

# Approaching Clinical Proteomics: Current State and Future Fields of Application in Cellular Proteomics

Rolf Apweiler,<sup>1,a</sup> Charalampos Aslanidis,<sup>2</sup> Thomas Deufel,<sup>3,b</sup> Andreas Gerstner,<sup>4,c</sup> Jens Hansen,<sup>2</sup> Dennis Hochstrasser,<sup>5,d</sup> Roland Kellner,<sup>6,a</sup> Markus Kubicek,<sup>7,e</sup> Friedrich Lottspeich,<sup>8,a</sup> Edmund Maser,<sup>9</sup> Hans-Werner Mewes,<sup>10</sup> Helmut E. Meyer,<sup>11,a</sup> Stefan Müllner,<sup>12,a,f</sup> Wolfgang Mutter,<sup>13,a</sup> Michael Neumaier,<sup>14,b</sup> Peter Nollau,<sup>15,b</sup> Hans G. Nothwang,<sup>16,b</sup> Fredrik Ponten,<sup>17</sup> Andreas Radbruch,<sup>18,c</sup> Knut Reinert,<sup>19</sup> Gregor Rothe,<sup>20,b</sup> Hannes Stockinger,<sup>21,c,f</sup> Attila Tárnok,<sup>22</sup> Mike J. Taussig,<sup>23</sup> Andreas Thiel,<sup>18,c</sup> Joachim Thiery,<sup>24,b</sup> Marius Ueffing,<sup>25,a</sup> Günther Valet,<sup>8,c</sup> Joel Vandekerckhove,<sup>26</sup> Christoph Wagener,<sup>15,b</sup> Oswald Wagner,<sup>27,e,f</sup> Gerd Schmitz<sup>2,a,b,f\*</sup>

<sup>1</sup>European Bioinformatics Institute, Hinxton, United Kingdom

<sup>2</sup>Institute of Clinical Chemistry and Laboratory Medicine, University of Regensburg, Regensburg, Germany

<sup>3</sup>Institute for Clinical Chemistry and Laboratory Diagnostics, University of Jena, Jena, Germany

<sup>4</sup>Department of Otorhinolaryngology/Surgery, University of Bonn, Bonn, Germany

<sup>5</sup>Department of Pathology/Clinical Chemistry, University of Geneva, Geneva, Switzerland

<sup>6</sup>Merck, Darmstadt, Germany

<sup>7</sup>Center for Bioinformatics and Clinical Proteomics, Medical University of Vienna, Vienna, Austria

<sup>8</sup>Max-Planck-Institute for Biochemistry, Martinsried/Munich, Germany

<sup>9</sup>Institute of Toxicology and Pharmacology, University of Kiel, Kiel, Germany

<sup>10</sup>Institute for Bioinformatics, GSF, Neuherberg, Germany

<sup>11</sup>Medical Proteomics Center, University of Bochum, Bochum, Germany

<sup>12</sup>Protagen, Dortmund, Germany

## • Abstract

Recent developments in proteomics technology offer new opportunities for clinical applications in hospital or specialized laboratories including the identification of novel biomarkers, monitoring of disease, detecting adverse effects of drugs, and environmental hazards. Advanced spectrometry technologies and the development of new protein array formats have brought these analyses to a standard, which now has the potential to be used in clinical diagnostics. Besides standardization of methodologies and distribution of proteomic data into public databases, the nature of the human body fluid proteome with its high dynamic range in protein concentrations, its quantitation problems, and its extreme complexity present enormous challenges. Molecular cell biology (cytomics) with its link to proteomics is a new fast moving scientific field, which addresses functional cell analysis and bioinformatic approaches to search for novel cellular proteomic biomarkers or their release products into body fluids that provide better insight into the enormous biocomplexity of disease processes and are suitable for patient stratification, therapeutic monitoring, and prediction of prognosis. Experience from studies of in vitro diagnostics and especially in clinical chemistry showed that the majority of errors occurs in the preanalytical phase and the setup of the diagnostic strategy. This is also true for clinical proteomics where similar preanalytical variables such as inter- and intra-assay variability due to biological variations or proteolytic activities in the sample will most likely also influence the results of proteomics studies. However, before complex proteomic analysis can be introduced at a broader level into the clinic, standardization of the preanalytical phase including patient preparation, sample collection, sample preparation, sample storage, measurement, and data analysis is another issue which has to be improved. In this report, we discuss the recent advances and applications that fulfill the criteria for clinical proteomics with the focus on cellular proteomics (cytoproteomics) as related to preanalytical and analytical standardization and to quality control measures required for effective implementation of these technologies and analytes into routine laboratory testing to generate novel actionable health information. It will then be crucial to design and carry out clinical studies that can eventually identify novel clinical diagnostic strategies based on these techniques and validate their impact on clinical decision making. © 2009 International Society for Advancement of Cytometry

## • Key terms

clinical proteomics; cellular proteomics; cytomics; cytoproteomics; proteome

<sup>13</sup>PROFOS, Regensburg, Germany

<sup>14</sup>Institute for Clinical Chemistry, University Hospital Mannheim, Mannheim, Germany

<sup>15</sup>Institute for Clinical Chemistry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>16</sup>Department of Animal Physiology, Technical University Kaiserslautern, Kaiserslautern, Germany

<sup>17</sup>Institute for Genetics and Pathology, University of Uppsala, Uppsala, Sweden

<sup>18</sup>German Rheumatism Research Center, Berlin, Germany

<sup>19</sup>Department of Mathematics and Computer Science, University of Berlin, Berlin, Germany

<sup>20</sup>Laboratory Center Bremen, Bremen, Germany

<sup>21</sup>Department of Molecular Immunology, University of Vienna, Vienna, Austria

<sup>22</sup>Department of Pediatric Cardiology, University of Leipzig, Leipzig, Germany

<sup>23</sup>Technology Research Group, The Babraham Institute Cambridge, Cambridge, United Kingdom

<sup>24</sup>Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig, Leipzig, Germany

<sup>25</sup>Institute of Human Genetics, GSF Neuherberg, Neuherberg, Germany

<sup>26</sup>Department of Medical Protein Research, University of Ghent, Ghent, Belgium

<sup>27</sup>Institute of Medical and Chemical Laboratory Diagnostics, University of Vienna, Vienna, Austria

<sup>a</sup>**Consensus document of the DGPF**—Deutsche Gesellschaft für Proteomforschung ([www.dgpf.org](http://www.dgpf.org)).

<sup>b</sup>**Consensus document of the DGKL**—Deutsche Vereinte Gesellschaft für Klinische Chemie und Laboratoriumsmedizin ([www.dgkl.de](http://www.dgkl.de)).

<sup>c</sup>**Consensus document of the DGfZ**—Deutsche Gesellschaft für Zytometrie ([www.dgfv.org](http://www.dgfv.org)).

<sup>d</sup>**Consensus document of the SSSC**—Schweizerische Gesellschaft für Klinische Chemie ([www.sssc.ch](http://www.sssc.ch)).

<sup>e</sup>**Consensus document of the OEGLMKC**—Österreichische Gesellschaft für Laboratoriumsmedizin und Klinische Chemie ([www.oeglmkc.at](http://www.oeglmkc.at)).

<sup>f</sup>**Consensus document of the Danubian Biobank Consortium** ([www.danubianbiobank.de](http://www.danubianbiobank.de))

Received 21 January 2009; Revision Received 28 April 2009; Accepted 6 July 2009

Additional Supporting Information may be found in the online version of this article

\*Correspondence to: Gerd Schmitz, Institute of Clinical Chemistry and Laboratory Medicine, University of Regensburg, Franz-Josef-Strauss Allee 11, Regensburg 93053, Germany

Email: [gerd.schmitz@klinik.uni-regensburg.de](mailto:gerd.schmitz@klinik.uni-regensburg.de)

Published online 8 September 2009 in Wiley InterScience ([www.interscience.wiley.com](http://www.interscience.wiley.com))

DOI: 10.1002/cyto.a.20779

© 2009 International Society for Advancement of Cytometry

**T**HE proteome of an organism is the protein complement of its genomic functionality. The proteome is highly dynamic and varies according to cell type and functional state of the cell. The proteome may reflect immediate and characteristic changes in response to disease processes and external stimulation. As the actual proteome, e.g., a set of all proteins present in a body fluid, cell, tissue, or organism at a certain point of time, represents only a subset of all possible gene products, the proteome can not be directly predicted from genome information. Any protein may exist in multiple forms that respond within a particular cell or between different cells via post-translational modifications or degradation processes that affect protein structure, localization, function, and turnover. Like the proteome, the proteolytic degradation products of the proteome, the so called low-molecular-weight (LMW) range proteome, may also have the potential to contain disease-specific information. The identified disease-specific peptides appear to be the fragments of endogenous high-abundant proteins, such as transthyretin or fragments of low abundance cellular and tissue proteins, such as BRCA-2 (1,2). Also, catabolic pathways lead to degradation of proteins and the extent of this degradation may depend on preanalytical variables like temperature or time of storage of the clinical sample.

Beyond protein target identification, emerging novel nanotechnology strategies that make use of these LMW biomarkers in vivo or ex vivo may enhance our ability to discover and characterize molecules for early disease detection or strati-

fication and will expand the prognostic capability of current proteomic modalities. However, they may also pose new problems related to accuracy and variation. In addition, pathophysiological processes involving proteolytic activities, e.g., tumor proteases, are detectable in the plasma peptidome (1). The discovery of sensitive and specific panels of biomarkers or single marker proteins holds the key to future diagnostic and therapeutic monitoring of complex diseases such as metabolic and vascular disease, cancer, and neurodegenerative disorders.

Clinical proteomics focuses on the clinical and analytical validation and implementation of novel therapy- or diagnosis-related markers that originate from preclinical studies designed to identify hits and leads in analogy to drug screening studies (fundamental research during the discovery phase is not discussed in this document). It also includes the selection, validation, and Standard Operating Procedure (SOP)-assessment of the most appropriate and robust method that can be integrated into the workflow of available analytical platforms in clinical laboratories. Hit picking of potential targets for clinical proteomics may follow either a top down approach applying high-throughput technologies to population-based studies or selected cohorts as case/control studies or from twin registers, to identify an unbiased approach, novel markers that have to be further characterized for sensitivity, specificity, and function. Alternatively, the bottom up approach follows search strategies such as protein-protein or metabolite interaction within already identified pathways or

around targets therein to find additional targets with higher sensitivity, disease specification, or stability.

Advances in proteomic analysis with regard to high-throughput and high-content analysis were needed for clinical proteomics. They have been achieved both in whole cell multiparameter flow cytometry and high-content microscopic imaging and in the field of affinity binding technologies and liquid phase technologies like MS/MS applied for cellular abstracts or fluid samples. High resolution liquid phase separation methods, along with progress in chemometry and biometry for large-scale data analysis, now offer the possibility to introduce these research tools into actual laboratory diagnostics to screen for risk factors, identify new disease-specific or stage-specific biomarkers and find novel cellular or fluid markers for therapeutic drug monitoring or new therapeutic targets. As a result, clinical proteomics has the potential to complement cytomics, genomics, metabolomics, lipidomics, glycomics, and transcriptomics, including splice variant analysis, in gaining a better understanding of disease processes and translate this complex knowledge into diagnostic tools for clinicians.

While metabolomics has already been established in clinical diagnosis for years, e.g., in newborn screening; clinical proteomics is only now on the verge to enter the hospital, and just like in the case of metabolomics, essential criteria for their successful use in a clinical environment have to be fulfilled. First, high-throughput analysis platforms must provide reproducible protein patterns in a clinically acceptable turnaround time and have to reach the instrumental stability of laboratory analysers that can operate at the technician level with a minimum of academic supervision necessary; platforms must also fit into the workflow of a clinical laboratory. Second, bioinformatics algorithms must include chemometry, data reduction, and conversion into actionable health information and they have to be robust and integrated into standard laboratory information systems. Third, preanalytical conditions have to be standardized and optimized for the development of clinically applicable tests. Before any work can start to establish novel biomarkers, a patient cohort, which can be used to address a specific clinical question and that is well-defined by all necessary anamnestic and physiologic parameters including age, sex, hormonal status, treatment, and hospitalization status must be identified and made available for study. For biomarker development, SOP-driven biobanks have to be integrated into the diagnostic workflow with validated storage conditions for all proteomic applications. The standardized and also SOP-driven preparation of patient samples is one of the most urgent challenges for reproducible and clinical useful results. Regulations must be established addressing medical-legal issues such as patient consent and commercial use of samples and intellectual property. This must be achieved at an international level to allow large-scale multi center studies that are a prerequisite for the task at hand.

Cytomics deals with the study of “cytomes” or cell systems at a single cell level to understand the functions and molecular architecture of the cytome (3,4). Using various cytometrical procedures including fluorescent microscoping imaging, multiparameter flow cytometry, novel transcriptomic, proteomic

planar, and bead-based arrays, the various components of a cell become accessible to high-throughput and high-content analysis.

In this review, we focus on cellular proteomics (cytoproteomics) and suggest accepted procedures for sample preparation and standardization of protocols. We summarize the preparative methods in the field of cytoproteomics and emphasize key clinical applications in the field. We conclude with bioinformatics approaches for the evaluation of proteomic data. Another review focussing on clinical proteomics in body fluids is in preparation and will also be published as a consensus document from the authors.

### PREANALYTICAL PHASE

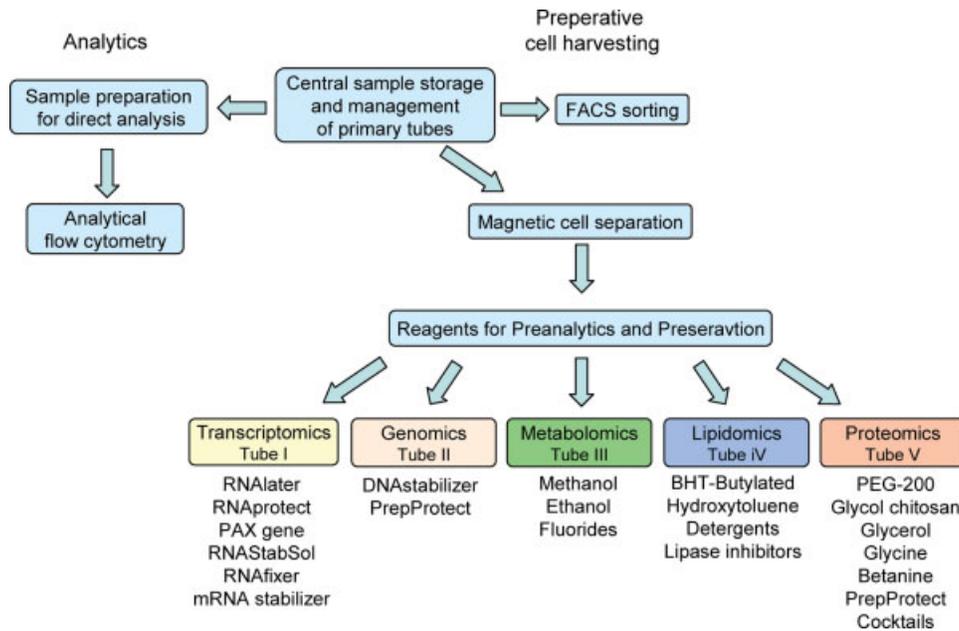
Sources of technical or sample variation are mainly found within the preanalytical phase; most prominent here is lack of standardized procedures for sample collection, including patient preparation, like fasting, specimen acquisition, handling, and storage, which accounts for more than 90% of the errors within the entire diagnostic process (5). Advances in genomics and proteomics have led to high expectations of clinical biomarker discovery; yet, for the successful generation of validated biomarkers, more attention must be directed towards the preanalytical stage such as sample collection, transport, preparation, and processing (6) but also at standardization and quality management. Automation of preanalytical sample processing using robotics and central sample storage and management helps to eliminate or minimize variations in sample quality (Fig. 1).

Furthermore, for cellular analysis, the sample material has to be preserved prior to consecutive analysis. There are a variety of reagents and methods available that facilitate proper handling, isolation without preactivation, and stabilization of the cellular material. Figure 1 summarizes the principles of cell harvesting by magnetic separation and the reagents available for stabilization. The choice of the reagents depends on the downstream application of either transcriptomic, genomic, metabolomic, lipidomic, or proteomic analysis.

The preclinical discovery phase may yield interesting biomarkers that require validation before any translation into clinical proteomics is feasible. The type of sample needed and their processing could be quite different for the two phases. In general, there is a major difference between the requirements of high-throughput proteomic profiling as a clinical proteomics approach and protein identification and in-depth characterization of single protein samples (low throughput) in the preclinical phase. Another point is the preparation of the patient, e.g., for parameters that depend on metabolic status where fasting before sample collection is essential; also other influencing factors must be accounted for, such as diurnal rhythms for parameters like peptide hormones or cytokines.

### CELLULAR PROTEOMICS (CYTOPROTEOMICS)

Blood cells offer unique insights into disease processes. Therefore, erythrocytes, granulocytes, monocytes, lympho-

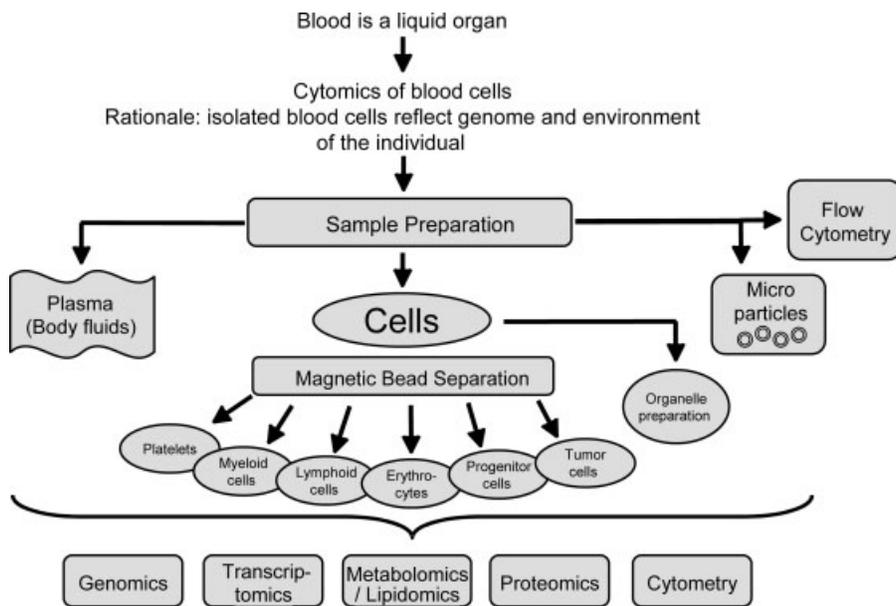


**Figure 1.** Automation of preanalytical sample processing and sample stabilization from primary tubes to high-throughput platform specific secondary tube formats containing stabilization reagents. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

cytes, and platelets are of special interest for clinical proteomics (7). Blood is a liquid organ and isolated blood cells reflect the environment and genome of the individual (Fig. 2).

Flow cytometry is currently widely used as an analytical tool for clinical cell analysis directly from anticoagulated whole blood and also for cell sorting to generate pure populations of cells from heterogeneous and highly integrated mixtures as are found in the majority of biological

environments. Elispot, slide based cytometry, and tissue arrays together with high-content screening microscopy are further upcoming techniques in cytoproteomics. The major challenge for preanalytical standardization is related to the use of fresh samples, either for direct multiparameter analysis of cellular proteomics in whole blood or body fluids without preseparation or for cell sorting and enrichment strategies for subsequent proteomic and functional genomic analysis.



**Figure 2.** Cytomics of the blood compartment.

### SAMPLE PREPARATION AND STANDARDIZATION FOR CLINICAL CYTOPROTEOMICS

For clinical analysis, the samples should be rapidly analyzed, because transportation and storage lead to artifacts like selective damage or aggregation of specific cell subpopulations or shedding of cell surface markers. Especially for activation markers or functional tests, the samples should be analyzed within 3 h. Immunophenotyping, including that for intracellular antigens, however, can be performed for up to 7 days following recently developed methods for the stabilization of cells without destruction of epitopes (8). Most parameters in clinical flow cytometry are performed from EDTA-blood, especially surface parameters for immunophenotyping. The EDTA-tube can be used in parallel for routine hematology and is available in all clinical settings. Many applications could also be done from blood anticoagulated using heparin or citrate. However, the current anticoagulants have major drawbacks for functional cell analysis, and the  $\text{Ca}^{2+}$  capturing anticoagulants disturb phosphoproteomics (9) because of the interference with cell signaling networks. In addition, the lysis-procedures release cytosolic products that coat and activate or inhibit other cell types. As an example heme/haptoglobin complexes in plasma bind to the CD163 cysteine-scavenger receptor of circulating monocytes and alter signal transduction (10). Therefore, novel anticoagulants and elimination of intact red blood cells (RBC) are necessary (11–14). The investigation of parameters, whose functions depend on calcium, like certain receptor ligand interactions, oxidative burst, or phagocytosis, heparinized blood has to be used. For cell preparation, the use of gradient centrifugation is almost completely substituted by erythrocyte lysis for routine applications before cytometric analysis (15). There are highly automated and standardized preparation robotics available for lysis and in part for complete sample handling ([www.beckmancoulter.com](http://www.beckmancoulter.com), [www.bdbioscience.com](http://www.bdbioscience.com)). For the analysis, highly standardized multiparameter flow cytometers with special software application for routine diagnosis ([www.beckmancoulter.com](http://www.beckmancoulter.com), [www.bdbioscience.com](http://www.bdbioscience.com), [www.miltenyibiotec.com](http://www.miltenyibiotec.com), [www.partec.com](http://www.partec.com)) are on the market. Up to eight parameters (FSC, SSC, and six different antibodies) are already implemented in routine diagnosis and CE-certified kits are available. Higher complexities using more than 15 parameters are available for more scientific approaches. Obviously, there is more information with high-content analysis of single cells, but there are also more problems due to antibody interactions like unspecific binding or cross-linking of antibodies. With increasing numbers of different fluorescent dyes, the compensation of the fluorescence signals in different detector channels is an important aspect. Most up-to-date analysers use digital compensation methods, which are highly standardized due to software driven compensation algorithms (16).

The strategy to validate antibody producing cell lines in so called “Cluster of Differentiation” workshops (CD-nomenclature) has generated more than 300 defined antibody specificities which are of great help for standardized clinical test development. A main topic of the working group “flow cyto-

metry and quantitative microscopy” of the DGKL is quality control and standardization. This working group is involved in the European Society for Clinical Cell Analysis (ESCCA), who already started to review and validate recommendations for flow cytometric multiparameter analysis. However, especially with regard to phenotyping and isolation of cells for genomic or proteomic analysis, standardization of techniques is still missing (17). Absolute quantification of the expression level of surface molecules is partly realized with stabilized cell lines (e.g., Daudi) or beads with certain amounts of epitopes, recognized by the specific antibodies used for cell staining (18). All these approaches have not yet reached a good interlaboratory comparability and need further standardization. The availability of antibodies against phosphorylated structures in signal transduction molecules recently has expanded the applicability of flow cytometry to the multiplex characterization of signaling patterns, e.g., during the activation of cells of the immune systems or during malignant transformation (9).

Combining analytical flow cytometry with magnetic isolation of specific cell populations for consecutive functional proteomic, transcriptomic, and metabolomic analysis will ultimately need robotics, which harvest cells from primary tubes and to transfer the separated cells into different preservation solutions for consecutive enabling technology platforms (Fig. 1).

### PREPARATIVE METHODS

#### Fluorescence-Activated Cell Sorting

Beside cell analysis, flow cytometry is also used for cell isolation using fluorescence-activated cell sorting (FACS) (19,20). Cell sorting allows the investigator to quantify several fluorescence and light-scattering parameters of individual particles and purify those events with the desired characteristics for further functional study. This technology can separate a heterogeneous cell suspension into purified fractions containing a single cell type with equivalent speed and accuracy. Different methods of capturing a particle of interest are implemented in different cytometers. Some systems use a so called catcher tube which is located in the upper portion of the flow cell. It moves in and out of the sample stream to collect a population of desired cells at a rate of up to 300 cells per second. The particles and cells can be physically separated and deposited onto a defined location for further analysis or culture even under sterile conditions. The advantage of this sorting method is a soft handling of cells, but the sorting speed is rather low. High-speed sorters use techniques like “stream in the air,” which allows much higher throughput with up to 100,000 cells per second, but cells are more stressed and may be less suitable for further analysis (21,22). Applications for high-speed cell sorting are commonly grouped into two categories—bulk sorting and rare event sorting. In bulk sorting, the sort operator is given a large number of cells of which a certain percentage is expressing a given phenotype and this population is desired for further study. This marker can be a surface antigen bound to a fluorescently labeled antibody, a transient transfection expressing a given protein of interest,

activated leukocytes or any detectable measure of cell state or phenotype. Whatever the measured parameter, the operator must screen through a large number of cells to collect a sufficient amount of material for the next step of analysis. For instance, an investigator may be interested in the top 5% of cells expressing a given intracellular cytokine to perform proteomic analysis. Typically, a fairly large number of cells are needed for these studies. With readily available mass-spectrometry methods, in the order of 10 million cells would be needed to detect a protein expressed at roughly 1000 copies/cell, corresponding to a total starting population of ~200 million cells. At a sort rate of 40,000 cells per second, this translates to 5000 s or around an hour and a half of sorting. This does not take into account material losses due to sample handling (e.g., centrifugation, etc.) or time losses due to technical issues.

### Magnetic Bead-Based Separation of Cells

Isolation techniques have to be considered before the analysis of the cell proteome. Magnetic bead-based techniques for the isolation of cell subsets according to the expression profile of membrane antigens provides a rapid and convenient approach to cell isolation from blood, bone marrow, and other sources of cells in suspension to select or deplete specific cell populations. These procedures can be performed without major technical instrumentation for low throughput applications, e.g., with Dynal magnetic bead systems (<http://www.invitrogen.com/dynal>). Disadvantages of this system are first the use of “large” (>5  $\mu\text{m}$ ) beads, which grossly activate some cells, for example monocytes and interfere with subsequent analysis of the cell by flow cytometry. A second major disadvantage is that there is no automation for high-throughput sample handling. In clinical proteomics, cell-based analysis requires high-throughput and rapid techniques for cell harvesting and collection, realized by two automated systems for cell isolation from whole blood without prior density gradient centrifugation. One system (RoboSep<sup>®</sup>; [https://www.stemcell.com/product\\_catalogue/robosep.asp](https://www.stemcell.com/product_catalogue/robosep.asp)) uses a lysing protocol prior to magnetic leukocyte preparation. This system has the major disadvantage that some activation pathways are influenced by the lysed erythrocyte membranes. Furthermore, yield and purity is inferior to the manual Dynal or automated Miltenyi System. The Miltenyi autoMACS<sup>™</sup> Pro and multiMACS<sup>™</sup> Separator<sup>®</sup> Systems need no lysing or other manipulation prior to magnetic bead separation ([www.miltenyibiotec.com](http://www.miltenyibiotec.com)). It is mandatory that the separation procedure does not induce major changes neither in the expression nor in the cellular localization of proteins. On the other hand, it is also necessary that the separation techniques do not result in unwanted cell loss, as it may occur under pathological conditions when cells float outside the density gradient fraction. In this regard two general approaches can be distinguished: (1) positive selection uses one or more cell specific markers for identification and separation, and therefore, the purity is usually very high—but is prone to alter cell function, whereas (2) negative selection depletes contaminating cells and leaves the target cell untouched, but could lack accuracy,

and high purity, and may result in poor recovery. Moreover, negative isolation has the drawback of frequent contamination with immature cells, e.g., reticulocytes or reticulated platelets. Based mostly on specific membrane-associated antigens, various cell subsets can be isolated. A major drawback of most available antibodies against cell surface proteins (Table 1), which are classified in the CD-nomenclature series, is their effect on cell activation. Therefore, there is need for the identification of other, non-activating cell specific targets with stable surface expression for affinity isolation. Table 1 summarizes common antibodies used for magnetic isolation of different cell types and the putative cell, RNA, protein, and lipid yield, if the cells are separated from 10 ml EDTA-blood. Sufficient cell counts for genomic, proteomic, or lipidomic analysis can be obtained from 10 to 20 ml of peripheral blood, depending on the cell count of the individual subset. Ready-to-use kits for positive or negative isolation of various cell populations are now commercially available. Acceptable results in terms of recovery, yield, and purity ( $\geq 90\%$ ) of the final cell population which is not yet reached with all systems (23).

The different methods for cell fractionation by size and density and highly selective affinity-based technologies including affinity chromatography, fluorescence-activated cell sorting (FACS), and magnetic cell sorting were recently reviewed (24).

### Microdissection of Cells

Microdissection offers the advantage to analyze directly the pathologically relevant cell type, which is often grossly underrepresented in the tissue samples investigated. For example, in a heterogeneous tumor tissue, tumor cells can be isolated specifically by microdissection, and the dissected cells can be analyzed quantitatively using for instance DIGE saturation labeling (25) or LC-MS (26). Applying DIGE saturation labeling in combination with two-dimensional gel electrophoresis and mass spectrometry, it has been shown for the first time that quantitative proteome analysis of 1,000 micro dissected cells is feasible (25). With this technique, a number of candidate biomarkers were identified and confirmed as pancreatic cancer cells compared to the cells in precursor lesions.

Laser-based microdissection offers one opportunity to separate functional tissue (27), but for proteomic analysis fine-needle microdissection is preferable, because fewer artifacts, especially thermal denaturation dependent on ablation energy, are introduced. The quality and quantity of tissue molecules critically depends on upstream tissue preparation (28). Snap-frozen specimens, for example, give a higher and better yield than fixed and embedded tissue specimens. Nevertheless, each preparation protocol has its own advantages and disadvantages, so the choice also depends on the consecutive type of analysis. Microdissection can be used to design proteome-based studies in combination with serology or in parallel to genomic and/or expression analysis for the identification of biomarkers and novel targets. Different systems for microscopic microdissection are available ([www.leica-microsystems.com](http://www.leica-microsystems.com), [www.arrayx.com](http://www.arrayx.com),

**Table 1.** Magnetic antibody separation of blood cells from 10 ml of whole blood

CELL FAMILY	CELL SUBSET	CAPTURE ANTIBODY/METHODS	CELL NUMBER	RNA (mg)	PROTEIN (mg)	LIPID (nmol)
Red Blood Cells	All	CD235a (Glycophorin A)	$5 \times 10^{10}$	0.3	500	250,000
	Immature (Reticulocytes)	FACS	$5 \times 10^8$	0.3	5	2,500
Platelets	All	CD42b (WF-Receptor)	$2.5 \times 10^9$	0.1	25	12,500
	Immature (Retic. Platelets)	FACS	$5 \times 10^7$	0.1	2.5	250
Myeloid cells	Neutrophils	CD 15 (Fucosyltransferase 4)	$3 \times 10^7$	3.0	3.0	15,000
	Basophils/Mastcells	Negative selection	$2.5 \times 10^5$	0.25	0.25	125
	Eosinophils	Negative selection	$1.5 \times 10^6$	1.5	1.5	750
	Monocytes/Phagocytes	CD 14 or Negative selection	$2.5 \times 10^6$	2.5	2.5	1,250
	Monocytes/APC	FACS	$0.5 \times 10^5$	0.5	0.5	250
Lymphoid cells	T-cells	CD 3+	$1 \times 10^7$	1.0	1.0	500
	T reg.	CD 4+/CD 25+	$5 \times 10^5$	5	0.5	250
	NK	CD 56 (NCAM-1)	$2 \times 10^6$	2.0	2.0	1,000
	B-cells	CD 19+	$2 \times 10^6$	2.0	2.0	1,000
Stem cells/Progenitors	Hematopoietic progenitors	CD133 (Prominin-1)	$5 \times 10^4$	0.5	0.05	25
	Endothelial progenitor cells	CD34+/KDR+ (VEGFR)	80	0.0008	0.00008	0.04
Non-hematopoietic cells	Tumor cells	Variable	Variable	Variable	Variable	Variable
	Fetal cells	FACS	50	0.0005	0.00005	0.02

www.qiagen.com/applications/microdissection, www.palm-microlaser.com). The standardization and automatization of these techniques is still in progress. Most preparative steps are still operator-dependent (29).

**PREPARATION OF CELLULAR SAMPLES AND APPLICATIONS**

**Mononuclear Cells (Lymphocytes, Monocytes, and Dendritic Cells)**

Monocytes are a target of proteomics, especially concerning clinical diagnosis and monitoring of atherosclerosis. Seong et al. (30) analyzed the effect of oxidative stress generated at sites of inflammation and injury on the proteomic profiles of monocytes. As a result, 28 identified proteins mainly involved in energy metabolism, translation, and mediation of protein folding, were over expressed. Dupont et al. (31) elaborated two-DE reference maps of human macrophage proteome and secretome to elucidate the macrophage dysfunctions involved in inflammatory, immunological, and infectious diseases. They showed that macrophages are involved in a wide array of biological functions, including cytoskeletal machinery, carbohydrate metabolism, apoptosis, and protein metabolism. Combined oligonucleotide microarray and proteomic approaches have been used to study genes associated with dendritic cell differentiation. Dendritic cells are antigen-presenting cells essential for the initiation of primary immune responses, they mostly derive from human monocytes. Protein analysis of these cells was done and about 4% of the protein spots separated by two-DE exhibited quantitative changes

during differentiation and maturation. The differentially expressed proteins were identified by MS and represent proteins with calcium binding, fatty acid binding, or chaperone activities, as well as cell motility functions (32).

Several studies report on the proteome of lymphocytes. From the current literature (PubMed), we created a database (Table 2, Supporting Information, www.LipidomicNet.org) summarizing the literature in which human blood cells have been studied with high-throughput transcriptomic (arrays, etc.) or proteomic (MALDI-TOF/SELDI-TOF/2D-gel electrophoresis, etc.) technologies. As shown in Table 2, for a variety of cells, transcriptomic and proteomic analyses have been performed thus linking genotype to phenotype, but still, there are cells that have not been yet analyzed in-depth on the proteomic level. While in Table 2 the total number of transcriptomic and proteomic publications related to individual cells is given, the detailed table in the Supporting Information displays the individual references and the type of the analysis, which involved either transcriptomic, proteomic, or both investigations.

Proteome databases of human helper T-cells were established using classical proteomics (33). The proteins differentially expressed in Th1 and Th2 cells are described in the article by Rautajoki et al. (34) Detailed proteomic studies have also been published on lymphoblastoid cells, and strategies for studying signaling pathways in lymphocytes, combining proteomics, and genomics, have been proposed (35). Signaling via immunoreceptors is orchestrated at specific plasma membrane microdomains, referred to as lipid rafts. Lipid rafts are dynamic assemblies floating freely in the surrounding mem-

**Table 2.** Current status of literature on high-throughput transcriptomic and proteomic analysis of cells from the blood compartment

CELL TYPE	SUBTYPE	TRANSCRIPTOME	PROTEOME	BOTH
Hematopoietic stem cells	Embryonic	4	6	2
	CFU-E	1		
	T-cell lineage	3		
	B-cell lineage		1	
	Hematopoietic	6	4	
Red blood cells	Mesenchymal	3	1	
	Reticulocytes	3	1	
	Red blood cells		6	
Platelets	Megakaryocytes	7		
	Platelets	3	11	
	Platelet microparticles		1	
Monocytes	Monocytes	40	7	2
	Macrophages	19	6	1
	Dendritic cells	20	4	2
Granulocytes	Neutrophils	6	6	2
	Eosinophils	1	2	
	Basophils	1		
Lymphocytes	T-cells	4	3	1
	B-cells	4	4	

The total number of publications per cell type in the field of transcriptomics and proteomics is shown. A more detailed list is given in the Supporting Information and in LipidomicNet ([www.lipidomicnet.org](http://www.lipidomicnet.org)).

branes of living cells. The proteins participating in lipid rafts in T lymphocytes have been studied with proteomics, and the subject was recently reviewed by Wollscheid et al. (36) MS was used by Li et al. (37) to specifically detect proteins depleted from rafts by cholesterol-disrupting drugs. The authors detected a large number of signaling molecules in lipid rafts and provided evidence for a connection between cytoskeletal proteins and lipid rafts.

### Granulocytes

Neutrophil activation is expressed by the production and release of inflammatory mediators, which induce inflammatory cascades leading to cell damage and dysfunction. The mediators involved in the inflammation processes are mainly proteins, making proteomics the ideal approach to investigate them. Functional proteomics also provides a good tool to elucidate the complex signal transduction network of inflammation to identify disease-associated targets and to improve current therapies. Genomic and proteomic studies were performed on activated neutrophils. The central pathway in the regulation of neutrophil function is the p38 mitogen-activated protein kinase signal transduction as shown by Singh et al. (38). In a recent study, Burlak et al. used subcellular proteomics to identify proteins associated with human neutrophil phagosomes. Proteins were identified by MALDI-TOF MS and/or LC-MS/MS analysis. They unexpectedly found enzymes typically associated with mitochondria in the phagosome fractions and conclude that neutrophil phagosomes have a till now unrecognized complexity (39).

### Erythrocytes and Reticulocytes

Because of its high abundance, high availability, and easy purification, red blood cells (RBC) have been the subject for many proteomic studies. In the last 6 years, more than 750 red blood cell proteins have been identified (40). RBCs can be pelleted by centrifugation, the supernatant can be discarded, and RBCs can be resuspended. The purity of erythrocytes can be further enhanced by eliminating the top RBC layers after centrifugation. Further density centrifugations, e.g., on CL5020 (Cederlane Laboratories, Hornby, ON, Canada), elimination of granulocytes by passing the RBC fraction through nylon nets and wash steps to eliminate plasma proteins should be done. The RBCs must be used then immediately to prepare membrane and cytoplasmic fractions. The quality of the preparations has to be determined by purity assessment of RBC samples. The packed RBC samples must be counted for white blood cells (WBCs), granulocytes, monocytes, platelets, reticulocytes, and RBCs, e.g., using a routine hematology analyser. No reticulocytes or other cell types should be found. Subproteomes of RBCs can be analyzed by membrane preparations and membrane extractions, soluble protein preparation, or cytoskeleton extractions (41). The study of Kakhniashvili et al. describes the proteome of red blood cells (42), where erythrocyte membranes as well as cytoplasm were analyzed by LC-MS/MS. The authors identified a total of 181 unique protein sequences in the membrane fractions and in the cytoplasmic fractions.

Because of the low frequency in peripheral blood, reticulocytes must be enriched for proteomic analysis. In contrast, for proteomic analysis of mature red blood cells (RBC), reticu-

locytes should be allowed to mature, e.g., by storing peripheral whole blood for 72 to 96 h at 4°C without shaking, although it has been suggested that the maturation of reticulocytes *in vitro* is limited at 4°C (43). Recent data on the analysis of the final material clearly shows that the procedures taken to eliminate reticulocytes is effective (41), but there may be problems with protein degradation during this time. Leukocytes must be eliminated for example with filters (Plasmodipur, Euro-diagnostics, Arnhem, The Netherlands).

### Platelets and Reticulated Platelets

Platelets are easily and rapidly activated after blood drawing. Therefore, collected blood should be immediately processed. Especially for platelets, recent studies have shifted away from global profiling to the analysis of subfractions of the proteome and the identification of changes induced upon blood cell activation. For example, these studies on platelets allowed the identification of many more platelet proteins than can be achieved by global profiling; giving a more complete view of the platelet proteome and the biological information obtained may be of more relevance. The two important subproteomes of the platelet, which have been extensively studied include the phosphoproteome and the secretome (44–46). Many signaling pathways in platelets are regulated by differential phosphorylation upon platelet activation. The application of new proteomic approaches on the protein phosphorylation events which occur during the activation of platelets lead to the discovery of new phosphorylated target proteins and new possible phosphorylation sites in already known phosphorylated proteins (47–49).

Like many other cells, platelets secrete proteins from pre-formed storage granules in response to stimuli. The platelets secretome has been analyzed in the supernatants isolated by differential centrifugation of low-dose thrombin-activated platelets. The supernatant contains secreted proteins but no membrane proteins. Using a complementary multidimensional chromatography approach, the secreted protein fraction may be digested with trypsin and the resulting peptides separated using strong cation exchange and reverse-phase chromatography before introduction into an ion-trap MS. Analysis of the secretome from thrombin-activated platelets has identified over 300 proteins that are secreted upon activation. For a number of proteins it was not known whether they are present in, or secreted by platelets, including secretogranin III, cyclophilin, and calumenin. These three secreted proteins have been identified in atherosclerotic lesions, suggesting a potential role in atherothrombosis (50).

### Circulating Microparticles

Cell-derived microparticles have been shown to be relevant in a number of diseases like thrombosis, cardiovascular disease, antiphospholipid syndrome, and systemic inflammation (SIRS, sepsis) (51–53). Microparticles have been shown to be a valuable tool in the field of red blood cell shape anomalies and dysfunctions. Furthermore, it is likely that microparticle release is part of the aging process of erythrocytes or platelets. It may coincide with apoptotic blebbing,

shape change, or loss of surface protein modifications. In this sense, the microparticle release is thought to be an innovative marker for chronic vascular diseases like atherosclerosis or other aging disorders.

Microparticles are defined as shedded membranous fragments or vesicles with a diameter of less than 1  $\mu\text{m}$  bearing on their surface markers of the parent cell (51). They are surrounded by a phospholipid bilayer that is mainly composed of phosphatidylcholine, sphingomyelin, and phosphatidylethanolamine. Microparticles differ from their parental cells in regard of the lipid composition and distribution between the two membrane leaflets. The asymmetrical phospholipid distribution with anionic phospholipids being confined to the inner leaflet usually changes during microparticle formation. Activation, as well as apoptosis, leads to calcium-dependant swelling, budding, and microparticle release (54). Finally, both cell activation and apoptosis result in disruption of the membrane skeleton structure that is necessary for surface blebbing and particle release. The principal sites of microparticle release seem to be cell protrusions that in regard to the cell type resemble microvilli, pseudopodia, filopodia, or proteopodia.

Microparticles are derived from a number of different cell types (55). In addition to its well-established role of platelet-derived microparticles in haemostasis, an important function in a variety of additional blood cells (e.g., red blood cells, monocytes, and lymphocytes) and in other cell types like endothelial and stromal progenitor cells has been proposed. Red blood cells have been known for a long time to shed membrane microparticles under stimulation. It could be shown that as for other microparticles this process depends on a rise in intracellular calcium (54). Additionally, they found that the microparticle release coincides with the formation of echinocytes, which is characterized by loss of the natural discoid shape to gain a more spherical form with regularly distributed cell membrane protrusions. The shape change as well as the microparticle release may resemble the *in vivo* events during cell aging. Circulating microparticles of leukocyte origin modulate cellular interactions through the upregulation of cytokines and cytoadhesins in endothelial cells and monocytes.

Antigenic markers measured by flow cytometry are mainly used to classify and further subdivide membrane microparticles (55,56). Alternative approaches include ELISA and solid phase capture assays. The advantage of flow cytometry lies in its capacity to distinguish different microparticle populations and subpopulations according to the expression of surface and intracellular antigens in a multiparameter approach. The sensitivity and specificity of flow cytometry could, therefore, be used to improve the analysis of blood and other cell types. In this sense it should be combined with the well-established technique of conventional analysers as outlined above.

### Organelles (Phagosome, Proteasome, Nucleosome)

The introduction of proteomic analysis like mass-spectrometry-based identification of proteins has created new opportunities for the study of organelle composition, process-

ing of transport intermediates, and large subcellular structures (Anderson and Mann, 2006). Traditional cell-biology techniques such as sucrose density gradient centrifugation are used to enrich these structures for proteomic analysis, and such analysis provides insights into the biology and functions of these structures. Beside a good upfront purification of the sample, the validation of sample purity done by electron microscopy, immuno-blotting, or other orthogonal methods to exclude significant cross-contamination during organelle separation is obligatory. Standardization of purification procedures is therefore urgently requested for a successful clinical application (57–60). A major challenge in the control of purity of the organelle fraction is the lack of specific markers for each organelle (61). The study of cell compartments provides a unique way to access and identify low abundance and organelle-specific proteins in a biological sample. The immunoisolation procedure on magnetic beads allows for the isolation of highly enriched fractions of certain organelles. A single analysis by MS enables the identification of thousands of peptides, leading to the formal identification of several hundred proteins (62).

The analysis of the proteome of extracellular integral membrane proteins in living cells is complicated because of their hydrophobic nature, their low abundance, and problems with sufficient purity. Therefore, a high-throughput platform has been developed, which is based on proteolytic digestion of whole intact cells (63). The resulting peptide fragments are subjected to liquid chromatology and tandem MS to generate the proteome of integral plasma membrane proteins. This method has been called PROCEED (PROteome of Cell Exposed Extracellular Domains).

Within the surface plasma membrane of cells morphologically distinct regions or domains, like microvilli, cell–cell junctions, clathrin coated pits, and lipid raft domains can be distinguished. Each of these domains is specialized for a particular function, such as nutrient absorption, cell–cell communication, and endocytosis. Lipid raft domains include caveolae, characterized by a distinctive membrane coat composed of caveolin-1, and rafts (64–69). Both have a high-content of cholesterol and sphingolipids, have a light buoyant density, and function in endocytosis and cell signaling.

Classical proteomics does not allow the detection of human endothelial caveolae and raft proteins as they represent less than 0.5% of the total cellular protein pool and less than 2% of total plasma membrane proteins. Therefore, fractionation techniques like cationic silica enrichment of caveolae or detergent solubilization with Triton X-100 at 4°C are needed to enable the subproteome analysis of these membrane microdomains (70). The analysis of these distinct regions can be put forward by high-throughput proteomic platforms together with modern techniques like stable isotope labeling (SILAC) (71,72).

Proteomic analysis of purified caveolae has identified a wide assortment of proteins that are localized to these structures. The first detailed proteomic analysis of caveolae was carried out in 1994. On the basis of their buoyancy and resistance to detergent solubilization, Lisanti et al. used sucrose density

ultracentrifugation to purify caveolae-rich membrane domains from murine lung tissue (73). This procedure allowed for the exclusion of  $\geq 98\%$  of an integral plasma membrane protein marker, while retaining  $\sim 85\%$  of total caveolin and  $\sim 55\%$  of GPI-linked marker proteins. This initial proteomic analysis of caveolae allowed for the first time large-scale characterization of caveolae-enriched protein constituents and provided the basis for many follow-up investigations into the functional significance of these proteins (74). For clinical proteomics, it is interesting that a long list of diseases could be associated with lipid rafts and raft associated proteins (75).

Phagosomes are key organelles for the innate ability of macrophages to participate in tissue remodeling, clear apoptotic cells, and restrict the spread of intracellular pathogens. Using a proteomic approach,  $>140$  proteins associated with latex bead-containing phagosomes had been identified. The elaboration of a two-dimensional gel database of  $>160$  identified spots allowed us to analyze how phagosome composition is modulated during phagolysosome biogenesis. The systematic characterization of phagosome proteins provided new insights into phagosome functions and the protein or groups of proteins involved in and regulating these functions. Different types of phagocytosis can be distinguished. For example, type I phagocytosis is primarily mediated by  $Fc\gamma$ -receptors, while type II phagocytosis is mediated by complement receptors, especially CR3. Another type of phagocytosis is associated with deep tubular invaginations. So far, more than 500 proteins have been identified in different phagosome preparations using MALDI-TOF MS and nano-electrospray MS/MS (57,76).

The secretion of proteins by exocytosis is an important subset of protein-trafficking events. Secretion can be either constitutive (occurring continuously) or regulated (occurring on demand) as a result of an extracellular signal. Both genetic and biochemical approaches have been combined to produce our current understanding of eukaryotic protein secretion, although there are clearly many questions that remain unanswered (77). Concerning clinical proteomics, it is of interest, that constituents of the exocytosis pathways are associated with multiple diseases. For example, members of the AP-3 pathway are involved in disorders of lysosome-related organelles such as the Hermansky-Pudlak syndrome complex, Chediak-Higashi syndrome, and the ceroid lipofuscinoses. This provides new opportunities to understand AP-3 pathway-related disorders and their relation to membrane phospholipid processing. Mutations in the ABCA1 gene, one member of the AP-3 pathway, are involved in dysregulated vesicular trafficking from the transgolgi compartment to the plasma membrane (78).

The nucleolus/nucleosome is a key organelle that coordinates the synthesis and assembly of ribosomal subunits and forms in the nucleus around the repeated ribosomal gene clusters. A quantitative analysis of the proteome of human nucleoli was successfully performed using mass-spectrometry-based organellar proteomics and stable isotope labeling. In vivo fluorescent imaging techniques are directly compared to endogenous protein changes (79,80). The flux of 489 endoge-

nous nucleolar proteins in response to three different metabolic inhibitors that each affect nucleolar morphology could be registered (81,82). Mitochondria are not only the major suppliers of ATP, but also significantly involved in different cellular processes, e.g., metabolism or apoptosis. The proteomic investigation of mitochondria, e.g., after stress exposure, offers the opportunity to specify their role in physiologic and pathologic processes (83). Both the nuclear- and mitochondrial-encoded proteins and their genes are summarized in MITOP (<http://www.mitop.de:8080/mitop2/>), a comprehensive database for genetic and functional information. The "Human disease catalogue" contains tables with a total of 110 human diseases related to mitochondrial protein abnormalities, sorted by clinical criteria and age of onset. MITOP should contribute to the systematic genetic characterization of the mitochondrial proteome in relation to human disease (81,84,85).

CLINICAL APPLICATIONS IN CYTOPROTEOMICS

Multiparameter Flow Cytometry

Multiparameter flow cytometry is widely used in the routine laboratory to generate clinical diagnostic information from complex heterogeneous mixtures such as human blood for multiple indications (23). The main indications for clinical flow cytometry can be classified according to the material used and divided into two major principles, the immunophenotyping and functional analysis of cells.

Immunophenotyping

Immunophenotyping has become a routine practice in the diagnosis and classification of leukaemias. For non-Hodgkin lymphoma, flow cytometry is one of several complementary techniques (86). Further important examples include the analysis of lymphocyte subpopulations, e.g., in the diagnosis of primary or acquired immunodeficiency, in the quantitation of haematopoietic stem cells and of residual leukocytes in erythrocyte preparations (15), and in the analysis of platelets (87). Diagnostic haematopathology depends on the applications of flow cytometric immunophenotyping (FCI) combined with immunohistochemical immunophenotyping. Selected cases may require additional cytogenetic and molecular studies for diagnosis. FCI offers the sensitive detection of antigens for which antibodies may not be available for paraffin-embedded immunohistochemical immunophenotyping. However, paraffin immunohistochemical immunophenotyping offers preservation of architecture and evaluation of expression of some proteins, which may not be available by flow cytometric immunophenotyping. These techniques should be used as complementary tools in diagnostic hematopathology. Types of specimens suitable for FCI include peripheral blood, bone marrow (BM) aspirates, and core biopsies, fine-needle aspirates (FNAs), fresh tissue biopsies, and all types of cell containing body fluids. There are many advantages of FCI compared to immunohistochemistry. For exam-

Table 3. Immunophenotyping in peripheral blood

PARAMETER	INDICATION
Lymphocyte subpopulations (T-, B-, NK-Cells)	Primary immunodeficiency syndromes, HIV, transplantation
Malignant lymphocytes	Non-Hodkin-lymphoma
CD34+ stem cells	Stem cell transplantation
Lymphocyte activation (CD38; HLA-DR)	HIV
T-cell repertoire	Omen-Syndrome
CD16, CD66b on granulocytes	Paroxysmal nocturnal hemoglobinuria
CD11 a/b on granulocytes	Leukocyte adhesion defects
HLA-DR on monocytes	Sepsis
LDL-receptor	Familial hypercholesterolemia

For the analysis of surface molecules, EDTA-blood is usually used and stored at room temperature (refer also "sample preparation").

ple, dead cells may be gated out of the analysis, weakly expressed surface antigens are detectable. Multicolour (2-, 3-, 4-) analysis can be performed, allowing for an accurate definition of the surface antigen profile of specific cells. Two simultaneous hematologic malignancies may be detected within the same tissue site. Tissue biopsy may be substituted by the relatively non-invasive diagnostic evaluation of body fluids. Disadvantages of FCI are for example that sclerotic BM may yield too few cells for adequate analysis. A markedly hypercellular or "packed" BM may yield too few cells for analysis. Sclerotic tissue may be difficult to suspend for individual cellular analysis. There is loss of architectural relationships. T-cell lymphomas that do not have an aberrant immunophenotype may not be detected but on the other hand, aberrant T-cell immunophenotype (i.e., absence or down-regulation of pan T-cell antigens, particularly CD7) does not necessarily indicate malignancy and may be observed in infectious mononucleosis or inflammatory disorders. This requires correlation with light microscopy as well as clinical data and additional ancillary studies (e.g., molecular/cytogenetic analysis), in some situations (88). A "Kompetenznetz Leukämien" is established in Germany to support standardization of clinical diagnosis (<http://www.kompetenznetz-leukaemie.de>). The diagnosis of many primary immunodeficiency diseases requires the use of several laboratory tests. Flow cytometry is applicable in the initial workup and subsequent management of diagnostics of primary immunodeficiency diseases related to the B-, T-, and NK-cell compartment (89). Further indications for immunophenotyping are listed in Table 3 (15).

For certain samples and functional assays, specific sample preparation procedures have to be considered as summarized in Table 4. In conclusion, multiparameter flow

**Table 4.** Sample preparation for different material and functional analysis in clinical diagnosis

SPECIMEN/BLOOD CELL TYPE	FUNCTIONAL TEST	INDICATION	SAMPLE PREPARATION
Reticulated platelets		Thrombocytopenia	Freshly drawn citrated blood, CTAD anticoagulated blood
Bone marrow		Leukemia-, lymphomadiagnosis	EDTA- or heparin-anticoagulated sample, erythrocyte lysing methods are preferable to density centrifugation
Bronchoalveolar lavage		Inflammatory, autoimmune lung diseases, sarcoidosis, collagenosis, fibrosis	Heparinized sample preferably stored at 4°C
Cerebrospinal fluid		Inflammatory, autoimmune diseases of the CNS	Preferably stored at 4°C
Neutrophiles	Oxidative burst	Chronic granulomatosis, sepsis, SIRS	Heparinized blood mandatory
Neutrophiles	Phagocytosis	Hereditary or transient defects in phagocytosis	Heparinized blood mandatory
Basophiles	Basophil degranulation	Allergic or pseudoallergic reactions	EDTA-or heparinized blood
Microparticle analysis		Thrombosis, cardiovascular disease, systemic inflammation (SIRS, sepsis)	Enrichment from anticoagulated blood by two differential centrifugation steps; flow cytometric measurement

cytometry is established and widely distributed in clinical routine analysis.

### Slide Based Cytometry

Slide based cytometry with fluorescence microscope based instruments is an analytical technique that allows rapid quantitative analysis of a high number of individual cells in suspension, cell culture, or in tissue sections. The specimens are tagged with fluorescent dyes and immobilized on a slide. There is a broad range of clinical applications for slide based cytometry systems (90) and flow cytometry and slide based cytometry analysis yield comparable results (91). Two important features of slide based cytometry that cannot be covered by flow cytometry are: (1) the quantitative analysis of tissue sections (92) and (2) the reanalysis of identical cells following restaining (93,94). Regarding (1), approaches have been proposed towards 3D-cytometry of tissue sections and in tissue cultures (95). Regarding (2), this approach enables quantification of up to 100 antigens per individual cell (96).

In conclusion, slide based cytometry instruments are for some years standard in pharmaceutical R and D, but until today, only few of them are used for clinical routine analysis. However, due to the rapid development in this field, a substantial role in the near future can be expected.

### High-Content Image Screens

Automated microscopy and advanced automated image analysis software offer the possibility to quantitatively correlate multiple markers to each other from a single cell within a large analyzed cell population (97–99). This could be useful to discover new pathways for target protein identification and gaining a better understanding of complex intracellular mechanisms. The development of novel algorithms for image

analysis and data analysis tools and employing automated confocal microscopy will shed new light on the High-Content Analysis (HCA) fields (100). Potential applications of such technologies for HCA screening lay in the use on intracellular high-definition imaging using, e.g., for G-protein receptor analysis (GPR ligand dependent internalization) for the analysis of signalomes (cytosol-PM translocation, cytosol/ER-nuclear translocation) (101).

In conclusion, high-content image screens appear to be a good candidate to fulfill high-throughput with sensitivity, low imprecision, and high potential to discover new clinical relevant pathways for drug developments.

### Tissue Arrays

Tissue arrays are assemblies of multiple patient tissue samples (e.g., tumors) prepared for multiplexed immunohistological serial analysis of tissue specimens (102–104). This technology promises to have a great potential for clinical research and diagnostics, especially in clinical oncology (105). It offers the following advantages: amplification of a scarce resource, experimental uniformity (tissues of multiple patients are treated in an identical manner), decreased assay volume, preservation of original blocks. Depending on the shape of the tissue sample and the method used to obtain it, multi-tissue array techniques may be classified into two different groups: rod-shaped tissue techniques and core tissue techniques. Some technical aspects should be considered when deciding which technique should be used: the number, size, and origin of tissue samples; the quality of paraffin wax, the distance between samples, and the depth in the receptor block; antigenicity preservation and block sectioning. European initiatives are established to design and standardize an infrastructure for a networked tumor tissue bank. The samples are collected

according to an SOP used at different institutes and the data are collected and distributed in a central database [TuBaFrost 4 (106); <http://genome.tugraz.at> (107)]. Antibody-based tissue profiling allows a streamlined approach for generating expression data both for normal and disease tissues. It is also possible to generate protein expression data on many different individual patients to evaluate heterogeneity of tissue profiles. In addition, specific antibodies directed to a particular target protein allow numerous functional assays to be performed ranging from conventional ELISA assays to detailed localization studies using fluorescent probes and protein capture experiments ("pull-down") for purification of specific proteins and their associated complexes for structural and biochemical analyses. The challenge for antibody-based proteomics is to move from a conventional protein-by-protein approach into a high-throughput mode to allow chromosome wide analysis. Technical challenges involve both the antigen production and the subsequent generation and characterization of the antibodies. In addition, methods for systematic protein profiling on a whole proteome level are lacking. Recently, Nilsson et al. showed that antibodies specific to human proteins can be generated in a high-throughput manner involving stringent affinity purification using recombinant protein epitope signature tags (PrESTs) as immunogens and affinity-ligands. The specificity of the generated affinity reagents, here called mono-specific antibodies (msAb), were validated with a novel protein microarray assay (108). The use of tissue microarrays (TMAs) generated from multiple biopsies combined into single paraffin blocks enabled high-throughput analysis of protein expression in various tissues and organs. Recently, Kampf et al. showed that high-throughput analysis of protein expression can be performed with a standard set of tissue microarrays representing both normal and cancer tissues (109). The TMA technology provides an automated array-based high-throughput technique where as many as 1000 paraffin-embedded tissue samples can be brought into one paraffin block in an array format. A comprehensive atlas of human protein expression patterns in normal and cancer tissues can be created by combining the methods mentioned earlier. A set of standardized TMAs can be produced to allow for rapid screening of a multitude of different tissues and cell types using immunohistochemistry. This approach could also quite effectively be used for generation of expression data for model animals such as mouse, rat, and chimpanzee. A valuable tool for medical and biological research can thus be envisioned as a complement to genome and transcript profiling data (110–113).

In conclusion, in clinical proteomics, tissue arrays are already in use with different complexity. The main applications of this techniques are screening purposes, quality control, diagnosis, biomarker validation, and teaching (103,114). In autoimmune diagnosis, for example commercially distributed tissue arrays ranging from two to more than 20 different tissue samples for the incubation with the serum of a single patient (<http://www.euroimmun.de>) are available. In cancer studies, multiplexed tissue arrays are in use for the screening for potential therapeutic antibodies, for the prediction of out-

come, prognosis, and for biomarker screening and validation (114).

### BIOINFORMATICS APPROACHES IN PROTEOMICS

Although there are numerous examples of single laboratory parameters, which allow good clinical support, like troponin I alone for the diagnosis of myocardial infraction, procalcitonin with a good predictive value for bacterial sepsis, or TSH in the diagnosis of thyroidal dysfunction, most proteomic studies indicated that a single biomarker may be not adequate for reliable diagnosis, staging, or prognosis of a disease. This immediately raises a question of how to combine several biomarkers to provide a diagnostic or predictive pattern? While a definitive answer is probably still far away, a number of approaches emerged.

Hierarchical decision tree-based classification methods, such as CART (115), were among the first algorithms to utilize the available information on multiple biomarkers. Another approach is heuristic clustering (116,117). However, empirical observations suggested that these approaches were not successful, because the number of incorrect predictions made by the classification algorithm increases with the complexity of the decision tree. Furthermore, the number of datasets available for training the decision tree was quite low, resulting in a lack of statistical significance beyond the second or third nodes of the tree.

Support Vector Machines (SVM) [for an example see (118)] appeared to be an excellent way to overcome some of these limitations due to the theoretical principles upon which they are based. Indeed, excellent empirical performance of SVM has been reported in a number of diverse applications. These approaches provided superior cross-validated predictive performance, but mixed results were obtained with blinded datasets. Reliable results have been obtained when the number of variables was less than 20 and substantial differences between the datasets existed. However, when the differences were more subtle, the number of support vectors had to be increased. This, in turn, resulted in the classifiers over-fitting (also referred to as "memorizing," a term often employed in Artificial Intelligence research) to the training set and thus in poor classification of blinded datasets (Mischak et al., unpublished data). To avoid such memorizing effects, the number of variables and dimensions has to be lowered.

One very important facet of the use of combinations of biomarkers when making a predictive diagnosis with a classification algorithm is to have a properly calibrated indication of the level of confidence in the predictions being made. In other words, a classification such as "this urine sample has been drawn from an individual with type II diabetes" should also have a numeric score indicating how likely, or probable, the classification is to be correct, i.e., with 90% confidence this urine sample has been drawn from an individual with type II diabetes. Clearly, 90% confidence is more reliable than a prediction with 50% confidence, especially if there are only two alternatives to be considered: disease presence versus disease absence (in which case 50% confidence indicates little more than random guessing). Having such confidence levels (also

referred to as probabilities) attached to a classification also enables unbalanced costs of miss-classification to be assigned in an optimal manner. For example, there may be far more serious consequences to incorrectly predicting the absence of a disease than to incorrectly predicting its presence. Although SVM's provide very encouraging classification performance on a range of difficult problems, they are devoid of any probabilistic semantics, thus are unable to provide levels of confidence attached to any classification, and the clinician is then left with no information as to how much the predictions should be trusted.

A promising probabilistic classification method which shares many of the positive characteristics of the SVM, but in addition provides the important levels of confidence with each classification prediction is based on the Gaussian Process. A general purpose and computationally efficient Gaussian Process based classification method has recently been developed (119), and has been successfully applied to the problem of correct prediction of BRCA1 and BRCA2 heterozygous genotypes (120). The probabilistic nature of Gaussian Process based classification methods provides a means of inferring optimally weighted combinations and possible selection of biomarkers and a detailed study of this capability is currently investigated.

No matter which of these approaches is being utilized, two basic considerations apply: (1) the number of independent variables should be kept minimal and should certainly be below the number of samples investigated, and (2) any such approach only becomes valid if applied towards an entirely blinded validation set, and it should be imperative to include such a blinded dataset in any report on potential biomarkers.

#### DATA ANALYSIS STRATEGIES

Mass-spectrometry-based proteomics has become an important component of biological and clinical research. The analysis of rather complex protein or peptide mixtures, most importantly derived from body tissues or body fluids challenges data acquisition, handling and data processing capabilities. Proteomic data are combined with patient or clinical data they are merged, compiled and integrated to very complex datasets. Thus, very heterogeneous types of data are generated with the need to be managed properly. As data handling, interpretation, validation, storage, and dissemination are critical with respect to ensuring their proper use, it is indispensable to develop formats as well as minimal requirements to ensure data quality.

Another challenge is the sheer size of proteomic datasets. Current genomic and proteomic analytical methods, while highly developed and powerful, are easily generating gigabyte comprehend datasets. The informatics, however, has to handle these data with respect to retrieving information from these data in a reasonable time and quality manner. For the various tasks including for example the identification of labeled and unlabeled peptide species or the elimination of background signal in the mass spectrum and chromatogram, a plethora of tools and approaches have been developed, although mature data standards are still missing (121). In addition, the wealth

of genomic or proteomic information accumulated from prior studies is almost never adequately used for the initial planning or interpretation of experiments or data from a given study as integration of data from outside remains challenging and tedious. One consequence of this situation is that in every proteomic experiment, the result is rediscovered (122). Using external resources for own research pipelines causes various problems in many cases, as remote access to external sources may be difficult and integration of the data retrieved laborious or hardware-dependent. In many cases, the compilation of the required external information has to be done in a manual fashion. Here, transparency, as well as efforts to generate the means to integrate previous data into a current dataset or experimental planning can avoid costly repetition of a study. As suggested by Mewes (personal communication, Workshop Clinical Proteomics, Martinsried 9/2006), it would be preferable to combine resources with interfaces (in any desirable programming language) and use these interfaces to convert generic data formats into standardized ones, send standardized information over the Web and thereby disseminate standardized and structured information to any interested clients located anywhere via Web Services. The HUPO-PSI has started these standardizing efforts already for proteomic datasets and the ProDaC consortium ([www.fp6-prodac.eu](http://www.fp6-prodac.eu)) will develop practical software tools to bring datasets into these standardized formats. However, at the end of generating/analyzing data and converting them into standardized file formats an export from local database systems into central data repository like PRIDE (<http://www.ebi.ac.uk/pride>) will be mandatory for private datasets as well as public datasets (after the results were published in scientific journals). These central repositories will be a fundamental required infrastructure to avoid rediscovering of already known results.

In many studies it is necessary to analyze a large number of individual proteomes and to compare the obtained results. In biomarker discovery or clinical studies, large cohorts of patients and normal individuals are required to detect protein patterns that consistently associate with a specific condition within a large background of proteins that may randomly fluctuate within the population tested. The risk to misinterpret correlating protein patterns as biomarkers is enormous. This is particularly true for the many types of cancer, but also other diseases such as Type 2 diabetes or heart failure where useful and adequate diagnostic markers with high specificity and sensitivity are lacking. Thus, the success of mass-spectrometry-based discover of candidate biomarkers is highly dependent on the ability to properly statistically handle and interpret these data with use of decoy databases and a defined false discovery rate. Besides the sheer volume of proteomic datasets the heterogeneity of patients, sample processing and data acquisition in multi-centered studies as well as heterogeneity in data formats, processing methods, software tools, and databases that are involved in the translation of spectral data into relevant and actionable information for scientists can greatly limit the value of such data. In the past, biomarker discovery has suffered from inconsistent data acquisition, statistical handling, and validation. Thus, clinical proteomics needs standard

operating procedures and guidelines for thorough data generation (123,124) consistent data analysis (125) as well as distinct training and validation of datasets (126).

These standards include:

- Standards and data formats for data acquisition (raw data), storage (mzXML/mzData), and exchange (PSI: <http://www.psicodev.info>)
- Public data repositories (Peptide Atlas, PRIDE, GPMDB, SwissProt/Uniprot)
- Standards for quality assessment (e.g., FDR, composite decoy databases, power analysis)
- Integration towards a “linear pipeline” (Aebersold, ProDaC)
- Integration in a complex database including biological information (systems oriented approach)

With respect to the interpretation of complex data from clinically oriented studies, rules of conduct have to be defined for linking profiling results with clinical phenotypes and allow proper statistical evaluation and interpretation of such protein profiles. Here, standards for data analysis and quality management have been recently established, standards for data reporting (HUPO-PSI, MCP, and other journal guidelines) have been set up and databases for proteomics data being structured as a task within large HUPO initiatives. Proteomic search algorithms and a public repository and peptide identifications for the proteomic community (PRIDE: <http://www.ebi.ac.uk/pride>) have been developed.

Data have to be qualified with respect to selectivity and specificity for a pathological process. Clustering variables, statistics, reproducibility, as well as standards and exchange formats and study design has to be carefully considered as a basis for proper data analysis. Here, sample retrieval, asservation, storage, and handling are some of the most important parameters to be strictly defined. In addition, protocols and standards have to be ensured to generate and secure reproducibility of the data. Recently established HUPO standards and SOPs for MS-based data acquisition and analysis can be considered as an important step towards that direction. A current challenge yet remaining is data heterogeneity, which can negatively affect proper integration of proteomics data with other biomedical data.

Future efforts in data management in the field of clinical proteomics have to consider that data and results are likely to be the application and disease specific.

A remaining problem with respect to handling clinically relevant data will be to allow utmost transparency and at the same time ensure data protection and intellectual property. Early teamwork between the clinical level, bioinformatics, medical informatics, and proteomic scientists is needed to overcome the current limitations.

Although grouped analysis of proteomics data is of high importance for the development of efficient therapies as well as for scientific purposes, it is equally essential to use such data to assess the therapy dependent disease progression of individual patients in advance that is at diagnosis, disease

relapse, or other time points of interest. This will favor the individualized treatment of patients in stratified patient groups as an important prerequisite to maximize therapeutic success and to minimize adverse drug reactions (personalized/individualized medicine) (127). One key question to address in this context relates to our ability to draw the right conclusions for (short-, mid-, or long-term) therapeutic approaches from the highly dynamic proteome patterns that are influenced by various states of disease.

With these aims, clinical proteomics studies have to comprise diseased patients as well as a reference group of healthy or otherwise defined individuals. An unknown validation set of diseased and reference individuals, including patients with unrelated diseases is required to prove the robustness of classification for unknown patient samples.

## CONCLUSION

Current approaches to understand the functional diversity of an organism preferentially strive for a systems approach whereby first the phenotypic classification of a specific cytochrome as the cellular systems and subsystems and functional components of the organism is achieved prior to an attempt to perform proteomic analysis. The fundamental basis for this approach has been well established in studies of cellular systems over many years. When combined with advanced proteomic approaches, this approach can achieve rapid and specific identification of a direct link between biomarkers and their functional roles in complex organisms (128). Integration of proteomics and cell-based technologies will allow the description of the molecular setup of normal and abnormal cell systems within a relational knowledge system, permitting the standardized discrimination of abnormal cell states in disease.

At the current stage, these methods are mostly qualitative and must be seen as exploratory approaches, aiming at advancing scientific knowledge within clinical studies rather than routine. In future, this may gradually change, with some of the methods as well as applications being established as clinical routine assays. As one of the consequences, individualized predictions of further disease course in patients (predictive medicine by cytomics) by characteristic discriminatory data patterns will then permit individualized therapies, identification of new pharmaceutical targets, and establishment of a standardized framework of relevant molecular alterations in disease (129). Most important for clinical proteomics and cytoproteomics on the verge of routine laboratory diagnosis is the control of preanalytical aspects. This is best met by a high degree of standardization, including SOPs and automated work stations for high-throughput sample preparation. Therefore, clinical proteomics and cytoproteomics have the potential to complement genomics, cytomics, metabolomics, lipidomics, glycomics, and transcriptomics including splice variant analysis for gaining better understanding of disease processes.

## LITERATURE CITED

1. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: Written in blood. *Nature* 2003;425:905.
2. Liotta LA, Petricoin EF. Serum peptidome for cancer detection: Spinning biologic trash into diagnostic gold. *J Clin Invest* 2006;116:26–30.

3. Gomase VS, Tagore S. Cytomics. *Curr Drug Metab* 2008;9:263–266.
4. Tarnok A. Slide-based cytometry for cytomics—A minireview. *Cytometry Part A* 2006;69A:555–562.
5. Lippi G, Salvagno GL, Montagnana M, Guidi GC. Reliability of the thrombin-generation assay in frozen-thawed platelet-rich plasma. *Clin Chem* 2006;52:1827–1828.
6. Conrads TP, Hood BL, Veenstra TD. Sampling and analytical strategies for biomarker discovery using mass spectrometry. *Biotechniques* 2006;40:799–805.
7. Thadikaran L, Siegenthaler MA, Crettaz D, Queloz PA, Schneider P, Tissot JD. Recent advances in blood-related proteomics. *Proteomics* 2005;5:3019–3034.
8. Bikoue A, Janossy G, Barnett D. Stabilized cellular immunofluorescence assay: CD45 expression as a calibration standard for human leukocytes 3. *J Immunol Methods* 2002;266:19–32.
9. Perez OD, Nolan GP. Phospho-proteomic immune analysis by flow cytometry: From mechanism to translational medicine at the single-cell level 3. *Immunol Rev* 2006;210:208–228.
10. Kristiansen M, Graversen JH, Jacobsen C, Sonne O, Hoffman HJ, Law SK, Moestrup SK. Identification of the hemoglobin scavenger receptor. *Nature* 2001;409:198–201.
11. Van AH, Bode C, Darius H, Diehm C, Encke A, Gulba DC, Haas S, Hacke W, Puhl W, Quante M, Riess H, Scharf R, Schellong S, Schror T, Schulte KL, Tebbe U. Anticoagulation: The present and future. *Clin Appl Thromb Hemost* 2001;7:195–204.
12. Proimos G. Platelet aggregation inhibition with glycoprotein IIb–IIIa inhibitors. *J Thromb Thrombolysis* 2001;11:99–110.
13. Tamada Y, Murata M, Hayashi T, Goto K. Anticoagulant mechanism of sulfonated polyisoprenes. *Biomaterials* 2002;23:1375–1382.
14. Fareed J, Jeske WP. Small-molecule direct antithrombins: Argatroban. *Best Pract Res Clin Haematol* 2004;17:127–138.
15. Sack U, Rothe G, Barlage S, Gruber R. Flow cytometry in clinical diagnostics. *J Lab Med* 2007;24:277.
16. Valet G, Tarnok A. Potential and challenges of a human cytochrome project. *J Biol Regul Homeost Agents* 2004;18:87–91.
17. Schwartz A, Marti GE, Poon R, Gratama JW, Fernandez-Repollet E. Standardizing flow cytometry: A classification system of fluorescence standards used for flow cytometry. *Cytometry* 1998;33:106–114.
18. Gratama JW, Dhautcourt JL, Mandy F, Rothe G, Barnett D, Janossy G, Papa S, Schmitz G, Lenkei R. Flow cytometric quantitation of immunofluorescence intensity: Problems and perspectives. European working group on clinical cell analysis 5. *Cytometry* 1998;33:166–178.
19. McCoy JP Jr. Basic principles of flow cytometry. *Hematol Oncol Clin North Am* 2002;16:229–243.
20. Tung JW, Parks DR, Moore WA, Herzenberg LA, Herzenberg LA. New approaches to fluorescence compensation and visualization of FACS data. *Clin Immunol* 2004;110:277–283.
21. Ashcroft RG, Lopez PA. Commercial high speed machines open new opportunities in high-throughput flow cytometry (HTFC). *J Immunol Methods* 2000;243:13–24.
22. Ibrahim SF. High-speed cell sorting: Fundamentals and recent advances. *Curr Opin Biotechnol* 2003;14:5–12.
23. Bernas T, Gregori G, Asem EK, Robinson JP. Integrating cytomics and proteomics. *Mol Cell Proteomics* 2006;5:2–13.
24. Dainiak MB, Kumar A, Galaev IY, Mattiasson B. Methods in cell separations. *Adv Biochem Eng Biotechnol* 2007;106:1–18.
25. Sitek B, Luttes J, Marcus K, Kloppel G, Schmiegel W, Meyer HE, Hahn SA, Stuhler K. Application of fluorescence difference gel electrophoresis saturation labeling for the analysis of microdissected precursor lesions of pancreatic ductal adenocarcinoma. *Proteomics* 2005;5:2665–2679.
26. Zang L, Palmer TD, Hancock WS, Sgroi DC, Karger BL. Proteomic analysis of ductal carcinoma of the breast using laser capture microdissection, LC-MS, and 16O/18O isotopic labeling. *J Proteome Res* 2004;3:604–612.
27. von EF, Melle C, Ernst G. Microdissecting the proteome. *Proteomics* 2007;7:2729–2737.
28. Erickson HS, Gillespie JW, Emmert-Buck MR. Tissue microdissection. *Methods Mol Biol* 2008;424:433–448.
29. Seliger B, Kellner R. Design of proteome-based studies in combination with serology for the identification of biomarkers and novel targets. *Proteomics* 2002;2:1641–1651.
30. Seong JK, Kim DK, Choi KH, Oh SH, Kim KS, Lee SS, Um HD. Proteomic analysis of the cellular proteins induced by adaptive concentrations of hydrogen peroxide in human U937 cells. *Exp Mol Med* 2002;34:374–378.
31. Dupont A, Tokarski C, Dekeyser O, Guihot AL, Amouyel P, Rolando C, Pinet F. Two-dimensional maps and databases of the human macrophage proteome and secretome. *Proteomics* 2004;4:1761–1778.
32. Richards J, Le NF, Hanash S, Beretta L. Integrated genomic and proteomic analysis of signaling pathways in dendritic cell differentiation and maturation. *Ann N Y Acad Sci* 2002;975:91–100.
33. Rosengren AT, Nyman TA, Lahesmaa R. Proteome profiling of interleukin-12 treated human T helper cells. *Proteomics* 2005;5:3137–3141.
34. Rautajoki K, Nyman TA, Lahesmaa R. Proteome characterization of human T helper 1 and 2 cells. *Proteomics* 2004;4:84–92.
35. Caron M, Imam-Sghiouar N, Poirier F, Le Caer JP, Labas V, Joubert-Caron R. Proteomic map and database of lymphoblastoid proteins. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002;771:197–209.
36. Wollscheid B, von Haller PD, Yi E, Donohoe S, Vaughn K, Keller A, Nesvizhskii AI, Eng J, Li XJ, Goodlett DR, Aebersold R, Watts JD. Lipid raft proteins and their identification in T lymphocytes. *Subcell Biochem* 2004;37:121–152.
37. Li N, Shaw AR, Zhang N, Mak A, Li L. Lipid raft proteomics: Analysis of in-solution digest of sodium dodecyl sulfate-solubilized lipid raft proteins by liquid chromatography-matrix-assisted laser desorption/ionization tandem mass spectrometry. *Proteomics* 2004;4:3156–3166.
38. Singh GB, Singh H. Functional proteomics with biolinguistic methods. n-Grams deliver sensitive portrayals of gene similarity. *IEEE Eng Med Biol Mag* 2005;24:73–80.
39. Burlak C, Whitney AR, Mead DJ, Hackstadt T, Deleo FR. Maturation of human neutrophil phagosomes includes incorporation of molecular chaperones and endoplasmic reticulum quality control machinery. *Mol Cell Proteomics* 2006;5:620–634.
40. Goodman SR, Kurdia A, Ammann L, Kakhniashvili D, Daescu O. The human red blood cell proteome and interactome. *Exp Biol Med (Maywood)* 2007;232:1391–1408.
41. Pasini EM, Kirkegaard M, Mortensen P, Lutz HU, Thomas AW, Mann M. In-depth analysis of the membrane and cytosolic proteome of red blood cells. *Blood* 2006;108:791–801.
42. Kakhniashvili DG, Bulla LA Jr, Goodman SR. The human erythrocyte proteome: Analysis by ion trap mass spectrometry. *Mol Cell Proteomics* 2004;3:501–509.
43. Neidez-Nguyen TM, Wajzman H, Marden MC, Bensidhoum M, Moncollin V, Giarratana MC, Kobari L, Thierry D, Douay L. Human erythroid cells produced ex vivo at large-scale differentiate into red blood cells in vivo. *Nat Biotechnol* 2002;20:467–472.
44. Ditttrich M, Birschmann I, Stuhlfelder C, Sickmann A, Herterich S, Nieswandt B, Walter U, Dandekar T. Understanding platelets. Lessons from proteomics, genomics and promises from network analysis. *Thromb Haemost* 2005;94:916–925.
45. Jung E, Heller M, Sanchez JC, Hochstrasser DF. Proteomics meets cell biology: The establishment of subcellular proteomes. *Electrophoresis* 2000;21:3369–3377.
46. Marcus K, Meyer HE. Two-dimensional polyacrylamide gel electrophoresis for platelet proteomics. *Methods Mol Biol* 2004;273:421–434.
47. Garcia A, Prabhakar S, Brock CJ, Pearce AC, Dwek RA, Watson SP, Hebestreit HF, Zitzmann N. Extensive analysis of the human platelet proteome by two-dimensional gel electrophoresis and mass spectrometry. *Proteomics* 2004;4:656–668.
48. Marcus K, Moebius J, Meyer HE. Differential analysis of phosphorylated proteins in resting and thrombin-stimulated human platelets. *Anal Bioanal Chem* 2003;376:973–993.
49. Moebius J, Zahedi RP, Lewandrowski U, Berger C, Walter U, Sickmann A. The human platelet membrane proteome reveals several new potential membrane proteins. *Mol Cell Proteomics* 2005;4:1754–1761.
50. Coppinger JA, Cagney G, Toomey S, Kislinger T, Belton O, McRedmond JP, Cahill DJ, Emili A, Fitzgerald DJ, Maguire PB. Characterization of the proteins released from activated platelets leads to localization of novel platelet proteins in human atherosclerotic lesions. *Blood* 2004;103:2096–2104.
51. Distler JH, Huber LC, Gay S, Distler O, Pisetsky DS. Microparticles as mediators of cellular cross-talk in inflammatory disease. *Autoimmunity* 2006;39:683–690.
52. George FD. Microparticles in vascular diseases. *Thromb Res* 2008;122 (Suppl 1):S55–S59.
53. Piccin A, Murphy WG, Smith OP. Circulating microparticles: Pathophysiology and clinical implications. *Blood Rev* 2007;21:157–171.
54. VanWijk MJ, VanBavel E, Sturk A, Nieuwland R. Microparticles in cardiovascular diseases. *Cardiovasc Res* 2003;59:277–287.
55. Shet AS. Characterizing blood microparticles: Technical aspects and challenges. *Vasc Health Risk Manag* 2008;4:769–774.
56. Jy W, Horstman LL, Jimenez JJ, Ahn YS, Biro E, Nieuwland R, Sturk A, gnat-George F, Sabatier F, Camoin-Jau L, Sampol J, Hugel B, Zoubairi F, Freyssinet JM, Nomura S, Shet AS, Key NS, Heibel RP. Measuring circulating cell-derived microparticles. *J Thromb Haemost* 2004;2:1842–1851.
57. Garin J, Diez R, Kieffer S, Dermine JF, Duclos S, Gagnon E, Sadoul R, Rondeau C, Desjardins M. The phagosome proteome: Insight into phagosome functions. *J Cell Biol* 2001;152:165–180.
58. Gruenberg J, Howell KE. Membrane traffic in endocytosis: Insights from cell-free assays. *Annu Rev Cell Biol* 1989;5:453–481.
59. Lamond AI, Earnshaw WC. Structure and function in the nucleus. *Science* 1998;280:547–553.
60. Verma R, Chen S, Feldman R, Schieltz D, Yates J, Dohmen J, Deshaies RJ. Proteasomal proteomics: Identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Mol Biol Cell* 2000;11:3425–3439.
61. Yates JR III, Gilchrist A, Howell KE, Bergeron JJ. Proteomics of organelles and large cellular structures. *Nat Rev Mol Cell Biol* 2005;6:702–714.
62. Brunet S, Thibault P, Gagnon E, Kearney P, Bergeron JJ, Desjardins M. Organelle proteomics: Looking at less to see more. *Trends Cell Biol* 2003;13:629–638.
63. Bledi Y, Inberg A, Linal M. PROCEED: A proteomic method for analyzing plasma membrane proteins in living mammalian cells. *Brief Funct Genomic Proteomic* 2003;2:254–265.
64. Kiss E, Nagy P, Balogh A, Szollosi J, Matko J. Cytometry of raft and caveola membrane microdomains: From flow and imaging techniques to high-throughput screening assays. *Cytometry Part A* 2008;73A:599–614.
65. Barlage S, Boettcher D, Boettcher A, Dada A, Schmitz G. High density lipoprotein modulates platelet function. *Cytometry Part A* 2006;69A:196–199.
66. Grandl M, Bared SM, Liebisch G, Werner T, Barlage S, Schmitz G. E-LDL and Ox-LDL differentially regulate ceramide and cholesterol raft microdomains in Human Macrophages. *Cytometry Part A* 2006;69A:189–191.
67. Orso E, Werner T, Wolf Z, Bandulik S, Kramer W, Schmitz G. Ezetimib influences the expression of raft-associated antigens in human monocytes. *Cytometry Part A* 2006;69A:206–208.

68. Wolf Z, Orso E, Werner T, Klunemann HH, Schmitz G. Monocyte cholesterol homeostasis correlates with the presence of detergent resistant membrane microdomains. *Cytometry Part A* 2007;71A:486–494.
69. Wolf Z, Orso E, Werner T, Boettcher A, Schmitz G. A flow cytometric screening test for detergent-resistant surface antigens in monocytes. *Cytometry Part A* 2006;69A:192–195.
70. Sprenger RR, Speijer D, Back JW, De Koster CG, Pannekoek H, Horrevoets AJ. Comparative proteomics of human endothelial cell caveolae and rafts using two-dimensional gel electrophoresis and mass spectrometry. *Electrophoresis* 2004;25:156–172.
71. Anderson RG, Jacobson K. A role for lipid shells in targeting proteins to caveolae, rafts, and other lipid domains. *Science* 2002;296:1821–1825.
72. Foster LJ, de Hoog CL, Mann M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proc Natl Acad Sci USA* 2003;100:5813–5818.
73. Lisanti MP, Scherer PE, Vidugiriene J, Tang Z, Hermanowski-Vosatka A, Tu YH, Cook RF, Sargiacomo M. Characterization of caveolin-rich membrane domains isolated from an endothelial-rich source: Implications for human disease. *J Cell Biol* 1994;126:111–126.
74. Cohen AW, Hnasko R, Schubert W, Lisanti MP. Role of caveolae and caveolins in health and disease. *Physiol Rev* 2004;84:1341–1379.
75. Simons K, Ehehalt R. Cholesterol, lipid rafts, and disease. *J Clin Invest* 2002;110:597–603.
76. Houde M, Bertholet S, Gagnon E, Brunet S, Goyette G, Laplante A, Princiotta MF, Thibault P, Sacks D, Desjardins M. Phagosomes are competent organelles for antigen cross-presentation. *Nature* 2003;425:402–406.
77. Alexander S, Srinivasan S, Alexander H. Proteomics opens doors to the mechanisms of developmentally regulated secretion. *Mol Cell Proteomics* 2003;2:1156–1163.
78. Schmitz G, Schambeck CM. Molecular defects in the ABCA1 pathway affect platelet function. *Pathophysiol Haemost Thromb* 2006;35:166–174.
79. Diaz LM, Maiya R, Sullivan MA, Han Y, Walton HA, Boehm SL, Bergeson SE, Mayfield RD, Morrisett RA. Sindbis viral-mediated expression of eGFP-dopamine D1 receptors in situ with real-time two-photon microscopic detection. *J Neurosci Methods* 2004;139:25–31.
80. Sanchez SA, Gratton E. Lipid-protein interactions revealed by two-photon microscopy and fluorescence correlation spectroscopy. *Acc Chem Res* 2005;38:469–477.
81. Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. Nuclear proteome dynamics. *Nature* 2005;433:77–83.
82. Andersen JS, Mann M. Organellar proteomics: Turning inventories into insights. *EMBO Rep* 2006;7:874–879.
83. Mathy G, Sluse FE. Mitochondrial comparative proteomics: Strengths and pitfalls. *Biochim Biophys Acta* 2008;1777:1072–1077.
84. Scharfe C, Zaccaria P, Hoertnagel K, Jaksch M, Klopstock T, Dembowski M, Lill R, Prokisch H, Gerbitz KD, Neupert W, Mewes HW, Meitinger T. MITOP, the mitochondrial proteome database: 2000 update. *Nucleic Acids Res* 2000;28:155–158.
85. Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schonfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C. The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc Natl Acad Sci USA* 2003;100:13207–13212.
86. Kaleem Z. Flow cytometric analysis of lymphomas: Current status and usefulness. *Arch Pathol Lab Med* 2006;130:1850–1858.
87. Michelson AD. Evaluation of platelet function by flow cytometry. *Pathophysiol Haemost Thromb* 2006;35:67–82.
88. Dunphy CH. Applications of flow cytometry and immunohistochemistry to diagnostic hematopathology. *Arch Pathol Lab Med* 2004;128:1004–1022.
89. Illoh OC. Current applications of flow cytometry in the diagnosis of primary immunodeficiency diseases. *Arch Pathol Lab Med* 2004;128:23–31.
90. Tarnok A, Gerstner AO. Clinical applications of laser scanning cytometry. *Cytometry* 2002;50:133–143.
91. Gerstner AO, Mittag A, Laffers W, Dahnert I, Lenz D, Bootz F, Boci J, Tarnok A. Comparison of immunophenotyping by slide-based cytometry and by flow cytometry. *J Immunol Methods* 2006;311:130–138.
92. Mosch B, Mittag A, Lenz D, Arendt T, Tarnok A. Laser scanning cytometry in human brain slices. *Cytometry Part A* 2006;69A:135–138.
93. Laffers W, Mittag A, Lenz D, Tarnok A, Gerstner AO. Iterative restaining as a pivotal tool for n-color immunophenotyping by slide-based cytometry 2. *Cytometry Part A* 2006;69A:127–130.
94. Mittag A, Lenz D, Gerstner AO, Tarnok A. Hyperchromatic cytometry principles for cytomics using slide based cytometry. *Cytometry Part A* 2006;69A:691–703.
95. Ecker RC, Tarnok A. Cytomics goes 3D: Toward tissomics. *Cytometry Part A* 2005;65A:1–3.
96. Schubert W. Exploring molecular networks directly in the cell. *Cytometry Part A* 2006;69A:109–112.
97. Gough AH, Johnston PA. Requirements, features, and performance of high content screening platforms. *Methods Mol Biol* 2007;356:41–61.
98. Korn K, Krausz E. Cell-based high-content screening of small-molecule libraries. *Curr Opin Chem Biol* 2007;11:503–510.
99. Krausz E. High-content siRNA screening. *Mol Biosyst* 2007;3:232–240.
100. Wollman R, Stuurman N. High-throughput microscopy: From raw images to discoveries. *J Cell Sci* 2007;120:3715–3722.
101. Heilker R. High content screening to monitor G protein-coupled receptor internalization. *Ernst Schering Found Symp Proc* 2006;229–247.
102. Bentzen SM, Buffa FM, Wilson GD. Multiple biomarker tissue microarrays: Bioinformatics and practical approaches. *Cancer Metastasis Rev* 2008;27:481–494.
103. Eguiluz C, Viguera E, Millan L, Perez J. Multitissue array review: A chronological description of tissue array techniques, applications and procedures. *Pathol Res Pract* 2006;202:561–568.
104. Guilar-Mahecha A, Hassan S, Ferrario C, Basik M. Microarrays as validation strategies in clinical samples: Tissue and protein microarrays. *OMICS* 2006;10:311–326.
105. Voduc D, Kenney C, Nielsen TO. Tissue microarrays in clinical oncology. *Semin Radiat Oncol* 2008;18:89–97.
106. Lopez-Guerrero JA, Riegman PH, Oosterhuis JW, Lam KH, Oomen MH, Spatz A, Ratcliffe C, Knox K, Mager R, Kerr D, Pezzella F, van DB, van dV, van BH, Morente MM, Alonso S, Kerjaschki D, Pammer J, Carbone A, Gloghini A, Teodorovic I, Isabelle M, Passioukov A, Lejeune S, Therasse P, van Veen EB, Dinjens WN, Llombart-Bosch A. TuBaFrost 4: Access rules and incentives for a European tumor bank. *Eur J Cancer* 2006;42:2924–2929.
107. Thallinger GG, Baumgartner K, Pirklbauer M, Uray M, Pauritsch E, Mehes G, Buck CR, Zatloukal K, Trajanoski Z. TAMEE: Data management and analysis for tissue microarrays. *BMC Bioinformatics* 2007;8:81.
108. Nilsson P, Paavilainen L, Larsson K, Odling J, Sundberg M, Andersson AC, Kampf C, Persson A, Al-Khalili SC, Ottosson J, Bjorling E, Hober S, Wernerus H, Wester K, Ponten F, Uhlen M. Towards a human proteome atlas: High-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* 2005;5:4327–4337.
109. Kampf C, Andersson AC, Wester K, Bjorling E, Uhlen M, Ponten F. Antibody-based tissue profiling as a tool for clinical proteomics. *Clin Proteomics* 2007;1:285–299.
110. Andersson AC, Stromberg S, Backvall H, Kampf C, Uhlen M, Wester K, Ponten F. Analysis of protein expression in cell microarrays: A tool for antibody-based proteomics. *J Histochem Cytochem* 2006;54:1413–1423.
111. Donahue MP, Rose K, Hochstrasser D, Vondrascer J, Grass P, Chibout SD, Nelson CL, Sinnavee P, Goldschmidt-Clermont PJ, Granger CB. Discovery of proteins related to coronary artery disease using industrial-scale proteomics analysis of pooled plasma. *Am Heart J* 2006;152:478–485.
112. Uhlen M, Bjorling E, Agaton C, Szigartyo CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnuson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skolleremo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan J, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Ponten F. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* 2005;4:1920–1932.
113. Uhlen M, Ponten F. Antibody-based proteomics for human tissue profiling. *Mol Cell Proteomics* 2005;4:384–393.
114. Lam JS, Belldgrun AS, Figlin RA. Tissue array-based predictions of pathobiology, prognosis, and response to treatment for renal cell carcinoma therapy. *Clin Cancer Res* 2004;10:6304S–6309S.
115. Breimann L, Friedman J, Olshen RA, Stone JC, editors. Classification and regression trees. 1984; Wadsworth Inc.
116. Appel R, Hochstrasser D, Roch C, Funk M, Muller AF, Pellegrini C. Automatic classification of two-dimensional gel electrophoresis pictures by heuristic clustering analysis: A step toward machine learning. *Electrophoresis* 1988;9:136–142.
117. Hernandez P, Gras R, Frey J, Appel RD. Popitam: Towards new heuristic strategies to improve protein identification from tandem mass spectrometry data. *Proteomics* 2003;3:870–878.
118. Burges CJC. A tutorial on support vector machines for pattern recognition. *Knowledge Discovery Data Mining* 1998;121–167.
119. Girolami M, Rogers S. Variational Bayesian multinomial probit regression with Gaussian process priors. *Neural Computation* 2006;18:1790–1817.
120. Kote-Jarai Z, Matthews L, Osorio A, Shanley S, Giddings I, Moreews F, Locke I, Evans DG, Eccles D, Williams RD, Girolami M, Campbell C, Eeles R. Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage. *Clin Cancer Res* 2006;12:3896–3901.
121. Lau KW, Jones AR, Swainston N, Siepen JA, Hubbard SJ. Capture and analysis of quantitative proteomic data. *Proteomics* 2007;7:2787–2799.
122. Pan S, Zhang H, Rush J, Eng J, Zhang N, Patterson D, Comb MJ, Aebersold R. High-throughput proteome screening for biomarker detection. *Mol Cell Proteomics* 2005;4:182–190.
123. Orchard S, Hermjakob H, Apweiler R. The proteomics standards initiative. *Proteomics* 2003;3:1374–1376.
124. Orchard S, Hermjakob H, Julian RK Jr, Runte K, Sherman D, Wojcik J, Zhu W, Apweiler R. Common interchange standards for proteomics data: Public availability of tools and schema. *Proteomics* 2004;4:490–491.
125. Stephan C, Reidegeld KA, Hamacher M, van HA, Marcus K, Taylor C, Jones P, Muller M, Apweiler R, Martens L, Korting G, Chamrad DC, Thiele H, Blugge M, Parkinson D, Binz PA, Lyall A, Meyer HE. Automated preprocessing pipeline for searching heterogeneous mass spectrometric data of the HUPO brain proteome project pilot phase. *Proteomics* 2006;6:5015–5029.
126. Dombon B, Aebersold R. Challenges and opportunities in proteomics data analysis. *Mol Cell Proteomics* 2006;5:1921–1926.
127. Valet G. Cytomics as a new potential for drug discovery. *Drug Discov Today* 2006;11:785–791.
128. Diks SH, Peppelenbosch MP. Single cell proteomics for personalized medicine. *Trends Mol Med* 2004;10:574–577.
129. Valet G, Leary JF, Tarnok A. Cytomics: New technologies: Towards a human cytoome project. *Cytometry Part A* 2004;59A:167–171.