Platelet Transcriptome Analysis*

Jutta Maria Rox  Jens Müller  Bernd Pötzsch

Institut für Experimentelle Hämatologie und Transfusionsmedizin, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

Key Words
Platelet · Megakaryocyte · RNA · Gene expression

Summary
Despite absence of nuclear DNA, human platelets contain small amounts of megakaryocyte derived mRNA. The development of microarray technology has allowed characterization of the transcripts represented in the platelet transcriptome. Its use, however, is limited by the exceptionally small yield of mRNA in platelets and the risk of leukocyte contamination of platelet preparations. This short review summarizes strategies to overcome these limitations. PCR-based amplification techniques have been successfully used to make available the small amounts of platelet RNA for gene expression profiling. Implementation of leukocyte filtration into the platelet isolation process sufficiently reduces the amount of contaminating leukocytes. The use of both techniques makes platelet transcriptome analysis applicable to single patients. This will improve our understanding of the molecular basis of inherited platelet disorders.

Introduction
Platelets play a critical role in hemostasis and are involved in wound repair and inflammatory processes. Contact to the subendothelial matrix induces platelet activation that is followed by release of cytokines and growth factors and formation of a vascular plug. Several inherited defects of platelet function resulting in bleeding diathesis are known. However, only in the minority of these diseases such as Glanzmann thrombasthenia and Bernard-Soulier syndrome the underlying molecular mechanisms have been characterized [1]. In addition, platelet activation is a main mechanism in the pathogenesis of thrombotic diseases such as stroke or myocardial infarction.
dial infarction. It has been shown that some platelet polymorphism may be associated with quantitatively altered cell surface expression of functionally important glycoproteins such as GPIa/IIa [2]. Although arteriosclerosis is a highly complex heterogeneous disorder with both genetic and environmental risk factors, genetic alterations of platelets may be functionally associated with disease risk. Furthermore, elucidating the molecular mechanisms of platelet activation may lead to new targets for antiplatelet drugs.

Circulating platelets are anucleated but contain small amounts of megakaryocyte derived mRNA [3]. The absence of nuclear DNA implies that they are not able to transcribe nuclear genes or regulate gene expression at the transcriptional level. On the other hand, rough endoplasmic reticulum and polyribosomes are present in platelets, and their ability for protein biosynthesis that can be stimulated by agonists like thrombin has been demonstrated [4]. However, the majority of regulation of platelet function occurs at the posttranslational level. Therefore, the value of characterizing the mRNA in platelets has been questioned, and little is known about platelet mRNAs.

**Gene Expression Analysis**

In an organism the genomes of all cells are identical. Different cells establish their individual identities through unique transcript and protein patterns that mirror their cell function. In opposite to DNA analysis of platelet-specific genes, transcriptome (and proteome) analysis of platelets is in a developmental state. Although major translational activity is lacking in platelets, mRNA analysis is able to indicate which gene products are present. The transcripts that can be identified will have been translated at some point during megakaryocyte/platelet biogenesis and thrombopoiesis. It has been shown that platelet gene expression profiling is able to predict the hitherto unknown presence of proteins in platelets [5]. In contrast to genomic analysis, RNA analysis can be used to investigate posttranscriptional and posttranslational modifications such as alternative splicing that influences protein function. Therefore, genomic, transcriptomic and proteomic approaches are required to elucidate the molecular structure of platelets.

Many techniques have been developed for detecting and quantitating gene expression, including Northern blotting and RT-PCR which are used to detect specific target transcripts in platelets [3, 6]. In contrast, serial analysis of gene expression (SAGE) and DNA microarray represent complementary techniques for cellular mRNA profiling [7]. SAGE is a sequencing based technique used to estimate the relative expression levels of thousand of transcripts by sequencing concatamers of short sequence tags derived from the RNA sample [8]. While SAGE represents an ‘open’ profiling strategy that can be used to identify novel genes, gene expression profiling by microarrays is limited to only those genes present on the gene chip.

Platelet gene expression profiles differ markedly dependent on the method used. Comparing SAGE and microarray platelet gene expression profiles, Gnatenko et al. [9] showed that 89% of the SAGE tags corresponded to mitochondrially derived genes. This can be explained by degradation of the megakaryocyte derived mRNA while the level of mitochondrially encoded transcripts is maintained. Thus, SAGE is a powerful tool for novel gene identification in platelets. Microarrays do not contain mitochondrial genes and therefore can be preferably used to compare the nuclearily transcribed megakaryocyte/platelet gene expression profiles of different platelet populations.

**Microarrays**

Microarrays are devices containing thousands of short DNA probes of specified sequences, spotted on a solid surface such as a glass slide. Each spot on the array corresponds to a particular probe that corresponds to a specified gene. The cDNA synthesized from the RNA sample selectively forms a stable, double stranded molecule with its complementary DNA probe. Compared to the traditional filter and blotting techniques, the advantage of microarray analysis consists of its higher throughput and greater precision [10, 11].

In a microarray experiment (semi)quantitative gene expression information for thousands of genes in parallel is generated, using one- or two-color fluorescent schemes. In one-color formats, expression profiles for each RNA sample are generated on a different array and then the different images are compared. Figure 1 shows a two-color strategy in which two RNA samples are reverse-transcribed into cDNAs and labeled separately with different fluorescent dyes (commonly, cyanine 3 (Cy3) and cyanine 5 (Cy5)). After mixing the samples together they are hybridized against the same microarray. The two populations compete for the same targets on the array. After hybridization and washing the array is scanned to determine spot intensity at the two corresponding wavelengths. Activated or repressed genes are visualized using a two-color graphical overlay. This allows comparisons of normal versus pathologic cell samples.

In a microarray experiment many variables influence the outcome of the experiment. Nucleic acid microarrays use short oligonucleotides (15–25 nucleotides), long oligonucleotides (50–120 nucleotides), or PCR-amplified cDNAs (100–3,000 nucleotides) as array. Short oligonucleotides sometimes lack the specificity required to ensure single-gene specificity in complex hybridizations. Additional probe design issues are able to cause underperformance of the microarray analysis, such as secondary structures within the target sequence or distance of the target sequence from the poly A tail. Variations in pin geometry and in exposure of different portions of the array to the hybridization solution are also able to introduce experimental error [12]. Microarray generated data are trans-
formed and normalized using a variety of software tools, including scatter plots, principal component analysis, cluster diagrams, self-organizing maps and other algorithms [13]. Using current techniques, there is 5–10% variation in signal intensities among replicate array elements on the same microarray and 10–30% variation among corresponding array elements on different microarrays [14]. However, the incompatibility of diverse array formats makes comparisons between different systems difficult [15]. To overcome this difficulty, standards for presenting and exchanging data derived from microarray studies have been introduced [16].

**Microarray Analysis of Platelet Transcripts**

Platelet transcriptome analysis is in a developmental stage due to technical issues related to two main problems: i) leukocyte contamination of platelet preparations and ii) the low level of residual cytoplasmatic mRNA in platelets.

**Leukocyte Depletion Methods**

Although it is not uncommon that tissue specimens used in gene expression analysis contain exogenous cells, contamination of platelet preparations with blood leukocytes has to be taken into consideration because of their disproportionate RNA concentrations. It has been estimated that a single platelet contains 0.2 fg total cellular RNA while a single leukocyte contains about 4-log as much (2–5 pg) [17]. Thus, a small number of contaminating leukocytes can significantly disturb a gene expression profile obtained from a platelet preparation. Current techniques that can be used to reduce the level of contaminating cells include differential centrifugation, filtration, magnetic beads coated with anti-leukocyte antibodies, and laser assisted microdissection [5, 9, 18, 19].

Contamination of platelet RNA by genomic DNA from blood leukocytes can be estimated by PCR amplification of leukocyte specific messages or genomic DNA [9, 18]. Figure 2 shows the content of CD45 transcripts of a platelet sample before and after one leukocyte filtration step measured by quantitative real time PCR. The commonly used leukocyte filtration technique typically achieves a 3-log leukocyte reduction that results in sufficiently pure platelet mRNA for reliable platelet gene expression profiling. More sophisticated methods such as laser assisted microdissection probably result in a more efficient leukocyte depletion. But compared to leukocyte filters commonly used in transfusion medicine their general applicability is markedly lower [20].

**RNA Amplification Strategies**

A limiting factor of the microarray technology is the need for large amounts of RNA per labeling reaction. Typically a minimum of about 20 µg total RNA is required [21]. The exceptionally low RNA concentration in platelets challenges analysis of the platelet transcriptome. Platelet apheresis offers the only chance to obtain sufficient platelet numbers from a single donor for profiling the platelet transcriptome individually. Due to their increased bleeding risk this approach cannot be performed in patients with quantitative or qualitative platelet defects [22]. Techniques that amplify the starting RNA and produce sufficient amounts of cDNA even from samples containing very small amounts of RNA may overcome these limitations. At present, two amplification strategies are used; both include reverse transcription which is commonly followed by linear in vitro transcription [23] or by exponential PCR amplification [24].

The use of an amplification procedure before gene expression analysis requires maintenance of the original message profile. Generally, linear amplification methods are likely to generate less population skewing than exponential amplification methods that can be used when the RNA is extremely limited [25]. To minimize the amount of whole blood needed for platelet gene expression analysis, we extensively investigated an exponential amplification procedure [26]. This technique uses the ability of Moloney’s murine leukemia virus reverse transcriptase to add several deoxyctydine (dC) residues when it reaches the ‘S’ end of the RNA template. First-strand cDNA

<table>
<thead>
<tr>
<th>Table 1. Synopsis of platelet gene expression studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probes</td>
</tr>
<tr>
<td>Platelet source</td>
</tr>
<tr>
<td>Amplification</td>
</tr>
<tr>
<td>RNA sample</td>
</tr>
<tr>
<td>Hybridization sample</td>
</tr>
<tr>
<td>Leukocyte depletion</td>
</tr>
<tr>
<td>Positive genes</td>
</tr>
</tbody>
</table>

Platelet Transcriptome Analysis

Transfus Med Hemother 2006;33:177–182
transcription is primed by an oligo(dexoxythymidine (dT)) primer. A second primer with a short stretch of deoxyguanosine (dG) residues at its 3’ end anneals to the polyC extensions and provides a template for further cDNA extension by the reverse transcriptase, which switches templates and incorporates the complement to the second primer. The known sequences of the oligonucleotides are then used as priming sites for end-to-end cDNA amplification by PCR. It has been shown that template switching ensures the generation of full-length cDNAs compared to conventional second-strand synthesis [27]. When comparing mRNA profiles obtained by microarray analysis of 10 µg of unamplified and 0.5 µg of amplified platelet RNA, we found a strong linear relationship of the corresponding signal intensities (R² = 0.915). Only 35 of 9,850 tested genes (0.4%) revealed discrepant results. Higher sensitivity due to the exponential amplification probably resulted in 27 genes with negative signals in the unamplified samples and positive signals in the amplified samples [26].

When using optimized conditions, even exponential amplification methods are able to preserve the original gene expression profile, making platelet RNA from single patients with platelets disorders available for gene expression studies.

Microarray Studies of Platelets Transcripts

So far three studies using two different commercial oligonucleotide microarray platforms were performed to characterize transcripts present in normal platelets (table 1). In all studies large quantities of platelets either from platelet concentrates [9, 18] or from whole blood [5] were needed to obtain sufficient amounts of platelet RNA. Methods for elimination of nucleated cells from the platelet preparations prior to RNA isolation differed markedly. Residual leukocyte contamination was assessed microscopically [5], using flow cytometry [9] or using PCR [9, 18] which is probably the most accurate method for determination of platelet purity. In two studies, in vitro transcription was performed prior to hybridization to amplify the initial platelet RNA [5, 9]. Despite the differences in starting material, leukocyte depletion methods and microarray techniques, the combined data indicate that human platelets contain a unique message profile, with a lower percentage of expressed transcripts than that obtained from other human cell types in which typically 30–50% of genes are present. This can probably be explained by the lack of ongoing nuclear gene transcription and the large proportion of continuously transcribed mitochondrial RNA in the anucleated platelet. Expressed genes were grouped on the basis of assigned gene annotation. As expected, genes associated with cell communication and signal trans-
Platelet Transcriptome Analysis

Fig. 3. Intra- and inter-individual variability of platelet vWF, F13A, and GAPDH gene expression. Whole blood (40 ml) was drawn weekly over a period of 3 weeks from three healthy volunteers (circles, individual 1; triangles, individual 2; squares, individual 3). After preparation of mRNA, vWF (closed symbols), GAPDH (open symbols), and F13A transcripts (open symbols with shadows) were quantified by RT-PCR. Results are shown as copy numbers (vWF and GAPDH) or units (F13A). Reprinted with permission from Clinical Chemistry 2004;50:2271–2278.

...reduction were over-represented, and genes involved in the protein biosynthesis (ribosomal proteins and translation factors) were under-represented. Genes with unknown function were also over-represented, indicating that many constituents of the platelet transcriptome have hitherto poorly characterized functions in platelet/megakaryocyte biology. The list of the most ‘abundant’ transcripts in human platelets as determined by microarray analysis contains messages for ubiquitous proteins (such as beta-2-microglobulin, beta-actin, ferritin and myosin light chain), those for proteins known to be highly expressed in platelets (such as platelet factor 4 and the cytokine RANTES), and for proteins hitherto not characterized in human platelets (such as neurogranin and clusterin). When comparing the data of two study groups [5, 9] in which the same microarray system was used, a 90% concordance of 50 most abundant platelet transcripts was found, indicating a high reproducibility of array analysis performed in different laboratories.

The findings obtained by microarray analysis were confirmed by RT-PCR for selected transcripts. Bugert et al. [18] performed RT-PCR of genes the signal intensities of which were unexpectedly in the negative range in microarray analysis, such as P-selectin and GPVI. Using RT-PCR, these transcripts were found to be weakly positive, confirming the well-known lower sensitivity of the hybridization based microarray technology compared to gene specific RT-PCR.

McRedmond et al. [5] compared the results of platelet microarray analysis with newly performed and previously published platelet proteomic studies. Despite the assumed platelet RNA decay after mainly megakaryocytic cellular gene transcription, there was a strong relation between transcriptional and proteomic datasets. It has been proven that positive results for platelet messages obtained by microarray analysis may predict the existence of platelet proteins. However, some platelet proteins such as fibrinogen are known to be taken up from plasma and may therefore not have any corresponding message in the platelet [28]. Furthermore, unspecific hybridization or cross-hybridization may result in false-positive microarray results [29]. Presumably, the relation between the transcriptome and proteome of platelets is lower than the relation between the platelet and megakaryocyte gene expression profile. Further studies are required to address the biological relevance of the transcripts identified by platelet gene expression profiling in platelet and megakaryocyte function.

Clinical Applications

Genetic factors have been proposed to explain inter-individual variation in normal platelet function [30]. However, data on intra- and inter-individual variability in platelet gene expression is limited. Figure 3 shows the variability of three platelet transcripts (von Willebrand factor (vWF), A-subunit of coagulation factor XIII (F13A), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH)) in three healthy volunteers over a period of 3 weeks [26]. Inter-individual differences remained stable over time, indicating a low intra-individual variability which is a prerequisite to compare platelet gene expression levels in individuals with phenotypic differences in platelet function.

Thus, it will be possible to include platelet gene expression profiling into clinical diagnostics. Gene profiling in an individual with a phenotypic platelet disorder may identify one or more genes the altered expression level of which accounts for these platelet characteristics. Probably, competitive hybridization with pooled normal platelet RNA will be required to identify specific up- or down-regulated genes (fig. 1). After identification of these genes, the initial gene profiling results have to be confirmed by independent methods. At best, identification of a genomic basis such as polymorphisms within the promoter region of the identified gene makes it possible to predict the platelet characteristics from the genotype. However, many sub-steps of this process are likely to be improved. Gene expression analysis using commercially available microarrays is restricted by cost. It would be preferable to use a platelet specific microarray covering known and/or suspected platelet transcripts than genome-wide microarrays. More efficient and reliable RNA amplification methods such as combined in vitro transcription and PCR based amplification will be developed to further reduce the blood volumes needed for platelet gene expression analysis.

Diseases with a molecular defect that is involved in genetic regulation of megakaryocytogenesis such as essential thrombocythemia and thromboembolic or hemorrhagic disorders.
that are known to be associated with platelet dysfunction are the most obvious pathological conditions to be studied by platelet gene expression profiling.

Conclusions

It has long been known that platelets contain residual megakaryocyte derived RNA transcripts and that these encode for platelet proteins. Analysis of the platelet transcriptome will characterize hitherto unknown genes that are involved in megakaryocyte and platelet function. Basic requirements for platelet gene expression analysis in individual patients are reliable leukocyte depletion and RNA amplification techniques. Comparisons between the platelet gene expression profiles of individuals with a phenotypic platelet disorder and of healthy individuals may lead to identification of the underlying genetic component. The identification of those genes may potentially allow the definition of new molecular markers as diagnostic and/or therapeutic targets.

References