

REVIEW

Open Access

# Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development

Maria Frantzi<sup>1,2\*</sup>, Akshay Bhat<sup>1,3</sup> and Agnieszka Latosinska<sup>2,3</sup>

## Abstract

Biomarker research is continuously expanding in the field of clinical proteomics. A combination of different proteomic-based methodologies can be applied depending on the specific clinical context of use. Moreover, current advancements in proteomic analytical platforms are leading to an expansion of biomarker candidates that can be identified. Specifically, mass spectrometric techniques could provide highly valuable tools for biomarker research. Ideally, these advances could provide with biomarkers that are clinically applicable for disease diagnosis and/ or prognosis. Unfortunately, in general the biomarker candidates fail to be implemented in clinical decision making. To improve on this current situation, a well-defined study design has to be established driven by a clear clinical need, while several checkpoints between the different phases of discovery, verification and validation have to be passed in order to increase the probability of establishing valid biomarkers. In this review, we summarize the technical proteomic platforms that are available along the different stages in the biomarker discovery pipeline, exemplified by clinical applications in the field of bladder cancer biomarker research.

**Keywords:** Clinical proteomics; Biomarkers; Verification; Validation; Mass spectrometry

## Introduction

Diseases with high complexity such as cancer are associated with increased incidence rates worldwide. Recent data reveal that approximately 7.6 million deaths caused by cancer occurred in 2008, with this number corresponding to 13% of all deaths [1]. Based on these numbers, there is substantial room for improvement in the current strategies for development of biomarkers capable of being introduced into clinical practice. According to the National Cancer Institute (<http://www.cancer.gov/>), a biomarker is defined as “*a molecule detected in body fluids or tissues that are associated with a special process (normal or abnormal), a condition or disease*”. Depending on the intended use, biomarkers can be distinguished on the following categories; diagnostic biomarkers which incorporate disease detection, prognostic that represent prediction of the course of a particular disease

(e.g. recurrence, progression and survival) and predictive that would allow for prediction of the response to treatment which could be subsequently applied in patient assessment [2-5]. In reference to malignant diseases, diagnosis at late stages generally results in poor clinical outcome [6,7]. The intended use of a cancer biomarker would hence be early stage diagnosis and/or prognosis. Thus, biomarkers that would enable early disease diagnosis are required, together with those that would provide prognostic values in disease status and predict an outcome of an illness prior to any treatment designed. Novel prognostic biomarkers may also help clinicians select an optimal therapeutic strategy for individuals, facilitating determination of the response to a specific treatment type. Successful introduction of biomarkers into routine clinical practice becomes the current motive in this research area and is expected to be beneficial to the patients and in health care systems. Discovering biomarkers is a multi-parameter process [8-10] and applying them to routine practice needs a proper consideration of multiple issues [11]. Additionally, since clinical needs differ among various diseases, biomarker development

\* Correspondence: [frantzi@mosaiques-diagnostics.com](mailto:frantzi@mosaiques-diagnostics.com)

<sup>1</sup>Mosaiques Diagnostics GmbH, Mellendorfer Strasse 7-9, D-30625 Hannover, Germany

<sup>2</sup>Biotechnology Division, Biomedical Research Foundation Academy of Athens, Soranou Ephessiou 4, 115 27 Athens, Greece

Full list of author information is available at the end of the article

including discovery, verification and validation cannot be restricted to a single methodology. In this article, we summarize the several challenges related to the biomarker research, as well as the available analytical platforms in the field of proteomics.

## Review

### Pipeline for biomarker development

For more than 40 years translational research in academia and industry has attempted to introduce novel biomarkers with clinical utility to improve the management of diseases, especially those with high social and economic burden such as cancer. Although the analytical platforms have expanded widely, especially in the case of Mass Spectrometry (MS), the vast majority of the published biomarker candidates are not introduced in the clinical practice, due to several issues. The lack of a gold standard can be a drawback since no test with an excellent performance is available to compare with the potential biomarkers. A risk of a biased assessment is thus present. Liu et al. recently addressed this issue by proposing a mathematical formula for covariate adjustment [12]. Several other causative factors have been reported to clarify this discrepancy either related to the disease background or the workflow of biomarker development. Along these lines, several levels of variability can be introduced starting with the disease heterogeneity [13,14], which occurs in complex diseases. Most importantly, the lack of valid biomarkers is often a result of an inadequate set up of the discovery and validation stages [15]. For this reason, a very careful design of the biomarker development is required from the discovery phase to subsequent verification and validation stages. At the same time, concurrent knowledge of the clinical background is needed and a clear target driven from the main clinical needs for the study [16,17].

### Clinical conceptions on the biomarkers study design

A good example to describe relevant issues on variations observed within the same disease is Urothelial Carcinoma. Bladder cancer (BCa) is a highly heterogeneous malignancy characterized by distinct clinical characteristics and molecular pathways [18]. Two independent molecular mechanisms that specifically trigger different phenotypes of urothelial carcinoma have been reported [18]. Ras/MAP kinase signaling activity is specifically involved in superficial phenotypes of the disease. Alterations in tumor suppressor activity of p53 and Rb and overexpression of EGFR and ErbB2, MMP-2 and MMP-9 are characteristic for muscle invasive phenotypes (MIBC) [19-21]. Additional challenges regarding bladder cancer could be that high risk non-muscle invasive BCa (NMIBC) tumour lesions are associated with poor outcome [22,23] and as for the tumor classification the

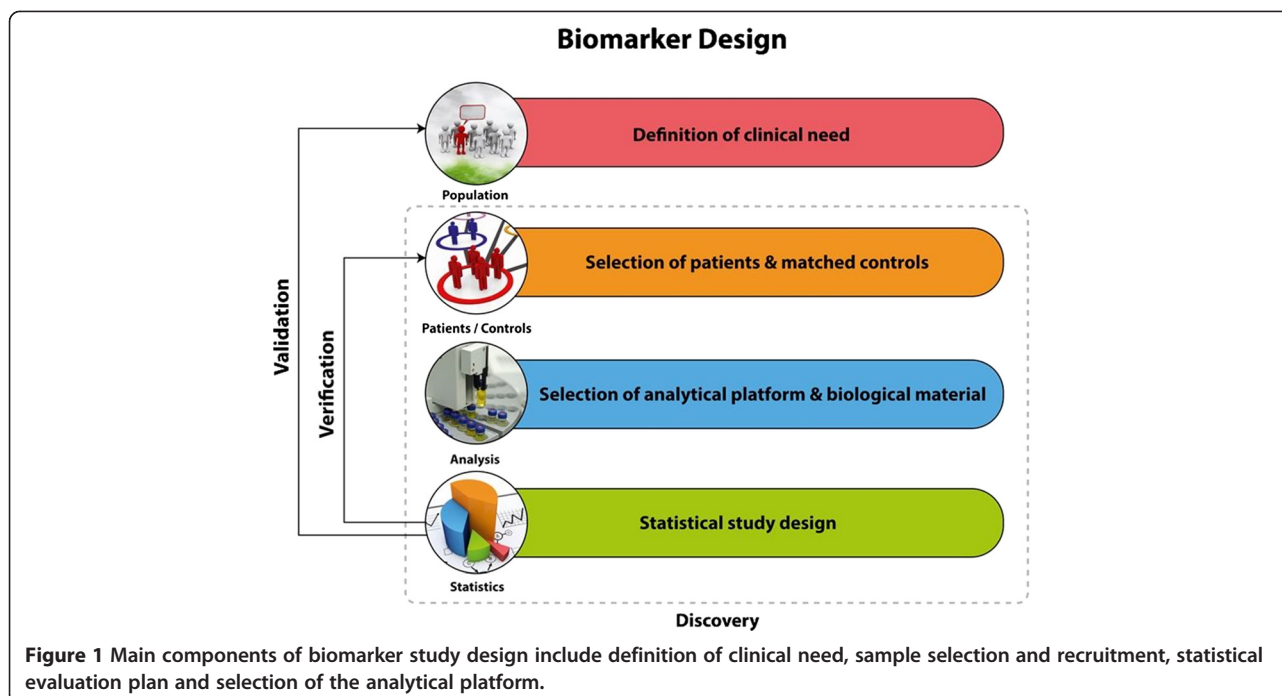
histological variations that can be observed within the same patient as the tumor progresses (intra-patient variation) [24]. These facts should be taken into account in the study design and in terms of clinical objective. The critical parts which have to be considered during the study design are represented in the Figure 1. A biomarker a priori can be applicable only for a specific context of use, for which its performance has been assessed. In the case of bladder cancer, a subgroup or a panel of cancer biomarkers with diagnostic potential that could successfully detect early stage events would be beneficial [25]. A further aim could be the investigation of prognostic indicators for treatment response. The introduction of guidelines in the different biomarker stages has been described as an approach of quality assessment of the biomarkers and standards for designing and reporting biomarker studies have been proposed [26]. For epidemiological studies certain requirements have been suggested in the context of STROBE-ME project [27]. For study design and requirements of predictive biomarkers PROBE standards have been proposed [28], while prognostic biomarkers should be in accordance to REMARK requirements [29].

### Sample biobanking

To develop a research finding into a clinical tool with diagnostic or prognostic value, a large number of biological samples and/or tissue specimens is required. Prerequisites include not only biological material resources, but also a very well-organized preservation domain to be retained, so called biobank. Ideally, a biobank should retain maximum quality of the biological material stored (following standardized protocols of sample handling), of associated clinical and demographical data, and it should be easily accessible and open to the scientific community [30]. Figure 2 depicts a rough outline of biobanking process [31]. An important issue is assigning a unique ID given to a sample [32], an appropriate database structure and management system. Such systems have been currently developed [33-35] mainly as laboratory informatics management applications (LIMS) that are built to tract samples from the initial steps of delivery.

### Sample collection and treatment

General considerations regarding the variability that can be related to the biological material that is under investigation as well as the sample treatment are also present [36,37]. Two articles published by Lescuyer et al. [38] and Good et al. [39] respectively, address the challenges related to the selection of the biological material for biomarker proteomics applications. In the article by Lescuyer et al. [38], advantages and limitations of the different biological fluids that can be applied in proteomics strategies in search for biomarkers are reported. The biological

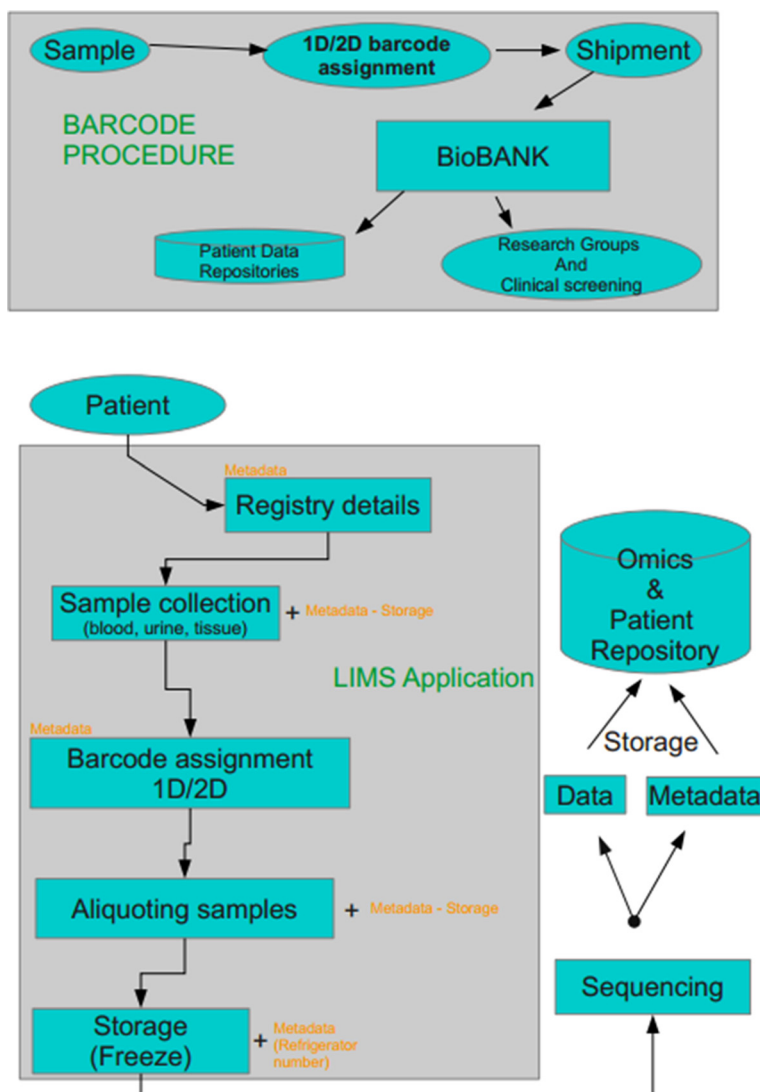


fluids could be separated into two distinct categories, based on their proximity to the organ of interest. For instance, biological fluids such as CSF, presenting with increased proximity with brain offer the opportunity for defining disease specific biomarkers, although the collection could be in a rather invasive manner. In contrast, in the second category belong body fluids that can be readily available in large amounts, such as urine. The amount of biological material, as well as the non-invasive way of collection, is a major advantage for biomarker studies. The drawback of the analysis of this type of biological fluids could be their heterogeneous content, especially due to the presence of several interfering compounds that are excreted [38]. In the article by Lescuyer et al. the authors also give guidance on selection of samples: especially the inclusion of a reference group that contain both healthy individuals, but also patients with closely related diseases or patients presenting similar symptoms. Further, the performance of the potential biomarkers should not be influenced by pre-analytical factors. To clarify this issue in every type of investigation factors such as storage conditions and proteolysis, for example, should be taken into account and reported [38]. In the follow up article by Good et al. [39], emphasis is put on the selection of the suitable biological fluid or 'sample source' as the very first step of the study design. The authors suggest the proper evaluation and reporting of the variability, introduced among others by sampling treatment prior to any analysis [39]. Variability can be limited by the application of approved standard protocols. In bladder cancer biomarker research, urine is a preferred source of biomarkers.

Protocols for urine collection and sample processing have been developed by European Kidney and Urine Proteomics (EuroKUP) and Human Kidney and Urine Proteome Project (HKUPP) (<http://www.eurokup.org/node/138>) [11].

#### Study design and evaluation of the analytical performance

Performance parameters of the analytical platform need to be well described and a quality control process must be in place. Recommendations for increased consistency through the application of standardized protocols have been already introduced by McGuire et al. [40] and Fiedler et al. [41]. Guidance may also be gained from a recent manuscript of assessing CE-MS platform performance [42] of the performance of a biomarker is frequently defined by its sensitivity and specificity. Sensitivity in this context is defined as the percentage of the true positive results and specificity to the percentage of true negative results. Skates et al. [43] in their recent study noted the importance of statistical design in biomarker studies. The aim of the above study was the establishment of the method for estimation of the sample size at the initial stage of biomarker development workflow to increase the probability that the selected putative biomarkers will pass the large scale validation in targeted population. The study was focused on the identification of ovarian biomarkers via proteomics approaches including Shotgun analysis of cyst fluids and MRM assay in plasma for discovery and verification, respectively. Therefore, a statistical model was constructed based on the multiple parameters such as distribution of proteins in individual sample, between biological or technical replicates. According



**Figure 2** Representative workflow of the typical procedure to be followed regarding the sample biobanking. This multistep process includes sample tracking by electronic system, as well as integration of patients clinical characteristics and demographic data. Finally, the deposition of acquired data in public repositories is presented.

to their model an initial cohort of 50 cases and 50 controls could successfully yield a good candidate and if independent verification is applied in a 5 times bigger cohort (250 cases/250 controls), then the chance of a biomarker to pass into clinical validation phase could be >90%. Along the same lines, Shariat et al. [44] reported the need of a well-designed statistical evaluation in the context of defining valid bladder cancer biomarkers, while Behrens and colleagues proposed the validation of bladder cancer biomarkers particularly in prospective studies that meet epidemiological criteria [45]. Collectively, a careful design of the study according to the main clinical needs is required. The initial study is probably best performed in a well phenotyped, predefined cohort and sample handling and analysis

should be performed according to strict guidelines. Subsequent mandatory confirmation of the results is best achieved in a prospective multicenter study in the population at risk [46]. Following the above suggestions will substantially increase the possibility that a candidate biomarker be successfully introduced into clinical practice [16,45].

### Technical considerations regarding the analytical set up for biomarker development

The complexity of the biological fluid may generate a need for a combination of different techniques such as fractionation approaches. In general, the basic requirements of the methodology that is selected are: simplicity of use, robustness, high accuracy and performance [47].

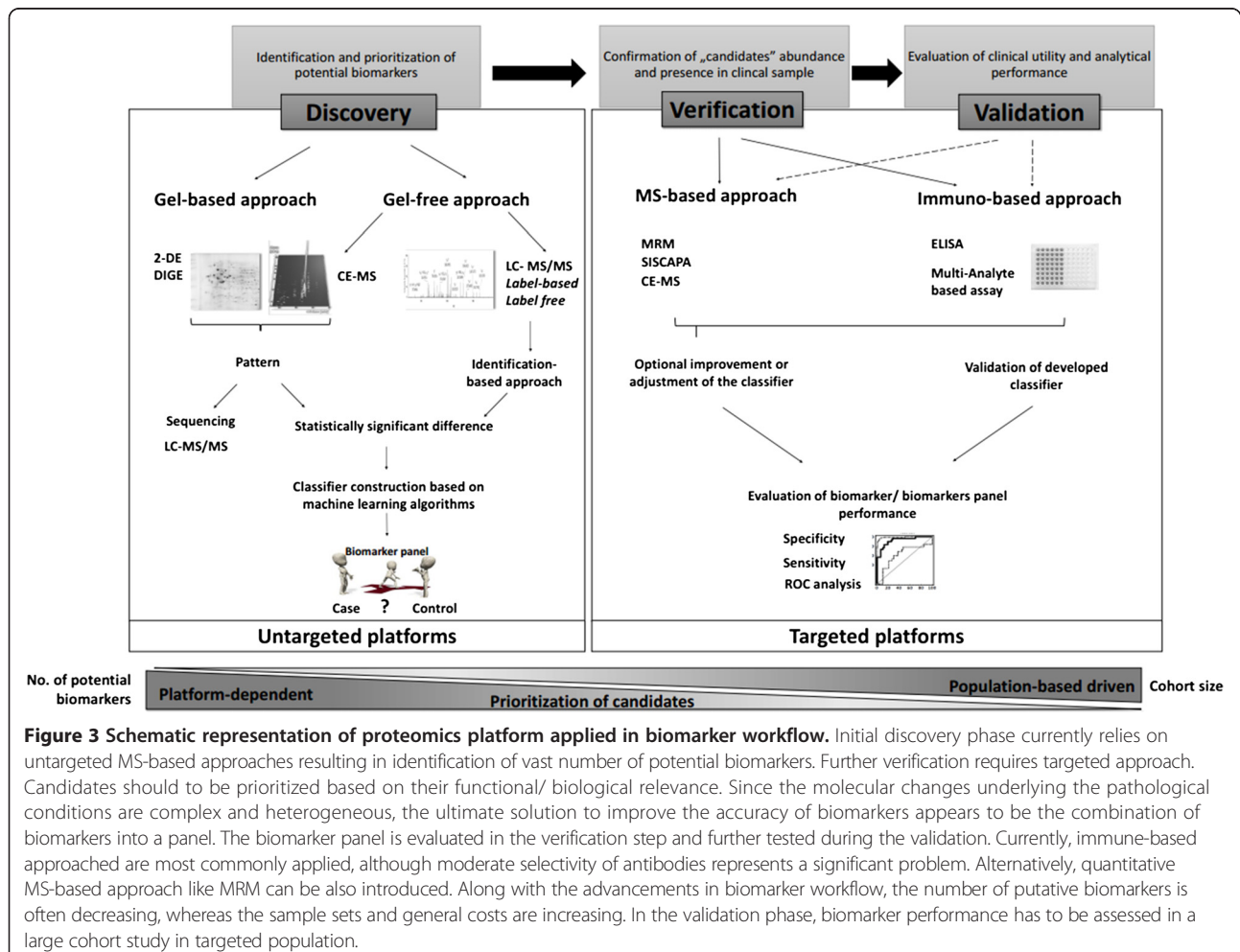
The biomarker workflow can be divided into 3 main parts: discovery, verification and validation. Depending on the specific aim of the study, various proteomic platforms can be applied from the unbiased discovery setting to the targeted quantification in the verification and validation stages. The basic characteristics of the objectives and platforms that can be employed at the different stages are summarized in the Figure 3. Below, a technical description of the available proteomic technologies for discovery, verification and validation stages is presented, together with certain recent applications particularly in the field of clinical proteomics in the quest of bladder cancer biomarker research.

### A) Discovery of biomarkers

Identification of biomarker candidates is the first step towards clinical implementation [48]. At the discovery phase, two major approaches can be distinguished: a knowledge-based approach wherein selection of biomarker candidates is based on the existing molecular mechanisms underlying the disease initiation or progression, or alternatively an

unbiased approach that involves untargeted identification of differentially expressed proteins between two analyzed groups [48]. Currently, MS-based proteomics techniques favor untargeted approaches in biomarker discovery that result in a substantial increase of novel biomarker candidates [49,50]. However, due to the limited number of analyzed samples, a high false discovery rate is observed and the reported differential expression frequently are a result of inter- or intra- individual variability. For example, 10% of urinary proteome is unique for a particular individual [51]. Small sample size at discovery phase leads to overestimation of accuracy of biomarker performance (i.e. sensitivity and specificity) and brings the reliability of findings into question [52]. Consequently, the confirmation of detected differences and further investigation of the impact of related diseases on a biomarker's performance is mandatory.

In general, the classical proteomics workflow includes protein separation using gel-based or gel free techniques followed by the identification by mass spectrometry. Issues that have to be taken into account include establishment of well characterized procedures and platforms



along with assessment of analytical variability in addition to the resolution being appropriate for the complexity of the analyzed material [52]. The most common specimen sources for biomarker discovery are body fluids (e.g. plasma, urine and cerebrospinal fluids) that are characterized by high complexity and wide dynamic range in protein concentration. In the case of plasma, the dynamic range is exceeding 10 orders of magnitude [53]. Despite recent advances in the MS-based techniques, the typical MS-based analysis covers only the dynamic range of three to four orders of magnitude [54]. Since the potential biomarker candidates often belong to the low abundance proteins their detection is hampered by the presence of the highly abundant proteins. This is especially evident in plasma where more than 95% of the total protein content comes from only 5 abundant proteins. In order to “dig the deep proteome” and increase coverage, several methods can be applied. One of the recommended strategies is reducing the sample complexity. This approach can be applied at several stages of the experimental workflow. Samples can be enriched for selected subcellular fractions or for specific type of proteins (based on post-translational modifications). Particularly, membrane ([55,56]), secreted [57,58], nuclear matrix [59] and phosphorylated proteins [60] and glycoproteins [61,62] are reported as a valuable source for biomarker identification. Moreover, in the case of plasma, depletion of highly abundant protein like albumin and IgG is widely applied. Several methods have been developed in order to improve the detection of proteins in the low concentration range. Comparison and/or evaluation of different depletion methods was a subject of several studies [63-66]. As an example Tu et al. [66] compared the protein content from plasma samples depleted using immunoaffinity chromatography with unfractionated samples. Applied methods (Multiple Affinity Removal System, Agilent Technologies, Inc.) enables removal of 7 or 14 most abundant plasma proteins. IEF-LC-MS/MS analysis revealed enrichment for 23 low abundant proteins in depleted fraction, which covered around 5-6% of total protein identifications [66]. As exemplified, even application of the depletion approach followed by shotgun analysis did not allow the identification of the majority of low abundance proteins. Additionally, some issues concerning the depletion approach include: a) reproducibility, b) co-depletion of proteins of interest, c) requirement of higher amount of starting material or d) removal of only selected highly abundant proteins. Alternative, extensive fractionation can be performed. Recently, Atanassov et al. [67] described the combination of three separation strategies namely 1D-PAGE, pIEF (peptide isoelectric focusing) and RP-HPLC as an effective methods for reaching deeper proteome. Authors analyzed the nuclear extract from HeLa cells using the

single separation methods (1D-PAGE and pIEF) as well as combined approach for the nuclear extract from HeLa cells. Increasing number of peptide and protein identifications were observed for the established three dimensional workflow (56228 peptides, 5260 proteins) in comparison to each individual approach (1D-PAGE-LC-MS/MS: 38557 peptides, 3540 proteins; pIEF-LC-MS/MS: 31113 peptides, 3945 proteins). Zhang et al. utilized a mix-bed ion-exchange chromatography (mixture of the strong cation and anion-exchange material) for protein separation prior to MudPIT analysis (multidimensional-protein identification technology) [68]. The workflow was tested using cell lysate from mammary tumor 4 T1. Using the classical MudPIT approach, 1292 proteins were identified, whereas application of 3D workflow enabled identification of 3084 proteins and covered 86% of identified proteins using classical MudPIT. Therefore, application of additional fractionation strategy can improve the proteome coverage. On the other hand, Zubarev et al. recently showed that the in-depth proteomic analysis can be performed using only one dimensional LC-MS/MS [69]. However, optimization is necessary, including sample preparation, chromatographic separation as well as MS analysis. The authors report that a 4 h procedure using a 50 cm column and MS/MS analysis enabled the identification of 37554 peptides corresponding to 4825 proteins (1% FDR at peptide and protein level). This covers around 50% of analyzed human malignant melanoma cell line (A375). Apparently, state-of-the-art MS-based platforms combined with optimized analytical workflow enable “deep” proteome coverage. Moreover, in comparison to the multidimensional procedures, the lower initial amount of starting material, shorter analytical time and cost-effectiveness are advantageous. In this part of review, we will summarize the current untargeted proteomics platforms for biomarker discovery along with the recent and representative examples of their application in bladder cancer biomarker discovery.

#### ***Gel-based proteomics platforms for biomarker discovery***

Two dimensional gel electrophoresis (2-DE) followed by MS serves as a classical approach in analysis of differentially expressed proteins [70-72]. In this method, two separation steps are conducted namely isoelectric focusing (IEF) and SDS-PAGE. First, proteins are separated according to their charge in an immobilized pH gradient (IPG) and subsequently based on their molecular mass in a polyacrylamide matrix. Afterwards, protein spots are visualized and the signal intensity is used for (semi)quantitative analysis [73,74]. 2-DE enables separation of up to 10000 proteins [75] along with detection of protein isoforms [76,77]. Particularly, analysis of post-translational modifications including phosphorylation and glycosylation is of paramount importance, since their alteration is frequently

related with pathological states. However, moderate reproducibility and limited detection for hydrophobic proteins (such as membrane), low abundance proteins, proteins above or below the pore size of the gel as well as proteins beyond the pH range of the IPG strips are shortcomings of 2-DE [78].

Difference gel electrophoresis (DIGE) is the recent advancement in traditional 2-DE [79]. In this approach three fluorescent cyanine dyes (Cy2, Cy3, and Cy5) of identical charge, similar molecular mass and different fluorescent properties, are used to label the proteins before separation [79,80]. The three samples: control, case and internal standard (combination of equal amounts of tested samples) are pooled after labeling and separated simultaneously. This reduces the number of gels as well as gel-to-gel variability, which is one of the major drawbacks of the classical approach [79,80]. Also, improved matching and more accurate quantification are achieved due to the presence of internal standard on all gels [80]. DIGE is more expensive and requires additional equipment like fluorescent scanner, but also offers higher sensitivity (0.5 fmol proteins) [80].

#### **Gel-free proteomics**

**Shotgun proteomics** To bypass the limitations of the classical gel-based approach, efforts have been focused on the development of gel-free strategies. Shotgun or “bottom-up” proteomics is based on the analysis of native or protease derived peptides followed by sequencing with tandem mass spectrometry (MS/MS). Keeping in mind that the complexity of the sample is high, to improve the proteome coverage extensive/multidimensional fractionations has to be conducted prior to MS/MS analysis. Currently, different fractionation strategies are applied including mostly chromatography (ion exchange, reverse phase etc.) [68,69], IEF [66] or combination of these techniques [67]. This issue was described in the introductory part in this section.

In general, the large-scale proteomic analysis by MS/MS includes the following parts: 1) peptide ionization, 2) separation of precursor ions based on the mass-to-charge ratio, 3) fragmentation, 4) analysis of daughter ions, and 5) data analysis including identification and quantification [81]. Two major quantification approaches can be employed namely: label-based (metabolic or chemical labeling) and label-free (spectral counting and intensity-based analysis) approaches. Both approaches along with their advantages and limitations were extensively reviewed by DeSouza et al. [82]. Briefly, peptides from particular samples are labeled with different tags, mixed and analyzed by MS. In metabolic labeling (e.g. SILAC) heavy isotope amino acids are incorporated during protein synthesis, while in chemical labeling (e.g. Isobaric tag for relative and absolute quantification (iTRAQ),

Isotope-coded affinity tags (iCAT) are applied after tryptic digestion. The mass shift introduced by the label is used to distinguish components derived from the different samples. The analytical variability on quantification is reported to significantly reduced in this experimental procedure in comparison the label-free approach where samples are measured separately, and differences in sample preparation as well as run-to-run variability can affect the results. Labeling is more expensive, since it requires isotope specific labeling reagents, and may also result in the introduction of artefacts, as well as a reduced coverage and dynamic range, due to the pooling.

Even though tandem mass spectrometry can result in high number of protein identifications, the false positive and false negative identifications are an inherent problem in shotgun experiments. Since shotgun proteomics is a peptide-based approach, the false protein identifications can occur as a result of incorrect assignment of fragmented ion spectra to peptide sequence as well as further inferring of protein identifications. Therefore, the false discovery rate is evaluated on peptide and protein level. Strategies to evaluate the false discovery rate for mass spectrometry results include searching against concatenated (chimeric database, composed from targeted and decoy database) [83] or decoy database [84], application of statistical models [85,86], or utilization of scoring system [87]. Methods for evaluation of assignment of peptides to protein sequences have been also described [88,89]. Of note: in general estimates of FDR are too optimistic [90], and the true FDR may be up to 10-fold higher.

**CE-MS** Another interesting strategy that can be used in the biomarker discovery phase relies on analysis of the low molecular weight proteome, also defined as “peptidome”. For the identification of biomarkers on the peptidomics level, capillary electrophoresis coupled to mass spectrometry (CE-MS) has been widely applied [91-93]. Analytical performance of CE-MS was described recently and several issues including precision, stability, limits of detection, reproducibility and intra-variability were addressed [42]. The developments and applications of this platform in clinical proteomics were reviewed recently [94,95]. Briefly, small proteins and peptides are separated through an electric field according to their charge and size. Currently, separation in an uncoated bare fused silica capillary column at low pH is the best practical solution [95]. Different capillary coatings have been proposed to improve the analytical performance (robustness, resolution, reproducibility), however due to coating instability none of these techniques has been routinely used [96]. After electrophoretic separation, analytes are ionized (mostly by electrospray ionization (ESI) followed by MS analysis. Two coupling approaches are generally employed:

sheathflow and sheathless interface. In general, all large-cohort studies are performed using the sheathflow approach, while sheathless interfaces currently are of limited robustness and reproducibility (as presented in the table from [95]). Detected peptides are characterized by the CE migration time, signal intensity and molecular mass. For identification, different platforms were tested e.g. CE-MALDI-TOF-TOF (Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry), CE-ESI-QTOF (Quadrupole time-of-flight mass spectrometer) [97] and CE-Orbitrap [98]. The migration time in CE is dependent only on mass and the number of basic, neutral and polar amino acids [97]. This fact can be used to assign the peptide sequences obtained from LC-MS/MS (Liquid Chromatography) analysis to peptide mass that is detected by CE-MS [97].

#### **Statistical data mining for proteomic biomarker discovery**

Current proteomic approaches enable assessment of thousands of biomarker candidates. As aforementioned, at this stage a high false positive rate is observed [10]. Therefore, robust statistical analysis that allows the determination of “true” as well as promising candidates for further verification is of paramount importance [99]. Dakna et al. examined different statistical tests to discover valid biomarkers from proteomic data [100]. The non-parametric Wilcoxon test was found best suited for analysis of proteomic data: after validation in an independent sample set the highest percentage of valid biomarkers were detected. Additionally, the authors showed that adjustment for multiple testing is mandatory to derive potential biomarkers that can be verified in an independent test set.

Obviously, selection of candidates should not be based only on statistical significance, but also on the ability to fulfill a specific clinical need [43,100]. However, poor statistical design at the early experimental stage results in lack of statistical power to select relevant candidates, due to insufficient number of samples analyzed [43,100].

#### **B) Verification of biomarkers**

Verification is mandatory to evaluate the findings obtained in the discovery phase. Although in discovery phase unbiased or untargeted approaches are applied to define a candidate or a set of candidate biomarkers, the verification phase represents a targeted approach for further evaluation of biomarkers sensitivity, specificity and predictive capabilities. More importantly, the verification is performed on the appropriate biological specimen that may be further used in clinical practice. In general, verification studies require larger patient populations than the discovery phase. Depending on the specific study design the analytical platform could be the same as in the discovery phase (CE-MS based classification) or vary

with preferable assays to be either mass spectrometry based [101,102] or protein binding assays [103].

#### **Protein binding assays**

Protein binding assays include both the traditional immunoaffinity based ELISA (Enzyme-linked Immunosorbent assay) or other multiplex assays and Protein microarrays [104]. The most widely used technique for protein quantification is ELISA. The advantages of the method are speed, sensitivity and specificity, and compatibility with standard clinical laboratory equipment, so that it can be applied in clinical routine. The selectivity depends on the antibody that is applied and furthermore it has to be evaluated in the specimen of interest. The additional limitation of ELISA is that it cannot provide a simultaneous quantitative analysis of multiple potential biomarkers. In order to obtain quantitative data via parallel analyses for multiple antigens, Multiplex immunoassays have been developed. Protein microarrays are designed to print specific antibodies or antigens onto a support surface, generally a slide or membrane. A single sample is hybridized to the array. The captured antigens or antibodies are subsequently detected [105]. Assay platforms such as MULTI-ARRAY (Meso Scale Discovery), Bio-Plex (Bio-Rad Laboratories), have been applied for Cytokine detection, while regarding the renal injury a panel of 7 Biomarkers based on Antibody assays has been proposed as biomarkers with improved potential to assess renal function [106]. A major risk in the multiplex arrays is the increased cross-reactivity, due to the presence of multiple antibodies, which are normally applied as a mixture. To improve the assay specificity, Juncker and his collaborators have developed a number of innovative platforms with improved performance [107-109]. As a solution to avoid the mixing of the reagents, this group attempted the application of glass slides in a device called “snap chip”. The antibodies are immobilized in a multiple arrangement on a glass slide, where the sample is also applied [107]. The above group also introduced the use of gel captured antibodies in alginate droplets to increase the sensitivity of the detection. In this study in 2011, the assay was evaluated using 6 proteins, 3 already reported cancer biomarkers, as well as 3 cytokines (CEA, HER2, ENG and TNF- $\alpha$ , IL-8, MIP/CCL4 respectively) [108]. A new methodology based on the co-localization of the primary captured antibodies and the secondary detection antibodies is now proposed, namely ACM or Antibody Colocalization Microarray. When compared with the classical singleplex ELISA and conventional multiplex sandwich assays, ACM was proven to decrease the level of cross-reactivity. However, this technique could be more complex as it requires precision in the alignment [109].

In general, immunoassays are widely used in U.S. Food and Drug Administration (FDA) approved devices for



cancer biomarkers, possibly also as a result of bias of the regulatory agencies towards conventional, but well characterized technologies. Many applications are reported, as reviewed by Fuzery et al. [8].

#### **Mass spectrometric quantitative approaches**

Apart from antibody-based technologies alternative methods for quantitative analysis and validation of potential biomarkers are quantitative MS methodologies, including the application of scanning techniques, such as multiple reaction monitoring (MRM) and other stable isotope labeling-based approaches such as SISCAPA (Stable Isotope Standards and Capture by Anti-Peptide Antibodies). The combination of high throughput capabilities of Mass Spectrometry, together with increased specificity and sensitivity that can be compared to immunoassays in some cases are the main advantages that make MS based applications very popular for quantitative validation studies [110]. MRM has the great advantage that an antibody is not required, but still awaits application in a clinical setting [111].

**Multiple reaction monitoring** MRM is the extended version of Selected Reaction Monitoring (SRM) [112]. The analysis is focused only on biomolecules of specific masses, while all others are excluded. Higher specificity is achieved by the isolation of a specific precursor ion, collision-induced fragmentation and the subsequent detection of the specific product ion after fragmentation. Triple quadrupole instruments are typically employed for this approach [111,113,114]. A recent application in biomarker characterization in tissue specimens from patients with colorectal carcinoma was introduced (Hyperplex MRM). In this study a combination of a strategy for relative quantification such as iTRAQ was conducted with an mTRAQ approach for absolute quantification. This resulted in increased robustness of the MRM approach since 4 different samples were labeled simultaneously and in increased validity of the quantification since relative and total quantities of the biomarkers could be achieved [115]. Another variation of MRM methodology is a peptide immunoaffinity enrichment technique coupled with stable isotope dilution mass spectrometry, called SISCAPA [116]. In this technique, one or more selected tryptic peptides with unique sequences representing the target protein, the “proteotypic” peptides, are enriched using anti-peptide antibodies bound to Protein G. A stable isotope dilution (SID) method is applied as an internal standard by the use of a defined quantity of spiked stable isotope- labeled peptide of the same sequence in a pre-defined quantity. The relative quantification of the peptides is indicative of the protein concentration in the sample. In this assay, the sensitivity and specificity of antibody binding is combined to

the versatility of MS, providing several advantages compared to the conventional immunoassays. Moreover, it provides the capability of analysis of multiple analytes in a single assay by combining antibodies in the enrichment step [117-119].

**Pre-treatment strategies** In order to decrease the limit of detection, different pre- treatment strategies can be combined, like enrichment of the peptides of interest, sample pre-fractionation and depletion of the high abundance peptides. For targeted peptide enrichment, specific anti-peptide antibodies can capture the peptides of interest in the way it described above in the SISCAPA approach [120].

Many studies have reported the value of sample pre-fractionation and/or depletion. Kuhn et al. [121] first applied this strategy to characterize C-reactive protein in serum of patients with rheumatoid arthritis upon depletion of the 3 most abundant proteins in serum: albumin, immunoglobulin G, and haptoglobin. Yang et al. conducted two-dimensional solid-phase extraction as fractionation step prior to quantification of somatropin in plasma samples [122]. Keshishian et al. reported a 1000-fold improvement of limit of detection (LOD) upon depletion of seven high abundant plasma proteins by strong cation exchange chromatography [123]. A range of the values for the limit of quantification (LOQ) was between 1 and 10 ng/ml and coefficient of variation (CV) of 3-15% was estimated [123]. Employing the SISCAPA methodology, Kuhn et al. enriched for troponin-I and interleukin-33 in plasma samples to characterize these proteins as cardiovascular biomarkers [124]. To assess the inter-laboratory performance of immunoaffinity enrichment coupled to MRM- MS, Kuhn et al. designed an inter-laboratory study based on the quantification of 8 predefined peptides from S100A7, S100A8, S100A12, and IL1RN proteins [125]. Coefficient of variation was calculated for replicates analyzed by the same system (intra-laboratory) and across different laboratories (inter-laboratory) [125]. Overall inter-laboratory CV was estimated below 25 at the LOQ level. Inter-laboratory CV for immuno-MRM particularly, was calculated to be 14%, while intra-laboratory CV for immuno-MRM was 7%, respectively [125].

#### **Data mining & statistical analysis**

Depending on the quantitative approach that has been followed, different methodologies can be followed for peak integration, data analysis and downstream statistical evaluation [126,127]. For relative quantification or differential expression purposes, data normalization has to be performed prior to every type of comparison. A widely used approach especially when using label-free proteomics is the one described by Jantos-Siwiy et al.

[128], where endogenous stable and abundant peptides are used as internal standards. Another approach is SRMstats which can be applied to adjust for the median of the logarithmic values of the intensities obtained by the heavy isotope labeled peptides [129]. Variation caused by the analytical process or sample treatment may be corrected with the same methodology proposed by Johnson et al. for microarrays studies [130], where parametric and non-parametric formulas are applied, taking into account the mean intensity and variance in each sample [130].

### C) Bioinformatics platforms in clinical proteomics

Knowledge of biological mechanisms is helpful in the interpretation of proteomic results [131]. The application of computational techniques in analysing information associated with biomolecules on a large-scale platform has now been firmly established as a discipline in molecular biology encompassing a wide subject area from structural biology, genomics to gene-expression studies. Biological data at the omics level from transcriptomics to proteomics and metabolomic profiles are being produced at a very high rate [132,133]. For such a surge in data, computing science has become indispensable to biological research especially in handling large quantities of data and probing the complex dynamics observed in nature [134]. The main aims of bioinformatics include:

- 1) The organization of the data in a way to allow researchers access existing information and to submit new entries. Some of which include GEO [135], ArrayExpress [136] and Human Proteinpedia (<http://www.humanproteinpedia.org/>),
- 2) The development of new tools and resources for data integration and analysis, for which expertise in computational theory as well as a thorough understanding of biology is required. Such examples are interaction databases like IntAct, BioGrid and databases related to diseases like OMIM, Oncomine and metabolomics databases like HMDB.
- 3) The application of these tools in data analysis and interpretation of the results based on a biological meaningful manner, for instance web based tools like String and Cytoscape for visualisation or AmiGO, KEGG, DAVID on the pathway level.

In particular for proteomics datasets, gene ontology and pathway annotations, as well as patient information should not only contain high confident data but should also be in a well-structured architecture to provide genuine data retrieval, coverage, and utility [137]. Some of the reliable protein/peptide and biological pathway resources used for proteomics profile processing in research and academic firms are described in Tables 1 and 2. After

**Table 1 List of reliable protein and peptide databases**

Databases for protein/Peptide data repository	
Protein/Peptide Database	Website/Link
UniProt/Swiss Prot	<a href="http://www.uniprot.org/">http://www.uniprot.org/</a>
Proteomics Identifications Database	<a href="http://www.ebi.ac.uk/pride/">http://www.ebi.ac.uk/pride/</a>
MEROPS	<a href="http://merops.sanger.ac.uk/">http://merops.sanger.ac.uk/</a>
PepBank	<a href="http://pepbank.mgh.harvard.edu/">http://pepbank.mgh.harvard.edu/</a>
PeptideAtlas	<a href="http://www.peptideatlas.org/">http://www.peptideatlas.org/</a>
ProteinProspector	<a href="http://prospector.ucsf.edu/prospector/mshome.htm">http://prospector.ucsf.edu/prospector/mshome.htm</a>
MassMatrix	<a href="http://www.massmatrix.net/mm-cgi/home.py">http://www.massmatrix.net/mm-cgi/home.py</a>

relying on annotated data sets from different databases (Table 2), the next steps are computational approaches in a systematic manner to analyse such integrated data. Computational approaches also provide means for inferring *in silico* and analysing changes in interactions and network dynamics [138]. Some of the computational tools for integrating proteomics datasets on a pathway level are:

- 1) Pathway analysis: KEGG [139], Ingenuity Pathway Analysis ([www.ingenuity.com](http://www.ingenuity.com)) MetaCore (<http://host.genego.com/metacore.php>)
- 2) Pathway mapping: Reactome [140], PathVisio [141], BioCyc plugin [142]
- 3) Gene Ontology analysis: ClueGO [143], BiNGO [144], FuncAssociate [145]
- 4) Network analysis: GeneMania [146], DisGeNet [147], EnrichmentMap [148], NetAtlas [149], NetworkAnalyzer (<http://med.bioinf.mpi-inf.mpg.de/netanalyzer/index.php>) [150], KUPNetViz [151]
- 5) Interactome mapping: iRefScape [152], MiMI [153], PanGIA (<http://prosecco.ucsd.edu/PanGIA/>), BioNetBuilder [154], Bisogenet [155], FunNetViz (<http://www.funnet.ws/>)
- 6) Metabolomics analysis: IDEOM [156], MAVEN, MetaCore, Beilstein, mzMatch [157]

### Applications of systems biology – disease diagnosis and treatment

Network based approaches to human diseases appear to have enormous potential in biological and clinical applications. To better understand the effects of cellular mechanisms on disease progression, identifying proteins and pathways that are related to disease may offer better targets for drug development. These advances may also lead to the selection of better and more accurate biomarkers that are associated with diseases and help with disease classification. Current systems-based approaches focus on identifying pathways that may be used to subtype a disease and develop treatments for

**Table 2 List of highly cited pathway databases for proteomic applications**

**Most cited repositories for biological pathways**

Pathway databases	Biological pathway	Website/Link
Reactome KnowledgeBase	Signal Transduction Pathway	<a href="http://www.reactome.org">http://www.reactome.org</a>
BioCarta Pathway Diagrams	Signal Transduction Pathway	<a href="http://www.biocarta.com/genes/index.asp">http://www.biocarta.com/genes/index.asp</a>
Pathway Commons	Signal Transduction Pathway	<a href="http://www.pathwaycommons.org/pc/">http://www.pathwaycommons.org/pc/</a>
Protein ANalysis THrough Evolutionary Relationships	Signal Transduction Pathway	<a href="http://www.pantherdb.org">http://www.pantherdb.org</a>
Protein Lounge	Signal Transduction Pathway	<a href="http://www.proteinlounge.com">http://www.proteinlounge.com</a>
WikiPathways	Signal Transduction Pathway	<a href="http://wikipathways.org/">http://wikipathways.org/</a>
Transcription Factor encyclopedia	Regulatory Pathways	<a href="http://www.cisreg.ca/cgi-bin/tfe/home.pl">http://www.cisreg.ca/cgi-bin/tfe/home.pl</a>
Transcription Regulatory Regions Database	Regulatory Pathways	<a href="http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/">http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/</a>
A Public Database of Transcription Factor and Regulatory Sequence Annotation	Regulatory Pathways	<a href="http://www.pazar.info/">http://www.pazar.info/</a>
Homo Sapiens Comprehensive Model Collection (HOCOMOCO)	Regulatory Pathways	<a href="http://autosome.ru/HOCOMOCO/index.php">http://autosome.ru/HOCOMOCO/index.php</a>
Transcription Factor Database	Regulatory Pathways	<a href="http://www.gene-regulation.com/index2.html">http://www.gene-regulation.com/index2.html</a>
Human Protein Reference Database	Protein-Protein Interactions	<a href="http://www.hprd.org/">http://www.hprd.org/</a>
Human Annotated and Predicted Protein Interaction Database	Protein-Protein Interactions	<a href="http://bioinformatics.iupui.edu/HAPPI/">http://bioinformatics.iupui.edu/HAPPI/</a>
Biomolecular Interaction Network Database	Protein-Protein Interactions	<a href="http://bond.unleashedinformatics.com/">http://bond.unleashedinformatics.com/</a>
Molecular Interaction Database	Protein-Protein Interactions	<a href="http://mint.bio.uniroma2.it/mint/">http://mint.bio.uniroma2.it/mint/</a>
Biological General Repository for Interaction Datasets	Protein-Protein Interactions	<a href="http://thebiogrid.org/">http://thebiogrid.org/</a>
Search Tool for the Retrieval of Interacting Genes/Proteins	Protein-Protein Interactions	<a href="http://string.embl.de/">http://string.embl.de/</a>

individual disease groups. Network modules have been used to predict patient survival, metastasis, invasion, drug response etc. [158-164]. For this purpose, a well characterised group of samples is required related to a disease subtype/stage, for example cancer metastasis to search among specific networks or so called sub-networks for potential biomarkers that enable disease classification [165]. Additionally, systems analysis may provide with insights in the molecular mechanisms underlying the diseases. This may be highly valuable in drug development by indicating correlation between the response to a drug and the responders' molecular background. An example of such an approach is the study by Chu and Chen, where a protein-interaction network was applied to investigate drug targets related to apoptosis [166].

**D) Validation of biomarker candidates**

The pivotal objective of the validation phase is to evaluate the clinical utility of the biomarker candidates [9]. Validation has to be performed in an independent, sufficiently large sample set also reflecting the heterogeneity of targeted population. This is mandatory also since the diagnostic accuracy is often generally overestimated in the model established in training set (groups of individuals used for discovery of biomarkers and development of the model) [52]. To demonstrate the clinical utility, validation studies have to be driven by the specific context of use and targeted population, since depending on

the clinical needs the biomarker has to fulfill different requirements regarding clinical performance (i.e. sensitivity and specificity). The accuracy of individual biomarker or biomarkers panel performance can be assessed by the ROC (receiver operating characteristics) analysis [167]. ROC curve represents a plot of true-positive rate (sensitivity, percentage of cancer patients who tested positive for biomarkers) versus false positive rate (FPR, percentage of healthy subject classified as having disease). Whereas specificity is defined as 1-FPR. In this method the area under the curve (AUC) is used as an indicator of the biomarker performance regarding the ability to distinguish between control and patients affected by disease. It is of paramount importance to take into account the false positives and false negatives in order to establish an optimal classification threshold at desired specificity and sensitivity level. Biomarkers utilized for screening should reveal high sensitivity and, frequently even much more important, a low level of false positives. On the other hand, specific diagnostic tests require high positive predictive values (PPV, percentage of diseased patient among all positive test result). Due to the fact that sensitivity and specificity do not provide the information about probability of disease occurrence, disease predictive values have to be assessed i.e. (PPV, PPN). However, these values are dependent on disease prevalence and can only be assessed in prospective studies [8]. Collectively, regardless of the clinical use, consequences from false positive and false negative cases have to be always considered as a benefit-to-harm ratio.

Ultimately, application of novel biomarker/biomarkers panel has to improve the outcome. A striking example of an unfortunate development is the prostate specific antigen (PSA) test. The applicability of this test for prostate cancer screening arose controversy in medicine [168]. It has been claimed that decrease of mortality for prostate cancer is the major benefit of PSA-based screening. However, Andriole et al. reported that the screening does not reduce the number of deaths for prostate cancer after 13-years follow up [169]. Moreover, the false-positive results (and they represent the majority in this case) have harmful consequences including invasive biopsy as well as following complication, overdiagnosis and overtreatment of disease [168].

Varied challenges are encountered at the validation stage including e.g. 1) samples quality and availability, 2) funding and 3) requirements of regulatory agencies [170]. Due to these facts, validation is a bottleneck in the biomarker development process [170]. In addition, the scientific reward in validation is moderate: validation studies are generally rejected in high impact journals. To also ensure robustness, validation should be performed in a multi-center study [52], and selected cohorts have to represent the population targeted with the biomarker. Bearing in mind that most promising candidates have to be tested in hundreds or even thousands of samples, validation requires quantitative, robust, (multiplex) and high throughput methodology. Not all of the platform applicable at previous stages can fulfil these requirements. Therefore, changing the platform can be necessary: biomarkers discovered using gel-based approach cannot be further validated by the same techniques, mainly due to its limited throughput. On the other hand, CE-MS serve as a good example of a technique, which can be applied for all biomarker development phases [91-93]. High reproducibility, high throughput and cost-effectiveness are reported [42]. These characteristic along with the need for developing biomarker panels, makes CE-MS an attractive platform for biomarker development.

Currently, antibody-based approaches are considered as the gold standard in clinical application, mostly since this is a technology well known to everybody. However, application of immunoassays is often hampered by the lack of high-quality antibodies. Additionally, the ELISA assay enables detection of a single antigen, a drawback for the validation of biomarker panels. Moreover, the low-dynamic range and high cost of development of ELISA based assays indicate a moderate utility of this approach in large scale validation studies. To bypass some of the limitations multi-analyte immunoassays have been applied [171] including planar array [172] and micro-bead assays [173,174]. In planar array, different antibodies are spotted on a flat surface, whereas in a

second type, antibodies are immobilized to varied microbeads [172]. Recently, Fu et al. [175] compared the analytical performance of five currently used multiplex immunoassays in the context of their application for validation of biomarkers (particularly cytokines) in serum. The MULTI-ARRAY (planar assay) and Bio-Plex (magnetic beads) are characterized by better performance than other tested multiplex assays i.e. A2 (Beckman Coulter), FAST Quant (Whatman Schleicher & Schuell BioScience), and Flow-Cytomix (Bender MedSystems), but this is also dependent on analyzed biospecimens (serum or purified cytokines) [175]. Application of multiplex assays enables high-throughput quantitative analysis and uses less sample volume. On the other hand, development of novel multiplex immunoassays is a challenging task. Currently, antibodies are commonly used as capture ligands, but aptamer ligands (oligonucleotides) may offer an alternative novel approach [176,177]. A promising alternative are MS-based (typically MRM) approaches, as described in details in previous section. After establishing clinical utility (which equals significant improvement over the current state of the art) [16], further assessment of analytical performance is required. In this case, the following issues have to be addressed: detection and quantification limit, precision, stability of analyte, specificity, interfering compounds etc. This topic was covered recently by Fuzery et al. [8] and it is beyond of the scope of this paper to review this issue in detail.

### **Application of proteomics approaches in BCa biomarker discovery**

Proteomic approaches have been applied at all stages of biomarker discovery workflow. This includes untargeted platforms for biomarker identification (gel-based and gel-free) [70,71,178-182] and targeted platforms for further verification and validation of biomarker candidates (MRM, multi-analyte assays) [106,108,115,183,184]. Additionally, to improve proteome coverage and identify low abundance protein, enrichment strategies have been also applied e.g. immobilized metal affinity (IMAC) [179,182] dual-lectin chromatography [185], or peptidomics approaches [92,93]. Bladder cancer is the second in incidence and mortality malignancy of the genitourinary system. At initial diagnosis, the majority of patients (75%) exhibit non-muscle invasive cancer (pTa, pT1, pTis), whereas the rest belongs to muscle invasive disease (pT2, pT3, pT4) [22]. The invasive phenotype results in significant decrease of the survival rate [22]. Additionally, high recurrence rate and cancer progression impose the requirement for lifelong monitoring of patients after treatment. Up to date, the gold standard for clinical diagnosis includes invasive cystoscopy and non-invasive voided urine cytology with limited sensitivity for detection of low grade tumors [186]. Although

some tests have been approved by FDA (e.g. NMP-22, BTA-TRAK, uCyt+), they seem to have no clinical utility [187-189]. Therefore, there is an urgent clinical need for application of novel non-invasive tests for early detection, patient monitoring and stratification. A vast number of potential biomarkers have been discovered using proteomics as well as genomics approaches. The detailed description of currently available bladder cancer biomarker candidates is beyond the scope of the manuscript and this topic was recently reviewed [190,191]. To give an overview on the current status of BCa proteomic biomarkers, the representative examples along with study design and potential clinical utility are described below and summarized in Table 3.

#### **Gel-based proteomics**

Chung et al. detected by 2-DE elevated levels of cofilin in BCa tissues vs. control urothelium. In total, 24 samples were analyzed resulting in identification of 12 differentially expressed proteins. The up-regulation of cofilin in BCa tissue specimens was confirmed by Western Blot and immunohistochemistry. Additionally, an antibody specific for phosphorylated Ser-3 of cofilin revealed elevated phosphorylation in BCa samples, especially in muscle-invasive BCa. In parallel, *in vitro* studies showed decreased EGF-induced migration in cofilin knock-down T24 cells. Collectively, both expression and phosphorylation of cofilin may be involved in BCa aggressiveness [178]. In another study, the urinary proteome was investigated in order to detect biomarkers for aggressive BCa. Zoidakis et al. analyzed urine samples from NMIBC and MIBC patients using enrichment by IMAC [182]. Collected enriched fractions were separated using 1D-SDS PAGE followed by *in-gel* digestion. Protein identification was performed using LC-MS/MS analysis. The study revealed aminopeptidase N, profilin-1 and myeloblastin as potential biomarker candidates. Further confirmation by Western blot or ELISA was conducted for selected proteins. Aminopeptidase N found to be down-regulated in MIBC, whereas Profilin-1 and myeloblastin were up-regulated in invasive cancer [182]. Orenes-Pinero has applied differential gel electrophoresis to investigate the urinary proteome of BCa patients ( $n = 7$ , positive cytology) and controls ( $n = 7$ , negative cytology) [181]. Differentially expressed proteins were identified by peptide mass fingerprinting using MALDI-TOF MS, including Regenerative protein (Reg-1), cytokeratins 1, 2 and 10, T-cell surface protein CD5 and prefoldin. Among these, only cytokeratin 1 was down-regulated in urine from BCa patients. Western blot analysis of eight bladder cancer cell line models (from non-invasive to metastatic) indicated the correlation between the levels of the proteins identified by proteomics and cancer progression. Quantitative analysis of urinary Reg-1 was evaluated by ELISA

( $n = 80$ ) and used for evaluation of diagnostic accuracy. Sensitivity and specificity at the level of 0.0038 ng/mL were 81.3% and 81.2%, respectively.

#### **Gel-free proteomics**

The shotgun approach followed by labeling or label-free quantification has also been widely implemented in biomarker discovery research. Frantzi et al. described urinary histone H2B and Zinc-finger 335 (NIF-1) as a potential progression marker for BCa. Urine from benign ( $n = 5$ ), non-invasive ( $n = 10$ ) and invasive cases ( $n = 5$ ) was enriched by IMAC and native peptides were analyzed by LC-MS/MS. In total, 1845 peptides were detected (638 precursor proteins). Differential regulation of histone H2B and NIF-1 were verified further by ELISA (urine,  $n = 166$ ) and immunohistochemistry (tissue samples,  $n = 32$ ) [179]. Apart from the label-free approach, labeling techniques have been also employed to discover biomarkers for BCa. Kato et al. used iTRAQ labeling to compare the proteome from bladder carcinoma urothelium ( $n = 6$ ) with paired normal tissues ( $n = 6$ ) [180]. 493 proteins were identified including 15 up-regulated proteins in cancer cases in comparison to adjacent normal samples (e.g. DDX39, B-cell receptor-associated protein 31, chaperonin containing TCP1, FK506 binding protein 4, S100 calcium binding protein A1). Immunohistochemistry ( $n = 303$ ) was used to verify the findings for protein which have not been previously evaluated. However, actin-related protein 3 homolog B was not verified, since the antibodies were not commercially available. This example indicates one of the drawbacks for application immune-based assays for protein verification: lack of specific antibodies. Authors found decreased expression of DDX39 with higher cancer stage and grade. In addition, low expression level of DDX39 significantly correlates with disease progression. Further functional analysis using siRNA assay was performed in bladder cancer cell line (T24). As a result, an increased invasion ability of cells transfected with si-DDX39 compared to control was observed. Therefore, reduced expression of this protein may serve as a biomarker to predict disease progression [180].

The search of potential biomarker candidates can be performed also at the peptide level. Briefly, peptidomic profiling was used for detection of urothelial carcinoma [93] as well as for prediction of MIBC [92]. In all these cases, biomarker panels were developed after analysis of a training set and further validated in independent test sets. CE-MS analysis was conducted for the discovery and initial validation phase. In a first study, Theodorescu et al. developed a 22 polypeptides panel for diagnosis of urothelial carcinoma [93]. The limited specificity (73%) was obtained in the test set (varied genitourinary disorders), whereas the sensitivity remained high (100%). Moreover, authors also advocated the application of

**Table 3 Representative examples of BCa biomarker candidates identified by proteomic approaches**

<i>Biomarker identification; Biomarker candidate/panels</i>	<i>Verification/Validation</i>	<i>Regulation</i>	<i>Potential clinical value; Biomarker performance</i>	<i>Ref.</i>
<b>Gel-based approaches</b>				
<b>2DE, n=24</b> Tissue: 6 normal urothelium, 9 NMIBC, 9 MIBC	<b>Western Blot</b> <b>Immunohistochemistry, n=24</b> <i>For both experiments, the same material was used as in the discovery phase.</i>	↑ in both NMIBC and MIBC;  ↑ phosphorylation level of cofilin in BCa tissue samples (most prominent in MIBC).	Predict cancer progression  Lack of evaluation of biomarker performance.	[178]
<b>Cofilin</b> <b>IMAC, 1-SDS-PAGE, n=35</b> Urine: Two pools from NMIBC, n <sub>1</sub> =9, n <sub>2</sub> =7 Two pools from MIBC, n <sub>3</sub> = 9, n <sub>4</sub> =10	<b>Western Blot</b> Aminopeptidase N, n=108 Myeloblastin, n=97	Aminopeptidase N  ↑ in MIBC	Biomarker for cancer aggressiveness  Lack of evaluation of biomarker performance.	[182]
<b>Aminopeptidase N, Myeloblastin, Profilin-1</b>	<b>ELISA</b> Profilin-1, n=82	Myeloblastin, Profilin 1  ↓ in MIBC		
<b>DIGE, n=14</b> Urine: 7 BCa (positive cytology), 7 controls (negative cytology)	<b>Western Blot</b> , 8 BCa cell line models <b>Tissue microarray, n=292</b> Primary urothelial cell carcinoma		Diagnosis, staging, outcome prognosis  <b>Detection of BCa:</b>	
<b>Reg- 1</b>	<b>ELISA, n=80</b> Urine: 32 BCa (positive cytology), 48 Controls (negative cytology)	↑ in BCa cases, association with stage	81.3% sensitivity 81.2% specificity	[181]
<b>Gel-free approaches</b>				
<b>LC-MS/MS, n=20</b> Urine: Benign (n=5), pTa, pT1 (n=10), pT2+ (n=5)	<b>ELISA, n = 166</b> Urine,  For H2B: n=147, For NIF-1: n = 158 <i>In both groups urine from benign, NIMB (Ta, T<sub>a</sub>) and MIBC (T<sub>2</sub>+) were included.</i>	↑ level of H2B with cancer stage in urine and tissue samples  ↓ level of NIF-1 with cancer stages (not agreement with urinary level)	Prediction of disease progression, discrimination of tumor stages  Lack of evaluation of biomarker performance.	[179]
<b>histone H2B, NIF-1</b>	<b>Immunohistochemistry, n=32</b> pTa, pT1, n=23, pT2+ n=9			

**Table 3 Representative examples of BCa biomarker candidates identified by proteomic approaches (Continued)**

<p><b>iTRAQ, n=12</b>            Tissue:            6 bladder cancer tissues (4 NMIBC, 2 MIBC)            and paired normal tissues;</p>	<p><b>Immunohistochemistry, n=303</b></p>	<p>↑ in 4/6 BCa samples in comparison to control (iTRAQ);             Inverse correlation to stage and histological grade progression (immunohistochemistry)</p>	<p>Prediction of disease progression             Lack of evaluation of biomarker performance.</p>	<p>[180]</p>
<p><b>DDX39</b>  <b>CE-MS, n=248</b>            Urine:            127 BCa patients, 121 Controls</p>	<p><b>CE-MS, n=130</b>            Urine,  <i>Test set: 68 NMIBC and 62 MIBC</i></p>	<p>↓ regulated in MIBC in comparison to NMIBC</p>	<p><b>Prediction of MIBC:</b>            81% sensitivity            57% specificity</p>	<p>[92]</p>
<p><b>4 polypeptide panel</b>  <b>CE-MS, n=79</b>            Urine:            46 BCa patients, 33 Controls  <b>22 polypeptides panel</b></p>	<p><b>CE-MS, n=366</b>            Urine,  <i>(Test set includes healthy controls, patients with non-malignant and malignant urological disorders)</i></p>	<p>Varied; 10 peptides ↑ in BCa;             12 ↓ in BCa in comparison to control</p>	<p><b>Detection of BCa:</b>            100 % sensitivity            73% specificity</p>	<p>[93]</p>

additional discriminatory panels (e.g. non-malignant disease vs. urothelial carcinoma) can increase the specificity level [93]. In a second study, a 4 polypeptides panel (fragments of membrane associated progesterone receptor component I, uromodulin, collagen  $\alpha$ -1 (I), Collagen  $\alpha$ -1 (III)) was reported by Schiffer et al. and enabled detection of MIBC with sensitivity of 81%. However, limited specificity was obtained (57%) [92]. Along the same avenue of research, reanalysis of existing and newly collected BCa peptidomics data is currently under investigation in our lab (n = 608 samples, 304 controls and 304 BCa patients). Several of previously detected peptides were confirmed as well as novel potential biomarkers were reported [95]. These data are currently rigorously verified in large, prospectively collected cohorts in the EU-funded BCMolMed ([www.bcmolmed.org](http://www.bcmolmed.org)) project.

To summarize, the number of detected biomarker candidates depends on utilized analytical platform at biomarker discovery stage e.g. 2-DE enables identification of over a dozen candidates, whereas this values can increased up to hundreds for MS-based approaches. In most cases immune-based assays were used to verify presence of selected protein. Only for the peptidomics markers, CE-MS was applied for all phases. However, in most cases only initial verification of detected biomarkers was conducted, but appropriate vigorous validation in a sufficiently large population is still outstanding. Therefore, in order to establish robust and accurate biomarker/panel of biomarkers, further validation has to be performed, driven by the clearly defined context of use and cohorts representing the targeted population. Also, apart from the peptidomics studies, only single proteins were used to evaluate the diagnostic accuracy, whereas a combination of biomarkers candidates may lead to substantial improvement of biomarker performance.

## Conclusions

Cancer is considered as a disease with high heterogeneity, increased incidence and mortality rates with a serious social and economic burden. The benefits from application of robust and accurate biomarkers in cancer management might result in significant improvement of clinical outcome via detection of cancer at early stages. An improvement in the therapeutic strategies based on the prognosis of the treatment response is also anticipated [192]. Multiple advances have been achieved regarding the proteomic technology that can be applied in major parts of the biomarker development: identification, verification and validation. Along the way of recent achievements in untargeted MS-based proteomic approaches, as well as targeted quantification proteomic strategies, the number of potential proteomic biomarkers has rapidly increased, as also indicated by the biomarkers candidates related to BCa summarized in this article

[193-196]. However, this apparent progress has not triggered successful implementation of novel biomarkers into clinical practice. Therefore, as pinpointed above, critical issues related to biomarker development should be taken into account to raise the awareness about difficulties encountered in the process. Shortcomings hampering the biomarker implementation include difficulties related to the definition of context of use, proper study design (selection of patients, statistical design), samples availability along with poor clinical characteristics, high sampling variability due to the lack of application of standardized protocol as well as the application of inappropriate statistics. Clinical implementation of biomarkers is complex and requires collaborative efforts between researchers from different fields and clinicians. In order to facilitate the translation into clinical utility and benefit for patients, various guidelines have been established to guide scientists in this endeavor [11,197].

Based on the literature published, it appears that numerous proteomic biomarkers do exist that will likely result in a substantial improvement of the current clinical situation [193,194,198].

Regarding Chronic Kidney Disease (CKD), a urinary biomarker model based on a panel of 273 peptides, as established after CE-MS analysis has been already well investigated in the context of early diagnosis of CKD. Good et al. [199] first proposed the above 273 biomarker model, reporting an AUC (area under the curve) value of 0.96 upon independent validation in an independent blinded cohort of 109 CKD samples and 34 urine samples from normal individuals [199]. The same performance (AUC of 0.96) was presented in a follow up study for the same model, using a multicentric validation approach including 137 urine samples (62 CKD patients and 75 normal controls) [200]. Importantly, Zurbig et al. [201] further evaluated the CKD273 peptide marker model for its diagnostic utility in a longitudinal study, where 316 urine samples were employed, including patients with diabetes type 1 and 2 [201]. In this study, the above model was able to predict the progression of normoalbuminuria to macroalbuminuria 5 years before onset, while the AUC value was estimated at 0.93, increased compared to the routinely used urinary albumin whose reported performance is 0.67 [201]. Finally, Andersen et al. [202] applied the CKD273 for characterization of the renoprotective treatment outcome in hypertensive type 2 diabetic patients treated with Ibersartan. In this study, urine samples were collected from patients undergoing Ibersartan treatment in two timepoints before and two years after treatment [202]. The changes in the peptide pattern of the treated patients are indicative of the possible utility of this model -and such proteomic biomarker approaches in general-, in the monitoring of the patients response to drug treatment [202]. The CKD273 classifier



is currently been implemented in a Multicentric European Trial, called PRIORITY where 3280 patients with diabetes type 2 are employed.

Another CE-MS derived peptide biomarker approach with increased potential to be implemented as a routine test for diagnosis of cholangiocarcinoma (CC) has been also described [203]. Due to the demanding clinical need for early detection of cholangiocarcinoma that increases the curative potential of a therapeutic treatment, Lankisch et al. [203] first proposed a peptide classifier as established by interpretation of CE-MS data derived from bile proteomic analysis [203]. Two models were presented enable to distinguish between patients with choledocholithiasis and malignant lesions as well as between cholangiocarcinoma and primary sclerosing cholangitis (PSC), a risk factor for cholangiocarcinoma. After independent validation, the first model was found able to distinguish between patients with gallstones and malignant lesions with sensitivity of 93% and specificity of 86%, while the second model classified PSC cases and CC cases with a sensitivity of 84% and specificity of 78%. Following a similar approach, Metzger et al. [204] introduced a urinary based peptide classifier in a follow up study, where the specific aim was the investigation of a non-invasive urinary test for early diagnosis of cholangiocarcinoma. In this case, an AUC value of 0,87 was reported with 83% sensitivity and 79% specificity, after validation in a cohort of 123 patients [204].

Based on the literature available, it appears that clinically useful proteomic biomarkers can be identified, and also validated, employing the technologies available today. Hence, unravelling this potential benefit would “only” require analysis of a sufficient number of samples using appropriate technologies. Assuming the availability of the required funds, the only limitation is the availability of samples. This major problem has not successfully been tackled by the generation of biobanks, these generally do not proved the support anticipated [197].

However, as recently also suggested Vlahou [17], combining efforts and testing multiple biomarkers in the same samples may be the most promising approach. Another hurdle in bringing the benefits to the patients apparently are the requirements by regulatory agencies, and cost as well as the reluctance of the public health systems to accept novel and beneficial approaches in medicine. Here, initiative from the relevant clinical professional societies and patients groups may be needed in combination with simplification and improvements of regulatory requirements, to enable timely implementation of highly beneficial developments to improve medicine and patient care.

#### Abbreviations

2-DE: Two dimensional gel electrophoresis; ACM: Antibody colocalization microarray; AUC: Area under the curve; BCa: Bladder Cancer;

CC: Cholangiocarcinoma; CE: Capillary Electrophoresis; CKD: Chronic kidney disease; CV: Coefficient of variation; DIGE: Difference gel electrophoresis; ELISA: Enzymed- linked Immunosorbent assay; ESI: Electrospray ionization; FDA: U.S. Food and Drug Administration; iCAT: Isotope- coded affinity tag; IEF: Isoelectric focusing; IMAC: Immobilized metal affinity chromatography; IPG: Immobilized pH gradient; iTRAQ: Isobaric Tag for relative and absolute quantification; LC: Liquid chromatography; LIMS: Laboratory informatics management applications; LOQ: Limit of quantification; MALDI: Matrix-assisted laser/desorption ionization; MIBC: Muscle invasive bladder cancer; MS: Mass spectrometry; MS/MS: Tandem mass spectrometry; NMIBC: Non-muscle invasive bladder cancer; NPV: Negative predictive value; PPV: Positive predictive value; PSC: Primary sclerosing cholangitis; QTOF: Quadrupole time-of-flight mass spectrometer; SILAC: Stable isotope labeling by amino acids in cell culture; SID: Stable isotope dilution; SISCAPA: Stable Isotope Standards and Capture by Anti-Peptide Antibodies; SRM/MRM: Single/Multiple reaction Monitoring.

#### Competing interests

Maria Frantzi and Akshay Bhat are employed by Mosaiques Diagnostics.

#### Authors' contributions

MF, AL and AB performed the literature review and contribute equally to the drafting of the manuscript. All authors read and approved the final manuscript.

#### Authors' information

All the authors namely MF, AB, AL are a PhD students involved in the European BCMolMed program (www.bcmolmed.org). The major goals of the project are identification of novel biomarkers related to bladder cancer invasion, verification of existing biomarkers as well as integration of data using system biology approach.

#### Acknowledgements

The work was supported by grant 317450 PITN-GA-2012-31750 BCMolMed (Molecular Medicine for Bladder Cancer) from the FP7 – PEOPLE – 2012 – ITN program.

#### Author details

<sup>1</sup>Mosaiques Diagnostics GmbH, Mellendorfer Strasse 7-9, D-30625 Hannover, Germany. <sup>2</sup>Biotechnology Division, Biomedical Research Foundation Academy of Athens, Soranou Ephessiou 4, 115 27 Athens, Greece. <sup>3</sup>Charité-Universitätsmedizin Berlin, Berlin, Germany.

Received: 20 November 2013 Accepted: 6 March 2014

Published: 29 March 2014

#### References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM: **Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008.** *Int J Cancer J Int Cancer* 2010, **127**:2893–2917.
2. Kulasingam V, Diamandis EP: **Strategies for discovering novel cancer biomarkers through utilization of emerging technologies.** *Nat Clin Prac Oncol* 2008, **5**:588–599.
3. Frantzi M, Makridakis M, Vlahou A: **Biomarkers for bladder cancer aggressiveness.** *Curr Opin Urol* 2012, **5**:390–396.
4. Fertig EJ, Slebos R, Chung CH: **Application of genomic and proteomic technologies in biomarker discovery.** *Am Soc Clin Oncol Educ Book/ASCO Am Soc Clin Oncol Meet* 2012, **32**:377–382.
5. Gupta S, Venkatesh A, Ray S, Srivastava S: **Challenges and prospects for biomarker research: a current perspective from the developing world.** *Biochim Biophys Acta.* in press.
6. Henley SJ, King JB, German RR, Richardson LC, Plescia M, Centers for Disease C, Prevention: **Surveillance of screening-detected cancers (colon and rectum, breast, and cervix) - United States, 2004–2006.** *Morb Mortal Wkly Rep Surveill Summ* 2010, **59**:1–25.
7. Iuga C, Seicean A, Iancu C, Buiga R, Kumar Sappa P, Volker U, Hammer E: **Proteomic identification of potential prognostic biomarkers in resectable pancreatic ductal adenocarcinoma.** *Proteomics.* in press.
8. Fuzery AK, Levin J, Chan MM, Chan DW: **Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges.** *Clin Proteomics* 2013, **10**:13.

9. Pavlou MP, Diamandis EP, Blasutig IM: **The long journey of cancer biomarkers from the bench to the clinic.** *Clin Chem* 2013, **59**:147–157.
10. Rifai N, Gillette MA, Carr SA: **Protein biomarker discovery and validation: the long and uncertain path to clinical utility.** *Nat Biotechnol* 2006, **24**:971–983.
11. Mischak H, Kolch W, Aivaliotis M, Bouyssié D, Court M, Dihazi H, Dihazi GH, Franke J, Garin J, Gonzalez de Peredo A, Iphöfer A, Jänsch L, Lacroix C, Makridakis M, Masselon C, Metzger J, Monsarrat B, Mrug M, Norling M, Novak J, Pich A, Pitt A, Bongcam-Rudloff E, Sivy J, Suzuki H, Thongboonkerd V, Wang LS, Zoidakis J, Züribig P, Schanstra JP, et al: **Comprehensive human urine standards for comparability and standardization in clinical proteome analysis.** *Proteomics Clin Appl* 2010, **4**:464–478.
12. Liu D, Zhou XH: **Covariate adjustment in estimating the area under ROC curve with partially missing gold standard.** *Biometrics* 2013, **69**:91–100.
13. Ehdaie B, Theodorescu D: **Molecular markers in transitional cell carcinoma of the bladder: new insights into mechanisms and prognosis.** *IJU* 2008, **24**:61–67.
14. Lopez E, Madero L, Lopez-Pascual J, Latterich M: **Clinical proteomics and OMICS clues useful in translational medicine research.** *Proteome Sci* 2012, **10**:35.
15. Diamandis EP: **The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem?** *BMC Med* 2012, **10**:87.
16. Mischak H, Vlahou A, Righetti PG, Calvete JJ: **Putting value in biomarker research and reporting.** *J Proteomics* 2014, **96**:A1–A3.
17. Vlahou A: **Network views for personalized medicine.** *Proteomics Clin Appl* 2013, **7**:384–387.
18. Volanis D, Papadopoulos G, Doumas K, Gkialas I, Delakas D: **Molecular mechanisms in urinary bladder carcinogenesis.** *J BUON* 2011, **16**:589–601.
19. Ewald JA, Downs TM, Cetnar JP, Ricke WA: **Expression microarray meta-analysis identifies genes associated with Ras/MAPK and related pathways in progression of muscle-invasive bladder transition cell carcinoma.** *PLoS One* 2013, **8**:e55414.
20. Mitra AP, Cote RJ: **Molecular pathogenesis and diagnostics of bladder cancer.** *Annu Rev Pathol* 2009, **4**:251–285.
21. Wu XR: **Urothelial tumorigenesis: a tale of divergent pathways.** *Nat Rev Cancer* 2005, **5**:713–725.
22. van den Bosch S, Alfred Witjes J: **Long-term cancer-specific survival in patients with high-risk, non-muscle-invasive bladder cancer and tumour progression: a systematic review.** *Eur Urol* 2011, **60**:493–500.
23. May M, Brookman-Amissah S, Roigas J, Hartmann A, Storkel S, Kristiansen G, Gilfrich C, Borchardt R, Hoschke B, Kaufmann O, Gunia S: **Prognostic accuracy of individual urothelial markers in noninvasive urinary bladder carcinoma: a multicentre study comparing the 1973 and 2004 World Health Organisation classifications.** *Eur Urol* 2010, **57**:850–858.
24. Dyrskjot L, Reinert T, Novoradovsky A, Zuiverloon TC, Beukers W, Zwarthoff E, Malats N, Real FX, Segersten U, Malmstrom PU, Knowles M, Hurst C, Sorge J, Borre M, Orntoft TF: **Analysis of molecular intra-patient variation and delineation of a prognostic 12-gene signature in non-muscle invasive bladder cancer; technology transfer from microarrays to PCR.** *Br J Cancer* 2012, **107**:1392–1398.
25. de Ronde JJ, Rigaiil G, Rottenberg S, Rodenhuis S, Wessels LF: **Identifying subgroup markers in heterogeneous populations.** *Nucleic Acids Res* 2013, **41**:1–10.
26. Tabb DL: **Quality assessment for clinical proteomics.** *Clin Biochem* 2012, **46**:411–420.
27. Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JP, Kirsch-Volders M, Matullo G, Phillips DH, Schoket B, Stromberg U, Vermeulen R, Wild C, Porta M, Vineis P: **Strengthening the reporting of Observational studies in Epidemiology-Molecular Epidemiology (STROBE-ME): an extension of the STROBE statement.** *Eur J Epidemiol* 2013, **26**:797–810.
28. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD: **Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design.** *J Natl Cancer Inst* 2008, **100**:1432–1438.
29. Altman DG, McShane LM, Sauerbrei W, Taube SE: **Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration.** *BMC Med* 2012, **10**:51.
30. Dillner J: **A basis for translational cancer research on aetiology, pathogenesis and prognosis: guideline for standardised and population-based linkages of biobanks to cancer registries.** *Eur J Cancer*. in press.
31. Malm J, Fehniger TE, Danmyr P, Vegvari A, Welinder C, Lindberg H, Appelqvist R, Sjodin K, Wieslander E, Laurell T, Hober S, Berven FS, Fenyö D, Wang X, Andrén PE, Edula G, Carlsohn E, Fuentes M, Nilsson CL, Dahlbäck M, Rezeli M, Erlinge D, Marko-Varga G: **Developments in biobanking workflow standardization providing sample integrity and stability.** *J Proteomics* 2013, **95**:38–45.
32. Marko-Varga G: **BioBanking as the central tool for translational medicine CTM issue 2013.** *Clin Transl Med* 2013, **2**:4.
33. Bruggeman M, Verheyen L, Vidmar T: **A dedicated LIMS for routine gamma-ray spectrometry.** *Appl Radiat Isot.* in press.
34. Morisawa H, Hirota M, Toda T: **Development of an open source laboratory information management system for 2-D gel electrophoresis-based proteomics workflow.** *BMC Bioinform* 2006, **7**:430.
35. Stephan C, Kohl M, Turewicz M, Podwojski K, Meyer HE, Eisenacher M: **Using laboratory information management systems as central part of a proteomics data workflow.** *Proteomics* 2010, **10**:1230–1249.
36. Aguilar-Mahecha A, Kuzyk MA, Domanski D, Borchers CH, Basik M: **The effect of pre-analytical variability on the measurement of MRM-MS-based mid- to high-abundance plasma protein biomarkers and a panel of cytokines.** *PLoS One* 2012, **7**:e38290.
37. Percy AJ, Parker CE, Borchers CH: **Pre-analytical and analytical variability in absolute quantitative MRM-based plasma proteomic studies.** *Bioanalysis* 2013, **5**:2837–2856.
38. Lescuyer P, Hochstrasser D, Rabilloud T: **How shall we use the proteomics toolbox for biomarker discovery?** *J Proteome Res* 2007, **6**:3371–3376.
39. Good DM, Thongboonkerd V, Novak J, Bascands JL, Schanstra JP, Coon JJ, Dominiczak A, Mischak H: **Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future.** *J Proteome Res* 2007, **6**:4549–4555.
40. McGuire JN, Overgaard J, Pociot F: **Mass spectrometry is only one piece of the puzzle in clinical proteomics.** *Brief Funct Genom Proteomic* 2008, **7**:74–83.
41. Fiedler GM, Ceglarek U, Leichte A, Thiery J: **Standardized preprocessing of urine for proteome analysis.** In *The Urinary Proteome*. Edited by Rai AJ. New York: Humana Press; 2010:47–63.
42. Mischak H, Vlahou A, Ioannidis JP: **Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience.** *Clin Biochem* 2013, **46**:432–443.
43. Skates SJ, Gillette MA, Labaer J, Carr SA, Anderson L, Liebler DC, Ransohof D, Rifai N, Kondratovich M, Tezak Z, Mansfield E, Oberg AL, Wright I, Barnes G, Gail M, Mesri M, Kinsinger CR, Rodriguez H, Boja ES: **Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies.** *J Proteome Res* 2013, **12**:5383–5394.
44. Shariat SF, Lotan Y, Vickers A, Karakiewicz PI, Schmitz-Drager BJ, Goebell PJ, Malats N: **Statistical consideration for clinical biomarker research in bladder cancer.** *Urol Oncol* 2010, **28**:389–400.
45. Behrens T, Bonberg N, Casjens S, Pesch B, Bruning T: **A practical guide to epidemiological practice and standards in the identification and validation of diagnostic markers using a bladder cancer example.** *Biochim Biophys Acta* 2013, **1844**:145–155.
46. Riley RD, Sauerbrei W, Altman DG: **Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis, and beyond.** *Br J Cancer* 2009, **100**:1219–1229.
47. Tabb DL, Vega-Montoto L, Rudnick PA, Variyath AM, Ham AJ, Bunk DM, Kilpatrick LE, Billheimer DD, Blackman RK, Cardasis HL, Carr SA, Clauser KR, Jaffe JD, Kowalski KA, Neubert TA, Regnier FE, Schilling B, Tegeler TJ, Wang M, Wang P, Whiteaker JR, Zimmerman LJ, Fisher SJ, Gibson BW, Kinsinger CR, Mesri M, Rodriguez H, Stein SE, Tempst P, Paulovich AG, et al: **Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry.** *J Proteome Res* 2010, **9**:761–776.
48. Frangogiannis NG: **Biomarkers: hopes and challenges in the path from discovery to clinical practice.** *Transl Res* 2012, **159**:197–204.
49. White NM, Masui O, Desouza LV, Krakovska O, Metias S, Romaschin AD, Honey RJ, Stewart R, Pace K, Lee J, Jewett MA, Bjarnason GA, Siu KW, Yousef GM: **Quantitative proteomic analysis reveals potential diagnostic markers and pathways involved in pathogenesis of renal cell carcinoma.** *Oncotarget* 2014, **5**:506–518.
50. Zou X, Feng B, Dong T, Yan G, Tan B, Shen H, Huang A, Zhang X, Zhang M, Yang P, Zheng M, Zhang Y: **Up-regulation of type I collagen during tumorigenesis of colorectal cancer revealed by quantitative proteomic analysis.** *J Proteomics* 2013, **94**:473–485.
51. Molina L, Salvat N, Ameer RB, Peres S, Sommerer N, Jarraya F, Ayadi H, Molina F, Granier C: **Analysis of the variability of human normal urine by**

- 2D-GE reveals a "public" and a "private" proteome. *J Proteomics* 2011, **75**:70–80.
52. Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, Bennett SE, Bischoff R, Bongcam-Rudloff E, Capasso G, Kuczyk MA, Mischak H, Vlahou A, Theodorescu D, Merseburger AS: **Recommendations for biomarker identification and qualification in clinical proteomics.** *Sci Transl Med* 2010, **2**(46):42.
53. Anderson NL, Anderson NG: **The human plasma proteome: history, character, and diagnostic prospects.** *Mol Cell Proteomics* 2002, **1**:845–867.
54. Zubarev RA: **The challenge of the proteome dynamic range and its implications for in-depth proteomics.** *Proteomics* 2013, **13**:723–726.
55. Leth-Larsen R, Lund RR, Ditzel HJ: **Plasma membrane proteomics and its application in clinical cancer biomarker discovery.** *Mol Cell Proteomics* 2010, **9**:1369–1382.
56. Yang Y, Toy W, Choong LY, Hou P, Ashktorab H, Smoot DT, Yeoh KG, Lim YP: **Discovery of SLC3A2 cell membrane protein as a potential gastric cancer biomarker: implications in molecular imaging.** *J Proteome Res* 2012, **11**:5736–5747.
57. de Wit M, Kant H, Piersma SR, Pham TV, Mongera S, van Berkel MP, Boven E, Ponten F, Meijer GA, Jimenez CR, Fijneman RJ: **Colorectal cancer candidate biomarkers identified by tissue secretome proteome profiling.** *J Proteomics* 2014, **99C**:26–39.
58. Makridakis M, Vlahou A: **Secretome proteomics for discovery of cancer biomarkers.** *J Proteomics* 2010, **73**:2291–2305.
59. Albrethsen J, Knol JC, Piersma SR, Pham TV, de Wit M, Mongera S, Carvalho B, Verheul HM, Fijneman RJ, Meijer GA, Jimenez CR: **Subnuclear proteomics in colorectal cancer: identification of proteins enriched in the nuclear matrix fraction and regulation in adenoma to carcinoma progression.** *Mol Cell Proteomics* 2010, **9**:988–1005.
60. Ha GH, Baek KH, Kim HS, Jeong SJ, Kim CM, McKeon F, Lee CW: **p53 activation in response to mitotic spindle damage requires signaling via BubR1-mediated phosphorylation.** *Cancer Res* 2007, **67**:7155–7164.
61. Chen J, Xi J, Tian Y, Bova GS, Zhang H: **Identification, prioritization, and evaluation of glycoproteins for aggressive prostate cancer using quantitative glycoproteomics and antibody-based assays on tissue specimens.** *Proteomics* 2013, **13**:2268–2277.
62. Li QK, Gabrielson E, Zhang H: **Application of glycoproteomics for the discovery of biomarkers in lung cancer.** *Proteomics Clin Appl* 2012, **6**:244–256.
63. Kullolli M, Warren J, Arampatzidou M, Pitteri SJ: **Performance evaluation of affinity ligands for depletion of abundant plasma proteins.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2013, **939**:10–16.
64. Patel BB, Barrero CA, Braverman A, Kim PD, Jones KA, Chen DE, Bowler RP, Merali S, Kelsen SG, Yeung AT: **Assessment of two immunodepletion methods: off-target effects and variations in immunodepletion efficiency may confound plasma proteomics.** *J Proteome Res* 2012, **11**:5947–5958.
65. Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS: **High-abundance protein depletion: comparison of methods for human plasma biomarker discovery.** *Electrophoresis* 2010, **31**:471–482.
66. Tu C, Rudnick PA, Martinez MY, Cheek KL, Stein SE, Slebos RJ, Liebler DC: **Depletion of abundant plasma proteins and limitations of plasma proteomics.** *J Proteome Res* 2010, **9**:4982–4991.
67. Atanassov I, Urlaub H: **Increased proteome coverage by combining PAGE and peptide isoelectric focusing: comparative study of gel-based separation approaches.** *Proteomics* 2013, **13**:2947–2955.
68. Zhang L, Yao L, Zhang Y, Xue T, Dai G, Chen K, Hu X, Xu LX: **Protein pre-fractionation with a mixed-bed ion exchange column in 3D LC-MS/MS proteome analysis.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2012, **905**:96–104.
69. Pirmoradian M, Budamgunta H, Chingkin K, Zhang B, Astorga-Wells J, Zubarev RA: **Rapid and deep human proteome analysis by single-dimension shotgun proteomics.** *Mol Cell Proteomics* 2013, **12**:3330–3338.
70. Chanthammachat P, Promwikom W, Pruegsanusak K, Roytrakul S, Srisomsap C, Chokchaichamnankit D, Svasti J, Boonyaphiphat PKS, Thongsuksai P: **Comparative proteomic analysis of oral squamous cell carcinoma and adjacent non-tumour tissue from Thailand.** *Arch Oral Biol* 2013, **58**:1677–1685.
71. Chen JH, Ni RZ, Xiao MB, Guo JG, Zhou JW: **Comparative proteomic analysis of differentially expressed proteins in human pancreatic cancer tissue.** *Hepatob Pancreat Dis Int* 2009, **8**:193–200.
72. Petrak J, Ivanek R, Toman O, Cmejla R, Cmejlova J, Vyoral D, Zivny J, Vulpe CD: **Deja vu in proteomics: a hit parade of repeatedly identified differentially expressed proteins.** *Proteomics* 2008, **8**:1744–1749.
73. Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, Postel W: **Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications.** *J Biochem Biophys Methods* 1982, **6**:317–339.
74. O'Farrell PH: **High resolution two-dimensional electrophoresis of proteins.** *J Biol Chem* 1975, **250**:4007–4021.
75. Marcus K, Joppich C, May C, Pfeiffer K, Sitek B, Meyer H, Stuehler K: **High-resolution 2DE.** *Methods Mol Biol* 2009, **519**:221–240.
76. Caldini A, Moneti G, Fanelli A, Bruschetini A, Mercurio S, Pieraccini G, Cini E, Ognibene A, Luceri F, Messeri G: **Epoetin alpha, epoetin beta and darbepoetin alfa: two-dimensional gel electrophoresis isoforms characterization and mass spectrometry analysis.** *Proteomics* 2003, **3**:937–941.
77. Zugaro LM, Reid GE, Ji H, Eddes JS, Murphy AC, Burgess AW, Simpson RJ: **Characterization of rat brain stathmin isoforms by two-dimensional gel electrophoresis-matrix assisted laser desorption/ionization and electrospray ionization-ion trap mass spectrometry.** *Electrophoresis* 1998, **19**:867–876.
78. Chevalier F: **Highlights on the capacities of "Gel-based" proteomics.** *Proteome Sci* 2010, **8**:23.
79. Unlu M, Morgan ME, Minden JS: **Difference gel electrophoresis: a single gel method for detecting changes in protein extracts.** *Electrophoresis* 1997, **18**:2071–2077.
80. Minden J: **Comparative proteomics and difference gel electrophoresis.** *BioTechniques* 2007, **43**:739–741. 743 passim.
81. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR III: **Protein analysis by shotgun/bottom-up proteomics.** *Chem Rev* 2013, **113**:2343–2394.
82. DeSouza LV, Siu KW: **Mass spectrometry-based quantification.** *Clin Biochem* 2013, **46**:421–431.
83. Jeong K, Kim S, Bandeira N: **False discovery rates in spectral identification.** *BMC Bioinform* 2012, **13**(Suppl 16):S2.
84. Kall L, Storey JD, MacCoss MJ, Noble WS: **Assigning significance to peptides identified by tandem mass spectrometry using decoy databases.** *J Proteome Res* 2008, **7**:29–34.
85. Keller A, Nesvizhskii AI, Kolker E, Aebersold R: **Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search.** *Anal Chem* 2002, **74**:5383–5392.
86. Ma K, Vitek O, Nesvizhskii AI: **A statistical model-building perspective to identification of MS/MS spectra with PeptideProphet.** *BMC Bioinform* 2012, **13**(Suppl 16):S1.
87. Moore RE, Young MK, Lee TD: **Qscore: an algorithm for evaluating SEQUEST database search results.** *J Am Soc Mass Spectrom* 2002, **13**:378–386.
88. Reiter L, Claassen M, Schrimpf SP, Jovanovic M, Schmidt A, Buhmann JM, Hengartner MO, Aebersold R: **Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry.** *Mol Cell Proteomics* 2009, **8**:2405–2417.
89. Teng B, Huang T, He Z: **Decoy-free protein-level false discovery rate estimation.** *Bioinformatics* 2014, **30**:675–681.
90. Li YF, Radivojac P: **Computational approaches to protein inference in shotgun proteomics.** *BMC Bioinform* 2012, **13**(Suppl 16):S4.
91. Frantzi M, Metzger J, Banks RE, Husi H, Klein J, Dakna M, Mullen W, Cartledge JJ, Schanstra JP, Brand K, Kuczyk MA, Mischak H, Vlahou A, Theodorescu D, Merseburger AS: **Discovery and validation of urinary biomarkers for detection of renal cell carcinoma.** *J Proteomics* 2013, **98C**:44–58.
92. Schiffer E, Vlahou A, Petrolekas A, Stravodimos K, Tauber R, Geschwend JE, Neuhaus J, Stolzenburg JU, Conaway MR, Mischak H, Theodorescu D: **Prediction of muscle-invasive bladder cancer using urinary proteomics.** *Clin Cancer Res* 2009, **15**:4935–4943.
93. Theodorescu D, Wittke S, Ross MM, Walden M, Conaway M, Just I, Mischak H, Frierson HF: **Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis.** *Lancet Oncol* 2006, **7**:230–240.
94. Stalmach A, Albalat A, Mullen W, Mischak H: **Recent advances in capillary electrophoresis coupled to mass spectrometry for clinical proteomic applications.** *Electrophoresis* 2013, **34**:1452–1464.
95. Latosinska A, Frantzi M, Vlahou A, Mischak H: **Clinical applications of capillary electrophoresis coupled to mass spectrometry in biomarker discovery: focus on bladder cancer.** *Proteomics Clin Appl* 2013, **7**:779–793.
96. Zuberovic A, Hanrieder J, Hellman U, Bergquist J, Wetterhall M: **Proteome profiling of human cerebrospinal fluid: exploring the potential of capillary electrophoresis with surface modified capillaries for analysis of complex biological samples.** *Eur J Mass Spectrom* 2008, **14**:249–260.

97. Zurbig P, Renfrow MB, Schiffer E, Novak J, Walden M, Wittke S, Just I, Pelzing M, Neuss C, Theodorescu D, Root KE, Ross MM, Mischak H: **Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation.** *Electrophoresis* 2006, **27**:2111–2125.
98. Mullen W, Albalat A, Gonzalez J, Zerefos P, Siwy J, Franke J, Mischak H: **Performance of different separation methods interfaced in the same MS-reflection TOF detector: a comparison of performance between CE versus HPLC for biomarker analysis.** *Electrophoresis* 2012, **33**:567–574.
99. Bramwell D: **An introduction to statistical process control in research proteomics.** *J Proteomics* 2013, **95**:3–21.
100. Dakna M, Harris K, Kalousis A, Carpentier S, Kolch W, Schanstra JP, Haubitz M, Vlahou A, Mischak H, Girolami M: **Addressing the challenge of defining valid proteomic biomarkers and classifiers.** *BMC Bioinform* 2010, **11**:594.
101. Liebler DC, Zimmerman LJ: **Targeted quantitation of proteins by mass spectrometry.** *Biochemistry* 2013, **52**:3797–3806.
102. Pham TV, Piersma SR, Oudgenoeg G, Jimenez CR: **Label-free mass spectrometry-based proteomics for biomarker discovery and validation.** *Expert Rev Mol Diagn* 2012, **12**:343–359.
103. Stoevesandt O, Taussig MJ: **Affinity proteomics: the role of specific binding reagents in human proteome analysis.** *Expert Rev Proteomics* 2012, **9**:401–414.
104. Brennan DJ, O'Connor DP, Rexhepaj E, Ponten F, Gallagher WM: **Antibody-based proteomics: fast-tracking molecular diagnostics in oncology.** *Nat Rev Cancer* 2010, **10**:605–617.
105. Caiazza RJ Jr, O'Rourke DJ, Barder TJ, Nelson BP, Liu BC: **Native antigen fractionation protein microarrays for biomarker discovery.** *Methods Mol Biol* 2011, **723**:129–148.
106. Ozer JS, Dieterle F, Troth S, Perentes E, Cordier A, Verdes P, Staedtler F, Mahl A, Grenet O, Roth DR, Wahl D, Legay F, Holder D, Erdos Z, Vlasakova K, Jin H, Yu Y, Muniappa N, Forest T, Clouse HK, Reynolds S, Bailey WJ, Thudium DT, Topper MJ, Skopek TR, Sina JF, Glaab WE, Vonderscher J, Maurer G, Chibout SD, et al: **A panel of urinary biomarkers to monitor reversibility of renal injury and a serum marker with improved potential to assess renal function.** *Nat Biotechnol* 2010, **28**:486–494.
107. Li H, Bergeron S, Juncker D: **Microarray-to-microarray transfer of reagents by snapping of two chips for cross-reactivity-free multiplex immunoassays.** *Anal Chem* 2012, **84**:4776–4783.
108. Li H, Leulmi RF, Juncker D: **Hydrogel droplet microarrays with trapped antibody-functionalized beads for multiplexed protein analysis.** *Lab Chip* 2011, **11**:528–534.
109. Pla-Roca M, Leulmi RF, Tourekanova S, Bergeron S, Laforte V, Moreau E, Gosline SJ, Bertos N, Hallett M, Park M, Juncker D: **Antibody colocalization microarray: a scalable technology for multiplex protein analysis in complex samples.** *Mol Cell Proteomics* 2012, **11**:M111. 011460.
110. Shi T, Su D, Liu T, Tang K, Camp DG II, Qian WJ, Smith RD: **Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics.** *Proteomics* 2012, **12**:1074–1092.
111. Yocum AK, Chinnaiyan AM: **Current affairs in quantitative targeted proteomics: multiple reaction monitoring-mass spectrometry.** *Brief Funct Genom Proteomic* 2009, **8**:145–157.
112. Ebhardt HA: **Selected reaction monitoring mass spectrometry: a methodology overview.** *Methods Mol Biol* 2014, **1072**:209–222.
113. Kitteringham NR, Jenkins RE, Lane CS, Elliott VL, Park BK: **Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, **877**:1229–1239.
114. Elschenbroich S, Kislinger T: **Targeted proteomics by selected reaction monitoring mass spectrometry: applications to systems biology and biomarker discovery.** *Mol Biosyst* 2011, **7**:292–303.
115. Yin HR, Zhang L, Xie LQ, Huang LY, Xu Y, Cai SJ, Yang PY, Lu HJ: **Hyperplex-MRM: a hybrid multiple reaction monitoring method using mTRAQ/iTRAQ labeling for multiplex absolute quantification of human colorectal cancer biomarker.** *J Proteome Res* 2013, **12**:3912–3919.
116. Anderson NL, Jackson A, Smith D, Hardie D, Borchers C, Pearson TW: **SISCAPA peptide enrichment on magnetic beads using an in-line bead trap device.** *Mol Cell Proteomics* 2009, **8**:995–1005.
117. Whiteaker JR, Lin C, Kennedy J, Hou L, Trute M, Sokal I, Yan P, Schoenherr RM, Zhao L, Voytovich UJ, Kelly-Spratt KS, Krasnoselsky A, Gafken PR, Hogan JM, Jones LA, Wang P, Amon L, Chodosh LA, Nelson PS, McIntosh MW, Kemp CJ, Paulovich AG: **A targeted proteomics-based pipeline for verification of biomarkers in plasma.** *Nat Biotechnol* 2011, **29**:625–634.
118. Whiteaker JR, Zhao L, Zhang HY, Feng LC, Piening BD, Anderson L, Paulovich AG: **Antibody-based enrichment of peptides on magnetic beads for mass-spectrometry-based quantification of serum biomarkers.** *Anal Biochem* 2007, **362**:44–54.
119. Kim K, Kim SJ, Yu HG, Yu J, Park KS, Jang IJ, Kim Y: **Verification of biomarkers for diabetic retinopathy by multiple reaction monitoring.** *J Proteome Res* 2010, **9**:689–699.
120. Makawita S, Diamandis EP: **The bottleneck in the cancer biomarker pipeline and protein quantification through mass spectrometry-based approaches: current strategies for candidate verification.** *Clin Chem* 2010, **56**:212–222.
121. Kuhn E, Wu J, Karl J, Liao H, Zolg W, Guild B: **Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and 13C-labeled peptide standards.** *Proteomics* 2004, **4**:1175–1186.
122. Yang Z, Hayes M, Fang X, Daley MP, Ettenberg S, Tse FL: **LC-MS/MS approach for quantification of therapeutic proteins in plasma using a protein internal standard and 2D-solid-phase extraction cleanup.** *Anal Chem* 2007, **79**:9294–9301.
123. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA: **Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution.** *Mol Cell Proteomics* 2007, **6**:2212–2229.
124. Kuhn E, Addona T, Keshishian H, Burgess M, Mani DR, Lee RT, Sabatine MS, Gerszten RE, Carr SA: **Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry.** *Clin Chem* 2009, **55**:1108–1117.
125. Kuhn E, Whiteaker JR, Mani DR, Jackson AM, Zhao L, Pope ME, Smith D, Rivera KD, Anderson NL, Skates SJ, Pearson TW, Paulovich AG, Carr SA: **Interlaboratory evaluation of automated, multiplexed peptide immunoaffinity enrichment coupled to multiple reaction monitoring mass spectrometry for quantifying proteins in plasma.** *Mol Cell Proteomics* 2011, **11**:M111. 013854.
126. Navarro P, Trevisan-Herraz M, Bonzon-Kulichenko E, Nunez E, Martinez-Acedo P, Perez-Hernandez D, Jorge I, Mesa R, Calvo E, Carrascal M, Hernandez ML, Garca F, Barcena JA, Ashman K, Abian J, Gil C, Redondo JM, Vazquez J: **General statistical framework for quantitative proteomics by stable isotope labeling.** *J Proteome Res* 2014, **13**:1234–1247.
127. Tucker G, Loh PR, Berger B: **A sampling framework for incorporating quantitative mass spectrometry data in protein interaction analysis.** *BMC Bioinform* 2013, **14**:299.
128. Jantos-Siwy J, Schiffer E, Brand K, Schumann G, Rossing K, Delles C, Mischak H, Metzger J: **Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease.** *J Proteome Res* 2009, **8**:268–281.
129. Chang CY, Picotti P, Huttenhain R, Heinzelmann-Schwarz V, Jovanovic M, Aebersold R, Vitek O: **Protein significance analysis in selected reaction monitoring (SRM) measurements.** *Mol Cell Proteomics* 2012, **11**:M111. 014662.
130. Johnson WE, Li C, Rabinovic A: **Adjusting batch effects in microarray expression data using empirical Bayes methods.** *Biostatistics* 2007, **8**:118–127.
131. Loman N, Watson M: **So you want to be a computational biologist?** *Nat Biotechnol* 2013, **31**:996–998.
132. Mayer P, Mayer B, Mayer G: **Systems biology: building a useful model from multiple markers and profiles.** *Nephrol Dial Transplant* 2012, **27**:3995–4002.
133. Reichhardt T: **It's sink or swim as a tidal wave of data approaches.** *Nature* 1999, **399**:517–520.
134. Luscombe NM, Greenbaum D, Gerstein M: **What is bioinformatics? A proposed definition and overview of the field.** *Methods Inf Med* 2001, **40**:346–358.
135. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Hout A, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A: **NCBI GEO: archive for functional genomics data sets—update.** *Nucleic Acids Res* 2013, **41**:D991–D995.
136. Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, Emam I, Farne A, Hastings E, Ison J, Keays M, Kurbatova N, Malone J, Mani R, Mupo A, Pedro Pereira R, Pilicheva E, Rung J, Sharma A, Tang YA, Ternert T, Tikhonov A, Welter D, Williams E, Brazma A, Parkinson H, Sarkans U: **ArrayExpress update—trends in database growth and links to data analysis tools.** *Nucleic Acids Res* 2013, **41**:D987–D990.

137. Zhang F, Drabier R: **IPAD: the integrated pathway analysis database for systematic enrichment analysis.** *BMC Bioinform* 2012, **13**(Suppl 15):S7.
138. Przytycka TM, Singh M, Slonim DK: **Toward the dynamic interactome: it's about time.** *Brief Bioinform* 2010, **11**:15–29.
139. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M: **The KEGG resource for deciphering the genome.** *Nucleic Acids Res* 2004, **32**:D277–D280.
140. Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L, Lewis S, Birney E, Stein L: **Reactome: a knowledgebase of biological pathways.** *Nucleic Acids Res* 2005, **33**:D428–D432.
141. van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, Evelo C: **Presenting and exploring biological pathways with PathVisio.** *BMC Bioinform* 2008, **9**:399.
142. Paley SM, Latendresse M, Karp PD: **Regulatory network operations in the pathway tools software.** *BMC Bioinform* 2012, **13**:243.
143. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J: **ClueGO: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.** *Bioinformatics* 2009, **25**:1091–1093.
144. Maere S, Heymans K, Kuiper M: **BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21**:3448–3449.
145. Berriz GF, Beaver JE, Cenik C, Tasan M, Roth FP: **Next generation software for functional trend analysis.** *Bioinformatics* 2009, **25**:3043–3044.
146. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q: **The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function.** *Nucleic Acids Res* 2010, **38**:W214–W220.
147. Bauer-Mehren A, Rautschka M, Sanz F, Furlong LL: **DisGeNET: a cytoscape plugin to visualize, integrate, search and analyze gene-disease networks.** *Bioinformatics* 2010, **26**:2924–2926.
148. Merico D, Isserlin R, Stueker O, Emili A, Bader GD: **Enrichment map: a network-based method for gene-set enrichment visualization and interpretation.** *PLoS One* 2010, **5**:e13984.
149. Yang L, Walker JR, Hogenesch JB, Thomas RS: **NetAtlas: a cytoscape plugin to examine signaling networks based on tissue gene expression.** *In Silico Biol* 2008, **8**:47–52.
150. Doncheva NT, Assenov Y, Domingues FS, Albrecht M: **Topological analysis and interactive visualization of biological networks and protein structures.** *Nat Protoc* 2012, **7**:670–685.
151. Moulos P, Klein J, Jupp S, Stevens R, Bascands JL, Schanstra JP: **The KUPNetViz: a biological network viewer for multiple -omics datasets in kidney diseases.** *BMC Bioinform* 2013, **14**:235.
152. Razick S, Mora A, Michalickova K, Boddie P, Donaldson IM: **iRefScape: a cytoscape plug-in for visualization and data mining of protein interaction data from iRefIndex.** *BMC Bioinform* 2011, **12**:388.
153. Gao J, Ade AS, Tarcea VG, Weymouth TE, Mirel BR, Jagadish HV, States DJ: **Integrating and annotating the interactome using the MiMi plugin for cytoscape.** *Bioinformatics* 2009, **25**:137–138.
154. Avila-Campillo I, Drew K, Lin J, Reiss DJ, Bonneau R: **BioNetBuilder: automatic integration of biological networks.** *Bioinformatics* 2007, **23**:392–393.
155. Martin A, Ochagavia ME, Rabasa LC, Miranda J, Fernandez-de-Cossio J, Bringas R: **BisoGenet: a new tool for gene network building, visualization and analysis.** *BMC Bioinform* 2010, **11**:91.
156. Creek DJ, Jankevics A, Burgess KE, Breitling R, Barrett MP: **IDEOM: an excel interface for analysis of LC-MS-based metabolomics data.** *Bioinformatics* 2012, **28**:1048–1049.
157. Scheltema RA, Jankevics A, Jansen RC, Swertz MA, Breitling R: **PeakML/mzMatch: a file format, Java library, R library, and tool-chain for mass spectrometry data analysis.** *Anal Chem* 2011, **83**:2786–2793.
158. Barabasi AL, Gulbahce N, Loscalzo J: **Network medicine: a network-based approach to human disease.** *Nat RevGenetics* 2011, **12**:56–68.
159. Chowdhury SA, Koyuturk M: **Identification of coordinately dysregulated subnetworks in complex phenotypes.** In *Pacific Symposium on Biocomputing 2010*. Edited by Altman RB, Dunker AK, Hunter L, Murray TA, Klein TE. Hackensack: World Scientific Publishing Co; 2011:133–144.
160. Chowdhury SA, Nibbe RK, Chance MR, Koyuturk M: **Subnetwork state functions define dysregulated subnetworks in cancer.** *J Comput Biol* 2011, **18**:263–281.
161. Dao P, Colak R, Salari R, Moser F, Davicioni E, Schonhuth A, Ester M: **Inferring cancer subnetwork markers using density-constrained biclustering.** *Bioinformatics* 2010, **26**:i625–i631.
162. Dao P, Wang K, Collins C, Ester M, Lapuk A, Sahinalp SC: **Optimally discriminative subnetwork markers predict response to chemotherapy.** *Bioinformatics* 2011, **27**:i205–i213.
163. Lee E, Chuang HY, Kim JW, Ideker T, Lee D: **Inferring pathway activity toward precise disease classification.** *PLoS Comput Biol* 2008, **4**:e1000217.
164. Muller FJ, Laurent LC, Kostka D, Ulitsky I, Williams R, Lu C, Park IH, Rao MS, Shamir R, Schwartz PH, Schmidt NO, Loring JF: **Regulatory networks define phenotypic classes of human stem cell lines.** *Nature* 2008, **455**:401–405.
165. Suthram S, Dudley JT, Chiang AP, Chen R, Hastie TJ, Butte AJ: **Network-based elucidation of human disease similarities reveals common functional modules enriched for pluripotent drug targets.** *PLoS Comput Biol* 2010, **6**:e1000662.
166. Chu LH, Chen BS: **Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets.** *BMC Syst Biol* 2008, **2**:56.
167. Baker SG: **The central role of receiver operating characteristic (ROC) curves in evaluating tests for the early detection of cancer.** *J Natl Cancer Inst* 2003, **95**:511–515.
168. Peres J: **Risks of PSA screening now better understood.** *J Natl Cancer Inst* 2013, **105**:1590–1592.
169. Andriole GL, Crawford ED, Grubb RL III, Buys SS, Chia D, Church TR, Fouad MN, Isaacs C, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Ragard LR, Clapp JD, Rathmell JM, Riley TL, Hsing AW, Izmirlian G, Pinsky PF, Kramer BS, Miller AB, Gohagan JK, Prorok PC: **Prostate cancer screening in the randomized prostate, lung, colorectal, and ovarian cancer screening trial: mortality results after 13 years of follow-up.** *J Natl Cancer Inst* 2012, **104**:125–132.
170. Brenner DE, Normolle DP: **Biomarkers for cancer risk, early detection, and prognosis: the validation conundrum.** *Cancer Epidemiol Biomarkers Prev* 2007, **16**:1918–1920.
171. Borrebaeck CA, Wingren C: **Antibody array generation and use.** *Methods Mol Biol* 2014, **1131**:563–571.
172. Kozak KR, Wang J, Lye M, Chuh J, Takkar R, Kim N, Lee H, Jeon NL, Lin K, Zhang C, Wong WL, DeForge LE: **Micro-volume wall-less immunoassays using patterned planar plates.** *Lab Chip* 2013, **13**:1342–1350.
173. Ostendorff HP, Awad A, Braunschweiger KI, Liu Z, Wan Z, Rothschild KJ, Lim MJ: **Multiplexed VeraCode bead-based serological immunoassay for colorectal cancer.** *J Immunol Methods* 2013, **400**:401:58–69.
174. Tekin HC, Gijb MA: **Ultrasensitive protein detection: a case for microfluidic magnetic bead-based assays.** *Lab Chip* 2013, **13**:4711–4739.
175. Fu Q, Zhu J, Van Eyk JE: **Comparison of multiplex immunoassay platforms.** *Clin Chem* 2010, **56**:314–318.
176. Cho M, Soo Oh S, Nie J, Stewart R, Eisenstein M, Chambers J, Marth JD, Walker F, Thomson JA, Soh HT: **Quantitative selection and parallel characterization of aptamers.** *Proc Natl Acad Sci USA* 2013, **110**:18460–18465.
177. Martinez O, Bellard E, Golzio M, Mechiche-Alami S, Rols MP, Teissie J, Ecochard V, Paquereau L: **Direct validation of aptamers as powerful tools to image solid tumor.** *Nucleic Acid Ther.* in press.
178. Chung H, Kim B, Jung SH, Won KJ, Jiang X, Lee CK, Lim SD, Yang SK, Song KH, Kim HS: **Does phosphorylation of cofilin affect the progression of human bladder cancer?** *BMC Cancer* 2013, **13**:45.
179. Frantzi M, Zoidakis J, Papadopoulos T, Zurbig P, Katafigiotis I, Stravodimos K, Lazaris A, Giannopoulou I, Ploumidis A, Mischak H, Mullen W, Vlahou A: **IMAC fractionation in combination with LC-MS reveals H2B and NIF-1 peptides as potential bladder cancer biomarkers.** *J Proteome Res* 2013, **12**:3969–3979.
180. Kato M, Wei M, Yamano S, Kakehashi A, Tamada S, Nakatani T, Wanibuchi H: **DDX39 acts as a suppressor of invasion for bladder cancer.** *Cancer Sci* 2012, **103**:1363–1369.
181. Orenes-Pinero E, Corton M, Gonzalez-Peramato P, Algaba F, Casal I, Serrano A, Sanchez-Carbayo M: **Searching urinary tumor markers for bladder cancer using a two-dimensional differential gel electrophoresis (2D-DIGE) approach.** *J Proteome Res* 2007, **6**:4440–4448.
182. Zoidakis J, Makridakis M, Zerefos PG, Bitsika V, Esteban S, Frantzi M, Stravodimos K, Anagnostou NP, Roubelakis MG, Sanchez-Carbayo M, Vlahou A: **Profilin 1 is a potential biomarker for bladder cancer aggressiveness.** *Mol Cell Proteomics* 2012, **11**:M111.009449.

183. Lee JY, Kim JY, Cheon MH, Park GW, Ahn YH, Moon MH, Yoo JS: **MRM validation of targeted nonglycosylated peptides from N-glycoprotein biomarkers using direct trypsin digestion of undepleted human plasma.** *J Proteomics* 2014, **98C**:206–217.
184. Lemoine J, Fortin T, Salvador A, Jaffuel A, Charrier JP, Choquet-Kastylevsky G: **The current status of clinical proteomics and the use of MRM and MRM (3) for biomarker validation.** *Expert Rev Mol Diagn* 2012, **12**:333–342.
185. Yang N, Feng S, Shedden K, Xie X, Liu Y, Rosser CJ, Lubman DM, Goodison S: **Urinary glycoprotein biomarker discovery for bladder cancer detection using LC/MS-MS and label-free quantification.** *Clin Cancer Res* 2011, **17**:3349–3359.
186. Raab SS, Grzybicki DM, Vrbin CM, Geisinger KR: **Urine cytology discrepancies: frequency, causes, and outcomes.** *Am J Clin Pathol* 2007, **127**:946–953.
187. Chang YH, Wu CH, Lee YL, Huang PH, Kao YL, Shiau MY: **Evaluation of nuclear matrix protein-22 as a clinical diagnostic marker for bladder cancer.** *Urology* 2004, **64**:687–692.
188. Halling KC, King W, Sokolova IA, Karnes RJ, Meyer RG, Powell EL, Sebo TJ, Cheville JC, Clayton AC, Krajnik KL, Ebert TA, Nelson RE, Burkhardt HM, Ramakumar S, Stewart CS, Pankratz VS, Lieber MM, Blute ML, Zincke H, Seelig SA, Jenkins RB, O'Kane DJ: **A comparison of BTA stat, hemoglobin dipstick, telomerase and Vysis UroVysion assays for the detection of urothelial carcinoma in urine.** *J Urol* 2002, **167**:2001–2006.
189. Sullivan PS, Nooraie F, Sanchez H, Hirschowitz S, Levin M, Rao PN, Rao J: **Comparison of ImmunoCyt, UroVysion, and urine cytology in detection of recurrent urothelial carcinoma: a "split-sample" study.** *Cancer* 2009, **117**:167–173.
190. Cheng L, Davison DD, Adams J, Lopez-Beltran A, Wang L, Montironi R, Zhang S: **Biomarkers in bladder cancer: translational and clinical implications.** *Crit Rev Oncol Hematol.* in press.
191. Kamat AM, Hegarty PK, Gee JR, Clark PE, Svatek RS, Hegarty N, Shariat SF, Xylinas E, Schmitz-Drager BJ, Lotan Y, Jenkins LC, Droller M, van Rhijn BW, Karakiewicz PI: **ICUD-EAU International Consultation on Bladder Cancer 2012: screening, diagnosis, and molecular markers.** *Eur Urol* 2013, **63**:4–15.
192. Cho WC: **Proteomics and translational medicine: molecular biomarkers for cancer diagnosis, prognosis and prediction of therapy outcome.** *Expert Rev Proteomics* 2011, **8**:1–4.
193. Chen CL, Lin TS, Tsai CH, Wu CC, Chung T, Chien KY, Wu M, Chang YS, Yu JS, Chen YT: **Identification of potential bladder cancer markers in urine by abundant-protein depletion coupled with quantitative proteomics.** *J Proteomics* 2013, **85**:28–43.
194. Grau L, Luque-Garcia JL, Gonzalez-Peramato P, Theodorescu D, Palou J, Fernandez-Gomez JM, Sanchez-Carbayo M: **A quantitative proteomic analysis uncovers the relevance of CUL3 in bladder cancer aggressiveness.** *PLoS One* 2013, **8**:e53328.
195. Shimwell NJ, Bryan RT, Wei W, James ND, Cheng KK, Zeegers MP, Johnson PJ, Martin A, Ward DG: **Combined proteome and transcriptome analyses for the discovery of urinary biomarkers for urothelial carcinoma.** *Br J Cancer* 2013, **108**:1854–1861.
196. Urquidí V, Rosser CJ, Goodison S: **Molecular diagnostic trends in urological cancer: biomarkers for non-invasive diagnosis.** *Curr Med Chem* 2012, **19**:3653–3663.
197. Mischak H, Ioannidis JP, Argiles A, Attwood TK, Bongcam-Rudloff E, Broenstrup M, Charonis A, Chrousos GP, Delles C, Dominiczak A, Dylag T, Ehrlich J, Egido J, Findeisen P, Jankowski J, Johnson RW, Julien BA, Lankisch T, Leung HY, Maahs D, Magni F, Manns MP, Manolis E, Mayer G, Navis G, Novak J, Ortiz A, Persson F, Peter K, Riese HH, et al: **Implementation of proteomic biomarkers: making it work.** *Eur J Clin Invest* 2012, **42**:1027–1036.
198. Majewski T, Spiess PE, Bondaruk J, Black P, Clarke C, Benedict W, Dinney CP, Grossman HB, Tang KS, Czerniak B: **Detection of bladder cancer using proteomic profiling of urine sediments.** *PLoS One* 2012, **7**:e42452.
199. Good DM, Zurbig P, Argiles A, Bauer HW, Behrens G, Coon JJ, Dakna M, Decramer S, Delles C, Dominiczak AF, Ehrlich JH, Eitner F, Fliser D, Frommberger M, Ganser A, Girolami MA, Golovko I, Gwinner W, Haubitz M, Herget-Rosenthal S, Jankowski J, Jahn H, Jerums G, Julian BA, Kellmann M, Kliem V, Kolch W, Krolewski AS, Luppi M, Massy Z, et al: **Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease.** *Mol Cell Proteomics* 2010, **9**:2424–2437.
200. Molin L, Seraglia R, Lapolla A, Ragazzi E, Gonzalez J, Vlahou A, Schanstra JP, Albalat A, Dakna M, Siwy J, Jankowski J, Bitsika V, Mischak H, Zurbig P, Traldi P: **A comparison between MALDI-MS and CE-MS data for biomarker assessment in chronic kidney diseases.** *J Proteomics* 2012, **75**:5888–5897.
201. Zurbig P, Jerums G, Hovind P, Macisaac RJ, Mischak H, Nielsen SE, Panagiotopoulos S, Persson F, Rossing P: **Urinary proteomics for early diagnosis in diabetic nephropathy.** *Diabetes* 2012, **61**:3304–3313.
202. Andersen S, Mischak H, Zurbig P, Parving HH, Rossing P: **Urinary proteome analysis enables assessment of renoprotective treatment in type 2 diabetic patients with microalbuminuria.** *BMC Nephrol* 2010, **11**:29.
203. Lankisch TO, Metzger J, Negm AA, Vosskuhl K, Schiffer E, Siwy J, Weismuller TJ, Schneider AS, Thedieck K, Baumeister R, Zurbig P, Weissinger EM, Manns MP, Mischak H, Wedemeyer J: **Bile proteomic profiles differentiate cholangiocarcinoma from primary sclerosing cholangitis and choledocholithiasis.** *Hepatology* 2011, **53**:875–884.
204. Metzger J, Negm AA, Plentz RR, Weismuller TJ, Wedemeyer J, Karlsen TH, Dakna M, Mullen W, Mischak H, Manns MP, Lankisch TO: **Urine proteomic analysis differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders.** *Gut* 2013, **62**:122–130.

doi:10.1186/2001-1326-3-7

**Cite this article as:** Frantzi et al.: Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clinical and Translational Medicine* 2014 **3**:7.

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)