UNIVERSITA' DEGLI STUDI DI UDINE

CORSO DI DOTTORATO DI RICERCA IN
MEDICINA CELLULARE E MOLECOLARE
CICLO XXIX

PhD Thesis

IDENTIFICATION OF NEW CANDIDATE BIOMARKERS FOR PROSTATE CANCER BY AFFINITY PROTEOMICS

PhD candidate: Dr. Elisa Pin

Supervisor: Prof. Francesco Curcio (Universita´ di Udine)
Co-supervisors: Prof. Peter Nilsson (KTH, Stockholm)
Dr. Mariaelena Pierobon (George Mason University, Manassas)

ANNO ACCADEMICO 2015/2016

ABSTRACT

Prostate cancer (PCA) is a complex malignancy that needs to be more thoroughly studied and understood at a molecular level to fill the current knowledge gap, and optimize diagnosis and to treatment. Prostate specific antigen (PSA) showed to be not specific for PCA, therefore a demand for novel specific biomarkers exists. The aim of our work was to identify new specific candidate biomarkers for PCA in tissue and plasma samples by means of affinity proteomics approaches such as reverse phase protein array and antigen arrays. Tissue samples are an invaluable source of biomarkers for cancer, but very limited in amount and requiring invasive procedures for collection. Still, they allow to directly profiling the molecular status of tumor itself. Beside tissue, a screening procedure on biological fluid such as plasma would be highly desirable, thanks to the less invasiveness and low-costs of samples collection. Among the biomarkers detectable in plasma are the autoantibodies. The first part of this thesis summarizes the current status of PCA epidemiology, treatment, and biomarkers research. Beside this, an overview of the affinity proteomics platforms available for biomarkers research, and the critical variables to consider in the biomarkers validation process are presented. The second part of the thesis reports the main results of two original studies where the author of the thesis is the main contributor. Paper I is based on the profiling of PCA tissue samples using RPPA. Our results indicate the feasibility of combining laser capture microdissection (LCM) and RPPA for evaluating the molecular architecture and cross-talking of epithelial and stromal compartments. Paper II is based on profiling the autoimmune response to PCA patients, comparing early and late stage of the disease. The authors identified and characterized the IgG reactivity toward a novel epitope for the candidate biomarker *prostein*. The data presented in this thesis provide two robust frameworks based on affinity proteomics platforms applied for protein profiling in tissue, and autoantibodies profiling in plasma in the context of PCA biomarkers discovery.

TABLE OF CONTENT

1. Prostate Cancer: biology, epidemiology and treatment

The human prostate is a walnut-sized exocrine gland of the male reproductive system. Its function is to produce alkaline substances that constitute the 30% of the semen volume and have the function to protect and prolong lifespan of sperm[1]. Anatomically, prostate can be subdivided into three morphologically distinct regions: the peripheral zone, the transition zone and the central zone. Benign prostatic hyperplasia (BPH) is a non-malignant growth, which commonly involves the transition zone in aging men. This zone surrounds the urethra as it passes through the prostate, and makes up about 20% of the prostate gland until the age of 40. On the other hand, prostate cancer (PCA) is a malignant growth that arises primarily from the prostate gland epithelial cells of the peripheral zone, and is therefore described as an *adenocarcinoma* [2-4] . The current PCA carcinogenesis model is based on a multistep process involving pre-neoplastic disorders. An initial inflammatory insult leads to proliferative inflammatory atrophy (PIA). PIA was then shown to merge with areas of prostatic intraepithelial neoplasia (PIN) and cancer tissue, providing evidences for the existence of a field cancerization effect and a step-wise progression[5].

According to the last data from the World Health Organization (WHO), PCA is the second cancer in men for incidence and the fifth cause of death by cancer in men worldwide, representing a public health burden[6]. However, the incidence of PCA worldwide is not homogeneous and its specific geographic and temporal patterns are still under study. Data from 2012 report an incidence-rate variation of 25-fold worldwide, with the highest rate in countries at higher socio-economical development[7]. In general, while PCA incidence has increased over time, cancer mortality has decreased, even though with relatively less variation and except for Africa where the highest rate of mortality is reported[8]. However, the described geographical and temporal distribution of PCA incidence most likely reflects the widespread of clinical practices such prostate specific antigen (PSA) testing and biopsy, rather than a real difference in incidence[9].

Age remains the main risk factor predisposing to PCA, which rarely occurs before 50 years and is more frequently diagnosed among men aged 65 and 74 with a median of 66 years [10]. Besides age and socio-economical situation, other established risk factors predisposing to PCA are familiarity [11], the presence of genetic polymorphisms and mutations [12], a fat-rich diet[13, 14], and low testosterone concentration in serum [15]. PCA incidence also varies by race and ethnicity, with the highest risk rate reported in black men. The reason is not yet fully understood, even though it is probably related both to genetic influences [16], and socio-economical reasons [17, 18].

PCA still remains a mostly indolent disease, with only 8% of the cases becoming clinically apparent[19] and almost 100% of them reaching the five-year survival rate when the disease is at local and regional stage; on the contrary, the survival rate falls to 30% in case of aggressive form with distant metastasis (e.g. bones, liver, brain, and lung) and relapse after treatment[7].

The current Food and Drug Administration (FDA) guidelines for the diagnosis of PCA recommend testing PSA concentration in blood, and digital rectal exploration (DRE) for men over 50 years of age. This well established screening practice is responsible for the improvement in early PCA diagnosis and the lowering of mortality rate, yet it is not free from drawbacks. In fact, the low specificity of PSA as a PCA biomarker, which is shown to increase also in prostatitis or BHP[20], requires further more invasive exams such as biopsy to exclude false positives. Beside this, false negatives also represent an issue; it was reported that the 15% of PCA patients with a Gleason score of 7 or higher have PSA level of 4 ng/ml or lower and are negative for DRE [20]. For these reasons, the role of PSA in PCA diagnosis has become controversial, and the existing data suggest it should be considered more as a prostate volume indicator rather than a biomarker of malignancy.

Besides the urgency for an early and specific PCA diagnosis, the disparity between the incidence and mortality rate highlights the importance to adequately stratify PCA patients and distinguish the aggressive forms from the indolent forms, that is to optimize treatment for the firsts and avoiding overtreatment of the latters. Currently used prognostic indicators for PCA are the classical TNM staging system[21], based on disease localization and spreading status, and the Gleason-grading method, which estimates tissue differentiation and histological pattern of cells differentiation[22]. Five basic Gleason grades

are used to generate a histological score ranging from 2 to 10, where a higher score correspond to cancer undifferentiation. Gleason grade has also been linked to clinical endpoints, such as clinical stage, progression and survival. Gleason grading of PCA needle biopsies is routinely used to plan patient management and prediction of response to radiotherapy and surgery.

To date, therapeutic approaches are defined dependently on PCA stage, Gleason-grading and on the age of the patient[23, 24]. For low grade PCA (<7) the first stratification of patients for treatment is based on age. Men aged above 70 with low Gleason score are routinely followed-up with watchful-waiting as standard-of-care approach, together with periodic PSA level checking, DRE and, when necessary, new biopsies. For younger patients with low Gleason score, prostatectomy is the treatment of choice thanks to the better outcome of the procedure, though is often accompanied by heavy side-effects like incontinence and impotence. The standard treatment for aggressive disease, independently of age, involves radical prostatectomy, hormonal therapy, radiation, chemotherapy, ultrasound treatment (HIFU) or cryosurgery. Regression is reported in 75% of the patients following surgery and pharmacological castration. However, the possible insurgence of a wide number of side effects such as osteoporosis, cognitive decline, cardiovascular morbidity, obesity, fatigue, and sexual dysfunctions, should be considered before therapy administration. Despite the development of differentiated treatment guidelines for high and low PCA grade, so far research has not been able to identify specific biomarkers able to complement Gleason scoring and improving the accuracy of PCA aggressiveness prediction. In fact, Gleason scoring proved to be effective for predicting the outcome and define the best treatment for the majority of PCAs with scores <6 and > 7, but it is poorly indicative for PCAs with Gleason 6 and 7, which are characterized by a largely unpredictable outcome. It is not unusual that PCAs with Gleason scores <6 progress towards aggressive disease, and in the same manner Gleason >7 can present with indolent disease. Another major concern in PCA treatment is the insurgence of androgen independency after 12–15 months from the beginning of the ablation therapy, which could lead to Androgen-independent PCA (AIPC) developing by clonal selection of hypermutated cells after the therapy. No treatment is currently available for AIPC which is usually characterized by a fatal outcome.

As for many other types of cancer, cellular heterogeneity found in PCA represents a challenge for treatment and diagnosis. It is widely accepted that PCA contains subpopulations of cells resistant to treatments that can be the initiators of metastasis[25, 26]. Notably, metastatic PCA cells showed both inter- and intra-individual genotype and phenotype variations [27], and genetic heterogeneity can be traced already in the primary carcinoma, as shown by deep molecular profiling [28]. Although not associated with a high mortality rate, PCA is a complex malignancy that needs to be more thoroughly studied and understood at a molecular level to improve on the currently available diagnostic and treatment strategies. The aim of our work was to identify new specific biomarkers for PCA by means of affinity proteomics approaches that will be discussed in details in the following sections.

2. The Proteome as a Source of Biomarkers

2.1 Proteins, Proteome and Proteomics

Since the Central Dogma of the Molecular Biology was stated by Francis Crick in 1953[29], enormous progress was made in our understanding of the genetic information flow from DNA to proteins in biological systems. The Human Genome Project (HGP) was launched in 1990 and completed in 2003[30], on the 50th anniversary of the DNA structure discovery by James D. Watson and Francis Crick. For the first time, the complete genetic blueprint that builds a human being, named *genome*, was fully readable. This opened the era of *genomics*, as the discipline that sequence, assemble, and analyze the genome structure and functions. To date, the Ensemble Database-version 86 counts 20,441 protein-encoding genes in the human genome. Each of these genes can give rise to several protein variants due to genomic recombination, open reading frames, and alternative splicing [31, 32]. DNA and genes contain all the genetic instructions for organisms' development and functioning but proteins are the true functional units of the cell. Proteins participate in all cellular processes from biochemical reactions to cell structure, and inter-cellular cross-talking. Their active or inactive status is dictated by post-translational modifications (PTMs)[33]. PTMs are chemical modifications that occur after protein translation, such as glycosylation, phosphorylation, ubiquitination, methylation, nitrosylation, acetylation, lipidation and proteolysis. PTMs can occur in every phase of the life of a protein. Moreover, whereas genetic mutations can only occur once per position, different PTMs may occur in tandem. PTMs are most often mediated by enzymes such as kinases, phosphatases, transferases and ligases, which add or remove functional groups, lipids or sugars to or from amino acid side chains, and proteases, which cleave specific sequences or regulatory subunits. Beside this, autocatalytic domains also allow proteins to self-modify.

The total complement of proteins present in a cell is known as its *proteome*, and the discipline that has the aim of deciphering it is known as *proteomics*.

The human proteome is incredibly complex and dynamic (Fig. 1). Proteins interact with other proteins and are organized in pathways and networks that regulate all cell functions and processes, from duplication to death. Indeed, protein-protein interactions are responsible for triggering signal transduction cascades to respond to stimuli coming from inside or outside the cell, and cause modifications of cellular processes. Networks of proteins adapt in response to multiple concomitant stimuli, determining an observable phenotype. Pathological phenotypes appear when variations occur at genetic, translational or post-translational level, determining an alteration in protein expression, conformation or activation that interferes with the normal interaction with other proteins. Indeed, beside the central dogma of the molecular biology, a new paradigm has been recently introduced where *network biology* or *system biology* represents the conjunction between genotype and phenotype[34].

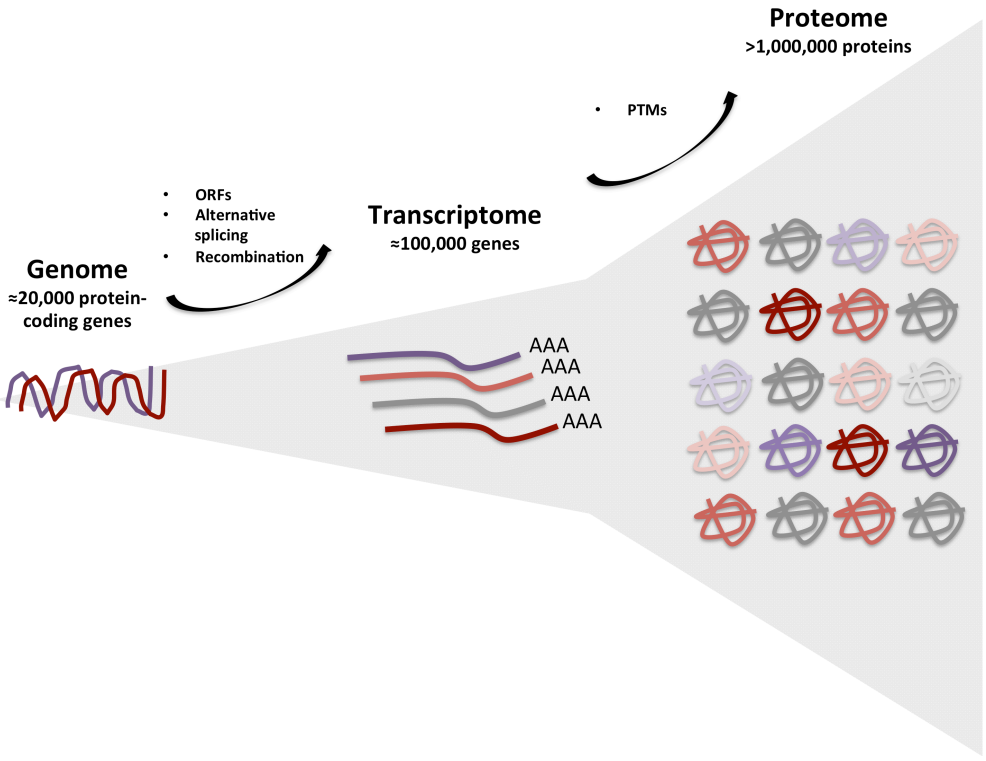

Figure 1. Complexity of the human proteome. While the genome counts around 20000 genes, the human proteome is estimated to be composed by over 1 million proteins. Changes at gene levels and different ORFs cause the production of many different transcripts from a single gene. At a post-traslational level are the PTMs that represent the major source of protein diversity.

PTMs are key regulators of protein-protein interaction or protein-biomolecules interaction, representing the base of signal transduction. Identifying and understanding PTMs is a critical step in the study of cell biology and disease diagnosis, treatment and prevention. Based on previous estimates, it is widely believed that the most abundant PTM is glycosylation. However, *phosphorylation* has a leading role as controller for the *kinome* and its aberrant activation in human cancers [35-37]. The vast majority (90%) of protein phosphorylation occurs on serine and threonine residues, while only 10% occurs on tyrosine. Growth factor signaling starts from tyrosine kinase receptors (e.g. EGFR, HER2 and VEGFR), which after binding to the ligand, dymerize and autophosphorylate, and trigger a cascade of events resulting in further interactions with downstream kinases [38-41].

Proteomics and its branches (e.g. phosphoproteomics) have a foreground role in clarifying molecular processes underlying normal and disease status, therefore representing an invaluable tool for the development of diagnostic, prognostic, therapeutic, and preventive medical applications. In 2010 Human Proteome Organization (HUPO) started an international effort named Human Proteome Project (HPP; http://www.thehpp.org/). The HPP mission is to generate the map of the human proteome and to become a source to help elucidating biological and molecular function, and improving diagnosis and treatment of diseases. Many research laboratories around the world participate in the project, which still represents a challenge in life science. Emerging technologies based on mass spectrometry (MS) and affinity proteomics are applied to reach this goal. The Human Protein Atlas (HPA) project is part of this effort, and aims to map the spatial distribution for the whole human proteome in tissues and cell lines by means of antibodies. Part of the work presented in this thesis is based on data generated by using antigen arrays obtained from the HPA. Further details about HPA are discussed in paragraph 3.1.1.

2.2 Protein Biomarkers for cancer and personalized medicine

The National Institute of Health (NIH) defines a biomarker as "a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease." Additionally, "A biomarker may be used to see how well the body responds to treatment for a disease or condition". All kind of biomolecules (e.g., proteins, nucleic acids, antibodies, peptides, etc.) can serve as biomarkers[42] but, as previously stated, proteins represent the best indicator to monitor a biological status or response to stimuli, therefore representing the ideal candidates for biomarkers.

A protein biomarker is a protein that can be measured in body fluids or tissues, and reflects a certain status of an individual. Protein biomarkers for cancer can be divided into diagnostic, prognostic, and predictive[43]. Diagnostic biomarkers are used to detect a disease in an individual, while prognostic biomarkers are helpful in defining the disease progression, aggressiveness and recurrence. Predictive biomarkers are useful after diagnosis, and allow selecting the optimal treatment by stratifying patients based on their responsiveness or non-responsiveness.

Proteins are the direct targets of new personalized treatments. The main targets of currently FDA approved monoclonal antibodies are kinases or proteins belonging to the signaling networks involved in cancer cells survival, proliferation and migration[44]. The activation or inhibition of a protein kinase gives rise to a cascade of events that modify the status of a pathway, or of the entire protein network modulating cellular status and processes [45]. As above mentioned, mapping the proteomic and phosphoproteomic network is fundamental to have a clear picture of the status of a cell and tissue.

Tissue samples are invaluable source of cancer biomarkers, even though they are very limited in amount and requiring invasive procedures for collection. To overcome these limits, a screening procedure based on biological fluid testing (e.g. plasma, serum, urine, saliva) would be highly desirable, due to the possibility of collecting much higher

volumes, with less invasiveness and low-costs. Biological fluids are not just practically convenient; they represent a real important source of proteins deriving from tissues by secretion, or released as a consequence of tissue damage or molecular changes related to a disease, included cancer. Disease conditions can deregulate the conventional processes of protein translocation and secretion, determining changes of protein levels in body fluids that do not necessary reflect an aberrant expression of the same protein at the tissue level. Such variation in protein levels into the circulation is therefore not predictable from tissues analysis alone. Yet, a mutated form of a protein – either at translational or post-translational level - may be expressed by tumor tissue at different levels compared to the normal tissue, but with a negligible resultant increment in the circulating level due to the greater mass of unaffected tissue releasing the native protein.

Autoantibodies, produced by the organism to recognize tumor-associated antigens (TAAs), are particularly notable amongst serological cancer biomarkers [46] and recently the scientific community invested a lot of effort in correlating their expression with clinical parameters, thus demonstrating their usefulness in diagnostic and prognostic processes[47].

It is easy to understand that an ideal protein cancer biomarker should originate from the affected tissue, be present with detectable concentrations in body fluids, correlate with the disease severity, and specific for the considered cancer type, with high sensitivity (low rate of false negatives) and high specificity (low rate of false positives)[48]. All these requirements are unlikely to be fulfilled by a single protein; beside this, the complexity of cancer also suggests that a panel-of-biomarkers would be more effective.

Proteomics and phosphoproteomics demonstrated to be highly promising for biomarkers discovery. The continuous technological improvement (e.g. throughput, multiplexing and sensitivity) have allowed the detection and quantification of proteins in many different body fluids and in tissue[49, 50]. The biomarker pipeline is commonly constituted by a series of preclinical phases (e.g., biomarker discovery, verification, evaluation) followed by a final clinical evaluation (Paulovich et al., 2008; Rifai et al., 2006; Rodriguez et al., 2010a; Surinova et al., 2011). Through this process, biomarkers are "filtered". Hundreds or thousands of targets are analyzed in only few samples at earlier phases, while a handful of targets are measured in thousands of samples at late phases (Fig.

2). An appropriate proteomic platform has to be chosen for each phase, based on throughput and multiplexing capability.

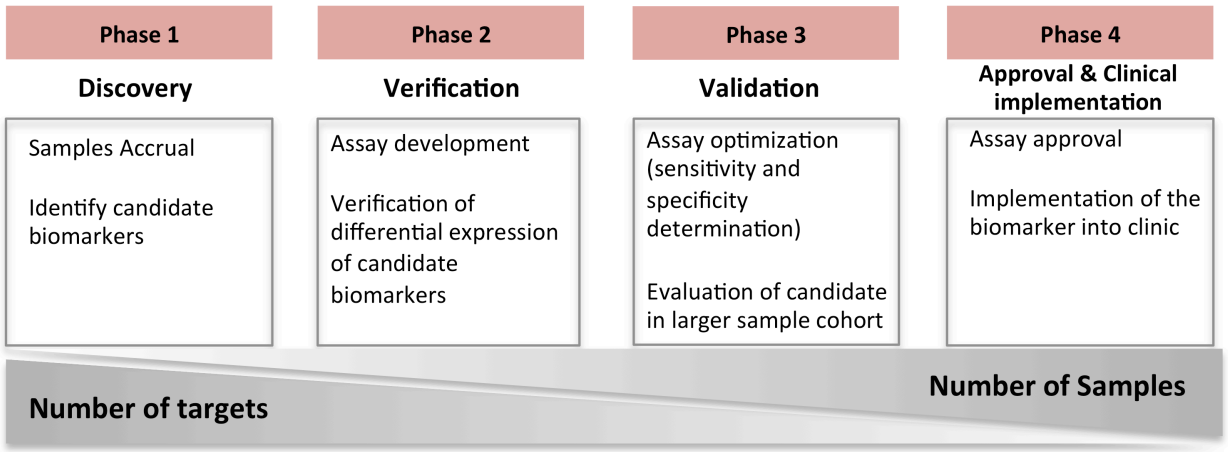

Figure 2. Biomarkers discovery and implementation into clinics. The diagram summarize the four phases in which the process from candidate biomarker discovery to implementation into clinics, passing through verification and validation.

During the last 15 years, thousands of potential protein biomarkers for cancer and other diseases have been published. However, some major challenges including pre-analytical, analytical, and post-analytical variables still represent a barrier to biomarkers validation and introduction in the clinical practice and personalized cancer treatment [51]. We will discuss these aspects in details in the following sections of this thesis.

3. Affinity Proteomics in Biomarkers Discovery and validation

3.1 Affinity Proteomics for a proteome-wide analysis

From a technological point of view, proteomics can be subdivided into two main branches. Mass-spectrometry (MS)-based proteomics and affinity-based proteomics represent two alternatives for approaching a proteomic question. MS has become the gold standard for analysis of complex protein samples, thanks to the availability of gene and genome sequence databases, and most notably thanks to the development of protein digestion and ionization methods which led to the 2002 Nobel prize in chemistry[52, 53]. Affinity proteomics, instead, is the field of proteomics based on the use of antibodies and other binding reagents as protein-specific detection probes[49, 54]. The two merge in immune-MS, which couple the selectivity of affinity enrichment to the certain protein identification provided by MS. This approach is promising in the detection of low-abundance protein biomarkers[55] and in the study of protein-protein interaction (PPI)[56]. However, MS is not exempt drawbacks, such as low throughput and sensitivity, and high costs.

In the medical area, affinity proteomics plays an important role in the identification of biomarkers, as well as for identifying new drug targets[49]. Affinity proteomics performance depends strictly on the availability of specific binders, and an availability of sensitive and high throughput detection methods.

Several different types of affinity reagents are available, but antibodies are still the most commonly used tools in affinity-based assays. Despite the commercial availability of around 2,000,000 antibodies, only 500,000 are unique and in many cases characterized by a variable quality. Published data report that the 25% of the commercially available antibodies fail specificity and are inconsistently performing across different application[57], though forcing each laboratory to run validation tests. Comprehensive resources collecting

quality-controlled and well-performing binders are required. Several antibody catalogues are available online. Antibodypedia is one such catalogue, and counts more than 1,800,000 antibodies covering 94% of the human protein-encoding genes[58]. Other resources include Biocompare (http://www. biocompare.com/), antibodies-online (http://www.antibo dies-online.com/) and Antibody Registry (http://www.antibodyregistry.org/). The feasibility of producing affinity-purified polyclonal antibodies directed toward the whole human proteome is represented by the HPA project (details in Paragraph 3.1.1). A major drawback of polyclonal antibodies is their finite availability; therefore, a big effort is now dedicated to the production of renewable affinity reagents, such as monoclonal antibodies, and recombinant affinity reagents. Among the latter we can cite antibody fragments[59], SOMAmers[60], nanobodies, camelids single-domain antibodies[61, 62], and affibodies[63]. The high cost of production still represent a barrier to the application of these affinity binders on a broad scale. Affinity binders can be applied to create affinity-based assays in several different formats. In the following sections I will discuss the two protein array formats applied in the work presented in this thesis: antigen arrays and reverse-phase protein array.

3.1.1 The Human Protein Atlas Project

A map of the human proteome, similar to the one for the genome, could lead to a better understanding of the molecular basis of diseases, fill the gap to explain complex pathological conditions, and improve the treatment strategy. For these reasons, large scale-efforts primarily based on mass-spectrometry technology have been made to identify all the protein products deriving from the translation of the human protein-coding genes[64-66]. Beside this, in 2003, the HPA project started with the aim to produce antibodies specific for all the human proteins and to create an "atlas" of the human proteome across cells, tissues and organs [67]. The atlas provides information about proteins expression and RNAseq data. Antibodies are produced based on a gene-centric approach. In-silico selected antigens, named protein epitope signature tag (PrEST), are used to immunize rabbits. Each PrEST consists of approximately 50-150 aminoacids and an affinity tag

(His_6-ABP), and has <60% homology to any other human protein-encoding gene. Also, all transmembrane domains of proteins are not considered[68, 69]. After selection, the target region in the gene is amplified from a total human cDNA pool, cloned and expressed in *Escherichia coli* as a recombinant fusion protein. The His-tag is used for the subsequent purification using affinity chromatography, while the albumin-binding protein (ABP) confers an increased solubility and immunogenicity[70, 71]. The purified PrEST sequence is verified by mass-spectrometry analysis, and used to immunize rabbits and purify the generated antibody[72]. PrESTs are also printed onto planar arrays in randomly selected sets each containing 384 antigens. Planar antigen arrays are then used to verify each antibody target and off-target binding, to determine its specificity.

The antibodies passing the criteria of specificity on PrEST array are then further analyzed in western blot using human cell line lysates, depleted plasma, as well as tissue lysates[73]. Only the antibodies confirming the presence of bands at the right molecular weight on western blot analysis are then utilized for protein expression profiling on human tissue microarrays (TMA) and human cell lines by immunohistochemistry (IHC)[73, 74]. Subsequently, the subcellular localization is evaluated on human cell lines by immunofluorescence (IF) and confocal microscopy[75]. More recently, these data have been supplemented with RNA-seq data from human organs and tissues [76]. All the mass-spectrometry, western blot images, IHC and IF images, and RNA-seq data are public available on the HPA portal (http://www.proteinatlas.org/). The current version 15 of the HPA has been released in April 2016 and counts 25,039 antibodies targeting 17,005 unique human proteins, corresponding to 86% of the human protein-coding genes.

Antigen arrays are systematically generated within the HPA, and represent also an invaluable tool to explore the autoantibody repertoire in body fluids with the aim to identify potential biomarker candidates [77].

HPA represents a standalone resource for the scientific community, enabling researchers from all over world to access to a wide range of data and reagents towards their proteins of interest, cell lines or tissue, complementing the mass-spectrometry based proteome maps[78].

3.2. Protein Arrays

Historically, IHC was the first affinity-based technique to be introduced in the study of proteins expression. Later on, two-dimensional gel electrophoresis (2D-GE) was developed and allowed the study protein expression and activation status, but also biomarkers discovery. These approaches are based on the classical 'one binder, one sample' setup. In the last 15 years affinity-based assays have been considerably expanded and adapted for proteome-wide scale studies, thanks to the introduction of miniaturized platform named *arrays,* which allow high-throughput and multiplex analyses. Protein microarrays derive from the DNA array format as introduced by Schena and colleagues in 1995[79], and later adapted to proteins. Over the past decade, protein microarrays greatly contributed to advances in proteomics and systems biology. They have been applied in numerous studies for uncovering the molecular bases of normal and pathological conditions, including protein-protein interaction, protein–DNA interactions, PTMs, signaling cascades, and pathogen–host interactions[80]. The great advantage of protein microarrays is their flexibility, which allows modulating the throughput and multiplexing. Several different formats of protein arrays exist[81]. The main distinction is based on the binder-sample setting. *Forward phase protein array* format (FFPA) is characterized by affinity-binders immobilized on a solid surface. Samples are usually diluted in assay buffer and flooded over the array, where analytes are captured by the immobilized binders. This assay format is characterized by medium to high multiplexing possibility, and low to medium throughput. FFPAs allow to test the expression pattern of up to thousands proteins in individual complex samples (e.g., body fluids, tissue lysate, cell line lysate) on planar format, or hundreds of proteins in hundreds of samples on bead format (e.g., 500 proteins in 384 samples using Luminex technology[82]). FFPA can be sub-classified based on the capture molecule immobilized on the solid surface (e.g., antibodies, antigens). Several different surfaces are available and allow the immobilization of the affinity binder while considering the conformational and functional characteristics that are needed for the analysis[83, 84]. Surfaces exist to guarantee the 3D conformation of antigens or proteins, or for directional immobilization, or adsorption. Once the analyte is captured, a detectable

signal is generated either by direct labeling of the sample with biotin and addition of streptavidin-phicoeritrine, or using a detection antibody conjugated to a fluorescent molecule (e.g. sandwich assay). Inversely to FPPA, *Reverse phase protein array* format (RPPA) is based on the immobilization of samples onto the array surface, and the following addition of capture antibodies in solution to test the expression or activation of specific proteins. This format allows to test up to thousands samples (high-throughput) for few targets (low to medium multiplexing).

When fluorescence is used as detection system, generated signals are measured by laser-based systems in both FPPA and RPPA formats. Dedicated software are able to quantify signals and convert them into numbers, generating output matrices that are used for data analysis. Genepix, Microvigene and xPONENT Luminex are the 3 used for the work here presented. Data analysis was run using several data analysis platforms, such as R, Graphpad and Jump.

A third type of array is *functional array*. Functional arrays are constructed with purified proteins, to enable testing various protein biochemical properties, such as interactions with proteins, nucleic acids, lipids, drugs, peptides, binding activities, and enzyme-substrate relationships interactions.

The following paragraphs describe principle and applications of antigen and peptide array, and RPPA.

3.2.1 Antigen and Peptide arrays

Antigen and peptide arrays are FFPA subtypes, where protein fragments or short peptides are immobilized on a solid surface, which can be planar or a micro-bead.

Antigen array technology is very useful for profiling the autoantibody repertoire in body fluids. Part of the results presented in this thesis (Paper II) were generated using antigen planar arrays routinely generated within the HPA project for the antibody validation process [72, 85]. Human protein fragments with average length of 100 aminoacids are immobilized onto a planar glass surface by means of non-contact printers [86]. Droplets

of 50-500 pl of diluted antigen are needed for generating spots sizes of 100-300 mm [87]. Planar platforms room tens-of-thousands of spots in each single slide, allowing a high multiplexing capability. Signal generation relies on fluorescence-based systems, where two secondary antibodies coupled to fluorofores are used. One antibody targets the protein fragment His-tag, and the resulting signal is used as quality control for antigen presence and spot shape. The second fluorescence conjugated antibody is an anti-human IgG, which generate a signal when a human antibody recognize one of the antigen printed onto the array. Signals are detected by means of a dual channel fluorescence scanner system (Agilent G2565 array scanner). The generated image is analyzed using dedicated soft wares (e.g., Genepix) allowing signal quantification. Within the HPA project, three different configurations of planar antigen arrays are available: 384 antigens printed in 21 sub-arrays (21 copies of the same 384 antigens collection), 11520 antigens in 2 sub-arrays, or 21120 antigens in 1 sub-array.

The second available platform for antigen arrays is bead-based, called *suspension bead array*. Protein fragments are immobilized on Magplex microspheres (Luminex), measuring 5.6 micrometer diameter and embedded in superparamagnetic particles [88]. Magnetic beads are very useful for assay automation, allowing automated plate washing and transfer, improving standardization and minimizing human errors. Microspheres are functionalized with carboxyl groups, which allows for covalent immobilization of the protein fragments via their primary amines. Luminex beads are color coded; 500 distinguishable beads identities (IDs) are obtained by mixing 3 dyes in different proportions. Each ID carries a different antigen type, and IDs can be mixed to form an array. A signal is generated by using a fluorescence labeled anti-human IgG antibody. A dual laser system integrated on a flow cytometry system (Flexmap 3D) is then used to measure the fluorescence intensity and identify bead identity. This system allows for semiquantitative methods.

Bead-based arrays are very flexible, offering the possibility to fully customize the array for each experiment. Another advantage is the possibility to avoid laborious image analysis process. However, the multiplex capability of bead-arrays is limited to the commercially available bead IDs, while the multiplexing capability of planar array is much higher. On the other hands, bead array allows a higher throughput compared to planar.

Planar antigen arrays can be considered as suitable for discovery studies and broad screenings, and antigens selected on this phase can be coupled to beads to create a targeted array and verify planar array results on a higher number of samples. This experimental design has been applied to study autoimmunity and to identify autoantibodies biomarkers in different type of pathology and body fluids [77, 89].

The first applications of antigen arrays for studying autoantibodies are dated at beginning of 2000, when the first arrays were created to test already known autoantibodies [90, 91]. After these proof-of-concept applications, antigen arrays have been applied to several hypothesis-driven and hypothesis-free studies in the context of autoimmune diseases, such as rheumatic diseases[92] and lupus erythematosus[93], but also in multiple sclerosis[94], Alzheimer's disease[95], and several types of cancer[96].

Limit of this type of array is represented by the fact that antigens represent only part of a protein, causing loss of information when the epitope is sited on a protein region not covered by the antigen, or when the epitope is conformational. Yet, the wide collection of antigens available within the HPA project represents a unique resource for studying autoimmunity.

However, antigens reactivity does not give any precise information about the epitope, which can be linear or conformational, and varies a lot in dimension[97]. When studying autoantibodies as biomarkers or for the development of cancer-specific vaccines, it is essential to fine map the reactive epitope and define the specificity of such reactivity[98]. Autoantibodies can be directed against intracellular or extracellular autoantigens, and this has been noticed to relate differently with pathogenicity[99]. In central nervous system related diseases, autoantibodies targeting intracellular antigens are believed to be biomarkers for an underlying tumor[100].

Epitope mapping methodology involves the use of soluble and immobilized overlapping peptide libraries, often in an array format [101]. Both planar and bead based formats are suitable, however bead-arrays are the method of choice when the mapping is related to one or a handful of short sequences (e.g., antigen). This approach can also be combined with alanine scanning[102], in which alanine substitutions are introduced into the synthetic peptides, allowing the study of the contribution of each single aminoacid in epitope recognition. An important application for epitope mapping is the localization of

epitopes for commercially available antibodies. This helps in defining which antibody to use for which application; when an assay is based on denaturated proteins, antibodies detecting linear epitopes are useful, while conformational epitopes are available only for proteins in native conformation. Maier and colleagues [103] published a high-throughput epitope-mapping of the vitamin-D receptor by using a recombinant peptide library consisting of 2304 overlapping peptides. More recently, Forsström and colleagues[104] described, for the first time, a peptide array based on parallel *in situ* photolithic synthesis of a total of 2.1 million overlapping peptides and covering all human proteins for the analysis of antibody specificity. A peptide bead-array can be generated by pre-coating beads with neutravidine, and affinity coupling biotin-labeled peptides. The signal detection method and read-out are the same as reported for antigen array. Also in this case, a limit of this mapping technology is the application only for mapping linear epitopes.

3.2.2 Reverse Phase Protein Microarray

RPPA is a sensitive, quantitative and robust technology allowing the analysis of protein expression and PTMs such as phosphorylation, glycosylation, ubiquitination and cleavage.

RPPA derives from miniaturized immunoassays[105], and the term "reverse" is used to distinguish it from FFPA. Contrary to FFPA, in RPPA is the sample to be immobilized onto the array. RPPA was described for the first time in 2001 by Paweletz and colleagues[106] as a powerful technology characterized by high-throughput and multiplexing possibility. Each array can host thousands of spots, depending on the desired dimension that can be set by using different array systems.

The optimal substrate for RPPA is a nitrocellulose coated glass slide. Nitrocellulose has high affinity for proteins, does not change their conformation, and gives low background[107]. RPPA samples are first denatured and then printed on the slide support by using a solid pin-based contact arrayer. Each pin has a flat end, and dips in a different sample loaded into a multiwell microplate. The sample is subsequently transferred onto the

nitrocellulose surface by contact. Several different types of pins are available with different diameters[108]. Only few picograms of total protein are needed for each array, allowing to print up to hundreds of arrays with just a few microliters of sample. The total number of samples printed on an array depends on number of replicates, sample dilutions, controls, and calibrators included in a specific experimental design (Fig. 3) [109, 110].

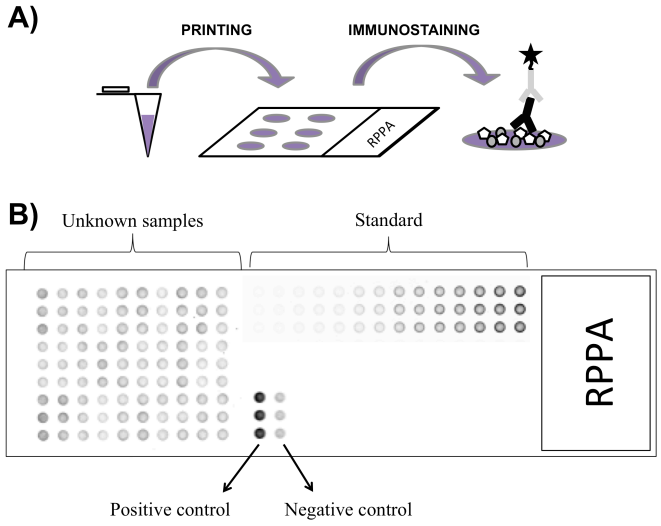

Figure 3. Reverse Phase Protein Array. A) The scheme represents the RPPA workflow, constituted by two main phases: the printing phase of sample onto nitrocellulose coated glass slides, and the immunostaining of the array with a primary antibody and a conjugated secondary antibody. B) Representation of how an RPPA slide appears after signal detection. In RPPA positive controls, negative controls and standard curves are always present together with unknown samples. All samples and standards are spotted in replicates (e.g., triplicates).

Several different types of samples can be tested using RPPA, such as whole tissue lysates, enriched tissue lysates (e.g., microdissected tissues) [111-113], cell lysates[114, 115], serum and plasma[116], and CSF [117]. Each RPPA spot represents the whole cellular proteome; therefore, the selection of highly specific detection antibodies is essential. RPPA antibodies validation is based on western blot, and only antibodies generating a single band at the right molecular weight are considered as validated [50]. Known positive and negative controls (e.g., cell or tissue lysates) for expression or activation of a certain

protein are tested in parallel to confirm the antibody specificity for the target.

Each array is probed with a detection antibody followed by a secondary antibody conjugated with enzymes. High sensitivity is ensured by using signal amplification systems based either on colorimetric or fluorimetric detection[118].

RPPA allows comparing the pattern of expression and/or activation of a certain protein across all samples tested in the same exact condition. A laser scanner and a dedicated software (e.g. MicroVigene, Vigenetech) are used to reveal signal intensity for each spot. In parallel to technological advancement, bioinformatics tools have been developed for data quality assurance, normalization and quantification[119]. Qualitative evaluation of each image and spot is manually performed by the operator. The analyte concentration is proportional to signal intensity after subtraction of background and negative control signal, and normalization to total protein amount. The normalized data can be used to run a wide range of analysis.

Each array reports data about expression and/or activation of a single target. Thanks to automatized immunostaining systems, several different arrays can be stained in parallel with different antibodies, allowing rebuilding the functional status of entire pathways driving cellular functions such as growth, survival and migration. Clarifying the entire cascade of events that leads to the activation and deactivation of proteins is the first step in identifying alterations in cell homeostasis that will drive the carcinogenesis processes. The great sensitivity of RPPA is a unique advantage when analyzing samples limited in their amount, such as needle biopsies in clinics. For this reason RPPA found applications in cancer-targeted therapy, but also in the assessment of treatment for other type of diseases, as well as for diagnostic and prognostic biomarker identification.

When analyzing tissue samples, the coupling of RPPA with laser capture microdissection (LCM) showed to be very powerful. LCM allows capturing pure (e.g., 95%) cellular populations that can be separately analyzed on RPPA for their molecular status. The variable composition of tissue samples can generate data misinterpretation when analyzing whole tissue lysates. LCM showed an invaluable tool for upfront enrichment of homogeneous cellular populations[111]. The results presented in this thesis (Paper I) have been generated by coupling LCM to RPPA to map the stroma-epithelium cross-talk.

RPPA can be calibrated including standard curves and internal positive and negative controls, allowing inter-experimental comparison and the generation of quantitative data[120]. RPPA showed to be particularly suitable for phosphoproteins measurements, representing a unique tool for rebuilding the activation status of key pathways in cancer development and progression[121-123], and response to treatment[124, 125]. The great effort in technical optimization allowed RPPA to find application in several clinical trials. Among these are the I-SPY1 and I-SPY2 clinical trials for the identification of new biomarkers and optimization of personalized treatment in breast cancer. I-SPY1 identified high levels of HER2 Y1248 in a subgroup of patients in absence of HER2 overexpression. This group showed also hyperactivation of HER3 and EGFR, indicating that these patients could get advantage from HER-targeted therapies[126, 127]. RPPA is also applied within the Side-Out´s Breast Cancer clinical trial, to find potential targets and select individualized treatments for patients with previously treated metastatic breast cancer. At the date when this thesis was written, the Phase III of this clinical trial was just launched[128].

4. Protein Biomarkers in Prostate Cancer

As above mentioned, the discovery of PSA revolutionized PCA diagnostics[129]. Since the introduction of the serum PSA test in clinical practice, the diagnosis of PCA has become more frequent. However, 50–75% of patients with elevated PSA level have a negative biopsy result for PCA[130]. Due to this low PSA specificity, several studies aimed to identify novel biomarkers for PCA.

High throughput proteomic platforms might be very promising in identifying and quantifying new biomarkers for PCA in body fluids and tissue[131]. The direct analysis of tissue might allow to get the molecular picture of PCA, and to uncover the processes leading to its genesis and progression. On the other hand, and as previously discussed, PCA diagnosis and treatment would highly benefit from a screening test procedure based on biological fluid, due to the minimally invasive, rapid and low-cost procedures needed to collect clinical samples. Beside plasma and serum samples, also urine, prostatic secretion and seminal plasma have been and keep being extensively investigated[132-134]. The aim of this paragraph is to summarize the main findings in PCA biomarkers discovery and validation.

4.1 Protein Biomarkers in Tissue

The multistep process leading to PCA through PIA and PIN pre-malignant states has been partially characterized and some common mutations were identified. Among these, PIA showed upregulation of Bcl-2 and GSTP1, and down-regulation of PTEN and CDKN1B. The progression to PIN is characterized by ETS transcription factor dysregulation and by the well-known TMPRSS2-ERG gene fusion[135, 136]. Later, PTEN deletion and RB1 loss represent key events at early phases of PCA oncogenesis [5]. Androgens (e.g., testosterone, androstenedione and dehydroepiandrosterone) regulate PCA growth by both stimulating cell proliferation and inhibiting apoptosis. They are released into the blood flow and circulate bound to albumin or sex-hormone binding protein

(SHBP). Thanks to their hydrophobicity, androgens enter prostate cells, where they are processed to the active form (e.g., dihydrotestosterone). Active androgens bind to androgen receptor (AR) and cause its conformational change and activation by phosphorylation and homodimerization. Activated AR is then transferred into the nucleus where binds to promoters and starts transcription of cell growth and survival promoting genes, and of PSA[137]. Therefore, the AR is the main candidate target for PCA treatment. More than 600 mutations have been reported for AR gene alone and many of them influence AR interactome and signaling transduction[138, 139]. Well-differentiated low-grade PCA contains glandular structures with tumor cells expressing known AR and PSA, while poorly differentiated late-grade PCAs are lacking in glandular structures and show more profound changes at a molecular level[2]. AR signaling cascade is therefore not sufficient to sustain tumor growth and progression, especially in late stage PCA where the tumor becomes androgen independent. Paracrine mediators secreted by stroma, such as IGF, FGF and EGF, are responsible for tumor epithelium-stroma cross-talk and causes the activation of other receptor tyrosine-kinases (RTKs) (e.g. IGFR, FGFR and EGFR) which together with AR sustain the PCA cells survival and proliferation at late stages of the disease[140, 141]. Beside this, at late PCA stages AR can be activated through alternative pathways and in absence of androgens[141]. The ligand-independent AR activation could take place through the phosphorylation induced by RTKs-activated AKT or MAPK [140], the protein kinase A signaling [142], or by the AR-STAT3 association induced by IL-6 [143]. All these molecules involved in different stages of PCA progression are potential targets for novel personalized therapies.

Tissue samples collected at all stages of PCA, from normal tissue to metastasis, have been profiled with the aim to identify new biomarkers. Already in 2001, Paweletz and collaborators demonstrated that SELDI-MS coupled to LCM enriched tissue samples are useful tools to analyze PCA tissue samples and to identify different molecular profiles from different cellular compartments (e.g., normal, tumor, invasion front)[144]. Later, the same technology combination showed that PCA-24 as an hypothetical biomarker able to distinguish PCA from normal tissue and BPH with high specificity and sensitivity[145]. Other studies reported signatures able to separately cluster BPH from PCA. Disulfide-isomerase, 14-3-3-protein, enoyl CoA-hydrase, prohibitin and B-tubulin β-2 showed to be

increased in PCA, while Keratin-II, desmin, HSP71, ATP-synthase-β-chain and creatine kinase- β-chain were increased in BPH[146]. Quantitative mass-spectrometry identified periostin to be up-regulated in PCA compared to BPH[147]. P53, c-MYC, AR and PSA, resulted to be the central hubs all known to be key regulators for PCa onset and progression and potential targets for therapy[148].

High-grade prostatic intraepithelial neoplasia (HGPIN) is a precursor of PCA. From a morphological point of view HGPIN is similar to PCA, but it retains a basal-cell layer[149]. Proteomic profiling of LCM enriched tissue identified GDF15 protein expression in 70% of PCA, 38% of HGPIN and absent in normal tissue, suggesting its role as a marker of early carcinogenesis[150]. Basal cells have also been specifically profiled since they are thought to play an important role in prostate carcinogenesis process, by being the physical barrier for PCA invasion of the adjacent stromal compartment[151, 152]. Retinoic acid-binding protein 2 was found to be down-regulated in basal cells of benign prostate compared with PCA and prostatitis[153].

As above mentioned in detail, depending on the grade of invasion and cell morphology, PCA is classified based on Gleason scores. Matched biopsies collected from BPI, Gleason 6 PCA and 23 Gleason 8+ PCA, have been analyzed using LCM coupled to mass-spectrometry, identifying lamin A as an hypothetical biomarker of differentiation and prognosis[154]. The higher the Gleason score, the higher is the grade of tissue invasion. EPLIN down-regulation was found to promote epithelial-mesenchymal transition (EMT) in PCA, and to correlate with metastases[155]. Previous studies report potential biomarkers for the detection of lymph-node metastasis (LNM). Collapsin response mediator protein-4 (CRPM4) was found to be a LNM suppressor[156]. Also a 6 protein panel (e.g., e-FABP5, MCCC2, PPA2, Ezrin, SLP2, and SM2) was found to be differentially expressed between PCA with and without LNM[157].

Variations at protein expression and activation can be due to genetic and epigenetic mutations. Several miRNAs participate in proteome modulation during PCA genesis and progression[158, 159]. MiRNAs are known to be involved in the regulation of AR signaling as well as p27[160, 161].

4.2 Protein Biomarkers in Body Fluids

Tumor invasion causes the destruction of tissue architecture and the release of proteins, first into the interstitial fluid, collected into the lymphatic system and finally into blood or other biological fluids, such as seminal fluid and urine. In normal conditions, PSA is expressed by prostatic epithelial cells and then secreted into the lumen of the gland as a component of seminal plasma (0.5–3.0 g/l). In PCA, the anatomic barrier between glands and capillary is disrupted causing PSA concentration in blood to raise from 0.5–2.0 ng/ml to 4–10 ng/ml in early stages of cancer, and up to 1 μg/ml in the late stages [48]. Beside influencing intracellular proteins, pathological conditions can also cause variation in proteins secretion, shedding of transmembrane proteins or part-of-proteins (e.g. extracellular domains), and variation in the anti-tumor response (e.g. C-reactive protein, immunoglobulins and chemokines).

Several studies focus on new candidate PCA biomarkers in biological fluids. A proteome profiling using MS of voided urine from PCA patients and individuals with negative biopsy allowed to identify a polypeptide panel confirming the presence of prostatic secretion, which can be the source of relevant prostate biomarkers in urine. These polypeptides resulted from Collagen α-1 (III) [642–659], Collagen α-1 (I) [699–725], and Psoriasis susceptibility 1 candidate gene 2 protein[162] (SPR1). This panel was validated on a blinded set of 213 samples (118 PCA and 95 negative biopsies), showing 89% sensitivity and 51% specificity. When age and percentage of free PSA were added to the signature, sensitivity reached 91% and specificity 69%[163]. The panel was then evaluated on routine clinical application, showing sensitivity 86% and specificity 59%[164].

A profiling of PCA serum samples performed by antibody microarrays for 180 different proteins identified five potential biomarkers (von Willebrand Factor, immunoglobulin M, Alpha1-antichymotrypsin, Villin and immunoglobulin G)[165].

Most of candidates biomarkers identified still need to be validated on large clinical cohorts; nevertheless the fact that many have been identified in different studies with different methods over the time increases the confidence in their ultimate utility. Among these are Beta 2 microglobulin [166-168], Zinc alpha2-glycoprotein (ZAG) [169, 170], Transforming growth factor-β1[171, 172], interleukin 6[173], CD90[174], Engrailed-2 (EN2) [175],

Fibronctin 1[168] and many other soluble factors and intracellular proteins involved in structural or metabolic functions.

Engrailed-2 (EN2) is a transcription factor expressed in PCA but not in normal prostate tissues, and it was detected in urine by western blot and ELISA. Levels of EN2 were 10.4-fold higher in PCA compared to controls. It showed sensitivity of 66% and a specificity of 88.2%, with AUC of 0.81 [175].

Beta-2-microglobulin (β2M) was reported as a potential marker of PCA aggressiveness. High levels of β2M were measured both in the serum[167] and prostatic secretion[176] of patients with metastatic PCA, and directly correlates with aggressive pathologic features in primary PCa specimens[166]. As a confirm, β2M was demonstrated to promote the epithelial-mesenchymal transition (EMT)[177].

Another candidate diagnostic biomarker for PCA is zinc alpha2-glycoprotein (ZAG). ZAG is a secreted protein, responsible for lipid degradation in adipocytes, expressed in blood, seminal fluid, urine, sweat and saliva. High levels of ZAG in seminal plasma are reported to be increased in PCA compared to healthy donors[178]. More recently, ZAG was reported to correlate also with PCA grade, with 1.3-fold increase in serum of Gleason score 7 PCA compared to Gleason score 5 [179]. This result was confirmed with ELISA in a larger cohort of samples, while IHC analysis demonstrated an inverse relationship; the possible explanation is the existence of two isoforms of ZAG in blood and seminal plasma.

PTEN loss is associated with poor prognosis in metastatic androgen independent PCA. In order to identify new prognostic factors, 79 proteins have been measured in the sera of 57 patients with meta- static cancer under hormone ablative therapy. Target proteins were chosen based on their abundance after PTEN loss in *in vivo* models. A combination of ELISA and SRM methods was used. A panel of predictors composed of five proteins was determined: THBS1, CRP, PVRL1, MME, and EFNA5. The panel showed an AUC of 0.96 and 0.94 for cumulative hazard ratios at 12- and 24-months[168, 180].

An interesting comprehensive study diagnostic and prognostic biomarkers to differentiate indolent from aggressive PCA was performed using 4 pools of sera respectively from BPH, localized non-progressing cancer, localized progressing cancer and metastatic cancer[168]. After immunodepletion of the 14 most abundant serum proteins the

pools were analyzed by iTRAQ. A hierarchical clustering data analysis showed a high similarity between BPH and non-progressive cancer profiles, while the metastatic group clustered separately from all the rest. An increased level of eukaryotic translation elongation factor 1 alpha 1 (eEF1A1) was observed in progressing and metastatic cancer patients compared to BPH. Afamin and fibronectin were classified as potential diagnostic biomarkers for low-grade cancer, since a 1.4 fold higher level was observed in the non-progressing cancer compared to the BPH. This study detected many already known PCA candidates biomarkers such as CRP, alpha-2-macroglobulin, zinc-alpha-2-glycoprotein, beta-2-microglobulin, fibronectin, and ceruloplasmin [181].

Recently, exosomes appeared as a novel non-invasive source of cancer biomarkers[182]. Several type of tumor-specific molecules can be found in exosomes isolated from biological fluids. Proteomic profiling of endosomes from urine identified hundreds of proteins originating from several tissues, included from prostate [183-185]. A recent study based on MS analysis compared 16 preoperative urine samples from PCA patients with 15 healthy controls, detecting 246 differentially expressed proteins[186]. Among these, 17 proteins showed sensitivity >60% at 100% specificity. TM256 protein had the highest sensitivity (94%) and when combined with a panel of other biomarkers, resulted to be able to fully differentiate PCA from healthy individuals. Another work based on immunoassays on exosomes extracted from urine to validate MS data, showed that flotillin 2, TMEM256, Rab3B and LAMTOR1 proteins have higher levels in PCA patients compared to healthy males. Flotillin 2 receiver operating characteristic curve showed AUC of 0.91, 88% sensitivity and 94% specificity[187].

Although all these findings require to be confirmed in larger cohorts of patients, these results support the feasibility of identifying highly sensitive and specific PCA biomarkers on urinary endosomes.

4.3 Autoimmunity as a source of biomarkers for PCA

Autoimmunity is the complex of immune responses developed by an organism toward its own proteins not recognized as *self*. Chemical exposures, infections and genetic

alterations are known causes of autoimmune diseases[188]. Cancer can also represent a cause of autoimmune response. Autoantibodies in cancer are thought to spontaneously develop due to antigens leakage from tumor tissue, production of new protein variants (neoantigens), or presence of proteins in the bloodstream in association with exosomes and microvescicles. Several studies confirm the usefulness of autoantibodies as biomarkers for diagnosis and prognosis of several different types of cancer[46]. Data about autoantibodies autoantibodies useful as biomarkers is still quite disperse in literature, but the interest on their application as biomarker candidates leaded to some effort in generating databases for the collection of all published autoantigens. One such database is AAgAtlas 1.0[189].

Autoimmunity has a polyclonal origin, where self-antigen presentation process is thought to be the same as for pathogenic antigens, with the immune system presenting the antigen as multiple and overlapping epitopes, each recognized by a different clone of lymphocytes[190]. The production of antibodies targeting TAAs represents a sort of "biological amplification" of protein biomarkers present at very low concentration in blood, even though sometimes the reactivity is not directly related to target concentration [191].

PCA is regarded as an immunogenic tumor. Literature reports quite some examples of autoantibodies useful in PCA diagnostic and prognostic process. An autoimmune signature able to distinguish PCA from controls with high specificity and sensitivity was published by Wang and colleagues already in the 2000[192]. Later on, another signature of 174 antigens representing proteins related to cytoskeleton, nucleus and RNA-associated resulted to be reactive exclusively in PCA serum samples, while no reactivity was identified in controls[191]. Among the autoantibodies reported to be useful in PCA diagnosis are those toward cyclin B1[193]. Furthermore, autoantibodies have been reported to be useful also in defining PCA prognosis. The reactivity toward PSA, Her2[194] and Fetuin-A[195] is enhanced in late PCA stages compared to early stages. Moreover, the identification of autoantibodies in PCA correlates to treatment failure after androgen deprivation and radiation therapy[196]. Taken together these data suggest autoantibodies in PCA should be further investigated as a potential source of biomarkers for both PCA diagnosis and aggressiveness prediction. However, particular attention is needed in characterizing the

specificity of such reactivity, since epitope are targets of cross-reacting autoantibodies or polyreactive antibodies[197-199].

5. Challenges in biomarkers discovery and validation

Despite the wide number of hypothetical biomarkers identified by the application of proteomic technologies, only very few have been validated and implemented in the clinical practice. As George Poste stated in 2011: "The dismal patchwork of fragmented research on disease-associated biomarkers should be replaced by a coordinated 'big science' approach"[200].

The process to develop and validate a biomarker is hard, with only 3-5% of the biomarkers are transferred to the clinical practice[201-203], and lab to lab variability present an important challenge in the process. Standardized protocols and an appropriate training are needed to reduce such variability. Furthermore, the choice of the right sample set for discovery and validation, the storage protocols, the samples complexity and limited amount, the technology limitations, but also the statistical analysis approach are some of the main drawbacks to the biomarkers validation[51]. All these variables may be added to the already inconstant nature of sample composition among individuals, resulting in data misinterpretation. The following sections will discuss the more relevant sources of variation in a protein biomarker identification process.

5.1 Study design

The study design is the first critical aspect a researcher needs to reflect on when planning a study for biomarkers identification. A deep knowledge of biology, analytical platforms, biochemical properties and statistics is needed to provide reliable results out of a biomarkers study. This process requires an interdisciplinary team constituted by clinicians, investigators and statisticians.

A clear biological question needs to be set and samples have to be selected accordingly, and avoiding any bias. Within the biomarkers research community there is widespread agreement that retrospective studies are less appropriate than randomized prospective studies, for which the hypotheses and analyses to be carried out are known in

advance[204]. However, under the appropriate circumstances, restrospective studies can also be very informative. One such case is the introduction of KRAS mutational screening in the decision making process for metastatic colorectal cancer treatment[205].

Untargeted screenings are useful when the research goal is to identify proteins or other molecules that are not already reported to be involved in a certain disease. Targeted analysis are more advisable for verification phase.

Statistics is helpful in determining the number of samples necessary to reach a good statistical power, allowing to make valid conclusions out of the analysis results[206, 207]. When the research aim is to identify differentially expressed or activated proteins in different conditions, samples need to be matched for relevant variables such as demographic variables (e.g., age and gender) to be comparable. An appropriate assay needs to be chosen and the samples collection, storage and preparation have to be compatible with it. Possibly, the process should count for a good automation level, to avoid human errors. Samples should be randomized to avoid any bias due to sample position or time-delay in the analysis. Including replicates in the analysis is also essential to guarantee statistical power and evaluate technical variation. Replicates are intended on different levels and the number of replicates depends on sample availability and on the assay platform used for the analysis. More analytical repetitions of the same sample can be included in an experimental run, or the entire experiment can be repeated on a different time with the same or modified layout.

Beside this, every analytical platform and assay have limits that inevitably affect the output data, and therefore need to be considered during the study design. The study design process is therefore complex and have the aim to face and give answer to a research question, by predicting and minimizing all the possible variables that could affect the results. The thoroughness in this phase is an essential requirement to guarantee the success of the research.

5.2 Pre-analytical variables

Biological samples are very complex matrices, which typically contain many

components, such as cells, proteins, nucleic acids, lipids, and metabolites. Moreover, the inter-individual heterogeneity – characterized by genetic diversity – makes the scenario very complex, challenging even more the identification of new biomarkers. If this intrinsic variation cannot be avoided, at least the procedures followed for samples collection, handling and storage need to be standardized across the different laboratories, and guidelines need to be established for high-quality translational research[200].

Guidelines change depending on the sample type, and on the type of assay. The mechanical process for collection causes a trauma to the sample and its components. Indeed, the cells in a tissue biopsy are alive and adapt to the trauma of excision, change of temperature, ischemia, and hypoxia[208, 209]. Such events induce modifications in pathways involved in the response to stress, in DNA repair system, and in protein phosphorylation/dephosphorylation in response to stress[210]. Biopsies can be collected intra-surgery or as image-directed needle biopsies. It is important to stop all cellular processes and stabilize proteins and PTMs in tissue samples to avoid imbalances of tissue molecular structure. The gold standard for preservation is represented by formalin-fixed and paraffin-embedded tissue (FFPE), which is a widely used method for tissue preparation for IHC analysis[211]. However, the slow process of tissue permeation by formalin causes a significant degradation of both proteins and nucleic acids[212]. To avoid these degradation processes, new types of fixative were introduced that optimize the protein and PTMs preservation and maintain the tissue morphology; among these are optimal cutting temperature (OCT) compound[210] and the one-step biomarker and histology preservative (BHP)[213]. Tissue fixation with OCT and BHP is followed by snap-freezing in liquid nitrogen. The delay between tissue excision, pathologic examination and sample storage is a major source of variation causing changes in proteome composition. The entire procedure requires from 30 minutes to several hours depending on organizational factors (e.g., laboratories and surgery rooms location, availability of a pathologist, and the number of concurrent cases), and during this time temperature fluctuation occurs that causes variations in the molecular network. The specimen size also represents a source of variation. In fact, bigger samples require longer time to allow the fixative to penetrate the tissue in depth and block the degradation processes[214].

When analyzing body fluids, other variables should be considered. For blood and

its derivatives (e.g., plasma and serum), the delay time between sample centrifugation and freezing does not affect the majority of the proteins[215, 216]; but for some, such as interleukines, chemokines and growth factors there can be a substantial variation[217]. Also, the composition of serum and plasma, the used collection tubes, and anti-coagulants can cause alterations in proteome composition of the sample[218, 219].

Other studies were developed to characterize the stability of stored samples, showing that repetitive freeze-thawing cycles influence the proteome composition more than long-term storage at -70 C[220, 221]. Freeze-thawing cycles are critical for the stability of the great majority of proteins, but seem not to affect much immunoglobulins stability.

It appears clear that standardized procedures for samples collection and storage are essential to guarantee high quality data production for the identifications of biomarkers.

5.3 Analytical variables

The outcome of an analysis that aims to detect and measure proteins that can possibly be useful as biomarkers strongly depends on sample preparation, and on the assay limits in terms of sensitivity and specificity.

Biomarkers are usually detected at low concentration, and identified in very complex matrixes. For these reasons the detection methods have to be high specific and sensitive, while the sample needs to undergo to a preparation that can be a further source of variation.

Tissues are constituted by many different cell types (e.g., fibroblasts, nerve cells, endothelial cells, infiltrating lymphocytes, epithelial cells, etc) that cross-talk with each other, and coordinate to sustain tumor growth and proliferation. Tissue heterogeneity represents a significant obstacle for identifying protein biomarkers specific for malignant cells. Indeed, different individuals with the same type of cancer, and even different biopsies of the same tumor, can have a peculiar distribution and different amount of each cell type. For this reason, upfront cellular enrichment is often needed to disassemble this complex tissue ecosystem and analyze the signaling network in different cell subpopulations separately[111, 123]. Over the years, many laboratories have coupled molecular

analyses with upfront LCM or cell sorting[222-224]. Another source of variation when working with tissue is the protein extraction process. Protein degradation or changes can occur during tissue lysis. It is therefore advisable to inhibit proteases, phosphatases and kinases at all phases of the analysis, from tissue collection to lysis. A variety of proteases and phosphatases inhibitors are commercially available, and can be applied to tissue lysates or tissue slides during LCM process, to stabilize the proteome.

On the other hand, the wide concentration range of proteins in blood, which spans up to 12 orders of magnitude, present a critical problem in the identification of blood biomarkers. [225]. Biomarkers in blood derive from tissues, pass through the endothelial barrier, and reach the circulation where they are diluted in a high amount of bio-fluid. To be able to detect low concentrated proteins, samples undergo to preparation such as depletion, fractionation, enrichment, enzymatic digestion or heating[226]. These can cause sample alteration in composition, and the analysis is further complicated by the fact that low molecular weight (LMW) proteins may exist in complexes with high molecular weight (HMW) and high abundance proteins such as albumin and immunoglobulins that are often discarded and removed *a priori* by depletion techniques[227, 228]. A possible way to overcome this issue is the introduction of hydrogel nanoparticles to enrich low abundant proteins [229, 230].

Sample handling and preparation helps in detection of proteins and can be different for different sample type and application. However, this might alter the protein composition in such extent that disease related profiles cannot be found. To reduce the impact of manual handling, new one-step sample preparation methods are developed[231].

Affinity-based methods strictly depend on the characteristics of the used binder. Antibodies need to have high sensitivity and specificity, and need to be validated. Antibodies are known to perform differently in different applications, therefore a result obtained with an array can frequently not being verified using other immunoassays (e.g., IHC). On the other hand, antigen-based arrays are also subjected to variation due to antigens conformation of different platforms. Antigens can have different length and can fold causing linear epitope inaccessibility or generation of conformational epitopes different from those of the native protein[232].

Finaly, another important source of variation is due to human error. Methods for biomarkers measurement should be highly robust, and standardized for dilution linearity, precision and intra- and inter-assay reproducibility. Analytically, in modern high-throughput laboratories automation is essential[233, 234].

5.4 Post-analytical variables

The post-analytical phases in the biomarkers evaluation process involves data analysis procedures, including normalization and interpretation. Post-analytical errors might affect biomarkers performance as much as pre-analytical and analytical variable do. Most of the times, variables in biology context are continuous, raising the problem to define cut-offs for clinical tests. Also, continuous data tend to be less reproducible than dichotomous parameters. For these reasons, a standardization is needed also for data analysis, and it can be reached only through a tight collaboration between clinicians, investigators and statisticians[51]. When working with –omics data it is necessary to be even more careful in data handling, due to the high amount of endpoints. Once generated, data undergo quality controls to inspect technical variation. One basic control is the evaluation of the coefficient of variation (CV) for the technical replicates included in the assay. Another step is the evaluation of data distribution using histograms, scatterplots or principal component analysis (PCA)[235]. When data is influenced by technical artifact, normalizations are needed. In general, normalization is used to adjust values so that they can be compared across the samples. The choice of a normalization method is critical because it can deeply affect the results of the analysis[236]. The most common methods for protein microarrays are probabilistic quotient normalization (PQN), variance stabilizing normalization, cyclic loess and robust linear model normalization [237-239]. Caution is also needed while removing outliers, since in many cases what it's defined as outlier could have a biological meaning. Furthermore, another issue is setting a clear cut-off for where outlier behavior starts. After data pre-processing, several statistical tools are available to compare protein expression or activation data deriving from microarray analysis between samples and different conditions. Among these are t-test, Wilcoxon rank sum test, linear models, fisher exact test, and others. Each of these tests has to be chosen based on intrinsic

properties of the data set, such as t-test, used only on normally distributed data. Statistical test measures the probability there is no difference between the compared groups. This measure is given by p-values. When p-value is small ($p<0.05$) there is high probability that the difference is real. When comparing microarray data, many statistical tests are run in parallel for each of the tested proteins. If the p-value cut-off for significance is 0.05, this means that when testing 1000 proteins by random chance alone we would expect 50 false positives. Bonferroni correction or false discovery rate (FDR) can be used to reduce false positives, but sometimes they can be too conservative and cause discard of true discoveries[240].

For each step within data analysis, a good visualization is helpful for result interpretation, especially with high-dimentional data. After normalization, histograms, boxplots and scatterplots can help for visualizing changes in data distribution. Unsupervised hierarchical clustering analysis and heatmaps can be used to explore patterns. After analysis of differential expression, volcano plots can show the significance and fold-change, and receiver operating characteristic (ROC) curve can visualize the relation between sensitivity and specificity of a classification.

There is a need for specifically designed database for protein microarray data archiving and sharing, and tailored standards for data processing and analyzing. Some examples are now available and likely will be implemented in the future, allowing for development of standard data processing pipeline[241].

6. Current investigation

As the title suggests, the work presented in this thesis is based on the profiling of PCA tissue and blood samples by using affinity proteomics platforms such as RPPA, antigen and peptide arrays. In the field of PCA research, the major lack is represented by the absence of biomarkers able to define the prognosis of patients. PCA diagnosis has shown substantial improvements in the last decade, while the wide spectrum of disease outcome that ranges from indolent disease to aggressive forms makes it urgent to introduce in the clinical practice new biomarkers to predict the patient prognosis and chose the best type of treatment. This will allow to early identify aggressive forms that need radical treatment, and at the same time to avoid heavy side-effects generated from overtreatment of patients with indolent PCA. The aim of the work here presented is to apply proteomic platforms, such as antigen planar and bead-based arrays, peptide bead arrays and RPPA, to characterize prostate cancer samples and to possibly identify new biomarker candidates in tissue and/or blood. As previously mentioned, proteins are the functional units of the cell, and the basis for communication between cells. Tissue cells can release proteins into the body fluid circuit by secretion or leakage, in response to stimulation or damage. Proteins are therefore the ideal biomarkers for several diseases, and also for PCA. In **Paper I** we wanted to unravel the mechanisms of communication that underlie PCA development and progression by disassembling tissue samples collected from PCA and adjacent normal tissue in its stromal and epithelial compartments using LCM technology, and profiling them separately using RPPA. Particular emphasis was given to the measurement of protein phosphorylation, to allow a picture of the activation status of important pathways expected to be involved in PCA driving processes. Tissue samples have the great advantage of giving a clear picture of the tumor molecular status. However, tissue samples can be obtained only through invasive procedures, therefore making them precious and limited. The identification of biomarkers in plasma/serum would allow avoiding such invasive procedures and providing higher amount of sample. A drawback in the use of plasma as a source of biomarkers is the low concentration of tissue proteins, requiring high sensitive methods to be detected. To some extent, cancer autoimmunity could be considered as a

natural amplification for low concentrated proteins in blood, even though the reactivity toward a certain protein is not always directly correlated to its expression levels or concentration. In **Paper II** we aimed to broadly screen plasma and serum samples of PCA patients and controls to study their autoimmune repertoire, and to identify autoimmune markers for the definition of PCA prognosis. To this purpose we applied antigen arrays from the HPA project collection, together with peptide arrays.

The most significant findings and the experimental design for each of the two presented papers are reported in the following sections. For further details we refer to original papers, which are not included in this thesis for copyright reasons. A discussion on faced challenges and perspectives is also present.

6.1 Paper I: Dissecting the PCA epithelium-stroma molecular architecture in tissue samples by using LCM and RPPA

This paragraph summarizes methods and results of a research study carried on by the author of this thesis' and collaborators. The research was published on Molecular Oncology in October 2016 as an original article that can be referred to for more details. Reference: Pin E, Stratton S, Belluco C. A pilot study exploring the molecular architecture of the tumor microenvironment in human prostate cancer using laser capture microdissection and reverse phase protein microarray. Mol Oncol. 2016. doi: 10.1016/j.molonc.2016.09.007. [Epub ahead of print]

Understanding the cross-talk between the tumor and stroma compartments in tissue is essential for identifying novel biomarkers to personalize and optimize the treatment of PCA patients, and in general for patients affected by solid cancers. This study tested the feasibility of using LCM and RPPA combination as a powerful workflow for unrevealing the signaling architecture of the tumor-stroma cross-talk by using PCA as a model. Biopsies from tumor and the adjacent normal-appearing tissue were collected from 18 PCA patients diagnosed as Gleason score 6-7. The choice of this subgroup of patients is justified by the urgency to define the prognosis of Gleason score 6-7 PCA patients.

Biopsies were embedded in OCT and snap-frozen within few minutes from collection to guarantee the best possible preservation of proteins and phosphorylations. After cutting the tissue in thin sections, the tumor epithelial cells, the normal-appearing epithelium, the stroma adjacent to the tumor epithelial cells, and the stroma surrounding the normal-appearing epithelium were isolated using LCM and separately profiled by RPPA platform. The choice to print stroma and epithelium samples onto different arrays and analyze them separately was due to the limited amount of microdissected material (<10000 cells per compartment), and to the fact that each compartment expresses specific proteins or showed activation of different pathways. For instance, the HER family proteins and their activation status were analyzed only in the epithelial compartment, while proteins such as Caveolin, the MMPs, and markers of immune activation like Zap70 Y319/Syk Y352 are known to be more involved in stromal activities, and were therefore analyzed only in this context. Beside these, proteins ubiquitous across different cell types, such as the members of the AKT-mTOR pathway, were analyzed both in the epithelium and stroma compartments. The expression or activation status of each target was compared between tumor and its normal appearing counterpart. Also, the correlation between targets was evaluated within and between compartments. This approach allowed the rebuilding of the molecular architecture of each compartment, and clarifying the interconnection between stromal and epithelial compartments.

All the endpoints found to be differentially expressed/activated between the normal and tumor compartments are summarized in Table 1. As expected, our results showed a decrease in the PTEN expression ($p=0.01$) and activation ($p=0.01$) in PCA epithelial compartment. Also, PTEN expression/activation showed a negative correlation with Akt S473 in both tumor and normal-appearing epithelium. The association was less strong in the malignant lesion, likely as a result of PTEN loss and other signaling derangements affecting the Akt pathway (Fig. 4). In agreement with the PTEN loss, we identified other events such as the overexpression of the anti-apoptotic protein Survivin and the suppression of Stat3 signaling through the modulation of the IL8-Stat3 signaling pathway. The tumor epithelium showed also an increase in B-Raf S445 and AR S650 phosphorylation compared to its normal-appearing counterpart, in agreement with tumor progression, and development of castration resistant PCA (Tab. 1).

Table 1. List of proteins/phosphoproteins that were statistically different between the tumor and normal-appearing tissue ($p < 0.05$). Panel A lists proteins that were statistically different between the tumor and normal-appearing epithelium. Panel B contains a list of proteins that were statistically different between the stroma surrounding the tumor and normal-appearing epithelium. For each analyte the trend in the tumor compartment compared to the normal counterpart, RPPA mean intensity values and 95% confidence interval (CI), percentage change of protein expression or activation in the tumor compared to the normal, the test performed, and the p-value are reported.

Protein	Trend in tumor	Tumor RPPA mean values	Normal RPPA mean values	% change in tumor tissue	95% CI	p-Value
Panel A: Epithelium compartment						
Androgen receptor S650	↑	24,513	10,134	+142%	4136/21,076	<0.01*
Acetyl-CoA carboxylase S79	↑	27,049	14,710	+83%	2612/23,973	0.01*
Adducin S662	↓	10,607	17,158	−38%	−12184/376	0.02*
B-Raf S445	↑	32,157	23,762	+35%	434/15,195	0.03*
Catenin-beta T41/S45	↓	7750	10,203	−24%	−5.974/−182	0.04**
CREB S133	↓	12,886	18,102	−29%	−8997/−1052	0.02*
ErbB2 Y1248	↓	8228	10,539	−22%	−5002/−353	0.03**
ErbB3 Y1289	↓	6989	10,534	−34%	−6307/−1148	<0.01*
HSP27 S82	↓	18,146	29,226	−38%	−20208/0	0.02*
Paxillin Y118	↓	1294	2544	−49%	−3612/0	0.03**
PTEN	↓	15,438	23,624	−35%	−15540/−472	0.01*
PTEN S380	↓	14,076	21,551	−35%	−13596/0	0.01*
P90RSK S380	↓	6537	10,991	−41%	−10320/−381	0.04**
Smad1 S463/465-Smad5 S463/465-Smad9 S465/467	↓	5245	10,134	−48%	−6959/−515	0.02**
Smad2 S245/250/255	↓	6626	8959	−26%	−4724/−29	0.04**
Stat3 Y727	↓	10,840	15,109	−29%	−8489/615	0.03*
Stat5 Y694	↓	6158	11,008	−44%	−8221/−1252	0.02**
Survivin	↑	9926	5829	+70%	1534/6588	<0.01**
VEGFR2 Y996	↓	17,322	19,753	−12%	−5362/524	0.03*
Panel B: Stroma compartment						
Akt S473	↑	19,990	9906	+102%	98/16,558	0.02*
E-Cadherin	↑	26,439	15,608	+69%	950/20,033	0.04*
Egr1	↑	4423	1580	+193%	144/2080	0.01**
eNOS S113	↑	7392	4924	+50%	827/2611	<0.01**
Erk 1/2 T202/Y204	↑	36,222	24,395	+48%	2269/21,263	0.03*
IL-8	↑	3675	2611	+41%	95/1357	0.02**
LDHA	↑	18,800	12,785	+47%	593/10,137	0.02*
MEK1/2 S217/221	↑	10,148	7874	+29%	300/3062	0.02**
TIMP2	↑	19,700	16,705	+18%	557/5385	0.01**

* p-value calculated based on two-sample t-test; ** p-value calculated based on Wilcoxon rank sum test

Regarding the stroma compartment, increased levels of IL-8, Egr-1, TIMP-2, and the hyperactivation of MEK, Erk, Akt and its downstream substrate eNOS were detected in the malignant compared to the normal-appearing counterpart (Tab. 1). These findings are in agreement with previous studies, and connecting them allowed us to rebuild the molecular cascade of events that leads PCA to become androgen-independent by the AR activation due to IL8 produced by stroma (Fig. 5).

Unexpectedly, our data also showed an increase in the activation level of the inflammatory molecules Stat3 Y727 and Stat5 Y694 along with different members of the

Smad family, and the RTKs ErbB2 Y1248 and ErbB3 Y1289 in the normal-appearing epithelim

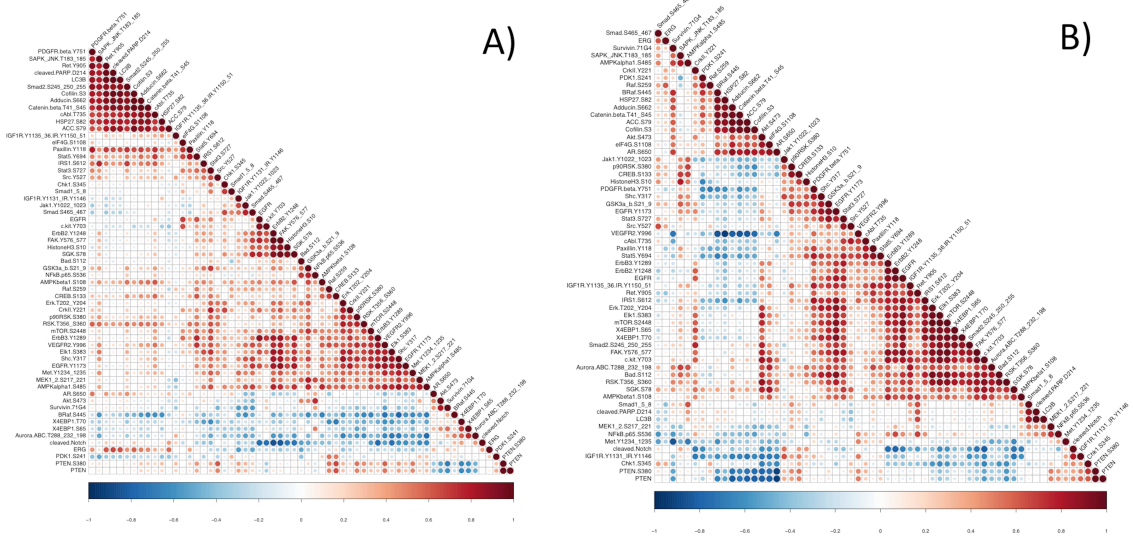

Figure 4. Correlation matrices of the different cellular compartments. In a color-coded scale from blue (low) to high (red) are represented the correlation coefficients for each pair of targets included in the analysis for A) tumor epithelium, and B) normal-appearing epithelium.

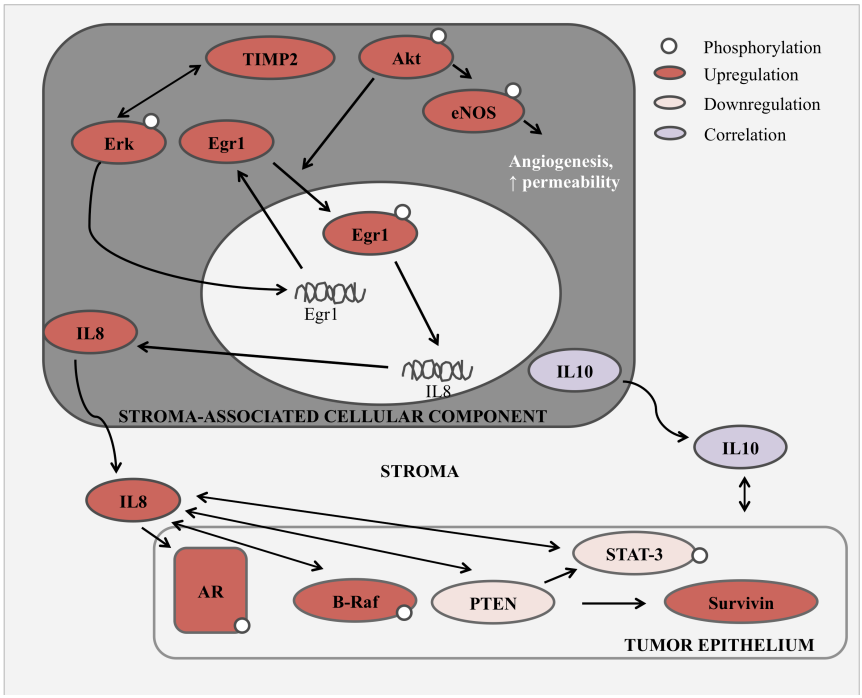

Figure 5. Protein network activation in PCA tumor microenvironment. The scheme represents selected proteins and phosphoproteins found to be differentially activated/expressed in the tumor epithelium and nearby tumor associated stroma.

To us, the explanation lays on the *field cancerization effect*[242], a biological process where morphologically normal-appearing areas close to the tumor were found to carry genetic alterations and deranged molecular structure; this is in agreement with the multi-step theory that attributes the clonal expansion, and the subsequent tumor development, to the accumulation of genetic alterations that can already occur in a still morphologically normal-appearing tissue. Supporting this, our results showed that the stroma surrounding the normal-appearing epithelium has a large number of strong positive correlations between targets involved in the immune response but also in the remodeling processes of extracellular matrix (e.g. Caveolin, MMP-9, IRAK1, Jak2, TGFb, TIMP-3, Zap70, Podoplanin) (Fig 6).

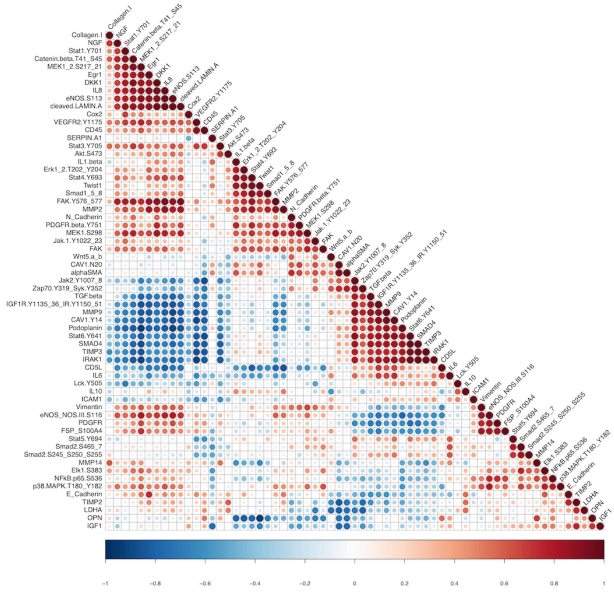

Figure 6. Correlation matrices of the normal-appearing stroma. In a color-coded scale from blue (low) to high (red) are represented the correlation coefficients for each pair of targets included in the analysis.

The correlation analysis of targets identified highly differential changes in the inflammatory/immune-response between the four different compartments. In the normal-appearing epithelium and surrounding microenvironment known feedback mechanisms regulating immuno-surveillance were more extensively maintained compared to the

malignant counterpart. The stroma surrounding the tumor cells showed a negative correlation between the IL-10 secretion and the lymphocyte activation status. IL-10 has also emerged as a major interaction node between epithelium and stroma of the malignant PCA tissue (Fig. 7).

Figure 7. Correlation analysis of tumor epithelium and mtched stroma. The diagram shows only correlations with coefficient > or equal to 0.75.

Based on the reported results, the workflow proposed in our study allowed us to explore the molecular network regulating the cross-talk between tumor and microenvironment during PCA growth and progression. Even if based on a low number of samples and limited by the field cancerization effect - which restricts our comparison to the signaling network of tumor cells and a normal-appearing but clearly already affected epithelium - our results indicate the feasibility of combining LCM and RPPA for separately profiling and evaluating the molecular architecture and cross-talking of epithelial and stromal compartments, needing very little clinical material. Further analyses on a larger number of samples are required to confirm our findings.

6.2 Paper II: Profiling the autoimmune repertoire in PCA by using planar and bead-based antigen and peptide arrays

This paragraph summarizes methods and results of a research study carried on by the author of this thesis' and collaborators. The research was published on the Journal of Proteome Research in October 2016 as an original article that can be refered to for more details. Reference: Pin E, Henjes F, Hong MG. Identification of a novel autoimmune peptide epitope of prostein in prostate cancer. J Proteome Res. 2016. DOI: 10.1021/acs.jproteome.6b00620 [Epub ahead of print].

Blood deriving biomarkers are desirable due to the less invasive sampling procedures, and the higher amount of available sample. In this study, we profiled the IgG repertoire of 589 plasma and serum samples from PCA patients with early and late-stage disease plus 20 controls (Tab. 2), to identify new potential autoimmune biomarkers for improving the definition of patient prognosis. The investigation was divided into three phases where we also combined planar antigen array, bead-based array and peptide arrays to validate our results (Fig. 8).

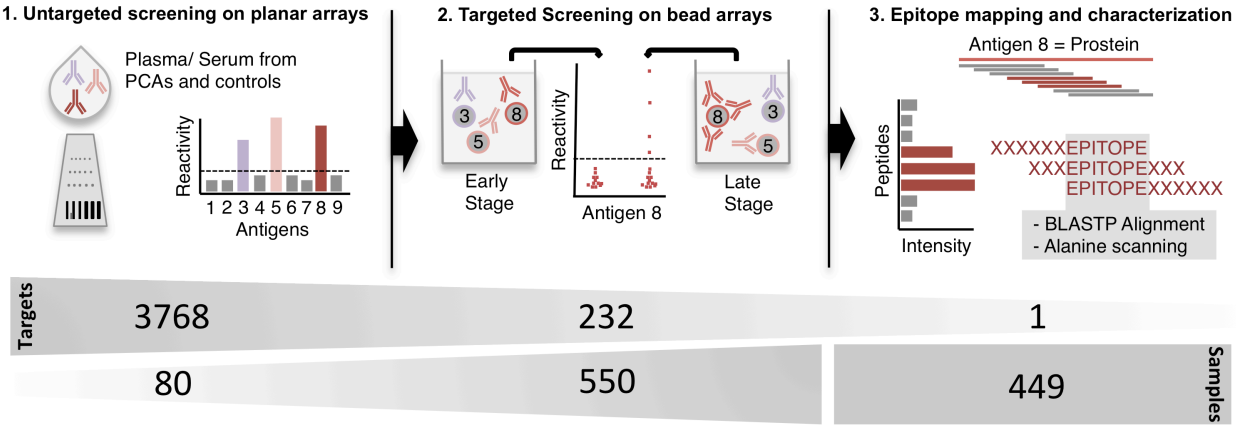

Figure 8. Study experimental design. The diagram represents the three phases in which the study was divided. A first untargeted screening phase was run on planar antigen arrays on 80 samples and testing reactivity toward 3768 protein fragments from the HPA collection. A selection of 161 antigens resulted reactive in phase 1, plus 71 additional targets selected from literature were included in phase 2. Protein fragments were coupled on magnetic beads to create a suspension bead antigen array. Reactivity toward these antigen was tested in 550 PCA plasma samples from early and late stage disease patients. Prostein representing fragment was highly reactive at late stages compared to early and was selected for further analysis to test the specificity of reactivity and to map the epitope.

Table 2. Clinical Data fof PCA Patients Included in Untargeted and Targeted Screening Phase[a]

variable	study sets	
name	**set 1**	**set 2**
collection time point	pre-diagnosis	diagnosis
sample type	serum	plasma
no. of patients	20	569
age (y)	66.7 ± 9.6	70.4 ± 8.6
PSA level*		
mean (ng/mL)		39.0 ± 241.2
<4 ng/mL N(%)		68 (12.0)
4−10 ng/mL N(%)		325 (57.2)
>10 ng/mL N(%)		175 (30.8)
Gleason grade N(%)*		
≤6		104 (43.5)
≥7		135 (56.5)
tumor stage N(%)*		
T1		345 (60.6)
T2	20 (100)	164 (28.9)
T3		49 (8.6)
T4		11 (1.9)
nodal stage N(%)*		
N0		24 (80.0)
N1		6 (20.0)
metastasis N(%)*		
M0		57 (76.0)
M1		18 (24)

[a] Clinical information are marked with (*) when the number of individuals did not correspond to total due to lack of data. Percentages are calculated considering only patients with available information. The Table does not include the 20 control subjects, for which no clinical information was available but age (66.6 ± 9.4).

A first broad untargeted screening phase was run with 80 age-matched samples divided in 20 healthy controls, 40 low stage PCA and 20 high stage PCA (from set 1 and 2 , Tab. 2). Using planar antigen microarrays produced within the HPA project, we tested the reactivity toward 3768 antigens, representing 3363 unique proteins. A high heterogeneity of reactivity was identified across the samples, characterized by a broad intensity range, with many antigens showing high intensity only in one single sample (Fig. 9).

Considering the high heterogeneity, a cut-off for antigens selection was set per sample, and by combining both intensity and frequency. Antigens with the highest intensities were considered valuable of inclusion in the second targeted screening phase only when a minimum of 2 samples passed the more stringent used cut-off (median + 250x MAD cut-off), while for antigen with lower intensity a higher frequency (minimum 10 reactive samples) was required for selection.

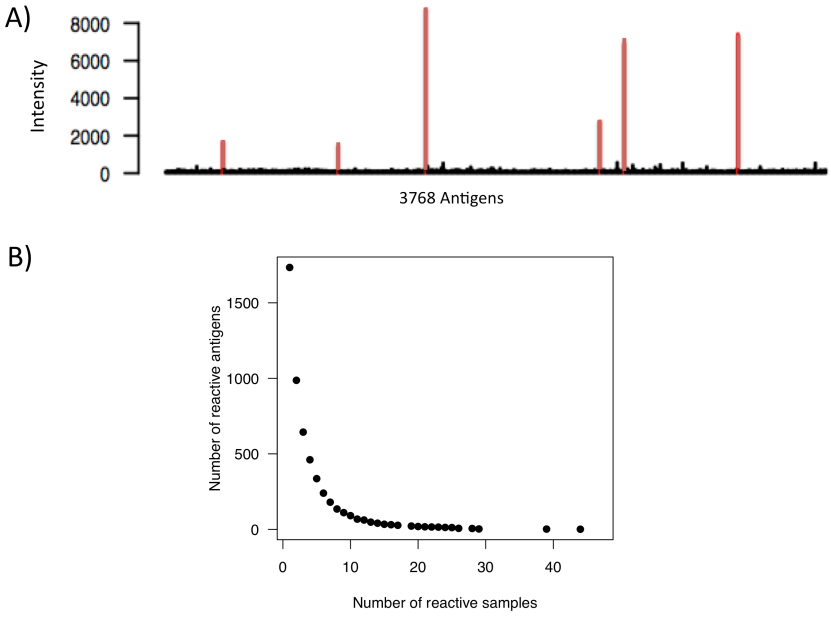

Figure 9. Heterogeneity of reactivity. (A) Example of IgG intensity profile for 1 samples over all 3768 antigens included in the untargeted screening phase. Red peaks represent antigens passing the cut-off and toward which IgG were reactive. (B) A large interindividual variation of reactivity profiles was observed, with most antigens recognized in single individuals only.

From this selection, 161 antigens were included in the second phase of the investigation. Beside these, other 71 antigens were selected representing interesting proteins from literature. In total, 201 unique proteins were included in the targeted screening phase. Reactivity toward the 232 protein fragments was evaluated using bead-based antigen arrays in 550 plasma samples collected from PCA cases at diagnosis (set 2; Tab. 2). Patients were classified in high T-stage (T3-T4) and low T-stage (T1-T2), and in high Gleason score (>7), median Gleason score (7) and low Gleason score (<7). Statistical analysis revealed that 3 antigens where more reactive at later stage of the disease, and they were representing IGFBP2, TBP and SLC45A3 - alias *prostein* – proteins (Tab. 3).

Table 3. Frequencies and Statistical Analysis Results for Antigens Identified To Be Differentially Reactive based on T-Stage and Gleason Score comparisons

Gene	Uniprot ID	Frequency of Reactivity			
		High Group*	Low Group*	Fisher Exact Test**	t-test***
High vs Low T-stage					
IGF2BP2	Q9Y6M1	33 (14)	18 (91)	0.025	0.18
SLC45A3	Q96JT2	26 (11)	14 (71)	0.044	0.49
High vs Low Gleason score					
SLC45A3	Q96JT2	27 (9)	10 (10)	0.044	0.35
TBP	P20226	21 (7)	5 (5)	0.016	0.39

* When the percentage and number of samples does not match the numbers reported in Tab. 1, it is due to exclusion of some samples from the analysis for technical criteria. ** Comparison of reactivity frequencies between the two groups. *** Comparison of mean of intensities between the two compared groups.

Among these 3 proteins we selected prostein, and analyzed its reactivity in deeper because it was confirmed to be more reactive at late stages of PCA both comparing high versus low T-stage and Gleason score (Tab.3 and Fig.10).

Also, prostein expression is restricted to prostate cells, making it a perfect PCA biomarker candidate. Further analyses were performed to identify the reactive epitope and test the specificity of this reactivity. Autoimmunity is complex, and often polyreactive antibodies or antibodies with specificity for other proteins with high homology could cross-react with an epitope. Being able to discriminate if the reactivity we identified is specific for prostein is essential for defining its role as a candidate biomarker. For this reason, both male and female samples collected pre-diagnosis were tested for reactivity toward the prostein antigen. Reactivity in men pre-diagnosis was identified at levels similar to very early stage PCA, possibly explained by the fact that prostein is expressed in both normal prostate and PCA or due to antibodies cross-reacting with the fragment. Female samples showed a median of signal slightly lower than men prediagnosis.

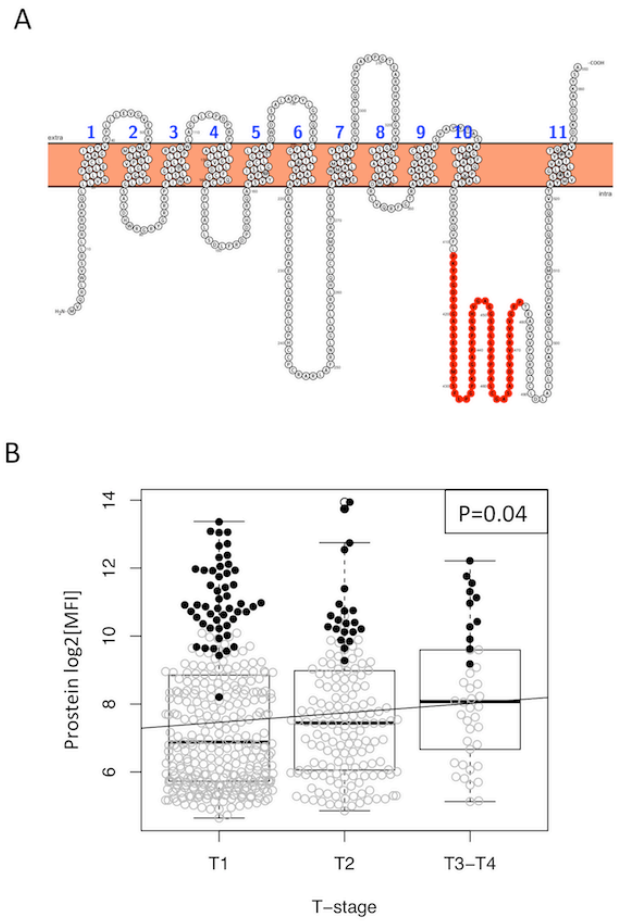

Figure 10. SLC45A3/Prostein structure and reactivity. (A) Secondary structure shows that prostein is a transmembrane protein with 11 transmembrane domains. Reactive antigen (red; aa 412–477) are localized in intracompartmental domains. (B) Linear model showing association between the reactive prostein antigen MFI and the tumor stage. ● = reactive samples. P-value for Anova test is reported in the graph.

An epitope mapping analysis using bead-based array with overlapping 15mer peptides to cover the whole prostein 66 aminoacids fragment sequence, showed that both men and women prediagnosis and PCA samples share the same epitope. The 9 aminoacid length epitope with sequence GPKPGAPFP covers the prostein sequence from aminoacid 434 to 442, in the C-terminal region (Fig. 11).

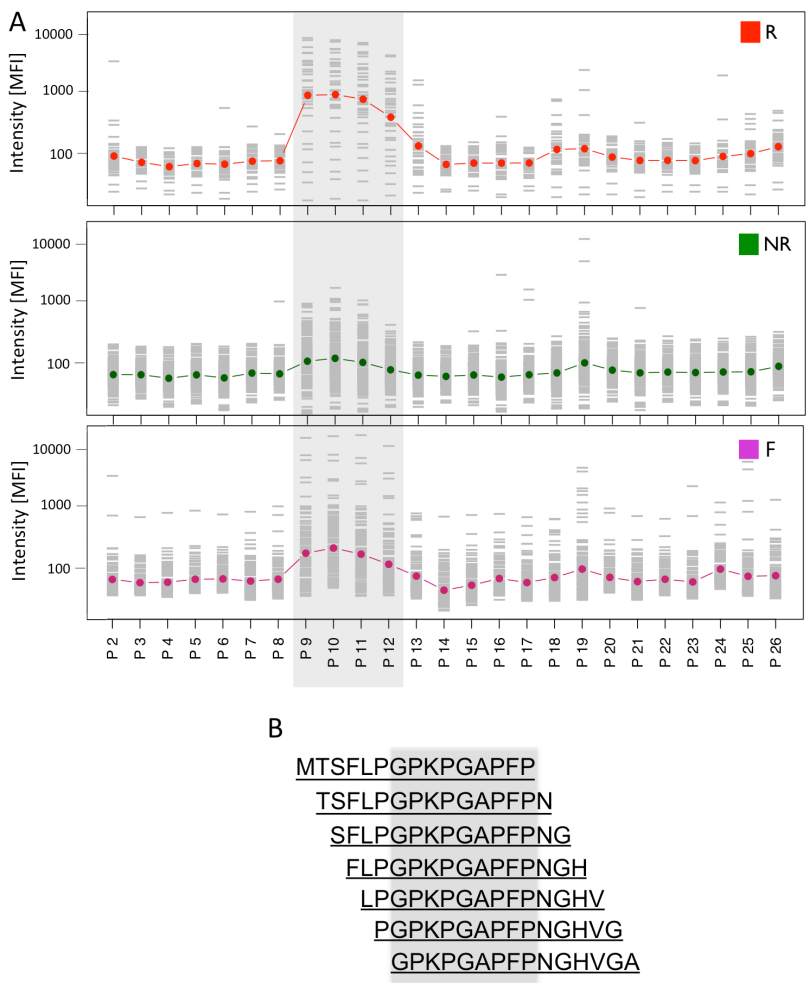

Figure 11. Prostein epitope identification. (A) Dot plots showing the MFI mean detected in reactive (R; N = 40), nonreactive (NR; N = 189) and female (F; N= samples, for epitope mapping using 15mers with 13 aa overlap to the next one. P1 was excluded from the analysis due to technical reasons. (B) Sequence of each of the statistically different peptides identified with the epitope refinement analysis using 15mers with 14 aa overlap to the next one. The epitope is highlighted in orange.

The high concentration of prolines, which are 4 out of 9 aminoacids, makes the epitope a target for natural polyreactive antibodies, which could explain the reactivity identified in female samples. Beside this, the reactivity in women could be explained also by the hypothetical presence of autoantibodies directed toward epitopes of TGIF2 which showed a sequence with 78% homology to prostein epitope. In parallel, also part of the men reactivity could be aspecific due to polyreactive antibodies, antibodies toward TGIF2, and also toward TMEM79, another prostate specific protein containing a sequence with

78% homology to the identified prostein epitope. In conclusion, our study demonstrates once again the high level of complexity of the immune system, and that it can be a source of biomarkers for PCA. These results also show that the deconvolution of such complexity is essential to identify useful and specific autoantibodies biomarkers. A limitation of our study is the use of protein fragments. This leads to lack of information about the presence of other epitopes on the same protein. Still, the availability of thousands of protein fragments is a unique resource for autoimmunity analysis. We believe that our approach based on cross-platform validation using antigen and peptides, planar and bead-based arrays, and complemented with bioinformatics analysis, could be a comprehensive and solid approach for deep characterization of the autoimmune repertoire.

6.3 Conclusions and future perspectives

The work here presented is based on profiling of proteins, phosphoproteins and autoantibodies within PCA by means of microarray technology, with the aim to identify new biomarkers. Thanks to their high sensitivity and mid to high throughput, both FPPA and RPPA are suitable for profiling biological samples and identifying biomarkers or molecular profiles related to a disease status. In **Paper I**, profiles related to proteins and phosphoproteins were evaluated in PCA tissue samples from patients with medium Gleason score (6-7). In particular, LCM and RPPA were coupled to separately profile the stroma and epithelium of both cancer and the surrounding normal-appearing epithelium, with the aim to rebuild the molecular structure of each compartment and study the epithelium-stroma cross-talking. Even though the number of enrolled patients was limited, the statistical analysis allowed to identify significant variations and correlations between targets in different compartments. Thanks to the high-sensitivity of RPPA, the expression and/or activation status of many targets was analyzed even when very limited amount of sample was available. Taken together, our analysis allowed to reveal differences in the activation of pathways between the different compartments. Beside differences, the study allowed also to rebuild a network that could explain PCA progression in Gleason score 6-7 patients by epithelium-stroma cross-talk.

In **Paper II** we identify a new IgG reactive epitope of prostein in PCA patients correlating with the stage of the disease. A general individual variation was observed for autoimmunity reactivity in PCA patients. This is in agreement with what present in literature, where reactivity toward human proteins are reported also for healthy individuals, and the majority of antigens show reactivity in a single individual or few individuals. Autoantibodies reactive toward the same antigen in many individuals are rare, making difficult to identify common patterns related to a disease. Beside this, cross-reacting autoantibodies or polyreactive antibodies can recognize the same epitope making the picture more complex, as highlighted also by our results on prostein. Reactivities are identified also in healthy samples, suggesting that an analysis of the baseline status in the general population is needed to profile autoantibodies in absence of symptoms, and possibly clean or normalized results in pathological conditions.

The data presented in this thesis provides a robust framework for protein profiling in tissue, and autoantibodies profiling in plasma in the context of PCA. Beside this, our data put some new insights on the molecular processes underlying PCA development and progression. The proposed workflows are based on careful evaluation of all the variables that could influence the results, from study design to data analysis. Our findings need to be validated in future studies involving a larger number of samples collected independently, hopefully contributing to increase knowledge, and optimizing the treatment of PCA.

1. H, G., Gray´s Anatomy - The anatomical basis of clinical practice 41st edition. 2015.
2. McNeal, J. E., Origin and development of carcinoma in the prostate. *Cancer* **1969,** *23* (1), 24-34.
3. Selman, S. H., The McNeal prostate: a review. *Urology* **2011,** *78* (6), 1224-8.
4. Tonry, C. L.; Leacy, E.; Raso, C.; Finn, S. P.; Armstrong, J.; Pennington, S. R., The Role of Proteomics in Biomarker Development for Improved Patient Diagnosis and Clinical Decision Making in Prostate Cancer. *Diagnostics (Basel)* **2016,** *6* (3).
5. Packer, J. R.; Maitland, N. J., The molecular and cellular origin of human prostate cancer. *Biochim Biophys Acta* **2016,** *1863* (6 Pt A), 1238-60.
6. Wong, M. C.; Goggins, W. B.; Wang, H. H.; Fung, F. D.; Leung, C.; Wong, S. Y.; Ng, C. F.; Sung, J. J., Global Incidence and Mortality for Prostate Cancer: Analysis of Temporal Patterns and Trends in 36 Countries. *Eur Urol* **2016**.
7. Brawley, O. W., Trends in prostate cancer in the United States. *J Natl Cancer Inst Monogr* **2012,** *2012* (45), 152-6.
8. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **2015,** *136* (5), E359-86.
9. Ferlay, J.; Shin, H. R.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D. M., Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* **2010,** *127* (12), 2893-917.
10. SEER Cancer Statistics Factsheets: Prostate Cancer. National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/statfacts/html/prost.html.
11. Albright, F.; Teerlink, C.; Werner, T. L.; Cannon-Albright, L. A., Significant evidence for a heritable contribution to cancer predisposition: a review of cancer familiality by site. *BMC Cancer* **2012,** *12*, 138.
12. Lynch, H. T.; Kosoko-Lasaki, O.; Leslie, S. W.; Rendell, M.; Shaw, T.; Snyder, C.; D'Amico, A. V.; Buxbaum, S.; Isaacs, W. B.; Loeb, S.; Moul, J. W.; Powell, I., Screening for familial and hereditary prostate cancer. *Int J Cancer* **2016,** *138* (11), 2579-91.
13. Stewart, S. B.; Freedland, S. J., Influence of obesity on the incidence and treatment of genitourinary malignancies. *Urol Oncol* **2011,** *29* (5), 476-86.
14. Venkateswaran, V.; Klotz, L. H., Diet and prostate cancer: mechanisms of action and implications for chemoprevention. *Nat Rev Urol* **2010,** *7* (8), 442-53.
15. Morgentaler, A.; Rhoden, E. L., Prevalence of prostate cancer among hypogonadal men with prostate-specific antigen levels of 4.0 ng/mL or less. *Urology* **2006,** *68* (6), 1263-7.
16. Murphy, A. B.; Ukoli, F.; Freeman, V.; Bennett, F.; Aiken, W.; Tulloch, T.; Coard, K.; Angwafo, F.; Kittles, R. A., 8q24 risk alleles in West African and Caribbean men. *Prostate* **2012,** *72* (12), 1366-73.
17. Taksler, G. B.; Keating, N. L.; Cutler, D. M., Explaining racial differences in prostate cancer mortality. *Cancer* **2012,** *118* (17), 4280-9.
18. Holmes, J. A.; Carpenter, W. R.; Wu, Y.; Hendrix, L. H.; Peacock, S.; Massing, M.; Schenck, A. P.; Meyer, A. M.; Diao, K.; Wheeler, S. B.; Godley, P. A.; Stitzenberg, K. B.; Chen, R. C., Impact of distance to a urologist on early diagnosis of prostate cancer among black and white patients. *J Urol* **2012,** *187* (3), 883-8.
19. Stamey, T. A.; Freiha, F. S.; McNeal, J. E.; Redwine, E. A.; Whittemore, A. S.; Schmid, H. P., Localized prostate cancer. Relationship of tumor volume to clinical significance for treatment of prostate cancer. *Cancer* **1993,** *71* (3 Suppl), 933-8.
20. Thompson, I. M.; Pauler, D. K.; Goodman, P. J.; Tangen, C. M.; Lucia, M. S.; Parnes, H. L.; Minasian, L. M.; Ford, L. G.; Lippman, S. M.; Crawford, E. D.; Crowley, J. J.; Coltman, C. A., Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med* **2004,** *350* (22), 2239-46.
21. Cancer, A. J. C. o., Purposes and Principles

of Cancer Staging. 2010.
22. Humphrey, P. A., Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol* **2004,** *17* (3), 292-306.
23. Nguyen-Nielsen, M.; Borre, M., Diagnostic and Therapeutic Strategies for Prostate Cancer. *Semin Nucl Med* **2016,** *46* (6), 484-490.
24. Trewartha, D.; Carter, K., Advances in prostate cancer treatment. *Nat Rev Drug Discov* **2013,** *12* (11), 823-4.

25. Shah, R. B.; Mehra, R.; Chinnaiyan, A. M.; Shen, R.; Ghosh, D.; Zhou, M.; Macvicar, G. R.; Varambally, S.; Harwood, J.; Bismar, T. A.; Kim, R.; Rubin, M. A.; Pienta, K. J., Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. *Cancer Res* **2004,** *64* (24), 9209-16.
26. Rycaj, K.; Cho, E. J.; Liu, X.; Chao, H. P.; Liu, B.; Li, Q.; Devkota, A. K.; Zhang, D.; Chen, X.; Moore, J.; Dalby, K. N.; Tang, D. G., Longitudinal tracking of subpopulation dynamics and molecular changes during LNCaP cell castration and identification of inhibitors that could target the PSA-/lo castration-resistant cells. *Oncotarget* **2016,** *7* (12), 14220-40.
27. Kumar, A.; Coleman, I.; Morrissey, C.; Zhang, X.; True, L. D.; Gulati, R.; Etzioni, R.; Bolouri, H.; Montgomery, B.; White, T.; Lucas, J. M.; Brown, L. G.; Dumpit, R. F.; DeSarkar, N.; Higano, C.; Yu, E. Y.; Coleman, R.; Schultz, N.; Fang, M.; Lange, P. H.; Shendure, J.; Vessella, R. L.; Nelson, P. S., Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat Med* **2016,** *22* (4), 369-78.
28. Cooper, C. S.; Eeles, R.; Wedge, D. C.; Van Loo, P.; Gundem, G.; Alexandrov, L. B.; Kremeyer, B.; Butler, A.; Lynch, A. G.; Camacho, N.; Massie, C. E.; Kay, J.; Luxton, H. J.; Edwards, S.; Kote-Jarai, Z.; Dennis, N.; Merson, S.; Leongamornlert, D.; Zamora, J.; Corbishley, C.; Thomas, S.; Nik-Zainal, S.; Ramakrishna, M.; O'Meara, S.; Matthews, L.; Clark, J.; Hurst, R.; Mithen, R.; Bristow, R. G.; Boutros, P. C.; Fraser, M.; Cooke, S.; Raine, K.; Jones, D.; Menzies, A.; Stebbings, L.; Hinton, J.; Teague, J.; McLaren, S.; Mudie, L.; Hardy, C.; Anderson, E.; Joseph, O.; Goody, V.; Robinson, B.; Maddison, M.; Gamble, S.; Greenman, C.; Berney, D.; Hazell, S.; Livni, N.; Fisher, C.; Ogden, C.; Kumar, P.; Thompson, A.; Woodhouse, C.; Nicol, D.; Mayer, E.; Dudderidge, T.; Shah, N. C.; Gnanapragasam, V.; Voet, T.; Campbell, P.; Futreal, A.; Easton, D.; Warren, A. Y.; Foster, C. S.; Stratton, M. R.; Whitaker, H. C.; McDermott, U.; Brewer, D. S.; Neal, D. E.; Group, I. P., Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat Genet* **2015,** *47* (4), 367-72.
29. Crick, F., Central dogma of molecular biology. *Nature* **1970,** *227* (5258), 561-3.
30. Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y. H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigó, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y. H.; Coyne, M.; Dahlke, C.; Mays, A.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M.; Pan, S.; Peck, J.; Peterson,

M.; Rowe, W.; Sanders, R.; Scott, J.; Simpson, M.; Smith, T.; Sprague, A.; Stockwell, T.; Turner, R.; Venter, E.; Wang, M.; Wen, M.; Wu, D.; Wu, M.; Xia, A.; Zandieh, A.; Zhu, X., The sequence of the human genome. *Science* **2001,** *291* (5507), 1304-51.
31. Black, D. L., Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **2003,** *72*, 291-336.
32. Ayoubi, T. A.; Van De Ven, W. J., Regulation of gene expression by alternative promoters. *FASEB J* **1996,** *10* (4), 453-60.
33. Khoury, G. A.; Baliban, R. C.; Floudas, C. A., Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* **2011,** *1*.
34. Bensimon, A.; Heck, A. J.; Aebersold, R., Mass spectrometry-based proteomics and network biology. *Annu Rev Biochem* **2012,** *81*, 379-405.
35. Stern, D. F., Phosphoproteomics for oncology discovery and treatment. *Expert Opin Ther Targets* **2005,** *9* (4), 851-60.
36. Guha, U.; Chaerkady, R.; Marimuthu, A.; Patterson, A. S.; Kashyap, M. K.; Harsha, H. C.; Sato, M.; Bader, J. S.; Lash, A. E.; Minna, J. D.; Pandey, A.; Varmus, H. E., Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. *Proc Natl Acad Sci U S A* **2008,** *105* (37), 14112-7.
37. Cui, Q.; Ma, Y.; Jaramillo, M.; Bari, H.; Awan, A.; Yang, S.; Zhang, S.; Liu, L.; Lu, M.; O'Connor-McCourt, M.; Purisima, E. O.; Wang, E., A map of human cancer signaling. *Mol Syst Biol* **2007,** *3*, 152.
38. Jin, Q.; Esteva, F. J., Cross-talk between the ErbB/HER family and the type I insulin-like growth factor receptor signaling pathway in breast cancer. *J Mammary Gland Biol Neoplasia* **2008,** *13* (4), 485-98.
39. Haura, E. B.; Zheng, Z.; Song, L.; Cantor, A.; Bepler, G., Activated epidermal growth factor receptor-Stat-3 signaling promotes tumor survival in vivo in non-small cell lung cancer. *Clin Cancer Res* **2005,** *11* (23), 8288-94.
40. Zandi, R.; Larsen, A. B.; Andersen, P.; Stockhausen, M. T.; Poulsen, H. S., Mechanisms for oncogenic activation of the epidermal growth factor receptor. *Cell Signal* **2007,** *19* (10), 2013-23.
41. Wiley, H. S., Trafficking of the ErbB receptors and its influence on signaling. *Exp Cell Res* **2003,** *284* (1), 78-88.
42. Henry, N. L.; Hayes, D. F., Cancer biomarkers. *Mol Oncol* **2012,** *6* (2), 140-6.
43. Etzioni, R.; Urban, N.; Ramsey, S.; McIntosh, M.; Schwartz, S.; Reid, B.; Radich, J.; Anderson, G.; Hartwell, L., The case for early detection. *Nat Rev Cancer* **2003,** *3* (4), 243-52.
44. Scott, A. M.; Wolchok, J. D.; Old, L. J., Antibody therapy of cancer. *Nat Rev Cancer* **2012,** *12* (4), 278-87.
45. Gschwind, A.; Fischer, O. M.; Ullrich, A., The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat Rev Cancer* **2004,** *4* (5), 361-70.
46. Damoiseaux, J.; Andrade, L. E.; Fritzler, M. J.; Shoenfeld, Y., Autoantibodies 2015: From diagnostic biomarkers toward prediction, prognosis and prevention. *Autoimmun Rev* **2015,** *14* (6), 555-63.
47. Kobold, S.; Luetkens, T.; Cao, Y.; Bokemeyer, C.; Atanackovic, D., Prognostic and diagnostic value of spontaneous tumor-related antibodies. *Clin Dev Immunol* **2010,** *2010*, 721531.
48. Kulasingam, V.; Diamandis, E. P., Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol* **2008,** *5* (10), 588-99.
49. Fredolini, C.; Byström, S.; Pin, E.; Edfors, F.; Tamburro, D.; Iglesias, M. J.; Häggmark, A.; Hong, M. G.; Uhlen, M.; Nilsson, P.; Schwenk, J. M., Immunocapture strategies in translational proteomics. *Expert Rev Proteomics* **2016,** *13* (1), 83-98.
50. Pierobon, M.; Wulfkuhle, J.; Liotta, L.; Petricoin, E., Application of molecular technologies for phosphoproteomic analysis of clinical samples. *Oncogene* **2015,** *34* (7), 805-14.
51. de Gramont, A.; Watson, S.; Ellis, L. M.; Rodón, J.; Tabernero, J.; Hamilton, S. R., Pragmatic issues in biomarker evaluation for targeted therapies in cancer. *Nat Rev Clin Oncol* **2015,** *12* (4), 197-212.
52. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.
53. Aebersold, R.; Mann, M., Mass-spectrometric exploration of proteome structure and function. *Nature* **2016,** *537* (7620), 347-55.
54. Stoevesandt, O.; Taussig, M. J., Affinity proteomics: the role of specific binding reagents in human proteome analysis. *Expert Rev Proteomics* **2012,** *9* (4), 401-14.
55. Ackermann, B. L.; Berna, M. J., Coupling immunoaffinity techniques with MS for quantitative analysis of low-abundance protein biomarkers. *Expert Rev Proteomics* **2007,** *4* (2), 175-86.

56. Dyson, M. R.; Zheng, Y.; Zhang, C.; Colwill, K.; Pershad, K.; Kay, B. K.; Pawson, T.; McCafferty, J., Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *Anal Biochem* **2011,** *417* (1), 25-35.
57. Egelhofer, T. A.; Minoda, A.; Klugman, S.; Lee, K.; Kolasinska-Zwierz, P.; Alekseyenko, A. A.; Cheung, M. S.; Day, D. S.; Gadel, S.; Gorchakov, A. A.; Gu, T.; Kharchenko, P. V.; Kuan, S.; Latorre, I.; Linder-Basso, D.; Luu, Y.; Ngo, Q.; Perry, M.; Rechtsteiner, A.; Riddle, N. C.; Schwartz, Y. B.; Shanower, G. A.; Vielle, A.; Ahringer, J.; Elgin, S. C.; Kuroda, M. I.; Pirrotta, V.; Ren, B.; Strome, S.; Park, P. J.; Karpen, G. H.; Hawkins, R. D.; Lieb, J. D., An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* **2011,** *18* (1), 91-3.
58. Björling, E.; Uhlén, M., Antibodypedia, a portal for sharing antibody and antigen validation data. *Mol Cell Proteomics* **2008,** *7* (10), 2028-37.
59. Whiteaker, J. R.; Zhao, L.; Frisch, C.; Ylera, F.; Harth, S.; Knappik, A.; Paulovich, A. G., High-affinity recombinant antibody fragments (Fabs) can be applied in peptide enrichment immuno-MRM assays. *J Proteome Res* **2014,** *13* (4), 2187-96.
60. Li, Y.; He, J.; Niu, Y.; Yu, C., Ultrasensitive electrochemical biosensor based on reduced graphene oxide-tetraethylene pentamine-BMIMPF6 hybrids for the detection of α2,6-sialylated glycans in human serum. *Biosens Bioelectron* **2015,** *74*, 953-9.
61. Muyldermans, S., Nanobodies: natural single-domain antibodies. *Annu Rev Biochem* **2013,** *82*, 775-97.
62. Groll, N.; Emele, F.; Poetz, O.; Rothbauer, U., Towards multiplexed protein-protein interaction analysis using protein tag-specific nanobodies. *J Proteomics* **2015,** *127* (Pt B), 289-99.
63. Renberg, B.; Nordin, J.; Merca, A.; Uhlén, M.; Feldwisch, J.; Nygren, P. A.; Karlström, A. E., Affibody molecules in protein capture microarrays: evaluation of multidomain ligands and different detection formats. *J Proteome Res* **2007,** *6* (1), 171-9.
64. Wilhelm, M.; Schlegl, J.; Hahne, H.; Gholami, A. M.; Lieberenz, M.; Savitski, M. M.; Ziegler, E.; Butzmann, L.; Gessulat, S.; Marx, H.; Mathieson, T.; Lemeer, S.; Schnatbaum, K.; Reimer, U.; Wenschuh, H.; Mollenhauer, M.; Slotta-Huspenina, J.; Boese, J. H.; Bantscheff, M.; Gerstmair, A.; Faerber, F.; Kuster, B., Mass-spectrometry-based draft of the human proteome. *Nature* **2014,** *509* (7502), 582-7.
65. Farrah, T.; Deutsch, E. W.; Omenn, G. S.; Sun, Z.; Watts, J. D.; Yamamoto, T.; Shteynberg, D.; Harris, M. M.; Moritz, R. L., State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. *J Proteome Res* **2014,** *13* (1), 60-75.
66. Bausch-Fluck, D.; Hofmann, A.; Bock, T.; Frei, A. P.; Cerciello, F.; Jacobs, A.; Moest, H.; Omasits, U.; Gundry, R. L.; Yoon, C.; Schiess, R.; Schmidt, A.; Mirkowska, P.; Härtlová, A.; Van Eyk, J. E.; Bourquin, J. P.; Aebersold, R.; Boheler, K. R.; Zandstra, P.; Wollscheid, B., A mass spectrometric-derived cell surface protein atlas. *PLoS One* **2015,** *10* (3), e0121314.
67. Uhlén, M.; Björling, E.; Agaton, C.; Szigyarto, C. A.; Amini, B.; Andersen, E.; Andersson, A. C.; Angelidou, P.; Asplund, A.; Asplund, C.; Berglund, L.; Bergström, K.; Brumer, H.; Cerjan, D.; Ekström, M.; Elobeid, A.; Eriksson, C.; Fagerberg, L.; Falk, R.; Fall, J.; Forsberg, M.; Björklund, M. G.; Gumbel, K.; Halimi, A.; Hallin, I.; Hamsten, C.; Hansson, M.; Hedhammar, M.; Hercules, G.; Kampf, C.; Larsson, K.; Lindskog, M.; Lodewyckx, W.; Lund, J.; Lundeberg, J.; Magnusson, K.; Malm, E.; Nilsson, P.; Odling, J.; Oksvold, P.; Olsson, I.; Oster, E.; Ottosson, J.; Paavilainen, L.; Persson, A.; Rimini, R.; Rockberg, J.; Runeson, M.; Sivertsson, A.; Sköllermo, A.; Steen, J.; Stenvall, M.; Sterky, F.; Strömberg, S.; Sundberg, M.; Tegel, H.; Tourle, S.; Wahlund, E.; Waldén, A.; Wan, J.; Wernérus, H.; Westberg, J.; Wester, K.; Wrethagen, U.; Xu, L. L.; Hober, S.; Pontén, F., A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* **2005,** *4* (12), 1920-32.
68. Berglund, L.; Björling, E.; Jonasson, K.; Rockberg, J.; Fagerberg, L.; Al-Khalili Szigyarto, C.; Sivertsson, A.; Uhlén, M., A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation. *Proteomics* **2008,** *8* (14), 2832-9.
69. Berglund, L.; Andrade, J.; Odeberg, J.; Uhlén, M., The epitope space of the human proteome. *Protein Sci* **2008,** *17* (4), 606-13.
70. Tegel, H.; Steen, J.; Konrad, A.; Nikdin, H.; Pettersson, K.; Stenvall, M.; Tourle, S.; Wrethagen, U.; Xu, L.; Yderland, L.; Uhlén, M.; Hober, S.; Ottosson, J., High-throughput protein production--lessons from scaling up from 10 to 288 recombinant proteins per week. *Biotechnol J* **2009,** *4* (1), 51-7.

71. Sjölander, A.; Nygren, P. A.; Stahl, S.; Berzins, K.; Uhlen, M.; Perlmann, P.; Andersson, R., The serum albumin-binding region of streptococcal protein G: a bacterial fusion partner with carrier-related properties. *J Immunol Methods* **1997,** *201* (1), 115-23.
72. Nilsson, P.; Paavilainen, L.; Larsson, K.; Odling, J.; Sundberg, M.; Andersson, A. C.; Kampf, C.; Persson, A.; Al-Khalili Szigyarto, C.; Ottosson, J.; Björling, E.; Hober, S.; Wernérus, H.; Wester, K.; Pontén, F.; Uhlen, M., Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* **2005,** *5* (17), 4327-37.
73. Andersson, A. C.; Strömberg, S.; Bäckvall, H.; Kampf, C.; Uhlen, M.; Wester, K.; Pontén, F., Analysis of protein expression in cell microarrays: a tool for antibody-based proteomics. *J Histochem Cytochem* **2006,** *54* (12), 1413-23.
74. Kampf, C.; Olsson, I.; Ryberg, U.; Sjöstedt, E.; Pontén, F., Production of tissue microarrays, immunohistochemistry staining and digitalization within the human protein atlas. *J Vis Exp* **2012,** (63).
75. Li, J.; Newberg, J. Y.; Uhlén, M.; Lundberg, E.; Murphy, R. F., Automated analysis and reannotation of subcellular locations in confocal images from the Human Protein Atlas. *PLoS One* **2012,** *7* (11), e50514.
76. Fagerberg, L.; Hallström, B. M.; Oksvold, P.; Kampf, C.; Djureinovic, D.; Odeberg, J.; Habuka, M.; Tahmasebpoor, S.; Danielsson, A.; Edlund, K.; Asplund, A.; Sjöstedt, E.; Lundberg, E.; Szigyarto, C. A.; Skogs, M.; Takanen, J. O.; Berling, H.; Tegel, H.; Mulder, J.; Nilsson, P.; Schwenk, J. M.; Lindskog, C.; Danielsson, F.; Mardinoglu, A.; Sivertsson, A.; von Feilitzen, K.; Forsberg, M.; Zwahlen, M.; Olsson, I.; Navani, S.; Huss, M.; Nielsen, J.; Ponten, F.; Uhlén, M., Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* **2014,** *13* (2), 397-406.
77. Ayoglu, B.; Mitsios, N.; Kockum, I.; Khademi, M.; Zandian, A.; Sjöberg, R.; Forsström, B.; Bredenberg, J.; Lima Bomfim, I.; Holmgren, E.; Grönlund, H.; Guerreiro-Cacais, A. O.; Abdelmagid, N.; Uhlén, M.; Waterboer, T.; Alfredsson, L.; Mulder, J.; Schwenk, J. M.; Olsson, T.; Nilsson, P., Anoctamin 2 identified as an autoimmune target in multiple sclerosis. *Proc Natl Acad Sci U S A* **2016,** *113* (8), 2188-93.
78. Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P. H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.; von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Pontén, F., Proteomics. Tissue-based map of the human proteome. *Science* **2015,** *347* (6220), 1260419.
79. Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O., Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **1995,** *270* (5235), 467-70.
80. Sutandy, F. X.; Qian, J.; Chen, C. S.; Zhu, H., Overview of protein microarrays. *Curr Protoc Protein Sci* **2013,** *Chapter 27*, Unit 27.1.
81. Uzoma, I.; Zhu, H., Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. *Genomics Proteomics Bioinformatics* **2013,** *11* (1), 18-28.
82. Chowdhury, F.; Williams, A.; Johnson, P., Validation and comparison of two multiplex technologies, Luminex and Mesoscale Discovery, for human cytokine profiling. *J Immunol Methods* **2009,** *340* (1), 55-64.
83. Sun, H.; Chen, G. Y.; Yao, S. Q., Recent advances in microarray technologies for proteomics. *Chem Biol* **2013,** *20* (5), 685-99.
84. Lee, J. R.; Magee, D. M.; Gaster, R. S.; LaBaer, J.; Wang, S. X., Emerging protein array technologies for proteomics. *Expert Rev Proteomics* **2013,** *10* (1), 65-75.
85. Sjöberg, R.; Sundberg, M.; Gundberg, A.; Sivertsson, A.; Schwenk, J. M.; Uhlén, M.; Nilsson, P., Validation of affinity reagents using antigen microarrays. *N Biotechnol* **2012,** *29* (5), 555-63.
86. Delehanty, J. B., Printing functional protein microarrays using piezoelectric capillaries. *Methods Mol Biol* **2004,** *264*, 135-43.
87. Espina, V.; Mehta, A. I.; Winters, M. E.; Calvert, V.; Wulfkuhle, J.; Petricoin, E. F.; Liotta, L. A., Protein microarrays: molecular profiling technologies for clinical specimens. *Proteomics* **2003,** *3* (11), 2091-100.
88. S, A., 2013.
89. Häggmark, A.; Hamsten, C.; Wiklundh, E.; Lindskog, C.; Mattsson, C.; Andersson, E.; Lundberg, I. E.; Grönlund, H.; Schwenk, J. M.; Eklund, A.; Grunewald, J.; Nilsson, P., Proteomic profiling reveals autoimmune targets in sarcoidosis. *Am J Respir Crit Care Med* **2015,** *191* (5), 574-83.

90. Robinson, W. H.; DiGennaro, C.; Hueber, W.; Haab, B. B.; Kamachi, M.; Dean, E. J.; Fournel, S.; Fong, D.; Genovese, M. C.; de Vegvar, H. E.; Skriner, K.; Hirschberg, D. L.; Morris, R. I.; Muller, S.; Pruijn, G. J.; van Venrooij, W. J.; Smolen, J. S.; Brown, P. O.; Steinman, L.; Utz, P. J., Autoantigen microarrays for multiplex characterization of autoantibody responses. *Nat Med* **2002,** *8* (3), 295-301.
91. Joos, T. O.; Schrenk, M.; Höpfl, P.; Kröger, K.; Chowdhury, U.; Stoll, D.; Schörner, D.; Dürr, M.; Herick, K.; Rupp, S.; Sohn, K.; Hämmerle, H., A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics. *Electrophoresis* **2000,** *21* (13), 2641-50.
92. Hueber, W.; Kidd, B. A.; Tomooka, B. H.; Lee, B. J.; Bruce, B.; Fries, J. F.; Sønderstrup, G.; Monach, P.; Drijfhout, J. W.; van Venrooij, W. J.; Utz, P. J.; Genovese, M. C.; Robinson, W. H., Antigen microarray profiling of autoantibodies in rheumatoid arthritis. *Arthritis Rheum* **2005,** *52* (9), 2645-55.
93. Fattal, I.; Shental, N.; Mevorach, D.; Anaya, J. M.; Livneh, A.; Langevitz, P.; Zandman-Goddard, G.; Pauzner, R.; Lerner, M.; Blank, M.; Hincapie, M. E.; Gafter, U.; Naparstek, Y.; Shoenfeld, Y.; Domany, E.; Cohen, I. R., An antibody profile of systemic lupus erythematosus detected by antigen microarray. *Immunology* **2010,** *130* (3), 337-43.
94. Ayoglu, B.; Häggmark, A.; Khademi, M.; Olsson, T.; Uhlén, M.; Schwenk, J. M.; Nilsson, P., Autoantibody profiling in multiple sclerosis using arrays of human protein fragments. *Mol Cell Proteomics* **2013,** *12* (9), 2657-72.
95. Nagele, E.; Han, M.; Demarshall, C.; Belinka, B.; Nagele, R., Diagnosis of Alzheimer's disease based on disease-specific autoantibody profiles in human sera. *PLoS One* **2011,** *6* (8), e23112.
96. Casiano, C. A.; Mediavilla-Varela, M.; Tan, E. M., Tumor-associated antigen arrays for the serological diagnosis of cancer. *Mol Cell Proteomics* **2006,** *5* (10), 1745-59.
97. Stave, J. W.; Lindpaintner, K., Antibody and antigen contact residues define epitope and paratope size and structure. *J Immunol* **2013,** *191* (3), 1428-35.
98. Platten, M.; Offringa, R., Cancer immunotherapy: exploiting neoepitopes. *Cell Res* **2015,** *25* (8), 887-8.
99. Sinmaz, N.; Nguyen, T.; Tea, F.; Dale, R. C.; Brilot, F., Mapping autoantigen epitopes: molecular insights into autoantibody-associated disorders of the nervous system. *J Neuroinflammation* **2016,** *13* (1), 219.
100. Graus, F.; Saiz, A.; Dalmau, J., Antibodies and neuronal autoimmune disorders of the CNS. *J Neurol* **2010,** *257* (4), 509-17.
101. Geysen, H. M.; Meloen, R. H.; Barteling, S. J., Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc Natl Acad Sci U S A* **1984,** *81* (13), 3998-4002.
102. Cunningham, B. C.; Wells, J. A., High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **1989,** *244* (4908), 1081-5.
103. Maier, R. H.; Maier, C. J.; Rid, R.; Hintner, H.; Bauer, J. W.; Onder, K., Epitope mapping of antibodies using a cell array-based polypeptide library. *J Biomol Screen* **2010,** *15* (4), 418-26.
104. Forsström, B.; Axnäs, B. B.; Stengele, K. P.; Bühler, J.; Albert, T. J.; Richmond, T. A.; Hu, F. J.; Nilsson, P.; Hudson, E. P.; Rockberg, J.; Uhlen, M., Proteome-wide epitope mapping of antibodies using ultra-dense peptide arrays. *Mol Cell Proteomics* **2014,** *13* (6), 1585-97.
105. Ekins, R. P., Multi-analyte immunoassay. *J Pharm Biomed Anal* **1989,** *7* (2), 155-68.
106. Paweletz, C. P.; Charboneau, L.; Bichsel, V. E.; Simone, N. L.; Chen, T.; Gillespie, J. W.; Emmert-Buck, M. R.; Roth, M. J.; Petricoin III, E. F.; Liotta, L. A., Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene* **2001,** *20* (16), 1981-9.
107. Tonkinson, J. L.; Stillman, B. A., Nitrocellulose: a tried and true polymer finds utility as a post-genomic substrate. *Front Biosci* **2002,** *7*, c1-12.
108. Rose, D., Microfluidic technologies and instrumentation for printing DNA microarrays. In Microarray Biochip Technology. . M. Schena,

ed. ed.; Eaton Publishing, Natick, Mass: 2000; pp 19-38.
109. Pin, E.; Federici, G.; Petricoin, E. F., Preparation and use of reverse protein microarrays. *Curr Protoc Protein Sci* **2014,** *75*, Unit 27.7.
110. Pierobon, M.; Vanmeter, A. J.; Moroni, N.; Galdi, F.; Petricoin, E. F., Reverse-phase protein microarrays. *Methods Mol Biol* **2012,** *823*, 215-35.

111. Silvestri, A.; Colombatti, A.; Calvert, V. S.; Deng, J.; Mammano, E.; Belluco, C.; De Marchi, F.; Nitti, D.; Liotta, L. A.; Petricoin, E. F.; Pierobon, M., Protein pathway biomarker analysis of human cancer reveals requirement for upfront cellular-enrichment processing. *Lab Invest* **2010,** *90* (5), 787-96.
112. Baldelli, E.; Haura, E. B.; Crinò, L.; Cress, D. W.; Ludovini, V.; Schabath, M. B.; Liotta, L. A.; Petricoin, E. F.; Pierobon, M., Impact of upfront cellular enrichment by laser capture microdissection on protein and phosphoprotein drug target signaling activation measurements in human lung cancer: Implications for personalized medicine. *Proteomics Clin Appl* **2015,** *9* (9-10), 928-37.
113. Chiechi, A.; Novello, C.; Magagnoli, G.; Petricoin, E. F.; Deng, J.; Benassi, M. S.; Picci, P.; Vaisman, I.; Espina, V.; Liotta, L. A., Elevated TNFR1 and serotonin in bone metastasis are correlated with poor survival following bone metastasis diagnosis for both carcinoma and sarcoma primary tumors. *Clin Cancer Res* **2013,** *19* (9), 2473-85.
114. Mazzone, M.; Selfors, L. M.; Albeck, J.; Overholtzer, M.; Sale, S.; Carroll, D. L.; Pandya, D.; Lu, Y.; Mills, G. B.; Aster, J. C.; Artavanis-Tsakonas, S.; Brugge, J. S., Dose-dependent induction of distinct phenotypic responses to Notch pathway activation in mammary epithelial cells. *Proc Natl Acad Sci U S A* **2010,** *107* (11), 5012-7.
115. Federici, G.; Gao, X.; Slawek, J.; Arodz, T.; Shitaye, A.; Wulfkuhle, J. D.; De Maria, R.; Liotta, L. A.; Petricoin, E. F., Systems analysis of the NCI-60 cancer cell lines by alignment of protein pathway activation modules with "-OMIC" data fields and therapeutic response signatures. *Mol Cancer Res* **2013,** *11* (6), 676-85.
116. Grote, T.; Siwak, D. R.; Fritsche, H. A.; Joy, C.; Mills, G. B.; Simeone, D.; Whitcomb, D. C.; Logsdon, C. D., Validation of reverse phase protein array for practical screening of potential biomarkers in serum and plasma: accurate detection of CA19-9 levels in pancreatic cancer. *Proteomics* **2008,** *8* (15), 3051-60.
117. Gyorgy, A. B.; Walker, J.; Wingo, D.; Eidelman, O.; Pollard, H. B.; Molnar, A.; Agoston, D. V., Reverse phase protein microarray technology in traumatic brain injury. *J Neurosci Methods* **2010,** *192* (1), 96-101.
118. Gallagher, R. I.; Silvestri, A.; Petricoin, E. F.; Liotta, L. A.; Espina, V., Reverse phase protein microarrays: fluorometric and colorimetric detection. *Methods Mol Biol* **2011,** *723*, 275-301.
119. Wachter, A.; Bernhardt, S.; Beissbarth, T.; Korf, U., Analysis of Reverse Phase Protein Array Data: From Experimental Design towards Targeted Biomarker Discovery. *Microarrays (Basel)* **2015,** *4* (4), 520-39.
120. Mueller, C.; Liotta, L. A.; Espina, V., Reverse phase protein microarrays advance to use in clinical trials. *Mol Oncol* **2010,** *4* (6), 461-81.
121. Baldelli, E.; Bellezza, G.; Haura, E. B.; Crinó, L.; Cress, W. D.; Deng, J.; Ludovini, V.; Sidoni, A.; Schabath, M. B.; Puma, F.; Vannucci, J.; Siggillino, A.; Liotta, L. A.; Petricoin, E. F.; Pierobon, M., Functional signaling pathway analysis of lung adenocarcinomas identifies novel therapeutic targets for KRAS mutant tumors. *Oncotarget* **2015,** *6* (32), 32368-79.
122. Irwin, M. E.; Johnson, B. P.; Manshouri, R.; Amin, H. M.; Chandra, J., A NOX2/Egr-1/Fyn pathway delineates new targets for TKI-resistant malignancies. *Oncotarget* **2015,** *6* (27), 23631-46.
123. Wulfkuhle, J. D.; Speer, R.; Pierobon, M.; Laird, J.; Espina, V.; Deng, J.; Mammano, E.; Yang, S. X.; Swain, S. M.; Nitti, D.; Esserman, L. J.; Belluco, C.; Liotta, L. A.; Petricoin, E. F., Multiplexed cell signaling analysis of human breast cancer applications for personalized therapy. *J Proteome Res* **2008,** *7* (4), 1508-17.
124. Colarossi, L.; Memeo, L.; Colarossi, C.; Aiello, E.; Iuppa, A.; Espina, V.; Liotta, L.; Mueller, C., Inhibition of histone deacetylase 4 increases cytotoxicity of docetaxel in gastric cancer cells. *Proteomics Clin Appl* **2014,** *8* (11-12), 924-31.
125. Lai, C. H.; Park, K. S.; Lee, D. H.; Alberobello, A. T.; Raffeld, M.; Pierobon, M.; Pin, E.; Petricoin Iii, E. F.; Wang, Y.; Giaccone, G., HSP-90 inhibitor ganetespib is synergistic with doxorubicin in small cell lung cancer. *Oncogene* **2014,** *33* (40), 4867-76.
126. Wulfkuhle, J. D.; Berg, D.; Wolff, C.; Langer, R.; Tran, K.; Illi, J.; Espina, V.; Pierobon, M.; Deng, J.; DeMichele, A.; Walch, A.; Bronger, H.; Becker, I.; Waldhör, C.; Höfler, H.; Esserman, L.; Liotta, L. A.; Becker, K. F.; Petricoin, E. F.; Investigators, I.-S. T., Molecular analysis of HER2 signaling in human breast cancer by functional protein pathway activation mapping. *Clin Cancer Res* **2012,** *18* (23), 6426-35.
127. Barker, A. D.; Sigman, C. C.; Kelloff, G. J.; Hylton, N. M.; Berry, D. A.; Esserman, L. J., I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin Pharmacol Ther* **2009,** *86* (1), 97-100.

128. Jameson, G. S.; Petricoin, E. F.; Sachdev, J.; Liotta, L. A.; Loesch, D. M.; Anthony, S. P.; Chadha, M. K.; Wulfkuhle, J. D.; Gallagher, R. I.; Reeder, K. A.; Pierobon, M.; Fulk, M. R.; Cantafio, N. A.; Dunetz, B.; Mikrut, W. D.; Von Hoff, D. D.; Robert, N. J., A pilot study utilizing multi-omic molecular profiling to find potential targets and select individualized treatments for patients with previously treated metastatic breast cancer. *Breast Cancer Res Treat* **2014,** *147* (3), 579-88.
129. Diamandis, E. P., Prostate-specific Antigen: Its Usefulness in Clinical Medicine. *Trends Endocrinol Metab* **1998,** *9* (8), 310-6.
130. McNaughton-Collins, M. F.; Barry, M. J., One man at a time--resolving the PSA controversy. *N Engl J Med* **2011,** *365* (21), 1951-3.
131. Pin, E.; Fredolini, C.; Petricoin, E. F., The role of proteomics in prostate cancer research: biomarker discovery and validation. *Clin Biochem* **2013,** *46* (6), 524-38.
132. Roobol, M. J.; Haese, A.; Bjartell, A., Tumour markers in prostate cancer III: biomarkers in urine. *Acta Oncol* **2011,** *50 Suppl 1*, 85-9.
133. Jia, X.; Chen, J.; Sun, S.; Yang, W.; Yang, S.; Shah, P.; Hoti, N.; Veltri, B.; Zhang, H., Detection of Aggressive Prostate Cancer-Associated Glycoproteins in Urine using Glycoproteomics and Mass Spectrometry. *Proteomics* **2016**.
134. Drabovich, A. P.; Saraon, P.; Jarvi, K.; Diamandis, E. P., Seminal plasma as a diagnostic fluid for male reproductive system disorders. *Nat Rev Urol* **2014,** *11* (5), 278-88.
135. Wang, J.; Cai, Y.; Ren, C.; Ittmann, M., Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. *Cancer Res* **2006,** *66* (17), 8347-51.
136. Mosquera, J. M.; Perner, S.; Genega, E. M.; Sanda, M.; Hofer, M. D.; Mertz, K. D.; Paris, P. L.; Simko, J.; Bismar, T. A.; Ayala, G.; Shah, R. B.; Loda, M.; Rubin, M. A., Characterization of TMPRSS2-ERG fusion high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin Cancer Res* **2008,** *14* (11), 3380-5.
137. Brinkmann, A. O.; Blok, L. J.; de Ruiter, P. E.; Doesburg, P.; Steketee, K.; Berrevoets, C. A.; Trapman, J., Mechanisms of androgen receptor activation and function. *J Steroid Biochem Mol Biol* **1999,** *69* (1-6), 307-13.
138. Gottlieb, B.; Beitel, L. K.; Wu, J. H.; Trifiro, M., The androgen receptor gene mutations database (ARDB): 2004 update. *Hum Mutat* **2004,** *23* (6), 527-33.
139. Paliouras, M.; Zaman, N.; Lumbroso, R.; Kapogeorgakis, L.; Beitel, L. K.; Wang, E.; Trifiro, M., Dynamic rewiring of the androgen receptor protein interaction network correlates with prostate cancer clinical outcomes. *Integr Biol (Camb)* **2011,** *3* (10), 1020-32.
140. Craft, N.; Shostak, Y.; Carey, M.; Sawyers, C. L., A mechanism for hormone-independent prostate cancer through modulation of androgen receptor signaling by the HER-2/neu tyrosine kinase. *Nat Med* **1999,** *5* (3), 280-5.
141. Feldman, B. J.; Feldman, D., The development of androgen-independent prostate cancer. *Nat Rev Cancer* **2001,** *1* (1), 34-45.
142. Nazareth, L. V.; Weigel, N. L., Activation of the human androgen receptor through a protein kinase A signaling pathway. *J Biol Chem* **1996,** *271* (33), 19900-7.
143. Chen, T.; Wang, L. H.; Farrar, W. L., Interleukin 6 activates androgen receptor-mediated gene expression through a signal transducer and activator of transcription 3-dependent pathway in LNCaP prostate cancer cells. *Cancer Res* **2000,** *60* (8), 2132-5.
144. Paweletz, C. P.; Liotta, L. A.; Petricoin, E. F., New technologies for biomarker analysis of prostate cancer progression: Laser capture microdissection and tissue proteomics. *Urology* **2001,** *57* (4 Suppl 1), 160-3.
145. Zheng, Y.; Xu, Y.; Ye, B.; Lei, J.; Weinstein, M. H.; O'Leary, M. P.; Richie, J. P.; Mok, S. C.; Liu, B. C., Prostate carcinoma tissue proteomics for biomarker discovery. *Cancer* **2003,** *98* (12), 2576-82.
146. Alaiya, A. A.; Al-Mohanna, M.; Aslam, M.; Shinwari, Z.; Al-Mansouri, L.; Al-Rodayan, M.; Al-Eid, M.; Ahmad, I.; Hanash, K.; Tulbah, A.; Bin Mahfooz, A.; Adra, C., Proteomics-based signature for human benign prostate hyperplasia and prostate adenocarcinoma. *Int J Oncol* **2011,** *38* (4), 1047-57.
147. Sun, C.; Zhao, X.; Xu, K.; Gong, J.; Liu, W.; Ding, W.; Gou, Y.; Xia, G.; Ding, Q., Periostin: a promising target of therapeutical intervention for prostate cancer. *J Transl Med* **2011,** *9*, 99.
148. Ummanni, R.; Mundt, F.; Pospisil, H.; Venz, S.; Scharf, C.; Barett, C.; Fälth, M.; Köllermann, J.; Walther, R.; Schlomm, T.; Sauter, G.; Bokemeyer, C.; Sültmann, H.; Schuppert, A.; Brümmendorf, T. H.; Balabanov, S., Identification of clinically relevant protein targets in prostate cancer with 2D-DIGE coupled mass spectrometry and systems biology network platform. *PLoS One* **2011,** *6* (2), e16833.

149. Montironi, R.; Mazzucchelli, R.; Lopez-Beltran, A.; Cheng, L.; Scarpelli, M., Mechanisms of disease: high-grade prostatic intraepithelial neoplasia and other proposed preneoplastic lesions in the prostate. *Nat Clin Pract Urol* **2007,** *4* (6), 321-32.
150. Cheung, P. K.; Woolcock, B.; Adomat, H.; Sutcliffe, M.; Bainbridge, T. C.; Jones, E. C.; Webber, D.; Kinahan, T.; Sadar, M.; Gleave, M. E.; Vielkind, J., Protein profiling of microdissected prostate tissue links growth differentiation factor 15 to prostate carcinogenesis. *Cancer Res* **2004,** *64* (17), 5929-33.
151. Liu, A.; Wei, L.; Gardner, W. A.; Deng, C. X.; Man, Y. G., Correlated alterations in prostate basal cell layer and basement membrane. *Int J Biol Sci* **2009,** *5* (3), 276-85.
152. Diaz, J. I.; Cazares, L. H.; Corica, A.; John Semmes, O., Selective capture of prostatic basal cells and secretory epithelial cells for proteomic and genomic analysis. *Urol Oncol* **2004,** *22* (4), 329-36.
153. Khamis, Z. I.; Iczkowski, K. A.; Sahab, Z. J.; Sang, Q. X., Protein profiling of isolated leukocytes, myofibroblasts, epithelial, Basal, and endothelial cells from normal, hyperplastic, cancerous, and inflammatory human prostate tissues. *J Cancer* **2010,** *1*, 70-9.
154. Skvortsov, S.; Schäfer, G.; Stasyk, T.; Fuchsberger, C.; Bonn, G. K.; Bartsch, G.; Klocker, H.; Huber, L. A., Proteomics profiling of microdissected low- and high-grade prostate tumors identifies Lamin A as a discriminatory biomarker. *J Proteome Res* **2011,** *10* (1), 259-68.
155. Zhang, S.; Wang, X.; Osunkoya, A. O.; Iqbal, S.; Wang, Y.; Chen, Z.; Müller, S.; Josson, S.; Coleman, I. M.; Nelson, P. S.; Wang, Y. A.; Wang, R.; Shin, D. M.; Marshall, F. F.; Kucuk, O.; Chung, L. W.; Zhau, H. E.; Wu, D., EPLIN downregulation promotes epithelial-mesenchymal transition in prostate cancer cells and correlates with clinical lymph node metastasis. *Oncogene* **2011,** *30* (50), 4941-52.
156. Gao, X.; Pang, J.; Li, L. Y.; Liu, W. P.; Di, J. M.; Sun, Q. P.; Fang, Y. Q.; Liu, X. P.; Pu, X. Y.; He, D.; Li, M. T.; Su, Z. L.; Li, B. Y., Expression profiling identifies new function of collapsin response mediator protein 4 as a metastasis-suppressor in prostate cancer. *Oncogene* **2010,** *29* (32), 4555-66.
157. Pang, J.; Liu, W. P.; Liu, X. P.; Li, L. Y.; Fang, Y. Q.; Sun, Q. P.; Liu, S. J.; Li, M. T.; Su, Z. L.; Gao, X., Profiling protein markers associated with lymph node metastasis in prostate cancer by DIGE-based proteomics analysis. *J Proteome Res* **2010,** *9* (1), 216-26.
158. Saraon, P.; Jarvi, K.; Diamandis, E. P., Molecular alterations during progression of prostate cancer to androgen independence. *Clin Chem* **2011,** *57* (10), 1366-75.
159. Khan, A. P.; Poisson, L. M.; Bhat, V. B.; Fermin, D.; Zhao, R.; Kalyana-Sundaram, S.; Michailidis, G.; Nesvizhskii, A. I.; Omenn, G. S.; Chinnaiyan, A. M.; Sreekumar, A., Quantitative proteomic profiling of prostate cancer reveals a role for miR-128 in prostate cancer. *Mol Cell Proteomics* **2010,** *9* (2), 298-312.
160. Galardi, S.; Mercatelli, N.; Giorda, E.; Massalini, S.; Frajese, G. V.; Ciafrè, S. A.; Farace, M. G., miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting p27Kip1. *J Biol Chem* **2007,** *282* (32), 23716-24.
161. Sun, T.; Wang, Q.; Balk, S.; Brown, M.; Lee, G. S.; Kantoff, P., The role of microRNA-221 and microRNA-222 in androgen-independent prostate cancer cell lines. *Cancer Res* **2009,** *69* (8), 3356-63.
162. Butler, G. S.; Overall, C. M., Proteomic identification of multitasking proteins in unexpected locations complicates drug targeting. *Nat Rev Drug Discov* **2009,** *8* (12), 935-48.
163. Theodorescu, D.; Schiffer, E.; Bauer, H. W.; Douwes, F.; Eichhorn, F.; Polley, R.; Schmidt, T.; Schöfer, W.; Zürbig, P.; Good, D. M.; Coon, J. J.; Mischak, H., Discovery and validation of urinary biomarkers for prostate cancer. *Proteomics Clin Appl* **2008,** *2* (4), 556-570.
164. Schiffer, E.; Bick, C.; Grizelj, B.; Pietzker, S.; Schöfer, W., Urinary proteome analysis for prostate cancer diagnosis: cost-effective application in routine clinical practice in Germany. *Int J Urol* **2012,** *19* (2), 118-25.
165. Miller, J. C.; Zhou, H.; Kwekel, J.; Cavallo, R.; Burke, J.; Butler, E. B.; Teh, B. S.; Haab, B. B., Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers. *Proteomics* **2003,** *3* (1), 56-63.
166. Mink, S. R.; Hodge, A.; Agus, D. B.; Jain, A.; Gross, M. E., Beta-2-microglobulin expression correlates with high-grade prostate cancer and specific defects in androgen signaling. *Prostate* **2010,** *70* (11), 1201-10.
167. Gross, M.; Top, I.; Laux, I.; Katz, J.; Curran, J.; Tindell, C.; Agus, D., Beta-2-microglobulin is an androgen-regulated secreted protein elevated in serum of patients with advanced prostate cancer. *Clin Cancer Res* **2007,** *13* (7), 1979-86.
168. Rehman, I.; Evans, C. A.; Glen, A.; Cross, S. S.; Eaton, C. L.; Down, J.; Pesce, G.; Phillips, J. T.; Yen, O. S.; Thalmann, G. N.; Wright, P. C.; Hamdy, F. C., iTRAQ identification of candidate serum biomarkers associated with metastatic progression of human prostate cancer. *PLoS One* **2012,** *7* (2), e30885.

169. Katafigiotis, I.; Tyritzis, S. I.; Stravodimos, K. G.; Alamanis, C.; Pavlakis, K.; Vlahou, A.; Makridakis, M.; Katafigioti, A.; Garbis, S. D.; Constantinides, C. A., Zinc α2-glycoprotein as a potential novel urine biomarker for the early diagnosis of prostate cancer. *BJU Int* **2012,** *110* (11 Pt B), E688-93.
170. Henshall, S. M.; Horvath, L. G.; Quinn, D. I.; Eggleton, S. A.; Grygiel, J. J.; Stricker, P. D.; Biankin, A. V.; Kench, J. G.; Sutherland, R. L., Zinc-alpha2-glycoprotein expression as a predictor of metastatic prostate cancer following radical prostatectomy. *J Natl Cancer Inst* **2006,** *98* (19), 1420-4.
171. Shariat, S. F.; Kim, J. H.; Andrews, B.; Kattan, M. W.; Wheeler, T. M.; Kim, I. Y.; Lerner, S. P.; Slawin, K. M., Preoperative plasma levels of transforming growth factor beta(1) strongly predict clinical outcome in patients with bladder carcinoma. *Cancer* **2001,** *92* (12), 2985-92.
172. Shariat, S. F.; Kattan, M. W.; Traxel, E.; Andrews, B.; Zhu, K.; Wheeler, T. M.; Slawin, K. M., Association of pre- and postoperative plasma levels of transforming growth factor beta(1) and interleukin 6 and its soluble receptor with prostate cancer progression. *Clin Cancer Res* **2004,** *10* (6), 1992-9.
173. Michalaki, V.; Syrigos, K.; Charles, P.; Waxman, J., Serum levels of IL-6 and TNF-alpha correlate with clinicopathological features and patient survival in patients with prostate cancer. *Br J Cancer* **2004,** *90* (12), 2312-6.
174. True, L. D.; Zhang, H.; Ye, M.; Huang, C. Y.; Nelson, P. S.; von Haller, P. D.; Tjoelker, L. W.; Kim, J. S.; Qian, W. J.; Smith, R. D.; Ellis, W. J.; Liebeskind, E. S.; Liu, A. Y., CD90/THY1 is overexpressed in prostate cancer-associated fibroblasts and could serve as a cancer biomarker. *Mod Pathol* **2010,** *23* (10), 1346-56.
175. Morgan, R.; Boxall, A.; Bhatt, A.; Bailey, M.; Hindley, R.; Langley, S.; Whitaker, H. C.; Neal, D. E.; Ismail, M.; Whitaker, H.; Annels, N.; Michael, A.; Pandha, H., Engrailed-2 (EN2): a tumor specific urinary biomarker for the early diagnosis of prostate cancer. *Clin Cancer Res* **2011,** *17* (5), 1090-8.
176. Drake, R. R.; Elschenbroich, S.; Lopez-Perez, O.; Kim, Y.; Ignatchenko, V.; Ignatchenko, A.; Nyalwidhe, J. O.; Basu, G.; Wilkins, C. E.; Gjurich, B.; Lance, R. S.; Semmes, O. J.; Medin, J. A.; Kislinger, T., In-depth proteomic analyses of direct expressed prostatic secretions. *J Proteome Res* **2010,** *9* (5), 2109-16.
177. Josson, S.; Nomura, T.; Lin, J. T.; Huang, W. C.; Wu, D.; Zhau, H. E.; Zayzafoon, M.; Weizmann, M. N.; Gururajan, M.; Chung, L. W., β2-microglobulin induces epithelial to mesenchymal transition and confers cancer lethality and bone metastasis in human cancer cells. *Cancer Res* **2011,** *71* (7), 2600-10.
178. Hassan, M. I.; Kumar, V.; Kashav, T.; Alam, N.; Singh, T. P.; Yadav, S., Proteomic approach for purification of seminal plasma proteins involved in tumor proliferation. *J Sep Sci* **2007,** *30* (12), 1979-88.
179. Byrne, J. C.; Downes, M. R.; O'Donoghue, N.; O'Keane, C.; O'Neill, A.; Fan, Y.; Fitzpatrick, J. M.; Dunn, M.; Watson, R. W., 2D-DIGE as a strategy to identify serum markers for the progression of prostate cancer. *J Proteome Res* **2009,** *8* (2), 942-57.
180. Halabi, S.; Small, E. J.; Kantoff, P. W.; Kattan, M. W.; Kaplan, E. B.; Dawson, N. A.; Levine, E. G.; Blumenstein, B. A.; Vogelzang, N. J., Prognostic model for predicting survival in men with hormone-refractory metastatic prostate cancer. *J Clin Oncol* **2003,** *21* (7), 1232-7.
181. Prior, C.; Guillen-Grima, F.; Robles, J. E.; Rosell, D.; Fernandez-Montero, J. M.; Agirre, X.; Catena, R.; Calvo, A., Use of a combination of biomarkers in serum and urine to improve detection of prostate cancer. *World J Urol* **2010,** *28* (6), 681-6.
182. Zijlstra, C.; Stoorvogel, W., Prostasomes as a source of diagnostic biomarkers for prostate cancer. *J Clin Invest* **2016,** *126* (4), 1144-51.
183. Principe, S.; Jones, E. E.; Kim, Y.; Sinha, A.; Nyalwidhe, J. O.; Brooks, J.; Semmes, O. J.; Troyer, D. A.; Lance, R. S.; Kislinger, T.; Drake, R. R., In-depth proteomic analyses of exosomes isolated from expressed prostatic secretions in urine. *Proteomics* **2013,** *13* (10-11), 1667-71.
184. Gonzales, P. A.; Zhou, H.; Pisitkun, T.; Wang, N. S.; Star, R. A.; Knepper, M. A.; Yuen, P. S., Isolation and purification of exosomes in urine. *Methods Mol Biol* **2010,** *641*, 89-99.
185. Gonzales, P. A.; Pisitkun, T.; Hoffert, J. D.; Tchapyjnikov, D.; Star, R. A.; Kleta, R.; Wang, N. S.; Knepper, M. A., Large-scale proteomics and phosphoproteomics of urinary exosomes. *J Am Soc Nephrol* **2009,** *20* (2), 363-79.
186. Øverbye, A.; Skotland, T.; Koehler, C. J.; Thiede, B.; Seierstad, T.; Berge, V.; Sandvig, K.; Llorente, A., Identification of prostate cancer biomarkers in urinary exosomes. *Oncotarget* **2015,** *6* (30), 30357-76.
187. Wang, L.; Skotland, T.; Berge, V.; Sandvig, K.; Llorente, A., Exosomal proteins as prostate cancer biomarkers in urine: From mass spectrometry discovery to immunoassay-based validation. *Eur J Pharm Sci* **2016**.

188. Teng, M. W.; Smyth, M. J., Cancer. Can cancer trigger autoimmunity? *Science* **2014,** *343* (6167), 147-8.
189. Wang, D.; Yang, L.; Zhang, P.; LaBaer, J.; Hermjakob, H.; Li, D.; Yu, X., AAgAtlas 1.0: a human autoantigen database. *Nucleic Acids Res* **2016**.
190. Klinman, D. M.; Steinberg, A. D., Systemic autoimmune disease arises from polyclonal B cell activation. *J Exp Med* **1987,** *165* (6), 1755-60.
191. Massoner, P.; Lueking, A.; Goehler, H.; Höpfner, A.; Kowald, A.; Kugler, K. G.; Amersdorfer, P.; Horninger, W.; Bartsch, G.; Schulz-Knappe, P.; Klocker, H., Serum-autoantibodies for discovery of prostate cancer specific biomarkers. *Prostate* **2012,** *72* (4), 427-36.
192. Wang, X.; Yu, J.; Sreekumar, A.; Varambally, S.; Shen, R.; Giacherio, D.; Mehra, R.; Montie, J. E.; Pienta, K. J.; Sanda, M. G.; Kantoff, P. W.; Rubin, M. A.; Wei, J. T.; Ghosh, D.; Chinnaiyan, A. M., Autoantibody signatures in prostate cancer. *N Engl J Med* **2005,** *353* (12), 1224-35.
193. Dai, L.; Li, J.; Ortega, R.; Qian, W.; Casiano, C. A.; Zhang, J. Y., Preferential autoimmune response in prostate cancer to cyclin B1 in a panel of tumor-associated antigens. *J Immunol Res* **2014,** *2014*, 827827.
194. McNeel, D. G.; Nguyen, L. D.; Storer, B. E.; Vessella, R.; Lange, P. H.; Disis, M. L., Antibody immunity to prostate cancer associated antigens can be detected in the serum of patients with prostate cancer. *J Urol* **2000,** *164* (5), 1825-9.
195. Mintz, P. J.; Rietz, A. C.; Cardó-Vila, M.; Ozawa, M. G.; Dondossola, E.; Do, K. A.; Kim, J.; Troncoso, P.; Logothetis, C. J.; Sidman, R. L.; Pasqualini, R.; Arap, W., Discovery and horizontal follow-up of an autoantibody signature in human prostate cancer. *Proc Natl Acad Sci U S A* **2015,** *112* (8), 2515-20.
196. Johnson, L. D.; Nesslinger, N. J.; Blood, P. A.; Chima, N.; Richier, L. R.; Ludgate, C.; Pai, H. H.; Lim, J. T.; Nelson, B. H.; Vlachaki, M. T.; Lum, J. J., Tumor-associated autoantibodies correlate with poor outcome in prostate cancer patients treated with androgen deprivation and external beam radiation therapy. *Oncoimmunology* **2014,** *3*, e29243.
197. Tchernychev, B.; Cabilly, S.; Wilchek, M., The epitopes for natural polyreactive antibodies are rich in proline. *Proc Natl Acad Sci U S A* **1997,** *94* (12), 6335-9.
198. Zhou, Z. H.; Tzioufas, A. G.; Notkins, A. L., Properties and function of polyreactive antibodies and polyreactive antigen-binding B cells. *J Autoimmun* **2007,** *29* (4), 219-28.
199. Nolen, B. M.; Lokshin, A. E., Autoantibodies for cancer detection: still cause for excitement? *Cancer Biomark* **2010,** *6* (5-6), 229-45.
200. Poste, G., Bring on the biomarkers. *Nature* **2011,** *469* (7329), 156-7.
201. DiMasi, J. A.; Feldman, L.; Seckler, A.; Wilson, A., Trends in risks associated with new drug development: success rates for investigational drugs. *Clin Pharmacol Ther* **2010,** *87* (3), 272-7.
202. Arrowsmith, J., Trial watch: Phase II failures: 2008-2010. *Nat Rev Drug Discov* **2011,** *10* (5), 328-9.
203. Arrowsmith, J., Trial watch: phase III and submission failures: 2007-2010. *Nat Rev Drug Discov* **2011,** *10* (2), 87.
204. Jones, D., Biomarker debate highlights retrospective challenge. *Nat Rev Drug Discov* **2009,** *8* (3), 179-80.
205. Allegra, C. J.; Jessup, J. M.; Somerfield, M. R.; Hamilton, S. R.; Hammond, E. H.; Hayes, D. F.; McAllister, P. K.; Morton, R. F.; Schilsky, R. L., American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy. *J Clin Oncol* **2009,** *27* (12), 2091-6.
206. Gilbert, P. B.; Janes, H. E.; Huang, Y., Power/sample size calculations for assessing correlates of risk in clinical efficacy trials. *Stat Med* **2016,** *35* (21), 3745-59.
207. Chen, D. T.; Huang, P. Y.; Lin, H. Y.; Haura, E. B.; Antonia, S. J.; Cress, W. D.; Gray, J. E., Strategies for power calculations in predictive biomarker studies in survival data. *Oncotarget* **2016**.
208. Espina, V.; Mueller, C.; Edmiston, K.; Sciro, M.; Petricoin, E. F.; Liotta, L. A., Tissue is alive: New technologies are needed to address the problems of protein biomarker pre-analytical variability. *Proteomics Clin Appl* **2009,** *3* (8), 874-882.
209. Spruessel, A.; Steimann, G.; Jung, M.; Lee, S. A.; Carr, T.; Fentz, A. K.; Spangenberg, J.; Zornig, C.; Juhl, H. H.; David, K. A., Tissue ischemia time affects gene and protein expression patterns within minutes following surgical tumor excision. *Biotechniques* **2004,** *36* (6), 1030-7.
210. Espina, V.; Edmiston, K. H.; Heiby, M.; Pierobon, M.; Sciro, M.; Merritt, B.; Banks, S.; Deng, J.; VanMeter, A. J.; Geho, D. H.; Pastore, L.; Sennesh, J.; Petricoin, E. F.; Liotta, L. A., A portrait of tissue

phosphoprotein stability in the clinical tissue procurement process. *Mol Cell Proteomics* **2008,** *7* (10), 1998-2018.
211. Lim, M. D.; Dickherber, A.; Compton, C. C., Before you analyze a human specimen, think quality, variability, and bias. *Anal Chem* **2011,** *83* (1), 8-13.
212. Nassiri, M.; Ramos, S.; Zohourian, H.; Vincek, V.; Morales, A. R.; Nadji, M., Preservation of biomolecules in breast cancer tissue by a formalin-free histology system. *BMC Clin Pathol* **2008,** *8*, 1.
213. Mueller, C.; Edmiston, K. H.; Carpenter, C.; Gaffney, E.; Ryan, C.; Ward, R.; White, S.; Memeo, L.; Colarossi, C.; Petricoin, E. F.; Liotta, L. A.; Espina, V., One-step preservation of phosphoproteins and tissue morphology at room temperature for diagnostic and research specimens. *PLoS One* **2011,** *6* (8), e23780.
214. Espina, V.; Mueller, C., Reduction of preanalytical variability in specimen procurement for molecular profiling. *Methods Mol Biol* **2012,** *823*, 49-57.
215. Ostroff, R.; Foreman, T.; Keeney, T. R.; Stratford, S.; Walker, J. J.; Zichi, D., The stability of the circulating human proteome to variations in sample collection and handling procedures measured with an aptamer-based proteomics array. *J Proteomics* **2010,** *73* (3), 649-66.
216. Qundos, U.; Hong, M. G.; Tybring, G.; Divers, M.; Odeberg, J.; Uhlen, M.; Nilsson, P.; Schwenk, J. M., Profiling post-centrifugation delay of serum and plasma with antibody bead arrays. *J Proteomics* **2013,** *95*, 46-54.
217. Ayache, S.; Panelli, M.; Marincola, F. M.; Stroncek, D. F., Effects of storage time and exogenous protease inhibitors on plasma protein levels. *Am J Clin Pathol* **2006,** *126* (2), 174-84.
218. Tort, A. B.; Dietrich, M. O.; Gonçalves, C. A.; Souza, D. O.; Portela, L. V., Influence of anticoagulants on the measurement of S100B protein in blood. *Clin Biochem* **2003,** *36* (8), 629-32.
219. Bowen, R. A.; Remaley, A. T., Interferences from blood collection tube components on clinical chemistry assays. *Biochem Med (Zagreb)* **2014,** *24* (1), 31-44.
220. Mitchell, B. L.; Yasui, Y.; Li, C. I.; Fitzpatrick, A. L.; Lampe, P. D., Impact of freeze-thaw cycles and storage time on plasma samples used in mass spectrometry based biomarker discovery projects. *Cancer Inform* **2005,** *1*, 98-104.
221. Lee, J. E.; Kim, S. Y.; Shin, S. Y., Effect of Repeated Freezing and Thawing on Biomarker Stability in Plasma and Serum Samples. *Osong Public Health Res Perspect* **2015,** *6* (6), 357-62.
222. Espina, V.; Wulfkuhle, J.; Liotta, L. A., Application of laser microdissection and reverse-phase protein microarrays to the molecular profiling of cancer signal pathway networks in the tissue microenvironment. *Clin Lab Med* **2009,** *29* (1), 1-13.
223. Liu, A., Laser capture microdissection in the tissue biorepository. *J Biomol Tech* **2010,** *21* (3), 120-5.
224. Xu, B. J., Combining laser capture microdissection and proteomics: methodologies and clinical applications. *Proteomics Clin Appl* **2010,** *4* (2), 116-23.
225. Abbatiello, S. E.; Schilling, B.; Mani, D. R.; Zimmerman, L. J.; Hall, S. C.; MacLean, B.; Albertolle, M.; Allen, S.; Burgess, M.; Cusack, M. P.; Gosh, M.; Hedrick, V.; Held, J. M.; Inerowicz, H. D.; Jackson, A.; Keshishian, H.; Kinsinger, C. R.; Lyssand, J.; Makowski, L.; Mesri, M.; Rodriguez, H.; Rudnick, P.; Sadowski, P.; Sedransk, N.; Shaddox, K.; Skates, S. J.; Kuhn, E.; Smith, D.; Whiteaker, J. R.; Whitwell, C.; Zhang, S.; Borchers, C. H.; Fisher, S. J.; Gibson, B. W.; Liebler, D. C.; MacCoss, M. J.; Neubert, T. A.; Paulovich, A. G.; Regnier, F. E.; Tempst, P.; Carr, S. A., Large-Scale Interlaboratory Study to Develop, Analytically Validate and Apply Highly Multiplexed, Quantitative Peptide Assays to Measure Cancer-Relevant Proteins in Plasma. *Mol Cell Proteomics* **2015,** *14* (9), 2357-74.
226. Pernemalm, M.; Lewensohn, R.; Lehtiö, J., Affinity prefractionation for MS-based plasma proteomics. *Proteomics* **2009,** *9* (6), 1420-7.
227. Bellei, E.; Bergamini, S.; Monari, E.; Fantoni, L. I.; Cuoghi, A.; Ozben, T.; Tomasi, A., High-abundance proteins depletion for serum proteomic analysis: concomitant removal of non-targeted proteins. *Amino Acids* **2011,** *40* (1), 145-56.
228. Granger, J.; Siddiqui, J.; Copeland, S.; Remick, D., Albumin depletion of human plasma also removes low abundance proteins including the cytokines. *Proteomics* **2005,** *5* (18), 4713-8.
229. Tamburro, D.; Fredolini, C.; Espina, V.; Douglas, T. A.; Ranganathan, A.; Ilag, L.; Zhou, W.; Russo, P.; Espina, B. H.; Muto, G.; Petricoin, E. F.; Liotta, L. A.; Luchini, A., Multifunctional core-shell nanoparticles: discovery of previously invisible biomarkers. *J Am Chem Soc* **2011,** *133* (47), 19178-88.

230. Fredolini, C.; Tamburro, D.; Gambara, G.; Lepene, B. S.; Espina, V.; Petricoin, E. F.; Liotta, L. A.; Luchini, A., Nanoparticle technology: amplifying the effective sensitivity of biomarker detection to create a urine test for hGH. *Drug Test Anal* **2009,** *1* (9-10), 447-54.
231. Kulak, N. A.; Pichler, G.; Paron, I.; Nagaraj, N.; Mann, M., Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* **2014,** *11* (3), 319-24.
232. Infantino, M.; Manfredi, M.; Meacci, F.; Sarzi-Puttini, P.; Ricci, C.; Atzeni, F.; Benucci, M., Anti-citrullinated peptide antibodies and rheumatoid factor isotypes in the diagnosis of rheumatoid arthritis: an assessment of combined tests. *Clin Chim Acta* **2014,** *436*, 237-42.
233. Sturgeon, C., Perspectives in Clinical Proteomics Conference: translating clinical proteomics into clinical practice. *Expert Rev Proteomics* **2010,** *7* (4), 469-71.
234. Sturgeon, C.; Hill, R.; Hortin, G. L.; Thompson, D., Taking a new biomarker into routine use--a perspective from the routine clinical biochemistry laboratory. *Proteomics Clin Appl* **2010,** *4* (12), 892-903.
235. Ringnér, M., What is principal component analysis? *Nat Biotechnol* **2008,** *26* (3), 303-4.
236. Quackenbush, J., Microarray data normalization and transformation. *Nat Genet* **2002,** *32 Suppl*, 496-501.
237. Bolstad, B. M.; Irizarry, R. A.; Astrand, M.; Speed, T. P., A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **2003,** *19* (2), 185-93.
238. Huber, W.; von Heydebreck, A.; Sültmann, H.; Poustka, A.; Vingron, M., Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **2002,** *18 Suppl 1*, S96-104.
239. Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H., Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* **2006,** *78* (13), 4281-90.
240. Xie, Y.; Pan, W.; Khodursky, A. B., A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data. *Bioinformatics* **2005,** *21* (23), 4280-8.
241. Xu, Z.; Huang, L.; Zhang, H.; Li, Y.; Guo, S.; Wang, N.; Wang, S. H.; Chen, Z.; Wang, J.; Tao, S. C., PMD: A Resource for Archiving and Analyzing Protein Microarray data. *Sci Rep* **2016,** *6*, 19956.
242. Dakubo, G. D.; Jakupciak, J. P.; Birch-Machin, M. A.; Parr, R. L., Clinical implications and utility of field cancerization. *Cancer Cell Int* **2007,** *7*, 2.