Advances and Applications of Antibody Arrays - Proteomic Profiling of Pancreatic Disease

Sandström Gerdtsson, Anna

2013

Advances and applications of antibody arrays

Proteomic profiling of pancreatic disease

Anna S Gerdtsson

LUND
UNIVERSITY

ACADEMIC THESIS

Which, by due permission of the Faculty of Engineering at Lund University, will be publicly defended at Hörsalen, Medicon Village, Scheelevägen 2, Lund, Friday 6th of December 2013 at 09.15 am.

Faculty opponent is Prof. Marta Sanchez-Carbayo,
CIC bioGUNE, Derio, Spain

Organization	Document name
LUND UNIVERSITY Department of Immunotechnology Medicon Village (Building 406) S-223 81 LUND SWEDEN	DOCTORAL DISSERTATION Date of issue 2013-12-06
Author(s) Anna S Gerdtsson	Sponsoring organization

Title and subtitle

Advances and applications of antibody arrays - Proteomic profiling of pancreatic disease

Abstract

Recombinant antibody microarrays have advanced into indispensable tools for large-scale, high-throughput multiplexed serum proteomics. This thesis, based upon five original papers, deals with the development of an in-house designed antibody microarray platform, and its applications for serum profiling of pancreatic disease.

Pancreatic cancer is the 4th deadliest cancer, with a 5-year survival rate of only 6%. In order to increase the survival of this deadly disease, novel diagnostic biomarkers for earlier detection will be essential. In paper I and II of this thesis, we identified candidate biomarker signatures for predicting pancreatic cancer among healthy controls and pancreatitis. Pancreatitis (pancreatic inflammation) is symptomatically highly similar to pancreatic cancer, and biomarkers able to discriminate pancreatic cancer from pancreatitis would be of great clinical value. Pancreatitis appears in mainly chronic, acute, and autoimmune manifestations, and like for pancreatic cancer, there is a lack of high-performing biomarkers for diagnosis and stratification. In paper III, we applied antibody microarrays for pancreatitis protein profiling, and presented tentative biomarker signatures for the three main subtypes of this disease.

In parallel to performing clinical applications of the antibody microarrays, technical efforts for improving and expanding the use of the platform have also been conducted. In paper IV, we studied the impact of the antibody-surface interplay, and evaluated different solid supports for antibody microarray production. We also took the first steps towards developing a user-friendly ELISA-like multiplexed biomarker assay, by presenting the first plate-based recombinant array-in-well set-up. In paper V, we designed protocols for an increased utility of the antibody microarray platform, to comprise not only targeting of proteins, but also serum/plasma profiling of glycan and carbonyl groups. Post-translational modification of proteins, like glycosylation and carbonylation (oxidation) is often altered in disease, and biomarkers based on differentiated levels of these modifications may complement traditional protein biomarkers. Proof-of-concept was demonstrated for preeclampsia, a common pregnancy disorder, for which the results indicated that particularly the level of carbonylation could be used for diagnosis and stratification.

In conclusion, the work in this thesis has contributed to an improved and increased utility of the recombinant antibody microarray technology, and demonstrated its use for serum proteomic profiling of pancreatic disease.

Key words

Proteomics, affinity proteomics, antibody microarray, pancreatic cancer, pancreatitis, biomarker, immunosignature

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language English
ISSN and key title	ISBN 978-91-7473-748-6
Recipient's notes	Number of pages 161 Price Security classification

Signature _____ Date _____

Advances and applications of antibody arrays

Proteomic profiling of pancreatic disease

Anna S Gerdtsson

LUND
UNIVERSITY

Faculty of Engineering
Department of Immunotechnology
ISBN 978-91-7473-748-6

Omslag Josefin Svenske

Printed in Sweden by Media-Tryck, Lund University
Lund 2013

Contents

My contribution to the papers

I I performed the experiments, analyzed the data and participated in writing the manuscript.

II I performed the experiments, analyzed the data and was main responsible for writing the manuscript.

III I performed the experiments, analyzed the data and was co-responsible for writing the manuscript.

IV I designed the experimental work, performed some of the experiments, supervised students in the practical work, and was the main writer of the manuscript.

V I supervised people in the experimental work. I analyzed the data and was co-responsible for writing the manuscript.

Original papers

This thesis is based upon the following papers, which are referred to in the text by their roman numerals (I-V).

I Wingren C., Sandström A., Segersvärd R., Carlsson A., Andersson R., Löhr M., Borrebaeck C.A.K. *Identification of serum biomarker signatures associated with pancreatic cancer.* Cancer Res. 2012, 72, 2481-90.

II Gerdtsson A.S., Säll A., Malats N., Real P., Persson H., Wingren C., Borrebaeck C.A.K. *Immunosignatures associated with pancreatic cancer identified in a large multicenter serum sample cohort using recombinant antibody microarrays.* Manuscipt.

III Sandström A., Andersson R., Segersvärd R., Löhr M., Borrebaeck C.A.K., Wingren C. *Serum protein profiling of pancreatitis using recombinant antibody microarrays reveals disease-associated biomarker signatures.* Proteomics Clin. Appl. 2012, 6, 486-96.

IV Gerdtsson A.S., Dexlin-Mellby L., Delfani P., Berglund E., Borrebaeck C.A.K., Wingren C. *Evaluation of solid supports for slide- and well-based recombinant antibody microarrays.* Manuscript.

V Gerdtsson A.S., Säll A., Åkesson A., Sjögren S.E.A., Borrebaeck C.A.K., Hansson S.R., Wingren C. *Protein, glycan and carbonyl profiling of crude proteomes using recombinant antibody microarrays.* Manuscript submitted for publication.

I have also contributed to the following scientific papers, not included in this thesis:

i Dexlin-Mellby L., Sandström A., Centlow M., Nygren S., Hansson S.R., Borrebaeck C.A.K., Wingren C. *Tissue proteome profiling of preeclamptic placenta using recombinant antibody microarrays.* Proteomics Clin. Appl. 2010, 4, 794-807.

ii Dexlin-Mellby L., Sandström A., Antberg L., Gunnarsson J., Hansson S.R., Borrebaeck C.A.K., Wingren C. *Design of recombinant antibody microarrays for membrane protein profiling of cell lysates and tissue extracts.* Proteomics 2011, 11, 1550-54.

Abbreviations

2-DE	Two-dimensional electrophoresis
Apo	Apolipoprotein
AUC	Area under curve
BRCA	Breast cancer susceptibility protein
C	Complement
CA	Carbohydrate antigen
CD	Cluster of differentiation
CDR	Complementary determining region
CEA	Carcinoembryonic antigen
CEACAM	Carcinoembryonic antigen-related cell adhesion molecule
CIMS	Context-independent motif-specific
CRP	C-reactive protein
CT	Computed tomography
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
ERCP	Endoscopic retrograde cholangiopancreatography
EUS	Endoscopic ultrasonography
Fab	Fragment antigen binding
FDA	Food and drug administration
FFPE	Formalin-fixed, paraffin-embedded
GAK	Cyclin G-associated kinase
GDF	Growth differentiation factor
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GPS	Global proteome survey
HADH	Hydroxyacyl-CoA dehydrogenase
HER	Human epidermal growth factor receptor
ICAM	Intercellular adhesion molecule
IGFBP	Insulin-like growth facor binding protein
IL	Interleukin
K-RAS	Kirsten rat sarcoma oncogene
LCN	Lipocalin

MALDI	Matrix assisted laser desorption/ionization
MAPK	Mitogen-activated protein kinase
MCP	Monocyte chemotactic protein
MRI	Magnetic resonance imaging
MRM	Multiple reaction monitoring
MS	Mass spectrometry
NIH	National institute of health
NHS	N-hydroxysuccinimide
OPG	Osteoprotegerin
PCA	Principal component analysis
PCR	Polymerase chain reaction
PSA	Prostate specific antigen
PTM	Post translational modification
Ras	Rat sarcoma
RCA	Rolling circle amplificaton
REG	Regenerating protein
RIA	Radioimmunoassay
ROC	Receiver operating characteristics
scFv	Single-chain fragment variable
SISCAPA	Stable isotope standard capture with antipeptide antibodies
SYCN	Syncollin
TGF	Transforming growth factor
TIMP	Tissue inhibitor of metalloproteinases
TNF	Tumor necrosis factor
TNFRSF	Tumor necrosis factor receptor superfamily
UPF	Up-frameshift suppressor
VEGF	Vascular endothelial growth factor

1. Introduction

Antibodies are naturally evolved in immune responses, where they act as specific binders for neutralizing invading pathogens. The human antibody repertoire is tremendously diverse as a result of somatic recombination and mutation events following infection, and high-affinity antibodies may be generated for essentially any given antigen. The intrinsic ability for combinatorial diversity and specificity for target antigens makes antibodies highly suitable for technical applications, where they can be used as binders for detecting molecules of interest in complex samples. *Antibody microarrays* are examples of one such application, and the work conducted in this thesis revolves around the antibody microarray technology, its design and its clinical applicability.

Antibody microarrays are highly multiplexed affinity proteomics assays, which can be employed for high-throughput protein expression profiling in the search for specific biomarker signatures (Borrebaeck & Wingren, 2007; Mustafa, Hoheisel, & Alhamdani, 2011; Sanchez-Carbayo, 2011). During the last decade, our research group have designed and optimized an antibody microarray platform based on human recombinant antibodies (Ingvarsson et al., 2007; Wingren, Ingvarsson, Dexlin, Szul, & Borrebaeck, 2007). The antibodies are selected from large scFv libraries (Soderlind et al., 2000) to target a range of different serum analytes, of which the vast majority are involved in immunoregulatory functions (Steinhauer, Wingren, Hager, & Borrebaeck, 2002).

A main focus of this thesis has been the application of antibody microarrays for identifying biomarker signatures for *pancreatic cancer*. Pancreatic cancer is the fourth most common cancer-related cause of death, with a 5-year survival rate of only 6% (Siegel, Naishadham, & Jemal, 2012). In fact, pancreatic cancer is one of the few cancers for which the survival has not been significantly improved over the last 40 years (American cancer society, Cancer Facts & Figures 2013). The high mortality of pancreatic cancer can be explained by late and vague presentation of symptoms and lack of sensitive methods for diagnosis, resulting in that most pancreatic tumors are detected at a late stage, where they have already metastasized and cannot be removed by surgery (Hidalgo, 2010). There is thus a great unmet clinical need for biomarkers for earlier diagnosis of this deadly disease. In **papers I and II**, we have presented candidate biomarker signatures for pancreatic cancer diagnosis, derived by comparing

protein profiles of pancreatic cancer patients to those of both normal healthy controls, as well as benign pancreatic conditions, such as *pancreatitis.*

Pancreatitis is an inflammatory state of the pancreas, which is symptomatically highly similar to, and difficult to separate from pancreatic cancer. Pancreatitis may appear as acute, chronic, or autoimmune manifestations, and as for pancreatic cancer, there is a lack of high-performing biomarkers for these conditions (Lippi, Valentino, & Cervellin, 2012; Lohr, 2007). In **paper III**, we have profiled the three different forms of pancreatitis, and identified biomarker signatures for discriminating pancreatitis from healthy controls, as well as stratifying acute, chronic and autoimmune pancreatitis.

Apart from applying the antibody microarray platform for biomarker discovery, I have also been involved in a number of technical evaluations of the platform, of which two studies are included in this thesis. In **paper IV**, different *solid supports* for antibody arrays have been assessed, and the first array-in-well set up has been presented, showing that plate-based antibody microarrays may be an option for future clinical implementations of the technology. In **paper V**, we have expanded the conventional antibody microarray set-up to target not only proteins, but also *glycan* and *carbonyl* groups. Post-translational modifications of proteins, like glycosylation and carbonylation (i.e. oxidation), are frequently altered in disease (Chandler & Goldman, 2013; Dalle-Donne, Giustarini, Colombo, Rossi, & Milzani, 2003), and serum profiling of such modifications may be an important complement to conventional protein biomarkers. The applicability of this novel set-up was demonstrated for preeclampsia, a common pregnancy disorder for which novel diagnostic and prognostic biomarkers would be of great clinical value (Forest et al., 2012).

In summary, this thesis deals with advances of the antibody microarray technology, recent efforts which have contributed both to expanding the utility of our in-house designed platform, as well as bringing the technology closer to clinical availability. In addition, the antibody microarrays have been applied for proteomic profiling of pancreatic cancer and pancreatitis, resulting in identification of candidate disease-specific biomarker signatures, which may in the long run lead to increased survival for many patients.

2. Cancer biomarkers

A *biomarker* is practically any characteristic that can be used as an indicator of disease, however in most cases it refers to biomolecules, such as *genes* or, as in our case, *proteins*, that are mutated or have an altered expression pattern in disease compared to controls. Biomarkers can be used not only for disease *diagnosis*, but also for e.g. *prognosis* of severity, *prediction* of therapeutic efficacy, or *stratification* of disease. All these types of biomarkers enable the use of *personalized medicine*, the molecular characterization of patients in order to identify the most beneficial treatment for each individual (Jorgensen, 2009; Langreth & Waldholz, 1999).

2.1 Clinically established cancer markers

Genomic technologies, including sequencing, gene microarrays and PCR, are well established methods, facilitating both the discovery and measurement of gene biomarkers. Examples of routinely measured gene markers are *BRCA1* and *BRCA2*, which if mutated indicate a significantly increased risk for breast and ovarian cancer in women (Miki et al., 1994; Wooster et al., 1995), and *EGFR*, which if mutated in lung cancer is a biomarker for beneficial treatment with the EGFR inhibitor gefitinib (Pao et al., 2004; Wakeling et al., 2002).

Gene marker analysis often provide simple, binary read-outs, such as mutated versus not mutated, or gene product being present versus absent. In contrast, protein biomarkers are generally more difficult to assess, as they rather have an altered level in disease, such that a cut-off level needs to be pinpointed. Protein markers are however beneficial considering that proteins are the actual functional molecules, and thus may provide a more accurate measure of the disease state. Moreover, it has been shown that the *translational* (protein) and *transcriptional* (mRNA) levels frequently are uncorrelated (Gygi, Rochon, Franza, & Aebersold, 1999). Consequently, both protein and gene markers are of value and may harbor complementary information.

To date, there are around twenty single protein cancer biomarkers that have been approved by the US Food and Drug Administration (FDA) (Anderson, 2010; Ludwig & Weinstein, 2005). The most recognized protein markers are perhaps

HER2, a breast cancer prognostic marker that indicate tumor aggressiveness and potential benefit from trastuzumab therapy (Schechter et al., 1984), and PSA, which is routinely used for prostate cancer screening (Carter et al., 1992; Thompson et al., 2004). For *pancreatic cancer*, which will be described in further detail in Chapter 4, the carbohydrate antigen CA19-9, has so far been the most prominent marker (Pleskow et al., 1989). However, not unlike many other single biomarkers, CA19-9 suffers from poor predictive value, and is today recommended solely for monitoring recurrence of CA19-9 positive tumors (Locker et al., 2006).

Biomarkers are often assessed by their sensitivity and specificity, where *sensitivity* is the ability of the marker to identify sick individuals in a population (true positives), and *specificity* the ability to identify non-sick individuals in a population (true negatives) (Altman & Bland, 1994). A good biomarker needs to have both high sensitivity and specificity, however there is always a trade-off between the two factors, and a certain cut-off level needs to be selected for which the sensitivity and specificity are reported. Alternatively, the performance of a biomarker may be illustrated by a *ROC-curve*, an approach which have been exploited by us in all biomarker studies included in this thesis. ROC-curves, generated by plotting sensitivity versus 1-specificity, demonstrate the biomarker accuracy over all sensitivity and specificity thresholds (Zweig & Campbell, 1993), and the area under the curve, or *AUC*, can be used as a measure of marker performance. An AUC value of 0.5 represent poor biomarker accuracy, "as good as guessing", and in the other end, an AUC of 1.0 represent perfect discrimination of cases and controls, corresponding to 100% sensitivity and specificity.

2.2 Multiplexed cancer biomarkers

Despite a tremendous number of discovery studies in which potential biomarkers have been identified, only a handful of these have managed to reach clinical implementation (Anderson, 2010; Issaq, Waybright, & Veenstra, 2011). These discouraging results have in recent years opted the field to move from single to *multiplexed markers*, for increased sensitivity and specificity (Brody, Gold, Lawn, Walker, & Zichi, 2010; Rifai, Gillette, & Carr, 2006). Considering that the same cellular mechanisms are involved in many different diseases, it is unlikely that a change in concentration of one single protein can be used as a high-performing marker for one specific type of cancer. Instead, multiplexed biomarker *panels* or *signatures* may provide synergistic effects in that they enable relatively unspecific (i.e.

deregulated in different types of cancers or related conditions) single-handed markers to be combined for a more sensitive and specific molecular fingerprint of the disease.

Microarray technologies are highly suitable for both the discovery and implementation of multiplexed biomarker panels. The first diagnostic microarray test to be FDA approved in 2007 was the MammaPrint, a signature of 70 genes that predicts the risk of recurrence in lymph node negative breast cancer (van 't Veer et al., 2002). Numerous multiplexed protein signatures have also been reported for various cancers (examples of pancreatic cancer protein signatures will be covered in Chapter 4). To the best of my knowledge, only one of these has so far been approved by the FDA, a 5-plex panel of ApoA1, β2-microglobulin, CA125-II, prealbumin, and transferrin for evaluating an ovarian mass for cancer prior to a planned surgery (Fung, 2010).

2.3 Biomarker discovery in serum

Apart from sensitive and specific, biomarker assays should preferably be *non-invasive.* Although biomarkers have been sought after and identified in different non-invasive sample specimens, such as urine, saliva, and sputum (Good et al., 2007; Hassanein et al., 2012; Sanchez-Carbayo, 2006; Truong, Yang, & Jarrard, 2013), *serum* and *plasma* are by far the most common sample formats for non-invasive biomarker analysis. Apart from being readily available, blood-derived samples are considered to be the most complete, containing both the entire blood proteome as well as secretion and leakage proteins from tissues, and (if present) tumors (Anderson & Anderson, 2002).

Serum and plasma differ in that serum is depleted of both red blood cells and coagulation factors, while plasma still contains coagulation factors but is prevented from clotting by adding an anti-coagulant (e.g. EDTA). Which of the two formats that is to be preferred has been much debated, and seems to depend on the technique at hand (Haab et al., 2005). There are, however, significant differences between the two (Liu et al., 2010; Schwenk et al., 2010), and thus biomarkers that have been discovered in e.g. serum, cannot not necessarily be validated in plasma. For the analysis of pancreatic disease (papers I-III) we have exclusively used serum, and for the case of simplicity I will refer to *serum proteins* for all blood-derived proteins.

Albeit beneficial in its non-invasiveness, availability and high protein content, the use of serum is also associated with several challenges. The serum proteome is highly complex, due to high concentration and heterogeneity. Serum proteins are generally present in many different forms due to e.g. precursors, degradation

products, splice variants, and different patterns of glycosylation and other post-translational modifications (Anderson & Anderson, 2002). A majority of the proteins are also assumed to be complex-bound to other proteins. In addition, the quantitative *dynamic range* of serum proteins spans over ten orders of magnitude (Fig. 1), with concentrations ranging from 35-50 mg/mL for albumin down to a few pg/mL for some cytokines (Anderson & Anderson, 2002; Hanash, Pitteri, & Faca, 2008). In fact, 99% of the protein mass in serum is made up of approximately 20 different proteins (Tirumalai et al., 2003), which consequently will mask the more low-abundant analytes, such as tissue leakage proteins and cytokines. Considering that potential biomarkers are most likely found among these more low-concentrated proteins (Haab et al., 2005; Surinova et al., 2011), highly sensitive technologies are required for their detection.

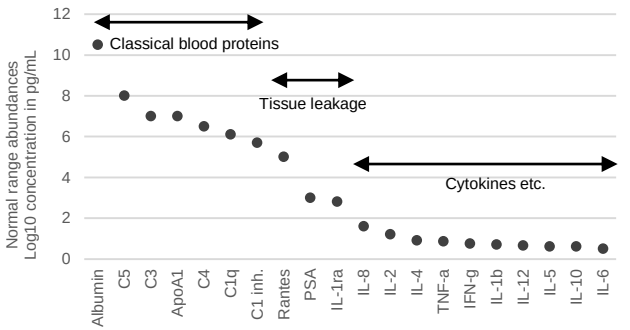


Figure 1. Dynamic range of blood proteins. Abundance is shown for selected proteins measured on our antibody microarrays. Albumin is plotted as a reference. *Values are obtained from the plasma proteome institute, and the plot is adopted from Anderson & Anderson (2002).*

2.4 Proteomics technologies for biomarker discovery

The large scale study of proteins, *proteomics*, began with the advent of *two-dimensional electrophoresis* (2-DE) (O'Farrell, 1975), which rapidly increased the protein coverage from single protein analysis. The 2-DE technology became the cornerstone for biomarker discovery, until advancements of *mass spectrometry* (MS), such as the introduction of *MALDI* (matrix-assisted laser desorption/ionization) (Karas & Hillenkamp, 1988), made MS the method of choice for large scale proteomic analysis. Despite tremendous additional development over the last decades, including the arrival of the *orbitrap* technology (Makarov, 2000), MS still suffers from relatively poor sensitivity in complex samples such as serum. Although the sensitivity may be increased by targeted approaches, such as *multiple reaction monitoring* (MRM) (Keshishian, Addona, Burgess, Kuhn, & Carr, 2007), highly sensitive MS-based methods are still not adapted for large-scale high-throughput biomarker discovery.

The sensitivity needed for identification of the most low-abundant serum proteins can so far only be achieved by the use of *affinity reagents*, such as antibodies. *Immunoassays* have been extensively used since the introduction of the radioimmunoassay (RIA), based on radioisotope labeling (Yalow & Berson, 1959), and later the enzyme-linked immunosorbent assay (ELISA) (Engvall & Perlmann, 1971), which still today is one of the most common diagnostic assays. *Monoclonal antibodies* were introduced a couple of years later (Kohler & Milstein, 1975), enabling the isolation of analytes from more complex mixtures, such as serum.

Successful combinations of MS and immunoassays have also been demonstrated. For example, *SISCAPA* (stable isotope standard capture with antipeptide antibodies), which couples antibody peptide enrichment for increased sensitivity to an MS-based read-out (Anderson et al., 2004). Another example is the elegant attempt to approach global proteome coverage presented by Olsson et al (Olsson, James, Borrebaeck, & Wingren, 2012; Olsson et al., 2011). In their *GPS* (global proteome survey) platform, antibodies directed against short (4-6 amino acids) N-terminal sequence motifs, where each motif may be shared by up to hundreds of proteins, are used for peptide capture and combined with MS-based detection and quantification. A similar approach, using 3-4 amino acid sequence motifs, has also been demonstrated (Poetz, Hoeppe, Templin, Stoll, & Joos, 2009).

The first attempts to *multiplex* immunoassays began in the mid 1980's with the *microspot* assay developed by Ekins and Chu (Ekins, Chu, & Biggart, 1990; Ekins & Chu, 1991). Since then, the field of *affinity proteomics* has rapidly advanced with numerous highly multiplexed microarray platforms. Apart from antibody

microarrays, antigen arrays and reverse-phase protein arrays have also been extensively utilized. *Antigen arrays* (Bussow et al., 1998; MacBeath & Schreiber, 2000) are generated by printing purified proteins or peptides onto solid supports, and are commonly deployed for profiling the autoantibody response in autoimmune diseases or cancer (Hudson, Pozdnyakova, Haines, Mor, & Snyder, 2007; Orenes-Pinero et al., 2010; Zhang, 2004). Commercial antigen arrays have also frequently been exploited, for example the ProtoArray (Schweitzer, Meng, Mattoon, & Rai, 2010), which in its current version constains over 9000 unique human proteins (www.lifetechnologies.com).

Instead of pure antigens, *reverse-phase protein arrays* (Paweletz et al., 2001) are composed of arrayed patient samples, which are probed with antibodies targeting proteins or proteins modifications such as phosphorylation, enabling the analysis of signaling networks in cancer (Silvestri et al., 2010; Wulfkuhle et al., 2003). Serum is however rarely used for reverse-phase arrays, as the technology does not support detection of low-abundant proteins in such complex sample format. Instead, tissue or cell lysates are the most frequently used sample specimens. While antigen and reverse-phase arrays are suitable for autoantibody profiling and functional studies using a relatively small amount of samples, antibody microarrays, which will be covered in chapter 3, have frequently been the affinity proteomics method of choice for large-scale protein biomarker discovery.

Like any technology, however, affinity proteomics also suffers from drawbacks. The main limitation compared to MS based proteomics has been the availability of antibodies, as commercial, well characterized, high-affinity binders may be hard to acquire. To this end, the use of recombinant antibody libraries, from which binders can be selected by display technologies, is an efficient way of generating antibodies against selected targets, and will be further described in the next chapter. Recently, there have been several initiatives for generating large sets of high quality affinity reagents at high throughput and low cost, including the *Affinomics*, *ProtomeBinders*, and *AffinityProteome* projects in the EU, as well as the *NIH* (National Institute of Health) *Common Fund's Protein Capture Reagents Program* and the *Human Antibody Initiative* in the US (Stoevesandt & Taussig, 2012). These projects, together with technical advances, such as the global proteome survey mentioned above, and efforts within the nanoarray technology which will be briefly described in Chapter 5, have the potential for increasing the protein coverage within affinity proteomics tremendously.

2.5 From discovery to clinical implementation

Although an overwhelming number of proteomics studies are being published each year, relatively few protein biomarkers are used in clinical practice, and the rate of clinical implementation of novel markers is actually decreasing (Rifai et al., 2006). This disappointing outcome can partly be explained by the long and costly process of biomarker assay development.

The protein biomarker pipeline has been described to consist of six separate phases (Rifai et al., 2006). First, proteins of interest are identified in a limited set of well-defined samples (*discovery*). Second, the identity and presence of candidate markers are confirmed by the use of alternative technologies (*qualification*), followed by validation of the markers in a larger number (hundreds) of samples with a broader range of controls (*verification*). Next, the assay characteristics, such as the linear range and reference interval for each analyte, is determined and made sure to meet the regulatory requirements (*assay optimization*), before the assay can be tested in thousands of samples that fully reflect the intended target population (*validation*), and finally refined to meet the regulatory standards for clinical tests (*commercialization*).

The majority of clinically approved protein biomarkers are measured using immunoassays, and for biomarkers discovered using MS, the technical transfer to an affinity reagent platform somewhere along this line, is also likely required (Hanash et al., 2008). In addition, MS discovery studies are frequently based on either pre-fractionated serum or plasma, or less complex sample formats, such as proximal fluids (e.g. pancreatic juice) or cell line models (Hanash et al., 2008). In contrast, affinity reagents can target low-abundant analytes in crude serum without the need for pre-fractionation, thus enabling discovery studies to be conducted using both the sample specimen *and* the technology intended for the final assay, potentially saving time and money spent on candidate markers, that otherwise may not withstand the platform transfer likely needed at a later stage of the biomarker development.

3. Antibody microarrays

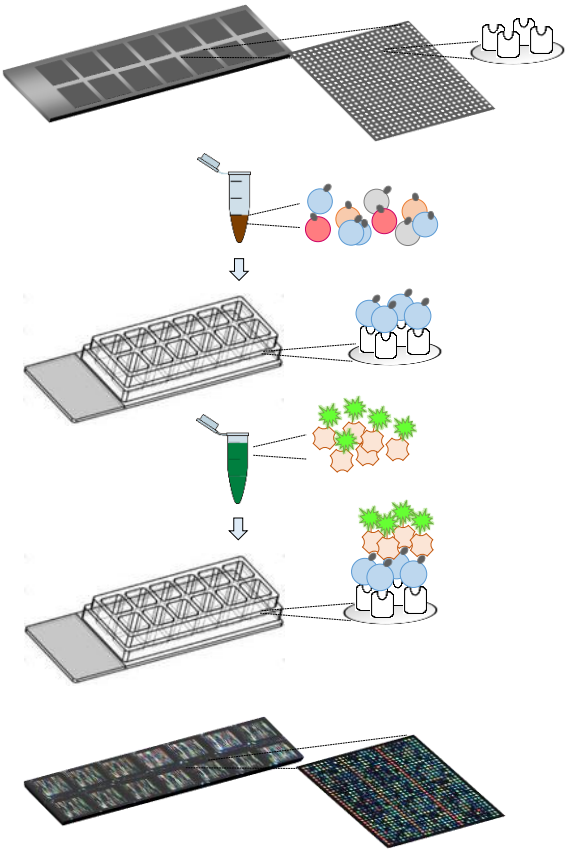
Antibody microarrays have over the last couple of decades evolved into indispensable tools for affinity proteomics. Owing to their ability for highly sensitive, multiplexed protein analysis in complex samples, as well as the capacity for high sample throughput, they have been widely applied for large-scale protein expression profiling in the search for disease-specific biomarkers.

3.1 The antibody microarray platform

The work in this thesis has been conducted using an in-house developed recombinant antibody microarray platform, designed and optimized for profiling of complex proteomes. (Ingvarsson et al., 2007; Steinhauer et al., 2002; Wingren et al., 2007). The platform set-up is illustrated in Fig. 2. In brief, the microarrays are produced by non-contact printing, using a robotic piezoelectric spotter equipped with glass capillaries dispensing approximately 300 pL droplets of purified, recombinant antibody solution onto a slide surface. The antibody droplets dry out immediately, forming spots with a diameter of ~120 µm on the solid support. Currently, the discovery arrays are composed of close to 300 antibodies, targeting around 100 different serum analytes.

Typically, the arrays are printed on day one, and used for sample analysis the following day. In the current set-up, up to ten slides are printed per day, with up to 14 subarrays per slide, thus enabling over a hundred samples to be analyzed per day and workstation. The slides are mounted in incubation gaskets with silicon superstructures that create individual subarray compartments. After the arrays have been blocked and washed, clinical samples (most often serum), which have been pre-labeled with biotin, are added and incubated to allow for protein binding to the arrayed antibodies. When the samples have been washed off, fluorophore-coupled streptavidin is applied, which binds to the biotin on the antibody-captured proteins. The arrays are then washed and dried, and slides are scanned in a confocal laser microarray scanner, creating high resolution images of the subarrays. The fluorescence intensity of each antibody spot is then quantified and used as a measure of the relative

level of the bound protein. This way, a protein binding profile is generated for each clinical sample, and by comparing sample groups of cases and controls, disease specific protein signatures may be identified.



1. Antibodies (scFvs) are printed onto a solid support in array patterns. Each array spot contains one unique antibody clone.

2. Serum samples that have been pre-labled with biotin are thawed on ice. Microarray slides are mounted in hybridization gaskets and blocked.

3. Blocking buffer is washed off and one serum sample is added to each array. The samples are incubated to allow the antibodies to bind their cognate proteins.

4. Unbound proteins are washed off and streptavidin-coupled fluorophore is added and incubated to allow the streptavidin to bind to the biotin on the captured proteins.

5. Unbound streptavidin is washed off. The slides are dried and scanned in a confocal laser microarray scanner. The fluorescence signal from each antibody spot is quantified and used as a measure of the relative abundance of protein in the sample.

Figure 2. Overview of the recombinant antibody microarray platform.

3.2 Immunosignaturing

Different strategies can be applied when deciding what antigens to target by an antibody microarray. Compared to MS-based discovery studies, which are largely *untargeted*, microarray analysis is by definition a *targeted* approach, in that the affinity reagents, and hence the proteins to be measured, have been selected in advance. Indeed, many microarrays consist of probes targeting disease associated antigens, such as tumor leakage proteins, known oncogene protein products, and previously reported cancer biomarkers (Sanchez-Carbayo, Socci, Lozano, Haab, & Cordon-Cardo, 2006; Schroder et al., 2010).

In this context, our approach has been to select antibodies that target proteins involved mainly in *immunoregulatory* functions, such as cytokines, chemokines, growth factors, adhesion molecules, and complement components (although more recent array expansions (paper II) also include e.g. enzymes and signaling molecules). The association between tumors and the immune system has long been recognized. Inflammation and immune response have been established as hallmarks of cancer (Hanahan & Weinberg, 2011), and it has been shown that inflammatory mediators may both promote and be promoted by tumors (Coussens & Werb, 2002; Lippitz, 2013).

Considering that the immune system is highly affected in any condition, the proteins targeted on our arrays are however not expected to show single-handed specificity towards a certain disease. Instead, it was postulated that *immunosignatures* could be identified and used as sensitive and specific fingerprints for individual diseases. Since then, such diagnostic and/or prognostic signatures have been delineated for a range of indications, including *H. pylori*-induced gastric cancer (Ellmark et al., 2006), pancreatic cancer (Ingvarsson et al., 2008), metastatic breast cancer (Carlsson et al., 2008; Carlsson, Wingren, et al., 2011), glioblastoma (Carlsson et al., 2010), as well as autoimmune diseases (Carlsson, Wuttge, et al., 2011). In paper I-III, this strategy has been further exploited for pancreatic cancer and pancreatitis.

3.3 Platform design – key parameters

Apart from the choice of what antigens to target, several technical key features need to be assessed and optimized in the design of a microarray platform (Borrebaeck & Wingren, 2009). As illustrated in Fig. 3, these include the choice of *affinity probes* and

solid support, both of which will influence the choice of *probe immobilization* technique. In parallel, the *sample labeling* approach should be selected and optimized. With these parameters in place, the assay protocol can be developed, including *array processing* (i.e. blocking, washing and sample buffers, incubation time, temperature, etc.), choice of *detection system*, and, finally, methods for *data preprocessing* (e.g. quality control and normalization), and *data analysis*. Some of these parameters will be described in further detail in this chapter.

Figure 3. Key parameters in the design of an affinity proteomics microarray platform.

3.3.1 Affinity probes

Although different types of affinity reagents, including *aptamers* (Cho, Collett, Szafranska, & Ellington, 2006; Gold et al., 2010) and *affibodies* (Renberg, Shiroyama, Engfeldt, Nygren, & Karlstrom, 2005), may been used for antigen capture, *antibodies* are undoubtedly the most frequent choice of affinity probe. Both *polyclonal* and *monoclonal* antibodies have been employed in array applications (Borrebaeck & Wingren, 2011; Stoevesandt & Taussig, 2007). Polyclonal antibodies, produced by immunization of animals, can be generated in high amounts at a relatively low cost. Nevertheless, monoclonal antibodies are generally preferred in array applications, as they are both renewable and specific for unique epitopes. Monoclonal antibodies have traditionally been produced using hybridoma technology, however advances in display methods have promoted the generation of *recombinant* antibodies.

One of the most frequently used system for generating and screening recombinant antibodies is *phage display* (McCafferty, Griffiths, Winter, & Chiswell, 1990), which also has been our method of choice. In phage display, fragments of full length antibodies, *Fab* (fragment antigen binding) or (as in our case) *scFv* (single-chain fragment variable), are displayed on phages in large antibody libraries. The library that has been the predominant source for our antibodies contains over 10^{10} unique antibody clones, which are all based on the same polypeptide framework, and differ only in the complementary determining region (CDR) loops that form the antigen binding site (Soderlind et al., 2000).

Using recombinant antibodies based on a fixed antibody structure is likely to contribute to a more homogenous on-chip behavior, compared to the use of protein probes of different sizes and charges. In addition, frameworks that demonstrate high stability in array applications may be selected (Dubel, Stoevesandt, Taussig, & Hust, 2010). The affinities of recombinant antibodies are typically in the nanomolar range and thus comparable to natural antibodies, but may be further increased by affinity maturation (Dubel et al., 2010). Recombinant antibodies can also be readily mutated for desired on-chip characteristics, such as increased stability (Nordström et al, manuscript in preparation), or by introducing functional groups for e.g. surface coupling (Steinhauer et al., 2006).

Following selection by phage display, the antibody genes are cloned into *E. coli* and can be produced and purified from either the periplasm or the cell supernatant. While antibody production used to be a tedious and time-consuming step, recent advances have dramatically increased the throughput, enabling the generation of hundreds of antibodies in less than a week by the use of robotics (Hust et al., 2009).

3.3.2 Solid support and antibody immobilization

Antibody arrays are of two main types, *planar arrays*, in which antibodies are printed on slides (Belov, de la Vega, dos Remedios, Mulligan, & Christopherson, 2001; Knezevic et al., 2001; Miller et al., 2003; Sanchez-Carbayo et al., 2006; Schroder et al., 2010), and *suspension arrays*, in which antibodies are immobilized on color-coded beads (Fulton, McDade, Smith, Kienker, & Kettman, 1997; Morgan et al., 2004; Schwenk et al., 2010). Bead-based arrays allow for highly automated analysis, using liquid handling systems, however, the level of multiplexing is limited to the number of colors that can be resolved in the detection system, which usually is based on flow cytometry. Moreover, it has been shown that bead based assays are more prone to cross reactivity, and have a narrower dynamic range than planar arrays (Ellington, Kullo, Bailey, & Klee, 2010).

The surface properties of the solid support will directly impact the assay performance, and the antibody-surface interplay has also been a focus of this thesis (paper IV). Solid supports should have a high binding capacity, preserve antibody activity and resist non-specific binding. Antibodies may be immobilized onto the solid support by *adsorption*, *affinity-coupling*, or by *covalent binding* to functional groups on the surfaces, in either a random or directed manner (Seurynck-Servoss et al., 2008; Steinhauer et al., 2005). In our set-up, a black polymer slide to which the antibodies are immobilized by adsorption, has been the surface of choice (Wingren et

al., 2007). Different solid supports for planar arrays, and their impact on the assay functionality will be discussed in further detail in Chapter 5.

Planar antibody arrays may be produced by either *contact* or *non-contact* printing. When antibodies are printed onto a surface, high reproducibility is a prerequisite, as well as high precision, considering that the spot-to-spot distance is usually only around 200 µm or less. Contact printers, equipped with solid metal pins, are often considered to be more rapid and easily cleaned than non-contact printers (Austin & Holway, 2011). Non-contact piezoelectric printing, which has been used for the production of our arrays, is however more reproducible, more easily controlled, allows for antibody recovery, and is often stated to be more biocompatible, i.e. not causing denaturation of the printed proteins (Delehanty, 2004).

Alternatives to antibody printing are demonstrated by *DNA-directed immobilization*, in which DNA is printed, allowing for antibodies tagged with complementary oligonucleotides to be directed to their corresponding array position (Wacker, Schroder, & Niemeyer, 2004), and *self-assembly arrays*, in which the antibodies are produced on-chip by the use of printed protein-encoding DNA and a cell-free expression system (Stoevesandt et al., 2011). These are appealing approaches in which the antibody purification step could potentially be circumvented, and work along this line is on-going also in our group (Wingren et al, unpublished observations).

3.3.3 Sample labeling and detection

Although the work in this thesis has been conducted using solely blood derived samples, including both serum (paper I-IV) and plasma (paper V), a range of different sample formats have been analyzed using antibody microarrays. For example, assay protocols for targeting urine (Kristensson et al., 2012), intact cells (Dexlin, Ingvarsson, Frendeus, Borrebaeck, & Wingren, 2008), tissue extracts and cell lysates (Dexlin-Mellby et al., 2011), and formalin-fixed paraffin-embedded (FFPE) tissue (Pauly et al., 2013), have been developed for our platform. For each type of sample, protocols regarding e.g. sample dilution, sample buffer, washing buffer, and labeling have been optimized.

Historically, different types of labeling reagents have been used for immunoassays, including radioisotypes, enzymes, and chemiluminescence. Today, most detection systems are however based on *fluorescence*, which is sensitive, safe, easy to use and offers stable signaling (Schaferling & Nagl, 2006). *Dual-color* approaches, in which two samples are labeled with two different fluorophores, e.g. Cy3 and Cy5,

and analyzed on the same microarray, was frequently applied in the early years of antibody microarray development, and are still not uncommon (Haab, Dunham, & Brown, 2001; Schroder et al., 2010). Dual-color labeling has also been assessed by our group, but was abandoned due to significant differences in labeling efficiencies of the different dyes tested (Wingren & Borrebaeck, 2008; Wingren et al., 2007).

The *biotin* and *streptavidin* reagents used by us are widely exploited in many different applications, owing to the high affinity binding of biotin to streptavidin (or avidin), and the minimal sample impact caused by the small biotin molecule. Sample proteins are tagged with biotin through *N*-hydroxysuccinimide (NHS)-coupling to primary amines, i.e. lysins and arginines. Each biotin molecule then captures one of the four binding sites on a streptavidin molecule, which in turn has been pre-coupled to a fluorophore. In paper V, we extended the use of our microarray platform by introducing labeling of glycans and carbonyls, which will be covered in Chapter 5.

Alternatively to direct labeling using fluorophore or biotin, a second, labeled antibody may be used for detection. Such *sandwich* antibody microarrays are often stated to be more specific than their single-capture counterparts, considering that the capture of two independent antibodies are required for signaling (Nielsen & Geierstanger, 2004). Moreover, the sample labeling step, which always involves a certain risk of affecting the sample, for instance by epitope masking, is circumvented in the sandwich approach. On the other hand, acquiring two antibodies targeting different epitopes of the same protein may be problematic, and in addition, sandwich arrays are limited to 30-50 different targets, as increased multiplexing have shown to introduce cross-reactivity of detection antibodies (Haab, 2005).

Sandwich microarrays have also been generated using *lectins*, carbohydrate binding proteins, as either capture or detection probes (or both), often in combination with protein specific antibodies (Haab & Yue, 2011; Heimburg-Molinaro, Song, Smith, & Cummings, 2011). Such *lectin-antibody arrays* may be applied for differential profiling of glycosylation in disease, as has been exemplified for pancreatic cancer (Yue et al., 2009; Yue et al., 2011). Highly informative, disease associated changes of both protein abundance and glycan alterations can be derived from these arrays. The drawbacks are the need for masking carbohydrates on the arrayed antibodies, the large amount of sample needed (one array per lectin is required), as well as the low affinity and cross reactivity frequently displayed by lectins. Moreover, the analysis and interpretation of the abundance of data generated from such set-ups, may be highly complex.

3.3.4 Data preprocessing

Unlike gene array analysis, there are yet no standardized way in how protein microarray data should be preprocessed. Instead, each research group has developed their own approach to data handling, based on experience and the nature of the dataset, i.e. the number of analytes and samples, and the scientific question at hand. Data preprocessing should however include a data distribution analysis, which will determine what statistical tools that can be applied in the downstream data analysis, as well as the exclusion of potential outliers. In addition, some sort of *normalization* of the data is more or less always required in order to reduce or eliminate non-biological sample variations. Commonly used normalization strategies include the use of spike-in antigens, reference methods (e.g. ELISA), and global normalization (Hamelinck et al., 2005; Park et al., 2003; Quackenbush, 2001). After thorough revision of several of these strategies, we have chosen a *semi-global* normalization approach, in which a subset of the analytes (15-20%) that show the smallest overall variation across all samples is used to calculate a scaling factor for each array (Ingvarsson et al., 2008). This method has been used in all studies in this thesis. The study in paper II was, however, conducted after an extensive platform update, which among other things, involved major changes in the array design. This resulted in that over 100 samples could be analyzed per day, compared to 10 samples in the previous set-up, and led to the observation of day-to-day variations, which could be eliminated by using a frequently applied *subtract group mean* strategy, in which the average signal of samples analyzed on the same day was calculated for each antibody and subtracted from the individual values. This was followed by an array-to-array normalization, using the semi-global approach described above.

3.4 Microarray data analysis

Traditionally, proteomics data analysis has been conducted in a *univariate* mode, i.e. evaluating one protein at a time. The differences in protein levels between cases and controls have been assessed using t-tests, and analytes with p-values below a certain threshold have been reported as significantly differentially expressed, and as having a potential discriminative power as a biomarker. Recently, it has however been increasingly recognized that more sophisticated statistical tools, able to interpret *multivariate* data such as that generated from antibody microarrays, may be a missing link for increasing the rate of biomarker panels being approved for clinical use

(Anderson, 2010). Considering that a certain level of correlation of disease-related protein expression is likely to be present, particularly when targeting highly interconnected immunoregulatory proteins, univariate analysis of variation may indeed not be the most appropriate option for identifying the optimal combination of analytes to include in a disease-specific immunosignature (Quackenbush, 2001). Instead, different multivariate models, which may be broadly divided into two categories based on unsupervised and supervised classification (Fig. 4), can be applied.

3.4.1 Approaches for sample classification

Unsupervised classification methods identify underlying patterns in the data and use these to cluster samples and/or variables without any prior knowledge of the true sample annotation (or antibody specificities). For example, *hierarchical clustering* creates branched trees of samples and/or antibodies based on the unbiased similarities in the array data (Eisen, Spellman, Brown, & Botstein, 1998). One should be aware, however, that branches will be generated even in completely random data; hence clusters do not necessarily reflect true biological patterns, and need to be carefully validated. Another unsupervised method, applied in papers II, III and V, is *principal component analysis* (PCA), which transforms multidimensional data into a limited set of orthogonal (uncorrelated) variables called components, where each component contains maximum data variability.

Unsupervised methods are suited for initial analysis and overview of datasets, and, as demonstrated in paper V for preeclampsia, for stratifying samples and identifying novel subgroups. They may also be deployed for quality control of array data. If unsupervised classification separates clinical samples according to their initial diagnosis, it implies that the array analysis was able to extract relevant information from the samples (provided that no technical bias has been introduced in the analysis).

In contrast, *supervised classification* is based on a priori information on sample annotation, and is applied to generate prediction models for distinguishing cases and controls. Supervised classification methods include regression analyses, random forests, and artificial neural networks, to mention a few (Cammann, Jung, Meyer, & Stephan, 2011; Kim et al., 2009). In papers I, II, III and V, we have applied a supervised learning method called *support vector machine* (SVM).

3.4.2 *Support vector machine analysis*

A support vector machine creates a *hyperplane* to separate two sample groups in a multidimensional space made up of the variable data, which in our case is composed of the antibody signals. New samples are predicted to belong to either cases or controls depending on what side of the hyperplane they are positioned in the space. In other words, the model is constructed by training on a representative dataset, before its predictive power can be evaluated in novel samples. The microarray data thus needs to be divided into *training set* and *test set*, and for accurate analysis it is imperative that the samples used to validate the model (test set) has not been involved in its construction (training set).

The separation of samples into training sets and test sets can be done in a number of ways. In paper I, we used two different approaches. First, we compared pancreatic cancer and healthy controls. Considering that the number of samples were relatively small (n=64), we reasoned that there was not enough data to perform adequate training and validation in two different datasets. Instead, we applied a *leave-one-out cross validation,* excluding one sample at a time from the dataset, and using the remaining samples to train the data. When the excluded sample was put back into the space, it was given a *decision value*, which corresponds to its distance to the hyperplane and was either negative or positive depending on what side of the plane it is positioned. This process was repeated until all samples have been given a decision value, after which the model performance could be evaluated. Second, we compared pancreatic cancer to the combined group of both healthy controls and pancreatitis. Here, we estimated that the number of samples (n=103) was sufficient to randomly subdivide the data into one training set and one test set. This approach is considered to be more stringent, since training is done only once, in comparison to the leave-one-out approach, in which a new hyperplane is created each time a sample is excluded, which may increase the risk for overtraining or *overfitting* the model.

3.4.3 *Identification of biomarker signatures*

Support vector machine models can be constructed using the entire dataset, i.e. information from all antibodies included on the array. Using hundreds of antibodies for disease classification will however not be optimal in a clinical assay setting, which is why condensed (≤30-plex), candidate biomarker signatures have been derived from the array data. In order to identify the combination of proteins that has the highest

classification power, we have taken on a *backward elimination* approach (Carlsson, Wingren, et al., 2011), and applied this in papers I and II.

The backward elimination algorithm excludes one antibody at the time from the data set, and performs leave-one-out cross validation of the samples using the remaining antibodies, with the subsequent generation of ROC-curves. The process is continued until all antibodies have been excluded once. The antibody that was excluded when the classification was performed with the highest accuracy (or rather the smallest error), is considered to be the least important feature, and is removed from the dataset. This procedure is then repeated until only a single antibody remains.

The order of eliminated antibodies can then be used to identify signatures of an appropriate number of proteins. The predictive value for each signature "length", as recorded from the elimination process, function as an estimate of the optimal number of analytes needed for high-performing classification. For example, the backward elimination performed in paper II showed that pancreatic cancer was readily discriminated from normal controls using only 4-10 antibodies, while much larger signatures, based on in average 67 antibodies, were needed for optimal classification of pancreatic cancer and pancreatitis, as these sample groups were much harder to separate.

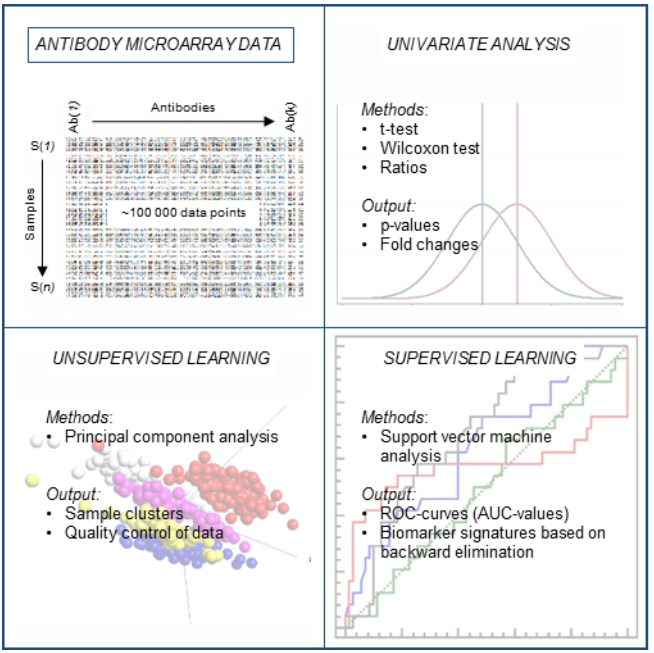


Figure 4. Overview of data analysis methods applied in this thesis.

4. Applications on pancreatic disease

The main purpose of this thesis has been applications of the antibody microarray technology for protein expression profiling of pancreatic disease. In paper I and II, we have applied antibody microarrays to analyze serum samples from patients with pancreatic cancer, pancreatitis, and healthy controls. By comparing protein patterns in these sample groups, biomarker signatures for classification of pancreatic cancer have been identified. In paper III, a similar approach was undertaken, but with a focus only on pancreatitis. Here, the array data was used to stratify chronic, acute and autoimmune pancreatitis, and to pinpoint potential biomarkers and biomarker signatures for each of these pancreatitis subtypes.

4.1 Pancreatic cancer

Pancreatic cancer affects approximately 1 in 10 000 in the US, making it the 12th most common type of cancer (Shaib, Davila, & El-Serag, 2006; Siegel et al., 2012). In Sweden, around 900 individuals are diagnosed with pancreatic cancer every year (www.cancerfonden.se). Even though it is not among the most frequent types of cancer, the median survival is only 6 months, and the 5-year survival rate is 6% (Michaud, 2004). Consequently, a close to equal number of deaths as new cases are reported each year, and pancreatic cancer is currently the 4th most common cancer related cause of death, surpassed only by lung, breast, and colon cancer (Siegel et al., 2012).

Surgery remains the only curative option for pancreatic cancer, although less than 15% of patients have surgically resectable disease at the time of diagnosis. The tumors are potentially resectable only in stage I and II pancreatic cancer (localized tumors), however when detected, the cancer has in most cases reached stage III or IV, with a tumor that has grown inoperable, and in about 50% of cases has formed distant metastases (stage IV) (Pannala, Basu, Petersen, & Chari, 2009).

The late detection is partly due to vague and late presented symptoms, including abdominal pain, jaundice, weight loss, and type II diabetes, symptoms of which many may be caused by other disorders than pancreatic cancer, including

pancreatitis. In addition, there is a lack of sensitive diagnostic methods and high-performing biomarkers that can detect pancreatic cancer at an early stage, and discriminate it from benign pancreatic conditions. Today, pancreatic cancer is diagnosed using mainly computed tomography (CT), endoscopic ultrasonography (EUS), endoscopic retrograde cholangiopancreatography (ERCP), and magnetic resonance imaging (MRI) (Cote, Smith, Sherman, & Kelly, 2013). These imaging methods are however both costly and rarely sensitive enough to detect premalignant lesions, e.g. pancreatic intraepithelial neoplasias, or early stage tumors that are still small and can be removed by surgery (Chu, Kohlmann, & Adler, 2010).

The pancreas is an abdominal gland with both *endocrine* function in the production of insulin and other hormones, and *exocrine* function in the secretion of digestive enzymes into the pancreatic juice for transportation to the small intestine. The pancreas consists of *head, body* and *tail*, with the head to the left, located in cavity of the duodenum, the body behind the stomach, and the tail protruding to the right towards the spleen (Fig. 5). Cancer may arise in all parts of the pancreas, including the endocrine compartment of the pancreatic islet cells. These neuroendocrine tumors are however rare and clinically distinct from other pancreatic cancers and has not been covered in this thesis.

The vast majority (95%) of pancreatic cancers are exocrine adenocarcinomas that arise in the pancreatic duct, a branched vessel that runs through the pancreas, and that carries the pancreatic juice to the common bile duct (Fig. 5). Most of the ductal cancers (80%) are located in the head of pancreas, and can sometimes be resected by the so called Whipple operation, unless the mesenteric blood vessels surrounding the pancreatic head also have been invaded. Tumors located in the *body* or *tail* of the pancreas may be removed by a procedure called distal pancreatectomy.

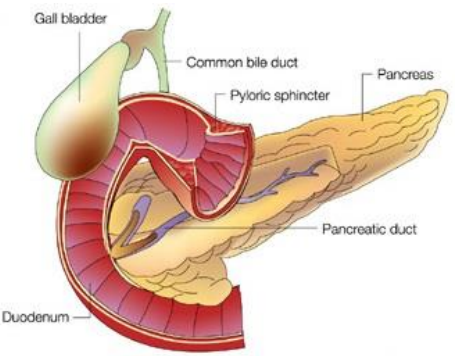


Figure 5. Anatomy of the pancreas.
Reproduced with permission from Nature Publishing Group

The body and tail tumors are however often more aggressive than those of the head, and have frequently metastasized at the time of diagnosis. Only cases with no evidence of metastasis or cancer invasion of the celiac artery, are considered for surgery (Hidalgo, 2010).

The causes of pancreatic cancer are largely unknown. The only factor that has been statistically established as causative is smoking, which increases the risk for pancreatic cancer 2-3 times (Hassan et al., 2007). Also, about 5-10% of pancreatic cancer patients have a family history of the disease, although the genetic predisposition behind familial pancreatic cancer has not fully been characterized. In general, pancreatic cancers are highly heterogeneous, and aberrantly mutated, with an average of over 60 genetic abnormalities per tumor (Hidalgo, 2010). Up to 90% of cases have mutations in *K-RAS*, an oncogene which if mutated produce persistently active Ras protein. Despite the high frequency, *K-RAS* is of limited use as a biomarker for pancreatic cancer, as its mutated form also is found in many other cancers including colorectal and lung cancer, as well as in pancreatitis and in up to 30% of other non-cancer controls (Parsons & Meng, 2009).

In fact, the only biomarker for pancreatic cancer that so far has been FDA approved for clinical use is *carbohydrate antigen 19-9* (CA19-9). CA19-9, or *sialyl Lewis a*, was first discovered over 30 years ago, as an elevated serum marker in colon and pancreatic cancer patients (Koprowski, Herlyn, Steplewski, & Sears, 1981). However, CA19-9 suffers from dismal predictive power and specificity for pancreatic cancer (Ballehaninna & Chamberlain, 2012), and is moreover completely absent in about 10% of the population (Lewis a and b negative individuals). Today, the use of CA19-9 is recommended solely for monitoring recurrence of CA19-9 positive tumors (Locker et al., 2006).

Many other markers have been suggested for pancreatic cancer, including CA242, CA125, CEA, CEACAM-1, CRP, DU-PAN2, GDF-15, haptoglobin, IGFBP-1, M2 pyruvate kinase, platelet factor 4, and serum amyloid A, to mention a few (Ballehaninna & Chamberlain, 2013; Duffy et al., 2010; Fry, Monkemuller, & Malfertheiner, 2008; Koopmann et al., 2006). Despite the low accuracy of CA19-9, it has not yet been outperformed by other markers in validated studies. There is thus a significant unmet clinical need for novel biomarkers for pancreatic cancer, preferably for early detection.

4.2 Biomarker panels for pancreatic cancer

Numerous multiplexed biomarker panels have also been identified for pancreatic cancer, including signatures based on transcripts, micro-RNAs, metabolites and autoantibodies (Ballehaninna & Chamberlain, 2013; Winter, Yeo, & Brody, 2013). The studies briefly reviewed here are restricted to those based on serum antigen panels, identified by the use of credible study design and multivariate analysis, with pre-validation in independent sample cohorts, or by cross-validation procedures.

Using focused, commercial bead-based arrays, Brand et al. recently identified a panel of CA19-9, ICAM-1 and OPG for discrimination of pancreatic cancer and healthy controls in large serum sample cohorts, but were unable to find a panel that improved the sensitivity of CA19-9 for discriminating pancreatic cancer from benign pancreatic conditions (Brand et al., 2011). Faca et al. used an appealing MS-based approach of identifying proteins of interest in a mouse model, with the subsequent validation of selected panels in human serum (Faca et al., 2008). A panel of LCN2, TIMP1, REG1A, REG3 and IGFBP4 was combined with CA19-9 for discrimination of early stage pancreatic cancer and controls (AUC 0.91) in a small sample set. In another MS-based study, Xue et al. found that the combination of CA19-9, ApoC1 and ApoA2 improved the diagnostic power over CA19-9 alone in pancreatic cancer when compared to both healthy (AUC 0.96) and disease controls (AUC 0.90) in small validation sets (Xue et al., 2010). Furthermore, Makawita et al. recently confirmed previously identified candidate biomarkers in two large serum sample cohorts using ELISA (Makawita et al., 2013). They found that a combination of CA19-9, SYCN and REG1B could distinguish early stage pancreatic cancer from disease free controls with an AUC of 0.87 and 0.90 in the two sample sets.

In an effort of profiling cytokine patterns of pancreatic cancer, Zeh et al. used a bead-based immunoassay and a cross-validation approach to identify multivariate cytokine panels for discriminating pancreatic cancer and healthy controls, as well as chronic pancreatitis with sensitivities and specificities in the range of 86-96% (Zeh et al., 2005). In another early effort of multivariate analysis, the group of Haab applied antibody microarrays and different classification models, with varying results (Orchekowski et al., 2005). This group has later changed focus from core protein biomarkers to the analysis of glycan epitope alterations in pancreatic cancer, using antibody-lectin sandwich microarrays, as was described in Chapter 3 (Haab et al., 2010; Yue et al., 2009).

The majority of reported multiplexed biomarker panels thus consist of 2-5 proteins, frequently medium to high abundant analytes, and often in combination

with CA19-9. In contrast, the candidate biomarker signatures identified by us in papers I and II have included up to 25 different proteins, with emphasis on low- to medium-abundant immunoregulatory serum proteins, thus representing key differences from most other multiplexed biomarker efforts. In paper I, we first identified an 18-plex protein signature for predicting pancreatic cancer samples from healthy controls. For pre-validation, we used a preexisting dataset from antibody microarray analysis of an independent serum sample cohort set (Ingvarsson et al., 2008), in which our signature demonstrated an AUC of 0.95. Second, we used a training set to identify a 25-plex signature for differentiating pancreatic cancer from the combined group of healthy and pancreatitis controls. This signature was pre-validated in a separate test set, in which it generated an AUC of 0.88.

In paper II, we undertook a similar approach, but in a significantly larger serum sample cohort, and using microarrays based on more than twice as many antibodies as in paper I. Here, ten different, randomly generated pairs of training and test sets were applied, and in each training set the most predictive 25-plex signature was identified, and evaluated in the corresponding test set. The ROC curves generated had an average AUC of 0.98 for the pancreatic cancer versus healthy controls signatures, and 0.67 for the pancreatic cancer versus benign controls (mainly chronic pancreatitis) signatures. The backward elimination analysis showed that 4-10 antibodies would be sufficient for close to perfect discrimination of cancer and healthy controls. However, further analysis showed that a much larger number of antibodies (average 67) was required for maximum separation of cancer and benign controls, indicating that small immuonosignatures (<10-plex) would likely not be highly specific for pancreatic cancer. Moreover, principal component analysis demonstrated that the serum samples in part could be separated based on the tumor location in the pancreas (head versus body/tail), which to the best of my knowledge has not previously been shown with proteomics. In summary, the signatures for predicting pancreatic cancer from normal controls that were derived by us (AUC 0.95-0.98) are among the best-performing pre-validated candidate biomarker panels that have been presented to date.

4.3 Study design

Sample cohorts used for biomarker discovery should ideally contain age- and gender-matched cases and controls, preferably collected in comparable numbers. The samples should be identically handled prior to analysis, be decoded in a way so that sample class is blind to the operator, and be applied on arrays in a random order, all to avoid introducing any technical biases (Ostroff et al., 2010; Ransohoff, 2005). As already has been emphasized, biomarker accuracy (e.g. sensitivity, specificity, AUC) should be evaluated in independent samples, by applying separate training and test sets.

In several biomarker studies, including those presented in this thesis, sensitivities and specificities (or AUC values) well above those achieved for CA19-9, have been reported for pancreatic cancer versus healthy controls. However, prior to conducting such biomarker discovery study, the purpose of any putative biomarkers should be clarified, and control groups should be applied accordingly. In line with this, we have used pancreatitis as one of the control groups for pancreatic cancer. Considering that pancreatic cancer and pancreatitis present with close to identical symptoms, diagnostic biomarkers that can separate these two conditions would indeed be of high clinical value.

If the biomarkers should be used for pancreatic cancer screening in the general population, healthy individuals would be the appropriate control group, and our results indicate that highly sensitive biomarker signatures for identifying pancreatic cancer from healthy controls can be derived from antibody microarray data. Because of the low prevalence of pancreatic cancer, a general screening test is however not likely to be cost effective (Larghi, Verna, Lecca, & Costamagna, 2009; Pannala et al., 2009). On the other hand, a recent study has demonstrated the cost-effectiveness of screening *high-risk* groups for pancreatic cancer, based on the sensitivity and specificity reported by us in paper I (Ghatnekar et al., 2013). Risk factors for pancreatic cancer are both *hereditary*, including a family history of disease, genetic predisposition (e.g. BRCA1 and BRCA2 mutations), Peutz-Jeghers syndrome, and hereditary pancreatitis, as well as *non-hereditary*, such as chronic pancreatitis, newly onset diabetes, *Helicobacter pylori* infection, smoking, obesity, and high age (Konstantinou, Syrigos, & Saif, 2013; Larghi et al., 2009). The prevalence of pancreatic cancer among smokers or elderly is still too low to justify screening in these population, however individuals with one or more of the other mentioned risk factors may be eligible for screening. The next step will thus be to explore the ability of our immunosignatures for predicting pancreatic cancer in high-risk groups. In particular individuals with a demonstrated hereditary predisposition, accounting for more than

10% of pancreatic cancer patients (Tersmette et al., 2001), and patients with new-onset diabetes, as recent data have shown that up to 80% of pancreatic cancer patients are either hyperglycaemic or diabetic (Pannala et al., 2009), may be considered as appropriate controls.

Moreover, pancreatic tumors should preferably be detected at an *early* and still resectable stage, which poses a major challenge on biomarker discovery studies, as early pancreatic cancer (stage I and II) clinical samples for obvious reasons rarely are available in large numbers. Instead, tentative biomarkers are generally reported in stage III and IV pancreatic cancer, with the supposition that many signature proteins would be deregulated already at an earlier cancer stage. In a current effort of identifying truly early stage markers, a set of serum samples collected one month up to five years *prior* to diagnosis of pancreatic cancer have been drawn from a prospective, population based sample cohort (Manjer et al., 2001). These samples have recently been analyzed on our antibody microarrays, together with healthy and chronic pancreatitis controls from the same cohort (Dexlin-Mellby and Wingren, ongoing studies), and we are currently awaiting the results from this study.

Another attractive option would be to analyze multiple samples from the same patient. Several studies have demonstrated that there are large biological variations in serum protein abundances between different individuals (Cava, Gonzalez, Pascual, Navajo, & Gonzalez-Buitrago, 2000; Tuxen, Soletormos, Petersen, Schioler, & Dombernowsky, 1999), which suggests that each patient is his or her own best control. By comparing samples drawn at different time points, base-line levels could thus be pinpointed for each patient, enabling disease-associated changes in protein abundance to be monitored over time in a truly personalized manner. This approach has yet not been pursued by us for pancreatic cancer, but hopefully such samples will be acquired for future studies.

4.4 Validation of biomarker signatures

An intelligent study design with proper pre-validation is a prerequisite for further verification studies of tentative biomarkers. As described in Chapter 2, the path from discovery study to clinical implementation of biomarkers is however long and tedious, involving thorough validation of both the technology and the proposed biomarkers in large independent sample cohorts. This process is even more complex when the technology is still evolving in parallel to that biomarker discovery studies are being performed, which often is the case. For instance, our platform has been applied in

numerous clinical studies, in parallel to which technical efforts aiming for further optimizing, improving, or expanding the microarray technology have been conducted (Borrebaeck & Wingren, 2009; Wingren & Borrebaeck, 2006).

The studies on pancreatic cancer profiling (papers I and II), were performed with a four year interval, over which the microarray platform went through a number of critical changes. For example, paper I was based on analysis using 121 antibodies, while in paper II, several new binders had been added to the repertoire, resulting in a total of 293 different antibodies. The arrays were also produced using different printers, and processed using highly different protocols. While the arrays used in paper I were printed one per slide, and processed in a semi-automated manner using a liquid handling system, a so called *protein array workstation* (PAW) which enabled 12 arrays to be processed each day, paper II was performed following several technical advances, including more rapid array printing, which contributed to a significantly increased throughput. Here, 13 arrays were printed on each slide and over hundred samples could be processed per workstation and day. One of the very first studies conducted on our antibody microarrays was also based on pancreatic cancer profiling (Ingvarsson et al., 2008), using yet another protocol differing from those applied in papers I and II. In addition, the data analysis strategies have been upgraded from simple differential expression analysis, to more advanced methods, such as the supervised classification approaches described in chapter 3.

However, in order to validate a biomarker signature in a novel sample cohort, the antibodies used and the protocol employed needs to be close to identical in the validation study as in the discovery study, or else the classification model will not be readily applicable in the new sample cohort. Moreover, thousands of samples may be required for true validation of clinical utility (Rifai et al., 2006). Particularly for multiplex biomarker panels, for which the disease associated changes in abundance of individual proteins may be relatively small, large sample cohorts are needed to achieve sufficient statistical power (Alonzo, Pepe, & Moskowitz, 2002). Hence, during the time needed to collect a sufficiently large, prospective serum cohort, particularly for early stage pancreatic cancer, the technology is likely to have changed, and the "validation" study may no longer be strictly comparable to the discovery study. Thus instead of taking on a conventional path of discovery followed by validation studies, we have conducted several consecutive studies on pancreatic cancer, each time in an improved technical setting, and with a larger, or more relevant sample cohort. This iterative discovery/validation process so far involves three analyzed (Ingvarsson et al 2008, paper I and II), one currently evaluated, and at least two more sample cohorts that are waiting in the pipeline, each adding a piece of the puzzle to the proteomic fingerprint of this complex disease.

Despite the technical and biological (sample based) differences, several of the same candidate biomarkers were identified both in paper I and II. For example, C1 inhibitor, C5, GM-CSF, IL-3, IL-4, MCP-1, and TGF-β1 appeared in the backward elimination signatures derived for pancreatic cancer versus pancreatitis in both these studies. In addition, the signatures in paper I and II also significantly overlapped with the profile presented by Ingvarsson et al. (Ingvarsson et al., 2008), with C1 inhibitor, C3, C5, CD40, Eotaxin, IL-4, MCP-1, TNF-β, and VEGF included in protein signatures from all three sample cohorts. However, several antibodies that were novel in paper II, including those targeting ApoA1, ApoA4, GAK, HADH2, MAPK1, TNFRSF3, and UPF3B, also showed potential for pancreatic cancer classification, replacing features of the signatures identified in the previous studies, and with a need for verification in an additional sample cohort. In the future, these immunosignatures will be further refined and validated, tested for presence in early stage pancreatic cancer, and evaluated for diagnostic potential in larger sample cohorts including the appropriate risk groups.

4.5 Proteomic profiling of pancreatitis

Pancreatitis is inflammation of the pancreas, believed to be caused by premature activation of pancreatic enzymes, which results in pancreatic cellular injury and inflammation (Vonlaufen, Wilson, & Apte, 2008). Pancreatitis may occur as acute, chronic, or autoimmune disease. *Acute pancreatitis*, which initially cause reversible pancreatic damage, may be triggered by obstruction of the pancreatic duct, often by gallstones or, in rare cases, by tumors. Approximately 10% of acute pancreatitis cases are classified as severe, with an intense inflammatory response that may even be lethal (Vonlaufen et al., 2008).

Chronic pancreatitis, in which the inflammatory state has become irreversible, may arise from recurrent episodes of acute attacks (Braganza, Lee, McCloy, & McMahon, 2011). Chronic pancreatitis has also often been associated with excessive alcohol use, although a significant part of the chronic pancreatitis cases appear for no apparent reason (idiopathic disease). In fact, the pancreatitis incidence even among heavy drinkers is low (<3%), and the alcohol intake need to be substantial (>5 drinks per day) to be associated with an increased risk of pancreatitis (Yadav et al., 2009). Consequently, it has been suggested that alcohol (often in combination with smoking) is merely a strong modifier of disease, and not the cause of susceptibility.

Moreover, the genetic predisposition for pancreatitis is becoming increasingly unraveled (Whitcomb, 2013).

In addition to the characterized genetic and environmental causes, pancreatitis may also have an autoimmune etiology. *Autoimmune pancreatitis* is a chronic disease that presents with the same symptoms as other forms of pancreatitis or pancreatic cancer, but that may be distinguished by increased levels of immunoglobulins and specific autoantibodies, as well as a response to steroid treatment. It has become increasingly clear that autoimmune pancreatitis may be subdivided into two types. Type 1 is more common worldwide, and appears to be the only subtype affecting the Asian population. It is characterized as a multi-organ disease with high IgG4 levels, while type 2 is restricted to the pancreas and diagnosed by a histological pattern of granulocyte epithelial lesions (Sah & Chari, 2012).

In general, there are several unmet clinical needs within the management of pancreatitis, including increased molecular understanding of disease, means of early diagnosis, as well as markers for predicting e.g. disease severity and the onset of chronic disease from recurrent acute pancreatitis (Whitcomb, 2013). Today, pancreatitis is diagnosed by the same imaging techniques (e.g. CT and MRI) used for pancreatic cancer, which suffer from insensitivity and high cost, as already been described. As for pancreatic cancer, there are currently no high-performing biomarkers in routine clinical use. Although the presence of elastase 1, pancreatic lipase, and amylase may indicate pancreatitis, these pancreatic enzymes are rarely detected at an early stage disease (Lohr, 2007). Serum IgG4 and autoantibodies are commonly measured for the diagnosis of autoimmune pancreatitis, however type 2 autoimmune pancreatitis can currently only be confirmed by histology. Pancreatitis biomarkers would thus be of great clinical value, both for diagnosis, and for stratifying the individual manifestations.

In paper III, we set out to meet these clinical needs by proteomic profiling of all three forms of pancreatitis and healthy controls, using our antibody microarray platform. First, it was demonstrated that pancreatitis could be readily discriminated from healthy controls by applying support vector machine classification in a leave-one-out cross validation manner (as described in Chapter 3). For each type of pancreatitis, signatures of significantly differentially expressed immunoregulatory analytes could also be derived. Next, the approach of support vector machine prediction, followed by protein signature identification was repeated, first for each subgroup versus all other samples (i.e. both other forms of pancreatitis and healthy controls), and second for comparing the individual subgroups to each other.

This extensive analysis generated a comprehensive depiction of the similarities and discrepancies of the three different forms of pancreatitis, based on the

immunoregulatory proteins measured on the antibody microarrays. It was shown that acute and autoimmune pancreatitis could be readily discriminated, while, as could be expected, chronic disease shared many features of both of these forms. Multiple tentative biomarkers could be suggested for each of the different subtypes. For example, the signatures for acute pancreatitis involved several complement proteins, while the signatures specific for chronic pancreatitis mostly contained cytokines. The analysis of autoimmune pancreatitis confirmed the notion of a highly heterogeneous disease. With a few exceptions, a significant part of the proteins showed decreased serum levels in the autoimmune condition compared to healthy controls or other pancreatitis. Unfortunately, no clinical data on type 1 or 2 autoimmunity was available, and thus these subtypes could not be stratified.

We concluded that profiling of the immunoregulatory proteome by the use of antibody microarrays showed high potential both for biomarker discovery and for stratification of pancreatitis. Although further validation in independent samples is needed, the protein signatures presented in paper III could prove useful as markers for pancreatitis, and may add to an improved understanding of the underlying molecular patterns of the disease(s), which potentially could provide a basis for personalized management of pancreatitis.

5. Technical advances of the antibody microarray assay

Apart from the clinical applications on pancreatic disease, the work in this thesis has also involved technical developments of the antibody microarray platform. In paper IV, different solid supports for antibody microarrays have been evaluated, including both source plates, planar slides and 96-well plates. In paper V, we have introduced an additional biotin reagent for detection of glycan and carbonyl groups in clinical samples, and demonstrated the applicability of this approach in a plasma sample cohort with preeclamptic patients and healthy pregnant controls. In addition to these efforts, several other on-going projects, of which a few will be mentioned in this chapter, show great potential for advancing the antibody microarray technology into a leading tool for biomarker discovery, and for bringing proteomics into the clinics.

5.1 Evaluation of solid supports

As briefly mentioned in Chapter 3, the *solid support* will have a tremendous impact on the antibody microarray assay (Kusnezow & Hoheisel, 2003). An extensive evaluation of the compatibility of different types of slides with our scFv antibodies has previously been performed by our group (Wingren et al., 2007). However, several years have elapsed since that study and similar studies by others (Angenendt, Glokler, Murphy, Lehrach, & Cahill, 2002; Angenendt, Glokler, Sobek, Lehrach, & Cahill, 2003; Kusnezow & Hoheisel, 2003; Seurynck-Servoss, White, Baird, Rodland, & Zangar, 2007), over which the availability of different surfaces and, in our case, both the printing procedure and the array design have changed. In addition, we recently observed high and inconsistent, non-specific protein binding to different *antibody source plates*, the 384-well plates in which the antibodies are loaded prior to printing, which in turn affected the printing performances. These factors, as well as an emerging need for developing a user-friendly immunoassay for clinical implementation of multiplexed biomarker signatures, motivated us to perform an updated evaluation of solid supports for recombinant antibody microarrays.

The use of recombinant antibodies offers several options for antibody immobilization. For example, in early versions of our microarray platform the antibodies were *affinity coupled* via histidine-tags to Ni^{2+}-nitriloacetic acid surfaces (Steinhauer et al., 2006). Later work however demonstrated that superior assay functionality (e.g. lower background and higher array density) was achieved when a black polymer slide, to which the antibodies are immobilized by *adsorption*, was used as solid support (Wingren et al., 2007).

In paper IV, we evaluated slides to which antibodies were immobilized either *covalently* to NHS- and epoxy coated surfaces, or by adsorption to different supports. Despite that the epoxy surfaces showed potential in means of high signals and excellent spot morphology, none of the slides tested outperformed the currently used black polymer surface, on which the assay demonstrated high reproducibility, low limit-of-detection, and high dynamic range, compared to the other surfaces that were assessed.

While the microarray solid support should have a high binding capacity, the antibody source plates should ideally be completely non-binding to avoid that precious antibody sample is adsorbed to the surface of the wells. The evaluation of different 384-well source plates showed large variations in antibody binding. A black polypropylene plate demonstrated high surface homogeneity and low protein binding compared to the other plates tested, and has since then been used as antibody source plate with good results.

Finally, we evaluated different flat-bottom 96-well plates for use as antibody microarray support. Most biomarkers in clinical use today are measured with ELISA assays, a standard technology that can be run in basically any clinical laboratory. Producing condensed microarrays in the bottom of 96-well plates thus seemed like a good option for developing a user-friendly, multiplexed biomarker assay. For this purpose, our first *array-in-well* design was presented in paper IV. The best performing set-up was based on clear polymer 96-well plates scanned in a confocal laser plate scanner. Although the limit of detection does not yet match the slide-based assay, in part due to the lack of sensitive, high-resolution plate scanners, as well as biocompatible plate based solid supports, the array-in-well assay is an important step towards the development of a biomarker signature test for clinical use.

5.2 Protein, glycan and carbonyl profiling

To date, the biomarker signatures identified by us have been based on differences in protein abundance between cases and controls. However, it has become increasingly recognized that also the *post-translational modifications*, which are natural and essential cellular processes for providing proteins with their final structure or functional ability, are frequently altered in disease compared to healthy. Profiling changes of the level and nature of post-translational modifications in disease thus may enable the discovery of novel biomarkers, which can complement traditional protein biomarkers. In paper V, we set out to target two of these modifications, namely glycosylation and carbonylation, using antibody microarray analysis.

Glycosylation is the addition of carbohydrate groups to specific amino acids of the protein polypeptide chain. The majority of human proteins are believed to be glycosylated in some way, making glycosylation the most common form of post-translational modification (Apweiler, Hermjakob, & Sharon, 1999). There are numerous examples of altered glycosylation patterns being used as indicators of disease (Chandler & Goldman, 2013), including increased levels of the CA19-9 (sialyl Lewis a) structure in pancreatic cancer, as described in previous chapters. *Carbonylation* is a less prevalent form of post-translational modification, in which the protein side chains are oxidized. Protein carbonylation may occur as a natural cellular process, or as a consequence of cellular stress caused by disease (Dalle-Donne et al., 2003), and has been associated with a range of conditions, predominantly Alzheimer's disease (Sultana & Butterfield, 2013), but also preeclampsia (Forest et al., 2012; Zusterzeel, Rutten, Roelofs, Peters, & Steegers, 2001).

Protein glycan and carbonyl groups are often targeted using *hydrazide* reagents. In the conventional microarray set-up, we label proteins with NHS-biotin targeting primary amines, i.e. lysins and arginines, as was described in Chapter 3. In paper V, a biotin-hydrazide reagent, which reacts with aldehydes and ketones (carbonyls), was added to the labeling procedure. The carbonyls may be preexisting as a result of disease or other oxidative effects on the clinical samples, or may be created by gentle oxidation of glycans. We labeled each clinical sample in three parallel batches; one using NHS-biotin (core protein biotinylation), and two using biotin-hydrazide, both with and without pre-oxidation of proteins (targeting glycan and carbonyl groups, respectively) (Fig. 6). Following sample labeling, the three preparations were applied on identical antibody microarrays, and the protein, glycan, and carbonyl signaling was assessed for each patient sample. First, we performed an evaluation and

optimization of the biotin-hydrazide labeling procedure, and second, proof-of-concept was demonstrated for plasma profiling of preeclampsia.

Preeclampsia is a pregnancy disorder that affects 3-8% of pregnancies worldwide, and is the most common cause of fetal and maternal pregnancy related death (Anderson, Olsson, Kristensen, Akerstrom, & Hansson, 2012). Preeclampsia is characterized by hypertension and proteinuria, but biomarkers for early detection and prediction are lacking. By deploying the novel labeling procedure, we analyzed protein, glycan and carbonyl profiles of plasma samples from preeclamptic patients, including late onset, early onset, and early onset with intra-uterine growth restriction preeclampsia, as well as normotensive pregnant women.

The study showed that three very distinct protein profiles were generated from the different labeling strategies. We used the glycan-to-protein and carbonyl-to-protein ratios to stratify the samples, and could conclude that tentative markers based on both glycan and carbonyl levels could be delineated for preeclampsia and its subtypes. The results implied that particularly the carbonyl level is altered in preeclampsia, and the carbonyl-to-protein ratios may potentially also be used to identify novel subtypes of disease.

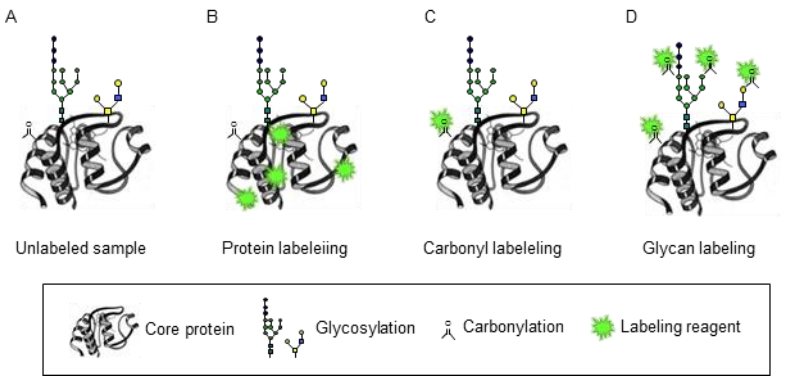


Figure 6. Schematic overview of the sample labeling approaches applied in paper V. A) Unlabeled glycosylated and carbonylated protein. B) Protein labeled with NHS-biotin, targeting primary amines in the core protein amino acid structure. C) Protein labeled with biotin-hydrazide, targeting carbonyl groups. D) Pre-oxidized protein labeled with biotin-hydrazide, targeting oxidized glycan groups (and preexisting carbonyl groups)

5.3 Current developments

Besides the technical advances that I have been involved in, including design of the array-in-well set-up (paper IV), and glycan and carbonyl array profiling (paper V), there are several more on-going projects for improving and expanding the antibody microarray technology, of which only a couple will be mentioned below.

Since the start of the antibody microarray era, there has been a constant aim for increasing the protein coverage. The level of multiplexing has gone from single-digit to the current set-ups of hundreds of antibodies, and is now moving forwards towards the measurement of thousands of proteins, thus slowly approaching a complete proteome coverage. In our group, this work has been performed on multiple levels. For example, increasing array density by going from micro to nano scale have been demonstrated by the first generation of *nanoarrays* (Wingren & Borrebaeck, 2007), which have been produced using dip-pen nanolithography technology (Petersson et al, submitted manuscript). In addition, the antibody repertoire is continually being expanded, for example by generation of the *context-independent motif-specific* (CIMS) antibodies used within the global proteome survey (GPS) briefly mentioned in Chapter 2, but also by the recent design and development of novel recombinant antibody libraries (Säll and Persson, unpublished data), which already have contributed to a significant increase of antibodies used in array applications. State-of-the-art robotics and liquid handling systems have also greatly increased the rate of selection and evaluation of novel binders from phage display libraries, as well as the subsequent production and purification of antibodies. Recent work has also involved the introduction of a non-natural amino acid to the antibodies for improved coupling design (Petersson et al, manuscript in preparation). These antibodies are currently used in a promising start-up project of analyzing arrays in solution coupled to a true quantitative read-out method, which might enable an even higher sample throughput combined with a highly robust detection system (Wingren et al, on-going studies).

In all, the above mentioned efforts for increased throughput and protein coverage have the potential for making antibody microarrays the obvious method of choice for proteomics discovery studies. In addition, the parallel clinical studies for refining and validating already identified protein signatures, together with the efforts of developing a highly user-friendly assay (e.g. the array-in-well set-up), show great potential for bringing the antibody microarray technology to clinical utility.

6. Concluding remarks

The field of proteomics has undergone tremendous technical advances during the last decade, resulting in improved protein detection in complex proteomes such as serum. Despite these technical efforts, few biomarkers are brought to the clinic, and most tentative protein markers never make it beyond the discovery phase. The reasons for this disappointing outcome are manifold. Proteomics technologies have suffered from poor reproducibility and sensitivity, causing difficulties in both the identification and verification of candidate markers. Moreover, the regulatory requirements on novel biomarkers have sharpened, resulting in lengthy and expensive processes of validating biomarkers for clinical utility.

In this aspect, recombinant antibody microarrays hold great potential for breaking new grounds within biomarker discovery. It has long been recognized that the use of affinity reagents offer superior sensitivity for capturing low-abundant serum proteins. The high sensitivity and reproducibility of immunoassays, combined with the increasing rate and decreasing cost with which new binders are generated, will significantly expand the protein coverage. The work in this thesis, represented by five original papers, revolves around this recombinant antibody microarray technology, its technical advances and clinical applications within pancreatic disease.

The urgent need for actionable results within pancreatic research has recently been acknowledged, with the Recalcitrant Cancer Research Act, (formerly known as the Pancreatic Cancer Research & Education Act) being signed into law in the US in January 2013, calling on the American National Cancer Institute (NCI) to develop frameworks for improving the survival of this deadly disease. To this end, earlier detection will be essential, and there is thus an immense clinical need for biomarkers for pancreatic cancer diagnosis. In paper I and II, we identified serum protein signatures for classification of pancreatic cancer among both healthy controls and pancreatitis (pancreatic inflammation). The pre-validation of the signatures in separate sample test sets showed that pancreatic cancer could be readily distinguished from healthy controls, with AUC values in the 0.95 to 0.98 range. The prediction of pancreatic cancer versus benign controls (mainly pancreatitis) was more challenging, resulting in AUC values up to 0.88. Even though the second study (paper II) was based on a significantly larger dataset, including both more samples and antibodies, and performed after an extensive technological assay update, the signatures identified

largely overlapped in the two studies. Next, these candidate biomarker signatures will be further refined and validated for early detection of pancreatic cancer in high-risk groups. In paper III, we performed proteomic profiling of the different forms of pancreatitis, to the best of my knowledge the first proteomic study of its kind. Here, extensive analysis for comparison of chronic, acute and autoimmune pancreatitis generated distinct protein signatures for the three individual manifestations. Although further validation is needed, these profiles may provide a deeper insight in the proteomic patterns behind these conditions, and may form the basis of biomarker signatures for the stratification and prediction of pancreatitis.

The two last studies of this thesis were of a more technical nature, aiming for developing and expanding the utility of the recombinant antibody microarray platform. In paper IV, we evaluated different surfaces for antibody microarray solid support, and also assessed the impact of the antibody source plates on the microarray production. We concluded that the hydrophilic black polymer surface that was already in use still provided superior assay functionality compared to novel slides that were assessed. Moreover, the first recombinant array-in-well design was presented, by which we have taken the first steps towards developing a more user-friendly, microarray-based biomarker immunoassay for clinical use. In paper V, we introduced a novel labeling procedure targeting core proteins, glycan and carbonyl groups. By profiling differential levels of both proteins and the post-translational modifications of glycosylation and carbonylation in disease, we hypothesized that novel and complementary biomarkers could be identified. The study included both the design and optimization of the assay protocol, and a first pilot application on preeclampsia, a common pregnancy disorder. The results demonstrated that different forms of preeclampsia e.g. early and late onset disease, could be distinguished, and indicated that particularly the level of oxidation (carbonylation) was altered in this disease.

In summary, this thesis demonstrates some of the advantages of using recombinant antibody microarrays for biomarker discovery, and their potential for clinical implementation of multiplexed biomarker assays. The technology has advanced immensely during the last decade, and also over the time in which the work in this thesis was conducted. My contribution to this progress has involved both optimizations and expansion of the microarray platform, and applications of the technology for biomarker discovery in pancreatic cancer and pancreatitis. Although further validation is needed, this thesis has added key pieces to the puzzle of proteomics in pancreatic diseases, and paved the way for future clinical utility of biomarker signatures, that in the end may provide benefit for many thousands of patients.

Populärvetenskaplig sammanfattning

Antikroppar är proteiner som ingår i kroppens immunförsvar, där de fungerar som bindare för att neutralisera patogena mikroorganismer eller ämnen som infekterat kroppen. Vid en infektion börjar immunceller producera antikroppar genom att på ett smart sätt sätta ihop olika gener. Varje enskild immuncell använder en viss kombination av gener och producerar därmed antikroppar som har en särskild *specificitet*. Antikroppsgenerna kan kombineras på ett nästan oändligt antal olika sätt, vilket innebär att kroppens immunceller kan generera bindare mot i princip vilken molekyl som helst. Antikropparnas förmåga att binda starkt och specifikt till olika molekyler är unik, och kan med fördel utnyttjas i olika tekniska applikationer. Ett exempel på en sådan teknik är antikroppschip, *antibody microarrays*, och allt arbete i den här avhandlingen kretsar kring den tekniken. Min forskning har syftat både till att använda antikroppschip inom sjukdomsdiagnostik, med fokus på sjukdomar i bukspottskörteln (artikel I-III), men även att på olika sätt vidareutveckla tekniken (artikel IV och V).

De antikroppschip som under det senaste decenniet har utvecklats av vår forskningsgrupp, tillverkas genom att syntetiska antikroppar sätts ned av en robot i mycket små volymer, $3 \cdot 10^{-9}$ l, som små prickar med en diameter av ca 0.1 mm, i ett förutbestämt matrismönster, en så kallad *array*, på en chipyta. Varje array innehåller hundratals olika antikroppar, och varje chip kan innehålla flera likadana sådana arrayer. Antikropparna har på förhand tagits fram från så kallade *antikroppsbibliotek*, som har skapats genom att på syntetisk väg härma immuncellernas förmåga att kombinera olika gener. Antikroppsbiblioteken innehåller flera miljarder unika antikroppar som till stor del ser likadana ut, de är baserade på samma stomme, men de skiljer sig åt i de aminosyror som används för att binda andra molekyler. Ur antikroppsbiblioteket kan man plocka fram antikroppar som binder specifikt till de molekyler man vill undersöka. Antikropparna produceras därefter i bakterier, och renas upp, innan de sätts på chip.

Antikroppschipen används för att analysera kliniska prover, vilka i mitt fall har varit uteslutande blodprover, även om tekniken även fungerar för andra provformat, såsom urin, celler eller vävnad. Varje patientprov analyseras på en enskild array, vars antikroppar fångar de proteiner i provet som de är specifika för. Proteinerna i patientproverna har på förhand märkts in med en fluorescerande molekyl, ett system

som gör att man kan mäta den relativa mängden protein som bundit till varje antikroppsprick genom att läsa in chipet i en laserscanner. På så sätt kan man för varje patient skapa en unik *proteinprofil*. Undersöker man hela grupper av patienter med en viss sjukdom, och jämför deras profiler mot exempelvis friska individer, kan man identifiera *proteinsignaturer*, kombinationer av proteiner som har förändrade blodnivåer i den sjuka gruppen jämfört med kontrollgruppen. Proteinsignaturerna kan då fungera som markörer, likt fingeravtryck, för den specifika sjukdomen. När sådana *biomarkörer* har identifierats måste de först valideras i andra, oberoende prover, innan de kan användas kliniskt för att exempelvis detektera, diagnosticera eller prognosticera sjukdomen.

En sjukdom där det finns ett stort behov av biomarkörer är cancer i bukspottskörteln, *pankreas*. Pankreascancer är en mycket aggressiv form av cancer, och trots att den inte är så vanlig (den drabbar ca 900 svenskar per år), är det den sjätte vanligaste cancerrelaterade dödsorsaken, och endast ca 6% av patienterna överlever 5 år efter diagnos. Anledningen till att så många dör av pankreascancer är att sjukdomen oftast upptäcks i ett alltför sent skede, när tumören har vuxit sig så stor, och ofta även bildat metastaser, att den inte går att operera bort. Den sena detektionen beror dels på att symptomen är vaga och lätt kan förväxlas med andra sjukdomar, dels på att det saknas biomarkörer för tidig diagnos av pankreascancer.

Pankreatit, eller inflammation i bukspottskörteln, är just en sådan sjukdom som pankreascancer lätt kan förväxlas med. Därför är det viktigt att biomarkörer för pankreascancer inte bara kan särskilja cancerpatienter från friska individer, utan även från patienter med pankreatit. I artikel I och II har vi identifierat sådana proteinsignaturer, vilka potentiellt kan fungera som specifika biomarkörer för pankreascancer. Även för pankreatit, som kan vara både av det *akuta*, *kroniska*, eller *autoimmuna* slaget, saknas det känsliga och specifika biomarkörer. I artikel III använde vi våra antikroppschip för att ta fram proteinprofiler för de olika formerna av pankreatit genom att jämföra olika provgrupper mot varandra, men också mot friska individer. De här signaturerna kan dels utgöra potentiella biomarkörer för diagnos av pankreatit, men kan också ge insikt i de molekylära processer som är involverade i de olika typerna av pankreatit.

Parallellt med att jag och andra har utfört kliniska studier med hjälp av våra antikroppschip, har tekniken fortsatt att utvecklas. Jag har varit involverad i flera sådana utvecklingsprojekt, varav två har resulterat i publikationer som presenteras i den här avhandlingen. I artikel IV analyserade vi bland annat chip producerade på olika ytor för att utvärdera vilken yta som lämpade sig bäst för våra antikroppar. Det visade sig att en polymeryta, till vilken antikropparna binder genom enkel adsorption, gav de bästa resultaten, i form av hög *reproducerbarhet* (låg mätvariation), *känslighet*

(låg detektionsgräns), och *dynamisk range* (stort mätområde). Vi visade även att vi kunde tillverka antikroppsarrayer i botten på brunnar i en 96-hålsplatta, något som kan underlätta teknikens användarvänlighet i en eventuell framtida klinisk implementering.

I artikel V undersökte vi möjligheten att expandera arraytekniken för att utöver proteiner även analysera *glykosylerings-* och *karbonyleringsmönster*. Glykosylering och karbonylering är exempel på *post-translationell modifiering*, PTM, av proteiner. PTM är en naturlig process som innebär att proteiner modifieras på olika sätt efter det att deras grundstruktur har skapats, det vill säga efter det att aminosyror har satts ihop till proteiner enligt den genbaserade koden som finns i vårt DNA. Glykosylering innebär att *kolhydratstrukturer* kopplas på proteinet, och karbonylering är en form av *oxidering* av proteinerna. PTM är viktig exempelvis för proteinstabilitet eller signalering, men det har även visat sig att modifieringarna kan påverkas i sjukdomstillstånd. På så sätt kan förändrade nivåer eller mönster av PTM fungera som markörer för sjukdomar, och kan komplettera traditionella biomarkörer baserade på förändringar av proteinnivåer. Genom att introducera en ny typ av reagens till vår analys visade vi att vi kunde mäta både nivåer av glykosylering och karbonylering för olika proteiner på våra antikroppschip. När vi hade visat att tekniken fungerade gjorde vi en första studie på *preeklampsi*, eller *havandeskapsförgiftning*, och kunde då identifiera både glykosylerings- och karbonyleringssignaturer som kunde särskilja patienter med preeklampsi från friska gravida kvinnor.

Sammanfattningsvis handlar den här avhandlingen om tillämpning och utveckling av en antikroppsbaserad microarrayteknik. Dels har min forskning syftat till att optimera och vidareutveckla tekniken, men främst till att applicera analysmetoden för att leta biomarkörer, med särskilt fokus på pankreascancer och pankreatit. Förhoppningsvis kan mina resultat och fortsatta studier i förlängningen leda till att nya biomarkörer för bättre och tidigare diagnos kan introduceras till sjukvården.

Acknowledgements

TACK…

… Christer, för att din dörr alltid står öppen och för att du alltid har en minut över (fast du egentligen inte har det). Du har en förmåga att se möjligheter i allt, och har lyckats entusiasmera och peppa mig även när det känts som svårast.

… Carl, för att du driver en sådan fantastisk institution, och för din inspirerande framåtanda. Ett särskilt tack för ditt stöd i pankreasprojektet, framför allt under de senaste månaderna.

… Linda, för att du introducerade mig till proteinchippens värld. Trots många sena kvällar och blodiga moderkakslabbar var det så pass roligt att göra exjobb för dig att det fick mig att vilja stanna på IT. Ett extra stort tack för alla värdefulla kommentarer till avhandlingen, och framför allt för riktigt effektiv peppning inför slutspurten.

… Anders, för att du har lärt mig allt jag kan om dataanalys. Min avhandling hade inte varit hälften så bra om jag inte hade delat rum med dig och kunnat ställa alla dumma frågor. Proteinchipgruppen hade inte heller varit där den är idag om det inte varit för din kreativitet och kunnighet.

… Malin, för allt trevligt sällskap i labbet och på konferenser, och för din förmåga att lyssna när jag har behövt ventilera stort som smått.

… övriga chippare, Niclas, Linn, Frida, Payam, Petter och Anna för samarbete, stöttning och trevliga stunder.

… mina duktiga exjobbare Anna Å, Sara, Anna S och Erica, för väl genomförda projekt och för att jag har fått utveckla min pedagogiska förmåga med er som försökskaniner.

… Carina, för att du bidrog med en erfaren pipetteringstumme i det sista pankreasprojektet, och Ann-Charlotte för proffsig proteinproduktionshjälp i det första projektet.

… mina kära doktorandkollegor för diskussioner och roligheter under de här åren. Särskilt tack till finaste Henrik och Sissela, jag har uppskattat ert sällskap hela vägen från KC.

… övriga kollegor på IT. Det är sannerligen en lyx att ha en arbetsplats med så många härliga arbetskamrater.

… mina bästa vänner (ingen nämnd ingen glömd) för att ni har försökt förstå vad jag jobbar med och för att ni tror att jag är smart. Jag saknar er alltid.

… Josefin, för det grafiska. Du är en stjärna.

… Gerdt & Annetje för visat stöd och intresse, och för sommarstugan som emellanåt skänkt välbehövlig vila från forskandet.

… min kära familj, för att ni finns. Tack mamma & pappa för att ni alltid har stöttat mig oavsett vad jag tagit mig för, och för att ni låtit mig gå min egen väg.

… Gillis, min skatt. Med dig känns allt så mycket lättare.

… min älskade Erik för alla diskussioner, allt stöd och all kärlek. Det finns ingen bättre!

References

Alonzo, T. A., Pepe, M. S., & Moskowitz, C. S. (2002). Sample size calculations for comparative studies of medical tests for detecting presence of disease. *Stat Med, 21*(6), 835-852.

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests. 1: Sensitivity and specificity. *BMJ, 308*(6943), 1552.

Anderson, N. L. (2010). The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clin Chem, 56*(2), 177-185. doi: 10.1373/clinchem.2009.126706

Anderson, N. L., & Anderson, N. G. (2002). The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics, 1*(11), 845-867.

Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., & Pearson, T. W. (2004). Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res, 3*(2), 235-244.

Anderson, U. D., Olsson, M. G., Kristensen, K. H., Akerstrom, B., & Hansson, S. R. (2012). Review: Biochemical markers to predict preeclampsia. *Placenta, 33 Suppl*, S42-47. doi: 10.1016/j.placenta.2011.11.021

Angenendt, P., Glokler, J., Murphy, D., Lehrach, H., & Cahill, D. J. (2002). Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal Biochem, 309*(2), 253-260.

Angenendt, P., Glokler, J., Sobek, J., Lehrach, H., & Cahill, D. J. (2003). Next generation of protein microarray support materials: evaluation for protein and antibody microarray applications. *J Chromatogr A, 1009*(1-2), 97-104.

Apweiler, R., Hermjakob, H., & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta, 1473*(1), 4-8.

Austin, J., & Holway, A. H. (2011). Contact printing of protein microarrays. *Methods Mol Biol, 785*, 379-394. doi: 10.1007/978-1-61779-286-1_25

Ballehaninna, U. K., & Chamberlain, R. S. (2012). The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J Gastrointest Oncol, 3*(2), 105-119. doi: 10.3978/j.issn.2078-6891.2011.021

Ballehaninna, U. K., & Chamberlain, R. S. (2013). Biomarkers for pancreatic cancer: promising new markers and options beyond CA 19-9. *Tumour Biol.* doi: 10.1007/s13277-013-1033-3

Belov, L., de la Vega, O., dos Remedios, C. G., Mulligan, S. P., & Christopherson, R. I. (2001). Immunophenotyping of leukemias using a cluster of differentiation antibody microarray. *Cancer Res, 61*(11), 4483-4489.

Borrebaeck, C. A. K., & Wingren, C. (2007). High-throughput proteomics using antibody microarrays: an update. *Expert Rev Mol Diagn, 7*(5), 673-686. doi: 10.1586/14737159.7.5.673

Borrebaeck, C. A. K., & Wingren, C. (2009). Design of high-density antibody microarrays for disease proteomics: key technological issues. *J Proteomics, 72*(6), 928-935. doi: 10.1016/j.jprot.2009.01.027

Borrebaeck, C. A. K., & Wingren, C. (2011). Recombinant antibodies for the generation of antibody arrays. *Methods Mol Biol, 785*, 247-262. doi: 10.1007/978-1-61779-286-1_17

Braganza, J. M., Lee, S. H., McCloy, R. F., & McMahon, M. J. (2011). Chronic pancreatitis. *Lancet, 377*(9772), 1184-1197. doi: 10.1016/S0140-6736(10)61852-1

Brand, R. E., Nolen, B. M., Zeh, H. J., Allen, P. J., Eloubeidi, M. A., Goldberg, M., . . . Lokshin, A. E. (2011). Serum biomarker panels for the detection of pancreatic cancer. *Clin Cancer Res, 17*(4), 805-816. doi: 10.1158/1078-0432.CCR-10-0248

Brody, E. N., Gold, L., Lawn, R. M., Walker, J. J., & Zichi, D. (2010). High-content affinity-based proteomics: unlocking protein biomarker discovery. *Expert Rev Mol Diagn, 10*(8), 1013-1022. doi: 10.1586/erm.10.89

Bussow, K., Cahill, D., Nietfeld, W., Bancroft, D., Scherzinger, E., Lehrach, H., & Walter, G. (1998). A method for global protein expression and antibody screening on high-density filters of an arrayed cDNA library. *Nucleic Acids Res, 26*(21), 5007-5008.

Cammann, H., Jung, K., Meyer, H. A., & Stephan, C. (2011). Avoiding pitfalls in applying prediction models, as illustrated by the example of prostate cancer diagnosis. *Clin Chem, 57*(11), 1490-1498. doi: 10.1373/clinchem.2011.166959

Carlsson, A., Persson, O., Ingvarsson, J., Widegren, B., Salford, L., Borrebaeck, C. A. K., & Wingren, C. (2010). Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients. *Proteomics Clin Appl, 4*(6-7), 591-602. doi: 10.1002/prca.200900173

Carlsson, A., Wingren, C., Ingvarsson, J., Ellmark, P., Baldertorp, B., Ferno, M., . . . Borrebaeck, C. A. K. (2008). Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays. *Eur J Cancer, 44*(3), 472-480. doi: 10.1016/j.ejca.2007.11.025

Carlsson, A., Wingren, C., Kristensson, M., Rose, C., Ferno, M., Olsson, H., . . . Borrebaeck, C. A. K. (2011). Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases. *Proc Natl Acad Sci U S A, 108*(34), 14252-14257. doi: 10.1073/pnas.1103125108

Carlsson, A., Wuttge, D. M., Ingvarsson, J., Bengtsson, A. A., Sturfelt, G., Borrebaeck, C. A. K., & Wingren, C. (2011). Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol Cell Proteomics, 10*(5), M110 005033. doi: 10.1074/mcp.M110.005033

Carter, H. B., Pearson, J. D., Metter, E. J., Brant, L. J., Chan, D. W., Andres, R., . . . Walsh, P. C. (1992). Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *JAMA, 267*(16), 2215-2220.

Cava, F., Gonzalez, C., Pascual, M. J., Navajo, J. A., & Gonzalez-Buitrago, J. M. (2000). Biological variation of interleukin 6 (IL-6) and soluble interleukin 2 receptor (sIL2R) in serum of healthy individuals. *Cytokine, 12*(9), 1423-1425. doi: 10.1006/cyto.2000.0714

Chandler, K., & Goldman, R. (2013). Glycoprotein disease markers and single protein-omics. *Mol Cell Proteomics, 12*(4), 836-845. doi: 10.1074/mcp.R112.026930

Cho, E. J., Collett, J. R., Szafranska, A. E., & Ellington, A. D. (2006). Optimization of aptamer microarray technology for multiple protein targets. *Anal Chim Acta, 564*(1), 82-90. doi: 10.1016/j.aca.2005.12.038

Chu, D., Kohlmann, W., & Adler, D. G. (2010). Identification and screening of individuals at increased risk for pancreatic cancer with emphasis on known environmental and genetic factors and hereditary syndromes. *JOP, 11*(3), 203-212.

Cote, G. A., Smith, J., Sherman, S., & Kelly, K. (2013). Technologies for imaging the normal and diseased pancreas. *Gastroenterology, 144*(6), 1262-1271 e1261. doi: 10.1053/j.gastro.2013.01.076

Coussens, L. M., & Werb, Z. (2002). Inflammation and cancer. *Nature, 420*(6917), 860-867. doi: 10.1038/nature01322

Dalle-Donne, I., Giustarini, D., Colombo, R., Rossi, R., & Milzani, A. (2003). Protein carbonylation in human diseases. *Trends Mol Med, 9*(4), 169-176.

Delehanty, J. B. (2004). Printing functional protein microarrays using piezoelectric capillaries. *Methods Mol Biol, 264*, 135-143. doi: 10.1385/1-59259-759-9:135

Dexlin-Mellby, L., Sandstrom, A., Antberg, L., Gunnarsson, J., Hansson, S. R., Borrebaeck, C. A. K, & Wingren, C. (2011). Design of recombinant antibody microarrays for membrane protein profiling of cell lysates and tissue extracts. *Proteomics, 11*(8), 1550-1554. doi: 10.1002/pmic.200900808

Dexlin, L., Ingvarsson, J., Frendeus, B., Borrebaeck, C. A. K, & Wingren, C. (2008). Design of recombinant antibody microarrays for cell surface membrane proteomics. *J Proteome Res, 7*(1), 319-327. doi: 10.1021/pr070257x

Dubel, S., Stoevesandt, O., Taussig, M. J., & Hust, M. (2010). Generating recombinant antibodies to the complete human proteome. *Trends Biotechnol, 28*(7), 333-339. doi: 10.1016/j.tibtech.2010.05.001

Duffy, M. J., Sturgeon, C., Lamerz, R., Haglund, C., Holubec, V. L., Klapdor, R., . . . Heinemann, V. (2010). Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. *Ann Oncol, 21*(3), 441-447. doi: 10.1093/annonc/mdp332

Eisen, M. B., Spellman, P. T., Brown, P. O., & Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A, 95*(25), 14863-14868.

Ekins, R., Chu, F., & Biggart, E. (1990). Multispot, multianalyte, immunoassay. *Ann Biol Clin (Paris), 48*(9), 655-666.

Ekins, R. P., & Chu, F. W. (1991). Multianalyte microspot immunoassay--microanalytical "compact disk" of the future. *Clin Chem, 37*(11), 1955-1967.

Ellington, A. A., Kullo, I. J., Bailey, K. R., & Klee, G. G. (2010). Antibody-based protein multiplex platforms: technical and operational challenges. *Clin Chem, 56*(2), 186-193. doi: 10.1373/clinchem.2009.127514

Ellmark, P., Ingvarsson, J., Carlsson, A., Lundin, B. S., Wingren, C., & Borrebaeck, C. A. K. (2006). Identification of protein expression signatures associated with Helicobacter pylori infection and gastric adenocarcinoma using recombinant antibody microarrays. *Mol Cell Proteomics, 5*(9), 1638-1646. doi: 10.1074/mcp.M600170-MCP200

Engvall, E., & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry, 8*(9), 871-874.

Faca, V. M., Song, K. S., Wang, H., Zhang, Q., Krasnoselsky, A. L., Newcomb, L. F., . . . Hanash, S. M. (2008). A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med, 5*(6), e123. doi: 10.1371/journal.pmed.0050123

Forest, J. C., Charland, M., Masse, J., Bujold, E., Rousseau, F., Lafond, J., & Giguere, Y. (2012). Candidate biochemical markers for screening of pre-eclampsia in early pregnancy. *Clin Chem Lab Med, 50*(6), 973-984. doi: 10.1515/cclm.2011.820

Fry, L. C., Monkemuller, K., & Malfertheiner, P. (2008). Molecular markers of pancreatic cancer: development and clinical relevance. *Langenbecks Arch Surg, 393*(6), 883-890. doi: 10.1007/s00423-007-0276-0

Fulton, R. J., McDade, R. L., Smith, P. L., Kienker, L. J., & Kettman, J. R., Jr. (1997). Advanced multiplexed analysis with the FlowMetrix system. *Clin Chem, 43*(9), 1749-1756.

Fung, E. T. (2010). A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. *Clin Chem, 56*(2), 327-329. doi: 10.1373/clinchem.2009.140855

Ghatnekar, O., Andersson, R., Svensson, M., Persson, U., Ringdahl, U., Zeilon, P., & Borrebaeck, C. A. K. (2013). Modelling the benefits of early diagnosis of pancreatic cancer using a biomarker signature. *International Journal of Cancer, 133*(10), 2392-2397. doi: Doi 10.1002/Ijc.28256

Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., . . . Zichi, D. (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One, 5*(12), e15004. doi: 10.1371/journal.pone.0015004

Good, D. M., Thongboonkerd, V., Novak, J., Bascands, J. L., Schanstra, J. P., Coon, J. J., . . . Mischak, H. (2007). Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future. *J Proteome Res, 6*(12), 4549-4555. doi: 10.1021/pr070529w

Gygi, S. P., Rochon, Y., Franza, B. R., & Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol, 19*(3), 1720-1730.

Haab, B. B. (2005). Antibody arrays in cancer research. *Mol Cell Proteomics, 4*(4), 377-383. doi: 10.1074/mcp.M500010-MCP200

Haab, B. B., Dunham, M. J., & Brown, P. O. (2001). Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol, 2*(2), RESEARCH0004.

Haab, B. B., Geierstanger, B. H., Michailidis, G., Vitzthum, F., Forrester, S., Okon, R., . . . Omenn, G. S. (2005). Immunoassay and antibody microarray analysis of the HUPO Plasma Proteome

Project reference specimens: systematic variation between sample types and calibration of mass spectrometry data. *Proteomics, 5*(13), 3278-3291. doi: 10.1002/pmic.200401276

Haab, B. B., Porter, A., Yue, T., Li, L., Scheiman, J., Anderson, M. A., . . . Simeone, D. M. (2010). Glycosylation variants of mucins and CEACAMs as candidate biomarkers for the diagnosis of pancreatic cystic neoplasms. *Ann Surg, 251*(5), 937-945. doi: 10.1097/SLA.0b013e3181d7738d

Haab, B. B., & Yue, T. (2011). High-throughput studies of protein glycoforms using antibody-lectin sandwich arrays. *Methods Mol Biol, 785*, 223-236. doi: 10.1007/978-1-61779-286-1_15

Hamelinck, D., Zhou, H., Li, L., Verweij, C., Dillon, D., Feng, Z., . . . Haab, B. B. (2005). Optimized normalization for antibody microarrays and application to serum-protein profiling. *Mol Cell Proteomics, 4*(6), 773-784. doi: 10.1074/mcp.M400180-MCP200

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674. doi: 10.1016/j.cell.2011.02.013

Hanash, S. M., Pitteri, S. J., & Faca, V. M. (2008). Mining the plasma proteome for cancer biomarkers. *Nature, 452*(7187), 571-579. doi: 10.1038/nature06916

Hassan, M. M., Bondy, M. L., Wolff, R. A., Abbruzzese, J. L., Vauthey, J. N., Pisters, P. W., . . . Li, D. (2007). Risk factors for pancreatic cancer: case-control study. *Am J Gastroenterol, 102*(12), 2696-2707. doi: 10.1111/j.1572-0241.2007.01510.x

Hassanein, M., Callison, J. C., Callaway-Lane, C., Aldrich, M. C., Grogan, E. L., & Massion, P. P. (2012). The state of molecular biomarkers for the early detection of lung cancer. *Cancer Prev Res (Phila), 5*(8), 992-1006. doi: 10.1158/1940-6207.CAPR-11-0441

Heimburg-Molinaro, J., Song, X., Smith, D. F., & Cummings, R. D. (2011). Preparation and analysis of glycan microarrays. *Curr Protoc Protein Sci, Chapter 12*, Unit12 10. doi: 10.1002/0471140864.ps1210s64

Hidalgo, M. (2010). Pancreatic cancer. *N Engl J Med, 362*(17), 1605-1617. doi: 10.1056/NEJMra0901557

Hudson, M. E., Pozdnyakova, I., Haines, K., Mor, G., & Snyder, M. (2007). Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. *Proc Natl Acad Sci U S A, 104*(44), 17494-17499. doi: 10.1073/pnas.0708572104

Hust, M., Steinwand, M., Al-Halabi, L., Helmsing, S., Schirrmann, T., & Dubel, S. (2009). Improved microtitre plate production of single chain Fv fragments in Escherichia coli. *N Biotechnol, 25*(6), 424-428. doi: 10.1016/j.nbt.2009.03.004

Ingvarsson, J., Larsson, A., Sjoholm, A. G., Truedsson, L., Jansson, B., Borrebaeck, C. A. K., & Wingren, C. (2007). Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins. *J Proteome Res, 6*(9), 3527-3536. doi: 10.1021/pr070204f

Ingvarsson, J., Wingren, C., Carlsson, A., Ellmark, P., Wahren, B., Engstrom, G., . . . Borrebaeck, C. A. K. (2008). Detection of pancreatic cancer using antibody microarray-based serum protein profiling. *Proteomics, 8*(11), 2211-2219. doi: 10.1002/pmic.200701167

Issaq, H. J., Waybright, T. J., & Veenstra, T. D. (2011). Cancer biomarker discovery: Opportunities and pitfalls in analytical methods. *Electrophoresis, 32*(9), 967-975. doi: 10.1002/elps.201000588

Jorgensen, J. T. (2009). New era of personalized medicine: a 10-year anniversary. *Oncologist, 14*(5), 557-558. doi: 10.1634/theoncologist.2009-0047

Karas, M., & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem, 60*(20), 2299-2301.

Keshishian, H., Addona, T., Burgess, M., Kuhn, E., & Carr, S. A. (2007). Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics, 6*(12), 2212-2229. doi: 10.1074/mcp.M700354-MCP200

Kim, B. K., Lee, J. W., Park, P. J., Shin, Y. S., Lee, W. Y., Lee, K. A., . . . Kim, C. W. (2009). The multiplex bead array approach to identifying serum biomarkers associated with breast cancer. *Breast Cancer Res, 11*(2), R22. doi: 10.1186/bcr2247

Knezevic, V., Leethanakul, C., Bichsel, V. E., Worth, J. M., Prabhu, V. V., Gutkind, J. S., . . . Krizman, D. B. (2001). Proteomic profiling of the cancer microenvironment by antibody arrays. *Proteomics, 1*(10), 1271-1278. doi: 10.1002/1615-9861(200110)1:10<1271::AID-PROT1271>3.0.CO;2-6

Kohler, G., & Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature, 256*(5517), 495-497.

Konstantinou, F., Syrigos, K. N., & Saif, M. W. (2013). Pancreatic cancer: what about screening and detection? *JOP, 14*(4), 312-315. doi: 10.6092/1590-8577/1645

Koopmann, J., Rosenzweig, C. N., Zhang, Z., Canto, M. I., Brown, D. A., Hunter, M., . . . Goggins, M. (2006). Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. *Clin Cancer Res, 12*(2), 442-446. doi: 10.1158/1078-0432.CCR-05-0564

Koprowski, H., Herlyn, M., Steplewski, Z., & Sears, H. F. (1981). Specific antigen in serum of patients with colon carcinoma. *Science, 212*(4490), 53-55.

Kristensson, M., Olsson, K., Carlson, J., Wullt, B., Sturfelt, G., Borrebaeck, C. A. K., & Wingren, C. (2012). Design of recombinant antibody microarrays for urinary proteomics. *Proteomics Clin Appl, 6*(5-6), 291-296. doi: 10.1002/prca.201100055

Kusnezow, W., & Hoheisel, J. D. (2003). Solid supports for microarray immunoassays. *J Mol Recognit, 16*(4), 165-176. doi: 10.1002/jmr.625

Langreth, R., & Waldholz, M. (1999). New era of personalized medicine: targeting drugs for each unique genetic profile. *Oncologist, 4*(5), 426-427.

Larghi, A., Verna, E. C., Lecca, P. G., & Costamagna, G. (2009). Screening for pancreatic cancer in high-risk individuals: a call for endoscopic ultrasound. *Clin Cancer Res, 15*(6), 1907-1914. doi: 10.1158/1078-0432.CCR-08-1966

Lippi, G., Valentino, M., & Cervellin, G. (2012). Laboratory diagnosis of acute pancreatitis: in search of the Holy Grail. *Crit Rev Clin Lab Sci, 49*(1), 18-31. doi: 10.3109/10408363.2012.658354

Lippitz, B. E. (2013). Cytokine patterns in patients with cancer: a systematic review. *Lancet Oncol, 14*(6), e218-228. doi: 10.1016/S1470-2045(12)70582-X

Liu, L., Aa, J., Wang, G., Yan, B., Zhang, Y., Wang, X., . . . Wu, Z. (2010). Differences in metabolite profile between blood plasma and serum. *Anal Biochem, 406*(2), 105-112. doi: 10.1016/j.ab.2010.07.015

Locker, G. Y., Hamilton, S., Harris, J., Jessup, J. M., Kemeny, N., Macdonald, J. S., . . . Asco. (2006). ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol, 24*(33), 5313-5327. doi: 10.1200/JCO.2006.08.2644

Lohr, J. M. (2007). What are the useful biological and functional markers of early-stage chronic pancreatitis? *J Gastroenterol, 42 Suppl 17*, 66-71. doi: 10.1007/s00535-006-1932-9

Ludwig, J. A., & Weinstein, J. N. (2005). Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer, 5*(11), 845-856. doi: 10.1038/nrc1739

MacBeath, G., & Schreiber, S. L. (2000). Printing proteins as microarrays for high-throughput function determination. *Science, 289*(5485), 1760-1763.

Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem, 72*(6), 1156-1162.

Makawita, S., Dimitromanolakis, A., Soosaipillai, A., Soleas, I., Chan, A., Gallinger, S., . . . Diamandis, E. P. (2013). Validation of four candidate pancreatic cancer serological biomarkers that improve the performance of CA19.9. *BMC Cancer, 13*(1), 404. doi: 10.1186/1471-2407-13-404

Manjer, J., Carlsson, S., Elmstahl, S., Gullberg, B., Janzon, L., Lindstrom, M., . . . Berglund, G. (2001). The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. *Eur J Cancer Prev, 10*(6), 489-499.

McCafferty, J., Griffiths, A. D., Winter, G., & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature, 348*(6301), 552-554. doi: 10.1038/348552a0

Michaud, D. S. (2004). Epidemiology of pancreatic cancer. *Minerva Chir, 59*(2), 99-111.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., . . . et al. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science, 266*(5182), 66-71.

Miller, J. C., Zhou, H., Kwekel, J., Cavallo, R., Burke, J., Butler, E. B., . . . Haab, B. B. (2003). Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers. *Proteomics, 3*(1), 56-63. doi: 10.1002/pmic.200390009

Morgan, E., Varro, R., Sepulveda, H., Ember, J. A., Apgar, J., Wilson, J., . . . Gaur, A. (2004). Cytometric bead array: a multiplexed assay platform with applications in various areas of biology. *Clin Immunol, 110*(3), 252-266. doi: 10.1016/j.clim.2003.11.017

Mustafa, S. A., Hoheisel, J. D., & Alhamdani, M. S. (2011). Secretome profiling with antibody microarrays. *Mol Biosyst, 7*(6), 1795-1801. doi: 10.1039/c1mb05071k

Nielsen, U. B., & Geierstanger, B. H. (2004). Multiplexed sandwich assays in microarray format. *J Immunol Methods, 290*(1-2), 107-120. doi: 10.1016/j.jim.2004.04.012

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem, 250*(10), 4007-4021.

Olsson, N., James, P., Borrebaeck, C. A. K., & Wingren, C. (2012). Quantitative proteomics targeting classes of motif-containing peptides using immunoaffinity-based mass spectrometry. *Mol Cell Proteomics, 11*(8), 342-354. doi: 10.1074/mcp.M111.016238

Olsson, N., Wingren, C., Mattsson, M., James, P., O'Connell, D., Nilsson, F., . . . Borrebaeck, C. A. K. (2011). Proteomic analysis and discovery using affinity proteomics and mass spectrometry. *Mol Cell Proteomics, 10*(10), M110 003962. doi: 10.1074/mcp.M110.003962

Orchekowski, R., Hamelinck, D., Li, L., Gliwa, E., vanBrocklin, M., Marrero, J. A., . . . Haab, B. B. (2005). Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer. *Cancer Res, 65*(23), 11193-11202. doi: 10.1158/0008-5472.CAN-05-1436

Orenes-Pinero, E., Barderas, R., Rico, D., Casal, J. I., Gonzalez-Pisano, D., Navajo, J., . . . Sanchez-Carbayo, M. (2010). Serum and tissue profiling in bladder cancer combining protein and tissue arrays. *J Proteome Res, 9*(1), 164-173. doi: 10.1021/pr900273u

Ostroff, R., Foreman, T., Keeney, T. R., Stratford, S., Walker, J. J., & Zichi, D. (2010). The stability of the circulating human proteome to variations in sample collection and handling procedures measured with an aptamer-based proteomics array. *J Proteomics, 73*(3), 649-666. doi: 10.1016/j.jprot.2009.09.004

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., . . . Wolmark, N. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med, 351*(27), 2817-2826. doi: 10.1056/NEJMoa041588

Paik, S., Tang, G., Shak, S., Kim, C., Baker, J., Kim, W., . . . Wolmark, N. (2006). Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol, 24*(23), 3726-3734. doi: 10.1200/JCO.2005.04.7985

Pannala, R., Basu, A., Petersen, G. M., & Chari, S. T. (2009). New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer. *Lancet Oncol, 10*(1), 88-95. doi: 10.1016/S1470-2045(08)70337-1

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., . . . Varmus, H. (2004). EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A, 101*(36), 13306-13311. doi: 10.1073/pnas.0405220101

Park, T., Yi, S. G., Kang, S. H., Lee, S., Lee, Y. S., & Simon, R. (2003). Evaluation of normalization methods for microarray data. *BMC Bioinformatics, 4*, 33. doi: 10.1186/1471-2105-4-33

Parsons, B. L., & Meng, F. (2009). K-RAS mutation in the screening, prognosis and treatment of cancer. *Biomark Med, 3*(6), 757-769. doi: 10.2217/bmm.09.95

Pauly, F., Dexlin-Mellby, L., Ek, S., Ohlin, M., Olsson, N., Jirstrom, K., . . . Wingren, C. (2013). Protein expression profiling of formalin-fixed paraffin-embedded tissue using recombinant antibody microarrays. *J Proteome Res.* doi: 10.1021/pr4003245

Paweletz, C. P., Charboneau, L., Bichsel, V. E., Simone, N. L., Chen, T., Gillespie, J. W., . . . Liotta, L. A. (2001). Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. *Oncogene, 20*(16), 1981-1989. doi: 10.1038/sj.onc.1204265

Pleskow, D. K., Berger, H. J., Gyves, J., Allen, E., McLean, A., & Podolsky, D. K. (1989). Evaluation of a serologic marker, CA19-9, in the diagnosis of pancreatic cancer. *Ann Intern Med, 110*(9), 704-709.

Poetz, O., Hoeppe, S., Templin, M. F., Stoll, D., & Joos, T. O. (2009). Proteome wide screening using peptide affinity capture. *Proteomics, 9*(6), 1518-1523. doi: 10.1002/pmic.200800842

Quackenbush, J. (2001). Computational analysis of microarray data. *Nat Rev Genet, 2*(6), 418-427. doi: 10.1038/35076576

Ransohoff, D. F. (2005). Bias as a threat to the validity of cancer molecular-marker research. *Nat Rev Cancer, 5*(2), 142-149. doi: 10.1038/nrc1550

Renberg, B., Shiroyama, I., Engfeldt, T., Nygren, P. K., & Karlstrom, A. E. (2005). Affibody protein capture microarrays: synthesis and evaluation of random and directed immobilization of affibody molecules. *Anal Biochem, 341*(2), 334-343. doi: 10.1016/j.ab.2005.03.039

Rifai, N., Gillette, M. A., & Carr, S. A. (2006). Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol, 24*(8), 971-983. doi: 10.1038/nbt1235

Sah, R. P., & Chari, S. T. (2012). Autoimmune pancreatitis: an update on classification, diagnosis, natural history and management. *Curr Gastroenterol Rep, 14*(2), 95-105. doi: 10.1007/s11894-012-0246-8

Sanchez-Carbayo, M. (2006). Antibody arrays: technical considerations and clinical applications in cancer. *Clin Chem, 52*(9), 1651-1659. doi: 10.1373/clinchem.2005.059592

Sanchez-Carbayo, M. (2011). Antibody microarrays as tools for biomarker discovery. *Methods Mol Biol, 785*, 159-182. doi: 10.1007/978-1-61779-286-1_11

Sanchez-Carbayo, M., Socci, N. D., Lozano, J. J., Haab, B. B., & Cordon-Cardo, C. (2006). Profiling bladder cancer using targeted antibody arrays. *Am J Pathol, 168*(1), 93-103. doi: 10.2353/ajpath.2006.050601

Schaferling, M., & Nagl, S. (2006). Optical technologies for the read out and quality control of DNA and protein microarrays. *Anal Bioanal Chem, 385*(3), 500-517. doi: 10.1007/s00216-006-0317-5

Schechter, A. L., Stern, D. F., Vaidyanathan, L., Decker, S. J., Drebin, J. A., Greene, M. I., & Weinberg, R. A. (1984). The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen. *Nature, 312*(5994), 513-516.

Schroder, C., Jacob, A., Tonack, S., Radon, T. P., Sill, M., Zucknick, M., . . . Hoheisel, J. D. (2010). Dual-color proteomic profiling of complex samples with a microarray of 810 cancer-related antibodies. *Mol Cell Proteomics, 9*(6), 1271-1280. doi: 10.1074/mcp.M900419-MCP200

Schweitzer, B., Meng, L., Mattoon, D., & Rai, A. J. (2010). Immune response biomarker profiling application on ProtoArray protein microarrays. *Methods Mol Biol, 641*, 243-252. doi: 10.1007/978-1-60761-711-2_14

Schwenk, J. M., Igel, U., Kato, B. S., Nicholson, G., Karpe, F., Uhlen, M., & Nilsson, P. (2010). Comparative protein profiling of serum and plasma using an antibody suspension bead array approach. *Proteomics, 10*(3), 532-540. doi: 10.1002/pmic.200900657

Seurynck-Servoss, S. L., Baird, C. L., Miller, K. D., Pefaur, N. B., Gonzalez, R. M., Apiyo, D. O., . . . Zangar, R. C. (2008). Immobilization strategies for single-chain antibody microarrays. *Proteomics, 8*(11), 2199-2210. doi: 10.1002/pmic.200701036

Seurynck-Servoss, S. L., White, A. M., Baird, C. L., Rodland, K. D., & Zangar, R. C. (2007). Evaluation of surface chemistries for antibody microarrays. *Anal Biochem, 371*(1), 105-115. doi: 10.1016/j.ab.2007.07.010

Shaib, Y. H., Davila, J. A., & El-Serag, H. B. (2006). The epidemiology of pancreatic cancer in the United States: changes below the surface. *Aliment Pharmacol Ther, 24*(1), 87-94. doi: 10.1111/j.1365-2036.2006.02961.x

Siegel, R., Naishadham, D., & Jemal, A. (2012). Cancer statistics, 2012. *CA Cancer J Clin, 62*(1), 10-29. doi: 10.3322/caac.20138

Silvestri, A., Colombatti, A., Calvert, V. S., Deng, J., Mammano, E., Belluco, C., . . . Pierobon, M. (2010). Protein pathway biomarker analysis of human cancer reveals requirement for upfront cellular-enrichment processing. *Lab Invest, 90*(5), 787-796. doi: 10.1038/labinvest.2010.47

Soderlind, E., Strandberg, L., Jirholt, P., Kobayashi, N., Alexeiva, V., Aberg, A. M., . . . Borrebaeck, C. A. K. (2000). Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat Biotechnol, 18*(8), 852-856. doi: 10.1038/78458

Steinhauer, C., Ressine, A., Marko-Varga, G., Laurell, T., Borrebaeck, C. A. K., & Wingren, C. (2005). Biocompatibility of surfaces for antibody microarrays: design of macroporous silicon substrates. *Anal Biochem, 341*(2), 204-213. doi: 10.1016/j.ab.2004.10.036

Steinhauer, C., Wingren, C., Hager, A. C., & Borrebaeck, C. A. K. (2002). Single framework recombinant antibody fragments designed for protein chip applications. *Biotechniques, Suppl*, 38-45.

Steinhauer, C., Wingren, C., Khan, F., He, M., Taussig, M. J., & Borrebaeck, C. A. K. (2006). Improved affinity coupling for antibody microarrays: engineering of double-(His)6-tagged single framework recombinant antibody fragments. *Proteomics, 6*(15), 4227-4234. doi: 10.1002/pmic.200600036

Stoevesandt, O., & Taussig, M. J. (2007). Affinity reagent resources for human proteome detection: initiatives and perspectives. *Proteomics, 7*(16), 2738-2750. doi: 10.1002/pmic.200700155

Stoevesandt, O., & Taussig, M. J. (2012). European and international collaboration in affinity proteomics. *N Biotechnol, 29*(5), 511-514. doi: 10.1016/j.nbt.2012.05.003

Stoevesandt, O., Vetter, M., Kastelic, D., Palmer, E. A., He, M., & Taussig, M. J. (2011). Cell free expression put on the spot: advances in repeatable protein arraying from DNA (DAPA). *N Biotechnol, 28*(3), 282-290. doi: 10.1016/j.nbt.2010.09.004

Sultana, R., & Butterfield, D. A. (2013). Oxidative modification of brain proteins in Alzheimer's disease: perspective on future studies based on results of redox proteomics studies. *J Alzheimers Dis, 33 Suppl 1*, S243-251. doi: 10.3233/JAD-2012-129018

Surinova, S., Schiess, R., Huttenhain, R., Cerciello, F., Wollscheid, B., & Aebersold, R. (2011). On the development of plasma protein biomarkers. *J Proteome Res, 10*(1), 5-16. doi: 10.1021/pr1008515

Tersmette, A. C., Petersen, G. M., Offerhaus, G. J., Falatko, F. C., Brune, K. A., Goggins, M., . . . Hruban, R. H. (2001). Increased risk of incident pancreatic cancer among first-degree relatives of patients with familial pancreatic cancer. *Clin Cancer Res, 7*(3), 738-744.

Thompson, I. M., Pauler, D. K., Goodman, P. J., Tangen, C. M., Lucia, M. S., Parnes, H. L., . . . Coltman, C. A., Jr. (2004). Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med, 350*(22), 2239-2246. doi: 10.1056/NEJMoa031918

Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., & Veenstra, T. D. (2003). Characterization of the low molecular weight human serum proteome. *Mol Cell Proteomics, 2*(10), 1096-1103. doi: 10.1074/mcp.M300031-MCP200

Truong, M., Yang, B., & Jarrard, D. F. (2013). Toward the detection of prostate cancer in urine: a critical analysis. *J Urol, 189*(2), 422-429. doi: 10.1016/j.juro.2012.04.143

Tuxen, M. K., Soletormos, G., Petersen, P. H., Schioler, V., & Dombernowsky, P. (1999). Assessment of biological variation and analytical imprecision of CA 125, CEA, and TPA in relation to monitoring of ovarian cancer. *Gynecol Oncol, 74*(1), 12-22. doi: 10.1006/gyno.1999.5414

Wacker, R., Schroder, H., & Niemeyer, C. M. (2004). Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavidin-biotin attachment: a comparative study. *Anal Biochem, 330*(2), 281-287. doi: 10.1016/j.ab.2004.03.017

Wakeling, A. E., Guy, S. P., Woodburn, J. R., Ashton, S. E., Curry, B. J., Barker, A. J., & Gibson, K. H. (2002). ZD1839 (Iressa): an orally active inhibitor of epidermal growth factor signaling with potential for cancer therapy. *Cancer Res, 62*(20), 5749-5754.

van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., . . . Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature, 415*(6871), 530-536. doi: 10.1038/415530a

Whitcomb, D. C. (2013). Genetic risk factors for pancreatic disorders. *Gastroenterology, 144*(6), 1292-1302. doi: 10.1053/j.gastro.2013.01.069

Wingren, C., & Borrebaeck, C. A. K. (2006). Antibody microarrays: current status and key technological advances. *OMICS, 10*(3), 411-427. doi: 10.1089/omi.2006.10.411

Wingren, C., & Borrebaeck, C. A. K. (2007). Progress in miniaturization of protein arrays--a step closer to high-density nanoarrays. *Drug Discov Today, 12*(19-20), 813-819. doi: 10.1016/j.drudis.2007.08.003

Wingren, C., & Borrebaeck, C. A. K. (2008). Antibody microarray analysis of directly labelled complex proteomes. *Curr Opin Biotechnol, 19*(1), 55-61. doi: 10.1016/j.copbio.2007.11.010

Wingren, C., Ingvarsson, J., Dexlin, L., Szul, D., & Borrebaeck, C. A. K. (2007). Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support. *Proteomics, 7*(17), 3055-3065. doi: 10.1002/pmic.200700025

Winter, J. M., Yeo, C. J., & Brody, J. R. (2013). Diagnostic, prognostic, and predictive biomarkers in pancreatic cancer. *J Surg Oncol, 107*(1), 15-22. doi: 10.1002/jso.23192

Vonlaufen, A., Wilson, J. S., & Apte, M. V. (2008). Molecular mechanisms of pancreatitis: current opinion. *J Gastroenterol Hepatol, 23*(9), 1339-1348. doi: 10.1111/j.1440-1746.2008.05520.x

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., . . . Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature, 378*(6559), 789-792. doi: 10.1038/378789a0

Wulfkuhle, J. D., Aquino, J. A., Calvert, V. S., Fishman, D. A., Coukos, G., Liotta, L. A., & Petricoin, E. F., 3rd. (2003). Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. *Proteomics, 3*(11), 2085-2090. doi: 10.1002/pmic.200300591

Xue, A., Scarlett, C. J., Chung, L., Butturini, G., Scarpa, A., Gandy, R., . . . Smith, R. C. (2010). Discovery of serum biomarkers for pancreatic adenocarcinoma using proteomic analysis. *Br J Cancer, 103*(3), 391-400. doi: 10.1038/sj.bjc.6605764

Yadav, D., Hawes, R. H., Brand, R. E., Anderson, M. A., Money, M. E., Banks, P. A., . . . North American Pancreatic Study, G. (2009). Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. *Arch Intern Med, 169*(11), 1035-1045. doi: 10.1001/archinternmed.2009.125

Yalow, R. S., & Berson, S. A. (1959). Assay of plasma insulin in human subjects by immunological methods. *Nature, 184 (Suppl 21),* 1648-1649.

Yue, T., Goldstein, I. J., Hollingsworth, M. A., Kaul, K., Brand, R. E., & Haab, B. B. (2009). The prevalence and nature of glycan alterations on specific proteins in pancreatic cancer patients revealed using antibody-lectin sandwich arrays. *Mol Cell Proteomics, 8*(7), 1697-1707. doi: 10.1074/mcp.M900135-MCP200

Yue, T., Maupin, K. A., Fallon, B., Li, L., Partyka, K., Anderson, M. A., . . . Haab, B. B. (2011). Enhanced discrimination of malignant from benign pancreatic disease by measuring the CA 19-9 antigen on specific protein carriers. *PLoS One, 6*(12), e29180. doi: 10.1371/journal.pone.0029180

Zeh, H. J., Winikoff, S., Landsittel, D. P., Gorelik, E., Marrangoni, A. M., Velikokhatnaya, L., . . . Lokshin, A. E. (2005). Multianalyte profiling of serum cytokines for detection of pancreatic cancer. *Cancer Biomark, 1*(6), 259-269.

Zhang, J. Y. (2004). Tumor-associated antigen arrays to enhance antibody detection for cancer diagnosis. *Cancer Detect Prev, 28*(2), 114-118. doi: 10.1016/j.cdp.2003.12.006

Zusterzeel, P. L., Rutten, H., Roelofs, H. M., Peters, W. H., & Steegers, E. A. (2001). Protein carbonyls in decidua and placenta of pre-eclamptic women as markers for oxidative stress. *Placenta, 22*(2-3), 213-219. doi: 10.1053/plac.2000.0606

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem, 39*(4), 561-577.