKs relative to EBs (fold change ≥ 2 ; $P < .05$), a selection of those encoding transmembrane proteins is shown in Table 1. Another 338 features (298 unique annotated transcripts) were up-regulated in EBs relative to MKs (Table 2). Complete lists of present genes, as well as all differentially expressed genes are available (Table S1, available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

Identification of novel MK plasma membrane proteins

The aim of this work was to identify novel or previously uncharacterized components of the MK and platelet plasma membranes. The list of genes upregulated in MKs relative to EBs was interrogated using the Ensembl Biomart tool^{25,26} to identify encoded proteins with transmembrane domains which identified 151 such transcripts (Table S2).

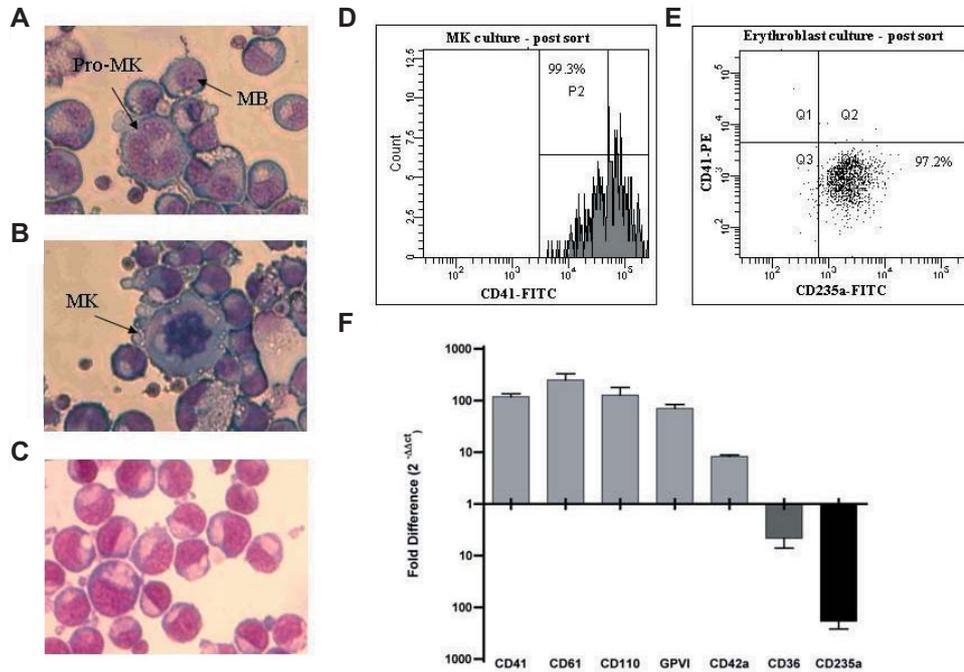


Figure 1: Characterization of MK and EB cultures

A-B. Megakaryocyte morphology, the MK cultures were predominantly in the megakaryoblast (MB) or promegakaryocyte (pro-MK; horse-shoe-shaped nucleus) stage of differentiation, although some large granular MKs (multilobed nucleus) were observed. **C.** Erythroblasts (EBs) were in the proerythroblast stage of differentiation. **D.** Flow cytometric analysis of MKs and EBs after sorting. After FACS sorting, 99.3% of the cells were positive for CD41. The parallel EB culture (**E.**) was sorted to a purity of 97.2% (based on CD235a expression). **F.** Confirmation of differential expression of lineage-associated markers by TaqMan real-time PCR. Data presented as fold difference in expression (corrected for GAPDH) between paired MK and EB cultures. Error bars represent the SEM of biologic (n=5) and technical (n=3) replicates combined. Values above the x-axis indicate genes up-regulated in MKs, whereas values below indicate genes up-regulated in EBs. Cells were stained using modified May-Grünwald eosin-methylene blue solution (Merck) and 0.5 mM KH_2PO_4 , pH 7.4. Images were acquired using a Leica DM 1L microscope (Leica, Mannheim, Germany) equipped with a Leica objective C, plan L 40X/0.50 numerical aperture PH2 with a total magnification of 400X. Images were captured using a Leica DC300 digital camera and Leica imaging software (IM500, version 1.20).

Gene ID	Gene Name	Fold difference, MK/EB
Integrins		
<i>ITGB3</i>	Platelet membrane glycoprotein IIIa	43.3
<i>ITGA2B</i>	Platelet membrane glycoprotein IIb	31.5
<i>ITGA6</i>	Integrin α -6 precursor	6.5
Immunoglobulin superfamily members		
<i>PECAM1</i>	Platelet endothelial cell adhesion molecule	47.3
<i>CD84</i>	CD84 antigen	12.3
<i>c6orf25</i>	G6B protein	10.4
<i>VSIG2</i>	V-set and immunoglobulin domain containing 2	10.1
<i>C6orf21</i>	G6F protein	9.2
<i>FCER1A</i>	High affinity immunoglobulin ϵ receptor α -subunit precursor	6.3
G protein-coupled receptors		
<i>SUCNR1</i>	G protein-coupled receptor 91 (GPR91)	7.2
<i>GPR171</i>	Probable G protein-coupled receptor H963	2.7
<i>GPR56</i>	G protein-coupled receptor 56	2.0
Tetraspannins		
<i>CD9</i>	CD9 antigen	23.9
<i>TSPAN15</i>	Tetraspan NET-7	19.0
<i>PHEMX</i>	Phemx protein	3.3
<i>TSPAN18</i>	Tetraspanin similar to uroplakin 1	2.8
<i>TSN9</i>	Tetraspan NET-5	2.6
Adaptor molecules		
<i>FCER1G</i>	High-affinity immunoglobulin ϵ receptor γ -subunit	16.0
<i>LAT2</i>	WBSCR5 protein, LAT2	13.1
<i>LAT</i>	Linker for activation of T cells	8.0
Leucine-rich repeat proteins		
<i>GP9</i>	Platelet glycoprotein IX	48.6
<i>GP1BA</i>	Platelet glycoprotein Ib α	23.1
<i>RTN4R</i>	Reticulon 4	3.4
<i>TLR9</i>	Toll-like receptor 9	2.3
Other Transmembrane domain-containing proteins		
<i>PF4</i>	Platelet factor 4	111.3
<i>SELP</i>	P-selectin	71.0
<i>PPBP</i>	Platelet basic protein	62.7
<i>TMEM40</i>	Transmembrane protein 40	19.8
<i>LST1</i>	Leukocyte-specific transcript 1 protein (B144 protein)	19.7
<i>GJA4</i>	Gap junction α -4	15.1
<i>PLXDC2</i>	Plexin domain-containing 2	13.1
<i>CAV1</i>	Caveolin-1	10.3
<i>MEST</i>	Mesoderm-specific transcript	8.6
<i>EPHNB1</i>	Ephrin-B1	8.2
<i>SLC44A1</i>	CDW92 antigen	6.6
<i>HG2A</i>	CD74 antigen	6.3
<i>CD40</i>	Tumor necrosis factor receptor superfamily member 5, CD40L receptor	6.0
<i>CCL5</i>	Small inducible cytokine A5	5.6
<i>ANPEP</i>	Aminopeptidase N (Myeloid plasma membrane glycoprotein CD13)	5.6
<i>CAV2</i>	Caveolin-2	5.5
<i>CD69</i>	CD69 antigen	4.7
<i>NINJ2</i>	Ninjurin 2	4.3
<i>LY6G5C</i>	G5C protein	3.7
<i>TNFRSF14</i>	Tumor necrosis factor receptor superfamily member 14	3.4
<i>PTAFR</i>	Platelet-activating factor receptor	2.9
<i>IL9R</i>	Interleukin-9 receptor	2.8
<i>ICAM2</i>	Intercellular adhesion molecule-2 (CD102 antigen)	2.6
<i>ADAM10</i>	ADAM 10	2.5
<i>NPR3</i>	Atrial natriuretic peptide clearance receptor	2.5
<i>TREML2</i>	Triggering receptor expressed on myeloid cells-like 2	2.2

Table 1: Transcripts encoding transmembrane domain containing proteins up-regulated in megakaryocytes

Transcripts are ranked within each structural category by the fold difference in expression between MKs and EBs as detected using cDNA microarrays.

*Transcripts were selected for further analysis.

HGNC identifier	Description	Fold difference, EB/MK
<i>ERAF</i>	Erythroid associated factor	142.3
<i>CA1</i>	Carbonic anhydrase I	140.9
<i>TAC3</i>	Neurokinin B precursor	116.6
<i>GDF15</i>	Growth/differentiation factor 15	73.45
<i>HBA2</i>	Haemoglobin, α 2	67.3
<i>PECI</i>	Peroxisomal 3,2-trans-enoyl-CoA isomerase	52.5
<i>PKLR</i>	Pyruvate kinase, isozymes R/L	50.0
<i>CCR7</i>	C-C chemokine receptor type 7	43.5
<i>TACSTD1</i>	Tumor-associated calcium signal transducer 1	38.8
<i>HBE1</i>	Hemoglobin ϵ chain	38.5
<i>NP_112177.1</i>	MSTP014	37.7
<i>SELENBP1</i>	Selenium-binding protein 1	36.0
<i>MARCH2</i>	Membrane-associated RING finger protein 2	35.1
<i>KRT17</i>	Keratin, type I cytoskeletal 17	32.4
<i>CTSL</i>	Cathepsin L	30.0
<i>KCNH2</i>	Potassium voltage-gated channel subfamily H member 2	29.0
<i>ELA2B</i>	Elastase 2B precursor	27.0
<i>FAM83A</i>	FAM83A	27.7
<i>CRYZ</i>	Quinone oxidoreductase	25.9
<i>TNFSF7</i>	Tumor necrosis factor ligand superfamily member 7	25.4

Table 2: Top 20 genes up-regulated in erythroblasts

Transcripts are ranked based on the fold difference between EBs and MKs as determined using cDNA microarrays.

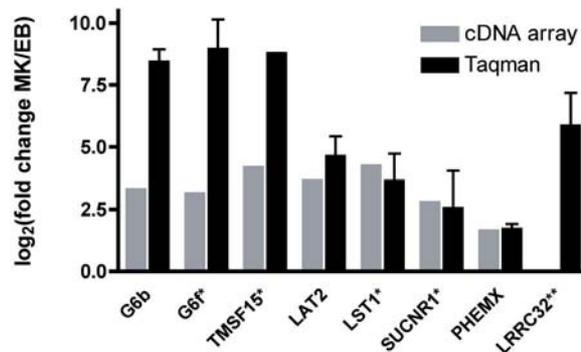


Figure 2: Confirmation of microarray results using TaqMan real-time PCR

Values for both cDNA array measurements (pale grey) and real-time PCR (black) are plotted as \log_2 (fold change). Results are shown as the average for the 5-paired MK/EB comparisons. Although all transcripts were detected by RT-PCR in all 5 of the MK samples, transcripts marked with an asterisk (*) were not detected in all of the EB cultures. Thus, the ratio is based on fewer than 5 comparisons (G6f, $n = 2$; TMSF15, $n = 1$; LST1 and SUCNR1, $n = 4$; all others $n = 5$). **LRRC32 was identified as differentially expressed in preliminary array experiments and by Taqman real-time PCR. Error bars represent SD of replicates.

We combined a thorough literature review with bioinformatical analysis of the encoded proteins to identify previously uncharacterized proteins that may have a modulatory role in platelet function. This search was restricted to Ig-SF members, GPCRs, tetraspannins, adaptor molecules, LRR-containing proteins, as well as proteins reported as having a signaling function in other cell types. By this process, we identified G6b, G6f, SUCNR1, TSPAN15, PHEMX, LAT2, and LST1 as candidates for validation. The LRR protein LRRC32 was also included for further investigation, as it was differentially expressed in MKs in preliminary microarray experiments.

Validation and blood cell expression of selected transcripts

The presence and differential expression of 8 novel transcripts encoding transmembrane proteins was investigated by quantitative RT-PCR. All transcripts tested were detected in the MK samples analyzed, and the direction of the differential expression observed on the microarrays was confirmed by the RT-PCR data. In the case of G6b (359-fold), G6f (514-fold), and TSPAN15 (459-fold) the fold differences observed were higher than those detected on the cDNA arrays (Figure 2).

The expression pattern of the selected transcripts in terminally differentiated peripheral blood cells was also determined by quantitative RT-PCR (Table 3). RNA was obtained from purified populations of platelets, B cells (CD19⁺), T cells (CD4⁺ and CD8⁺), monocytes (CD14⁺), and granulocytes (CD16⁺). The G6b transcript was detected in MKs and platelets and was substantially more abundant in the former than in CD4⁺ T-cells, monocytes, and B-cells (258-fold, 1234-fold, and 77-fold more abundant, respectively) and not detected in CD8⁺ T-cells and granulocytes. The transcript encoding G6f was not detected in any other cells except MKs and platelets, and the transcript encoding SUCNR1 was detected in MKs, platelets, and monocytes (24-fold less abundant than in MKs). LRRC32 mRNA was detected in MKs and to a lesser extent (16-fold less) in B-cells. TSPAN15 was also more abundant in MKs than any other cell type tested but showed high levels of expression in CD4⁺ T-cells and

Transcript name	Structural features of protein	Platelet	CD4 ⁺ T _H cell	CD8 ⁺ T _C cell	CD14 ⁺ monocyte	CD16 ⁺ granulocytes	CD19 ⁺ B cell
<i>G6B</i>	IgSF	2.7	258.7	ND	1234.7	ND	77.7
<i>G6F</i>	IgSF	2.4	ND	ND	ND	ND	ND
<i>SUCNR1</i>	GPCR	-1.5	ND	ND	24.0	ND	ND
<i>TSPAN15</i>	Tetraspannin	19.6	2.5	30.3	ND	ND	ND
<i>PHEMX</i>	Tetraspannin	12.7	1.8	-1.1	2.0	3.2	5.6
<i>LST1</i>	TM protein	1.3	-13.4	1.6	-47.9	-385.3	-8.5
<i>LAT2</i>	Adaptor molecule	ND	2.8	ND	-4.0	-9.6	-9.3
<i>LRRC32</i>	LRR	ND	ND	ND	ND	ND	13.0

Table 3: Transcript distribution in purified populations of blood cells

Table shows fold difference ($2^{-\Delta\Delta Ct}$) relative to MK RNA. Positive values indicate that the transcript was more abundant in MKs than the other cell type tested. Negative values indicate higher expression in cell type tested relative to MKs. ND indicates not detected in cell type tested.

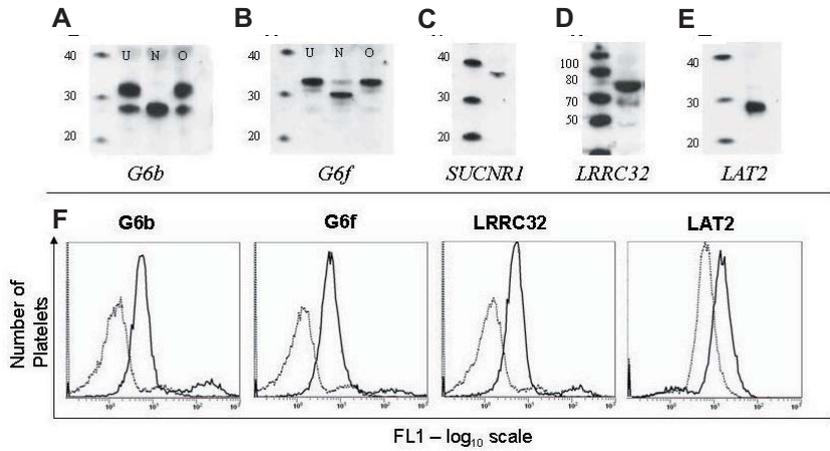


Figure 3: Expression of novel transmembrane proteins in platelets
A-B. G6b and G6f were detected in platelets. *N*-glycosidase (lanes marked N), but not *O*-glycosidase (lanes marked O) treatment of a platelet lysate shows that both G6b and G6f are *N*-glycosylated platelet proteins. The first lane in each blot is a molecular weight marker (MagicMark; Invitrogen; sizes in kDa). **C-E.** Detection of SUCNR1, LRRC32, and LAT2 in platelets by Western blot. **F.** Flow cytometric detection of novel proteins in platelets. Dotted line shows fluorescence detection of a matched preimmune serum (or in the case of LAT2, murine IgG1), and solid line shows the fluorescence using antisera/antibodies against the cognate antigen.

is thus not platelet specific. PHEMX, LST1, and LAT2 expression was pleiomorphic, corresponding to their previously confirmed expression in other blood cell types.^{27,28,29} Thus G6b, G6f, LRRC32, and SUCNR1 appear to be the most MK restricted of the transcripts detected.

Confirmation of protein expression in platelets and other blood cells

By Western blotting, we were able to demonstrate the presence of all proteins tested in platelets. Two bands were detected for G6b (26 and 32 kDa; Figure 3A, lane U),

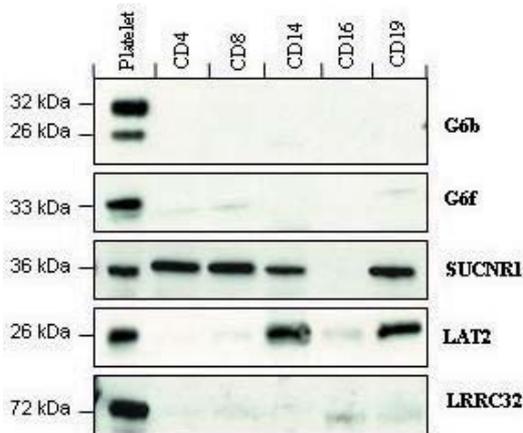


Figure 4: Protein distribution in purified populations of blood cells
 G6b, G6f, and LRRC32 are restricted to platelets, whereas LAT2 is also detected in monocytes (CD14⁺) and B cells (CD19⁺). The succinate receptor, SUCNR1, was detected in all cells tested except granulocytes.

and *N*-glycosidase treatment demonstrated that the higher molecular weight band represents an *N*-glycosylated form of the G6b protein (Figure 3A, lane N). G6f was detected as a 33-kDa protein (Figure 3B, lane U), which when treated with *N*-glycosidase was detected at 31 kDa (Figure 3B, lane N). Neither G6b nor G6f are *O*-glycosylated in platelets (Figures 3A-B, lanes marked O). Although the transcripts encoding LRRC32 (72 kDa) and LAT2 (29 kDa) could not be detected in platelets, the corresponding proteins were detected by Western blot (Figure 3C-D). A single band of the correct size was also observed for SUCNR1 (36 kDa; Figure 3E). Furthermore, the expression of G6b, G6f, and LRRC32 on the platelet surface was confirmed by flow cytometry on fixed platelets (Figure 3F). LAT2 was detected in permeabilized platelets (Figure 3F), whereas SUCNR1 could not be detected by flow cytometry because of the unsuitability of the antibody for flow cytometry (data not shown).

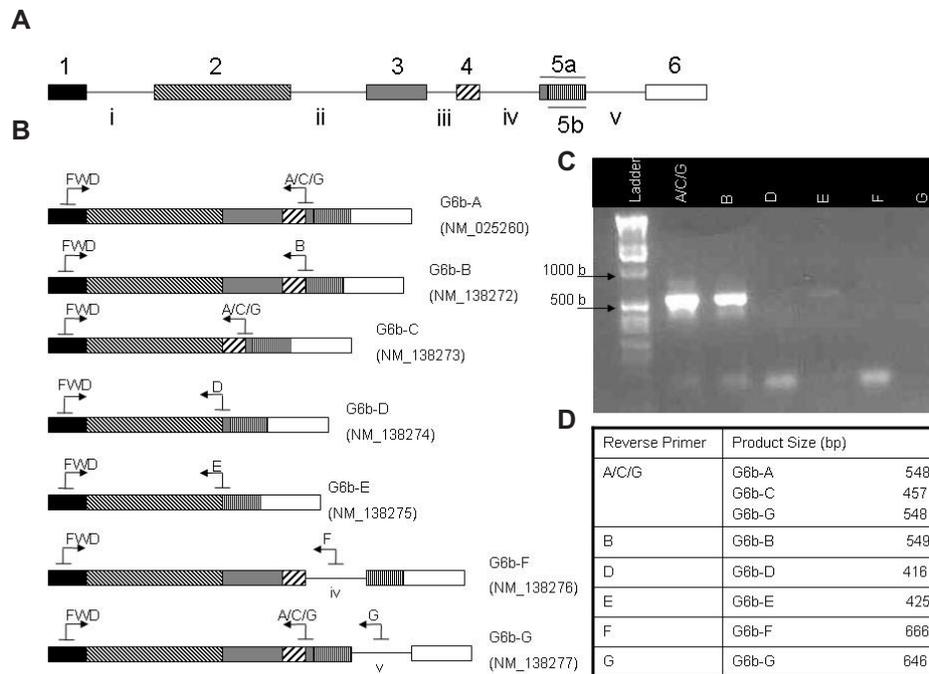


Figure 5: Alternative splicing of G6b in platelets

A. The G6b (*c6orf25*) gene consists of 6 exons, with the fifth exon containing 2 potential splice acceptor sites (exon 5a or 5b). **B.** Seven splice variants of the G6b transcript are annotated in the Entrez Gene database. Transcripts encoding G6b-F and -G contain the intronic sequences IV and V, respectively. Splice variant-specific primer binding sites are represented by arrows. **C.** RT-PCR was used to determine which G6b transcripts are expressed in platelets. The reverse primer G6b-A/C/G potentially amplifies G6b-A, -C, and -G (lane 2); however, only G6b-A is detected because the 457-bp band corresponding to G6b-C is not observed nor is any band detected with the G6b-G-specific primer (lane 7). G6b-B (lane 3) is the only other splice variant detected. **D.** Expected sizes of PCR products.

Western blot analysis of the purified populations of blood cells is largely in agreement with the transcript expression data, with G6b, G6f, and LRRC32 proteins undetected in blood cells other than platelets (Figure 4). However, in contrast with the restricted detection of the transcript, the SUCNR1 protein was in T-cells, monocytes, and B-cells, whereas the LRRC32 protein is restricted to platelets and LAT2 is expressed in platelets, monocytes, and B-cells (Figure 4).

Analysis of G6b splice variants present in platelets

The *C6orf25* gene encodes 7 splice variants of the G6b transcript and protein, including transmembrane and secreted isoforms (Figure 5A-B). Splice variant specific PCR primers were used to determine which of these isoforms were most abundant in platelets. Using this approach, we could detect only G6b-A and G6b-B splice variants in platelets (Figure 5C).

Succinate potentiates the effect of ADP at physiologically relevant concentrations

On its own, even high doses of succinate (1 mM) do not cause platelet aggregation. However, in combination with a suboptimal dose of ADP (1 μ M) the effect of succinate is profound, with a substantial increase in secondary aggregation (Figure 6A-B). We then investigated the effect of physiologically relevant concentrations of succinate on platelet aggregation. Final aggregation in response to 1 μ M ADP averaged of 17.3% (\pm 9.8%) in the 4 donors tested. In the presence of 50 and 100 μ M succinate, final aggregation rose slightly to 24.3% (\pm 9.1%) and 35.8% (\pm 15.5%), respectively. At higher doses of succinate (0.3-1 mM), final aggregation was higher in all donors tested, averaging 82.2% (\pm 8.3%) to 86.9% (\pm 5.1%) aggregation (Figure 6C-D). The presence of succinate has a small effect on primary aggregation in response to ADP, but the main effect is on secondary aggregation (Figure 6D). The same effect was seen when platelets were stimulated with minimally activating doses of CRP-XL or TRAP-6 and succinate (Figure 6E and 6F, respectively).

Discussion

Through comparative gene expression profiling of *in vitro*-differentiated MKs and EBs, we identified subsets of genes that are differentially expressed between the 2 cell types. The initial aim of this work was the identification of novel membrane proteins which modulate platelet function. In total 516 unique genes were up-regulated in MKs when compared to EBs, and 298 genes were up-regulated in the EB. Although the MKs used in this study were differentiated from cord blood derived CD34⁺ HPCs, we would expect considerable overlap between the genes expressed in these cells and those expressed in primary bone marrow MKs. The cultured cells are morphologically and phenotypically similar to committed MK progenitors, and a panel of MK- and platelet-specific transcripts are significantly up-regulated in MKs over EBs.

We used publicly available dataset studies to confirm some of our findings. The first published platelet transcriptome study³⁰ identified 2147 transcripts as “present” in the platelet using Affymetrix U95 Av2 microarrays. A total of 157 transcripts (30%) were shared between this list and the MK-up-regulated fraction. The greatest overlap between this and other studies was seen when the data was compared with the study of normal MKs derived from bone marrow CD34⁺ cells, using Affymetrix U133A

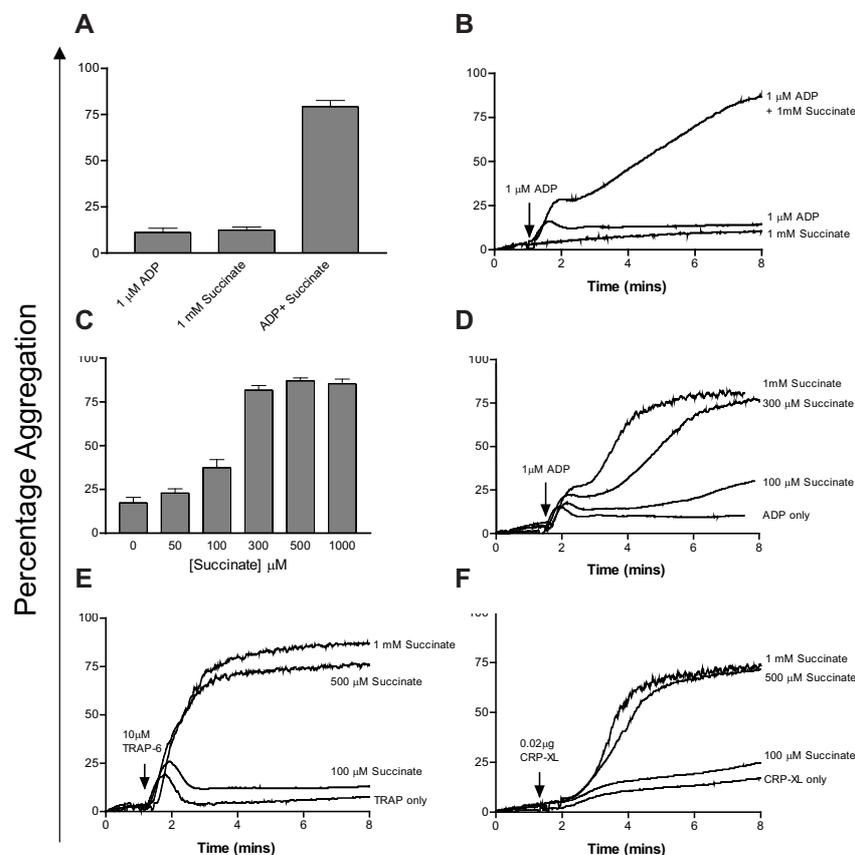


Figure 6: Succinate potentiates platelet aggregation in response to low doses of agonists

A. Succinate (1 mM) alone has no effect on platelets, whereas a combination of succinate (1 mM) and ADP (1 μ M) significantly increases final aggregation. Error bars are calculated based on duplicate measurements of final aggregation in 4 individuals. **B.** A representative aggregometry trace from this experiment. **C.** Investigation of the dose response to succinate and ADP. Platelets were costimulated with 1 μ M ADP and varying concentrations of succinate (0 - 1000 μ M) and final aggregation was measured. Data are shown as the average of triplicate measurements at each concentration for each of 4 donors; error bars show the SEM of the values obtained. **D.** A representative aggregometry trace from this experiment. **E-F.** Similar concentrations of succinate also potentiate the effect of low doses of TRAP-6 (10 μ M) and CRP-XL (0.02 μ g) on platelet aggregation.

microarrays (GEO dataset GSM15648)³¹ in which approximately 69% (358 transcripts) of the MK-up-regulated transcripts described here were identified as “present”. Less overlap with published platelet proteome studies was observed, because in part of the restricted scale of proteomic studies. Of the 434 proteins described in platelets in the Oxford Glycoproteomics database,^{32,33} 63 were also detected as up-regulated in the MK in these experiments (12% of the up-regulated transcripts). A similar comparison

with the more recent study of Martens *et al.*,³⁴ in which 641 platelet proteins were identified, showed that 99 of these platelet proteins were among the MK-up-regulated genes (19% of all up-regulated genes). Taken together, the comparison between this study and those mentioned above shows a 73% overlap (379 of 516 genes) between the MK-up-regulated gene fraction and other published genomic studies of the platelet and MK. The significant overlap between these studies analyzed using considerably different experimental approaches reinforces the validity of *in vitro*-differentiated MKs as a model for MK transcriptome analysis. A similar comparative analysis of the genes up-regulated in EBs showed that 10 % of the encoded proteins were detected in a recent proteomics study of red blood cells.³⁵ An additional comparison with the HEMBASE Database^{36,37} which consists of 3 erythroid EST libraries, confirmed that 112 (37.5 %) of the 298 EB-up-regulated transcripts had been previously observed in erythroid cells.

Bioinformatical filtering of the MK-up-regulated gene list identified 151 novel transcripts that encode putative transmembrane domain-containing proteins. From this list, genes were selected for further investigation based on 2 criteria. First, the gene must encode a protein that is unknown or poorly characterized in the human platelet. Second, the protein must have structural features or a reported function that suggest a role in platelet function. To this end, we selected 8 transcripts for validation by Taqman real-time PCR. Two Ig-SF proteins (G6b^{21,38} and G6f^{39,40}), 2 tetraspanins (TSPAN15⁴¹ and PHEMX^{27,42,43}), a single GPCR (SUCNR1),¹⁰ an adaptor molecule (LAT2),^{44,45} an immunomodulatory protein (LST1),²⁹ and the LRR-containing protein LRRC32^{46,47} were all confirmed to be up-regulated in MKs when compared to EBs. The expression pattern of these transcripts in other blood cells, platelets, T-cells (CD4⁺ and CD8⁺), B-cells, monocytes, and granulocytes, was then examined to determine if any of these transcripts were uniquely expressed in the MK/platelet lineage. The transcripts encoding G6b, G6f, LRRC32, and SUCNR1 are the most platelet restricted. Using either murine polyclonal antisera generated against *E coli*-expressed recombinant protein (G6b, G6f, and LRRC32) or commercially available antibodies (LAT2, SUCNR1), we were able to confirm that these proteins are expressed in platelets. Western blot analysis of protein expression in other blood cells confirmed that G6b, G6f, and LRRC32 are platelet specific; we have also confirmed this result by flow cytometry on whole blood (data not shown). Unexpectedly, the SUCNR1 protein appears to be expressed in all cells but granulocytes. The discrepancy between transcript detection and protein detection remains to be clarified, although examination of the GNF gene expression database⁴⁸ suggests that the SUCNR1 transcript is abundant in CD34⁺ progenitor cells. The protein may therefore be expressed during hematopoiesis and remain present in terminally differentiated cells, even in the absence of sustained transcription of the gene.

The genes encoding G6b and G6f are located within close proximity to each other (separated by 5.5 kb [kilobase]) in the MHC class III locus on chromosome 6; however, their function is unknown,⁴⁹ although G6b has been shown to bind heparin with high affinity.³⁸ Both have recently been detected in proteomic studies of the platelet.^{40,34} Our work corroborates this by confirming expression of the G6b and G6f transcript in MKs and platelets, as well as confirming protein expression on the platelet surface. Our data also suggests the expression of both is restricted to platelets and MKs. Furthermore, we show that G6b is expressed in platelets in both an *N*-glycosylated and nonglycosylated form, consistent with observations made in G6b-transfected cell

lines,²¹ whereas G6f appears to be present only in its glycosylated form. In addition, we demonstrate that splice variants G6b-A and G6b-B of this gene are expressed in platelets. Both encode transmembrane isoforms of the protein, one of which, G6b-B, contains 2 immunoreceptor tyrosine inhibitory motif (ITIM) domains in its cytoplasmic tail and has been shown to interact with the SH2-containing protein phosphatases SHP-1 and SHP-2 on phosphorylation.²¹ These features suggest a potential inhibitory role in platelet function, and in agreement with this we have demonstrated that the cross-linking of G6b using polyclonal antiserum inhibits platelet aggregation.⁵⁰ That 2 splice variants with identical extracellular domains, but different cytoplasmic tails, are expressed on platelets may have functional consequences in modulating the downstream signal upon ligand binding.

G6f has been shown to recruit the Src homology 2 domain containing proteins Grb2 and Grb7 on phosphorylation and activate the Ras-MAP kinase pathway, inducing phosphorylation and activation of ERK (p44/42 MAPK) in K562 cells,³⁹ suggesting it could potentially have an activatory role in platelet or MK signaling. A recent study has indeed shown that the cytoplasmic tail of G6f undergoes phosphorylation and recruits Grb2 in CRP-stimulated platelets.⁵¹

The LRR protein LRRC32 and the adaptor molecule LAT2 were also identified in platelets in this study. Transcripts encoding both were expressed in MKs, and, although neither of the transcripts could be detected in platelets, we have demonstrated that the proteins are present. Because platelets cannot synthesize nuclear mRNA, it is not surprising that there is some discrepancy between the detection of transcript and protein in platelets. LRRC32 (GARP, GARPIN) is a 72-kDa protein which consists of 21 extracellular LRRs and a 13-amino acid cytoplasmic tail, the expression of which is restricted to platelets and megakaryocytes. As a structural homologue of GPIIb and GPV, the protein has previously been observed in murine platelets, MKs, and endothelial cells.^{46,47} No function has yet been identified for LRRC32, but the absence of a long cytoplasmic tail suggests it operates as part of a multiprotein complex. LAT2 (WBSCR5, LAB, or NTAL) has structural and functional similarity to LAT.^{44,52} LAT2 is expressed in B cells, natural killer cells, and mast cells, in which it is coexpressed with LAT.^{53,45} LAT and LAT2 have been shown to have both synergistic and antagonistic roles in mast cell signaling,^{45,28} and the same may apply in platelets.

The *SUCNR1* gene is located on chromosome 3, in close proximity to *P2RY12* and *P2RY1* as part of a cluster of 7 genes encoding GPCRs.⁵⁴ The gene encodes a novel GPCR that has the citric acid cycle intermediate succinate as its ligand. In *SUCNR1*-transfected cells, succinate induced calcium flux and inhibition of cAMP synthesis, consistent with it being coupled to G_i/G_o and Gq signaling pathways.¹⁰

The fact that platelets express a receptor for succinate may be of physiologic significance. More than 2 decades ago, it was shown that succinate could potentiate the effect of several platelet agonists, including ADP, epinephrine, and the endoperoxide analog U46619.⁵⁵ We have shown that physiologically relevant doses of succinate potentiate the effect of a low dose (1 μ M) of ADP on platelet aggregation. In some of the donors we tested, partial potentiation was seen at doses as low as 50 and 100 μ M succinate, and in all donors maximal secondary aggregation was observed between 300 and 500 μ M. Succinate also had the same effect on suboptimal doses of the GPVI-specific ligand CRP-XL^{56,57,58} and the PAR-1 receptor agonist TRAP-6, suggesting that this costimulatory effect is not restricted to ADP. Although further

work is underway in our laboratories to clarify the molecular events that mediate this response, Huang *et al.*⁵⁵ have reported that succinate inhibits cAMP synthesis in platelets, an observation which is consistent with signal transduction through a G_i-coupled GPCR. In SUCNR1-transfected cells, a 5 μ M dose of succinate was required to give a half-maximal calcium flux response.¹⁰ In the platelet aggregation studies described herein, succinate doses of 100 μ M or greater were required to elicit a reproducible secondary response. That succinate alone cannot induce a response, but requires another agonist to initiate aggregation is also of interest, suggesting it may act as a secondary signal in platelet activation.

The concentration of succinate in normal circulating plasma is between 5.7 and 90 μ M,^{59,60,57,61,62,63} with succinate levels shown to increase both during exercise (up to 125 μ M)⁶⁰ and in bacterially infected blood cultures (up to 5 mM).⁶⁴ Succinate is released from papillary muscles into the extracellular space in response to cardiac hypoxia^{65,66} and can also induce hypertension via a SUCNR1-dependent mechanism.¹⁰ The possibility of a connection between elevated succinate levels, enhanced platelet activity, and hypoxia is thus a promising avenue for future research into the role of platelets in the pathologic consequences of atherosclerosis and atherothrombosis.

In conclusion, we have shown that comparative gene expression profiling of MKs and EBs can be used to identify novel transcripts and proteins expressed in their circulating progeny. Furthermore, this experimental approach appears to give a high probability of identifying transcripts and proteins with lineage-restricted expression. A focussed approach to the analysis of this microarray data has allowed us to generate and test novel hypotheses about megakaryocyte and platelet biology in health and disease.

Acknowledgements

We thank Dr. Richard Farndale, Department of Biochemistry, University of Cambridge, for the kind gift of CRP-XL. We would also like to acknowledge I. W. de Jong for technical assistance, F. P. J. Mul for cell sorting, and Dr Paul Lyons for assistance in purifying the peripheral blood cell populations.

This work was supported by the European Union 6th Framework Programme (LSHM-CT-2004-503485).

Reference list

- 1 McDonald TP and Sullivan PS. Megakaryocytic and erythrocytic cell lines share a common precursor cell. *Exp Hematol* 1993; **21**: 1316-1320.
- 2 Lemarchandel V, Ghysdael J, Mignotte V, Rahuel C, Romeo PH. GATA and Ets cis-acting sequences mediate megakaryocyte-specific expression. *Mol Cell Biol* 1993; **13**: 668-676.
- 3 Orkin SH. GATA-binding transcription factors in hematopoietic cells. *Blood* 1992; **80**: 575-581.
- 4 Martin DI, Zon LI, Mutter G, Orkin SH. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature* 1990; **344**: 444-447.
- 5 Romeo PH, Prandini MH, Joulin V, Mignotte V, Prenant M, Vainchenker W, Marguerie G, Uzan G. Megakaryocytic and erythrocytic lineages share specific transcription factors. *Nature* 1990; **344**: 447-449.
- 6 Cantor AB and Orkin SH. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene* 2002; **21**: 3368-3376.
- 7 Debili N, Coulombel L, Croisille L, Katz A, Guichard J, Breton-Gorius J, Vainchenker W. Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow.

- Blood* 1996; **88**: 1284-1296.
- 8 Vannucchi AM, Paoletti F, Linari S, Cellai C, Caporale R, Ferrini PR, Sanchez M, Migliaccio G, Migliaccio AR. Identification and characterization of a bipotent (erythroid and megakaryocytic) cell precursor from the spleen of phenylhydrazine-treated mice. *Blood* 2000; **95**: 2559-2568.
 - 9 Macaulay IC, Carr P, Gusnanto A, Ouwehand WH, Fitzgerald D, Watkins NA. Platelet genomics and proteomics in human health and disease. *J Clin Invest* 2005; **115**: 3370-3377.
 - 10 He W, Miao FJ, Lin DC, Schwandner RT, Wang Z, Gao J, Chen JL, Tian H, Ling L. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. *Nature* 2004; **429**: 188-193.
 - 11 Petalidis L, Bhattacharyya S, Morris GA, Collins VP, Freeman TC, Lyons PA. Global amplification of mRNA by template-switching PCR: linearity and application to microarray analysis. *Nucleic Acids Res* 2003; **31**: e142.
 - 12 Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3-
 - 13 Xue Y, Haas SA, Brino L, Gusnanto A, Reimers M, Talibi D, Vingron M, Ekwall K, Wright AP. A DNA microarray for fission yeast: minimal changes in global gene expression after temperature shift. *Yeast* 2004; **21**: 25-39.
 - 14 Dobbin KK, Shih JH, Simon RM. Comment on 'Evaluation of the gene-specific dye bias in cDNA microarray experiments'. *Bioinformatics* 2005; **21**: 2803-2804.
 - 15 Dobbin KK, Kawasaki ES, Petersen DW, Simon RM. Characterizing dye bias in microarray experiments. *Bioinformatics* 2005; **21**: 2430-2437.
 - 16 Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002; **30**: e15.
 - 17 Kerr MK, Afshari CA, Bennett L., et al. Statistical analysis of a gene expression microarray experiment with replication. *Statistica Sinica* 2002; **12**: 203-217.
 - 18 Repsilber D, Fink L, Jacobsen M, Blasing O, Ziegler A. Sample selection for microarray gene expression studies. *Methods Inf Med* 2005; **44**: 461-467.
 - 19 Bates D and Sarkar S. The lme4 package. <http://Imqesalquspbr/CRAN/doc/packages/lme4pdf>. 2006.
 - 20 Benjamini Y and Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B* 1995; **57**: 289-300.
 - 21 de Vet EC, Aguado B, Campbell RD. G6b, a novel immunoglobulin superfamily member encoded in the human major histocompatibility complex, interacts with SHP-1 and SHP-2. *J Biol Chem* 2001; **276**: 42070-42076.
 - 22 van den Oudenrijn S, von dem Borne AE, de Haas M. Differences in megakaryocyte expansion potential between CD34(+) stem cells derived from cord blood, peripheral blood, and bone marrow from adults and children. *Exp Hematol* 2000; **28**: 1054-1061.
 - 23 Williams N and Levine RF. The origin, development and regulation of megakaryocytes. *Br J Haematol* 1982; **52**: 173-180.
 - 24 Chasis JA and Mohandas N. Red blood cell glycoporphins. *Blood* 1992; **80**: 1869-1879.
 - 25 Southcott MJ, Tanner MJ, Anstee DJ. The expression of human blood group antigens during erythropoiesis in a cell culture system. *Blood* 1999; **93**: 4425-4435.
 - 26 European Bioinformatics Institute and Cold Spring Harbor Laboratory. Biomart. <http://www.biomart.org/biomart.martview/>. Accessed June 2006.
 - 27 Nicholson RH, Pantano S, Eliason JF, Galy A, Weiler S, Kaplan J, Hughes MR, Ko MS. Phemx, a novel mouse gene expressed in hematopoietic cells maps to the imprinted cluster on distal chromosome 7. *Genomics* 2000; **68**: 13-21.
 - 28 Rivera J. NTAL/LAB and LAT: a balancing act in mast-cell activation and function. *Trends Immunol* 2005; **26**: 119-122.
 - 29 Rollinger-Holzinger I, Eibl B, Pauly M, Griesser U, Hentges F, Auer B, Pall G, Schratzberger

- P, Niederwieser D, Weiss EH, Zwierzina H. LST1: a gene with extensive alternative splicing and immunomodulatory function. *J Immunol* 2000; **164**: 3169-3176.
- 30 Gnatenko DV, Dunn JJ, McCorkle SR, Weissmann D, Perrotta PL, Bahou WF. Transcript profiling of human platelets using microarray and serial analysis of gene expression. *Blood* 2003; **101**: 2285-2293.
- 31 Tenedini E, Fagioli ME, Vianelli N, Tazzari PL, Ricci F, Tagliafico E, Ricci P, Gugliotta L, Martinelli G, Tura S, Baccarani M, Ferrari S, Catani L. Gene expression profiling of normal and malignant CD34-derived megakaryocytic cells. *Blood* 2004; **104**: 3126-3135.
- 32 Garcia A, Prabhakar S, Brock CJ, Pearce AC, Dwek RA, Watson SP, Hebestreit HF, Zitzmann N. Extensive analysis of the human platelet proteome by two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* 2004; **4**: 656-668.
- 33 O'Neill EE, Brock CJ, von Kriegsheim AF, Pearce AC, Dwek RA, Watson SP, Hebestreit HF. Towards complete analysis of the platelet proteome. *Proteomics* 2002; **2**: 288-305.
- 34 Martens L, Van Damme P, Van Damme J, Staes A, Timmerman E, Ghesquiere B, Thomas GR, Vandekerckhove J, Gevaert K. The human platelet proteome mapped by peptide-centric proteomics: a functional protein profile. *Proteomics* 2005; **5**: 3193-3204.
- 35 Pasini EM, Kirkegaard M, Mortensen P, Lutz HU, Thomas AW, Mann M. In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood* 2006; **108**: 791-801.
- 36 Goh SH, Lee YT, Bouffard GG, Miller JL. Hembase: browser and genome portal for hematology and erythroid biology. *Nucleic Acids Res* 2004; **32**: D572-D574.
- 37 Gubin AN, Njoroge JM, Bouffard GG, Miller JL. Gene expression in proliferating human erythroid cells. *Genomics* 1999; **59**: 168-177.
- 38 de Vet EC, Newland SA, Lyons PA, Aguado B, Campbell RD. The cell surface receptor G6b, a member of the immunoglobulin superfamily, binds heparin. *FEBS Lett* 2005; **579**: 2355-2358.
- 39 de Vet EC, Aguado B, Campbell RD. Adaptor signalling proteins Grb2 and Grb7 are recruited by human G6f, a novel member of the immunoglobulin superfamily encoded in the MHC. *Biochem J* 2003; **375**: 207-213.
- 40 Lewandrowski U, Moebius J, Walter U, Sickmann A. Elucidation of N-glycosylation sites on human platelet proteins: a glycoproteomic approach. *Mol Cell Proteomics* 2006; **5**: 226-233.
- 41 Serru V, Dessen P, Boucheix C, Rubinstein E. Sequence and expression of seven new tetraspans. *Biochim Biophys Acta* 2000; **1478**: 159-163.
- 42 Robb L, Tarrant J, Groom J, Ibrahim M, Li R, Borobakas B, Wright MD. Molecular characterisation of mouse and human TSSC6: evidence that TSSC6 is a genuine member of the tetraspanin superfamily and is expressed specifically in haematopoietic organs. *Biochim Biophys Acta* 2001; **1522**: 31-41.
- 43 Tarrant JM, Groom J, Metcalf D, Li R, Borobakas B, Wright MD, Tarlinton D, Robb L. The absence of Tssc6, a member of the tetraspanin superfamily, does not affect lymphoid development but enhances in vitro T-cell proliferative responses. *Mol Cell Biol* 2002; **22**: 5006-5018.
- 44 Brdicka T, Imrich M, Angelisova P, Brdickova N, Horvath O, Spicka J, Hilgert I, Luskova P, Draber P, Novak P, Engels N, Wienands J, Simeoni L, Osterreicher J, Aguado E, Malissen M, Schraven B, Horejsi V. Non-T cell activation linker (NTAL): a transmembrane adaptor protein involved in immunoreceptor signaling. *J Exp Med* 2002; **196**: 1617-1626.
- 45 Gilfillan AM and Tkaczyk C. Integrated signalling pathways for mast-cell activation. *Nat Rev Immunol* 2006; **6**: 218-230.
- 46 Ollendorff V, Noguchi T, deLapeyriere O, Birnbaum D. The GARP gene encodes a new member of the family of leucine-rich repeat-containing proteins. *Cell Growth Differ* 1994; **5**: 213-219.
- 47 Roubin R, Pizette S, Ollendorff V, Planche J, Birnbaum D, deLapeyriere O. Structure and developmental expression of mouse Garp, a gene encoding a new leucine-rich repeat-containing protein. *Int J Dev Biol* 1996; **40**: 545-555.

-
- 48 Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 2004; **101**: 6062-6067.
 - 49 Mallya M, Campbell RD, Aguado B. Transcriptional analysis of a novel cluster of LY-6 family members in the human and mouse major histocompatibility complex: five genes with many splice forms. *Genomics* 2002; **80**: 113-123.
 - 50 Newland SA, Macaulay IC, Floto AR, de Vet EC, Ouwehand WH, Watkins NA, Lyons PA, Campbell DR. The novel inhibitory receptor G6B is expressed on the surface of platelets and attenuates platelet function in vitro. *Blood* 2007; **109**: 4806-4809.
 - 51 Garcia A, Senis YA, Antrobus R, Hughes CE, Dwek RA, Watson SP, Zitzmann N. A global proteomics approach identifies novel phosphorylated signaling proteins in GPVI-activated platelets: involvement of G6f, a novel platelet Grb2-binding membrane adapter. *Proteomics* 2006; **6**: 5332-5343.
 - 52 Wonerow P and Watson SP. The transmembrane adapter LAT plays a central role in immune receptor signalling. *Oncogene* 2001; **20**: 6273-6283.
 - 53 Chiesa S, Mingueneau M, Fuseri N, Malissen B, Raulet DH, Malissen M, Vivier E, Tomasello E. Multiplicity and plasticity of natural killer cell signaling pathways. *Blood* 2006; **107**: 2364-2372.
 - 54 Wittenberger T, Schaller HC, Hellebrand S. An expressed sequence tag (EST) data mining strategy succeeding in the discovery of new G-protein coupled receptors. *J Mol Biol* 2001; **307**: 799-813.
 - 55 Huang EM, McGowan EB, Detwiler TC. Succinate potentiates the action of platelet agonists. *Thromb Res* 1984; **36**: 1-8.
 - 56 Gibbins JM, Okuma M, Farndale R, Barnes M, Watson SP. Glycoprotein VI is the collagen receptor in platelets which underlies tyrosine phosphorylation of the Fc receptor gamma-chain. *FEBS Lett* 1997; **413**: 255-259.
 - 57 Kehrel B, Wierwille S, Clemetson KJ, Anders O, Steiner M, Knight CG, Farndale RW, Okuma M, Barnes MJ. Glycoprotein VI is a major collagen receptor for platelet activation: it recognizes the platelet-activating quaternary structure of collagen, whereas CD36, glycoprotein IIb/IIIa, and von Willebrand factor do not. *Blood* 1998; **91**: 491-499.
 - 58 Morton LF, Hargreaves PG, Farndale RW, Young RD, Barnes MJ. Integrin alpha 2 beta 1-independent activation of platelets by simple collagen-like peptides: collagen tertiary (triple-helical) and quaternary (polymeric) structures are sufficient alone for alpha 2 beta 1-independent platelet reactivity. *Biochem J* 1995; **306 (Pt 2)**: 337-344.
 - 59 Forni LG, McKinnon W, Lord GA, Treacher DF, Peron JM, Hilton PJ. Circulating anions usually associated with the Krebs cycle in patients with metabolic acidosis. *Crit Care* 2005; **9**: R591-R595.
 - 60 Hochachka PW and Dressendorfer RH. Succinate accumulation in man during exercise. *Eur J Appl Physiol Occup Physiol* 1976; **35**: 235-242.
 - 61 Krebs HA. Chemical composition of blood plasma and serum. *Annu Rev Biochem* 1950; **19**: 409-430.
 - 62 Kushnir MM, Komaromy-Hiller G, Shushan B, Urry FM, Roberts WL. Analysis of dicarboxylic acids by tandem mass spectrometry. High-throughput quantitative measurement of methylmalonic acid in serum, plasma, and urine. *Clin Chem* 2001; **47**: 1993-2002.
 - 63 Nordmann J and Nordmann R. Organic acids in blood and urine. *Adv Clin Chem* 1961; **4**: 53-120.
 - 64 Wust J. Presumptive diagnosis of anaerobic bacteremia by gas-liquid chromatography of blood cultures. *J Clin Microbiol* 1977; **6**: 586-590.
 - 65 Sanborn T, Gavin W, Berkowitz S, Perille T, Lesch M. Augmented conversion of aspartate and glutamate to succinate during anoxia in rabbit heart. *Am J Physiol* 1979; **237**: H535-H541.
 - 66 Taegtmeyer H. Metabolic responses to cardiac hypoxia. Increased production of succinate by rabbit papillary muscles. *Circ Res* 1978; **43**: 808-815.
-