Cancer impacts each patient and family differently. Our current understanding of the disease is primarily limited to clinical hallmarks of cancer, but many specific molecular mechanisms remain elusive. Genetic markers can be used to determine predisposition to tumor development, but molecularly targeted treatment strategies that improve patient prognosis are not widely available for most cancers. Individualized care plans, also described as personalized medicine, still must be developed by understanding and implementing basic science research into clinical treatment. Proteomics holds great promise in contributing to the prevention and cure of cancer because it provides unique tools for discovery of biomarkers and therapeutic targets. As such, proteomics can help translate basic science discoveries into the clinical practice of personalized medicine. Here we describe how biological mass spectrometry and proteome analysis interact with other major patient care and research initiatives and present vignettes illustrating efforts in discovery of diagnostic biomarkers for ovarian cancer, development of treatment strategies in lung cancer, and monitoring prognosis and relapse in multiple myeloma patients. *Molecular & Cellular Proteomics* 7:1780–1794, 2008.

The discovery of the causative genetic underpinnings of cancer has been a focus of biomedical research for decades. The multigenic nature of cancer has hindered progress in understanding the underlying mechanisms that lead to a specific disease phenotype. Recent advances in high throughput technologies, which evaluate tens of thousands of genes or proteins in a single experiment, are providing new methods for identifying biochemical determinants of the disease process. To facilitate these technologies, the correlation of specific phenotypes to individual genotypes is key to leveraging the use of model organisms and patient samples in cancer research. Integration of these data allows cancer researchers to ask complex questions about the mechanism of specific disease manifestations and to retrieve data sets containing disparate data that can be further analyzed using statistical methods to reveal new insights that should be further investigated.

With the comprehensive cataloging of human genes and links between gene function and disease, the future of medicine looks toward mechanistic personalized medicine approaches to cure diseases such as cancer. Using arrays that can profile gene expression, many groups have been able to define gene expression signatures related to diagnosis (cancer versus benign, subtype of leukemia, etc.), prognosis (likelihood of cure), and prediction (probability of response to therapy). Although most of these approaches remain in the research domain, some have been thrust into the mainstream of standard clinical practice, e.g., Oncotype DX® for prediction of breast cancer recurrence. Proteomics will be next in line to deliver new tools to help patients with cancer live longer and have a better quality of life.

Proteomics is an emerging field that can make unique contributions to the prevention and cure of cancer. From strength in protein sequence analysis to broad scale cataloging of proteins and post-translational modifications, a wide variety of proteomics tools are available to effect changes in patient care. Proteomics has the advantage over genomics-based assays because of direct examination of the molecular machinery of cell physiology, including protein expression, sequence variations and isoforms, post-translational modification, and protein-protein complexes. However, certain disadvantages also exist, including (i) stringent requirements for sample collection, preparation, and analysis, (ii) lack of amplification procedures similar to PCR that can allow assay development using limited biological starting material, (iii) requirements for purification strategies to enrich samples for intended work (e.g., phosphoprotein analysis), and (iv) costs necessary for staffing and equipping a shared resource or clinical laboratory able to perform the required assays. Nonetheless proteomics techniques should be implemented with basic clinical medicine along with DNA- and/or mRNA-based profiling strategies to enhance cancer screening, diagnosis, treatment, and follow-up.

An overview of the potential of these cutting edge technologies in the development of personalized medicine has recently been presented by Dalton and Friend (1). Here we build
on that foundation and illustrate roles for proteomics in the interaction between research and clinical practice with specific vignettes. To visualize how proteomics may contribute to the development of personalized medicine, researchers must have an understanding of the patient’s journey from cancer diagnosis through treatment as shown in Fig. 1A. Cancer may be detected through routine check-ups, self-exams, or following the presentation of specific symptoms. Any or all of these factors may contribute to the diagnosis of the incoming patient. At this point, staging and molecular profiling will also be performed using samples obtained by tumor biopsy as well as blood and/or urine collection. The development of personalized cancer care has several goals that impact current and future patients: (i) identify needs of the individual patient, (ii) identify biomarkers to predict needs and risks, (iii) develop and implement methods for minimally invasive patient sampling, (iv) match the right treatment to each patient, (v) improve the performance of clinical trials through molecular profiling, and (vi) raise the standard of care by partnering with other hospitals and clinical care centers.

Using the diagnosis, staging, and molecular profiles, a physician can assess the patient’s prognosis and predict potentially effective therapies. The patient should be directed to treatment regimens based on drugs with the proper mechanism of action. The outcome of this step is directed treatment, which optimizes the patient’s survival chances and quality of life. Cancer survivors must be monitored for relapse or recurrence as well as the development of new cancers; frequently they will (re-)enter screening programs. As a consequence of a cancer diagnosis, members of the patient’s family and caregivers may choose to enroll in a screening program as well. The discovery, development, and implementation of biomarkers for ongoing monitoring are also critical to clinical practice.

Communication of the challenges in treatment enables researchers to contribute to clinical practice (Fig. 1B). The interaction between clinicians and researchers must be very strong for iterative examination of clinical practice and the development of personalized medicine. Here specific case studies illustrate potential roles for proteomics in improving patient care. Targeted and broad scale proteomics experi-
ments have been implemented for the discovery of diagnostic biomarkers in ovarian cancer (Vignette I). Phosphoproteomics contributes to preclinical models for directed tyrosine kinase inhibitor treatment of lung cancer (Vignette II). Detection of disease progression in multiple myeloma with quantitative mass spectrometry illustrates aspects of ongoing patient assessment (Vignette III). Finally we discuss institutional infrastructure and an example of successful implementation molecular biomarkers into personalized cancer care.

VIGNETTE I: PROTEIN BIOMARKER DISCOVERY: EARLY DETECTION OF OVARIAN CANCER

Most patients with ovarian cancer have widespread metastatic disease at initial diagnosis largely because of the inability to detect ovarian cancer at an early stage (2–5). There is currently no proven, effective method for early detection of ovarian cancer through biomarkers, imaging, or other means (6–11). The most commonly used biomarker for ovarian cancer, CA125 (12), is elevated in only about 50% of stage I ovarian cancer cases (8). Beyond the lack of effective detection, there is no accurate method for diagnosis of ovarian cancer short of surgery; even among symptomatic women, tissue evaluation by a pathologist is the only reliable way to distinguish between women with benign and malignant disease. Because of these limitations, ovarian cancer is detected at later stages where patients have very poor prognosis and few treatment options. Early detection significantly improves patient outcomes.

The need for determining additional biomarkers for early detection of ovarian cancer that complement CA125 has been reviewed in great detail from many perspectives, recently including proteomics. Discovery of appropriate biomarkers would enable population screening and personalized care. An extensive list of candidates has been prepared by Williams et al. (13) that includes proteins, glycans, lipids, and metabolites. However, Jacobs and Menon (14) describe the difficulties inherent in screening for ovarian cancer; in particular, the required specificity would need to be essentially 100% to produce a biomarker with sufficient positive predictive value. None of the existing candidates have been implemented in patient care strategies, so discovery efforts continue. Although these reviews shed light on the challenges and prior candidate biomarkers, here we illustrate how early detection of disease has been a proving ground for proteomics strategies and discuss the successes of these cancer biomarker discovery efforts.

Because of the difficult nature of this task and the desire to have unbiased approaches, researchers have applied the entire menu of proteomics tools to finding novel candidate biomarkers. Often these experiments are used as a proving ground for analytical technology, and the rigorous requirements for sample collection, processing, and analysis are determined retrospectively for improvement in the next round of sample analysis. The ongoing development and complexity of proteomics as well as calls to standardize experiments across institutions have raised many dilemmas for investigators. The challenge in broad scale plasma proteome analysis is reflected in the fact that targeting methods rarely have the depth to detect clinically relevant molecules released from a tumor, and proteome cataloging experiments are not practical for case-control studies of sufficient population to detect statistically significant differences. Furthermore each proteomics experiment will have specific strengths and weaknesses because of the method of protein or peptide selection and the type of visualization or detection chosen (15–17).

Mass spectrometry profiling and intact protein separations, including two-dimensional gel electrophoresis (2DE) and multidimensional liquid chromatography, have been used to target differences between cancer patients and controls; proteomes from healthy controls and cancer patients have been cataloged using LC-MS/MS shotgun sequencing. Each technique will be reviewed to illustrate its strengths and the data resources that have been produced. Although pattern analysis is integral to the detection and targeting of candidate biomarkers, approaches that rely on fingerprinting alone, without identified target molecules, will be omitted because the addition of novel diagnostic molecules will bring the most value to patient care.

Differential Display Techniques: Mass Spectrometry Profiling—MS profiling has been shown to be an effective method for detecting differences in plasma of cancer patients and controls. This method is attractive because the approach can be rapid, and parallel processing enables the high throughput required for clinical sample analysis. Implementation appears deceptively simple. Despite controversy around the initial report using selection by surface retentate chromatography and mass analysis with MALDI MS to fingerprint ovarian cancer patients, controls, and patients with benign disease (18), MS profiling approaches have substantially improved even in light of the limited number of proteins or peptides that can be detected (19). After initial investigations of chemical fractionation methods using reverse phase, ion exchange, and IMAC, most of the components of the low molecular weight serum/plasma proteome were found to be intact highly abundant proteins or proteolytic fragments of plasma proteins. This method of selection and detection is limited in sensitivity, peak capacity, and dynamic range; therefore, it is unlikely to detect components at levels below 1 μM unless it is combined with immunoprecipitation (20).

---

1 The abbreviations used are: 2DE, two-dimensional gel electrophoresis; MM, multiple myeloma; MGUS, monoclonal gammopathy of undetermined significance; MRM, multiple reaction monitoring; EIC, extracted ion chromatogram; TKI, tyrosine kinase inhibitor; ITIH4, inter-α-trypsin inhibitor heavy chain H4; 2D, two-dimensional; SH2, SRC homology 2; GAP, GTPase-activating protein; STAT, signal transducers and activators of transcription; EGFR, epidermal growth factor receptor; ERK, extracellular signal-regulated kinase; SPEP, serum protein electrophoresis; IFE, immunofixation electrophoresis; HER2, human epidermal growth factor receptor 2.
Nevertheless MS profiling has detected differences between ovarian cancer patients and controls, including the α subunit of haptoglobin (21) and a panel consisting of apolipoprotein A-I, transthyretin, and inter-α-trypsin inhibitor heavy chain H4 (ITIH4) (22). Subsequent work using capture by immobilized antibodies shows variations in ITIH4 processing in several types of cancer when compared with controls (23). Further investigation of transthyretin has revealed differences in redox modifications including cysteinylation and glutathionylation (24). Clearly the extensive characterization of targeted molecules produces the most value because specific molecular markers can be determined. In addition, MS profiling highlighted the role of protease activity in serum and plasma that may also be used to distinguish cancer patients from controls (25, 26).

**Differential Display Techniques: 2DE and Multidimensional LC Protein Separations**—Higher molecular weight proteins have been analyzed by 2DE and multidimensional liquid phase separations. These approaches enable the assessment of changes to intact molecules including proteolytic processing and other post-translational modifications like glycosylation or phosphorylation. After protein identification, these modifications can be investigated further in the hope of developing an assay for detecting the specific differentially expressed isoforms.

In fluorescence DIGE, each sample is labeled with a different fluorescent probe, combined, and separated by isoelectric focusing and SDS-PAGE. Differences in protein expression can be distinguished as spots that are either red or green; complete overlap is represented by yellow (Fig. 2A). This approach has been popular because of the ease of interpreting the fluorescence images, the high number of protein spots (typically > 1,000), and the fact that the cancer and control samples are processed together. In addition, post-translational modifications can be targeted directly in 2D gel approaches using antibodies or specific fluorescent stains that recognize phospho- and glycoproteins.

In ovarian cancer biomarker discovery, 2DE techniques have been used to examine the isoforms of abundant serum proteins, indicating differences in phosphorylated fibrinogen α (27) as well as haptoglobin and transferrin (28). Another study presented several potential protein biomarkers, including complement components, serum glycoproteins, serum protease inhibitors, transferrin, and afamin (29). The last of these markers was further verified by ELISAs and compared with C-reactive protein and CA125. DIGE or targeted staining techniques in 2D gel analysis can provide information about post-translational modifications to abundant proteins in the plasma.

Multidimensional liquid phase separations complement 2DE. Proteome fractionation approaches have evolved to
include immunodepletion of the most abundant plasma proteins (e.g. top six or top 12 removal) followed by ion exchange, liquid phase isoelectric focusing or chromatofocusing, and reverse phase separations of the lower abundance components. The intensity of the proteins in the final separation, whether reverse phase chromatography or SDS-PAGE, is compared with select targets for protein identification as shown in Fig. 2B. These data were acquired using the commercialized version of 2D protein LC (PF2D, Beckman Coulter). The targeted fractions are recovered, digested with trypsin, and submitted for protein identification using LC-MS/MS peptide sequencing. Current results implicate abundant plasma proteins as candidate markers for ovarian cancer. Even in mouse models with enormous tumor burden, the most prevalent differences correspond to host response or abundant plasma proteins (30).

Protein Catalogs Created by Shotgun Sequencing as Resources for Biomarker Discovery—The role of LC-MS/MS shotgun sequencing in the proteome analysis of biofluids has created resources that can be exploited for biomarker discovery. Through a series of analytical improvements, the human plasma proteome has been extensively cataloged by Smith and co-workers (31–33). In addition, the Human Proteome Organisation (HUPO) plasma proteome project has created a reference for more than 3,000 proteins identified in plasma (34–36), including the corresponding gene ontology terms (37). Extensive sequencing efforts have now identified more than 1,500 proteins from the human urinary proteome (38) using healthy samples. In addition, ascites fluid from ovarian cancer patients has been extensively analyzed in a recent publication by Gortzak-Uzan et al. (39). Each of these protein catalogs could be scanned to reveal candidate biomarkers based on disease etiology or organ site.

In addition to creating protein catalogs, (semi-)quantitative measurements can be derived from LC-MS/MS analysis of groups of patients and controls. The peptide counting statistics, which describe the number of peptides or tandem mass spectra assigned to sequences from a given protein, can be used to estimate the relative amount of a protein in a complex mixture as shown for haptoglobin in Fig. 2C. The intensity of the intact peptide in the mass spectra can also be used to quantify the relative amount of protein in each sample; the peak areas are calculated for each individual peptide using extracted ion chromatograms (EICs) as shown for two peptides from haptoglobin in Fig. 2D. Although the primary goal of LC-MS/MS is the catalog of proteins generated by peptide sequence assignments, EIC analysis can provide excellent data on the relative expression level of proteins in cancer patients’ and controls’ plasma. Furthermore the protein and its representative peptides need only be identified with high confidence in one of the samples. With reproducible chromatography and accurate mass measurement, the peaks in the other samples can be correlated prior to EIC analysis. This approach has been termed “label-free” proteomics, and it has been widely and effectively used to characterize biological and clinical samples (40).

Narrowing the Field of Candidate Biomarkers with Quantitative Mass Spectrometry—A strong case can be made that extensive verification efforts using quantitative mass spectrometry should be the next step for biomarker development for ovarian cancer. Using molecules described above in curated reviews (13) or from tissue proteome analyses (41–44) investigators could apply absolute quantification to evaluate many candidates in the same sample in a single analytical experiment. Peptides detected during protein identification experiments can be immediately used for quantitative mass spectrometry analysis. Multiple reaction monitoring is typically used to specifically quantify individual peptides, which represent their proteins of origin (an additional description is included in Vignette III). Even in complex matrices like plasma, individual peptides can be monitored effectively (45–47). Multiplexing strategies have also proven to be effective; Anderson and Hunter (48) used LC-multiple reaction monitoring (MRM) to develop a quantitative assay for 53 plasma proteins, illustrating the breadth of targets that could be accessed in a single analysis. The quantities of the proteins can be plotted by sample group, illustrating the potential utility in separating cancer patients from controls, as shown in Fig. 2E. Overlapping distributions (left) will not make effective candidates; proteins expressed at higher levels in most cancer patients (right) can be further validated by quantitative mass spectrometry or immunoassays using larger sample groups. Many of the current candidate biomarkers for detection of ovarian cancer could be better defined or ruled out using quantitative assays, including MRM or other high resolution LC-MS techniques.

VIGNETTE II: PRECLINICAL MODELING OF TREATMENT STRATEGIES: EXAMINATION OF ONCOGENIC TYROSINE KINASE SIGNALING AND TYROSINE KINASE INHIBITION USING PHOSPHOPROTEOMICS

Following diagnosis, each patient is placed on a particular treatment regimen. At present, few if any of the broadly deployed strategies are molecularly driven. Proteomics can be used for preclinical modeling and probing archived tissue sections for biomarkers of response or resistance. These investigations begin by matching the appropriate proteomics tools to the clinical problem and relevant biological pathways. The application of phosphoproteomics holds great promise for understanding oncogenic signaling pathways and developing biomarkers that could be predictive for patient outcome on specific drug regimens. The focus here is to describe how proteomics technology can be applied to the study of tyrosine kinase signaling pathways and tyrosine kinase inhibitors (TKIs) in cancer. The important points are: (i) signaling pathways are assembled in distinct modules, (ii) many of the targets within these modules have inhibitors moving toward clinical use, (iii) assays that predict function/activity/dependence of these modules may allow for personalized therapy, (iv)
the existence of interchangeable modules produces complexity in signaling pathways, and (v) existence of redundant modules and/or complex networks suggests the need for combinatorial strategies for future clinical trials. Proteomics can have a major impact on identifying these modules, developing pharmacodynamic assays, and unraveling the complexity of signaling networks.

**Targeting Oncogenic Kinase Signaling Pathways**—Normal cell physiology is controlled by proteins within the cell that act together similar to an electric circuit to ensure normal cell behavior. Cancer is a disease where these circuits are dysregulated or rearranged in such a way that the output drives the cell toward excessive growth and spread to unintended areas of the body. Knowledge of these individual molecules and cohesive signaling modules can help identify proteins involved in driving a particular cancer cell and suggest combinatorial therapeutic strategies. Critical components of these circuits are signaling proteins called kinases that act as relays and regulate the activity of other important genes and proteins. Protein kinase signaling pathways regulate the “hallmarks of cancer” including cell growth, survival, invasion/metastasis, and angiogenesis (49). Not surprisingly, it has been known for quite some time that aberrant kinase signaling can lead to tumorigenesis. Notable examples of tyrosine kinases driving cancer are viral SRC and the breakpoint cluster region protein and Abelson murine leukemia viral oncogene homolog 1 (BCR-ABL) fusion protein that displays constitutive kinase activity in chronic myelogenous leukemia (50).

Considerable enthusiasm continues to focus on targeting aberrant kinase pathways in lung cancer. Both tyrosine and serine/threonine kinases are under investigation as are the pathways they regulate. Sequencing of the human genome identified nearly 100 tyrosine kinase proteins, some of which are known to be involved in the pathogenesis of cancer as well as other tyrosine kinases with potential (as yet undefined) roles. Tyrosine kinases can either span the cellular membrane and become activated by extracellular ligands (receptor tyrosine kinase) or exist as intracellular proteins activated by intracellular events (non-receptor tyrosine kinases). The catalytic core subunit of tyrosine kinases recruits ATP and phosphorylates a tyrosine residue on substrate (downstream) proteins. In some cases, the tyrosine kinase autophosphorylates itself on specific sites, leading to enhanced function as well as producing a potential biomarker for activated kinase. For example, SRC family members can both phosphorylate downstream substrates as well as autophosphorylate itself on tyrosine 419. Because this event leads to enhanced catalytic activity, the degree of autophosphorylation serves as a biomarker of SRC activity in tumor cells. Phosphorylated tyrosine residues on substrate proteins change cellular physiology by modifying protein functions, including enzyme activity, subcellular localization, and/or protein-protein interactions. Important for relaying signaling, phosphorylated tyrosine sites can act as docking sites for proteins containing SRC homology 2 (SH2) domains (51). The human genome encodes ~110 distinct SH2 domain-containing proteins, and although these domains are generally conserved they still retain enough variability to lead to specificity in signaling (52). SH2-containing proteins have diverse functions including adaptors (Grb2), scaffolds (Shc), kinases (SRC), phosphatases (Shp2), Ras signaling (RasGAP), transcription (STAT), ubiquitination (Cbl), cytoskeletal function (Tensin), and phospholipid second messenger signaling (Polo-like kinase). Signaling originating from tyrosine kinases pass through individual substrate proteins and interactions with SH2 proteins, which are ultimately linked to downstream effectors. These common sets of effector pathways include Ras/Raf/mitogen-activated protein kinase/extracellular signal-regulated kinase kinase (MEK)/ERK signaling modules, STAT signaling modules, phosphatidylinositol 3-kinase/Akt/mammalian target of rapamycin (mTOR) signaling molecules, protein kinase C modules, and others. These effector cascades regulate downstream proteins, some of which include transcription factors (DNA-binding proteins) that modulate gene expression. As a whole, dysregulation of these pathways alters cellular physiology and produces malignant behavior in cells.

The expression of individual tyrosine kinases, substrates of individual tyrosine kinases, SH2 domain proteins, components of effector cascades, and genomic alterations in lung cancer cells allow for modular signaling networks to be created that can be unique for each particular cell or tumor. Thus, tumor cells produce complex signal transduction networks that can be attacked at different points in an attempt to either kill tumor cells or revert the cells to benign function. The major hurdle is to determine which set of signaling proteins is active and relevant for an individual’s tumor. Thus, although we have a deeper understanding of how signaling networks are created, assays to determine the active signaling proteins in a patient’s tumor require further development.

The future of cancer treatment will be based on personalized approaches that identify the critical molecules necessary for tumor growth and survival and match patients to appropriate molecularly directed therapy. If practicing clinicians can match a kinase inhibitor with an individual patient survival would dramatically improve, and toxicity could be reduced. Perhaps the best example is the use of imatinib for BCR-ABL-dependent chronic myelogenous leukemia (53). The success of imatinib was heavily influenced by the knowledge of BCR-ABL signaling and its critical importance to leukemia cells. Similar stories include the use of imatinib for gastrointestinal stromal tumors (50), Herceptin for HER2-overexpressing breast cancer, and gefitinib/erlotinib for EGFR mutant-driven lung cancers. However, given the apparent low rates of kinase mutations in the human genome from cancers, obvious mutations in key kinases may be rare, and a large group of patients may have tumors driven by kinases that are not mutated (54). In addition, patients without activating mutations may benefit from kinase inhibitors, such as the case of
lung cancer patients treated with erlotinib (55–57). Improved outcomes may also appear with logical combinations of inhibitors that target important network “hubs” that regulate tumor survival. Complex proteomics approaches enable discovery of biomarker panels for tumors without obvious genomic mutations in critical tyrosine kinases.

**Identifying Biomarkers That Predict Clinical Outcome following TKI Treatment**—The following sections describe tools and techniques as well as current results that define tyrosine kinase signaling networks before and after drug treatment and assist in development of personalized therapy. A flowchart that pairs proteomics experiments with biochemistry/molecular biology, animal models, and early phase clinical trials is shown in Fig. 3. Chemical proteomics can be used to identify drug targets by affinity chromatography and subsequent protein identification. Modulation of kinase activity can be measured by the amount of autophosphorylation and modification of specific downstream substrates using phosphotyrosine selection and LC-MS/MS; quantitative mass spectrometry measurements can also be incorporated to evaluate the magnitude of the changes.

**Identifying Novel Drug-Kinase Interactions through Chemical Proteomics**—Because of the conserved nature of tyrosine kinase domains in tyrosine kinases, small molecule inhibitors designed to inhibit one tyrosine kinase protein can often be “dirty” molecules and have effects on other tyrosine kinase proteins. For example, imatinib was found to have inhibitory effects on c-Kit, and this finding was exploited for the successful treatment of gastrointestinal stromal tumors (58). Studies examining the binding of compounds to individual tyrosine kinases reveal a spectrum of specificity ranging from compounds that bind to few tyrosine kinases to compounds that bind to numerous (>20) tyrosine kinases (59). In addition to binding partner identification, these studies also have the ability to derive quantitative information regarding inhibitor binding and selectivity (60). More recent studies have examined entire libraries of tyrosine kinase inhibitors to produce novel drug-protein interactions that can be exploited for future therapeutic benefit. Similar studies have also highlighted the ability of chemical proteomics approaches, derivatizing the drugs to a stationary phase for affinity chromatography, to identify serine/threonine kinases bound to TKIs as well as non-kinase substrates of TKIs (61, 62). Thus, in the future as mapping of an individual’s tumor tyrosine kinase profile becomes available, it may be possible to match tumor tyrosine kinase dependence with compounds or mixtures of compounds that bind and inhibit the driver kinases. It is likely that existing compounds have inhibitory actions beyond that of their original design, and these could be used for individual patients. Adverse effects of compounds could also be related to off-target inhibition identified through such screens.
Use of Emerging Technologies and Biomarkers to Identify Aberrant Kinase Signaling and Predict Response to Targeted Therapies—Identifying patients that will benefit from kinase inhibitors remains a critical problem. Some examples of possible assays to predict sensitivity to kinase inhibitors include mutation analysis on genomic DNA, evaluation of gene amplification (fluorescence in situ hybridization), immunohistochemistry, and gene expression analysis (63, 64). Emerging technologies that produce robust proteomics analysis will further characterize signaling pathways that can be exploited for therapeutic purposes and may provide additional information relevant for patient selection and/or monitoring. MS-based proteomics may be helpful to identifying tumor cells dependent on kinases for growth and/or survival (65). However, successful implementation requires enrichment because phosphorylated tyrosine residues (Tyr(P)) represent only 0.5% of the total phosphorylated amino acids within a cell (66). Proteomics techniques have been coupled with anti-Tyr(P) antibodies to purify Tyr(P) proteins or proteolytic Tyr(P) peptides for LC-MS/MS analysis (Fig. 4). Phosphotyrosine proteomics has been used to characterize protein networks and pathways downstream of oncogenic HER2 and BCR-ABL (67–69). These methods can also be used to identify novel tyrosine phosphorylation sites and identify oncogenic proteins resulting from activating mutations in protein tyrosine kinases (68–71). The data can then be used in either expert literature curation or machine learning techniques to synthesize network models that can be further evaluated (67). These methodologies can be coupled with TKIs or other compounds to further understand their effect on protein networks. Identification of critical tyrosine kinase proteins in an important oncogenic network may also suggest “druggable” targets that can be entered into therapeutic discovery research.

To illustrate the utility of such an approach, a global survey of phosphotyrosine signaling was performed in both lung cancer cell lines and primary tumors (72). This analysis identified a number of previously identified tyrosine kinases important in lung cancer including HER family proteins, hepatocyte growth factor receptor, vascular endothelial growth factor receptor, IGF-1R, and SRC as well as tyrosine kinases not recognized to be important in lung cancer pathogenesis such as human homolog of avian virus, anaplastic lymphoma kinase, adhesion-related kinase and platelet-derived growth factor receptor. From these experiments, novel causative agents, such as fusion proteins incorporating ALK and ROS and aberrant platelet-derived growth factor receptor α activation, were identified; furthermore sensitivity to imatinib was shown in a small subset of cell lines and tumors. Finally clustering analysis suggests distinct groups of tumors expressing active tyrosine kinases and substrate proteins thus offering the possibility of identifying tumor subsets driven by groups of tyrosine kinases and subsets of patients appropriate for combinations of tyrosine kinase inhibitors in clinical trials. This approach suggests that proteomics analysis can discern individual tumor wiring circuitry that can be exploited for therapeutic benefit.

As a complement to cataloging phosphotyrosine proteins, SH2 profiling can be used to examine phosphotyrosine signaling in cancer cells. As discussed above, tyrosine phosphorylation on proteins serves as docking sites for proteins with SH2 domains. Screening assays using nearly the entire complement of human SH2 domains have been used to profile phosphotyrosine signaling in cancer cells (52). This approach was able to discern differences in SH2 profiles (and therefore phosphotyrosine signaling) in cells transformed by distinct oncoproteins. Expansion of this approach to improve feasibility and enable quantitative analysis could allow for wide scale profiling of cancers (73).

Quantitative Approaches to Evaluate TKI Target Modulation—Although identification of tyrosine kinases in the active state in an individual tumor strongly complements studies showing that a particular TKI can bind that target, the two results do not necessarily predict that the drug can down-regulate the target activation state in vivo and that loss of the target function will translate into an effect on tumor physiology.
(i.e., halt cell proliferation and/or initiate cell death). The typical approach is to treat either tumor cell lines or xenografts with TKI and then examine the effects on tumor growth along with analysis of critical downstream signaling pathways such as phosphatidylinositol 3-kinase/Akt, STATs, and mitogen-activated protein kinase (MAPK)/ERK signaling. However, such an approach is biased because it only examines known pathways and requires reagents including specific phosphoantibodies. Mass spectrometry provides an approach whereby all detectable activated kinases and substrate phosphoproteins are curated; the changes in all of these phosphoproteins can be examined in the same experiment. This approach could discern unique mechanisms of drug action acting through pathways not currently believed to be either drug targets or outside of canonical signaling pathways. In addition, off-target drug effects may be elucidated that (i) predict toxicity of compounds and (ii) identify additional drug targets involved in other disease processes. A number of quantitative mass spectrometry approaches have been reported and are continuing to be improved for characterizing phosphosignaling including stable isotope labeling by amino acids in cell culture (SILAC), isobaric tags for relative and absolute quantification (iTRAQ), ICAT, MRM, and label-free approaches such as spectral counting and peak area calculations from EICs (74–77). One approach and representative data are shown in Fig. 4. Lung cancer cell lines are treated with TKI (black) or vehicle controls (white). Lysates are prepared; the proteins are denatured, reduced, alkylated, and digested with trypsin. After buffer exchange, the Tyr(P) peptides are isolated by immunoprecipitation with immobilized antibodies. After elution, cataloging experiments are performed with LC-MS/MS. In addition to sequence assignments, the phosphorylation sites are identified. Each site can be quantified by EICs and calculating the total peak area for that peptide. By comparison of the peptide ion signals, the relative amounts of phosphorylation in the vehicle control (white) and drug treatment (black) groups can be determined as revealed in Fig. 4.

Ultimately the incorporation of these MS-based data sets into a coherent preclinical model will impact patient care. The flowchart shown in Fig. 3 represents one approach incorporating both preclinical and clinical information to better understand tyrosine kinase inhibitors and their impact on patient outcome. Direct binding targets of tyrosine kinase inhibitors can be discovered through chemical proteomics, and the expression (or level of modification) of these putative targets can be examined through shotgun phosphoproteomics. Modulation of these targets can be examined through quantitative methods described above using clinically achievable concentrations of compounds identified through early phase clinical trials in patients. Inhibition of targets ultimately changes the downstream network of signaling pathways and affects hallmarks of cancer (growth, survival, angiogenesis, etc.). These studies also have the potential to lead to novel assays that can feed back to the clinic, such as pharmacodynamic assays in tumor tissue or blood to examine target modulation in vivo. However, major challenges still remain. First, the amount of data gathered by these approaches identifies a number of phosphotyrosine sites (or other post-translational modifications) that have unclear biological significance. For some proteins such as EGFR and SRC, a large body of published work has defined the functional significance of each Tyr(P) site. For many other proteins, this is not the case. Thus, more basic biochemical experiments are still necessary and important to define the functional significance of certain post-translational modifications. Second, the interactions between different phosphoproteins in the signaling networks are poorly defined. Cancer circuitry likely has tremendous plasticity that allows for rewiring depending on changes in expression patterns, localization, environmental cues, and post-translational modifications. Even networks downstream of well studied tyrosine kinases such as EGFR and SRC can be complicated. Third, use of MS-based assays in patient materials is difficult given the limited amount of tumor obtained from core biopsies and the large amounts of starting material required for immunoaffinity purification of tyrosine phosphorylated peptides. Tumor collection procedures are also critical to maintain post-translational modifications, such as phosphorylation; furthermore tumor heterogeneity may complicate interpretation of results. Methods for assessing each cell in a population and sampling small numbers of cells are just now becoming available, including phosphospecific flow cytometry (78) and capillary isoelectric focusing coupled to antibody-based detection (79). Thus, translation of MS-based assays to other user-friendly assays with robust quantitative characteristics will be important to develop and validate for use in the clinical setting.

**VIGNETTE II: PATIENT MONITORING WITH QUANTITATIVE MASS SPECTROMETRY: ASSESSMENT OF MYELOMA PROGRESSION AND RELAPSE**

Following treatment and partial or complete remission, patients will be reassessed at regular intervals to check for disease relapse or recurrence. Clinical assays based on quantitative mass spectrometry can also play a role in ongoing patient management as illustrated with monitoring multiple myeloma (MM), which is a cancer of the plasma cell that develops primarily in the elderly population. The progression of the tumor is well understood, and it can be diagnosed by the presence of multiple myeloma cells in the bone marrow and detected by the amount of antibody secretion from the clonal population of plasma cells. A premalignant condition known as monoclonal gammopathy of undetermined significance (MGUS) develops at certain rates in the United States population: 3% at age 50, 5% at age 70, and 7% by age 85; ~1% of MGUS patients progress to multiple myeloma on an annual basis (80). The molecular causes for progression from MGUS to MM are unknown. After the onset of the cancer, multiple myeloma patients suffer from several symptoms, including calcium dysregulation, renal failure, anemia, and bone
lesions. A diagnosis of multiple myeloma is established using blood and urine tests. For advanced stage patients, complete skeletal surveys are also used to examine the damage caused by multiple myeloma in the bone marrow. Staging with serum calcium, creatinine, hemoglobin, and most importantly the concentration of the "monoclonal serum protein" was established in 1975 by Durie and Salmon (81). The International Staging System determined in 2005 uses those markers as well as serum albumin and \( \beta_2 \)-microglobulin (82). The survival statistics indicate the importance of early detection and proper staging and show the devastating impact of multiple myeloma. Stage I patients have median survival times of 62 months, stage II patients have median survival times of 45 months, and stage III patients have median survival times reduced to 29 months (82). Novel therapeutic strategies include proteasome inhibition with agents like bortezomib (83, 84), which can be used as a targeted therapy; treatment is more effective for patients with myelomas that secrete high levels of monoclonal antibodies (85). Patient monitoring strategies present significant challenges particularly in the detection of MGUS patients most likely to develop multiple myeloma and ongoing assessment of relapse or recurrence in previously treated multiple myeloma patients. Many MM patients who have undergone treatment are repetitively checked at 2–4-week intervals, leading to high numbers of clinic visits and collection of large volumes of blood.

In the following paragraphs, methods for patient sampling and detection of the monoclonal serum protein are presented from a process chemistry standpoint. Process chemists use extensive background knowledge of synthesis, analysis, and engineering to redesign industrial assembly lines or improve individual steps in manufacturing. In this case, the current clinical methods for analyzing serum from multiple myeloma patients are reviewed; a quantitative mass spectrometry assay for monoclonal proteins is developed and assessed for its value in clinical implementation.

Because each plasma cell secretes a unique antibody, the replication of the tumor cell and the progression of the disease can be monitored by measuring the serum concentration of the monoclonal antibody it produces. Initial qualitative measurements are made using serum protein electrophoresis (SPEP) and dye visualization (see Fig. 5A). Separation of the serum proteins is achieved, isolating albumin from four regions of globulins, termed \( \alpha_1 \), \( \alpha_2 \), \( \beta \), and \( \gamma \), described by the differences in their migration relative to albumin. Normally antibodies migrate into the \( \gamma \) region but are low in intensity compared with albumin and present only as diffuse bands (Fig. 5A, left). The monoclonal antibodies produced in high concentration by multiple myeloma cells can be visualized as a single narrow, discrete, dark band usually in the \( \gamma \) region of the gel (Fig. 5A, right). Patients with abnormally high levels of protein in the \( \gamma \) region can be diagnosed with multiple myeloma after the type of antibody is defined as monoclonal using immunofixation electrophoresis (IFE), which is a separation similar to SPEP but with specific detection for each antibody chain (Fig. 5B). Typical screens test for immunoglobulin G, A, and M heavy chains as well as \( \kappa \) and \( \lambda \) light chains. Immunoglobulin D or E myelomas are very rare; when suspected, lanes of the standard IFE are replaced, enabling specific detection of IgD or IgE heavy chain proteins. In the example, the patient has a tumor that produces an IgG \( \kappa \) monoclonal protein (Fig. 5B). The combination of these two tests establishes the relative amount and type of the antibody that is secreted by the multiple myeloma tumor cells. These gel-based techniques have recently been complemented by capillary array instruments that can analyze eight samples in parallel, greatly increasing the throughput and lowering the amount of sample preparation necessary (tubes of serum are simply loaded into the instrument, which automatically dilutes each sample in the buffer used for capillary electrophoresis).

Even using this new technology, SPEP and IFE are performed separately.

A quantitative mass spectrometry method could replace these methods with a single analysis; this technique is the subject of a forthcoming manuscript. Example bands from SPEP have been processed for protein identification using LC-MS/MS. The detection of peptides from the constant regions of the antibody indicates the peptides that could be developed for quantitative monitoring as shown in Fig. 5C for ALPAPIEK detected from immunoglobulin G heavy chains. After generating peptides for monitoring each of the types of antibodies, a comprehensive method for antibody measurement was made. Briefly minute volumes of serum (1–10 \( \mu l \)) can be processed for detection of each of the antibody chains; G, A, M, \( \kappa \), and \( \lambda \). After protein denaturation with urea, disulfide reduction, and cysteine alkylation, trypsin digestion is performed. The sample is then diluted and analyzed with LC-MRM on a triple quadrupole mass spectrometer (Fig. 5D). The instrument selectively quantifies peptides by filtering the \( m/z \) of the intact species in the first quadrupole (Q1), fragmenting the molecules in the second quadrupole (Q2), and filtering the \( m/z \) of a particular fragment in the third quadrupole (Q3). Each of these peptide and fragment pairs is known as a transition; the instrument measures each transition as part of a cycle, continuously moving from one to the next. For each peptide, multiple transitions are monitored; the coincidence detection of multiple fragments from the peptide increases the confidence in the assignment. Each protein should be quantified using more than one peptide. Although several rules for peptide selection have been put forward, selection of peptide in biological or clinical context frequently deviates from those guidelines. Examples of quantification with LC-MRM are shown in Fig. 5, E and F, for the ALPAPIEK peptide from immunoglobulin G and DSTYLSSTLTLK from the \( \kappa \) light chain. The ion signals corresponding to the \( y_6 \) and

---

y6 ions of ALPAPIEK were detected at 22 min in Fig. 5E; y4 is also monitored. The y8, y11, and y12 ion signals for DSTYSLSSTLTLSK were detected at 28 min as shown in Fig. 5F. These ion signals were obtained from the sample used for the SPEP and IFE, illustrated with the schematic diagrams in Fig. 5, A and B. This quantitative mass spectrometry technique is currently being implemented in parallel with SPEP and IFE in a clinical trial.

Pending successful completion of this evaluation, the quantitative mass spectrometry assay will be used in a broader research context. Its use would be clearly advantageous in animal models where limited amounts of blood serum can be obtained. However, because of the ease of use of SPEP combined with IFE, the clinical impact of such a test may be limited because of the costs associated with and the training necessary for successful mass spectrometry analysis. In the long run, as robotics and automated sample processing become less expensive and more widely applied, this quantitative mass spectrometry assay could be deployed in the clinic. The implementation of a single quantitative test would provide advantages over the qualitative tests currently used to follow multiple myeloma patients. The speed and parallel processing that could be achieved with automated sample handling and MS detection would also significantly improve the throughput of patient samples. The adoption of this technique at a tertiary cancer center could enable surrounding primary care physicians and hospitals to send samples to a centralized facility for processing and analysis. Point of care patient sampling could be performed with rapid turnaround of results to the treating physician (~1 day) even at a centralized facility.
CONCLUSIONS

Proteomics can have great impact on clinical cancer care in the near future. To be successful, experimental design remains the most important step. To answer a specific question, the right methods have to be selected; the required specimens must be collected and stored with the proper techniques. Although a wide variety of tools and techniques are available, directed analysis gives the best chance for proteomics to assist in the delivery of personalized medicine. In cancer, both success, in the form of survivors, and failure, as patients that succumb to disease, must be measured as clinical practice is iteratively improved.

Potential roles for proteomics have been presented using different clinical vignettes. Impact can be made in diverse areas including biomarker discovery, molecular prescription of targeted therapy regimens, and patient monitoring. Each of these descriptions should indicate the difficulty associated with clinical proteomics as well as the importance of close collaboration between researchers and clinicians. The discovery of novel biomarkers is extremely challenging even in the stage of experiment design and technique selection. Phosphoproteomics is making contributions to preclinical modeling for molecularly driven therapy in lung cancer; the combination of chemical proteomics, phosphoproteomics, and SH2 interactions provided detailed mechanistic views of the protein-protein interactions and signaling networks targeted by tyrosine kinase inhibitors. Finally quantitative mass spectrometry can change paradigms in patient sampling for monitoring biomarkers as shown for multiple myeloma. Even with a diagnostic and prognostic biomarker for this disease, many challenges remain for treatment and patient care.

All of these research efforts move forward because of institutional initiatives and support as well as strong community involvement. The contributions of the patients and their families should not be overlooked; many consent to tissue collection and data sharing as they enter cancer prevention, screening, and treatment regimens. At the H. Lee Moffitt Cancer Center and Research Institute, Lifetime Cancer Screening and Total Cancer Care initiatives enroll healthy controls and cancer patients, respectively. Data from healthy people that develop cancer and those that survive the disease create powerful resources for research opportunities in discovery, translation, and delivery of personalized cancer care. Methods for specimen banking and coalescing clinical and research data play an important role in these institutional initiatives; similar to the evaluation of patient care, these institutional protocols must be iteratively evaluated and improved.

Translating protein-based biomarkers into the clinic requires a close collaboration between oncologists, pathologists, statisticians, and clinical trials staff along with infrastructure for real time biomarker analysis. Such an approach can be observed in a recent study using tailored chemotherapy for the treatment of advanced lung cancer based on the expression of two genes (86). Trial participation required a dedicated biopsy of the tumor specifically for gene expression analysis performed by real time quantitative PCR. Predetermined values for gene expression were used for decisions regarding usage of the chemotherapy drugs. The goal of this study was not to compare different treatments but to demonstrate that upfront patient selection with tailored treatment will result in improved outcome. This was found to be the case with data indicating an unprecedented 12-month improvement in survival. The conclusion from this experience is that real time therapeutic decision making based on biomarkers for patients with advanced non-small cell lung cancer is feasible and promising for improvement in patient outcome. However, this approach currently represents a boutique therapy because the expression analysis techniques used require a substantial infrastructure and, therefore, may not be readily accessible to the vast majority of patients. Thus, the development of a more generally applicable methodology based on technology familiar to clinical laboratories and pathologists, such as immunohistochemistry, is desirable. To overcome these limitations, using an automated, quantitative, immunofluorescence-based technique (accurate quantitative analysis (AQUA)), protein expression (rather than mRNA) was assessed and developed into a more applicable assay system for further development (87).

However, several issues should be addressed before a general recommendation for implementation of a biomarker-based therapeutic approach can be given. First, it is crucial to corroborate the prognostic and predictive impact of these biomarkers in independent data sets under controlled conditions. Second, it is important to test the general feasibility of a tailored treatment approach in a multicenter trial and to verify promising phase II outcome data in a large community-based trial. Thus, randomized and multisite clinical trials are necessary to produce more convincing data suggesting the utility of such an approach.

Acknowledgments—We thank Bin Fang, Wei Guan, Victoria Izumi, Umut Oguz, C. Eric Thomas, the Haura laboratory, Tricia Holtje, and Amy Koomen for contributions. Proteomics at Moffitt is supported by the United States Army Medical Research and Material Command under Award DAMD17-02-2-0051 for a National Functional Genomics Center, the NCI, National Institutes of Health under Award P30-CA076292 as a Cancer Center Support Grant, and the Moffitt Foundation.

* This work was supported, in whole or in part, by National Institutes of Health Grants R01-CA106414 (to R. S.), R01-CA123174 (to E. B. H.), and R01-CA102726 and U01-CA101222 (to G. B.) from the NCI. This work was also supported by the National Functional Genomics Center (to E. B. H. and J. M. K.) and American Cancer Society Grant CRTG-00-196-01-CCE (to R. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
† To whom correspondence should be addressed: SRB3 Molecular Oncology and Proteomics, H. Lee Moffitt Cancer Center and Research Inst., University of South Florida, 12902 Magnolia Dr., Tampa, FL 33612. Tel.: 813-745-8524; Fax: 813-745-3829; E-mail: john.koomen@moffitt.org.

REFERENCES


60. R80


