NMR-based Metabolomics: Global Analysis of Metabolites to Address Problems in Prostate Cancer

Matthew J. Roberts
Centre for Clinical Research / Centre for Advanced Imaging
The University of Queensland, Australia

Horst Joachim Schirra
Centre for Advanced Imaging
The University of Queensland, Australia

Martin F. Lavin
Centre for Clinical Research
The University of Queensland, Australia
Radiation Biology and Oncology
Queensland Institute of Medical Research, Australia

Robert A. Gardiner
Centre for Clinical Research
The University of Queensland, Australia
Department of Urology
Royal Brisbane and Women's Hospital, Queensland, Australia

1 Introduction

Cancer significantly contributes to the worldwide burden of disease and premature death across many countries. Consequently, current oncology research focuses on discovering and validating new biomarkers to improve early detection. These efforts are of worldwide importance in detecting significant cancer while it is still localised and in lessening associated morbidities and death.

Biomarkers have mostly been sourced from non- or minimally invasive biofluids, such as blood, urine, and biopsy tissue. Traditionally, biomarkers were limited to circulating end-products of altered cellular function in cancer. However, technology advances and emergence of the –omics sciences have improved analysis of genes, gene expression, proteins and metabolites alike – on both an individual and system-wide scale. This field of research, termed "systems biology", has allowed for molecules at all levels of the cellular hierarchy to be considered as biomarkers. Continuous improvements in sensitivity, resolution and precision of these analytical techniques produces large datasets, allowing for simultaneous characterisation of, ideally all, compounds in a single sample. Subsequent statistical analysis of these datasets and their interpretation with respect to cellular function is the basis of the different -omics technologies, such as genomics, transcriptomics, proteomics and metabolomics.

In this chapter, we will describe principles and processes that are involved in investigating biological or clinical problems with nuclear magnetic resonance (NMR)-based metabolomics - an approach that involves the global analysis of metabolites. In writing for the scope of this book, we have broken this chapter into three sections: (1) First we will describe and illustrate the methods commonly used in NMR-based metabolomics, including spectral processing, data treatment and subsequent statistical analysis. (2) Secondly, we will use prostate cancer (PCa) as a case study to illustrate how NMR-based metabolomics can be applied to a clinical problem. PCa is the second most common type of cancer and the sixth leading cause of cancer-related death worldwide (Center *et al*., 2012; Ferlay *et al*., 2010; Siegel *et al*., 2012). The diagnosis of prostate cancer is currently problematic for a number of reasons that include lack of sensitive and specific tumour markers as well as limitations due to morbidity inherent with the biopsy diagnosis process. Furthermore, many patients harbour early prostate cancer with insignificant tumours that may not progress to produce clinical problems. (3) Lastly, we will briefly outline the future directions for the role of NMR-based metabolomics, including personalized medicine and integration with other –omics datasets, in order to create a holistic, systems biology approach to solving clinical problems.

Outlining the processes, applications and potential of metabolomics will be of assistance to biostatisticians and bioinformaticians who may be interested in expanding into this area of research. Similarly, we aim to inspire scientists and clinicians who are interested in applying this approach to a scientific or clinical problem.

2 Metabolomics: History and Methods

2.1 What is Metabolomics?

Metabolomics has been highlighted as a technique that is unique and exciting in biomarker discovery (Abate-Shen & Shen, 2009; Bino *et al.*, 2004; Nicholson & Lindon, 2008). It is the quantification of all small molecular weight metabolites to accurately define the metabolite composition of a biological sample (Fiehn, 2002). The term "metabolomics" is often used interchangeably with "metabonomics", which

is defined as 'the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli' (Nicholson *et al.,* 1999). The technical approach to both metabonomics and metabolomics is similar, involving measurement and analysis of metabolite data for a given sample. Conceptually, however, the objective and application of these techniques is slightly different: Metabolomics seeks to describe the composition of complex biological samples, while metabonomics aims to map and understand the change of a biological system in response to external or artificial stimuli. In the current chapter, we will use the term "metabolomics", while both metabolomics and metabonomics will be discussed. The advent and development of metabolomics/metabonomics has largely been possible due to advances in analytical techniques, such as nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS), as well as chromatographic separation techniques. As the authors have expertise in NMR spectroscopy, the focus of this chapter will be on NMR-based techniques within the field of metabolomics. We refer the reader to several excellent reviews that detail the application of MS in metabolomics and systems biology (Lu *et al.*, 2008; Pasikanti *et al.*, 2008; Theodoridis *et al.*, 2008)

2.2 Historical Perspective

While "metabolomics" is a recently coined term, the analysis of metabolic end products is a long-practiced ancient scientific process. About 2000-1500 BCE, analysis of urine by human taste or animal behaviour (due to the high urinary glucose concentration) helped diagnose patients with diabetes mellitus (van der Greef & Smilde, 2005). In 1506, a "urine wheel" by Ulrich Pinder linked crude changes detected by human senses (colour, smell, taste) with various medical conditions (Pinder, 1506; Weiss & Kim, 2012). Although qualitative measures of metabolism have been performed for centuries, the origins of quantitative studies stem from the measurement of insensible perspiration and other hydration losses in medieval Italy (Eknoyan, 1999). Technological limitations in analytical chemistry hindered further advances until the early 20th century. At this time, the development of sensitive analytical methods allowed quantification of key compounds/metabolites in urine and other sample types (Simoni *et al.*, 2002). This quantitative approach continued to develop with the introduction of various analytical techniques, such as mass spectrometry (Griffiths, 2008; Thomson, 1912) and NMR (Bloch *et al.*, 1946; Freeman, 1995; Purcell *et al.*, 1946), and with the application of these techniques to metabolic research (Gates & Sweeley, 1978; Hoult *et al.*, 1974). The integration of medical science and analytical chemistry at this time led to a greater understanding of metabolic perturbations in medical conditions, e.g. kidney stones among many others (Reginato & Kurnik, 1989). Further improvements in metabolite profiling in the latter half of the 20th century were aided by advances in chemometrics, the foundation behind data analysis in analytical chemistry (Geladi & Esbensen, 1990). Appropriate data processing and interpretation was achieved by multivariate statistical methods, which will be outlined in more detail below.

2.3 Modern Metabolomics and its Varieties/Applications

Metabolomics continues to evolve as a field and is increasingly used in a variety of applications. Initially, biofluids were analyzed specifically to quantify metabolic perturbations due to drug toxicity, disease, and other internal and external influences. For example, perturbations in steroid metabolism were used for anti-doping testing during the Los Angeles Summer Olympics in 1984 (Fitch, 2008). In the following section, the diverse applications of metabolomics are briefly described; more comprehensive accounts are available in other sources (Duarte & Gil, 2012; Ma *et al.*, 2012b; Ng *et al.,* 2011; Rhee & Gerszten, 2012; Spratlin *et al.*, 2009).

2.3.1 Biofluid and Excretion Analysis

Metabolomics has been used to quantify endogenous metabolites in many human biofluids, with those most commonly analyzed being urine and blood (serum, plasma). Analysis of urine metabolites has shown early promise in diagnosing kidney (Ganti & Weiss, 2011) and bladder tumors (Hyndman *et al.*, 2011), as well as more systemic conditions such as type 2 diabetes mellitus (Salek *et al.*, 2007). Indeed, distinct serum metabolite patterns have been characterized for abnormal clinical states including breast cancer (Oakman *et al.*, 2011), leukemia (MacIntyre *et al.*, 2010), sepsis and acute lung injury (Serkova *et al.*, 2011), coronary artery disease (Brindle *et al.*, 2002) and obesity (Oberbach *et al.*, 2011; Xie *et al.*, 2012). Furthermore, cardiovascular health has been assessed by metabolomic analysis of feces, linking perturbations in the metabolite profiles of gut flora to the metabolic syndrome and dyslipidemia (Wang *et al.*, 2011). Seminal fluid and expressed prostatic secretions (EPS), have been used to characterize disturbed metabolism in prostate cancer (Averna *et al.*, 2005; Kline *et al.*, 2006; Lynch *et al.*, 1994, 1997; Serkova *et al.*, 2008) and infertility (Deepinder *et al.*, 2007; Hamamah *et al.*, 1993, 1998). Studies on cerebrospinal fluid have associated metabolite changes with brain tumors and neurodegenerative disorders, such as multiple sclerosis and Alzheimer's disease (Han *et al.*, 2011). Salivary metabolomics has been used to investigate oral cancer and pre-malignant changes, such as leukoplakia (Wei *et al*., 2011). Thus, this minimally invasive approach has enormous potential in providing valuable scientific and clinical information for medical professionals and researchers.

2.3.2 NMR Spectroscopy of Tissues and *in vivo* Imaging Techniques

High-resolution magic-angle spinning (HR-MAS) NMR spectroscopy can be used to perform non-destructive metabolite profiling of tissue or other solid samples (Keifer, 2007). That means that after HR-MAS NMR spectroscopy, further testing of tissue samples can be performed such as histopathological evaluation, the current gold standard in disease diagnosis, or genome and protein sequencing. As a result, HR-MAS NMR has been used to investigate a number of disease states (Brown *et al.*, 2012; Dittrich *et al.*, 2012; Kurhanewicz *et al.*, 2002; Martínez-Bisbal *et al.*, 2004; Maxeiner *et al.*, 2010; Millis *et al.*, 1997; Sitter *et al.*, 2002).

Metabolomic analysis can also be performed directly *in vivo*, largely owing to advances in magnetic resonance spectroscopic imaging and positron emission tomography (PET). A standard clinical magnetic resonance imaging (MRI) scan uses similar physical concepts to NMR, but takes many scans across a section of living tissue. This data is processed to produce an anatomically correct image based on physical properties of the tissue. A magnetic resonance spectroscopy imaging sequence is able to produce NMR-like spectra for a targeted volume segment in the body, allowing for visualization of metabolite content in that anatomical location. Metabolic alterations measured *in vivo* have been shown to correlate with histopathology (Delongchamps *et al.*, 2011; Kwock *et al.*, 2006). Furthermore, *in vivo* metabolomics is being used to monitor the response to various therapies, such as radiotherapy and chemotherapy (Coy *et al.*, 2011; Lodi & Ronen, 2011).

Different MRI techniques allow the investigation of different phenomena. Dynamic contrast-enhanced MRI uses the uptake and elimination of contrast agents, such as gadolinium, to distinguish between different tissues. The use of dynamic contrast-enhanced MRI in oncology is based on the premise that cancer cells have a higher metabolism, and thus a higher uptake and elimination of gadolinium contrast (Vargas *et al.*, 2012). Diffusion weighted imaging, initially used to investigate connectivity between different brain regions (Behrens *et al.*, 2003), uses slower water diffusion in cancerous tissues compared

with surrounding healthy cells due to a higher nuclear content and cellular density coupled with extracellular changes (Lim & Tan, 2012). Recently, these MRI techniques have been combined into multiparametric MRI, which has been shown to increase accuracy in cancer detection (Hoeks *et al.*, 2011). The value of these *in vivo* applications and their role in oncology is commonly described and reviewed in radiology literature (Hoeks *et al.*, 2011; Hoh *et al.*, 1997; Koh & Collins, 2007; Padhani, 2002).

Further understanding of altered metabolism in cancer and identification of abnormal pathways facilitates imaging using PET in combination with computed tomography (CT) via PET/CT (Basu *et al.*, 2011; Jones & Price, 2012). After identifying metabolites, that are either preferentially used or upregulated within particular pathways, nuclear isotopes can be chemically attached either to these metabolites or to metabolite analogues. The emission of positrons from these isotopes can then be measured as gamma rays and superimposed on a CT scan during PET/CT scanning. For example, most cancer cells display heightened glycolysis. Thus, fluoro-deoxy-glucose, containing a radiolabeled positron emitter such as ^{18}F, can be administered and taken up by cancer cells, which are highlighted (Costello & Franklin, 2005). In addition, PET is able to distinguish specific cancers, e.g. ^{11}C-choline PET is used to detect prostate cancer (Evangelista *et al.*, 2012; Reske *et al.*, 2006).The application of PET in other clinical scenarios is diverse, but widespread use is limited by logistical and financial constraints (Fletcher *et al.*, 2008; Kelloff *et al.*, 2005; Rohren *et al.*, 2004).

Recently, a novel method has been proposed to perform *in vivo* metabolomics during surgery by using real-time MS analysis of the smoke produced from electric cautery to biochemically recognize malignant/diseased tissue in which macroscopic changes are not present (Balog *et al.*, 2010; Kinross *et al.*, 2011). Although major development is required before clinical use, the initial concept is intriguing in its potential to improve surgical accuracy and treatment outcomes following cancer surgery.

2.4 Integration with other -Omics Sciences

Metabolites are part of the complex and interconnected cellular hierarchy involving DNA, RNA, proteins and metabolites, as outlined in Figure 1. Consequently, providing an understanding of the mutual relationships between genomic, transcriptomic, proteomic and metabolomic data is a major aim of systems biology.

Integration of multi-omics data sets is already providing insight into biological processes. This integration is enabled by the availability of new statistical methods to correlate information contained in multiple large datasets (Rantalainen *et al.*, 2006). Furthermore, ever increasing genome-wide association studies (GWAS) are identifying multiple risk loci associated with various disease states. However, the penetrance of these loci, and therefore their relevance, remains unclear. By integrating metabolomics and other –omics sciences with GWAS, it is anticipated that identification of loci with a high penetrance/phenotypic manifestations will be unveiled (Suhre *et al.*, 2011; Weckwerth, 2011). Integration of –omics data sets has major potential in oncology (Casado-Vela *et al.*, 2011) and some studies have used a targeted approach in relating datasets obtained by different analytical methods (Rantalainen *et al.*, 2006). A recently published study related metabolomic changes to genomic disturbances in PCa tissue to demonstrate alterations in m-aconitase and acetyl citrate lyase. Phospholipase A2 group VII and choline kinase α were responsible for altered citrate and choline levels, respectively (Bertilsson *et al.*, 2012). Other integrative works investigated colorectal cancer (Ma *et al.*, 2012a), heart failure (Lin *et al.*, 2011) and other diseases (Adamski, 2012; Chen *et al.*, 2012). These studies show impressive proof of concept

of this new approach. As a result, multivariate statistical analysis and integration of large and multi-omics data sets are a valuable strategy for further investigation.

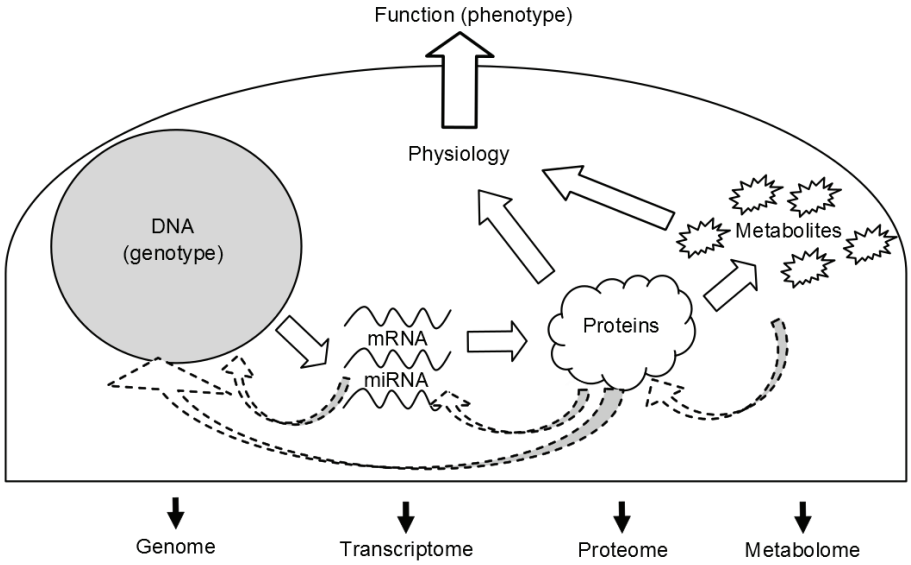

Figure 1: Major targets for exploratory analysis in systems biology. The flow of information and biochemical processes between various levels in cellular organisation illustrates the progression from genotype to phenotype (solid arrows). The individual levels of cellular organisation are also regulated by complex intrinsic feedback mechanisms (dashed arrows). Adapted from (Roberts *et al*., 2011)

Another key frontier in systems biology is the creation of genome-wide *in silico* models of cellular metabolism that are able to incorporate and integrate multi-omic data (Bordbar & Palsson, 2012). Such reconstructed metabolic networks have already suggested improvements for targeted treatment strategies of cholesterol homeostasis in human cellular models (Mo *et al*., 2007).

The first genome-scale model of human metabolism published in 2007 was a promising milestone (Duarte *et al*., 2007). The next step in metabolic modeling in oncology will be to construct cancer-specific metabolic models that will incorporate –omics data. This daunting task requires improved curation and annotation of genome databases, as well as integration of high quality –omics datasets from studies of specific cancer types, or of other disease states to create disease-specific reconstructions. Public accessibility and maintenance of data from publicly funded research is critical to these efforts (Field *et al*., 2009), as is data management (Sansone *et al*., 2012) and interpretation (Leader *et al*., 2011).

2.5 Analytical Techniques

Of the many analytical techniques that are used in metabolomics to investigate physiological and pathological states, NMR spectroscopy and gas chromatography (GC) or liquid chromatography (LC) combined with MS are the two most commonly used methods. Both have low running costs, are diverse in sample type and allow for accurate metabolite identification. Other techniques in use are ultra-

performance LC-MS, inductively coupled plasma MS, Fourier-transform MS, Fourier-transform infrared spectrometry and thin layer chromatography (Zhang *et al*., 2012). We will briefly describe the processes involved in MS-based metabolomics, with more extensive reviews available elsewhere (Dunn *et al.*, 2005; Dunn & Ellis, 2005; Lu *et al.*, 2008; Mishur & Rea, 2012; Pasikanti *et al.*, 2008; Theodoridis *et al.*, 2008). Subsequently, we will focus on NMR-based metabolomics, describing the basic principles and statistical approaches that are currently in use.

2.5.1 GC-/LC-MS

Mass spectrometry detects ionized compounds in biological samples according to their mass/charge (m/z) ratio following chromatographic (e.g. GC or LC) separation, and metabolites can be identified in the resulting mass spectrum with reference to internal standards (Wilson *et al.*, 2005). LC- and GC-MS are well used techniques in metabolite analysis and have similar sensitivity, with the major difference being that GC requires more sample preparation (derivatization) and higher analysis temperatures, thus LC may be preferred for this reason (Issaq *et al.*, 2009; Mishur & Rea, 2012). We will briefly outline the basic processes in MS-based metabolomics, which are sample preparation, separation (via liquid-/gas-chromatography), ionization, mass analysis and detection, and finally, data processing.

Sample preparation for MS is dependent on the type of sample. Simple biofluid preparation often involves removing macromolecules through protein precipitation and centrifugation or filtration. Similarly, sampling of the exometabolome (metabolites secreted by cells or organisms into the growth medium) is straightforward. In contrast, to obtain intracellular metabolites, tissues or cells need to be extracted in an appropriate solvent system (Mishur & Rea, 2012). As different solvent systems are biased towards particular classes of metabolites (e.g. polar extraction systems yielding predominantly polar metabolites), the exact choice of solvent depends upon the metabolites of interest. To achieve consistency between samples, internal standards are typically added during/after extraction (Mishur & Rea, 2012).

In GC-MS, derivatization is a further necessary preparation step, which is applied to non-volatile metabolite classes, such as amino and organic acids, sugars, amines and lipids, to render them volatile and thermally stable for GC (Dunn *et al.*, 2005; Dunn & Ellis, 2005; Mishur & Rea, 2012). Derivatization can introduce bias towards individual metabolites if the derivatizing agents are not provided in excess, as the derivatizing reactions have different efficiencies with different metabolites (Mishur & Rea 2012). In addition, metabolites with multiple exchangeable protons will create multiple derivatization products that will show up as separate peaks, thus complicating the final mass spectrum. For GC-MS, electron impact ionization is almost exclusively used (Dunn *et al.*, 2005).

LC-MS is rapidly replacing GC-MS as method of choice in metabolomics, as both methods are similarly sensitive, but sample preparation for LC-MS is simpler, because derivatization is not required (Griffin & Shockcor, 2004). LC typically runs as reverse-phase high-performance LC (HPLC), or recently even as ultra-performance LC (UPLC), and electrospray ionization is typically used in LC-MS systems (Mishur & Rea, 2012). Electrospray ionization MS can run in positive or negative ionization mode, and because individual metabolites are generally only detected in one of those two modes, both ionization modes need to be run to improve coverage of the metabolome (Dunn & Ellis, 2005).

In the resulting mass spectrum, metabolites are quantified by external calibration or by comparison with internal standards (Dunn & Ellis, 2005). GC-MS experiments may also require the use of deconvolution software to adequately analyse overlapping chromatographic peaks (Dunn & Ellis, 2005). To allow comparable results between experiments, data may also undergo further pre-treatment steps, including

spectral alignment and automated picking of metabolite peaks (Dettmer *et al.*, 2007). The subsequent multivariate analysis of processed MS data is similar to data obtained by NMR-spectroscopy.

2.5.2 NMR Spectroscopy

NMR spectroscopy is a quantitative technique used to accurately determine metabolite concentrations in samples. Chemical compounds in biological samples are identified by their characteristic peak patterns and signal positions in the NMR spectrum with the aid of online databases (Wishart *et al.*, 2009) (http://www.metabolomicssociety.org/database). More comprehensive accounts of NMR theory and application in metabolomics are available in dedicated texts (Ross *et al.*, 2007). A comparison of the strengths and limitations of NMR and MS is provided in Table 1.

Technique	Advantages	Limitations
NMR	- high reproducibility - high resolution - non-destructive - quantitative - low running costs - minimal sample preparation/no derivatization - unbiased metabolite profile - analysis of tissue (HR-MAS) - translation to *in vivo* (MRI) - rapid analysis - ability for automation - structural identification (2D, 3D)	- low sensitivity - peak overlap - libraries of limited use due to complex matrix - long acquisition times for heteronuclear techniques, e.g. ^{13}C - high initial capital cost - reduced availability
GC-MS	- high sensitivity - large linear range - robust - identification of wide range of metabolites (wider range with LC-MS) - analysis of complex biofluids - non-targeted - established databases - widely available and comparably low capital cost - preferred for targeted analysis	- slow - sample unable to be re-used - requires chemical derivatization - potentially multiple derivatization products for metabolites - many analytes thermally unstable - metabolite weight limitation (<1400 Da)
LC-MS	- high sensitivity - high reproducibility - large linear range - no chemical derivatization needed	- slow - limited commercial libraries - sample unable to be re-used - generation of adducts - higher capital cost (HPLC-MS)

Table 1: Comparison of the advantages and disadvantages of NMR, GC-MS, and LC-MS (Issaq *et al.*, 2009; Shepherd *et al.*, 2011; Shulaev, 2006; Zhang *et al.*, 2012).

2.6 Processing of NMR Data

Processing of NMR data comprises four steps: Fourier transformation, phase correction, baseline correction and calibration. Fourier transformation transforms the raw real-time data into the frequency domain, and phase correction corrects the phase of the resulting NMR spectrum. Baseline correction ensures a constant zero baseline across an NMR spectrum, and calibration is needed to ensure a consistent chemical shift scale/axis across all spectra. Inadequate processing introduces artifacts that confound statistical analysis and jeopardize data integrity (see Figure 2). Minimum standards for reporting and processing have been outlined (Goodacre *et al*., 2007), and continue to be a good guide for authors publishing metabolomics research.

Usually, NMR spectra require a phase correction following Fourier transformation in order to achieve pure absorptive line shapes for all peaks in a NMR spectrum. Where possible, phasing should be performed automatically or by the same operator across all samples to ensure consistency. Incorrect phasing can distort peak integrals and thus, the subsequent multivariate statistical analysis (MVSA; see Figure 2A/2B).

Baseline correction is the third critical processing step in producing consistent, comparable and reliable data in NMR spectroscopy. At a minimum, the y-offset of the entire spectrum is corrected to be zero. However, baseline corrections are often more complex, using spline, polynomial or other mathematical functions to accomplish a zero baseline over the whole spectrum. As signal intensities are calculated with reference to zero, inadequate baseline correction will distort spectral peak intensities (Figure 2B/2C), and compromise the subsequent MVSA.

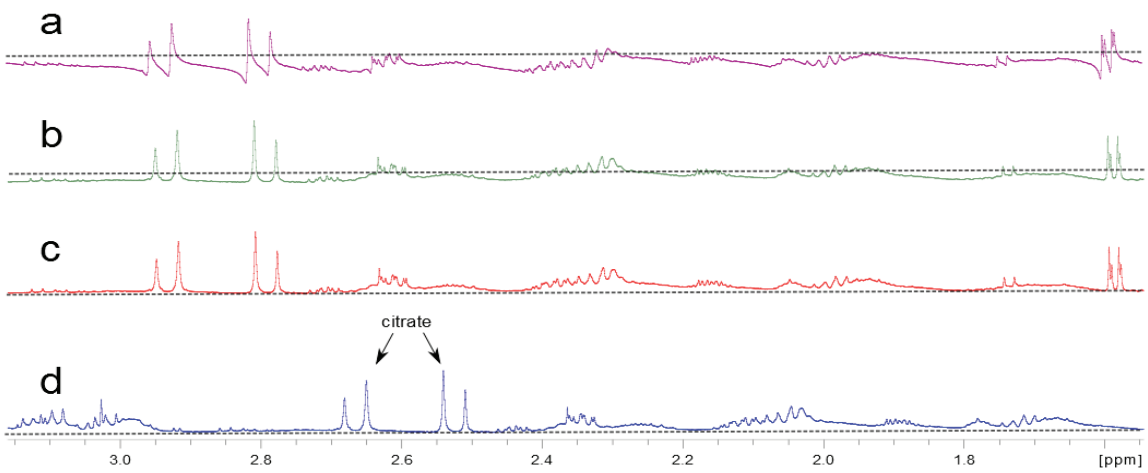

Figure 2: Sequential steps of processing NMR spectra. Shown is a spectrum of human ejaculate in the region of citrate (2.45 – 2.65 ppm) as example. Correct baseline position is illustrated with dotted lines. a - Spectrum after Fourier transformation. The phases of the NMR signals are partly dispersive and in need of phase correction. b - Spectrum after phase correction. Globally reduced metabolite peak intensities and negative values for the baseline occur across the entire spectrum. c - Spectrum after baseline correction, but requiring calibration to a chemical shift standard. Chemical shift values are incorrect across the entire spectrum, resulting in incorrect metabolite identification. d - Correctly processed spectrum suitable for data reduction.

Importantly, the chemical shift axis of each NMR spectrum must be adequately calibrated using a chemical shift standard, such as (deuterated) 4,4-dimethyl-4-silapentane-1-sulfonic acid-d_6 (DSS). In samples containing a high protein content (such as plasma), DSS cannot be used as internal standard. Thus, other endogenous metabolites that are present across all samples, such as lactate, glucose or formate, are used as internal reference. The alternative is to use DSS as an *external* standard by either inserting a capillary with DSS in deuterium oxide (D_2O) into the NMR tube, or inserting the sample in a capillary into a tube containing DSS in D_2O. The use of 4,4-dimethyl-4-silapentane-1-ammonium trifluoroacetate (DSA) has also been suggested as chemical shift standard that is not affected by protein binding (Alum *et al.*, 2008). Chemical shift calibration ensures consistent global alignment of spectra in a metabolomics data set which is critical for statistical analysis. In addition, correct spectral alignment is required for reliable metabolite identification. However, even in correctly calibrated spectra, individual peaks can still exhibit differences in chemical shift between individual spectra due to differences in sample pH and ionic strength. These can be corrected post-processing by various automatic peak alignment procedures (Anderson *et al.*, 2011; Giskeødegård *et al.*, 2010; MacKinnon *et al.*, 2012; Savorani *et al.*, 2010; Staab *et al.*, 2010; Wright *et al.*, 2012).

2.7 Statistical Pre-processing

2.7.1 Data Reduction

After processing of raw NMR data, further processing steps are needed to prepare data for MVSA, which are usually termed "statistical pre-processing". Reducing the full resolution data into small segments of equal width, called bins, or "buckets" (Figure 3), is the most widespread method of data reduction in chemometrics (Wishart, 2008). Compared with analysis at full resolution, this method considerably reduces the size of the data matrix in MVSA, and is particularly helpful when peak positions or widths vary slightly due to changes in pH, ionic strength or other factors. However, due to decreasing data resolution, bucketing can complicate metabolite identification following data analysis. Other pre-processing methods that can be used, particularly in targeted metabolomics, include deconvolution, peak-picking, and weighting factors (Goodacre *et al.*, 2007).

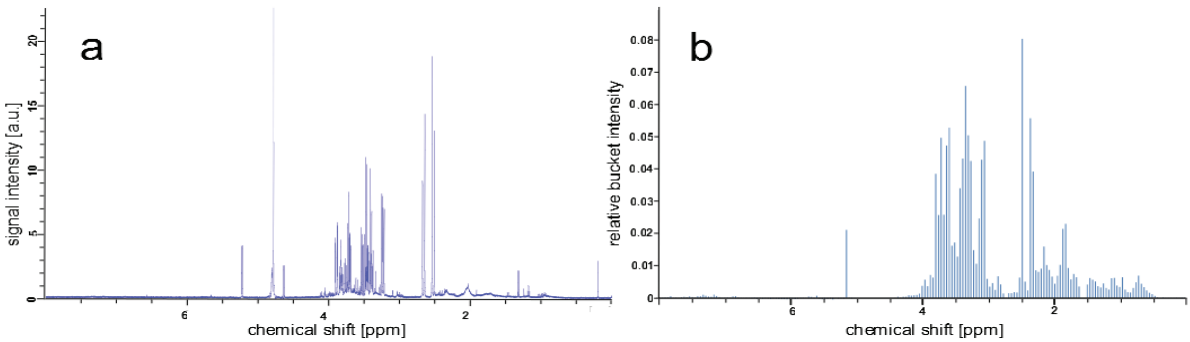

Figure 3: Illustration of data reduction by "bucketing". a - NMR spectrum of human ejaculate with suitable pre-processing (segment shown). b – Same spectrum segmented/data reduced into buckets of 0.04 ppm width across the region 8 – 0 ppm. The region around the water signal from 5.08 – 4.52 ppm was excluded due to artifacts from imperfect water suppression. Note that the area in each individual bucket is integrated across the spectral width of 0.04 ppm and then normalized, yielding intensities similar to a histogram.

2.7.2 Normalization

After data reduction, data need to be normalized to produce data that are comparable between samples (Craig *et al*., 2006). Normalization is a row operation in the data matrix, and different normalization methods are used to obtain the best representation of the data. In total integral normalization, or normalizing to total intensity, the spectral intensity in each bucket is divided by the total intensity of each spectrum. This procedure normalizes differences between spectra due to sample concentration/dilution, e.g. due to different water content between samples. However, total integral normalization is vulnerable to distortions when one or a few intense signals change considerably between spectra.

Another method involves normalization to an internal reference compound. For metabolomics analysis of urine, normalization to creatinine has been widely used (Akira *et al*., 2008; Jentzmik *et al*., 2010). For physiological reasons, urine creatinine is believed to be a suitable indicator of urine concentration as creatinine excretion is constant. However, creatinine normalization has limitations because it will be confounded by any background pathophysiology that alters serum creatinine levels or creatinine excretion, such as in kidney disease. In these cases, creatinine normalization is not suitable. Furthermore, variations in chemical properties within the sample can distort creatinine alignment, so that other metabolites, especially creatine, will overlap with the creatinine signal and thus impede the proper measurement of creatinine concentration (Ross *et al*., 2007).

Probabilistic quotient normalization (PQN) is a method that reduces variation caused by large changes in the intensity of one or a few signals across samples, as shown in Figure 4. Thus, PQN can overcome the main weakness of total integral normalization. In PQN, which is usually performed after total integral normalization, each variable (bucket) in a spectrum is first divided by the intensity of the same variable in a reference spectrum. Afterwards the full spectrum is divided by the median of these quotients. This procedure is repeated for all spectra in a data set, using the same reference spectrum (Dieterle *et al*., 2006).

Recently, Kohl *et al*. compared many normalization methods, with some derived from genomic data analysis, and recommended more advanced methods, such as quantile normalization for datasets of $n \geq 50$, as well as Cubic Spline Normalization and Variance Stabilisation Normalization (Kohl *et al*., 2012).

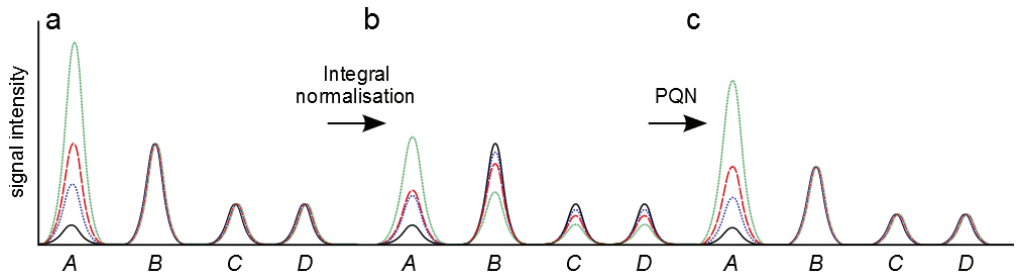

Figure 4: Schematic depiction of spectra containing four NMR signals *A*-*D*, following two different normalization methods. a – raw spectra with marked intensity variation present in the first peak *A* and identical intensities of remaining peaks between all spectra. b - Integral normalization (normalization to total intensity) reduces variation, and therefore influence of the dominant signal *A*, but also alters relative intensities of the smaller, previously identical signals. c - Probabilistic quotient normalization (PQN) partially reduces variation and influence of the larger signal *A*, while maintaining the original relationships between smaller peaks to allow optimal comparison during MVSA.

2.7.3 Scaling Effects

Following normalization, metabolomic data must be appropriately scaled, or transformed in a column operation in a way that changes how much signals of large and small intensity, respectively, influence the data analysis (Craig *et al*., 2006; Goodacre *et al*., 2007). The objective is to reduce noise and maximize information content in the data. Inappropriate scaling may lead to results that highlight parts of the data unrelated to a biological factor, thus compromising the analysis and biological interpretation of the data. In metabolomics statistical analysis, three scaling methods are largely used.

Centre scaling, or mean centering, subtracts the mean value of each variable/bucket from the original data of that bucket (Craig *et al*., 2006). This method is the least manipulative, and is also best at minimizing background noise, but large relative variations in small signals may not be detected. Mean centering is usually performed mandatorily. Thus, this method is also sometimes referred to as "no scaling", as no further scaling is performed after mean centering.

Univariate, or unit variance, scaling divides the raw data obtained after mean centering by the standard deviation of each variable. Univariate scaling gives each variable equal weighting, such that variables with small absolute but large relative variation are highlighted, but this also means that background noise and other unrelated data variation may be overemphasized and thus confound the analysis (Craig *et al*., 2006).

Pareto scaling is performed by dividing each variable by the square root of its standard deviation (Erikson *et al*., 1999). This is the recommended scaling method for NMR-based metabolomic data, as it is able to increase the weighting on metabolites with smaller amplitudes, but does not overemphasize the influence of background noise.

2.8 Statistical Analysis

Statistical analysis of metabolomic data depends on the biological question studied and the design of the particular project, thus, the choice of data analysis methods varies between different projects. In addition, the methods for data analysis are continually evolving. Nevertheless, there is a core set of methods of univariate and multivariate statistical analysis that is in use for metabolomics, and minimum reporting standards for data analysis have been established (Goodacre *et al*., 2007; Liberati *et al*., 2009; Schulz *et al*., 2010).

2.8.1 Univariate Statistical Analysis

The role of univariate analysis in metabolomics is largely of a targeted nature. An example would be where metabolites of interest have been identified by MVSA, and detailed analysis of statistical significance of the individual metabolites is desired. Basic univariate methods can be used to analyze whether or not individual metabolites are significantly different between two classes. However, as with any statistical analysis, the distribution of the data determines the type of analysis used. If the data are normally distributed, t-tests, z-tests, and analysis of variance (ANOVA) may be used. In cases where the distribution is not normal, non-parametric methods such as the Kruskal-Wallis test are used (Goodacre *et al*., 2007). However, given the high number of variables within a metabolomics dataset, use of multiple hypothesis testing corrections, such as Bonferroni correction or false discovery rate/Benjamini-Hochberg are absolutely imperative (Noble, 2009). This means that, to be significant, *p*-values need to be much

smaller (e.g. $p \leq 5 \times 10^{-5}$) after correcting for multiple hypothesis testing compared to standard univariate statistical analysis (Broadhurst & Kell, 2006).

2.8.2 Multivariate Statistical Analysis (MVSA)

Modern day metabolomics is largely based on data sets incorporating many variables, between several hundred in the case of bucketed data and up to 65,536 if 1D-NMR spectra are used at full resolution. Thus, MVSA methods which simultaneously analyze all these variables are preferred (Broadhurst & Kell, 2006). MVSA determines whether there are inherent patterns or groupings within the data that correspond to biological states and also which variables are important in discriminating between the different groupings. Thus, this approach is well suited to analyzing metabolomics datasets, where the aim is to correlate multiple metabolite changes with alterations in biology.

Approach	Method	Advantages	Disadvantages
Unsupervised	Principal components analysis (PCA)	• Simplifies data • Describes variation in original data without bias • Groups samples with similar metabolite profiles	• Variation might be unrelated to biological question • Influenced by *any* confounders
Supervised	Partial least squares (PLS)	• Simplifies data • Groups samples with similar metabolite profiles • Extracts variation that is correlated with external data/identifiers • more directed to the biological question	• Possibility of introducing bias • Require rigorous validation
	Orthogonal projections to latent structures (OPLS)	• As PLS • Removes orthogonal (unrelated) variation • Improved knowledge extraction	
	2-way OPLS (O2PLS)	• As OPLS • Two-way data correlation between ***X*** and ***Y*** • Potential for unsupervised analysis (when analyzing two large datasets without external ***Y*** table)	
	Kernel OPLS	• As OPLS • Improved model prediction	
	O*n*PLS	• As O2PLS • Simultaneous data correlation from multiple (*n*) matrices	• Possibility of introducing bias • Require rigorous validation • Not widely available (commercial/open source)
	Bi-modal O*n*PLS	• As O*n*PLS • Data correlation between variables (columns) and samples (rows)	
	O*n*PLS path modelling	• As O*n*PLS • Linkage of matrices along statistically related paths	

Table 2: Summary of multivariate statistical analysis methods used in metabolomics for information recovery. The advantages and disadvantages of unsupervised and supervised methods are outlined.

There are two general classes of MVSA methods: unsupervised methods, which analyze patterns within a data matrix ***X***, and supervised methods in which the patterns in ***X*** are correlated with other external data (e.g. clinical data) contained in a ***Y*** matrix or ***Y*** table. The advantages and disadvantages of the most common MVSA methods are summarized in Table 2, and will be discussed in the following sections. Some examples of software programs, both commercial and free/open-source, that are available to perform MVSA are given in Table 3.

Software Package	Use in NMR/MS	Reference/Source
Commercial		
SIMCA	Both	http://www.umetrics.com
MATLAB	Both	http://www.mathworks.com
MarkerLynx/MassLynx	MS	http://www.waters.com
STATISTICA Data Miner	Both	http://www.statsoft.com
AMIX	Both	http://www.bruker.com
Agilent Mass Profiler Professional	MS	http://metabolomics.chem.agilent.com
Progenesis CoMet	MS	http://www.nonlinear.com
Free/Open Source		
R	Both	http://www.r-project.org
Metaboanalyst	Both	http://www.metaboanalyst.ca (Xia *et al*., 2012)
MAVEN	MS	http://maven.princeton.edu (Melamud *et al*., 2010)
MZmine	MS	http:// mzmine.sourceforge.net (Pluskal *et al*., 2010)
MeltDB	MS	https://meltdb.cebitec.uni-bielefeld.de/ (Neuweger *et al*., 2008)
MetabolomeExpress	MS	https://www.metabolome-express.org (Carroll *et al*., 2010)

Table 3: Examples of software packages used to perform multivariate statistical analysis in metabolomics.

2.8.3 Unsupervised Methods in Multivariate Statistical Analysis

Often, unsupervised analysis methods are initially used in MVSA as they are excellent data exploration tools that can be either used to simplify the data (dimensionality reduction – principal- or independent-components analysis) or to group samples with similar metabolite patterns (clustering – hierarchal, partitional). Although many different methods exist, principal components analysis (PCA) is the most commonly used method of unsupervised analysis in metabolomics.

PCA simplifies the original multivariate data which form a swarm of data points in a high-dimensional statistical space, by projecting them down into a new space with comparatively few dimensions called principal components (PCs). These PCs are latent variables that describe the variation in the original data. The first PC indicates the direction in which most variation occurs in the data. Subsequent PCs are all orthogonal to each other and sorted in order of descending amount of variation. This arrangement describes the majority of variation in the data within the first few PCs. PCA is visualized by two types of plots, the scores and loadings plots (Figure 5).

The scores plot illustrates the relationship and similarity of samples to each other, and allows for inspection of groupings and outliers. Outliers can be visually identified on the scores plot, or statistically defined as being outside the Hotelling's 95% confidence range across all components. Further statistical validation can be obtained using the residual variance of the model, known as distance to model plot

(Erikson *et al*., 1999; Trygg *et al*., 2007). The variables (buckets, ultimately metabolites) attributable to each component in the model are illustrated in the loadings plot (see Figure 5; Denkert *et al*., 2006; Pan *et al*., 2007).

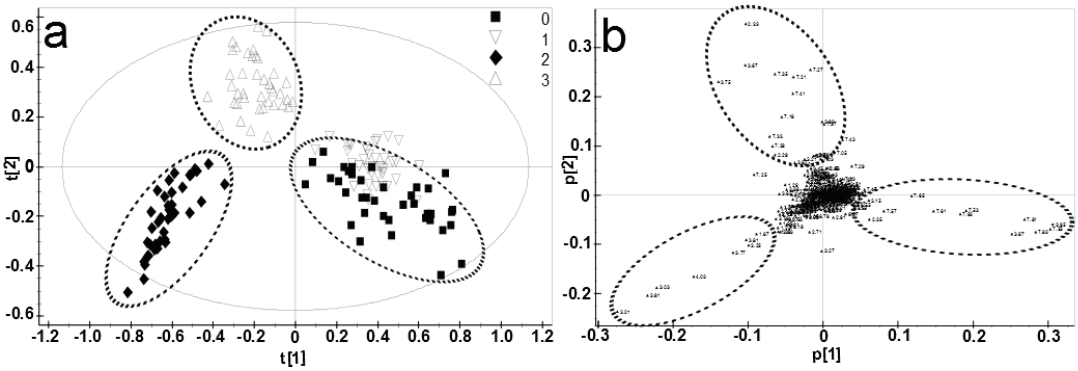

Figure 5: Visualization of a typical multivariate analysis (e.g. PCA, PLS etc.). Sheep urine samples before and after road transport of 12 and 48 hours are shown as an example (Li *et al*., 2010). a – Scores plot – shows any relationships between samples, such as the presence of separate groups or outliers. This scores plot shows similarity of the pre-transport groups for both transport durations (open inverted triangle = 12 h, black squares = 48 h), and differences between both post-transport groups (48 hours = black diamonds, 12 hours = open triangles). b – Loadings plot – displays the relationship of influential variables that are responsible for the position of outliers or groups seen in the scores plot. Note that the positions/directions of groupings/outliers in the scores plot and responsible variables in the loadings plot correspond to each other (Trygg *et al*., 2007).

The inherent advantage of unsupervised methods is that they are unbiased, i.e. they detect *any* statistical variation in the data, whether or not it is related to the underlying biological effect (e.g. differences between cancer and non-cancer samples). This property is also their most noticeable limitation, because when confounding effects are larger than the biological effect, unsupervised analyses will predominantly show the effects of these confounding factors. Distorting variation may also come from uncorrelated background variation, or from noise. For this reason, robust experimental design that limits confounding factors and appropriate pre-processing prior to data analysis are vitally important to ensure meaningful results.

2.8.4 Supervised Methods in Multivariate Statistical Analysis (MVSA)

MVSA can be improved by including external data, such as clinical data, in a ***Y*** table or ***Y*** matrix. This data inclusion then makes it possible to use a different class of MVSA methods, which are called supervised analysis methods. The biggest advantage of supervised methods is that they can identify the variation (and associated variables) in the biological data that is *correlated* (or co-varies) with the external data, i.e. they improve information recovery and thus interpretation of the biological data. The main supervised methods used in metabolomics-based biomedical research are partial least squares (PLS) and

orthogonal projections to latent structures (OPLS) (Gu *et al*., 2011; Pan *et al*., 2007; Trygg & Lundstedt, 2007).

PLS is a method that seeks to identify correlation between the dataset matrix (***X***) and one or multiple variables (contained in ***Y***). ***Y*** data may be categorical (class identities, e.g. healthy vs. disease) or continuous (blood pressure, height etc.). If the external variable(s) are qualitative, then the method will discriminate between the corresponding classes and is known as PLS discriminant analysis (PLS-DA) (Trygg *et al*., 2007). Figure 6 illustrates this distinction. Supervised analyses in metabolomics can be affected by systematic variation that is unrelated to the class, as this affects any correlation found by the analysis method.

0.5	0.46	0.42	0.38	0.34	0.3	0.26	STATUS	SerumPSA
-0.00026	-0.00029	-0.00022	-0.00024	-0.0002	-0.00015	-0.00023	A	6.7
-0.00079	-0.00081	-0.00078	-0.00081	-0.00087	-0.0008	-0.00081	A	5.1
-0.00034	-0.00033	-0.00045	-0.00041	-0.00045	-0.00045	-0.00049	A	8.2
6.37E-07	3.44E-05	1.58E-05	-2.4E-05	-2.1E-05	1.13E-06	-3.8E-05	A	3.4
1.67E-05	-1.7E-05	-4.2E-05	-6.9E-05	-9E-05	-8.1E-05	-9.6E-05	A	5.4
-0.00025	-0.00031	-0.00031	-0.00031	-0.00031	-0.0003	-0.00034	A	11
-0.00048	-0.00053	-0.00052	-0.00054	-0.00058	-0.00059	-0.00062	A	8.2
3.64E-05	5.5E-05	2.67E-05	-3.4E-06	-2E-05	2.86E-05	-3.7E-05	A	6
0.00017	0.000109	0.000103	6.84E-05	3.05E-05	4.6E-05	3.34E-05	A	2.5
-0.00037	-0.00037	-0.00036	-0.00041	-0.00041	-0.00041	-0.00044	A	8.3
1.51E-05	-3.6E-05	-6.5E-05	-9.2E-05	-0.00013	-0.00015	-0.00016	A	4.7
7.17E-05	4.02E-05	1.94E-05	5.16E-06	-9.5E-06	-4E-05	-6.1E-05	A	2.2
-0.00056	-0.00053	-0.0005	-0.00047	-0.00049	-0.00047	-0.00045	A	6.4
3.4E-05	-1.2E-05	5.92E-06	-6.8E-05	-6.5E-05	-8.5E-05	-0.00011	A	0.78
-0.00013	-0.00018	-0.00014	-0.00019	-0.00016	-0.00013	-0.0002	B	4.1
0.000497	0.000455	0.000421	0.000418	0.000276	0.000294	0.000266	B	7.9
0.000208	0.000193	0.000153	0.000141	0.000113	0.00011	5.38E-05	B	5.6
-0.00016	-0.00019	-0.00014	-0.0002	-0.00023	-0.00024	-0.0002	B	6.7
-0.00023	-0.0002	-0.0002	-0.00022	-0.00024	-0.00021	-0.00022	B	2.7
-0.00067	-0.00069	-0.00067	-0.00066	-0.00081	-0.00074	-0.00071	B	5.2
-0.00013	-0.0001	-0.00016	-0.00016	-0.00014	-0.00014	-0.00013	B	10.5
-0.00039	-0.00039	-0.00041	-0.00043	-0.00038	-0.0004	-0.00039	B	5.4
0.000446	0.000427	0.000348	0.00034	0.000294	0.000285	0.000267	B	5.6
1.33E-05	-3.8E-06	-2E-05	-2.1E-05	-4.1E-05	-3.7E-05	-4.4E-05	C	8.6
0.000109	7.26E-05	8.2E-05	8.7E-05	6.78E-05	1.65E-05	-9.6E-06	C	5.6
-8.5E-05	-0.0001	-0.00011	-0.00017	-0.00018	-0.00016	-0.00016	C	3.6
5.66E-05	7.49E-05	2.74E-05	-2.4E-05	-2.1E-06	-9.9E-06	-9.6E-05	C	2.8
0.000171	0.000123	0.000103	8.17E-05	6.53E-05	6.1E-05	3.62E-05	C	7.3
-0.00023	-0.00034	-0.00026	-0.00013	-0.0003	-0.00032	-0.00027	C	8.9
-0.00013	-7.7E-05	-0.00015	-0.00015	-0.00015	-0.00013	-0.00017	C	11.9
X							***Y*** PLS-DA	***Y*** PLS

Figure 6: Data included in a PLS analysis. The ***X*** table/matrix comprises the metabolomic data following pre-processing. Depending on the objective of the analysis, the ***Y*** table/matrix can include continuous (PLS) or categorical (PLS-DA) data to which the ***X*** data are correlated.

OPLS is an improvement on PLS that separates variation in the data into two parts: one that is correlated with the biological factor(s) and one that is unrelated/orthogonal, i.e. OPLS separates ***X*** into variation that is predictive of ***Y*** and variation that is orthogonal to ***Y*** (Brindle *et al*., 2002; Trygg *et al*., 2002, 2003, 2007; Wagner *et al*., 2005). A further development of OPLS is 2-way OPLS (O2PLS). While OPLS only correlates data in ***X*** with ***Y***, O2PLS is able to correlate ***X*** and ***Y*** with each other in both directions (Trygg & Wold, 2003). In addition, individual variables from an O2PLS analysis can be visualized as a bivariate 1D loadings plot facilitating identification of potential metabolites (Cloarec *et al*., 2005). Both, OPLS and O2PLS have recently been preferred to PLS, as separation and correlation of predictive variation to the ***Y*** table has been shown to optimize discriminant analysis, improving overall knowledge extraction (Bylesjö *et al*., 2006; Pinto *et al*., 2012; Wiklund *et al*., 2008). This is because both predictive and orthogonal variation can be examined, which may provide more detailed insight into the factors in-

fluencing the biological system (Kirwan *et al*., 2012). Furthermore, O2PLS can be used to correlate two different data sets with each other, e.g. metabolomic and proteomic datasets in an animal model of prostate cancer (Rantalainen *et al*., 2006). If applied in this way, O2PLS is essentially an unsupervised analysis that is able to correlate variables of different datasets, providing further insight into related structures and pathways in altered metabolic states (Bylesjö *et al*., 2007).

Different extensions of OPLS or O2PLS have been published, including kernel-OPLS, which improves model prediction (Rantalainen *et al*., 2007). O*n*PLS is an extension of O2PLS which determines correlation not only between two, but multiple (*n*) matrices, allowing for integration of any number of datasets for a given study (Löfstedt & Trygg, 2011). Bi-modal O*n*PLS is an extension to O*n*PLS that is not only able to analyze orthogonal variation in variables (columns), but also in samples (rows) (Löfstedt *et al*., 2012a). This bi-modal approach should provide more informed data analysis, of both, the variables associated to the biological question, and of confounding factors associated to particular samples. Finally, O*n*PLS path modeling is a method of linking multiple matrices along a set of paths that flow between data blocks. These paths are assumed to be due to a specific causative mechanism, e.g. changes over time, and are able to extract the minimum number of predictive components that have maximum covariance and correlation (Löfstedt, *et al*., 2012b). Use of these recent extensions of O2PLS is not yet widespread, but highly promising in improving metabolomic data analysis.

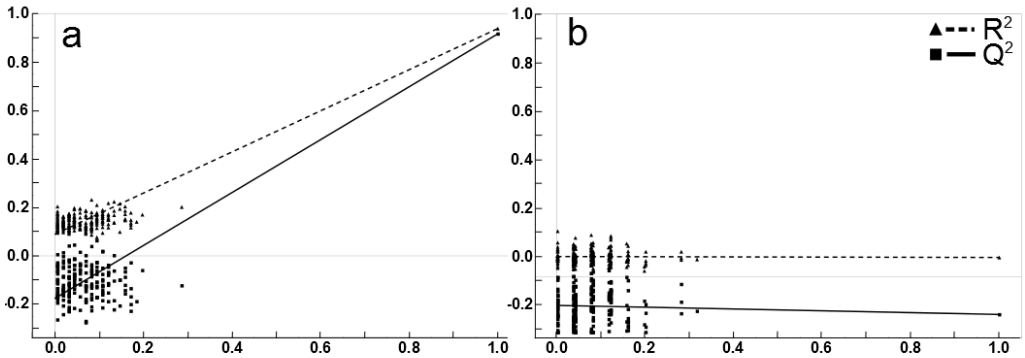

Figure 7: Validation by permutation analysis. a – Example of a valid model (original R^2 and Q^2 plotted on right side of panel), with permutations resulting in models that are less predictive (plotted on left side of panel). The *x*-axis indicates the distance of the permuted model to the original model, and the *y*-axis indicates R^2 and Q^2. b – Example of an invalid model, with permutations resulting in models with similar or improved predictability.

One inherent problem of any supervised MVSA method is that, because they attempt to correlate the experimental data (***X***) with external data (***Y***), they are prone to introducing bias in the analysis. This can happen due to overemphasis of spurious correlations in the data that are only coincidental and not caused by biology. As a result, MVSA models have to be rigorously validated when compared to unsupervised analysis methods. The value of a validated supervised model may be higher than that of a model originating from unsupervised methods because supervised methods are more directed toward the biological question. There are several methods of validation. The gold standard is the use of an independent set of data to test the predictive power of the original (training) set of data, combined with external cross-

validation of the training data set. An established alternative is permutation analysis in which the data in the ***Y*** table are repeatedly permuted at random and the model recalculated with the permuted ***Y*** data (Westerhuis *et al*., 2008). If the model is stable and correlations to ***Y*** are only of biological origin, randomization and permutation of ***Y*** data will reduce the fit and predictability of the model (see Figure 7A). Weak models in which correlations to ***Y*** are due to chance, rather than biology, will produce permuted models that may provide similar or superior prediction than the original model and are thus invalid (see Figure 7B).

Measures of validity in this context are the R^2 value, which measures goodness of fit, and Q^2, which measures model prediction ability. However – similar to cross-validation – permutation analysis becomes less trustworthy the lower the ratio between number of samples (n) and variables (k) is. The turning point may be near a ratio of n/k of $< 0.02 – 0.04$, although this is not applicable for all data sets, and each study has to be evaluated on its own merits (Rubingh *et al*., 2006). In situations where validation via permutation analysis is not easily accessible, cross-validated ANOVA can be used which uses cross-validated *predictive* residuals using two degrees of freedom for each component, and is more reliable than ANOVA which uses *fitted* residuals (Eriksson *et al*., 2008).

3 Current Evidence: Metabolomics in Prostate Cancer

3.1 Prostate Cancer Pathophysiology

Prostate cancer (PCa) is the most common internal cancer in men worldwide and is more prevalent and lethal in Western countries (Siegel *et al*., 2012). Continually evolving methods for early PCa detection have improved outcomes due to earlier treatment and a better prognosis for patients. Current methods of detection (serum prostate-specific antigen (PSA) and/or digital rectal examination) leading to diagnosis (via trans-rectal ultrasound (TRUS) guided biopsy) require improvement due to limited diagnostic sensitivity and specificity. Improved methods will help to avoid morbidity in men for whom a diagnosis of PCa remains elusive due to limitations and problems associated with TRUS-guided biopsy, as is the current situation. Thus, PCa pathogenesis has been extensively studied to facilitate the discovery of new methods for determining the presence of PCa.

The prostate gland sits in the pelvis below the bladder and in front of the rectum. It is a secretory gland that contributes to the seminal fluid component of ejaculate/semen to facilitate sperm motility and egg fertilization *in utero*. The secretory portion of the gland is called the peripheral zone (PZ), and constitutes 70% of the gland volume. The epithelium within the PZ secretes prostatic fluid, which contains proteins, such as prostatic acid phosphatase and prostatic specific antigen (PSA), and metabolites, such as citrate and polyamines (e.g. spermine) (Costello & Franklin, 2009). Furthermore, prostatic cells have been shown to be present in EPS and ejaculate, which makes both biofluids suitable media for molecular analysis (Gardiner *et al*., 1996).

Citrate production, after sequestration, by PZ epithelium results in a higher citrate concentration in EPS when compared with blood plasma (Costello & Franklin, 2009). This process is facilitated by zinc-dependent truncation of the tricarboxylic acid cycle by inhibiting the enzyme m-aconitase, as shown in Figure 8.

ZIP1 is the primary transporter for zinc ions in PZ epithelium, and is expressed by the *ZIP1* gene, which has consequently been described as a tumor suppressor gene in PCa (Costello & Franklin, 2006).

The expression of ZIP1 and other zinc transporters recently has been described as being regulated by the micro-RNA cluster miR-183-96-182 (Mihelich *et al*., 2011). Zinc ions inhibit m-aconitase, which converts citrate to isocitrate, the first step of the citric acid cycle. As a result, the preferential sequestration of zinc ions in the PZ epithelium causes citric acid cycle truncation, producing an increased glucose requirement within the PZ epithelium.

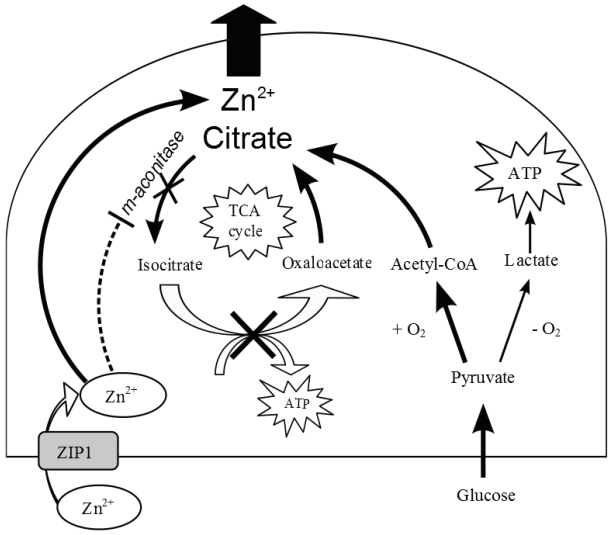

Figure 8: Pertinent physiology of the healthy PZ epithelium. Biochemical reactions are shown by solid/open arrows and regulatory interactions by dashed arrows. When healthy, ZIP1 mediated uptake of zinc inhibits isomerization of citrate to isocitrate by m-aconitase. The result is high intracellular concentrations of zinc ions and citrate, which are secreted to aid in fertilization. Adapted from (Costello & Franklin, 2009; Roberts *et al*., 2011).

The resulting high citrate and Zn^{2+} concentrations in PZ epithelium are reflected in EPS. Citrate is important in seminal ion homeostasis, and is the predominant regulator of calcium ions, which are important in the motility, metabolism and fertilization functions of sperm (Owen & Katz, 2005). Levels of zinc ions are correlated with those of other cations, such as calcium and magnesium, but are considerably higher in concentration. In semen, these cations are largely redistributed in binding to negatively charged seminal vesicle proteins, such as seminogelins, which are vital in regulating sperm function (de Lamirande, 2007). In seminal fluid, zinc ions are bound mostly to metallothionein, with changes in levels of zinc being paralleled by those of this protein which is mostly derived from the prostate itself (Suzuki *et al*., 1994).

3.1.2 The Malignant Prostate

Malignant transformation of cells is the result of irreversible genetic alterations, most commonly due to mutations. Specific to PCa, malignant transformation impairs Zn^{2+} accumulation, removing zinc-mediated inhibition of m-aconitase. The result is completion of the citric acid cycle and increased ATP production via oxidative phosphorylation. This is reflected by low zinc and citrate concentrations present in PZ epithelium and prostatic fluid, which have been investigated as potential biomarkers (Costello &

Franklin, 2009). Further alterations in gene expression impair normal mitochondrial functioning. Coupled with the relatively rapid division and increased basal metabolic rate in cancer cells, increased glycolysis and lactate fermentation in the presence of oxygen occurring in the malignant state increases glucose uptake, as well as proteolysis and subsequent alanine production. Pyruvate is produced in excess of what can be processed by the tricarboxylic acid cycle, and is converted to lactate. This is known as the Warburg effect, and is seen as a marker of advanced disease in prostate and other cancers (Bayley & Devilee, 2012; Warburg, 1956). This process is outlined in Figure 9. Furthermore, increased membraneogenesis accompanying increased cellular proliferation adds to the changes in the metabolite profile with malignant transformation, and requires synthesis of choline and creatine, which have been shown to be elevated in malignant prostate tissues (Noworolski *et al*., 2008).

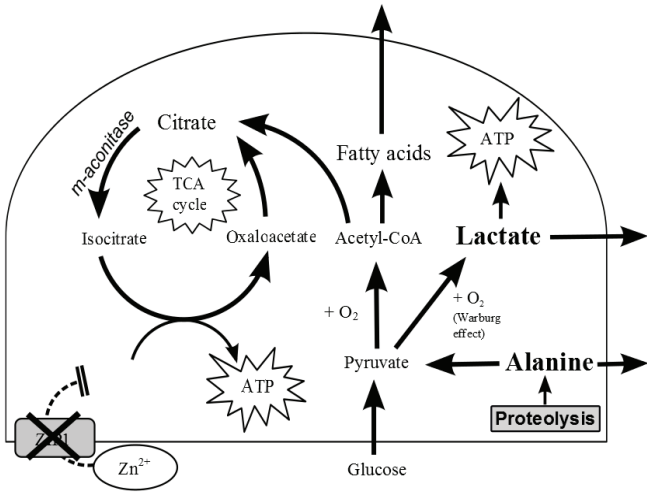

Figure 9: Pathophysiology of the PZ epithelium after malignant transformation. Impaired zinc uptake reduces inhibition of *m-aconitase*, resulting in citrate isomerization and completion of the TCA cycle. Alanine is produced secondary to proteolysis and lactate as a consequence of the Warburg effect. Adapted from (Costello & Franklin, 2009; Israël & Schwartz, 2011; Roberts *et al*., 2011).

3.2 Individual Biomarkers

A single biomarker that is able to confirm the presence of an altered biological process or indicates progression of a disease is a valuable asset in prompting appropriate management for any medical condition to improve the outcome for a particular patient. For instance, extremely high serum levels of the human hormone β-chorionic gonadotropin (β-hCG, a marker normally used in pregnancy) in a male patient with a small testicular mass strongly indicate the presence of choriocarcinoma. While this is an example of an ideal biomarker, such biomarkers do not currently exist for most scenarios in oncology, particularly in PCa (Cole, 2009).

3.2.1 Serum

The most widely used biomarker for PCa is serum human kallikrein 3, also known as prostatic-specific antigen (PSA). PSA is a serine protease that is normally secreted in seminal fluid to catalyze proteolysis

of seminal proteins, such as seminogelin (Lilja, 1985). PSA is elevated in blood in the presence of PCa as well as with other prostatic conditions, such as bacterial prostatitis and benign prostatic hyperplasia (BPH). Despite not being specific for cancer, PSA is clinically valued and widely used (Clarke *et al*., 2010).

The normal range of serum PSA, based on population studies, is defined as <1.0 ng/ml. Serum PSA is not only frequently elevated in other disease states, but can also change in the absence of pathology due to other confounders, such as racial and environmental variables (Henderson *et al*., 1997; Marks *et al*., 2006; Vollmer, 2004). In biomarker research, often a cut-off point is derived from studies to determine the optimal sensitivity (i.e. the ability of the test to accurately predict true negative patients) and specificity (i.e. the ability of the test to predict accurate true positive patients). A vast body of evidence has shown that a safe cut-off value for PSA does not exist (Schröder *et al*., 2000; Thompson *et al*., 2004). Safety in this context refers to a level that is low enough to detect the majority of men with cancer, but not so low as to cause extensive and unnecessary investigation of men without cancer. This finding of the absence of a safe cut-off value for PSA has been used as one of the major arguments against population screening for PCa using PSA (Catalona *et al*., 2012; Moyer, 2012).

Serum PSA is a clinically valued test when used with discrimination, and various adaptations have been discovered and trialed with varying success, though none has been considered superior to total serum PSA itself (Auprich *et al*., 2012; Roobol *et al*., 2009). Examples include free to total PSA levels, PSA velocity and doubling time (time course of an increase by a factor of two), PSA density (serum PSA in relation to the prostate volume determined by TRUS) and, most recently, the prostate health index which incorporates serum PSA, pro-PSA and percentage free PSA (Catalona *et al*., 2011; Hori *et al*., 2012; Lughezzani *et al*., 2012).

Other serum biomarkers in PCa diagnosis vary in type and size, from circulating tumor cells (Danila *et al*., 2011; Doyen *et al*., 2012), to microRNAs (Catto *et al*., 2011; Sevli *et al*., 2010), with small molecules and ions, such as sarcosine and zinc, having yielded inconsistent results as markers (Daragó *et al*., 2011; Li *et al*., 2005; Lucarelli *et al*., 2012; Struys *et al*., 2010).

3.2.2 Urine

Ideally, the perfect marker of PCa is sourced from a non-invasive sample/procedure and indicates both the presence and nature of the disease. Currently, the best urinary marker for PCa is PCA3, formerly known as differential display clone 3 (DD3) (Salagierski & Schalken, 2012). The PCA3 test relies on a patient having had a firm digital rectal examination or prostatic massage just before micturition with the flow of urine flushing dislodged prostatic cells in the prostatic urethra to beyond the external meatus with the void for collection, so there is some licence involved in calling this a urine test. PCA3 is a non-coding RNA which has been shown to be highly expressed in and specific for prostatic tissue (Bussemakers *et al*., 1999; de Kok *et al*., 2002; Landers *et al*., 2005). PCA3 in urine is expressed as a ratio to PSA RNA, and improves detection compared with serum PSA. Use of recently described PCA3 isoforms may further improve results (Clarke *et al*. 2009; Haese *et al*., 2008; Hessels *et al*., 2003). PCA3 also contributes to and has been recommended for clinical decision making for men with previous negative biopsies but in whom clinical suspicion is high (Tombal *et al*., 2012), however its role in PCa detection has yet to be established clinically. Inclusion of the TMPRSS2: ERG fusion gene also has been reported to improve detection of PCa with reference to biopsy (Tomlins *et al*., 2011).

3.2.3 Seminal Fluid

The main concern with tests using EPS is that both firm DRE / prostatic massage and TRUS biopsy target the posterior part of the prostate, and neglect anterior and anterolateral aspects of the gland in which up to 30% of PCa's are sited (Quann *et al*., 2010; Samaratunga *et al*., 2007). In contrast, seminal fluid contains a prostatic component, which is the result of global smooth muscle contraction, and thus theoretically reflects the pathological status of the whole gland. Furthermore, the ability to produce seminal fluid via ejaculation is an indicator of cardiovascular status, as erectile dysfunction is a known event in deteriorating cardiovascular status (Chew *et al*., 2011; Schouten *et al*., 2008). Thus, men who produce seminal fluid are expected to have a more favorable mortality outcome following intervention with curative intent for PCa than men who are impotent since cardiovascular disease is the commonest cause of patient demise in this population. Prostatic tissue and prostatic fluid show similar levels of citrate and Zn^{2+}, further suggesting that prostatic fluid reflects intraprostatic pathophysiological status (Zaichick *et al*., 1996).

3.2.4 Metabolite Changes in Seminal Fluid

Historically, changes in citrate and Zn^{2+} in PCa have been the most pronounced and easily detectable in prostatic and seminal fluid (Cooper & Imfeld, 1959). Metabolite profiling and recent metabolomic analysis of seminal fluid and EPS have discovered alterations in other metabolites, summarized in Table 3. Disturbed zinc homeostasis removes the inhibition of m-aconitase, resulting in citrate oxidation in the citric acid cycle. This causes luminal Zn^{2+} and citrate depletion. Zinc depletion only occurs in PCa, and has been shown to be a stable indicator of PCa status and progression (Zaichick *et al*., 1996).

Although citrate levels are altered in other pathophysiological states, such as BPH and prostatitis (Cooper & Farid, 1964), reduced citrate concentrations in histologically benign prostatic tissue is considered to precede microscopic evidence of PCa (Dittrich *et al*., 2012). In poorly differentiated tumors, these normally abundant metabolites are present in very low concentrations (Kurhanewicz *et al*., 1993). This metabolite relationship in PCa has also been correlated with the Gleason histological scoring system and is more accurate than serum PSA (Kline *et al*., 2006). Such biochemical changes reflect early neoplastic processes that may not be histologically identifiable, a concept familiar in oncology as the "field effect" (Costello & Franklin, 2009). This further supports the role of metabolomics in identifying significant metabolic alterations in pre-malignant tissue.

Other metabolite changes seen in oncology that are not prostate-specific are also present in seminal fluid. Disturbed synthesis and intracellular depletion of polyamines, such as spermine, are reflected in prostatic fluid (Cheng *et al*., 2001; Kline *et al*., 2006; Serkova *et al*., 2008; van der Graaf *et al*., 2000). The prostate contains the highest levels of spermine in the body, and disturbances in ornithine-decarboxylase in polyamine metabolism have been a hypothesized mechanism for spermine depletion (Mohan *et al*., 1999; Simoneau *et al*., 2008; van der Graaf *et al*., 2000). A role of increased reactive oxygen species production by increased expression of spermine oxidase in PCa has linked inflammation with PCa carcinogenesis (Goodwin *et al*., 2008). Levels of myo-inositol, a molecule involved in membrane biosynthesis, have also been shown to be reduced in prostatic fluid (Lynch & Nicholson, 1997; Serkova *et al*., 2008).

Some changes in metabolite levels in prostatic tissue are not reflected in prostatic fluid. Choline is upregulated in PCa tissue, both *in vitro* and *in vivo*, being hypothesized as another metabolite involved in membrane biosynthesis. The use of choline as a marker in prostatic fluid is compromised by the endoge-

nous conversion of phosphocholine (from the seminal vesicles) to choline catalyzed by prostatic acid phosphatase (from the prostate) shortly following ejaculation. This produces a biological artefact in choline concentration. Lactate and alanine are also increased in PCa tissue as part of the Warburg effect. However, spermatozoa utilize fructose from the seminal vesicles and glucose via glycolysis to produce ATP to fuel flagellal movement *in utero* to aid fertilization, resulting in varying levels of lactate and alanine as metabolic by-products. This illustrates a confounding factor between external cellular components and intraprostatic metabolites.

PCa-induced change	Metabolite	Role (normal)	Alteration hypothesis	Reference
Increase	Choline	Membrane phospholipid precursor	Increased membraneogenesis	(DeFeo & Cheng, 2010; Swanson *et al*., 2003, 2006, 2008)
	Lactate	End product of anaerobic glycolysis	Warburg effect	(DeFeo & Cheng, 2010; Swanson *et al*., 2006; Tessem *et al*., 2008)
	Alanine	End product of anaerobic glycolysis	Warburg effect	(DeFeo & Cheng, 2010; Swanson *et al*., 2006; Tessem *et al*., 2008)
	Omega-6 fatty acids	Cell membrane biosynthesis, fatty acid oxidation	Altered gene Expression	(Stenman *et al*., 2009)
	Cholesterol	Membrane biosynthesis, androgen regulated	Increased cell Turnover	(Thysell *et al*., 2010)
	Sarcosine	Glycine metabolism, purine synthesis	Cell invasion	(Sreekumar *et al*., 2009)
	(Choline + creatine) / citrate	Metabolite ratio	Increased ratio	(Kurhanewicz *et al*., 1995; van Asten *et al*., 2008)
	Choline / citrate	Metabolite ratio	Increased ratio	(van Asten *et al*., 2008)
	Choline / creatine	Metabolite ratio	Increased ratio	(van Asten *et al*., 2008)
Decrease	Citrate	Ion homeostasis, pH buffer	m-aconitase activation	(Serkova *et al*., 2008)
	Spermine	Polyamine synthesis	Oxidative stress, enzyme alteration	(Serkova *et al*., 2008; van der Graaf *et al*., 2000)
	Myo-inositol	Membrane biosynthesis		(Serkova *et al*., 2008)
	Citrate / spermine	Metabolite ratio	Decreased ratio	(Lynch & Nicholson, 1997)
	Citrate / creatine	Metabolite ratio	Decreased ratio	(van Asten *et al*., 2008)

Table 3: Summary of metabolite changes in prostate cancer. Changes generic to cancer, such as lactate and alanine, are listed together with changes specific to prostate physiology, such as citrate, sarcosine and spermine. Adapted from (Roberts *et al*., 2011).

3.3 Metabolomics in Prostate Cancer Diagnosis: Finding the Best Combination

Metabolite concentrations and ratios have aided in distinguishing PCa from benign prostates (Swanson *et al*., 2003, 2008; Tessem *et al*., 2008; van Asten *et al*., 2008). Yet, despite promising preliminary results,

there is currently no test available that is accepted as an accurate, stand-alone diagnostic or screening test. With improved data acquisition and processing technology, the concept of using entire metabolic profiles as a large-scale combination of biomarkers has become feasible. Furthermore, metabolite profiles have been shown to be more sensitive as predictors of PCa, and in predicting metastatic potential (Hricak *et al*., 2007) (Cheng *et al*., 2005; Mazaheri *et al*., 2008; Wu *et al*., 2010). This concept has been demonstrated by metabolomic imaging, in which multivoxel MR spectra of intact prostates were analysed with MVSA. This was able to detect highly significant changes between the global metabolite profiles of benign and malignant prostate tissue without the need to identify specific metabolites (Wu *et al*., 2010). Similar relationships were demonstrated using freshly frozen PCa tissue when microarray gene expression data and metabolomic data were combined using PLS, providing further insight into mechanisms of metabolite alterations in PCa (Bertilsson *et al*., 2012). In another study, metabolomic profiling provided an accurate prediction of biochemical recurrence of PCa, that is a rise in serum PSA, following intervention (Maxeiner *et al*., 2010). This illustrates the potential of metabolomics as a suitable method for monitoring PCa behavior following clinical interventions (Roberts *et al*., 2011).

The concept of multiple markers in cancer diagnosis has been examined for some time. Specific to PCa, panels of molecular and protein-based markers have been used to improve serum PSA-based PCa detection (Cao *et al*., 2011; Cuperlovic-Culf *et al*., 2010; Talesa *et al*., 2009). The most widely publicized and promising appear to be the combination of PCA3 and TMPRSS2-ERG fusion transcripts in post-massage urine as previously described. Multiple studies have shown improved sensitivity when combining PCA3 and TMPRSS2-ERG compared with PCA3 or serum PSA alone (Hessels *et al*., 2003; Tomlins *et al*., 2011). Multiple mRNA markers (Gal*N*Ac-T3, PSMA, Hepsin and PCA3) in malignant prostate tissue have been able to provide optimal detection rates (Clarke *et al*., 2010; Landers *et al*., 2005, 2008). Other researchers have attempted to combine single markers of different origins to improve PCa diagnosis. For example, a multiplex model utilizing gene-, protein- and metabolite-based targets for PCa outperformed any single biomarker (Cao *et al*., 2011). However, although these studies are promising in improving PCa diagnosis, many are impractical for use in a clinical setting, mostly due to financial and logistical constraints. Thus, in addition to improved accuracy, a further potential benefit of using metabolomics in PCa diagnosis is a reduction in cost and logistical requirements for each sample, although a high initial capital equipment financial outlay is required. Further limitations of this approach are outlined below.

3.4 Limiting/Confounding Factors

As previously discussed, detecting PCa is difficult. This is due to many confounding factors relating to the pathophysiology of PCa, but also due to concomitant prostatic disease mimicking PCa. In the majority of studies to date, the greatest interference arises from concomitant pathophysiology of the prostate, such as BPH and prostatitis. Serum PSA is known to be elevated in BPH: androgens contributing to BPH development drive PSA synthesis which is mirrored in serum levels. Prostatitis associated with inflammation and prostatic cell lysis, results in increased release of intracellular PSA into the bloodstream, elevating serum PSA. It is for these reasons that serum PSA lacks sensitivity (detecting many false positives) and that there is sustained criticism directed toward serum PSA testing.

Tissue and prostatic fluid levels of citrate were reported to be initially promising in PCa diagnosis compared with serum PSA, but were observed in the past to be depleted in prostatitis (Averna *et al*., 2005; Kavanagh *et al*., 1982; Kline *et al*., 2006). Notwithstanding similar potential dilemmas as those

experienced with serum PSA, diminished levels of both zinc and citrate in these samples may provide improved sensitivity in PCa detection, although conclusive evidence in conjunction with PSA elevations is lacking to date (Daragó *et al*., 2011; Kavanagh *et al*., 1982; Zaichick *et al*., 1996). The relationship of citrate depletion in tissues following radiotherapy or hormonal therapy also relates to biochemical recurrence, defined as a rising serum PSA following intervention with curative intent (Menard *et al*., 2001; Mueller-Lisse *et al*., 2001). Furthermore, reduced specificity of serum PSA when compared with metabolite diagnosis may be due to pre-malignant disturbances in metabolic homeostasis that are not histologically visible (Costello & Franklin, 2009). The underlying issue of biochemical characterisation preceding histopathology creates ongoing uncertainty, as tissue histopathology is the current gold standard for PCa diagnosis.

Intra- and extracellular citrate levels are also known to increase in BPH, thus citrate estimation in biopsy tissue may be unreliable. In these circumstances, prostatic fluid may be a more appropriate sample to use since prostatic fluid is produced mostly in the peripheral zone, which is also where most PCa is located (Costello & Franklin, 2009), whereas BPH develops in the transition (central) zone of the prostate. A small proportion of transition zone tumors may be missed, but in the large majority of cases these are less aggressive and therefore of less significance (Grignon & Sakr, 1994).

4 Future Directions

4.1 Pharmacometabolomics and Theranosis: Towards Personalized Medicine

Pharmacometabolomics seeks to predict the metabolic response to exogenous therapeutic agents prior to or during drug administration, and theranosis is the identification and monitoring of optimal treatments for patients as guided by diagnostic tests (Clayton *et al*., 2006; DeNardo & DeNardo, 2012; Nicholson *et al*., 2012). Both are important aspects, given the recently emerging evidence of inter-individual differences in drug pharmacokinetics, being the ability of an individual to absorb, distribute, metabolize and excrete an administered drug (Suhre *et al*., 2011). The result is reduced therapeutic efficacy, but may also be responsible for toxicity and adverse drug effects. The etiology of these differences in drug metabolism is diverse and is well understood in only a limited number of circumstances (e. g. Cytochrome p450 enzyme family (De Gregori *et al*., 2010)). Gender, age, race and concomitant diseases have been suggested as inherent factors, but have yet to be substantiated (Nebert *et al*., 2003). External factors such as dietary and lifestyle habits, as well as toxin exposure, may also have a large influence on therapeutic efficacy. Furthermore, less obvious but important factors may contribute, such as altered gut flora in various circumstances (Dumas *et al*., 2006; Ley *et al*., 2006; Wang *et al*., 2011). As such, metabolic phenotypes are diverse and complex due to these many influencing factors (Holmes *et al*., 2008). Although genetic profiling across different disorders is important, metabolite profiling promises to better reflect the phenotype of disease states, and advanced analysis between both methods may help to identify genes with significant penetrance.

As has been illustrated, the carcinogenic changes in various cancers will cause common changes to individual metabolic profiles that can be investigated with metabolomics. In contrast, each individual patient will exhibit inherently different metabolic profiles, while also responding differently to therapeutic interventions (Wilson, 2009). Thus, the concept of personalized medicine, where treatments are tailored to an individual's personal metabolic or genetic phenotype, is one that is exciting, and important in

advancing medical treatments (Loscalzo *et al*., 2007). Using metabolite profiles as a representation of metabolic phenotype promises to enable theranosis by providing the most useful information for predicting inter-individual variation that will guide and assess efficacy of treatment outcomes.

Early research focused on genetic predisposition to cancer or alterations in drug-metabolizing enzymes (Yong *et al*., 2006). Recent research has focused on identifying varying metabolic profiles that indicate significantly affected drug metabolism, with links to altered gut flora homeostasis (Clayton *et al*., 2006, 2009). This research related individual background urinary metabolic phenotype to biological and therapeutic outcomes of drug metabolism. Other research has identified alterations in metabolite profiles to illustrate pharmacokinetics and early toxicity, important in preventing adverse outcomes from drug toxicity (Winnike *et al*., 2010). Recently, this approach has been applied to surgery, and allows for personalized pre-, intra- and post-operative care to improve patient outcomes (Kinross *et al*., 2011; Mirnezami *et al*., 2012). The concept of pharmacometabolomics can also be applied to outcome prediction, similar to that used in GWAS, with evidence of serum metabolite levels to be predictive of body mass following chemotherapy for breast cancer (Keun *et al*., 2009).

Theoretically, pharmacometabolomics has advantages over pharmacogenomics in representing the phenotype resulting from multiple genetic effects. However, it is believed that a combination of approaches will provide best prediction and outcomes (Nicholson *et al*., 2011). Given current variations in efficacy, toxicity and adverse outcomes of treatments, developing personalized medicine is imperative to provide better medical care to patients while also reducing health budget costs. Thus, the pharmacometabolomic approach is one that has potential to change the therapeutic landscape not only in oncology, but across all fields of medicine.

4.2 Metabolomics to Elucidate Biological Mechanisms

As outlined, metabolomics has been useful in displaying changes in metabolites in various healthy and pathological states. The analysis of metabolites illustrates the end product of normally functioning or disturbed cellular processes and mechanisms. Thus, analysis of changes in metabolite profiles can lead to insights about the underlying biochemical or biological mechanisms. This has been demonstrated in different areas, including but not limited to, drug toxicity, cancer and plant studies (Bylesjö *et al.*, 2007, 2008; Klenø *et al.*, 2004; Rantalainen *et al.*, 2006).

For example, metabolomics could explain how altered STAT5 signaling as a result of truncated intracellular domains of growth hormone receptor in liver tissue leads to late-onset obesity, as systemic metabolite changes were consistent with globally altered metabolism contributing to obesity (Schirra *et al.*, 2008). A similar approach was used for data obtained from a human prostate cancer xenograft model in mice measured by NMR-based metabonomics and proteomics (two dimensional difference gel electrophoresis). Pathway analysis was used to link altered protein expression to changes in amino acids, which contributed to the metabolic phenotype (Rantalainen *et al.*, 2006).

4.3 Integration with other –omics

As previously outlined, the –omics approach to sample analysis provides data sets that require complex statistical analysis to extract meaningful information. In isolation, each –omics field provides insight into that particular level of cell function, and interactions and influences causing the results are hypothesized based on previous research or logical thinking. Thus, appropriate integration of –omics datasets has become an important step in providing meaningful information in systems biology (Crockford *et al*., 2006;

Bylesjö *et al.*, 2008). This approach was e.g. used in insulin resistant mice using NMR-based metabonomics and genomics (quantitative trait locus mapping), and showed altered gut metabolites that were linked with genomic alterations (Dumas *et al.*, 2007). In both studies, large datasets were used to determine which metabolites were similarly affected by alterations in precursor compounds.

A suggested method that uses O2PLS for integration of large datasets for optimal information recovery is outlined in Figure 10 (Bylesjö *et al.*, 2008), using the example of a study that has data from transcriptomics, proteomics and metabolomics experiments. As O2PLS is able to extract information from two datasets at a time, the method for correlating multiple datasets is naturally a multi-step procedure. In the first step the joint variation between two of the three datasets (e.g. transcript and metabolite data) is extracted by O2PLS. This means using one of the datasets as $\boldsymbol{X}$ matrix in O2PLS and the other one as $\boldsymbol{Y}$ matrix. Note that in this case, O2PLS is effectively run as *unsupervised* analysis, because only two experimental datasets are correlated against each other, without including a set of external metadata. In the second step, the joint variation between transcript and metabolite data obtained in step 1 is then correlated with the third dataset (proteomics) in a second O2PLS, which will yield the variation common to all three datasets. In the final step, the joint variation that is common to all three datasets is deflated from the original datasets in a series of three parallel O2PLS analyses, to produce variation that is specific to each dataset.

It is trivial to extend this scheme to more than three datasets by adding on further O2PLS steps between steps 2 and 3 that each time introduce a further set of experimental data into the analysis. It should also be noted that it is prudent to repeat the first two steps with different orders of combining the three datasets – e.g. transcript and metabolite data first, then including protein data, *versus* metabolite and protein data first, then including transcript data, etc. – in order to rule out that potential slight imperfections in the symmetry behavior of O2PLS might cause secondary effects on the data analysis.

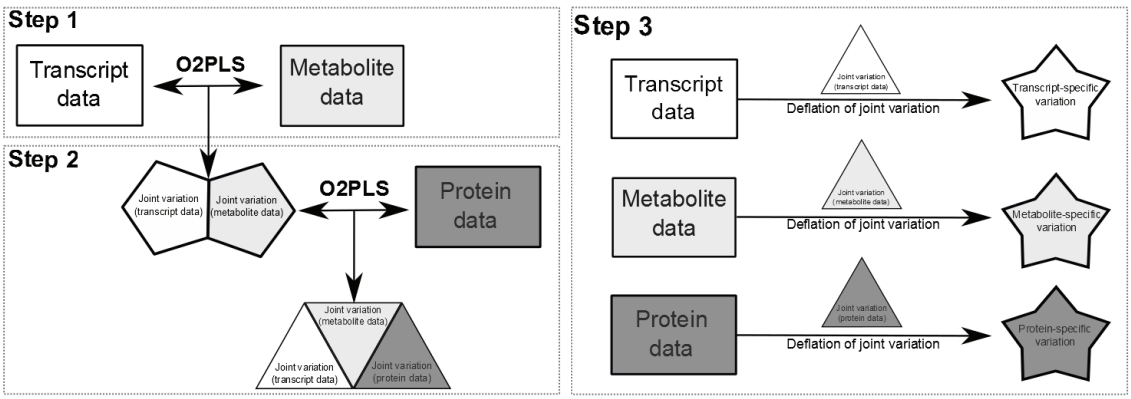

Figure 10: Graphic representation of stepwise data integration of multiple –omics datasets with O2PLS. In the first step, joint variation between two –omics datasets (e.g. transcript and metabolite data) is determined. Using O2PLS in a second step, this joint variation is then correlated with the third –omics dataset (e.g. protein data) to determine variation that is joint to all three sets of data. The third step removes the joint variation between all three individual datasets to produce variation that is specific to each dataset. This dataset-specific variation may be important in helping to address the biological question. Adapted from (Bylesjö *et al.*, 2008).

4.4 Use of Computational Modeling

As is widely highlighted, the current approach to metabolomics including valid statistical analysis, metabolite identification, and biological interpretation is highly time-consuming. As such, the quest to develop computerized methods of metabolite analysis and identification is underway (Aggio *et al*., 2011; Tulpan *et al*., 2011). Following metabolite identification, the next step is to determine the relationship and similarities, if any, of the identified metabolites to metabolic pathways, of which some preliminary programming applications have been released to address this issue (Aggio *et al*., 2010; Bebek & Yang, 2007; Leader *et al*., 2011).

Even more promising is the development of genome-scale computer models of metabolic networks. Extensive work has been completed on bacteria such as *Escherichia coli*, with *in silico* simulation reported to mimic experimental changes (Edwards *et al*., 2001). Application of these reconstructed networks is more complicated in eukaryotes, such as human cells, due to the complex cellular and organismic organisation, including intra- and extra-cellular regulation and interactions. Despite these, an initial model human cell was constructed to provide a general baseline in expression and response to biological variables (Bordbar & Palsson, 2012; Mo *et al*., 2007). Furthermore, a reconstruction of healthy liver cells was combined with whole-body pharmacokinetics to investigate multiple levels in biological organisation and provide mechanistic insights into for various drug-induced scenarios (Krauss *et al*., 2012). Alterations to such models to accurately reflect cancer and other pathophysiological states by incorporating known and emerging evidence will better describe the response of these cells (Jerby & Ruppin, 2012). Depending on the type of model, spatiotemporal processes and interactions within cells that may be undefinable or difficult to quantify are currently difficult to incorporate and apply to an artificial model (Materi & Wishart, 2007). Further development of these reconstructed networks may occur via integration of –omics data sets, and research in this field is continuing (Joyce & Palsson, 2006). Computational modeling by incorporating multiple data sets represents a logical and informative, yet challenging, approach to oncology research to guide pharmaceutical development strategies. The result will be better informed treatment approaches and improved treatment outcomes for these patients.

5 Conclusions

Metabolomics is a novel, modern and robust scientific approach that has shown great advances across many fields in biomedical research. The application of metabolomics to differing fields in medical science, including pathophysiology insight, drug development and *in vivo* imaging make it unique from all other approaches. Further research and collaboration to develop reconstructed networks, via integration of many terabytes of –omics data, is the next frontier in providing valuable insights to advance medical research and treatments in various human disease states.

References

Abate-Shen, C., & Shen, M. M. (2009). Diagnostics: The prostate-cancer metabolome. Nature, 457(7231), 799-800.

Adamski, J. (2012). Genome-wide association studies with metabolomics. Genome Medicine, 4(4), 34.

Aggio, R. B., Ruggiero, K., & Villas-Bôas, S. G. (2010). Pathway Activity Profiling (PAPi): from the metabolite profile to the metabolic pathway activity. Bioinformatics, 26(23), 2969-2976.

Aggio, R. B. M., Villas-Bôas, S. G., & Ruggiero, K. (2011). Metab: an R package for high-throughput analysis of metabolomics data generated by GC-MS. Bioinformatics, 27(16), 2316-2318.

Akira, K., Masu, S., Imachi, M., Mitome, H., Hashimoto, M., & Hashimoto, T. (2008). ^{1}H NMR-based metabonomic analysis of urine from young spontaneously hypertensive rats. Journal of Pharmaceutical and Biomedical Analysis, 46(3), 550-556.

Alum, M. F., Shaw, P. A., Sweatman, B. C., Ubhi, B. K., Haselden, J. N., & Connor, S. C. (2008). 4,4-Dimethyl-4-silapentane-1-ammonium trifluoroacetate (DSA), a promising universal internal standard for NMR-based metabolic profiling studies of biofluids, including blood plasma and serum. Metabolomics, 4(2), 122-127.

Anderson, P. E., Mahle, D. A., Doom, T. E., Reo, N. V., DelRaso, N. J., & Raymer, M. L. (2011). Dynamic adaptive binning: an improved quantification technique for NMR spectroscopic data. Metabolomics, 7(2), 179-190.

Auprich, M., Augustin, H., Budäus, L., Kluth, L., Mannweiler, S., Shariat, S. F., Fisch, M., Graefen, M., Pummer, K., & Chun, F. K. H. (2012). A comparative performance analysis of total prostate-specific antigen, percentage free prostate-specific antigen, prostate-specific antigen velocity and urinary prostate cancer gene 3 in the first, second and third repeat prostate biopsy. BJU International, 109(11), 1627-1635.

Averna, T. A., Kline, E. E., Smith, A. Y., & Sillerud, L. O. (2005). A decrease in ^{1}H nuclear magnetic resonance spectroscopically determined citrate in human seminal fluid accompanies the development of prostate adenocarcinoma. The Journal of Urology, 173(2), 433-438.

Balog, J., Szaniszlo, T., Schaefer, K. C., Denes, J., Lopata, A., Godorhazy, L., Szalay, D., Balogh, L., Sasi-Szabo, L., Toth, M., & Takats, Z. (2010). Identification of biological tissues by rapid evaporative ionization mass spectrometry. Analytical Chemistry, 82(17), 7343-7350.

Basu, S., Kwee, T. C., Surti, S., Akin, E. A., Yoo, D., & Alavi, A. (2011). Fundamentals of PET and PET/CT imaging. Annals of the New York Academy of Sciences, 1228, 1-18.

Bayley, J. P., & Devilee, P. (2012). The Warburg effect in 2012. Current Opinion in Oncology, 24(1), 62-67.

Bebek, G., & Yang, J. (2007). PathFinder: mining signal transduction pathway segments from protein-protein interaction networks. BMC Bioinformatics, 8:335.

Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J. M., Matthews, P. M. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature Neuroscience. 6, 750 – 757.

Bertilsson, H., Tessem, M. B., Flatberg, A., Viset, T., Gribbestad, I., Angelsen, A., & Halgunset, J. (2012). Changes in gene transcription underlying the aberrant citrate and choline metabolism in human prostate cancer samples. Clinical Cancer Research, 18(12), 3261-3269.

Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B. J., Mendes, P., Roessner-Tunali, U., Beale, M. H., Trethewey, R. N., Lange, B. M., Wurtele, E. S., & Sumner, L. W. (2004). Potential of metabolomics as a functional genomics tool. Trends in Plant Science, 9(9), 418-425.

Bloch, F., Hansen, W. W., & Packard, M. (1946). Nuclear Induction. Physical Review, 69(3-4), 127.

Bordbar, A., & Palsson, B. Ø. (2012). Using the reconstructed genome-scale human metabolic network to study physiology and pathology. Journal of Internal Medicine, 271(2), 131-141.

Brindle, J. T., Antti, H., Holmes, E., Tranter, G., Nicholson, J. K., Bethell, H. W., Clarke, S., Schofield, P. M., McKilligin, E., Mosedale, D. E., & Grainger, D. J. (2002). Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using ^{1}H-NMR-based metabonomics. Nature Medicine, 8(12), 1439-1444.

Broadhurst, D. I., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics, 2(4), 171-196.

Brown, M. V., McDunn, J. E., Gunst, P. R., Smith, E. M., Milburn, M. V., Troyer, D. A., & Lawton, K. A. (2012). Cancer detection and biopsy classification using concurrent histopathological and metabolomic analysis of core biopsies. Genome Medicine, 4:33.

Bussemakers, M. J. G., van Bokhoven, A., Verhaegh, G. W., Smit, F. P., Karthaus, H. F. M., Schalken, J. A., Debruyne, F. M. J., Ru, N., & Isaacs, W. B. (1999). DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. Cancer Research, 59(23), 5975-5979.

Bylesjö, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2006). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. Journal of Chemometrics, 20(8-10), 341-351.

Bylesjö, M., Eriksson, D., Kusano, M., Moritz, T., & Trygg, J. (2007). Data integration in plant biology: the O2PLS method for combined modeling of transcript and metabolite data. The Plant Journal, 52(6), 1181-1191.

Bylesjö, M., Nilsson, R., Srivastava, V., Grönlund, A., Johansson, A. I., Jansson, S., Karlsson, J., Moritz, T., Wingsle, G., & Trygg, J. (2008). Integrated Analysis of Transcript, Protein and Metabolite Data To Study Lignin Biosynthesis in Hybrid Aspen. Journal of Proteome Research, 8(1), 199-210.

Cao, D. L., Ye, D. W., Zhang, H. L., Zhu, Y., Wang, Y. X., & Yao, X. D. (2011). A Multiplex Model of Combining Gene-Based, Protein-Based, and Metabolite-Based With Positive and Negative Markers in Urine for the Early Diagnosis of Prostate Cancer. Prostate, 71(7), 700-710.

Carroll, A. J., Badger, M. R., & Harvey Millar, A. H. (2010). The MetabolomeExpress Project: enabling web-based processing, analysis and transparent dissemination of GC/MS metabolomics datasets. BMC Bioinformatics, 11:376.

Casado-Vela, J., Cebrián, A., Gómez del Pulgar, M. T., & Lacal, J. C. (2011). Approaches for the study of cancer: towards the integration of genomics, proteomics and metabolomics. Clinical & Translational Oncology, 13(9), 617-628.

Catalona, W. J., Partin, A. W., Sanda, M. G., Wei, J. T., Klee, G. G., Bangma, C. H., Slawin, K. M., Marks, L. S., Loeb, S., Broyles, D. L., Shin, S. S., Cruz, A. B., Chan, D. W., Sokoll, L. J., Roberts, W. L., van Schaik, R. H. N., & Mizrahi, I. A. (2011). A Multicenter Study of [-2]Pro-Prostate Specific Antigen Combined With Prostate Specific Antigen and Free Prostate Specific Antigen for Prostate Cancer Detection in the 2.0 to 10.0 ng/ml Prostate Specific Antigen Range. The Journal of Urology, 185(5), 1650-1655.

Catalona, W. J., D'Amico, A. V., Fitzgibbons, W. F., Kosoko-Lasaki, O., Leslie, S. W., Lynch, H. T., Moul, J. W., Rendell, M. S., & Walsh, P. C. (2012). What the U.S. Preventive Services Task Force Missed in Its Prostate Cancer Screening Recommendation. Annals of Internal Medicine, 157(2), 137-138 & W-27.

Catto, J. W. F., Alcaraz, A., Bjartell, A. S., De Vere White, R., Evans, C. P., Fussel, S., Hamdy, F. C., Kallioniemi, O., Mengual, L., Schlomm, T., & Visakorpi, T. (2011). MicroRNA in prostate, bladder, and kidney cancer: a systematic review. European Urology, 59(5), 671-681.

Center, M. M., Jemal, A., Lortet-Tieulent, J., Ward, E., Ferlay, J., Brawley, O., & Bray, F. (2012). International Variation in Prostate Cancer Incidence and Mortality Rates. European Urology, 61(6), 1079-1092.

Chen, R., Mias, G. I., Li-Pook-Than, J., Jiang, L., Lam, H. Y. K., Chen, R., Miriami, E., Karczewski, K. J., Hariharan, M., Dewey, F. E., Cheng, Y., Clark, M. J., Im, H., Habegger, L., Balasubramanian, S., O'Huallachain, M., Dudley, J. T., Hillenmeyer, S., Haraksingh, R., Sharon, D., Euskirchen, G., Lacroute, P., Bettinger, K., Boyle, A. P., Kasowski, M., Grubert, F., Seki, S., Garcia, M., Whirl-Carrillo, M., Gallardo, M., Blasco, M. A., Greenberg, P. L., Snyder, P., Klein, T. E., Altman, R. B., Butte, A. J., Ashley, E. A., Gerstein, M., Nadeau, K. C., Tang, H., & Snyder, M. (2012). Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell, 148(6), 1293-1307.

Cheng, L. L., Wu, C. L., Smith, M. R., & Gonzalez, R. G. (2001). Non-destructive quantitation of spermine in human prostate tissue samples using HRMAS ^{1}H NMR spectroscopy at 9.4 T. FEBS Letters, 494(1-2), 112-116.

Cheng, L. L., Burns, M. A., Taylor, J. L., He, W., Halpern, E. F., McDougal, W. S., & Wu, C. L. (2005). Metabolic characterization of human prostate cancer with tissue magnetic resonance spectroscopy. Cancer Research, 65(8), 3030-3034.

Chew, K. K., Gibson, N., Sanfilippo, F., Stuckey, B., & Bremner, A. (2011). Cardiovascular mortality in men with erectile dysfunction: increased risk but not inevitable. The Journal of Sexual Medicine, 8(6), 1761-1771.

Clarke, R. A., Schirra, H. J., Catto, J. W., Lavin, M. F., & Gardiner, R. A. (2010). Markers for Detection of Prostate Cancer. Cancers, 2(2), 1125-1154.

Clarke, R.A., Zhao, Z., Guo, A.Y., Roper A.K., Teng, L., Fang, Z.M., Samaratunga, H.M., Lavin M.F., Gardiner R.A. (2009). New genomic structure for prostate cancer specific gene PCA3 within BMCC1: implications for prostate cancer detection and progression. PLoS ONE, 4 (3), e4995.

Clayton, T. A., Lindon, J. C., Cloarec, O., Antti, H., Charuel, C., Hanton, G., Provost, J. P., Le Net, J. L., Baker, D., Walley, R. J., Everett, J. R., & Nicholson, J. K. (2006). Pharmaco-metabonomic phenotyping and personalized drug treatment. Nature, 440(7087), 1073-1077.

Clayton, T. A., Baker, D., Lindon, J. C., Everett, J. R., & Nicholson, J. K. (2009). Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. Proceedings of the National Academy of Sciences, 106(24), 14728-14733.

Cloarec, O., Dumas, M. E., Trygg, J., Craig, A., Barton, R. H., Lindon, J. C., Nicholson, J. K., & Holmes, E. (2005). Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. Analytical Chemistry, 77(2), 517-526.

Cole, L.A. (2009). Human chorionic gonadotropin and associated molecules. Expert Review of Molecular Diagnostics, 9 (1), 51-73.

Cooper, J. F., & Farid, I. (1964). The role of citric acid in the physiology of the prostate. Lactate/citrate ratios in benign and malignant prostatic homogenates as an index of prostatic malignancy. The Journal of Urology, 92, 533-536.

Cooper, J. F., & Imfeld, H. (1959). The role of citric acid in the physiology of the prostate: a preliminary report. The Journal of Urology, 81(1), 157-164.

Costello, L. C., & Franklin, R. B. (2005). 'Why do tumor cells glycolyse?': From glycolysis through citrate to lipogenesis. Molecular and Cellular Biochemistry, 280(1-2), 1-8.

Costello, L. C., & Franklin, R. B. (2006). The clinical relevance of the metabolism of prostate cancer; zinc and tumor suppression: connecting the dots. Molecular Cancer, 5:17.

Costello, L. C., & Franklin, R. B. (2009). Prostatic fluid electrolyte composition for the screening of prostate cancer: a potential solution to a major problem. Prostate Cancer and Prostatic Diseases, 12(1), 17-24.

Coy, S. L., Cheema, A. K., Tyburski, J. B., Laiakis, E. C., Collins, S. P., & Fornace, A., Jr. (2011). Radiation metabolomics and its potential in biodosimetry. International Journal of Radiation Biology, 87(8), 802-823.

Craig, A., Cloarec, O., Holmes, E., Nicholson, J. K., & Lindon, J. C. (2006). Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets. Analytical Chemistry, 78(7), 2262-2267.

Crockford, D. J., Holmes, E., Lindon, J. C., Plumb, R. S., Zirah, S., Bruce, S. J., Rainville, P., Stumpf, C. L., & Nicholson, J. K. (2006). Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies. Analytical Chemistry, 78(2), 363-371.

Cuperlovic-Culf, M., Belacel, N., Davey, M., & Ouellette, R. J. (2010). Multi-gene biomarker panel for reference free prostate cancer diagnosis: determination and independent validation. Biomarkers, 15(8), 693-706.

Danila, D. C., Fleisher, M., & Scher, H. I. (2011). Circulating tumor cells as biomarkers in prostate cancer. Clinical Cancer Research, 17(12), 3903-3912.

Daragó, A., Sapota, A., Matych, J., Nasiadek, M., Skrzypińska-Gawrysiak, M., & Kilanowicz, A. (2011). The correlation between zinc and insulin-like growth factor 1 (IGF-1), its binding protein (IGFBP-3) and prostate-specific antigen (PSA) in prostate cancer. Clinical Chemistry and Laboratory Medicine, 49(10), 1699-1705.

De Gregori, M., Allegri, M., De, G. S., Garbin, G., Tinelli, C., Regazzi, M., Govoni, S., & Ranzani, G. N. (2010). How and why to screen for CYP2D6 interindividual variability in patients under pharmacological treatments. Current Drug Metabolism, 11(3), 276-282.

de Kok, J. B., Verhaegh, G. W., Roelofs, R. W., Hessels, D., Kiemeney, L. A., Aalders, T. W., Swinkels, D. W., & Schalken, J. A. (2002). $DD3^{PCA3}$, a very sensitive and specific marker to detect prostate tumors. Cancer Research, 62(9), 2695-2698.

de Lamirande, E. (2007). Semenogelin, the main protein of the human semen coagulum, regulates sperm function. Seminars in Thrombosis and Hemostasis, 33(1), 60-68.

Deepinder, F., Chowdary, H. T., & Agarwal, A. (2007). Role of metabolomic analysis of biomarkers in the management of male infertility. Expert Review of Molecular Diagnostics, 7(4), 351-358.

DeFeo, E. M., & Cheng, L. L. (2010). Characterizing human cancer metabolomics with ex vivo ^{1}H HRMAS MRS. Technology in Cancer Research & Treatment, 9(4), 381-391.

Delongchamps, N. B., Rouanne, M., Flam, T., Beuvon, F., Liberatore, M., Zerbib, M., & Cornud, F. (2011). Multiparametric magnetic resonance imaging for the detection and localization of prostate cancer: combination of T2-weighted, dynamic contrast-enhanced and diffusion-weighted imaging. BJU International, 107(9), 1411-1418.

DeNardo, G. L., & DeNardo, S. J. (2012). Concepts, consequences, and implications of theranosis. Seminars in Nuclear Medicine, 42(3), 147-150.

Denkert, C., Budczies, J., Kind, T., Weichert, W., Tablack, P., Sehouli, J., Niesporek, S., Könsgen, D., Dietel, M., & Fiehn, O. (2006). Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. Cancer Research, 66(22), 10795-10804.

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. Mass Spectrometry Reviews, 26(1), 51-78.

Dieterle, F., Ross, A., Schlotterbeck, G., & Senn, H. (2006). Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in ^{1}H NMR Metabonomics. Analytical Chemistry, 78(13), 4281-4290.

Dittrich, R., Kurth, J., Decelle, E. A., DeFeo, E. M., Taupitz, M., Wu, S., Wu, C. L., McDougal, W. S., & Cheng, L. L. (2012). Assessing prostate cancer growth with citrate measured by intact tissue proton magnetic resonance spectroscopy. Prostate Cancer and Prostatic Diseases, 15(3), 278-282.

Doyen, J., Alix-Panabières, C., Hofman, P., Parks, S. K., Chamorey, E., Naman, H., & Hannoun-Lèvi, J. M. (2012). Circulating tumor cells in prostate cancer: a potential surrogate marker of survival. Critical Reviews in Oncology/Hematology, 81(3), 241-256.

Duarte, I. F., & Gil, A. M. (2012). Metabolic signatures of cancer unveiled by NMR spectroscopy of human biofluids. Progress in Nuclear Magnetic Resonance Spectroscopy, 62, 51-74.

Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R., & Palsson, B. Ø. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proceedings of the National Academy of Sciences, 104(6), 1777-1782.

Dumas, M.E., Barton, R. H., Toye, A., Cloarec, O., Blancher, C., Rothwell, A., Fearnside, J., Tatoud, R., Blanc, V., Lindon, J. C., Mitchell, S. C., Holmes, E., McCarthy, M. I., Scott, J., Gauguier, D., & Nicholson, J. K. (2006). Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. Proceedings of the National Academy of Sciences, 103(33), 12511-12516.

Dumas, M. E., Wilder, S. P., Bihoreau, M. T., Barton, R. H., Fearnside, J. F., Argoud, K., D'Amato, L., Wallis, R. H., Blancher, C., Keun, H. C., Baunsgaard, D., Scott, J., Sidelmann, U. G., Nicholson, J. K., & Gauguier, D. (2007). Direct quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. Nature Genetics, 39(5), 666-672.

Dunn, W. B., Bailey, N. J. C., & Johnson, H. E. (2005). Measuring the metabolome: current analytical technologies. Analyst, 130(5), 606-625.

Dunn, W. B., & Ellis, D. I. (2005). Metabolomics: Current analytical platforms and methodologies. Trends in Analytical Chemistry, 24(4), 285-294.

Edwards, J. S., Ibarra, R. U., & Palsson, B. Ø. (2001). In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nature Biotechnology, 19(2), 125-130.

Eknoyan, G. (1999). Santorio Sanctorius (1561-1636) - founding father of metabolic balance studies. American Journal of Nephrology, 19(2), 226-233.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C. & Wold, S. (2006). Multi- and megavariate data analysis. Second Edition. Umetrics AB. Umeå, Sweden.

Eriksson, L., Trygg, J., & Wold, S. (2008). CV-ANOVA for significance testing of PLS and OPLS® models. Journal of Chemometrics, 22(11-12), 594-600.

Evangelista, L., Guttilla, A., Zattoni, F., Muzzio, P. C., & Zattoni, F. (2012). Utility of Choline Positron Emission Tomography/Computed Tomography for Lymph Node Involvement Identification in Intermediate- to High-risk Prostate Cancer: A Systematic Literature Review and Meta-analysis. European Urology, 63(6), 1040-1048.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., & Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International Journal of Cancer, 127(12), 2893-2917.

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. Plant Molecular Biology, 48(1-2), 155-171.

Field, D., Sansone, S.A., Collis, A., Booth, T., Dukes, P., Gregurick, S. K., Kennedy, K., Kolar, P., Kolker, E., Maxon, M., Millard, S., Mugabushaka, A.M., Perrin, N., Remacle, J. E., Remington, K., Rocca-Serra, P., Taylor, C. F., Thorley, M., Tiwari, B., & Wilbanks, J. (2009). 'Omics Data Sharing. Science, 326(5950), 234-236.

Fitch, K. D. (2008). Androgenic-anabolic steroids and the Olympic Games. Asian Journal of Andrology, 10(3), 384-390.

Fletcher, J. W., Djulbegovic, B., Soares, H. P., Siegel, B. A., Lowe, V. J., Lyman, G. H., Coleman, R. E., Wahl, R., Paschold, J. C., Avrill, N., Einhorn, L. H., Suh, W. W., Samson, D., Delbeke, D., Gorman, M., & Shields, A. F. (2008). Recommendations on the use of F-18-FDG PET in oncology. Journal of Nuclear Medicine, 49(3), 480-508.

Freeman, R. (1995). A short history of NMR. Chemistry of Heterocyclic Compounds, 31(9), 1004-1005.

Ganti, S., & Weiss, R. H. (2011). Urine metabolomics for kidney cancer detection and biomarker discovery. Urologic Oncology, 29(5), 551-557.

Gardiner, R. A., Samaratunga, M. L., Gwynne, R. A., Clague, A., Seymour, G. J., & Lavin, M. F. (1996). Abnormal prostatic cells in ejaculates from men with prostatic cancer – a preliminary report. British Journal of Urology, 78(3), 414-418.

Gates, S. C., & Sweeley, C. C. (1978). Quantitative metabolic profiling based on gas chromatography. Clinical Chemistry, 24(10), 1663-1673.

Geladi, P., & Esbensen, K. (1990). The start and early history of chemometrics: Selected interviews. Part 1. Journal of Chemometrics, 4(5), 337-354.

Giskeødegård, G. F., Bloemberg, T. G., Postma, G., Sitter, B., Tessem, M. B., Gribbestad, I. S., Bathen, T. F., & Buydens, L. M. C. (2010). Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods. Analytica Chimica Acta, 683(1), 1-11.

Goodacre, R., Broadhurst, D., Smilde, A. K., Kristal, B. S., Baker, J. D., Beger, R., Bessant, C., Connor, S., Capuani, G., Craig, A., Ebbels, T., Kell, D. B., Manetti, C., Newton, J., Paternostro, G., Somorjai, R., Sjöström, M., Trygg, J., & Wulfert, F. (2007). Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics, 3(3), 231-241.

Goodwin, A. C., Jadallah, S., Toubaji, A., Lecksell, K., Hicks, J. L., Kowalski, J., Bova, G. S., De Marzo, A. M., Netto, G. J., & Casero, R. A., Jr. (2008). Increased spermine oxidase expression in human prostate cancer and prostatic intraepithelial neoplasia tissues. Prostate, 68(7), 766-772.

Griffin, J. L., & Shockcor, J. P. (2004). Metabolic profiles of cancer cells. Nature Reviews Cancer, 4(7), 551-561.

Griffiths, J. (2008). A Brief History of Mass Spectrometry. Analytical Chemistry, 80(15), 5678-5683.

Grignon, D.J., & Sakr, W.A., (1994) Zonal origin of prostatic adenocarcinoma: are there biologic differences between transition zone and peripheral zone adenocarcinomas of the prostate gland? Journal of Cellular Biochemistry, Supplement 19, 267-269.

Gu, H. W., Pan, Z. Z., Xi, B. W., Asiago, V., Musselman, B., & Raftery, D. (2011). Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: Application to the detection of breast cancer. Analytica Chimica Acta, 686(1-2), 57-63.

Haese, A., de la Taille, A., van Poppel H., Marberger, M., Stenzl, A., Mulders, P. F. A., Huland, H., Abbou, C. C., Remzi, M., Tinzl, M., Feyerabend, S., Stillebroer, A. B., van Gils, M. P. M. Q., & Schalken, J. A. (2008). Clinical utility of the PCA3 urine assay in European men scheduled for repeat biopsy. European Urology, 54(5), 1081-1088.

Hamamah, S., Seguin, F., Barthelemy, C., Akoka, S., Le Pape, A., Lansac, J., & Royere, D. (1993). ^{1}H nuclear magnetic resonance studies of seminal plasma from fertile and infertile men. Journal of Reproduction and Fertility, 97(1), 51-55.

Hamamah, S., Seguin, F., Bujan, L., Barthelemy, C., Mieusset, R., & Lansac, J. (1998). Quantification by magnetic resonance spectroscopy of metabolites in seminal plasma able to differentiate different forms of azoospermia. Human Reproduction, 13(1), 132-135.

Han, X., Rozen, S., Boyle, S. H., Hellegers, C., Cheng, H., Burke, J. R., Welsh-Bohmer, K. A., Doraiswamy, P. M., & Kaddurah-Daouk, R. (2011). Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. PLoS ONE, 6(7), e21643.

Henderson, R. J., Eastham, J. A., Culkin, D. J., Kattan, M. W., Whatley, T., Mata, J., Venable, D., & Sartor, O. (1997). Prostate-Specific Antigen (PSA) and PSA Density: Racial Differences in Men Without Prostate Cancer. Journal of the National Cancer Institute, 89(2), 134-138.

Hessels, D., Klein Gunnewiek, J. M. T., van Oort, I., Karthaus, H. F. M., van Leenders, G. J. L., van Balken, B., Kiemeney, L. A., Witjes, J. A., & Schalken, J. A. (2003). DD3^{PCA3}-based molecular urine analysis for the diagnosis of prostate cancer. European Urology, 44(1), 8-16.

Hoeks, C. M. A., Barentsz, J. O., Hambrock, T., Yakar, D., Somford, D. M., Heijmink, S. W. T. P. J., Scheenen, T. W. J., Vos, P. C., Huisman, H., van Oort, I. M., Witjes, J. A., Heerschap, A., & Fütterer, J. J. (2011). Prostate Cancer: Multiparametric MR Imaging for Detection, Localization, and Staging. Radiology, 261(1), 46-66.

Hoh, C. K., Schiepers, C., Seltzer, M. A., Gambhir, S. S., Silverman, D. H. S., Czernin, J., Maddahi, J., & Phelps, M. E. (1997). PET in oncology: Will it replace the other modalities? Seminars in Nuclear Medicine, 27(2), 94-106.

Holmes, E., Wilson, I. D., & Nicholson, J. K. (2008). Metabolic phenotyping in health and disease. Cell, 134(5), 714-717.

Hori, S., Blanchet, J. S., & McLoughlin, J. (2012). From prostate-specific antigen (PSA) to precursor PSA (proPSA) isoforms: a review of the emerging role of proPSAs in the detection and management of early prostate cancer. BJU International, 112(6), 717-728.

Hoult, D. I., Busby, S. J. W., Gadian, D. G., Radda, G. K., Richards, R. E., & Seeley, P. J. (1974). Observation of tissue metabolites using ^{31}P nuclear magnetic resonance. Nature, 252(5481), 285-287.

Hricak, H., Choyke, P. L., Eberhardt, S. C., Leibel, S. A., & Scardino, P. T. (2007). Imaging prostate cancer: a multidisciplinary perspective. Radiology, 243(1), 28-53.

Hyndman, M. E., Mullins, J. K., & Bivalacqua, T. J. (2011). Metabolomics and bladder cancer. Urologic Oncology, 29(5), 558-561.

Israël, M., & Schwartz, L. (2011). The metabolic advantage of tumor cells. Molecular Cancer, 10:70.

Issaq, H. J., Van, Q. N., Waybright, T. J., Muschik, G. M., & Veenstra, T. D. (2009). Analytical and statistical approaches to metabolomics research. Journal of Separation Science, 32(13), 2183-2199.

Jentzmik, F., Stephan, C., Miller, K., Schrader, M., Erbersdobler, A., Kristiansen, G., Lein, M., & Jung, K. (2010). Sarcosine in urine after digital rectal examination fails as a marker in prostate cancer detection and identification of aggressive tumors. European Urology, 58(1), 12-18.

Jerby, L., & Ruppin, E. (2012). Predicting Drug Targets and Biomarkers of Cancer via Genome-Scale Metabolic Modeling. Clinical Cancer Research, 18(20), 5572-5584.

Jones, T., & Price, P. (2012). Development and experimental medicine applications of PET in oncology: a historical perspective. The Lancet Oncology, 13(3), e116-125.

Joyce, A. R., & Palsson, B. Ø. (2006). The model organism as a system: integrating 'omics' data sets. Nature Reviews Molecular Cell Biology, 7(3), 198-210.

Kavanagh, J. P., Darby, C., & Costello, C. B. (1982). The response of seven prostatic fluid components to prostatic disease. International Journal of Andrology, 5(5), 487-496.

Keifer, P. A. (2007). High-Resolution MAS for Liquids and Semisolids. Encyclopedia of Magnetic Resonance: John Wiley & Sons, Ltd.

Kelloff, G. J., Hoffman, J. M., Johnson, B., Scher, H. I., Siegel, B. A., Cheng, E. Y., Cheson, B. D., O'Shaughnessy, J., Guyton, K. Z., Mankoff, D. A., Shankar, L., Larson, S. M., Sigman, C. C., Schilsky, R. L., & Sullivan, D. C. (2005). Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development. Clinical Cancer Research, 11(8), 2785-2808.

Keun, H. C., Sidhu, J., Pchejetski, D., Lewis, J. S., Marconell, H., Patterson, M., Bloom, S. R., Amber, V., Coombes, R. C., & Stebbing, J. (2009). Serum molecular signatures of weight change during early breast cancer chemotherapy. Clinical Cancer Research 15(21), 6716-6723.

Kinross, J. M., Holmes, E., Darzi, A. W., & Nicholson, J. K. (2011). Metabolic phenotyping for monitoring surgical patients. The Lancet, 377(9780), 1817-1819.

Kirwan, G. M., Johansson, E., Kleemann, R., Verheij, E. R., Wheelock, Å. M., Goto, S., Trygg, J., & Wheelock, C. E. (2012). Building Multivariate Systems Biology Models. Analytical Chemistry, 84(16), 7064-7071.

Klenø, T. G., Kiehr, B., Baunsgaard, D., & Sidelmann, U. G. (2004). Combination of 'omics' data to investigate the mechanism(s) of hydrazine-induced hepatotoxicity in Rats and to identify potential biomarkers. Biomarkers, 9(2), 116-138.

Kline, E. E., Treat, E. G., Averna, T. A., Davis, M. S., Smith, A. Y., & Sillerud, L. O. (2006). Citrate concentrations in human seminal fluid and expressed prostatic fluid determined via ^{1}H nuclear magnetic resonance spectroscopy outperform prostate specific antigen in prostate cancer detection. The Journal of Urology, 176(5), 2274-2279.

Koh, D.M., & Collins, D. J. (2007). Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. American Journal of Roentgenology, 188(6), 1622-1635.

Kohl, S. M., Klein, M. S., Hochrein, J., Oefner, P. J., Spang, R., & Gronwald, W. (2012). State-of-the art data normalization methods improve NMR-based metabolomic analysis. Metabolomics, 8(1), S146-S160.

Krauss, M., Schaller, S., Borchers, S., Findeisen, R., Lippert, J., & Kuepfer, L. (2012). Integrating Cellular Metabolism into a Multiscale Whole-Body Model. PLoS Computational Biology, 8(10), e1002750.

Kurhanewicz, J., Dahiya, R., Macdonald, J. M., Chang, L. H., James, T. L., & Narayan, P. (1993). Citrate alterations in primary and metastatic human prostatic adenocarcinomas: ^{1}H magnetic resonance spectroscopy and biochemical study. Magnetic Resonance in Medicine, 29(2), 149-157.

Kurhanewicz, J., Vigneron, D. B., Nelson, S. J., Hricak, H., MacDonald, J. M., Konety, B., & Narayan, P. (1995). Citrate as an in vivo marker to discriminate prostate cancer from benign prostatic hyperplasia and normal prostate peripheral zone: detection via localized proton spectroscopy. Urology, 45(3), 459-466.

Kurhanewicz, J., Swanson, M. G., Nelson, S. J., & Vigneron, D. B. (2002). Combined magnetic resonance imaging and spectroscopic imaging approach to molecular imaging of prostate cancer. Journal of Magnetic Resonance Imaging, 16(4), 451-463.

Kwock, L., Smith, J. K., Castillo, M., Ewend, M. G., Collichio, F., Morris, D. E., Bouldin, T. W., & Cush, S. (2006). Clinical role of proton magnetic resonance spectroscopy in oncology: brain, breast, and prostate cancer. The Lancet Oncology, 7(10), 859-868.

Landers, K. A., Burger, M. J., Tebay, M. A., Purdie, D. M., Scells, B., Samaratunga, H., Lavin, M. F., & Gardiner, R. A. (2005). Use of multiple biomarkers for a molecular diagnosis of prostate cancer. International Journal of Cancer, 114(6), 950-956.

Landers, K. A., Samaratunga, H., Teng, L., Buck, M., Burger, M. J., Scells, B., Lavin, M. F., & Gardiner, R. A. (2008). Identification of claudin-4 as a marker highly overexpressed in both primary and metastatic prostate cancer. British Journal of Cancer, 99(3), 491-501.

Leader, D. P., Burgess, K., Creek, D., & Barrett, M. P. (2011). Pathos: A web facility that uses metabolic maps to display experimental changes in metabolites identified by mass spectrometry. Rapid Communications in Mass Spectrometry, 25(22), 3422-3426.

Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). Microbial ecology: human gut microbes associated with obesity. Nature, 444(7122), 1022-1023.

Li, J., Wijffels, G., Yu, Y., Nielsen, L. K., Niemeyer, D. O., Fisher, A. D., Ferguson, D. M., & Schirra, H. J. (2010). Altered Fatty Acid Metabolism in Long Duration Road Transport: An NMR-based Metabonomics Study in Sheep. Journal of Proteome Research, 10(3), 1073-1087.

Li, X. M., Zhang, L., Li, J., Li, Y., Wang, H. L., Ji, G. Y., Kuwahara, M., & Zhao, X. J. (2005). Measurement of serum zinc improves prostate cancer detection efficiency in patients with PSA levels between 4 ng/mL and 10 ng/mL. Asian Journal of Andrology, 7(3), 323-328.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. British Medical Journal, 339, b2700.

Lilja, H. (1985). A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. The Journal of Clinical Investigation, 76(5), 1899-1903.

Lim, K. S., & Tan, C. H. (2012). Diffusion-weighted MRI of adult male pelvic cancers. Clinical Radiology, 67(9), 899-908.

Lin, D., Hollander, Z., Meredith, A., Stadnick, E., Sasaki, M., Cohen Freue, G., Qasimi, P., Mui, A., Ng, R. T., Balshaw, R., Wilson-McManus, J. E., Wishart, D., Hau, D., Keown, P. A., McMaster, R., & McManus, B. M. (2011). Molecular signatures of end-stage heart failure. Journal of Cardiac Failure, 17(10), 867-874.

Lodi, A., & Ronen, S. M. (2011). Magnetic resonance spectroscopy detectable metabolomic fingerprint of response to antineoplastic treatment. PLoS ONE, 6(10), e26155.

Löfstedt, T., & Trygg, J. (2011). OnPLS-a novel multiblock method for the modelling of predictive and orthogonal variation. Journal of Chemometrics, 25(8), 441-455.

Löfstedt, T., Eriksson, L., Wormbs, G., & Trygg, J. (2012a). Bi-modal OnPLS. Journal of Chemometrics, 26(6), 236-245.

Löfstedt, T., Hanafi, M., Mazerolles, G., & Trygg, J. (2012b). OnPLS path modelling. Chemometrics and Intelligent Laboratory Systems, 118, 139-149.

Loscalzo, J., Kohane, I., & Barabasi, A. L. (2007). Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. Molecular Systems Biology, 3:124.

Lu, X., Zhao, X., Bai, C., Zhao, C., Lu, G., & Xu, G. (2008). LC–MS-based metabonomics analysis. Journal of Chromatography B, 866(1–2), 64-76.

Lucarelli, G., Fanelli, M., Larocca, A. M. V., Germinario, C. A., Rutigliano, M., Vavallo, A., Selvaggi, F. P., Bettocchi, C., Battaglia, M., & Ditonno, P. (2012). Serum sarcosine increases the accuracy of prostate cancer detection in patients with total serum PSA less than 4.0 ng/ml. The Prostate, 72(15), 1611-1621.

Lughezzani, G., Lazzeri, M., Larcher, A., Lista, G., Scattoni, V., Cestari, A., Buffi, N. M., Bini, V., & Guazzoni, G. (2012). Development and internal validation of a Prostate Health Index based nomogram for predicting prostate cancer at extended biopsy. The Journal of Urology, 188(4), 1144-1150.

Lynch, M. J., Masters, J., Pryor, J. P., Lindon, J. C., Spraul, M., Foxall, P. J., & Nicholson, J. K. (1994). Ultra high field NMR spectroscopic studies on human seminal fluid, seminal vesicle and prostatic secretions. Journal of Pharmaceutical and Biomedical Analysis, 12(1), 5-19.

Lynch, M. J., & Nicholson, J. K. (1997). Proton MRS of human prostatic fluid: correlations between citrate, spermine, and myo-inositol levels and changes with disease. Prostate, 30(4), 248-255.

Ma, Y., Zhang, P., Wang, F., Liu, W., Yang, J., & Qin, H. (2012a). An integrated proteomics and metabolomics approach for defining oncofetal biomarkers in the colorectal cancer. Annals of Surgery, 255(4), 720-730.

Ma, Y., Zhang, P., Yang, Y., Wang, F., & Qin, H. (2012b). Metabolomics in the fields of oncology: a review of recent research. Molecular Biology Reports, 39(7), 7505-7511.

MacIntyre, D. A., Jiménez, B., Jantus Lewintre, E., Reinoso Martín, C., Schäfer, H., García Ballesteros, C., Ramon Mayans, J., Spraul, M., García-Conde, J., & Pineda-Lucena, A. (2010). Serum metabolome analysis by ^{1}H-NMR reveals differences between chronic lymphocytic leukemia molecular subgroups. Leukemia, 24(4), 788-797.

MacKinnon, N., Ge, W., Khan, A. P., Somashekar, B. S., Tripathi, P., Siddiqui, J., Wei, J. T., Chinnaiyan, A. M., Rajendiran, T. M., & Ramamoorthy, A. (2012). Variable Reference Alignment: An Improved Peak Alignment Protocol for NMR Spectral Data with Large Intersample Variation. Analytical Chemistry, 84(12), 5372-5379.

Marks, L. S., Andriole, G. L., Fitzpatrick, J. M., Schulman, C. C., & Roehrborn, C. G. (2006). The interpretation of serum prostate specific antigen in men receiving 5α-reductase inhibitors: a review and clinical recommendations. The Journal of Urology, 176(3), 868-874.

Martínez-Bisbal, M. C., Martí-Bonmatí, L., Piquer, J., Revert, A., Ferrer, P., Llácer, J. L., Piotto, M., Assemat, O., & Celda, B. (2004). ^{1}H and ^{13}C HR-MAS spectroscopy of intact biopsy samples *ex vivo* and *in vivo* ^{1}H MRS study of human high grade gliomas. NMR in Biomedicine, 17(4), 191-205.

Materi, W., & Wishart, D. S. (2007). Computational systems biology in drug discovery and development: methods and applications. Drug Discovery Today, 12(7-8), 295-303.

Maxeiner, A., Adkins, C. B., Zhang, Y., Taupitz, M., Halpern, E. F., McDougal, W. S., Wu, C. L., & Cheng, L. L. (2010). Retrospective analysis of prostate cancer recurrence potential with tissue metabolomic profiles. Prostate, 70(7), 710-717.

Mazaheri, Y., Shukla-Dave, A., Hricak, H., Fine, S. W., Zhang, J., Inurrigarro, G., Moskowitz, C. S., Ishill, N. M., Reuter, V. E., Touijer, K., Zakian, K. L., & Koutcher, J. A. (2008). Prostate cancer: identification with combined diffusion-weighted MR imaging and 3D ^{1}H MR spectroscopic imaging – correlation with pathologic findings. Radiology, 246(2), 480-488.

Melamud, E., Vastag, L., & Rabinowitz, J. D. (2010). Metabolomic Analysis and Visualization Engine for LC−MS Data. Analytical Chemistry, 82(23), 9818-9826.

Menard, C., Smith, I. C., Somorjai, R. L., Leboldus, L., Patel, R., Littman, C., Robertson, S. J., & Bezabeh, T. (2001). Magnetic resonance spectroscopy of the malignant prostate gland after radiotherapy: a histopathologic study of diagnostic validity. International Journal of Radiation Oncology *Biology* Physics, 50(2), 317-323.

Mihelich, B. L., Khramtsova, E. A., Arva, N., Vaishnav, A., Johnson, D. N., Giangreco, A. A., Martens-Uzunova, E., Bagasra, O., Kajdacsy-Balla, A., & Nonn, L. (2011). miR-183-96-182 Cluster Is Overexpressed in Prostate Tissue and Regulates Zinc Homeostasis in Prostate Cells. The Journal of Biological Chemistry, 286(52), 44503-44511.

Millis, K. K., Maas, W. E., Cory, D. G., & Singer, S. (1997). Gradient, high-resolution, magic-angle spinning nuclear magnetic resonance spectroscopy of human adipocyte tissue. Magnetic Resonance in Medicine, 38(3), 399-403.

Mirnezami, R., Kinross, J. M., Vorkas, P. A., Goldin, R., Holmes, E., Nicholson, J., & Darzi, A. (2012). Implementation of molecular phenotyping approaches in the personalized surgical patient journey. Annals of Surgery, 255(5), 881-889.

Mishur, R. J., & Rea, S. L. (2012). Applications of mass spectrometry to metabolomics and metabonomics: Detection of biomarkers of aging and of age-related diseases. Mass Spectrometry Reviews, 31(1), 70-95.

Mo, M. L., Jamshidi, N., & Palsson, B. Ø. (2007). A genome-scale, constraint-based approach to systems biology of human metabolism. Molecular BioSystems, 3(9), 598-603.

Mohan, R. R., Challa, A., Gupta, S., Bostwick, D. G., Ahmad, N., Agarwal, R., Marengo, S. R., Amini, S. B., Paras, F., MacLennan, G. T., Resnick, M. I., & Mukhtar, H. (1999). Overexpression of Ornithine Decarboxylase in Prostate Cancer and Prostatic Fluid in Humans. Clinical Cancer Research, 5(1), 143-147.

Moyer, V. A. (2012). Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Annals of Internal Medicine, 157(2), 120-134.

Mueller-Lisse, U. G., Swanson, M. G., Vigneron, D. B., Hricak, H., Bessette, A., Males, R. G., Wood, P. J., Noworolski, S., Nelson, S. J., Barken, I., Carroll, P. R., & Kurhanewicz, J. (2001). Time-dependent effects of hormone-deprivation therapy on prostate metabolism as detected by combined magnetic resonance imaging and 3D magnetic resonance spectroscopic imaging. Magnetic Resonance in Medicine, 46(1), 49-57.

Nebert, D. W., Jorge-Nebert, L., & Vesell, E. S. (2003). Pharmacogenomics and "individualized drug therapy": high expectations and disappointing achievements. American Journal of Pharmacogenomics, 3(6), 361-370.

Neuweger, H., Albaum, S. P., Dondrup, M., Persicke, M., Watt, T., Niehaus, K., Stoye, J., & Goesmann, A. (2008). MeltDB: a software platform for the analysis and integration of metabolomics experiment data. Bioinformatics, 24(23), 2726-2732.

Ng, D. J. Y., Pasikanti, K. K., & Chan, E. C. Y. (2011). Trend analysis of metabonomics and systematic review of metabonomics-derived cancer marker metabolites. Metabolomics, 7(2), 155-178.

Nicholson, J. K., Lindon, J. C., & Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica, 29(11), 1181-1189.

Nicholson, J. K., & Lindon, J. C. (2008). Systems biology: Metabonomics. Nature, 455(7216), 1054-1056.

Nicholson, J. K., Wilson, I. D., & Lindon, J. C. (2011). Pharmacometabonomics as an effector for personalized medicine. Pharmacogenomics, 12(1), 103-111.

Nicholson, J. K., Everett, J. R., & Lindon, J. C. (2012). Longitudinal pharmacometabonomics for predicting patient responses to therapy: drug metabolism, toxicity and efficacy. Expert Opinion on Drug Metabolism & Toxicology, 8(2), 135-139.

Noble, W. S. (2009). How does multiple testing correction work? Nature Biotechnology, 27(12), 1135-1137.

Noworolski, S. M., Vigneron, D. B., Chen, A. P., & Kurhanewicz, J. (2008). Dynamic contrast-enhanced MRI and MR diffusion imaging to distinguish between glandular and stromal prostatic tissues. Magnetic Resonance Imaging, 26(8), 1071-1080.

Oakman, C., Tenori, L., Claudino, W. M., Cappadona, S., Nepi, S., Battaglia, A., Bernini, P., Zafarana, E., Saccenti, E., Fornier, M., Morris, P. G., Biganzoli, L., Luchinat, C., Bertini, I., & Di Leo, A. (2011). Identification of a serum-detectable metabolomic fingerprint potentially correlated with the presence of micrometastatic disease in early breast cancer patients at varying risks of disease relapse by traditional prognostic methods. Annals of Oncology, 22(6), 1295-1301.

Oberbach, A., Blüher, M., Wirth, H., Till, H., Kovacs, P., Kullnick, Y., Schlichting, N., Tomm, J. M., Rolle-Kampczyk, U., Murugaiyan, J., Binder, H., Dietrich, A., & von Bergen, M. (2011). Combined Proteomic and Metabolomic Profiling of Serum Reveals Association of the Complement System with Obesity and Identifies Novel Markers of Body Fat Mass Changes. Journal of Proteome Research, 10(10), 4769-4788.

Owen, D. H., & Katz, D. F. (2005). A review of the physical and chemical properties of human semen and the formulation of a semen simulant. Journal of Andrology, 26(4), 459-469.

Padhani, A. R. (2002). Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. Journal of Magnetic Resonance Imaging, 16(4), 407-422.

Pan, Z. Z., Gu, H. W., Talaty, N., Chen, H. W., Shanaiah, N., Hainline, B. E., Cooks, R. G., & Raftery, D. (2007). Principal component analysis of urine metabolites detected by NMR and DESI-MS in patients with inborn errors of metabolism. Analytical and Bioanalytical Chemistry, 387(2), 539-549.

Pasikanti, K. K., Ho, P. C., & Chan, E. C. Y. (2008). Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. Journal of Chromatography B, 871(2), 202-211.

Pinder, U. U. B. (1506). Epiphanie Medicorum. Speculum videndi urinas hominum. Clavis aperiendi portas pulsuum. Berillus discernendi causas & differentias febrium. Sodalitas Celtica: Nuremberg.

Pinto, R. C., Trygg, J., & Gottfries, J. (2012). Advantages of orthogonal inspection in chemometrics. Journal of Chemometrics, 26(6), 231-235.

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics, 11: 395.

Purcell, E. M., Torrey, H. C., & Pound, R. V. (1946). Resonance Absorption by Nuclear Magnetic Moments in a Solid. Physical Review, 69(1-2), 37-38.

Quann, P., Jarrard, D. F., & Huang, W. (2010). Current prostate biopsy protocols cannot reliably identify patients for focal therapy: correlation of low-risk prostate cancer on biopsy with radical prostatectomy findings. International Journal of Clinical and Experimental Pathology, 3(4), 401-407.

Rantalainen, M., Cloarec, O., Beckonert, O., Wilson, I. D., Jackson, D., Tonge, R., Rowlinson, R., Rayner, S., Nickson, J., Wilkinson, R. W., Mills, J. D., Trygg, J., Nicholson, J. K., & Holmes, E. (2006). Statistically Integrated Metabonomic–Proteomic Studies on a Human Prostate Cancer Xenograft Model in Mice. Journal of Proteome Research, 5(10), 2642-2655.

Rantalainen, M., Bylesjö, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2007). Kernel-based orthogonal projections to latent structures (K-OPLS). Journal of Chemometrics, 21(7-9), 376-385.

Reginato, A. J., & Kurnik, B. (1989). Calcium oxalate and other crystals associated with kidney diseases and arthritis. Seminars in Arthritis and Rheumatism, 18(3), 198-224.

Reske, S. N., Blumstein, N. M., Neumaier, B., Gottfried, H. W., Finsterbusch, F., Kocot, D., Möller, P., Glatting, G., & Perner, S. (2006). Imaging prostate cancer with ^{11}C-choline PET/CT. Journal of Nuclear Medicine, 47(8), 1249-1254.

Rhee, E. P., & Gerszten, R. E. (2012). Metabolomics and cardiovascular biomarker discovery. Clinical Chemistry, 58(1), 139-147.

Roberts, M. J., Schirra, H. J., Lavin, M. F., & Gardiner, R. A. (2011). Metabolomics: a novel approach to early and noninvasive prostate cancer detection. Korean Journal of Urology, 52(2), 79-89.

Rohren, E. M., Turkington, T. G., & Coleman, R. E. (2004). Clinical applications of PET in oncology. Radiology, 231(2), 305-332.

Roobol, M. J., Schröder, F. H., Crawford, E. D., Freedland, S. J., Sartor, A. O., Fleshner, N., & Andriole, G. L. (2009). A Framework for the Identification of Men at Increased Risk for Prostate Cancer. The Journal of Urology, 182(5), 2112-2122.

Ross, A., Schlotterbeck, G., Dieterle, F., & Senn, H. (2007). Chapter 3 - NMR Spectroscopy Techniques for Application to Metabonomics. In John C. Lindon, Jeremy K. Nicholson & Elaine Holmes (Eds.), The Handbook of Metabonomics and Metabolomics (pp. 55-112). Amsterdam: Elsevier Science B.V.

Rubingh, C. M., Bijlsma, S., Derks, E. P. P. A., Bobeldijk, I., Verheij, E. R., Kochhar, S., & Smilde, A. K. (2006). Assessing the performance of statistical validation tools for megavariate metabolomics data. Metabolomics, 2(2), 53-61.

Salagierski, M., & Schalken, J. A. (2012). Molecular diagnosis of prostate cancer: *PCA3* and *TMPRSS2:ERG* gene fusion. The Journal of Urology, 187(3), 795-801.

Salek, R. M., Maguire, M. L., Bentley, E., Rubtsov, D. V., Hough, T., Cheeseman, M., Nunez, D., Sweatman, B. C., Haselden, J. N., Cox, R. D., Connor, S. C., & Griffin, J. L. (2007). A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. Physiological Genomics, 29(2), 99-108.

Samaratunga, H., Yaxley, J., Kerr, K., McClymont, K., & Duffy, D. (2007). Significance of minute focus of adenocarcinoma on prostate needle biopsy. Urology, 70(2), 299-302.

Sansone, S. A., Rocca-Serra, P., Field, D., Maguire, E., Taylor, C., Hofmann, O., Fang, H., Neumann, S., Tong, W., Amaral-Zettler, L., Begley, K., Booth, T., Bougueleret, L., Burns, G., Chapman, B., Clark, T., Coleman, L. A., Copeland, J., Das, S., de Daruvar, A., de Matos, P., Dix, I., Edmunds, S., Evelo, C. T., Forster, M. J., Gaudet, P., Gilbert, J., Goble, C., Griffin, J. L., Jacob, D., Kleinjans, J., Harland, L., Haug, K., Hermjakob, H., Ho Sui, S. J., Laederach, A., Liang, S., Marshall, S., McGrath, A., Merrill, E., Reilly, D., Roux, M., Shamu, C. E., Shang, C. A., Steinbeck, C., Trefethen, A., Williams-Jones, B., Wolstencroft, K., Xenarios, I., & Hide, W. (2012). Toward interoperable bioscience data. Nature Genetics, 44(2), 121-126.

Savorani, F., Tomasi, G., & Engelsen, S. B. (2010). icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. Journal of Magnetic Resonance, 202(2), 190-202.

Schirra, H. J., Anderson, C. G., Wilson, W. J., Kerr, L., Craik, D. J., Waters, M. J., & Lichanska, A. M. (2008). Altered metabolism of growth hormone receptor mutant mice: a combined NMR metabonomics and microarray study. PLoS ONE, 3(7), e2764.

Schouten, B. W., Bohnen, A. M., Bosch, J. L. H. R., Bernsen, R. M. D., Deckers, J. W., Dohle, G. R., & Thomas, S. (2008). Erectile dysfunction prospectively associated with cardiovascular disease in the Dutch general population: results from the Krimpen Study. International Journal of Impotence Research, 20(1), 92-99.

Schröder, F. H., van der Cruijsen-Koeter, I., de Koning, H. J., Vis, A. N., Hoedemaeker, R. F., & Kranse, R. (2000). Prostate cancer detection at low prostate specific antigen. Journal of Urology, 163(3), 806-811.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. British Medical Journal, 340:c332.

Serkova, N. J., Gamito, E. J., Jones, R. H., O'Donnell, C., Brown, J. L., Green, S., Sullivan, H., Hedlund, T., & Crawford, E. D. (2008). The metabolites citrate, myo-inositol, and spermine are potential age-independent markers of prostate cancer in human expressed prostatic secretions. The Prostate, 68(6), 620-628.

Serkova, N. J., Standiford, T. J., & Stringer, K. A. (2011). The emerging field of quantitative blood metabolomics for biomarker discovery in critical illnesses. American Journal of Respiratory and Critical Care Medicine, 184(6), 647-655.

Sevli, S., Uzumcu, A., Solak, M., Ittmann, M., & Ozen, M. (2010). The function of microRNAs, small but potent molecules, in human prostate cancer. Prostate Cancer and Prostatic Diseases, 13(3), 208-217.

Shepherd, L. V. T., Fraser, P., & Stewart, D. (2011). Metabolomics: a second-generation platform for crop and food analysis. Bioanalysis, 3(10), 1143-1159.

Shulaev, V. (2006). Metabolomics technology and bioinformatics. Briefings in Bioinformatics, 7(2), 128-139.

Siegel, R., Naishadham, D., & Jemal, A. (2012). Cancer statistics, 2012. CA: A Cancer Journal for Clinicians, 62(1), 10-29.

Simoneau, A. R., Gerner, E. W., Nagle, R., Ziogas, A., Fujikawa-Brooks, S., Yerushalmi, H., Ahlering, T. E., Lieberman, R., McLaren, C. E., Anton-Culver, H., & Meyskens, F. L., Jr. (2008). The effect of difluoromethylornithine on decreasing prostate size and polyamines in men: results of a year-long phase IIb randomized placebo-controlled chemoprevention trial. Cancer Epidemiology, Biomarkers & Prevention, 17(2), 292-299.

Simoni, R. D., Hill, R. L., & Vaughan, M. (2002). Analytical Biochemistry: the Work of Otto Knuf Olof Folin on Blood Analysis. Journal of Biological Chemistry, 277(20), e9.

Sitter, B., Sonnewald, U., Spraul, M., Fjösne, H. E., & Gribbestad, I. S. (2002). High-resolution magic angle spinning MRS of breast cancer tissue. NMR in Biomedicine, 15(5), 327-337.

Spratlin, J. L., Serkova, N. J., & Eckhardt, S. G. (2009). Clinical applications of metabolomics in oncology: a review. Clinical Cancer Research, 15(2), 431-440.

Sreekumar, A., Poisson, L. M., Rajendiran, T. M., Khan, A. P., Cao, Q., Yu, J., Laxman, B., Mehra, R., Lonigro, R. J., Li, Y., Nyati, M. K., Ahsan, A., Kalyana-Sundaram, S., Han, B., Cao, X., Byun, J., Omenn, G. S., Ghosh, D., Pennathur, S., Alexander, D. C., Berger, A., Shuster, J. R., Wei, J. T., Varambally, S., Beecher, C., & Chinnaiyan, A. M. (2009). Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. Nature, 457(7231), 910-914.

Staab, J. M., O'Connell, T. M., & Gomez, S. M. (2010). Enhancing metabolomic data analysis with Progressive Consensus Alignment of NMR Spectra (PCANS). BMC Bioinformatics, 11:123.

Stenman, K., Hauksson, J. B., Gröbner, G., Stattin, P., Bergh, A., & Riklund, K. (2009). Detection of polyunsaturated omega-6 fatty acid in human malignant prostate tissue by 1D and 2D high-resolution magic angle spinning NMR spectroscopy. Magnetic Resonance Materials in Physics, Biology and Medicine, 22(6), 327-331.

Struys, E. A., Heijboer, A. C., van Moorselaar, J., Jakobs, C., & Blankenstein, M. A. (2010). Serum sarcosine is not a marker for prostate cancer. Annals of Clinical Biochemistry, 47(3):282.

Suhre, K., Shin, S. Y., Petersen, A. K., Mohney, R. P., Meredith, D., Wägele, B., Altmaier, E., Deloukas, P., Erdmann, J., Grundberg, E., Hammond, C. J., de Angelis, M. H., Kastenmuller, G., Köttgen, A., Kronenberg, F., Mangino, M., Meisinger, C., Meitinger, T., Mewes, H. W., Milburn, M. V., Prehn, C., Raffler, J., Ried, J. S., Römisch-Margl, W., Samani, N. J., Small, K. S., Wichmann, H. E., Zhai, G., Illig, T., Spector, T. D., Adamski, J., Soranzo, N. & Gieger, C., (2011). Human metabolic individuality in biomedical and pharmaceutical research. Nature, 477(7362), 54-60.

Suzuki, T., Suzuki, K., Nakajima, K., Otaki, N., & Yamanaka, H. (1994). Metallothionein in human seminal plasma. International Journal of Urology, 1(4), 345-348.

Swanson, M. G., Vigneron, D. B., Tabatabai, Z. L., Males, R. G., Schmitt, L., Carroll, P. R., James, J. K., Hurd, R. E., & Kurhanewicz, J. (2003). Proton HR-MAS spectroscopy and quantitative pathologic analysis of MRI/3D-MRSI-targeted postsurgical prostate tissues. Magnetic Resonance in Medicine, 50(5), 944-954.

Swanson, M. G., Zektzer, A. S., Tabatabai, Z. L., Simko, J., Jarso, S., Keshari, K. R., Schmitt, L., Carroll, P. R., Shinohara, K., Vigneron, D. B., & Kurhanewicz, J. (2006). Quantitative analysis of prostate metabolites using ^{1}H HR-MAS spectroscopy. Magnetic Resonance in Medicine, 55(6), 1257-1264.

Swanson, M. G., Keshari, K. R., Tabatabai, Z. L., Simko, J. P., Shinohara, K., Carroll, P. R., Zektzer, A. S., & Kurhanewicz, J. (2008). Quantification of choline- and ethanolamine-containing metabolites in human prostate tissues using ^{1}H HR-MAS total correlation spectroscopy. Magnetic Resonance in Medicine, 60(1), 33-40.

Talesa, V. N., Antognelli, C., Del Buono, C., Stracci, F., Serva, M. R., Cottini, E., & Mearini, E. (2009). Diagnostic potential in prostate cancer of a panel of urinary molecular tumor markers. Cancer Biomarkers, 5(6), 241-251.

Tessem, M. B., Swanson, M. G., Keshari, K. R., Albers, M. J., Joun, D., Tabatabai, Z. L., Simko, J. P., Shinohara, K., Nelson, S. J., Vigneron, D. B., Gribbestad, I. S., & Kurhanewicz, J. (2008). Evaluation of lactate and alanine as metabolic biomarkers of prostate cancer using ^{1}H HR-MAS spectroscopy of biopsy tissues. Magnetic Resonance in Medicine, 60(3), 510-516.

Theodoridis, G., Gika, H. G., & Wilson, I. D. (2008). LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. Trends in Analytical Chemistry, 27(3), 251-260.

Thompson, I. M., Pauler, D. K., Goodman, P. J., Tangen, C. M., Lucia, M. S., Parnes, H. L., Minasian, L. M., Ford, L. G., Lippman, S. M., Crawford, E. D., Crowley, J. J., & Coltman, C. A., Jr. (2004). Prevalence of Prostate Cancer among Men with a Prostate-Specific Antigen Level ≤4.0 ng per Milliliter. New England Journal of Medicine, 350(22), 2239-2246.

Thomson, J. J. (1912). Rays of positive electricity. Philosophical Magazine Series 6, 24(140), 209-253.

Thysell, E., Surowiec, I., Hörnberg, E., Crnalic, S., Widmark, A., Johansson, A. I., Stattin, P., Bergh, A., Moritz, T., Antti, H., & Wikström, P. (2010). Metabolomic Characterization of Human Prostate Cancer Bone Metastases Reveals Increased Levels of Cholesterol. PLoS ONE, 5(12), e14175.

Tombal, B., Ameye, F., de la Taille, A., de Reijke, T., Gontero, P., Haese, A., Kil, P., Perrin, P., Remzi, M., Schröder, J., Speakman, M., Volpe, A., Meesen, B., & Stoevelaar, H. (2012). Biopsy and treatment decisions in the initial management of prostate cancer and the role of PCA3; a systematic analysis of expert opinion. World Journal of Urology, 30(2), 251-256.

Tomlins, S. A., Aubin, S. M. J., Siddiqui, J., Lonigro, R. J., Sefton-Miller, L., Miick, S., Williamsen, S., Hodge, P., Meinke, J., Blase, A., Penabella, Y., Day, J. R., Varambally, R., Han, B., Wood, D., Wang, L., Sanda, M. G., Rubin, M. A., Rhodes, D. R., Hollenbeck, B., Sakamoto, K., Silberstein, J. L., Fradet, Y., Amberson, J. B., Meyers, S., Palanisamy, N., Rittenhouse, H., Wei, J. T., Groskopf, J., & Chinnaiyan, A. M. (2011). Urine *TMPRSS2:ERG* fusion transcript stratifies prostate cancer risk in men with elevated serum PSA. Science Translational Medicine, 3(94), 94ra72.

Trygg, J., & Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). Journal of Chemometrics, 16(3), 119-128.

Trygg, J., & Wold, S. (2003). O2-PLS, a two-block (X–Y) latent variable regression (LVR) method with an integral OSC filter. Journal of Chemometrics, 17(1), 53-64.

Trygg, J., Holmes, E., & Lundstedt, T. (2007). Chemometrics in Metabonomics. Journal of Proteome Research, 6(2), 469-479.

Trygg, J., & Lundstedt, T. (2007). Chapter 6 - Chemometrics Techniques for Metabonomics. John C. Lindon, Jeremy K. Nicholson & Elaine Holmes (Eds.), The Handbook of Metabonomics and Metabolomics (pp. 171-199). Amsterdam: Elsevier Science B.V.

Tulpan, D., Léger, S., Belliveau, L., Culf, A., & Čuperlović-Culf, M. (2011). MetaboHunter: an automatic approach for identification of metabolites from ^{1}H-NMR spectra of complex mixtures. BMC Bioinformatics, 12:400.

van Asten, J. J. A., Cuijpers, V., Hulsbergen-van de Kaa, C., Soede-Huijbregts, C., Witjes, J. A., Verhofstad, A., & Heerschap, A. (2008). High resolution magic angle spinning NMR spectroscopy for metabolic assessment of cancer presence and Gleason score in human prostate needle biopsies. Magnetic Resonance Materials in Physics, Biology and Medicine, 21(6), 435-442.

van der Graaf, M., Schipper, R. G., Oosterhof, G. O., Schalken, J. A., Verhofstad, A. A., & Heerschap, A. (2000). Proton MR spectroscopy of prostatic tissue focused on the detection of spermine, a possible biomarker of malignant behavior in prostate cancer. Magnetic Resonance Materials in Physics, Biology and Medicine, 10(3), 153-159.

van der Greef, J., & Smilde, A. K. (2005). Symbiosis of chemometrics and metabolomics: past, present, and future. Journal of Chemometrics, 19(5-7), 376-386.

Vargas, H. A., Wassberg, C., Akin, O., & Hricak, H. (2012). MR Imaging of Treated Prostate Cancer. Radiology, 262(1), 26-42.

Vollmer, R. T. (2004). Race and the linkage between serum prostate-specific antigen and prostate cancer: a study of American veterans. American Journal of Clinical Pathology, 122(3), 338-344.

Wagner, S., Scholz, K., Donegan, M., Burton, L., Wingate, J., & Völkel, W. (2005). Metabonomics and Biomarker Discovery: LC–MS Metabolic Profiling and Constant Neutral Loss Scanning Combined with Multivariate Data Analysis for Mercapturic Acid Analysis. Analytical Chemistry, 78(4), 1296-1305.

Wang, Z., Klipfell, E., Bennett, B. J., Koeth, R., Levison, B. S., DuGar, B., Feldstein, A. E., Britt, E. B., Fu, X., Chung, Y. M., Wu, Y., Schauer, P., Smith, J. D., Allayee, H., Tang, W. H. W., DiDonato, J. A., Lusis, A. J., & Hazen, S. L. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature, 472(7341), 57-63.

Warburg, O. (1956). On the origin of cancer cells. Science, 123(3191), 309-314.

Weckwerth, W. (2011). Unpredictability of metabolism–the key role of metabolomics science in combination with next-generation genome sequencing. Analytical and Bioanalytical Chemistry, 400(7), 1967-1978.

Wei, J., Xie, G., Zhou, Z., Shi, P., Qiu, Y., Zheng, X., Chen, T., Su, M., Zhao, A., & Jia, W. (2011). Salivary metabolite signatures of oral cancer and leukoplakia. International Journal of Cancer, 129(9), 2207-2217.

Weiss, R. H., & Kim, K. (2012). Metabolomics in the study of kidney diseases. Nature Reviews Nephrology, 8(1), 22-33.

Westerhuis, J. A., Hoefsloot, H. C. J., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J. J., van Duijnhoven, J. P. M., & van Dorsten, F. A. (2008). Assessment of PLSDA cross validation. Metabolomics, 4(1), 81-89.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfries, J., Moritz, T., & Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. Analytical Chemistry, 80(1), 115-122.

Wilson, I. D., Plumb, R., Granger, J., Major, H., Williams, R., & Lenz, E. M. (2005). HPLC-MS-based methods for the study of metabonomics. Journal of Chromatography B, 817(1), 67-76.

Wilson, I. D. (2009). Drugs, bugs, and personalized medicine: Pharmacometabonomics enters the ring. Proceedings of the National Academy of Sciences, 106(34), 14187-14188.

Winnike, J. H., Li, Z., Wright, F. A., Macdonald, J. M., O'Connell, T. M., & Watkins, P. B. (2010). Use of pharmaco-metabonomics for early prediction of acetaminophen-induced hepatotoxicity in humans. Clinical Pharmacology and Therapeutics, 88(1), 45-51.

Wishart, D. S. (2008). Quantitative metabolomics using NMR. Trends in Analytical Chemistry, 27(3), 228-237.

Wishart, D. S., Knox, C., Guo, A. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J. G., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhutdinov, R., Li, L., Vogel, H. J., & Forsythe, I. (2009). HMDB: a knowledgebase for the human metabolome. Nucleic Acids Research, 37, D603-D610.

Wright, A. J., Buydens, L. M. C., & Heerschap, A. (2012). A phase and frequency alignment protocol for ^{1}H MRSI data of the prostate. NMR in Biomedicine, 25(5), 755-765.

Wu, C. L., Jordan, K. W., Ratai, E. M., Sheng, J., Adkins, C. B., Defeo, E. M., Jenkins, B. G., Ying, L., McDougal, W. S., & Cheng, L. L. (2010). Metabolomic imaging for human prostate cancer detection. Science Translational Medicine, 2(16), 16ra18.

Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D., & Wishart, D. S. (2012). MetaboAnalyst 2.0—a comprehensive server for metabolomic data analysis. Nucleic Acids Research, 40(W1), W127-W133

Xie, B., Waters, M. J., & Schirra, H. J. (2012). Investigating potential mechanisms of obesity by metabolomics. Journal of Biomedicine and Biotechnology, 2012, Article ID 805683, 1-10.

Yong, W. P., Innocenti, F., & Ratain, M. J. (2006). The role of pharmacogenetics in cancer therapeutics. British Journal of Clinical Pharmacology, 62(1), 35-46.

Zaichick, V. Y., Sviridova, T. V., & Zaichick, S. V. (1996). Zinc concentration in human prostatic fluid: normal, chronic prostatitis, adenoma and cancer. International Urology and Nephrology, 28(5), 687-694.

Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. Analyst, 137(2), 293-300.