



Platelet genomics and proteomics in human health and disease

Iain C. Macaulay,¹ Philippa Carr,¹ Arief Gusnanto,² Willem H. Ouwehand,^{1,3} Des Fitzgerald,⁴ and Nicholas A. Watkins^{1,3}

¹Department of Haematology, University of Cambridge, Cambridge, United Kingdom. ²Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge, United Kingdom. ³National Blood Service Cambridge, Cambridge, United Kingdom.

⁴Molecular Medicine Laboratory, Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland.

Proteomic and genomic technologies provide powerful tools for characterizing the multitude of events that occur in the anucleate platelet. These technologies are beginning to define the complete platelet transcriptome and proteome as well as the protein-protein interactions critical for platelet function. The integration of these results provides the opportunity to identify those proteins involved in discrete facets of platelet function. Here we summarize the findings of platelet proteome and transcriptome studies and their application to diseases of platelet function.

Introduction

Normal hemostasis balances between pro- and antithrombotic behaviors, with platelets playing a pivotal role. Platelets are involved in maintaining vascular integrity by sensing and responding to endothelial damage. Platelets also have additional roles in wound healing and repair as well as activation of inflammatory and immune responses (1). The central role of the platelet in maintaining hemostasis raises the possibility that small genotypic variations may have dramatic phenotypic effects. Investigation of genetic variants that modulate platelet function will increase understanding of platelet biology and identify risk predictors for diseases involving platelets.

Several diseases resulting from disorders in platelet function are well characterized. In most cases these are rare, monogenic disorders, associated with well-defined phenotypes. Examples include Glanzmann thrombasthenia and Bernard-Soulier syndrome (2). The monogenic nature of these diseases means the underlying molecular defects can be identified via a reductionist approach focused on a single gene or its products. The study of rare allelic variants of these genes has contributed significantly to understanding platelet biology.

It is clear that platelets play a significant role in common diseases — notably in atherothrombosis and coronary artery disease (CAD). Atherothrombosis and CAD are the outcome of a complex interaction between genes and the environment, but it is clear that variations in platelet activity modulate thrombus formation. For example, large-scale clinical trials of oral “antiplatelet” drugs demonstrate that minor variations in platelet function can dramatically increase the risk of myocardial infarction (3). Evidence suggests that single nucleotide polymorphisms are associated with changes in platelet function (4, 5); however, conflicting results have been obtained when these polymorphisms are tested for disease association (6). It is more realistic to consider multiple contributing factors with the subtle effects of numerous genetic polymorphisms

combining to create a range of platelet response within the normal population. To understand how variation at the gene transcription, translation, and protein levels perturbs platelet function from a normal to a prothrombotic phenotype, a more holistic approach is required. Recent technological and methodological advances now make this possible.

Genomics without a genome: the platelet transcriptome

Platelets are produced in the bone marrow from megakaryocytes as cytoplasmic fragments without genomic DNA (7). This renders them incapable of transcription of nuclear material, and platelets were thought to have no synthetic capacity. However, platelets retain a small but functionally significant amount of megakaryocyte-derived RNA as well as the proteins and molecular machinery necessary for translation. Furthermore, platelets can respond to physiological stimuli using biosynthetic processes that are regulated at the level of protein translation (8, 9), demonstrating a functional role for platelet mRNA (10, 11).

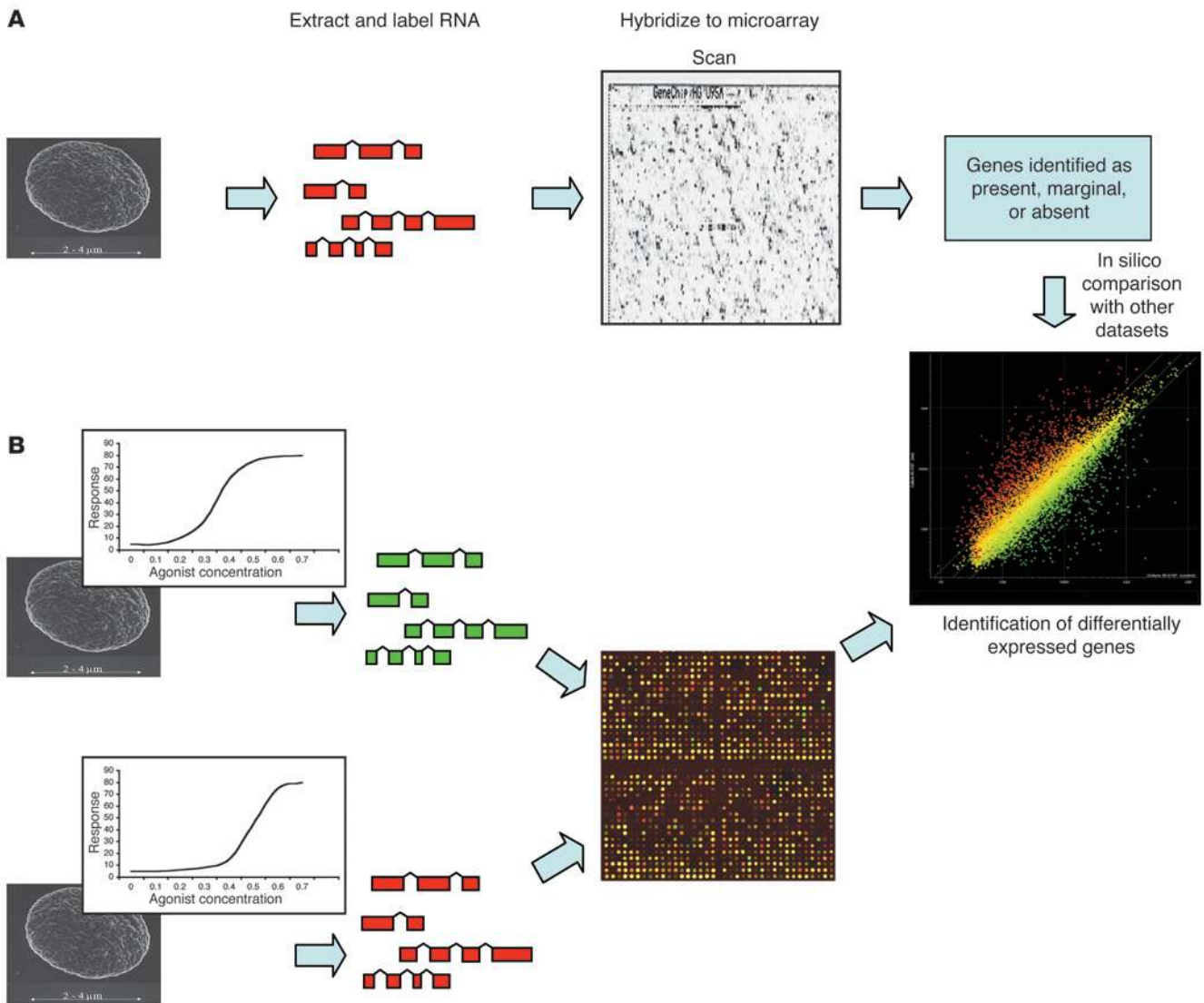
Initial studies using platelet-derived mRNA in the 1980s involved the construction of cDNA libraries (12) that identified many transcripts present in platelets. The development of PCR and its application to platelet biology facilitated the characterization of platelet transcripts and was instrumental in defining many monogenic platelet disorders and the human platelet antigens (2, 13, 14). Recently, transcription profiling methods, such as serial analysis of gene expression (SAGE) and microarray technology, have been applied to more fully characterize transcripts in platelets (15–18).

Microarray technology, the fundamentals of which are described in Figure 1 and reviewed elsewhere (19), represents a rapid, semiquantitative system for gene expression profiling. The number of published microarray studies using platelet-derived mRNA is limited, but those that have been performed suggest that between 15% and 32% of genes studied are present in platelets (16–18) (Table 1). Several well-characterized platelet genes encode the most abundant transcripts (e.g., *glycoprotein Iba*, *glycoprotein IIb*, and *platelet factor 4*). Numerous transcripts from genes involved in cytoskeletal organization were also identified (16–18). Extrapolating from the published studies, and assuming that the arrays have an unbiased coverage of the approximately 22,000 genes in the human genome (20), it could be estimated that as many as 3,000–6,000 transcripts are present in platelets. This figure is sur-

Nonstandard abbreviations used: 2D-GE, 2-dimensional gel electrophoresis; DRM, detergent-resistant membrane; LC, liquid chromatography; MALDI, matrix-assisted laser-desorption ionization; MS, mass spectrometry; MS/MS, tandem MS; PEAR1, platelet endothelial aggregation receptor 1; PV, polycythemia vera; SAGE, serial analysis of gene expression; TOF, time-of-flight.

Conflict of interest: The authors have declared that no conflict of interest exists.

Citation for this article: *J. Clin. Invest.* 115:3370–3377 (2005). doi:10.1172/JCI26885.

**Figure 1**

Platelet transcript profiling by microarray. Generally, 1 of 2 approaches can be used to identify expressed transcripts in platelets. **(A)** With single-channel oligonucleotide arrays, mRNA is isolated from platelets, labeled, and hybridized to the microarray. The array is then processed and scanned, and genes are identified as “present,” “marginal,” or “absent.” Comparisons between samples can then be made in silico to identify differentially expressed genes. **(B)** In a 2-channel experiment, RNA from 2 individuals is isolated, each sample is labeled with a different fluorescent dye, and the 2 are compared directly on a microarray. In this case, the individuals have a different dose-response curve for a platelet agonist. Differentially expressed genes can then be directly identified.

prisingly high considering the limited amount of mRNA in platelets. Estimates suggest that each platelet contains just 0.002 fg mRNA (~12,500-fold less than a nucleated cell [ref. 21]). The low level of RNA increases the potential for interference from contaminating cells, emphasizing the need to obtain pure platelets prior to RNA isolation. A limited overlap with the leukocyte transcriptome (20% of the top 50) has been observed, suggesting that 25% of the detected genes are restricted to platelets (17).

The results of microarray analysis of the platelet transcriptome suggest a significant ontological bias toward metabolism, receptors, and signaling activities, categories that are concordant with known platelet function. Genes involved in the immune response were also present, consistent with the role of the plate-

let in inflammation and immunity (18). In addition, a number of highly expressed transcripts previously unknown in the platelet, such as *neurogranin* (a PKC substrate) and *clusterin* (a complement lysis inhibitor), were identified (17). Recently, reduced platelet expression of *clusterin* mRNA was demonstrated in patients with systemic lupus erythematosus (22); however, the significance of this remains to be determined.

A number of unexpected transcripts were identified in the published platelet microarray studies. In each study, erythroid transcripts were observed; this might suggest reticulocyte contamination, but residual expression from precursor cells remains a possibility. A number of histone transcripts were also present; however, the significance of this remains unclear.



Table 1
Summary of platelet transcript profiling studies

	Study		
	1	2	3
Microarray	HG-U95Av2	Pan 10K	HG-U95Av2
Manufacturer	Affymetrix	MWG Biotech	Affymetrix
No. of genes assayed	12,599	9,850	12,599
No. of positive genes	2,147	1,526	3,978–4,022
Percentage positive	13–17	15.5	31
Reference	(17)	(18)	(16)

SAGE in the platelet

Despite the lack of genomic DNA, the platelet does not appear to be entirely transcriptionally silent. Platelets contain a large number of mitochondria, each of which contains several copies of its own 16-kb circular genome that may be actively transcribed in platelets. The extent of this continued mitochondrial transcription was dramatically demonstrated by Gnatenko et al., who performed SAGE analysis in parallel with microarrays (17). Though lower in throughput than microarray analysis, SAGE is an open and quantitative strategy for transcriptome profiling (23).

When applied to the platelet, 89% of 2,033 SAGE tags were mitochondrial in origin (17). The remaining 11% corresponded to 126 unique genes, half of which were not on the microarray (17). Given this excess of mitochondrial transcripts, more than 300,000 SAGE tags may be required to completely characterize the platelet transcriptome. Subtractive SAGE is an alternative technique that may permit full characterization of the nonmitochondrial transcripts; however, with full genome expression arrays, the development of high-throughput sequencing technologies, and unbiased tiling path expression arrays (24–26), it seems possible that SAGE may be surpassed in future studies (11).

Studies of the platelet transcriptome, as well as that of the megakaryocyte (27, 28), have thus far given a comprehensive overview of the categories of genes important in platelet function. However, transcriptome analysis is not without its limitations. These studies produce long lists of transcripts that are classified as present or provide information on their relative expression levels (16–18) but give no information about localization, interactions, posttranslational modifications, or activation state of gene products. Thus there are several levels of cell function to which transcriptome analysis is oblivious. An integrated approach that identifies novel transcripts and confirms protein expression and functional significance is required.

Proteomics

While the transcriptome of a cell defines those genes that are expressed, it is the encoded proteins that perform the majority of cellular functions. The human genome encodes an estimated 22,000 genes, but the number of functional proteins that can be generated through alternative splicing and posttranslational modifications is estimated to be at least 50 times higher (29). The characterization of a cell's proteome represents a far more complex task than definition of the transcriptome.

Two-dimensional gel electrophoresis (2D-GE) (30), which separates proteins by both size and charge, is the basis of many current proteomic techniques but has been used for many years to study platelet biology (31, 32). Initial studies using reducing and nonreducing gels

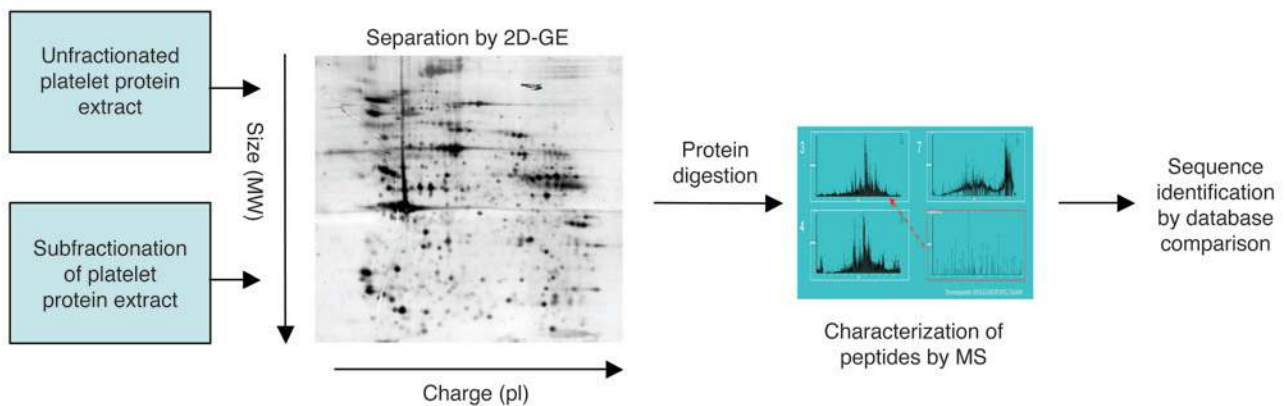
resulted in the naming of many of the platelet glycoproteins and the identification of protein defects in patients with inherited bleeding disorders (2). Non-gel-based separation techniques such as multidimensional liquid chromatography (LC) and Multidimensional Protein Identification Technology (MudPIT) are now replacing 2D-GE. These techniques have the advantages that they can be automated and are able to detect membrane and basic proteins (33, 34). The development and application of electrospray and matrix-assisted laser-desorption ionization (MALDI), which permits the ionization of large biomolecules (35), led to significant advances in proteomic science. These ionization techniques can be applied in combination with mass spectrometry (MS) to study the platelet proteome. MS instruments, such as MALDI-time-of-flight-MS (MALDI-TOF-MS) and complex tandem MS (MS/MS) machines, allow the unambiguous identification of proteins from mixtures, permitting the identification of hundreds rather than tens of platelet proteins.

Global and focused platelet proteomics

The description of the platelet proteome has used 2 general approaches (Figure 2) that involve either the global cataloguing of proteins present in resting platelets (16, 36) or the characterization of “subproteomes,” and changes within them in response to stimulation (37–43). Initial studies of the cytosolic platelet proteome using 2D-GE over a wide pI range (pI 3–10) identified 186 proteins, a number of which were phosphorylated in resting platelets (44). This represented a 10-fold increase over the number found in previous studies that had not used mass spectrometry.

More recent studies have shifted away from global profiling to the analysis of subfractions of the proteome and the identification of changes induced upon platelet activation (37–43). These focused studies allow the identification of many more platelet proteins than can be achieved by global profiling, giving a more complete view of the platelet proteome. For example, a study of the platelet proteome using a narrow, acidic pI range (pI 4–5) identified the protein products of 123 different genes (36). Extension of the same study into the pI 5–11 range identified an additional 760 protein features representing 311 different genes (38). Despite significant effort, this only represented the characterization of 54% of the pI 5–11 2D-GE proteome. These combined studies highlight the abundance of signaling (24%) and cytoskeletal proteins (15%) present in the platelet (38), mirroring the results obtained with microarrays and also identify several proteins not previously reported in platelets, such as the adapter molecule SH3GL1, an SH3 domain-containing Grb2-like protein, and RSU1, a regulator of G protein signaling. In addition, 15 predicted proteins were identified, demonstrating the utility of this approach in the identification of novel platelet proteins. Most recently, novel non-gel-based protein separation techniques (45, 46) have been applied to the study of the whole platelet proteome (47). This study identified 641 platelet proteins, 404 of which were novel or not previously described in the platelet; thus, this represents the largest catalogue of platelet proteins to date.

The platelet proteome can be broken down further by subcellular prefractionation prior to LC or 2D-GE (Figure 2), allowing the detection of low-abundance proteins that are masked in whole-cell analysis. These studies have focused on cytoskeleton-associated (42) and detergent-resistant membrane-associated (DRM-associated) proteins (48). Following platelet activation, several proteins, including tropomyosin, myosin, and caldesmon, translocate to the cytoskeleton through interactions with F-actin. A study by Gevaert et al. identified additional proteins that translocate to the cytoskeleton

**Figure 2**

Analyzing the platelet proteome. 2D-GE has been extensively applied to the characterization of the platelet proteome. Published studies have analyzed the platelet proteome as a whole (36, 38, 44) or in subfractions (16, 40, 42). Proteins are separated on the gel by charge (pI) and molecular weight (MW). The protein spots are then digested with a proteolytic enzyme, typically trypsin, and the resultant peptides are characterized by mass spectrometry to generate partial sequence information for the protein. This information can then be used to identify the protein by comparison of the sequence with databases of known protein and nucleic acid sequences.

upon activation, including a number of actin-related proteins, heat shock proteins, and coronin-like protein (42). DRMs are sphingolipid- and cholesterol-rich domains in the plasma membrane (49) where receptors and downstream signaling molecules are highly concentrated. Several receptors on the surface of platelets show enhanced signal transduction when associated with DRMs (50). It is anticipated that the proteomic analysis of DRMs will identify proteins that are recruited to the rafts upon platelet activation (48).

Functional platelet proteomics

The platelet can rapidly respond to a variety of agonists, and the effect of these on the platelet proteome has been studied. Like many other cells, platelets secrete proteins from preformed storage granules in response to stimuli (51). Analysis of the secretome from thrombin-activated platelets has identified over 300 proteins that are secreted upon activation (41). A number of novel proteins, including secretogranin III, a monocyte chemoattractant precursor, were identified that may represent a group of proteins that mediate functions secondary to blood clot formation. In addition, a number of the secreted proteins have been identified in atherosclerotic lesions, suggesting a potential role in atherothrombosis (41).

Posttranslational modifications

Many signaling pathways in platelets are regulated by protein phosphorylation (52), and recent proteomic studies have focused on the identification of proteins that are differentially phosphorylated upon platelet activation (37, 40). Early studies identified phosphorylated proteins in thrombin-activated platelets (43, 44, 53), and enrichment for phosphotyrosine-containing proteins by immunoprecipitation detected 67 proteins that were differentially phosphorylated between resting and thrombin-activated platelets. Through the application of MALDI-TOF, the identity of 10 of these, including FAK and SYK, was confirmed (40). In a separate study that compared the unfractionated platelet proteomes from resting platelets and platelets activated by thrombin receptor-activating peptide (TRAP), 62 differentially regulated protein features were identified (37). Characterization of the phosphorylated targets in TRAP-activated platelets led to the

identification of a number of novel platelet proteins, such as the adapter protein downstream of tyrosine kinase-2 (DOK2), which is phosphorylated downstream of inside-out signaling through $\alpha_{IIb}\beta_3$ and also after glycoprotein VI activation. Additionally, activation-dependent phosphorylation of RGS10 and RGS18, 2 regulators of G protein signaling, was demonstrated (37). These proteins were not thought to undergo any posttranslational phosphorylation.

Tyrosine phosphorylation is just one posttranslational modification that is important for platelet function. Glycosylation, another key modification, has been studied in different cell types (54, 55), and the application of glycoproteomic analysis to the platelet will be of interest.

Other future directions include the application of isotope-coded affinity tags, which allows quantitative differential proteomics (56) that will aid the identification of proteins expressed differentially between phenotypically different platelets.

Integration of transcriptome and proteome data

The integration of proteomic and genomic studies is critical, but integrating data from different platforms is challenging. The correlation of protein and transcript levels is weak in nucleated cells (57) and poorly defined in platelets. McRedmond et al. performed a qualitative correlation of their platelet transcript profile with the proteome by identifying a confirmed set of 82 proteins secreted by thrombin-activated platelets (16). Seventy of these 82 proteins were represented on the microarray, and messages corresponding to 69% were detected. In the same study, a comparison of the transcript profile with those found in previously published platelet proteomics studies showed similar levels of correlation with proteins in the pI range 4–5 (68%) and tyrosine-phosphorylated platelet proteins (69%) (16). Thus, while it appears that transcript and protein abundance may correlate poorly, presence of transcript appears to be associated with presence of protein.

There is a clear discrepancy between the numbers of expressed genes identified by transcriptomics and confirmation of the protein in platelets. Given the diversity of proteins generated by posttranslational modification, coupled with the lack of tran-

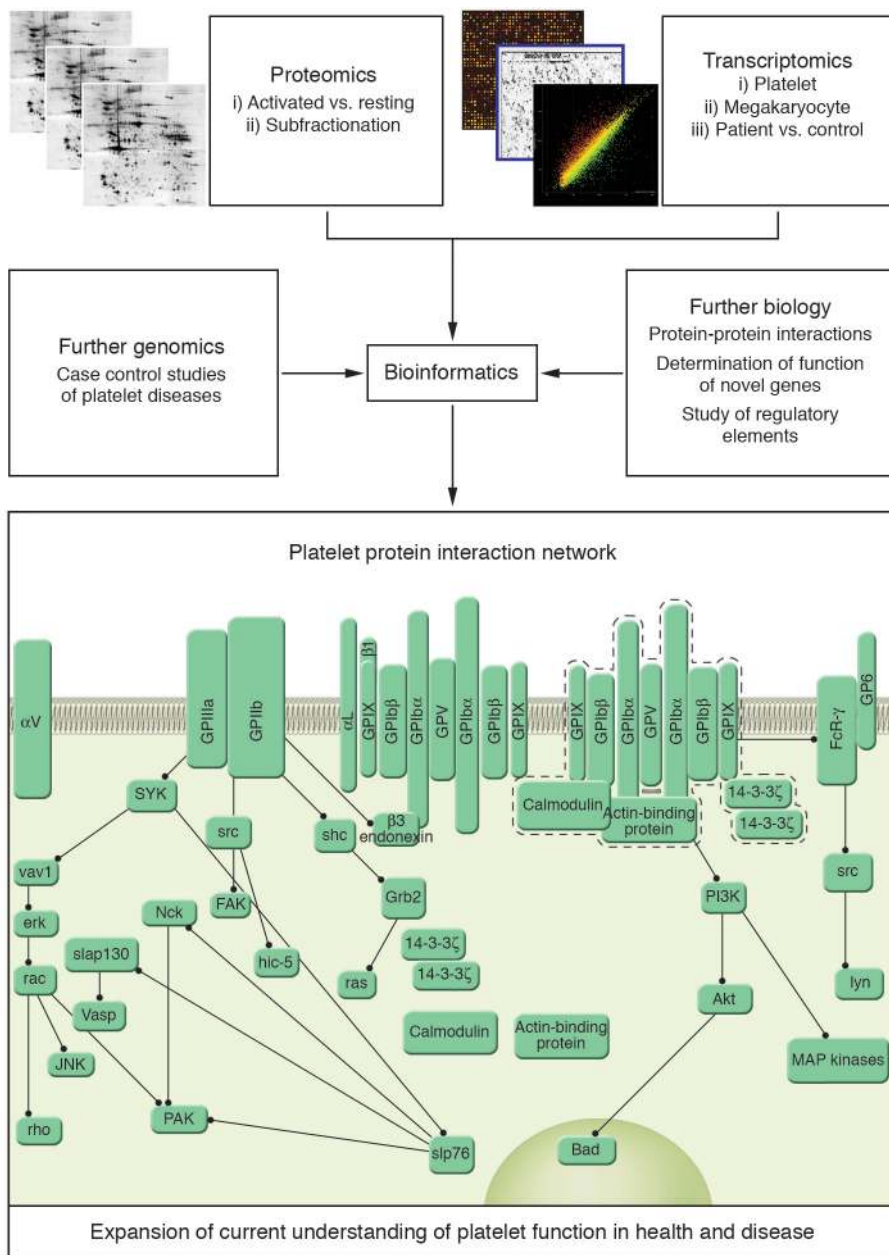


Figure 3

Integration of proteome and transcriptome data. Global studies of the platelet transcriptome and proteome will allow complete characterization of the platelet system. The functional significance of changes in both proteome and transcriptome can be considered by comparison of healthy and diseased individuals, as well as those individuals with extreme platelet phenotypes. Further genomics, involving case-control genotyping studies, will permit determination of the genetic components that underlie normal variation in platelet function. Finally, bioinformatic integration of these approaches, along with the study of protein function using model organisms or *in vitro* examination of protein interactions, will expand upon current models of platelet function in health and disease. This map was generated using CellDesigner version 2.0 (<http://www.systems-biology.org/index.html>).

scription in the platelet, one might expect more proteins than transcripts in the platelet. However, current experimental evidence runs contrary to this, with tens or hundreds of proteins being identified in platelet proteomic studies but thousands of transcripts being detected. To a large extent, this can be explained by the limitations of proteomic analysis; however, the application of novel techniques should reduce this discrepancy by increasing the number of identified proteins.

The recent study of Nanda et al. has applied both oligonucleotide microarrays and phosphoproteome analysis to identify and characterize a novel platelet transmembrane receptor (58). Using bioinformatic analysis of the platelet transcript profile, proteins with predicted transmembrane domains were identified. In a parallel proteomic study, phosphoproteins were purified from aggregated platelets by affinity chromatography, separated using

1-dimensional gel electrophoresis, and identified using LC/MS/MS. Platelet endothelial aggregation receptor 1 (PEAR1), which signals secondary to platelet-platelet interactions, was identified using both approaches. This study highlights how proteomics and transcriptomics can be combined to identify novel proteins with a role in platelet function. A combined analysis of published transcriptomic and proteomic datasets could lead to the identification of additional novel proteins present in platelets. However, such an analysis is not trivial, as data can be in different formats and access to raw data can be restricted. The submission of data to public databases at the time of publication will facilitate the integration of datasets and ensure that standardized documentation detailing the experimental approach is available. In the case of microarray data, public databases exist and there are clear guidelines, published by the Microarray Gene Expression Data



Society (www.mged.org), about the minimal data required for submission (59). For proteomic data, the guidelines and databases are currently under development.

The collation of proteomic and transcriptomic data will also allow a more complete characterization of the platelet through the identification of novel regulatory mechanisms and the construction of protein interaction networks (Figure 3). The cataloguing of large numbers of genes and proteins associated with the platelet lineage may also permit in silico study of regulatory elements responsible for megakaryocyte maturation and platelet formation. Software such as TFBScluster (60), which has been used to identify *cis*-regulatory sequences controlling blood and endothelial development (61), could also be applied to megakaryocyte-restricted genes to identify specific transcription factor signatures. Furthermore, novel bioinformatics approaches that visualize complex biochemical networks will expand current models of hemostasis to include proteins and processes identified in transcriptomic and proteomic studies (62).

Platelet proteomics and genomics in complex disease

Given the availability of whole-genome transcriptome platforms, robust RNA amplification techniques (15), and novel proteomics technologies, platelet biologists are now in a position to begin characterizing the full complement of transcripts and functionally relevant protein fractions in an individual's platelets. This allows a shift in focus away from determining the components of the platelet system to asking fundamental questions about the mechanisms that determine interindividual variation in platelet function. One of the difficulties in analyzing transcriptomic and proteomic data is the sheer volume of data generated from a single experiment. Comparative studies reduce the complexity of results by allowing the researcher to focus in on differentially expressed genes and proteins. Such studies, which are in their infancy, require large populations of individuals with well-defined clinical phenotypes and standardized protocols. The importance of this is highlighted by a comparison of 2 published microarray studies (16, 17), which used identical microarray platforms. While the most abundant transcripts in both experiments appear similar, there is an almost 2-fold difference in the number of transcripts identified as "present" (Table 1) because of differences in sample preparation and analysis parameters. In the study by McRedmond et al. (16), the analysis accounted for the low signal intensities obtained when platelet RNA was analyzed using microarrays.

The application of microarrays to the study of complex disease requires many statistical considerations (reviewed in Chui and Churchill, ref. 63). The development and application of sound statistical methods for the analysis of microarray data is critical to obtaining meaningful results (64, 65). Many microarray-based studies are underpowered and lack sufficient technical and biological replication to infer statistically significant datasets.

Genomics and proteomics in hematological disease

Transcript profiling is being used to investigate a number of platelet disorders, although no published studies have focused directly on the platelet. Comparative transcript profiling of neutrophil RNA from gray platelet syndrome (GPS) sufferers and healthy individuals showed that genes involved in the biosynthesis of cytoskeleton proteins are upregulated in GPS (66). However, the significance of these findings remains to be elucidated. Microarrays have also been applied to study myeloproliferative disorders such

as polycythemia vera (PV) and essential thrombocythemia (ET) (67–70). These acquired disorders are clonal HSC malignancies characterized by hypersensitivity to numerous cytokines. Transcript profiling of megakaryocytes differentiated in vitro from HSCs suggests that megakaryocytes from ET patients are more resistant to apoptosis. In this study (68), proapoptotic genes, such as *Bcl-2-associated X protein (BAX)* and *Bcl-2/adenovirus E1B 19-kDa-interacting protein 3 (BNIP3)*, were downregulated in patients. Similarly, transcript profiling of samples from PV patients has identified a number of genes that are over- or underexpressed in these patients relative to controls. Transcript profiling of granulocyte RNA identified 253 genes that were upregulated and 391 that were downregulated relative to controls. In a similar study (70), antiapoptotic and survival factors were upregulated in granulocytes from PV patients, suggesting that these factors may promote cell survival. Recently, an acquired mutation in the JH2 pseudokinase domain of the *JAK2* gene has been identified in more than 80% of PV patients (71–73). This discovery was made using 3 different but complementary techniques, none of which used transcriptome or proteome technologies. It is highly unlikely that the 1849G→T mutation, encoding a V617F substitution, would have been identified directly by transcript profiling; however, the effect of a constitutively active kinase that is directly linked to cytokine signaling should be detectable by expression profiling. It would be interesting to reconsider the transcript profiling results in light of the identification of the *JAK2* mutation. That the studies identified a signature rather than the cause of the disease highlights the fact that the identification of differential expression is not enough. To be truly informative, such studies must consider not only changes in transcript abundance, but also the underlying causes of these changes. It is more likely that complex diseases will be characterized by the integration of transcriptomic, proteomic, and genomic studies.

Conclusions

Advances in technology and methodology of high-throughput experimental approaches, such as genomics, proteomics, and bioinformatics, provide an opportunity to describe the whole platelet system in health and disease. These technologies currently provide a limited catalogue of transcripts and proteins, but the application of additional experimental approaches, including the use of model organisms, may eventually allow the construction of a platelet "molecular interaction map." The ability to construct such maps is already contributing to our understanding of basic biology in model organisms (74–76), and such approaches have recently been applied to mammalian systems (77). This integrated understanding of platelet biology will allow a more complete characterization of the role of platelets in complex diseases and provide unique phenotypes for the discovery of genetic traits. The catalogues of genes that are being developed for the platelet and the megakaryocyte are the first step toward the identification of genetic variants that play a role in determining both interindividual variation in platelet response and disease risk. These insights will be gained from candidate gene and whole-genome genotyping studies of individuals with well-characterized platelet phenotypes.

While the identification of a "molecular signature" of prothrombotic platelet phenotypes, analogous to those identified by the application of proteomic and genomic studies of various cancers (78, 79), may seem fanciful, it would be careless to dismiss such a possibility. It is inevitable that the cost of these technologies will decrease as they become more widely applied. This will allow large-



er transcriptome and proteome studies to be carried out, which will, in turn, permit a detailed characterization of the variation of platelet gene expression within a large population of individuals, and thus investigation of the association, if any, between platelet phenotype and gene expression.

The development of diagnostic or therapeutic tools based, either directly or indirectly, on the application of proteomic and genomic analysis of platelets is still a long way off. However, these technologies are being applied more and more widely in the field of platelet

biology. The extent to which such approaches will affect either the diagnosis or the treatment of platelet-associated diseases remains to be seen; however, it seems their impact on our understanding of basic platelet biology will be profound.

Address correspondence to: Nicholas A. Watkins, Department of Haematology, University of Cambridge, Long Road, Cambridge CB2 2PT, United Kingdom. Phone: 44-1223-548101; Fax: 44-1223-548136; E-mail: naw23@cam.ac.uk.

1. Weyrich, A.S., and Zimmerman, G.A. 2004. Platelets: signaling cells in the immune continuum. *Trends Immunol.* **25**:489–495.
2. Nurden, A.T., and Nurden, P., editors. 2002. *Inherited disorders of platelet function*. Academic Press. London, United Kingdom. 681–700.
3. Shields, D.C., et al. 2002. The contribution of genetic factors to thrombotic and bleeding outcomes in coronary patients randomised to IIb/IIIa antagonists. *Pharmacogenomics J.* **2**:182–190.
4. Hetherington, S.L., et al. 2005. Dimorphism in the P2Y1 ADP receptor gene is associated with increased platelet activation response to ADP. *Arterioscler. Thromb. Vasc. Biol.* **25**:252–257.
5. Joutsu-Korhonen, L., et al. 2003. The low-frequency allele of the platelet collagen signalling receptor glycoprotein VI is associated with reduced functional responses and expression. *Blood.* **101**:4372–4379.
6. Casas, J.P., Hingorani, A.D., Bautista, L.E., and Sharma, P. 2004. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch. Neurol.* **61**:1652–1661.
7. Italiano, J.E., Jr., and Shivdasani, R.A. 2003. Megakaryocytes and beyond: the birth of platelets. *J. Thromb. Haemost.* **1**:1174–1182.
8. Kieffer, N., Guichard, J., Farcet, J.P., Vainchenker, W., and Breton-Gorius, J. 1987. Biosynthesis of major platelet proteins in human blood platelets. *Eur. J. Biochem.* **164**:189–195.
9. Booyse, F.M., and Rafelson, M.E., Jr. 1968. Studies on human platelets. I. Synthesis of platelet protein in a cell-free system. *Biochim. Biophys. Acta.* **166**:689–697.
10. Weyrich, A.S., et al. 1998. Signal-dependent translation of a regulatory protein, Bcl-3, in activated human platelets. *Proc. Natl. Acad. Sci. U. S. A.* **95**:5556–5561.
11. Denis, M.M., et al. 2005. Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets. *Cell.* **122**:379–391.
12. Wicki, A.N., et al. 1989. Isolation and characterization of human blood platelet mRNA and construction of a cDNA library in lambda gt11. Confirmation of the platelet derivation by identification of GPIb coding mRNA and cloning of a GPIb coding cDNA insert. *Thromb. Haemost.* **61**:448–453.
13. Metcalfe, P., et al. 2003. Nomenclature of human platelet antigens. *Vox Sang.* **85**:240–245.
14. Newman, P.J., et al. 1988. Enzymatic amplification of platelet-specific messenger RNA using the polymerase chain reaction. *J. Clin. Invest.* **82**:739–743.
15. Rox, J.M., et al. 2004. Gene expression analysis in platelets from a single donor: evaluation of a PCR-based amplification technique. *Clin. Chem.* **50**:2271–2278.
16. McRedmond, J.P., et al. 2004. Integration of proteomics and genomics in platelets: a profile of platelet proteins and platelet-specific genes. *Mol. Cell. Proteomics.* **3**:133–144.
17. Gnatenko, D.V., et al. 2003. Transcript profiling of human platelets using microarray and serial analysis of gene expression. *Blood.* **101**:2285–2293.
18. Bugert, P., Dugrillon, A., Gunaydin, A., Eichler, H., and Kluter, H. 2003. Messenger RNA profiling of human platelets by microarray hybridization. *Thromb. Haemost.* **90**:738–748.
19. Holloway, A.J., van Laar, R.K., Tothill, R.W., and Bowtell, D.D. 2002. Options available—from start to finish—for obtaining data from DNA microarrays II. *Nat. Genet.* **32**(Suppl.):481–489.
20. International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature.* **431**:931–945.
21. Fink, L., et al. 2003. Characterization of platelet-specific mRNA by real-time PCR after laser-assisted microdissection. *Thromb. Haemost.* **90**:749–756.
22. Wang, L., et al. 2004. Transcriptional down-regulation of the platelet ADP receptor P2Y(12) and clusterin in patients with systemic lupus erythematosus. *J. Thromb. Haemost.* **2**:1436–1442.
23. Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science.* **270**:484–487.
24. Reinartz, J., et al. 2002. Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. *Brief Funct. Genomic. Proteomic.* **1**:95–104.
25. Brenner, S., et al. 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* **18**:630–634.
26. Bertone, P., et al. 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science.* **306**:2242–2246.
27. Kim, J.A., et al. 2002. Gene expression profile of megakaryocytes from human cord blood CD34(+) cells ex vivo expanded by thrombopoietin. *Stem Cells.* **20**:402–416.
28. Shim, M.H., Hoover, A., Blake, N., Drachman, J.G., and Reems, J.A. 2004. Gene expression profile of primary human CD34+CD38lo cells differentiating along the megakaryocyte lineage. *Exp. Hematol.* **32**:638–648.
29. Hochstrasser, D.F., Sanchez, J.C., and Appel, R.D. 2002. Proteomics and its trends facing nature's complexity. *Proteomics.* **2**:807–812.
30. Gorg, A., et al. 2000. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis.* **21**:1037–1053.
31. Jenkins, C.S., et al. 1976. Platelet membrane glycoproteins implicated in ristocetin-induced aggregation. Studies of the proteins on platelets from patients with Bernard-Soulier syndrome and von Willebrand's disease. *J. Clin. Invest.* **57**:112–124.
32. Gravel, P., et al. 1995. Human blood platelet protein map established by two-dimensional polyacrylamide gel electrophoresis. *Electrophoresis.* **16**:1152–1159.
33. Washburn, M.P., Wolters, D., and Yates, J.R., 3rd. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **19**:242–247.
34. Link, A.J., et al. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **17**:676–682.
35. Pandey, A., and Mann, M. 2000. Proteomics to study genes and genomes. *Nature.* **405**:837–846.
36. O'Neill, E.E., et al. 2002. Towards complete analysis of the platelet proteome. *Proteomics.* **2**:288–305.
37. Garcia, A., et al. 2004. Differential proteome analysis of TRAP-activated platelets: involvement of DOK-2 and phosphorylation of RGS proteins. *Blood.* **103**:2088–2095.
38. Garcia, A., et al. 2004. Extensive analysis of the human platelet proteome by two-dimensional gel electrophoresis and mass spectrometry. *Proteomics.* **4**:656–668.
39. Marcus, K., and Meyer, H.E. 2004. Two-dimensional polyacrylamide gel electrophoresis for platelet proteomics. *Methods Mol. Biol.* **273**:421–434.
40. Maguire, P.B., et al. 2002. Identification of the phosphotyrosine proteome from thrombin activated platelets. *Proteomics.* **2**:642–648.
41. Coppinger, J.A., et al. 2004. Characterization of the proteins released from activated platelets leads to localization of novel platelet proteins in human atherosclerotic lesions. *Blood.* **103**:2096–2104.
42. Gevaert, K., Eggermont, L., Demol, H., and Vandekerckhove, J. 2000. A fast and convenient MALDI-MS based proteomic approach: identification of components scaffolded by the actin cytoskeleton of activated human thrombocytes. *J. Biotechnol.* **78**:259–269.
43. Immler, D., et al. 1998. Identification of phosphorylated proteins from thrombin-activated human platelets isolated by two-dimensional gel electrophoresis by electrospray ionization-tandem mass spectrometry (ESI-MS/MS) and liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS). *Electrophoresis.* **19**:1015–1023.
44. Marcus, K., Immler, D., Sternberger, J., and Meyer, H.E. 2000. Identification of platelet proteins separated by two-dimensional gel electrophoresis and analyzed by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and detection of tyrosine-phosphorylated proteins. *Electrophoresis.* **21**:2622–2636.
45. Gevaert, K., et al. 2003. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat. Biotechnol.* **21**:566–569.
46. Gevaert, K., et al. 2004. Reversible labeling of cysteine-containing peptides allows their specific chromatographic isolation for non-gel proteome studies. *Proteomics.* **4**:897–908.
47. Martens, L., et al. 2005. The human platelet proteome mapped by peptide-centric proteomics: a functional protein profile. *Proteomics.* **5**:3193–3204.
48. Maguire, P.B., Foy, M., and Fitzgerald, D.J. 2005. Using proteomics to identify potential therapeutic targets in platelets. *Biochem. Soc. Trans.* **33**:409–412.
49. Simons, K., and Ikonen, E. 1997. Functional rafts in cell membranes. *Nature.* **387**:569–572.
50. Wonerow, P., et al. 2002. Differential role of glycolipid-enriched membrane domains in glycoprotein VI- and integrin-mediated phospholipase Cgamma2 regulation in platelets. *Biochem. J.* **364**:755–765.
51. Reed, G.L., editor. 2002. *Platelet secretion*. Academic Press. London, United Kingdom. 181–195.
52. Santos, M.T., et al. 2000. Participation of tyrosine phosphorylation in cytoskeletal reorganization, alpha(IIb)beta(3) integrin receptor activation, and aspirin-insensitive mechanisms of thrombin-stimulated human platelets. *Circulation.* **102**:1924–1930.
53. Marcus, K., Moebius, J., and Meyer, H.E. 2003. Differential analysis of phosphorylated proteins in



- resting and thrombin-stimulated human platelets. *Anal. Bioanal. Chem.* **376**:973–993.
54. Zhang, H., Li, X.J., Martin, D.B., and Aebersold, R. 2003. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat. Biotechnol.* **21**:660–666.
55. Kaji, H., et al. 2003. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat. Biotechnol.* **21**:667–672.
56. Gygi, S.P., et al. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**:994–999.
57. Hack, C.J. 2004. Integrated transcriptome and proteome data: the challenges ahead. *Brief Funct. Genomic. Proteomic.* **3**:212–219.
58. Nanda, N., et al. 2005. Platelet-endothelial aggregation receptor 1 (PEAR1), a novel epidermal growth factor repeat-containing transmembrane receptor, participates in platelet contact-induced activation. *J. Biol. Chem.* **280**:24680–24689.
59. Brazma, A., et al. 2001. Minimum information about a microarray experiment (MIAME): toward standards for microarray data. *Nat. Genet.* **29**:365–371.
60. Donaldson, I.J., Chapman, M., and Gottgens, B. 2005. TFBScluster: a resource for the characterization of transcriptional regulatory networks. *Bioinformatics.* **21**:3058–3059.
61. Donaldson, I.J., et al. 2005. Genome-wide identification of cis-regulatory sequences controlling blood and endothelial development. *Hum. Mol. Genet.* **14**:595–601.
62. Joshi-Tope, G., et al. 2005. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* **33**:D428–D432.
63. Cui, X., and Churchill, G.A. 2003. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol.* **4**:210.
64. Culhane, A.C., Perriere, G., and Higgins, D.G. 2003. Cross-platform comparison and visualisation of gene expression data using co-inertia analysis. *BMC Bioinformatics.* **4**:59.
65. Culhane, A.C., Thioulouse, J., Perriere, G., and Higgins, D.G. 2005. MADE4: an R package for multivariate analysis of gene expression data. *Bioinformatics.* **21**:2789–2790.
66. Hyman, T., et al. 2003. Use of a cDNA microarray to determine molecular mechanisms involved in grey platelet syndrome. *Br. J. Haematol.* **122**:142–149.
67. Goerttler, P.S., et al. 2005. Gene expression profiling in polycythaemia vera: overexpression of transcription factor NF-E2. *Br. J. Haematol.* **129**:138–150.
68. Tenedini, E., et al. 2004. Gene expression profiling of normal and malignant CD34-derived megakaryocytic cells. *Blood.* **104**:3126–3135.
69. Pellagatti, A., et al. 2004. Gene expression profiling in the myelodysplastic syndromes using cDNA microarray technology. *Br. J. Haematol.* **125**:576–583.
70. Pellagatti, A., et al. 2003. Gene expression profiling in polycythemia vera using cDNA microarray technology. *Cancer Res.* **63**:3940–3944.
71. Baxter, E.J., et al. 2005. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet.* **365**:1054–1061.
72. James, C., et al. 2005. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature.* **434**:1144–1148.
73. Kralovics, R., et al. 2005. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N. Engl. J. Med.* **352**:1779–1790.
74. Tong, A.H., et al. 2004. Global mapping of the yeast genetic interaction network. *Science.* **303**:808–813.
75. Li, S., et al. 2004. A map of the interactome network of the metazoan *C. elegans*. *Science.* **303**:540–543.
76. Giot, L., et al. 2003. A protein interaction map of *Drosophila melanogaster*. *Science.* **302**:1727–1736.
77. Barrios-Rodiles, M., et al. 2005. High-throughput mapping of a dynamic signaling network in mammalian cells. *Science.* **307**:1621–1625.
78. van 't Veer, L.J., et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* **415**:530–536.
79. van de Vijver, M.J., et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**:1999–2009.