

Transcript profiling of human platelets using microarray and serial analysis of gene expression

Dmitri V. Gnatenko, John J. Dunn, Sean R. McCorkle, David Weissmann, Peter L. Perrotta, and Wadie F. Bahou

Human platelets are anucleate blood cells that retain cytoplasmic mRNA and maintain functionally intact protein translational capabilities. We have adapted complementary techniques of microarray and serial analysis of gene expression (SAGE) for genetic profiling of highly purified human blood platelets. Microarray analysis using the Affymetrix HG-U95Av2 approximately 12 600-probe set maximally identified the expression of 2147 (range, 13%-17%) platelet-expressed transcripts, with approximately 22% collectively involved in metabolism and receptor/signaling, and an overrepresentation of genes with unassigned function (32%). In contrast, a modified SAGE protocol using the Type IIS restriction enzyme

MmeI (generating 21-base pair [bp] or 22-bp tags) demonstrated that 89% of tags represented mitochondrial (mt) transcripts (enriched in 16S and 12S ribosomal RNAs), presumably related to persistent mt-transcription in the absence of nuclear-derived transcripts. The frequency of non-mt SAGE tags paralleled average difference values (relative expression) for the most "abundant" transcripts as determined by microarray analysis, establishing the concordance of both techniques for platelet profiling. Quantitative reverse transcription-polymerase chain reaction (PCR) confirmed the highest frequency of mt-derived transcripts, along with the mRNAs for neurogranin (NGN, a protein kinase C substrate) and

the complement lysis inhibitor clusterin among the top 5 most abundant transcripts. For confirmatory characterization, immunoblots and flow cytometric analyses were performed, establishing abundant cell-surface expression of clusterin and intracellular expression of NGN. These observations demonstrate a strong correlation between high transcript abundance and protein expression, and they establish the validity of transcript analysis as a tool for identifying novel platelet proteins that may regulate normal and pathologic platelet (and/or megakaryocyte) functions. (Blood. 2003;101:2285-2293)

© 2003 by The American Society of Hematology

Introduction

Human blood platelets play critical roles in normal hemostatic processes and pathologic conditions such as thrombosis, vascular remodeling, inflammation, and wound repair. Generated as cytoplasmic buds from precursor bone marrow megakaryocytes, platelets are anucleate and lack nuclear DNA, although they retain megakaryocyte-derived mRNAs.^{1,2} Platelets contain rough endoplasmic reticulum and polyribosomes, and they retain the ability for protein biosynthesis from cytoplasmic mRNA.³ Quiescent platelets generally display minimal translational activity, although newly formed platelets such as those found in patients with immune thrombocytopenic purpura (ITP) synthesize various α -granule and membrane glycoproteins (GPs), including GPIb and GPIIb/IIIa ($\alpha_{IIb}\beta_3$). Furthermore, stimulation of quiescent platelets by agonists such as α -thrombin increases protein synthesis of various platelet proteins, including Bcl-3.⁴ Like nucleated cells, the rapid translation of preexisting mRNAs may be regulated by integrin ligation to extracellular matrices.⁵ In the case of platelets, the primary integrin involved in this process appears to be $\alpha_{IIb}\beta_3$ with cooperative signals mediated by the collagen receptor $\alpha_2\beta_1$.^{6,7}

Integrin-mediated platelet protein synthesis appears to be regulated at the level of translation initiation involving the eukaryotic initiation factor 4E (eIF4E). Instead of directly influencing eIF4E activity via posttranslational modifications (ie, phosphorylation), platelet eIF4E activity best correlates with its spatial redistribution to the mRNA-enriched cytoskeleton.⁸ Furthermore, because protein translation is partially inhibited by the immunosuppressant rapamycin, it suggests that adhesion- and/or aggregation-induced outside-in-signaling function to regulate protein synthesis through the mTOR (mammalian target of rapamycin) pathway.^{6,8,9}

Despite the biologic importance of platelets and their intact protein synthetic capabilities, remarkably little is known about platelet mRNAs. Younger platelets contain larger amounts of mRNA with a greater capacity for protein synthesis, as determined by using fluorescent nucleic acid dyes such as thiazole orange.¹⁰ This assay has been used as a quantitative determinant of younger or "reticulated" platelets (RPs). Indeed increased reticulated platelets are typically found in patients with conditions associated with rapid platelet turnover such as ITP; typically RP percentages in

From the Department of Medicine, Department of Pathology, and Program in Genetics, State University of New York, Stony Brook; Biology Department, Brookhaven National Laboratory, Upton, NY; and Department of Pathology, Robert Wood Johnson Medical Center, New Brunswick, NJ.

Submitted September 16, 2002; accepted November 3, 2002. Prepublished online as *Blood* First Edition Paper, November 14, 2002; DOI 10.1182/blood-2002-09-2797.

Supported by grants HL49141 and HL53665, by a Veteran's Administration REAP award (W.F.B.), and by National Institutes of Health Center grant MO1 10710-5 to the University Hospital General Clinical Research Center. W.B. is an Established Investigator of the American Heart Association. Studies at

Brookhaven National Laboratory were supported by a Laboratory Directed Research and Development award (J.J.D.) and by the Offices of Biological and Environmental Research, and of Basic Energy Sciences (Division of Energy Biosciences) of the US Department of Energy.

Reprints: Wadie F. Bahou, Division of Hematology, HSCT15-040, State University of New York at Stony Brook, Stony Brook, NY 11794-8151; e-mail: wbahou@notes.cc.sunysb.edu.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

© 2003 by The American Society of Hematology

such patients approach 10% to 20% of all platelets, considerably higher than in healthy control subjects.¹¹ Interestingly, high RPs have been associated with enhanced thrombotic risk when identified in patients with thrombocytosis,¹⁰ suggesting that quantitatively increased mRNA levels may be associated with the prothrombotic phenotype. Whether this is related to globally altered gene expression profiles or to select changes more evident during situations of rapid platelet turnover remains unknown. Certainly, technical limitations of this assay limit its utility in defining prothrombotic genotypes,¹⁰⁻¹² and it cannot identify differentially expressed genes that may be causally implicated in disordered platelet phenotypes.

Toward the goal of defining the molecular anatomy of the platelet genome, we have adapted complementary techniques of microarray and serial analysis of gene expression (SAGE) for genetic profiling of highly purified human blood platelets. Microarray technology represents a “closed” profiling strategy limited by the target genes imprinted onto gene chips. In contrast, SAGE is an “open” architectural system that can be used to identify novel genes and to quantify differentially expressed mRNAs.¹³⁻¹⁵ The sequence of each tag along with its positional location uniquely identifies the gene from which it is derived, and differentially expressed genes can be identified in a quantitative manner because the tag frequency reflects the mRNA level at the time of cellular harvest and analysis. By using both technologies, we have identified a number of previously uncharacterized genes that appear to be expressed in human platelets, while simultaneously establishing the dominant frequency of mitochondrial-expressed genomes comprising the platelet mRNA pool. These observations provide a panoramic overview of the platelet transcriptome, while additionally providing insights into the molecular pathways regulating platelet (and/or megakaryocyte) function in normal and pathologic conditions.

Materials and methods

Reagents and supplies

Thermus aquaticus (Taq) polymerase was purchased from (Roche, Indianapolis, IN), T4 DNA ligase was purchased from Invitrogen (Carlsbad, CA), and restriction enzymes were from New England Biolabs (Beverly, MA), except for *MmeI*, which was obtained from the Center for Technology Transfer (Gdansk, Poland). All oligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) 3-channel synthesizer and are listed in Table 1. Monoclonal antibodies used for flow cytometric analysis included the FITC (fluorescein isothiocyanate)-conjugated anti-CD41 ($\alpha_{IIb}\beta_3$) immunoglobulin G1 (IgG1; Immunotech, Miami, FL); phycoerythrin (PE)-conjugated antiglycophorin (IgG2; Becton Dickinson Pharmingen, San Diego, CA); and peridinin chlorophyll protein (PERCP)-conjugated anti-CD45 (IgG1; Becton Dickinson Pharmingen).

Platelet isolation, purification, and immunodetection

All human subjects provided informed consent for an IRB (Institutional Review Board)-approved protocol completed in conjunction with the General Clinical Research Center at Stony Brook University Hospital. Peripheral blood (20 mL) from healthy volunteers drawn into 2 mL of 4% sodium citrate (0.4% vol/vol final concentration) was used to isolate erythrocytes by differential centrifugation (1500g) or to isolate pure leukocytes by density-gradient centrifugation as previously described.¹⁶ Platelets collected from healthy volunteers by apheresis were used within 24 hours of collection. After addition of 2 mM EDTA (ethylenediaminetetraacetic acid), apheresis-derived platelets from a single donor were centrifuged at 140g for 15 minutes at 25°C. To minimize leukocyte contamination, only the upper 9/10 of the platelet-rich plasma (PRP) was

used for gel filtration over a BioGel A50M column (1000 mL total volume) equilibrated with HBMT (HEPES-buffered modified Tyrodes buffer: 10 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid) pH 7.4, 150 mM NaCl, 2.5 mM KCl, 0.3 mM NaH₂PO₄, 12 mM NaHCO₃, 0.2% bovine serum albumen [BSA], 0.1% glucose, 2 mM EDTA). Gel-filtered platelets (GFPs) were subsequently filtered through a 5- μ m nonwetting nylon filament filter (BioDesign, Carmel, NY) at 25°C and harvested by centrifugation at 1500g for 10 minutes at 25°C. Platelets were gently and thoroughly resuspended in 10 mL HBMT buffer and incubated with 120 μ L murine monoclonal anti-CD45 antibody conjugated to magnetic microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany) on a rotating platform for 45 minutes at 25°C. Magnetic separation columns were used to capture CD45⁺ cells (leukocyte fraction) by positive selection (MACS II; Miltenyi Biotec). Purified platelets were concentrated by centrifugation at 1500g and immediately used for total RNA isolation.

The efficiency of platelet purification was documented at each step by flow cytometry.¹⁷ Briefly, aliquots containing 2×10^6 platelets were incubated with saturating concentrations of FITC-conjugated anti-CD41, PE-conjugated antiglycophorin, and PERCP-conjugated anti-CD45 for 15 minutes in the dark at 25°C, washed with phosphate-buffered saline (PBS), and fixed in PBS/1% formalin. Samples were analyzed using a FACScan (fluorescence-activated cell sorter scan) flow cytometer (Becton Dickinson) using CELLQuest software designed to quantify the number of CD45⁺ and glycophorin-positive events in the sample (expressed as the number of events per 100 000 CD41⁺ events). For some experiments, fixed platelets were permeabilized with 0.1% Triton-X/PBS for 30 minutes at 25°C prior to the addition of primary antibodies, all as previously described.¹⁷

Platelet protein detection was completed by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and immunoblot analysis as previously described, using the species-specific horseradish peroxidase-conjugated secondary antibody and enhanced chemiluminescence.¹⁶ Antibodies included the anticlustarin monoclonal antibody (Quidel, Santa Clara, CA; 1:1000 primary and 1:10 000 secondary) and the antineurogranin rabbit polyclonal antibody (Chemicon International, Temecula, CA; 1:1000 primary and 1:10 000 secondary).

Molecular analyses and microarray profiling

Purified, individual cell fractions were resuspended in 10 mL Trizol reagent (Invitrogen), transferred into diethylpyrocarbonate (DEPC)-treated Corex (Springfield, MA) tubes, and serially purified and precipitated by using isopropanol essentially as previously described.¹⁶ Total cellular RNA was harvested by centrifugation at 12 500g for 20 minutes at 4°C, washed 2 times with 75% ethanol (10 mL/tube), and resuspended in 100 μ L DEPC-treated water. Platelet mRNA quantitation was performed by using fluorescence-based real-time PCR (polymerase chain reaction) technology (TaqMan Real-Time PCR; Applied Biosystems, Foster City, CA). Oligonucleotide primer pairs were generated by using Primer3 software (www.genome.wi.mit.edu), designed to generate approximately 200-base pair (bp) PCR products at the same annealing temperature, and are outlined in Table 1. Purified platelet mRNA (4 μ g) was used for first-strand cDNA synthesis using oligo(dT) and SuperScript II reverse transcriptase (Invitrogen). For real-time reverse transcription (RT)-PCR analysis, the RT reaction was equally divided among primer pairs and used in a 40-cycle PCR reaction for each target gene by using the following cycle: 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, and 71°C for 10 seconds (40 cycles total). mRNA levels were quantified by monitoring real-time fluorometric intensity of SYBR green I. Relative mRNA abundance was determined from triplicate assays performed in parallel for each primer pair and was calculated by using the comparative threshold cycle number (Δ -Ct method) as previously described.¹⁸

Gene expression profiles were completed by using the approximately 12 600-probe set HG-U95Av2 gene chip (Affymetrix, Santa Clara, CA). Total cellular RNA (5 μ g) was used for cDNA synthesis by using SuperScript Choice system (Life Technologies, Rockville, MD) and an oligo(dT) primer containing the T7 polymerase recognition sequence (Primer S1; Table 1), followed by cDNA purification using GFX spin columns. In vitro transcription was completed in the presence of biotinylated ribonucleotides by using a BioArray HighYield RNA Transcript

Table 1. Oligonucleotide primers

Primer	Gene and primer direction	Sequence (5' - 3')	Nucleotide Position
S1	Oligo (dT)	5'-Bn-GGCCAGTGAATTGTAATACGACTCCTATAGGGAGGCGG- (dT) ₂₄ -3'	—
Cassette A	SAGE	5'-TTTGGATTGCTGGTTCGAGTACAACCTAGGCTTAATCCGACATG-3' 3'-*CCTAAACGACCAGCTCATGTTGATCCGAATAAGGCTp-5'	—
Cassette B	SAGE	5'-pTTCAATGGCGGAGACGTCGCCACTAGTGTCCGAACTGACTA*-3' 3'-NNAAGTACCGCCTCTGCAGGCGGTGATCACAGCGTTGACTGAT-5'	—
S2	SAGE	5'-Bn-GGATTTGCTGGTTCGAGTACA-3'	—
S3	SAGE	5'-Bn-TAGTCAGGTGCGACACTAGTGGC-3'	—
GP4	Glycoprotein IIB [F]	5'-AGGGCTTTGAGAGACTCATCTGTA-3'	2094-2117
GP5	Glycoprotein IIB [R]	5'-ACAATCTTGTCTGTTTGGATTCTG-3'	2301-2279
GP6	Glycoprotein IIIA [F]	5'-TATAAAGAGGCCACGCTACCTTC-3'	2335-2358
GP7	Glycoprotein IIIA [R]	5'-CACTTCCACATACTGACATTTCTCC-3'	2532-2509
PAR18	PAR1 [F]	5'-AATGTCAGTTCTGATATGGAAGCA-3'	2585-2608
PAR19	PAR1 [R]	5'-CCCAAATGTTCAAACCTCTTTAGC-3'	2776-2753
SR8	16S rRNA [F]	5'-TGCAAAGGTAGCATAATCACTTGT-3'	2586-2609
SR9	16S rRNA [R]	5'-GTTTAGGACCTGTGGGTTTGTAG-3'	2785-2762
NADH10	NADH2 [F]	5'-CTAGCCCCATCTCAAATCATATAC-3'	4875-4898
NADH11	NADH2 [R]	5'-AATGGTTATGTTAGGGTTGTACGG-3'	5075-5052
THYM12	Thymosin β4 [F]	5'-AAGACAGAGACGCAAGAGAAAAAT-3'	135-158
THYM13	Thymosin β4 [R]	5'-GCAGCACAGTCATTTAAACTTGAT-3'	336-313
CLUS14	Clusterin [F]	5'-CCAACAGAATTCATACGAGAAGG-3'	1006-1028
CLUS15	Clusterin [R]	5'-CGTTATATTTCTGGTCAACCTCT-3'	1222-1199
NRG16	Neurogranin [F]	5'-GCCCTTTTAGTTAGTTCTGCAGTC-3'	1351-1374
NRG17	Neurogranin [R]	5'-TTTTCTTTAAGTGAGTGTGCTTGG-3'	1567-1544
TCR18	β-chain [F] T-cell receptor	5'-CCACAACATATGTTTGGTATCGT-3'	131-153
TCR19	β-chain [R] T-cell receptor	5'-CTAGCACTGCAGATGTAGAAGCT-3'	332-310
CD4520	CD45 [F]	5'-GCTCAGAATGGACAAGTA-3'	3771-3788
CD4521	CD45 [R]	5'-CACACCCATACACATACA-3'	4280-4261

[F] indicates forward (sense) strand; [R], reverse (antisense) strand; Bn, biotin; p, a phosphorylated 5' end (cassettes A and B); underlining, *Nla*III sites in cassettes A and B; arrows, corresponding sequence for S2 and S3 within cassettes A and B, respectively; bold, the *Mme*I site; and N, A, C, T, or G, nucleotide position based on the following accession numbers: glycoprotein IIB (J02764), glycoprotein IIIA (M35999), PAR1 (M62424), 16S rRNA and NADH2 (NC_001807), thymosin β4 (M17733), clusterin (M25915), neurogranin (X99076), TCR β-chain (AF043182), CD45 (Y00638).

*Indicates an amino-modified 3' end in both cassettes; —, not applicable.

Labeling Kit (Enzo Diagnostics, Farmingdale, NY), and, after metal-induced fragmentation, 15 μg biotinylated cRNA was hybridized to the HG-U95Av2 oligonucleotide probe array for 16 hours at 45°C. After washing, the cRNA was detected with streptavidin-phycoerythrin (Molecular Probes, Eugene, OR) and analysis was completed by using a Hewlett-Packard Gene Array Scanner (Affymetrix). The average difference value (AD) for each probe set was quantified using MAS 4.01 software (Affymetrix), calculated as an average of fluorescence differences for perfectly matched versus single-nucleotide mismatched 25-mer oligonucleotides (16 to 20 oligonucleotide pairs per probe set). The software is designed to exclude “positive calls” in the presence of high average differences with associated high mismatch intensities.

SAGE profiles

Platelet SAGE libraries were generated essentially as previously described,¹³ modified as outlined in Figure 1 for the use of *Mme*I as the tagging enzyme.¹⁹ This type IIS restriction enzyme cleaves 20 of 18 bp past its nonpalindromic (TCCRAC) recognition sequence, thereby generating longer tags (21- or 22-mer) than those obtained using *Bsm*FI as the standard tagging enzyme (13-14 bp tags). These longer *Mme*I-generated tags potentially provide for more definitive “tag-to-gene” identification and are particularly useful in characterizing expression patterns in the absence of complete genomic sequence data (comprehensive methods detailed in Dunn et al¹⁹). Briefly, poly(A) mRNA was isolated from 10 μg total platelet RNA using the oligo-(dT) S1 primer conjugated to magnetic beads (DynaL Biotech, Lake Success, NY), followed by cDNA synthesis using Super-Script II reverse transcriptase (Invitrogen). The cDNA was then digested

with the restriction enzyme *Nla*III (anchoring enzyme), ligated to cassette A using T4 DNA ligase, and, after the beads were extensively washed, the cDNA was digested with *Mme*I to release the tags from the beads. After purification, tags were ligated to degenerate cassette B linkers (specifically

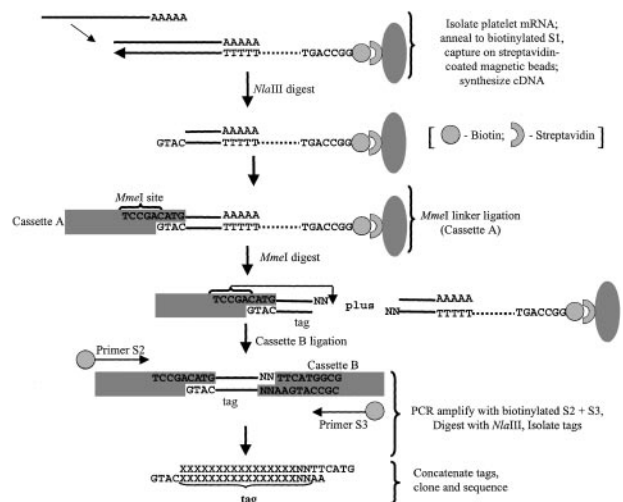


Figure 1. Schema outlining the modified SAGE protocol used in platelet analyses. The final tags are flanked by the *Nla*III (anchoring enzyme) CATG sequence, thereby providing tag-to-gene identification when exported to a relational database (refer to “Bioinformatic analyses” and Table 1 for details).

designed to anneal to the nonuniform *MmeI* overhangs), and PCR-amplified using biotinylated primers S2 and S3 for 30 cycles (95°C for 30 seconds; 58°C for 30 seconds; 72°C for 30 seconds) using Platinum Taq DNA polymerase (Gibco BRL). A fraction (20%) of the pooled PCR products were then subjected to one round of linear amplification using primer pair S2/S3, followed by a second round of 25 amplifications using primer S2 alone (95°C for 30 seconds, 58°C for 30 seconds, 72°C for 30 seconds). Primer S3 was subsequently added for one cycle (95°C for 2.5 minutes, 58°C for 30 seconds, 72°C for 5 minutes); the latter steps were collectively adapted to exclude heteroduplex formation.¹⁸ Unincorporated primers were removed by incubation with 200 U *Escherichia coli* exonuclease I for 60 minutes at 37°C. PCR products were then pooled and digested with *NlaIII* to release tags, and biotinylated linker arms were cleared using streptavidin-coated immunoaffinity magnetic beads (Dynal Biotech). Tags were concatamerized using 5 U/ μ L T4 DNA ligase, and products more than 100 bp were isolated by size-fractionation in low-melting agarose gels. The DNA was purified by GFX spin columns, and the concatamers were cloned into the *SphI* site of pZero (Invitrogen). After transformation into *E coli* TOP10 cells, recombinant clones were isolated and sequenced in 96-well microtiter plates using an ABI 377 sequencer and ABI Prism BigDye terminator chemistry (Perkin-Elmer Applied Biosystems, Branchburg, NJ).

Bioinformatic analyses

Functional grouping of genes determined to be present by Affymetrix MAS 4.01 software was performed using a dChip program linked to the National Center for Biotechnology LocusLink, which is an annotated reference database for genes and their postulated functions.²⁰ Of the approximately 12 600-probe sets represented on the Affymetrix HG-U95Av2 Gene chip, functional annotations exist for approximately 8100 with the remainder categorized as unknown. Microarray data were visualized and analyzed using BRB-ArrayTools software (Version 2.1), kindly developed and provided by Dr Richard Simon and Amy Peng (linus.nci.nih.gov/BRB-ArrayTools.html). A logarithmic (base 2) transformation was applied to the average difference values for individual data sets for determination of microarray concordancies. Discordancy was defined as a 2-log difference in the maximum log intensities between individual experiments.

SAGE tags were extracted by using in-house SAGE software uniquely modified to identify *MmeI* tags. The software ensures that only unambiguous 21- to 22-bp tag sequences are extracted for transcript profiling. Tags with ambiguities (Ns), lengths other than 21 or 22 bp, or with ambiguous orientations were extracted to separate files for manual editing or further examination. Finalized data were exported to a relational database for tag quantification and genetic identification.²⁰

Results

Platelet purification

To ensure that the RNA profiles accurately represented those of circulating blood platelets, a number of complementary methods were implemented to remove contaminating nucleated leukocytes. Purification methods incorporating gel filtration, a 5- μ m leukocyte reduction filter, and magnetic CD45 immunodepletion allowed for the cumulative enrichment of highly purified platelets. The efficacy of this purification method was initially established by using peripheral blood platelet-rich plasma as the starting material. The final product contained no more than 3 to 5 leukocytes per 1×10^5 platelets as determined by parallel flow cytometric analysis, representing an approximate 450-fold reduction of nucleated leukocytes. These results correlated well with molecular evidence for leukocyte depletion as determined by RT-PCR using both CD45 and T-cell receptor β -chain (TCR β) primers (see Figure 2). Because the total RNA yield from peripheral blood platelets was insufficient for microarray studies, we adapted the protocol to platelet apheresis donors with nearly identical final purity (Figure

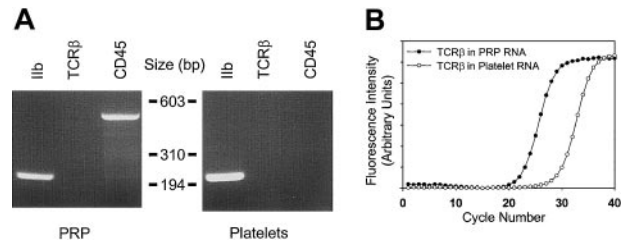


Figure 2. Determination of platelet purity. (A) Total cellular RNA (1.8 μ g) from platelet-rich plasma (PRP) or purified platelets from a single apheresis donor were analyzed by RT-PCR (35 cycles) using oligonucleotide primers specific for glycoprotein IIb (GPIIb), T-cell receptor β -chain (TCR β), or CD45; 10 μ L of the 50 μ L reactions were analyzed by ethidium-stained agarose gel electrophoresis. Minimal to no TCR β gene product was visually evident only in PRP. Size markers corresponding to *HaeIII*-restricted ϕ X174 DNA are shown. (B) Real-time RT-PCR was completed by using 1.8 μ g total RNA and TCR β -specific oligonucleotide primers optimized for quantitative analysis by real-time PCR.¹⁸ On the basis of parallel determinations using RNA isolated from known amounts of purified leukocyte standards, the leukocyte-depletion protocol represents an approximate 2.5-log purification from the starting PRP. Results are representative of one complete set of experiments repeated on 2 separate occasions, and data points represent the mean from triplicate wells, with standard errors of the mean (SEM) less than 1% (not shown).

2). The platelet recovery was nearly 65% of the starting material, yielding approximately 2.3×10^{11} platelets from an initial apheresis pack containing approximately 3.6×10^{11} platelets. The bulk of the losses occurred during the initial centrifugation and filtration steps. The purification protocol was less effective at removing erythrocytes, although there were less than 50 glycoprotein-positive cells per 1×10^5 platelets after the final purification step. Nonetheless, these cells represent unlikely sources for contaminating cellular RNA (see “Cellular microarray analysis” below).

Cellular microarray analysis

The purified platelet RNA was sufficient for microarray studies and was used for cRNA generation and hybridization to the Affymetrix HG-U95Av2 GeneChip. The anatomic profile of platelet RNAs from 3 healthy male donors was determined by using Affymetrix software. Of the 12 599 probe sets imprinted onto the chip, a maximum of 2147 (17%) transcripts were computationally identified as “present” by the Affymetrix software, 152 (1.2%) were equivocal, and nearly 82% were absent. As a fraction of the total genes present on the chip, the percentage of platelet-expressed genes (15%-17%) was generally lower than that obtained from other human cell types in which 30% to 50% of genes are present as determined by Affymetrix software (J. Schwedes, personal communication, May 2002). The “limited number” of platelet-expressed transcripts presumably reflects the lack of ongoing gene transcription in the anucleate platelet. Because less than 1% of circulating red blood cells contain residual RNA, it is unlikely that any of these transcripts are erythrocyte derived, although this was formally addressed by isolating total cellular RNA from 20 mL of whole blood (corresponding to an ~ 3 -log fold excess of erythrocytes than that identified in our final sample). The total cellular yield of RNA from this starting material was approximately 250 ng, suggesting that less than 1 ng erythrocyte-derived RNA was present in the purified platelet preparations. Despite this, however, both α - and β -globin transcripts—along with both the ferritin heavy and light chains—were identified as abundant transcripts (Table 2). Although the most parsimonious explanation would be residual contaminating reticulocytes, this is not supported by our erythrocyte contamination estimates, and their significance remains unresolved.

As a means of better dissecting the molecular anatomy of the platelet, expressed genes were grouped on the basis of assigned

Table 2. Top 50 human platelet-expressed genes

Accession no.	Gene symbol	AD values, range*	Gene transcript†	Leukocyte expression‡
M17733	TMSB4X	140 142-307 852	Thymosin β 4 mRNA, complete cds	+
X99076	NRGN	101 510-148 279	Neurogranin gene	+
M25079	HBB	40 839-229 556	β -globin mRNA, complete cds	+
M25915	CLU	84 720-140 246	Complement cytolysis inhibitor (clusterin) complete cds	-
J04755	FTHP1	82 980-148 621	Ferritin H processed pseudogene, complete cds	-
D78361	OAZ1	73 098-118 140	mRNA for ornithine decarboxylase antizyme	-
X04409	GNAS	77 761-94 781	mRNA for coupling protein G(s) α -subunit (alpha-S1)	-
M25897	PF4	62 811-126 908	Platelet factor 4 mRNA, complete cds	-
AB021288	B2M	61 689-108 921	β 2-microglobulin	+
X00351	ACTB	25 143-73 775	mRNA for β -actin	-
D21261	TAGLN2	76 687-101 931	mRNA for KIAA0120 gene	+
AL031670	FTLL1	69 865-99 966	Ferritin, light polypeptide 1	+
U59632	GPIIB	41 404-110 328	Platelet glycoprotein I β b chain mRNA	-
M21121	CCL5	47 308-106 399	T-cell-specific protein (RANTES) mRNA, complete cds	-
X13710	GPX1	41 318-96 878	Unspliced mRNA for glutathione peroxidase	-
J00153	HBA1	21 326-144 201	Alpha globin gene cluster on chromosome 16	+
M22919	MYL6	46 337-106 833	Nonmuscle/smooth muscle alkali myosin light chain gene	+
L20941	FTH1	52 787-74 763	Ferritin heavy chain mRNA, complete cds	-
J03040	SPARC	51 156-74 261	SPARC/osteonectin mRNA, complete cds	-
X56009	GNAS	45 543-72 096	GSA mRNA for α subunit of GsGTP binding protein	-
X58536	HLA	31 183-82 613	mRNA for major HLA class I locus C heavy chain	+
M54995	PPBP	46 571-67 169	Connective tissue activation peptide III mRNA	-
U34995	GAPD	35 095-70 250	Normal keratinocyte subtraction library mRNA, clone H22a	+
L40399	MLM3	32 107-73 364	Clone zap112 (mutL protein homolog 3) mRNA	-
X77548	NCOA4	31 452-61 036	cDNA for RFG (RET proto-oncogene RET/PTC3)	-
U90551	H2AFL	35 086-51 892	Histone 2A-like protein (H2A/I) mRNA	-
M11353	H3F3A	31 614-55 813	H3.3 histone class C mRNA	-
Z12962	RPL41	36 003-54 853	mRNA for homologue to yeast ribosomal protein L41	+
X06956	TUBA1	20 988-61 798	HALPHA 44 gene for α -tubulin	-
AB028950	TLN1	24 571-58 611	mRNA for KIAA 1027 protein	-
Y12711	PGRMC1	33 680-43 174	mRNA for putative progesterone binding protein	-
M16279	MIC2	30 894-48 166	Integrated membrane protein (MIC2) mRNA	-
D78577	YWHAH	24 785-50 437	Brain 14-3-3 protein β -chain	-
AF070585	TOP3B	20 027-67 945	Clone 24675, unknown cDNA	-
AA524802	Unknown	23 846-39 481	CDNA, IMAGE clone 954213	-
AB009010	UBC	28 745-38 389	mRNA for polyubiquitin UbC	+
X57985	H2AFQ	21 678-52 108	Genes for histones H2B.1 and H2A	-
X54304	MLCB	25 733-34 109	mRNA for myosin regulatory light chain	-
M14539	F13A1	23 691-48 474	Factor XIII subunit α -polypeptide mRNA, 3' end	-
AI540958	Unknown	24 872-41 118	cDNA, PEC 1.2_15_HOI.r 5' end /clon	-
AL050396	FLNA	13 634-55 235	cDNA DKFZp 586K1720	-
X56841	HLA-E	12 890-49 327	Nonclassical MHC class I antigen gene	-
M26252	PKM2	15 450-47 786	TCB (cytosolic thyroid hormone-binding protein)	-
M14630	PTMA	19 314-45 088	Prothymosin alpha mRNA	-
AF045229	RGS10	19 156-34 243	Regulator of G protein signaling 10 mRNA	-
AA477898	Unknown	16 863-44 756	cDNA, Z&34f08.rl 5' end	-
X95404	FL1	15 216-37 456	mRNA for nonmuscle type cofilin	-
M34480	ITGA2B	8 627-45 495	Platelet glycoprotein IIb (GPIIb) mRNA	-
Z83738	H2BFE	18 001-31 306	HH2B/e gene	-
L19779	H2AFO	17 319-38 951	Histone H2A.2 mRNA, complete cds	-

*Gene expression quantifications were calculated as the average difference (AD) value (matched versus mismatched oligonucleotides) for each probe set using Affymetrix GeneChip software, version 4.01. The range of values from 3 distinct platelet microarrays is shown; the normalization value for all microarray analyses was 250.

†Transcripts are rank-ordered (highest to lowest) using BRB-ArrayTools software by log-intensities of AD values obtained from 3 different healthy donors; 33 of the top 40 transcripts were listed among the top 50 in all 3 microarray sets.

‡Leukocyte expression was determined by microarray analysis using purified peripheral blood leukocytes, followed by construction of rank-intensity plots for comparison to platelet top 50 transcripts.²⁰ Top leukocyte-derived transcripts identified within the ranked top 50 platelet transcripts are depicted by a (+) present, or (-) absent.

cds indicates coding sequence.

gene annotations, and this analysis was used to provide a panoramic definition of the platelet transcriptome. Of the genes that could be cataloged within assigned "clusters," those involved in metabolism (11%) and receptor/signaling (11%) represented the largest groups. Also evident in these analyses is the relatively large percentage of genes involved in functions unrelated to these key groups (ie, miscellaneous, 25%), and the overrepresentation of genes with unknown function (32%) as annotated by Affymetrix

and RefSeq databases.²¹ These results identify a vast array (nearly one half) of platelet genes (and gene products) that presumably have important, but poorly characterized functions, in platelet and/or megakaryocyte biology.

Although microarray analysis is not truly quantitative, rank-ordering using the mean log-intensities from 3 independent microarray analyses allowed for the categorization of the top platelet transcripts (Table 2). Computational analyses demonstrated that

only 10 of the top 100 genes were discordant among the 3 platelet microarrays, although 71 of 100 genes were discordant between platelet and leukocyte arrays. An inventory of the top 50 platelet genes is listed in Table 2, which also delineates those found to be highly expressed in peripheral blood leukocytes by parallel microarray experiments with this purified cellular fraction (data not shown). Further analysis of these cell subsets demonstrated that approximately 25% ($n = 547$) of the total platelet transcripts were platelet restricted. Furthermore, only 10 of the 50 most highly expressed genes were found to overlap, confirming the distinct cellular profiles of each transcriptome. Of the 12 overlap genes, 3 corresponded to globin or ferritin chains (again suggesting the presence of contaminating reticulocytes in both purified fractions), and another 4 were involved in actin cytoskeletal reorganization and human leukocyte antigen (HLA) expression, gene products that regulate critical functions in both cell types. Given the importance of cytoskeletal reorganization in downstream platelet activation events, it is not unexpected that components of the actin machinery system would demonstrate prominent transcript expression. Previous estimates suggest that 20% to 30% of the total platelet proteome is comprised of actin with other components such as actin-binding protein, myosin, and talin accounting for an additional 2% to 5% of the total protein.^{1,22} The mRNAs encoding the actin-related machinery are overrepresented in our microarray analysis, with 8 such transcripts found among the 50 highest platelet-expressed genes. Interestingly thymosin β 4 demonstrated the highest expression pattern. In unstimulated platelets, 30% to 40% of actin is polymerized as F-actin,²² whereas the balance of actin monomers (G-actin) are polymerization inhibited by sequestering proteins such as profilin (100 μ M) and thymosin β 4 (600 μ M).²³ The high thymosin β 4 transcript expression not only correlates with its known abundance in platelets but also supports the importance of actin inhibitory proteins in maintaining the nonstimulated state of circulating platelets.

Platelet SAGE analyses

Although these initial studies identified the distribution and relative expression patterns of the genes within the Affymetrix data set, they do not allow for analyses of genes that are unrepresented by these oligonucleotide chips. Unlike closed microarray profiling strategies, SAGE is an open architectural system that is ideally suited for novel gene and pathway identification. Accordingly, the platelet RNA used for microarray studies was used for platelet SAGE. A total of 2033 tags were initially cataloged, of which 1800 (89%) corresponded to mitochondrial-derived genes. These results were quite different from those obtained by microarray analyses, but the discrepancy can be resolved by the nonrepresentation of the mitochondrial genome on the gene chip. The mitochondrial genome is a compact approximately 16.6-kilobase (kb) sequence encoding 13 genes and 2 ribosomal subunits.²⁴ Primary mitochondrial transcripts are polycistronic and typically contain premature termination or unpredictable splice sites, resulting in multiple polyadenylated transcripts from individual genes.^{24,25} Indeed, the overall distribution of platelet-derived mitochondrial SAGE tags is quite similar to that found in muscle.²⁵ All 13 genes containing *Nla*III sites were detected, whereas neither of the non-*Nla*III-containing genes were identified (nicotinamide adenine dinucleotide [NADH] dehydrogenase subunit 4L and adenosine triphosphatase [ATPase] 8). Most of the tags were from the 16S and 12S ribosomal RNAs—which collectively accounted for 68% of the total mitochondrial tags—with the fewest tags represented by NADH dehydrogenase subunits 3, 5, 6, and cytochrome c oxidase 1

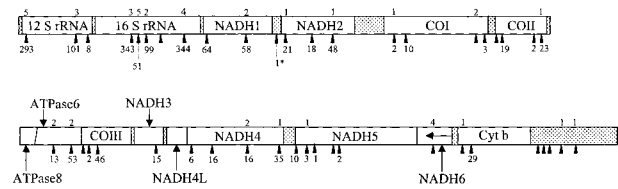


Figure 3. Schema of the mitochondrial genome with SAGE tag distributions (only tags with identical matches are displayed). The abundance of the SAGE tags ($n = 1800$) at individual *Nla*III sites (arrows) within the mitochondrial heavy strand is shown on the bottom, whereas those tags corresponding to the mitochondrial light strand are delineated above the arrows (the presence of an unaccompanied arrow implies no SAGE tags at that *Nla*III site). The gene products of mt-DNA (RefSeq accession no. NC_001807) are delineated by the open rectangles, whereas stippled boxes represent tRNA genes and control regions (the single tag represented by the [*] refers to mitochondrial transfer RNA-serine). Note that NADH6 is encoded by the light strand and that there are no *Nla*III sites within the ATPase8 gene segment. CO[n], cytochrome c oxidase subunit; Cyt b, cytochrome b.

(Figure 3). The NADH dehydrogenase subunit 6 RNA is the only mRNA encoded by the light (L) strand of mitochondrial DNA and was the least abundantly detected transcript.

The unusually high preponderance of mitochondrial-derived genes is not inconsistent with the known enrichment of these genomes in human platelets,^{1,24} and presumably reflects persistent transcription from the mitochondrial (mt) genome in the absence of nuclear-derived transcripts. This overrepresentation of mtDNA in platelets is considerably greater than that of its closest cell type (skeletal muscle), in which mt genomes represent approximately 20% to 25% of all SAGE tags.²⁵ Interestingly, the energy metabolism of platelets is not dissimilar from that of skeletal muscle, both cell types actively using glycolysis and large amounts of glycogen for ATP generation.²⁶ Like muscle, platelets are metabolically adapted to rapidly expend large amounts of energy required for aggregation, granule release, and clot retraction. Similar to the situation in all eukaryotic cells, platelet mitochondria represent the primary source of ATP, which is generated from oxidative phosphorylation reactions occurring within these organelles. Mitochondria are also responsible for most of the toxic reactive oxygen species generated as by-products of oxidative phosphorylation and are central regulators of the apoptotic process in other cellular types. The mtDNA encodes polypeptides found within 4 of the 5 multifunctional complexes that regulate oxidative phosphorylation within the platelet mitochondria.²⁷ Whether the continued generation of these polypeptides has a role in platelet energy metabolism and/or the apoptotic mechanisms regulating platelet survival remains speculative, although not inconsistent with our observations.

Comparative analysis of SAGE and microarray transcript abundance

Complete SAGE libraries require the sequencing of up to 30 000 tags for an exhaustive cataloging of individual mRNAs, especially those with limited copy numbers.^{13,28} Given the preponderance of mt-derived transcripts, comparable sampling would have required sequence analysis of nearly 300 000 SAGE tags, an inordinate number for comprehensive analysis of the platelet transcriptome. For platelets, alternative methodologies incorporating subtractive SAGE will be required for more comprehensive transcript profiling.²⁹ Our initial sampling of nonmitochondrial genes remains informative, however, and entirely consistent with the results of platelet microarray studies. As shown in Table 3, SAGE tags for the genes encoding thymosin β 4, β 2-microglobulin, neurogranin, and the platelet glycoprotein Ib β polypeptide were among the most frequently identified platelet genes, similar to the rank-ordered results determined by microarray analysis. To formally confirm the

Table 3. SAGE-identified nonmitochondrial tags

Frequency	CATG + SAGE tags*	Accession no.†	Gene	Microarray‡
26	GTTGTGGTAACTCTGGT	NM_004048.1	β2-microglobulin (B2M), mRNA	PPP
21	TTGGTGAAGGAAGAAGT	NM_021109.1	Thymosin β4, X chromosome (TMSB4X), mRNA	P
8	AGCTCCGACGCCAGGTC	NM_002620.1	Platelet factor 4 variant 1 (PF4V1), mRNA	p
8	AGCTCCGACGCCGGT	NM_002619.1	Platelet factor 4 (PF4), mRNA	P
7	TGTATAAGACAACCTC	NM_002704.1	Proplatelet basic protein (β-thromboglobulin)	Pp
5	GGGCAACAATGCGGTCCA	NM_000407.1	Glycoprotein I bβ polypeptide, mRNA	P
3	AGGTAATAAAAGGTAAT	NM_003512.1	H2A histone family, member L (H2AFL), mRNA	P
3	AGTGGCAAGTAAATGGC	NM_021914.2	Cofilin 2 (muscle) (CFL2), mRNA	N/A
3	TGACTGTGCTGGGTTGG	NM_006176.1	Neurogranin (protein kinase C substrate, RC3) mRNA	P
3	TTGGGGTTTCCTTTACC	NM_002032.1	Ferritin, heavy polypeptide 1 (FTH1), mRNA	P
2	CCCTTGTGACTACCTAT	NM_025158.1	Hypothetical protein FLJ22251 (FLJ22251), mRNA	N/A
2	CCTGTAACCCAGCTAC	NM_032779.1	Hypothetical protein FLJ14397 (FLJ14397), mRNA	N/A
2	CTTGTAGTCCCAGCTAC	NM_017962.1	Hypothetical protein FLJ20825 (FLJ20825), mRNA	N/A

*Unique tags identified more than once.

†Refers to the RefSeq accession no.²¹ Note that this number does not necessarily correspond to the accession no. provided by Affymetrix software annotations (Table 1).

‡Presence (P) or absence (A) is based on results from 3 distinct platelet microarray experiments. Capitalized "P" designates a gene that is in the top 50 on all 3 microarray experiments, whereas small "p" designates those transcripts not in the top 50. Two of the genes (β2-microglobulin and β-thromboglobulin) are represented by 3 and 2 probe sets, respectively, on the HG-U95Av2 gene chip; for β2-M, all 3 probe sets were in the top 50 genes, whereas for thymosin β4 1 of 2 was in the top 50 for all experiments (the other probe set was in the top 75 for all experiments). N/A indicates oligonucleotide not present on Affymetrix HG-U95Av2 gene chip.

results independently obtained by SAGE and microarray analysis, quantitative RT-PCR was completed by using oligonucleotide primers specific for 2 abundant mitochondrial transcripts, 16S rRNA and NADH2 thymosin β4 (high-abundance by microarray and SAGE), 2 incompletely characterized high-abundance transcripts (neurogranin and clusterin; see "Protein immunoanalysis of platelet clusterin and neurogranin"), a low-abundant transcript (T-cell receptor β-polypeptide), and the genes encoding proteins with well-established quantitative determinations (ie, glycoprotein α_{IIb}β₃ [~50 000 receptors/platelet]; protease-activated receptor-1 (PAR1) [~1800 receptors/platelet]).¹ As shown in Figure 4, these analyses reveal excellent concordance between SAGE and microarray studies, demonstrating the predominant frequency of the mitochondrial-derived 16S rRNA/NADH2 transcripts, with incrementally lower expression of other transcripts as initially demonstrated by microarray (16S > NADH2 > thymosin β4 > neurogranin > clusterin > α_{IIb}β₃ > PAR1 > TCRβ).

Given the small number of nonmitochondrial SAGE tags available for analysis (n = 233), limited conclusions can be drawn using traditional (nonsubtraction) platelet SAGE libraries as pre-

sented here. Overall, a total of 126 unique tags were identified, the majority of which (94) were represented only once. Of the total unique tags, nearly one half represented novel genes not present on the Affymetrix U95Av2 GeneChip. Of the genes with unique tags identified more than once, there was excellent concordance with microarray expression analysis, with nearly all of the SAGE tags in Table 3 corresponding to platelet top 75 microarray transcripts. The platelet factor (PF) 4 variant represents a single aberration because this was rank-ordered approximately 350 by microarray, although its SAGE tag frequency was identical to that of the predominant PF4 transcript. The lack of extensive nonmitochondrial SAGE sampling precludes any further extrapolations from this apparent aberration. Of note, a subset of these tags had long poly(A) tracts, although they all corresponded to genes identified in the RefSeq database.²¹ We cannot exclude the possibility of a SAGE artifact for this small subset of tags (~2%, representing 46 of 2033 tags), although the authenticity of the vast majority of tags (~98%) clearly validates the methodology. These tags are most likely explained by the unique biology of the platelet (ie, mRNA decay in the absence of de novo transcription) or to mRNA degradation occurring during the extensive purification methods. In summary, even with a remarkably limited sampling, the power of this approach in gene identification of relatively abundant and less abundant transcripts is evident. It is clear, however, given the unique molecular anatomy of the platelet (ie, abundance of mitochondrial transcripts), that SAGE adaptations will be required for more comprehensive genetic profiling.²⁹

Protein immunoanalysis of platelet clusterin and neurogranin

Although most of the "most abundant" transcripts would conform to a priori predictions for platelet-expressed mRNAs, a number of transcripts were identified that had been poorly characterized in human platelets. To further establish the authenticity of highly expressed transcripts such as clusterin and neurogranin, confirmatory protein analyses were completed. As shown in Figure 5, both proteins were clearly detected in purified platelet lysates; furthermore, their cellular platelet distributions conformed to those predicted based on previously proposed functions. Note for example that clusterin—functionally characterized as a complement lysis inhibitor able to block the terminal complement reaction—is primarily expressed on the extracellular platelet membrane.³⁰

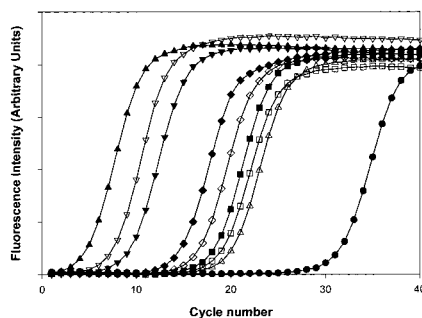


Figure 4. Quantitative real-time RT-PCR analysis of platelet transcripts. Real-time RT-PCR was completed by using purified platelet RNA and oligonucleotide primer pairs specifically designed using Primer3 software to generate similarly-sized (~200-bp) PCR products, optimized to the same annealing temperature. In graph, (□) represents 16S, (▾) represents NADH2, (▿) represents thymosin, (◊) represents clusterin, (◆) represents neurogranin, and (●) represents TCRβ. Curves are representative of one complete set of experiments (repeated twice), and line plots reflect average determinations from 3 wells performed in parallel with SEM less than 1% for all data points.

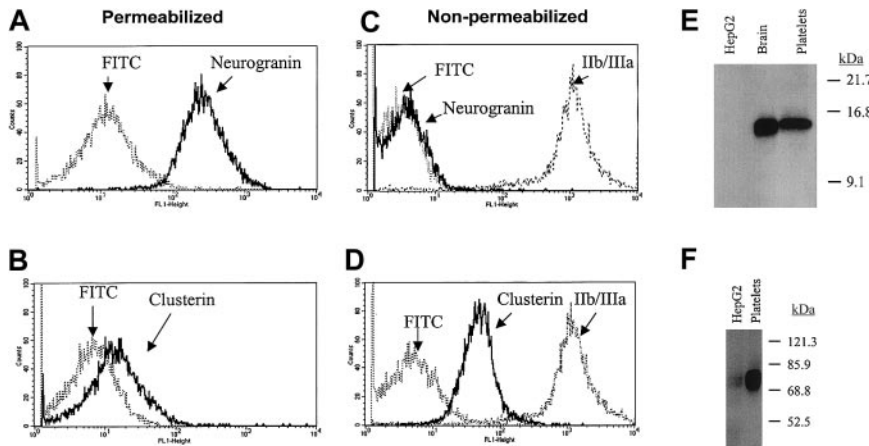


Figure 5. Immunocytometry and immunoblot analysis of platelet neurogranin and clusterin. (A-D) Gel-filtered platelets were either fixed in 3.7% formaldehyde (nonpermeabilized) or fixed with permeabilization in the presence of 0.1% Triton-X, followed by flow cytometric analysis using anticusterin, anti-IIb/IIIa, or antineurogranin antibodies and the FITC-conjugated species-specific secondary antibody (in C, the FITC-conjugated antirabbit and antimouse controls are essentially superimposed). (E-F) Ten micrograms of solubilized HepG2 cells (hepatocyte cell line), human brain, or purified platelet lysates were analyzed by SDS-PAGE,¹⁷ and immunoblot analysis were completed by using 1:1000 dilutions of either antineurogranin (18% SDS-PAGE) or anticusterin (8% SDS-PAGE) antibodies. The anticusterin antibody recognized 2 platelet immunoreactive species under shorter exposure. Although the relative neurogranin and clusterin protein abundances are suboptimally quantified by these analyses, platelet clusterin appears to demonstrate considerable expression when compared with that previously identified in hepatocytes.³¹

Given the importance of complement activation in platelet destruction, the prominent expression of cell-surface clusterin might suggest a role for this protein in normal and pathologic events regulating platelet survival. Interestingly, a clusterin-deficient knockout mouse has been generated that demonstrates enhanced cardiac dysfunction in a model of autoimmune myocarditis.³¹ Although these mice apparently have normal baseline hemograms (B. Aronow, personal communication, October 2002), it remains unestablished if they would be predisposed to immune-type thrombocytopenia in systemic models of autoimmunity.

Similarly, the gene encoding an intracellular effector protein that may have key roles in downstream platelet activation events has now been demonstrated to have abundant transcript expression in human platelets. Neurogranin is a highly expressed platelet transcript with its gene product demonstrating a primarily intracellular pattern of distribution. Neurogranin is generally described as a brain-specific, Ca²⁺-sensitive calmodulin-binding phosphoprotein that is preferentially expressed in neuronal cell bodies and dendrites.^{32,33} It is a specific protein kinase C (PKC) substrate that can also be modified by nitric oxide and other oxidants to form intramolecular disulfide bonds. Both its phosphorylation and oxidation state attenuate its binding affinity for calmodulin.³³ In stimulated platelets, PKC generation is linked to various activation pathways such as calcium-regulated kinases, mitogen-activated protein (MAP) kinases, and receptor tyrosine kinases.¹ Thus, these observations suggest that platelet neurogranin may function as a previously unidentified component of a PKC-dependent activation pathway coupled to one (or more) of these effector proteins.

Discussion

These data provide documentation for a unique platelet mRNA profile that may provide a tool for analyzing platelet molecular networks. Nonetheless, the molecular analysis of the platelet transcriptome may be confounded by the constant decay of mRNAs in the absence of new gene transcription, a situation that may, for example, limit the identification of low-abundance transcripts. Similarly, because the circulating platelet pool contains

a mixed population of variably aged platelets, a "static" mRNA profile represents an average of this heterogeneous blood pool. Despite these potential limitations, the combination of genomic and proteomic technologies are likely to provide powerful tools for the global analysis of platelet function. Current strategies for cataloging "whole cellular proteomes" are generally accomplished by using 2 developing methodologies: (1) high resolution 2-dimensional polyacrylamide gel electrophoresis (2-DE) with mass spectrometric sequence identification,³⁴ and (2) microcapillary liquid chromatography with tandem mass spectrometry (μ LC-MC/MC).³⁵ Further modifications of both procedures have been devised for direct comparative studies between 2 cellular proteomes. The introduction of 2-DE differential gel electrophoresis has now made it possible to detect and quantify differences between experimental sample pairs resolved on the same 2-dimensional gel.³⁶ Likewise, the application of isotope-coded affinity tags to μ LC-MC/MC represent a novel means of quantitative analyses between cellular proteomes.³⁷ The success of both approaches relies on the availability of comprehensive genomic databases and mathematical algorithms for optimal protein identification. Indeed, mathematical modeling studies have demonstrated the need to delineate both protein and mRNA expression levels for optimal definition of intracellular networks.³⁸ Our data present an initial framework for delineating platelet function by defining the molecular anatomy of human platelets, information that is likely to provide important clues into the dynamic protein interactions regulating normal and pathologic platelet functions. Furthermore, because the platelet transcriptome mirrors the mRNAs derived from precursor megakaryocytes, these analyses may provide insights into the biochemical and molecular events regulating megakaryocytopenesis and/or proplatelet formation.

Acknowledgments

We thank Dr Maureen Krause, Jean Wainer, and Lesley Scudder for assistance with some of the experiments; John Schwedes (University DNA microarray facility) with the microarray analysis; and Ms Shirley Murray for manuscript preparation.

References

- Steinberg P, Hill R. Platelets and megakaryocytes. In: Le R, et al, eds. *Wintrobe's Clinical Hematology*. Baltimore, MD: Williams & Wilkins; 1999.
- Newman P, Gorski J, White G, Gidwitz S, Cretney C, Aster R. Enzymatic amplification of platelet-specific messenger RNA using the polymerase chain reaction. *J Clin Invest*. 1988;82:739-743.
- Kieffer N, Guichard J, Farcet J, Vainchenker W, Breton-Gorius J. Biosynthesis of major platelet proteins in human blood platelets. *Eur J Biochem*. 1987;164:189-195.
- Weyrich A, Dixon D, Pabla R, et al. Signal-dependent translation of a regulatory protein, Bcl-2, in activated human platelets. *Proc Natl Acad Sci U S A*. 1998;95:5556-5561.
- Benecke BJ, Ben Ze'ev A, Penman S. The control of mRNA production, translation and turnover in

- suspended and reattached anchorage-dependent fibroblasts. *Cell*. 1978;14:931-939.
6. Pabla R, Weyrich AS, Dixon DA, et al. Integrin-dependent control of translation: engagement of integrin α IIb β 3 regulates synthesis of proteins in activated human platelets. *J Cell Biol*. 1999;144:175-184.
 7. Chicurel ME, Singer RH, Meyer CJ, Ingber DE. Integrin binding and mechanical tension induce movement of mRNA and ribosomes to focal adhesions. *Nature*. 1998;392:730-733.
 8. Lindemann S, Tolley N, Eyre J, Kraiss L, Mahoney T, Weyrich A. Integrins regulate the intracellular distribution of eukaryotic initiation factor 4E in platelets. *J Biol Chem*. 2001;276:33947-33951.
 9. Brown EJ, Schreiber SL. A signaling pathway to translational control. *Cell*. 1996;86:517-520.
 10. Rinder H, Schuster J, Rinder C, Wang C, Schweidler H, Smith B. Correlation of thrombosis with increased platelet turnover in thrombocytosis. *Blood*. 1998;91:1288-1294.
 11. Richards E, Baglin T. Quantitation of reticulated platelets: methodology and clinical application. *Br J Haematol*. 1995;91:445-451.
 12. Robinson M, Mackie I, Khair K, et al. Flow cytometric analysis of reticulated platelets: evidence for a large proportion of non-specific labelling of dense granules by fluorescent dyes. *Br J Haematol*. 1998;100:351-357.
 13. Velculescu V, Zhang L, Vogelstein B, Kinzler K. Serial analysis of gene expression. *Science*. 1995;270:484-487.
 14. Zhang L, Zhou W, Velculescu V, et al. Gene expression profiles in normal and cancer cells. *Science*. 1997;276:1268-1272.
 15. Morin PJ, Sparks AB, Korinek V, et al. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science*. 1997;275:1787-1790.
 16. Bahou W, Campbell A, Wicha M. cDNA cloning and molecular characterization of MSE55: a novel human serum constituent protein that displays bone marrow stromal endothelial cell-specific expression. *J Biol Chem*. 1992;267:13986-13992.
 17. Bahou W, Collier B, Potter C, Norton K, Kutok J, Goligorsky M. The thrombin receptor extracellular domain contains sites crucial for peptide ligand-induced activation. *J Clin Invest*. 1993;91:1405-1413.
 18. Heid C, Stevens J, Livak K, Williams P. Real-time quantitative PCR. *Genome Res*. 1996;6:986-994.
 19. Dunn J, McCorkle S, Praisman L, et al. Genome signature tags (GSTs): a system for profiling genomic DNA. *Nucleic Acid Res*. 2001;29:137-140.
 20. Kroll T, Wolf S. Ranking: a closer look on globalization methods for normalization of gene expression arrays. *Nucleic Acids Res*. 2002;30:e50.
 21. Pruitt KD, Maglott DR. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res*. 2001;29:137-140.
 22. Fox JE, Boyles JK, Reynolds CC, Phillips DR. Actin filament content and organization in unstimulated platelets. *J Cell Biol*. 1984;98:1985-1991.
 23. Safer D, Elzinga M, Nachmias VT. Thymosin beta 4 and Fx, an actin-sequestering peptide, are indistinguishable. *J Biol Chem*. 1991;266:4029-4032.
 24. Wallace DC. Mouse models for mitochondrial disease. *Am J Med Genet*. 2001;106:71-93.
 25. Welle S, Bhatt K, Thornton C. Inventory of high-abundance mRNAs in skeletal muscle of normal men. *Genome Res*. 1999;9:506-513.
 26. Karpatkin S, Charnatz A, Langer RM. Glycogenesis and glyconeogenesis in human platelets. Incorporation of glucose, pyruvate, and citrate into platelet glycogen; glycogen synthetase and fructose-1,6-diphosphatase activity. *J Clin Invest*. 1970;49:140-149.
 27. Raha S, Robinson BH. Mitochondria, oxygen free radicals, and apoptosis. *Am J Med Genet*. 2001;106:62-70.
 28. Yu J, Zhang L, Hwang P, Rago C, Kinzler K, Vogelstein B. Identification and classification of p53-regulated genes. *Proc Natl Acad Sci U S A*. 1999;96:14517-14522.
 29. Wang E, Miller L, Ohnmacht G, Liu E, Marincola F. High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol*. 2000;18:157-159.
 30. Kirszbaum L, Sharpe JA, Murphy B, et al. Molecular cloning and characterization of the novel, human complement-associated protein, SP-40,40: a link between the complement and reproductive systems. *EMBO J*. 1989;8:711-718.
 31. McLaughlin L, Zhu G, Mistry M, et al. Apolipoprotein J/clusterin limits the severity of murine autoimmune myocarditis. *J Clin Invest*. 2000;6:1105-1113.
 32. Martinez DA, Perez JL, Bernal J, Coloma A. Structure, organization, and chromosomal mapping of the human neurogranin gene (NRGN). *Genomics*. 1997;41:243-249.
 33. Wu J, Li J, Huang K, Huang F. Attenuation of PKC and PKA signal transduction in the neurogranin knockout mouse. *J Biol Chem*. 2002;277:19498-19505.
 34. Gygi S, Rochon Y, Franza B, Aebersold R. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*. 1999;19:1720-1730.
 35. Link A, Eng J, Schieltz DM, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol*. 1999;17:676-682.
 36. Unlü M, Morgan M, Minden J. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*. 1997;18:2071-2077.
 37. Gygi S, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*. 1999;17:994-999.
 38. Hatizmanikatis V, Lee K. Dynamical analysis of gene networks requires both mRNA and protein expression information. *Metabol Eng*. 1999;1:275-281.