Faculty of Medicine

IN SEARCH FOR POTENTIAL GENETIC AND PROTEOMIC BIOMARKERS IN LUNG CANCER AND IN *STAPHYLOCOCCUS AUREUS* BACTERAEMIA BY MASS SPECTROMETRY

Annina Rostila

Doctoral dissertation, to be presented for public discussion with the permission of the Faculty of Medicine of the University of Helsinki, in lecture room Porthania PIII, Porthania building, Yliopistonkatu 3, Helsinki,
on September 26, 2020, at 12 noon.

Helsinki 2020

Department of Biochemistry and Developmental Biology
University of Helsinki
Helsinki, Finland

Doctoral School in Health Sciences (DSHealth)
Doctoral Programme in Biomedicine (DPBM)
Unigrafia
Helsinki 2020

ISBN 978-951-51-6367-7 (paperback)
ISBN 978-951-51-6368-4 (PDF)

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

Supervised by

Docent Sisko L Anttila, MD, PhD

Department of Pathology

Helsinki University Central Hospital

Helsinki, Finland

Docent Anne Puustinen, PhD

Finnish Institute for Verification of the Chemical Weapons Convention

Department of Chemistry

University of Helsinki

Helsinki, Finland

Docent Marc Baumann, MD, PhD

Department of Biochemistry and Developmental Biology

University of Helsinki

Helsinki, Finland

Reviewed by

Docent Steffen Ohlmeier, PhD

Faculty of Biochemistry and Molecular Medicine

University of Oulu

Oulu, Finland

Petri Kouvonen, PhD

Thermo Fisher Scientific / University of Turku

Turku, Finland

Official Opponent

Professor Outi Kuittinen, MD, PhD

Institute of Clinical Medicine

University of Eastern Finland

Kuopio, Finland

Victory is reserved to those who are willing to pay the price.

- Sun Tzu

ABSTRACT

Aims of the study

The aims of this were to: 1) identify CRP-gene SNPs associated with clinical outcome, and blood CRP-levels in SA bacteraemia, 2) identify protein biomarkers for early diagnosis of lung cancer in high-risk individuals, such as tobacco smokers, and individuals significantly exposed to asbestos, and 3) evaluate and compare the use of different biological fluids, plasma and sputum, in proteomic biomarker discovery with mass spectrometry.

Study populations and methods

To study the effect of CRP gene polymorphisms on plasma (or serum) CRP levels, DNA was extracted from patients (n=145) with SA positive blood cultures. SNP genotyping was done with MALDI-TOF.

Identification of potential protein biomarkers in lung cancer was done using plasma (n=205) and sputum (n=123) samples obtained from former and current smokers, and from former and current smokers with lung cancer, asbestos exposure, or both. 2D SDS-PAGE MS was used to study the differential protein abundancy in plasma samples, 2D-DIGE MS was used for the sputum samples. Immunoblotting and ELISA were used to validate the results.

Results

CRP gene SNP rs3091244 with minor allele A, and SNP rs3093075 with minor allele T, together in strong linkage disequilibrium, were significantly associated with high maximal plasma CRP levels during the first week of illness. Individuals with the recessive A-allele of triallelic SNP rs3091244, had a median of maximal plasma CRP level of 282 mg/L (IQR 169 mg/L), whereas individuals with alleles T or C had a median of maximal CRP level of 179 mg/L (IQR 148 mg/L).

28 differentially abundant plasma proteins were found across four study groups, and 22 sputum proteins were differentially abundant across three study groups. Upregulation of proteins involved in redox-reactions, PRDX1 (plasma), TXN (sputum), and GAPDH (sputum), associated with lung cancer. Low plasma levels, but high sputum levels of PRDX2 were associated with lung cancer. S100A8/9 from sputum was strongy associated with sputum GAPDH, thus could associate with lung cancer. TPM3 and TPM4 were scarce in plasma from individuals with non-metastasized lung cancer. TPM4 was also associated with asbestos exposure. SAA abundance in plasma was a good marker for lung cancer but lacked specificity. ECM1 was validated from sputum but did not reach statistically significant differences in its abundance between study groups.

Conclusions

Based on the studies comprising this thesis, the following conclusions can be drawn: 1) The A-allele of SNP rs3091244 affects the CRP levels in the blood during SAB more than the other alleles, without any significant association with *e.g.* mortality. This should be considered when treating SAB patients; 2) PRDX1 (plasma), PRDX2 (sputum), and TXN (sputum), have the highest potential to become predictive or diagnostic biomarkers for lung cancer. PRDX2 and TXN interact with other cancer-related proteins, GAPDH and S100A8/9, thus their molecular mechanisms are complex; 3) Abundant TMP4 in plasma best indicates exposure to asbestos; 4) Mass spectrometry, and its coupled methods (*e.g.* gel electrophoresis), are powerful tools in finding genome and proteome alterations; 5) Sputum and plasma are suitable for biomarker discovery, but more studies are needed to make sputum more easily available for high-throughput mass spectrometry. In conclusion, these studies are in concordant with the study aims, and complement the molecular background on SAB, lung cancer and asbestosis, and suggest feasible predictive or diagnostic biomarkers for future studies with mass spectrometry.

ABSTRACT IN FINNISH – TIIVISTELMÄ

Tutkimuksen tavoitteet

Tämän väitöskirjan tavoitteena oli: 1) tunnistaa CRP-geenin snippien assosiaatioita stafylokokkisepsiksen kliiniseen lopputulemaan sekä veren CRP-pitoisuuksiin, 2) tunnistaa proteiineja, jotka voisivat toimia biomarkkereina keuhkosyövän varhaisvaiheen diagnosoinnissa korkean riskin yksilöissä, eli tupakoijissa sekä asbestille altistuneissa sekä 3) arvioida ja vertailla plasman ja yskoksen soveltuvuutta proteiinibiomarkkerien tutkimiseen massaspektrometrialla.

Aineisto ja menetelmät

CRP-geenin snippien vaikutusta veren CRP-pitoisuuksiin tutkittiin stafylokokkisepsikseen sairastuneista henkilöistä (n=145). Snipit määritettiin monistetusta genomisesta DNA:sta massaekstensionmenetelmällä ja snippien määrittämiseen käytettiin MALDI-TOF-massaspektrometriaa.

Proteiinibiomarkkereita tutkittiin plasmasta (n=205) ja yskoksestä (n=123) keuhkosypään sairastuneilla, asbestille altistuneilla, sekä oireettomilla verrokeilla. Kukin ryhmä jakautui edelleen tupakoijiin ja tupakoinnin lopettaneisiin. Muutamalla tutkittavalla oli sekä keuhkosypä että altistus asbestille. Proteiinien ilmenemistä eri ryhmissä tutkittiin plasmasta 2D SDS-PAGE MS:lla ja yskoksestä 2D-DIGE MS:lla. Proteiinien ilmenemiserot validoitiin Western Blot-menetelmällä sekä ELISA:lla.

Tulokset

CRP geenin SNP rs3091244 alleeli A ja SNP rs3093075 alleeli T olivat keskenään kytkentäepätasapainossa ja assosioituivat tilastollisesti merkitsevästi veren maksimaalisiin CRP-pitoisuuksiin ensimmäisen sairausviikon aikana. SNP rs3091244 A-alleelin kantajilla oli ensimmäisen sairausviikon maksimaalisten CRP-pitoisuuksien mediaani 282 mg/L (IQR 169 mg/L), kun C- ja T-alleelin kantajilla vastaava luku oli 179 mg/L (IQR 148 mg/L).

Redox-reaktioissa toimivat proteiinit, PRDX1 (plasma), TXN (yskös) ja GAPDH (yskös) ilmenivät runsaina keuhkosyövässä. PRDX2:a oli vähän keuhkosyöpäryhmään kuuluvien plasmassa, mutta paljon saman ryhmän ysköksissä. S100A8/9 korreloi merkitsevästi proteiinin GAPDH kanssa, mistä syystä S100A8/9 voi olla merkittävässä roolissa myös keuhkosyövässä. Proteiineja TPM3 ja TPM4 oli plasmassa niukasti niillä, joiden keuhkosyöpä ei ollut metastasoitunut. TPM4 korreloi asbestialtistuksen kanssa. ECM1 validoitiin, mutta sen runsaudessa ei havaittu eroja eri tutkimusryhmien välillä.

Johtopäätökset

Tähän väitöskirjaan kuuluvien osajulkaisujen tulosten perusteella voidaan todeta seuraavat johtopäätökset: 1) Snipin rs3091244 A-alleeli vaikuttaa veren CRP-tasoihin muita alleeleja voimakkaammin SAB:n aikana, vaikka esim. kuolleisuudessa ei ole eri alleelien kantajilla merkittävää eroja. Tämä tulisi huomioida hoitopäätöksiä tehtäessä; 2) Proteiineista PRDX1 (plasma), PRDX2 (yskös) ja TXN (yskös) voisi tulla keuhkosyövän ennustavia tai diagnostisia biomarkkereita. PRDX2 ja TXN korreloivat merkittävästi ysköksen proteiinien GAPDH ja S100A8/9 kanssa, joista voisi myös olla hyötyä keuhkosyövän biomarkkereina; 3) Plasman proteiinia TPM4 voisi käyttää merkittävän asbestialtistuksen biomarkkerina; 4) Massaspektrometria, sekä siihen yhdistettynä esim. geelielektroforeesi, on käyttökelpoinen menetelmä proteomin ja genomin muutosten havaitsemiseen; 5) Yskös ja plasma ovat biomarkkeritutkimukselle sopivia näytetyyppejä, mutta lisää tutkimusta tarvitaan, jotta ysköksiä voitaisiin helposti käyttää massaspektrometriassa. Yhteenvetona voidaan todeta, että johtopäätökset tukevat tutkimukselle asetettuja tavoitteita ja täydentävät molekyylibiologista ymmärrystä stafylokokkisepsiksestä, keuhkosyövästä ja asbestialtistuksesta. Tulokset lisäksi ehdottavat ennustavia tai diagnostisia biomarkkereita massaspektrometrialla toteutettavia jatkotutkimuksia varten.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Mölkänen T*, Rostila A*, Ruotsalainen E, Alanne M, Perola M, Järvinen A. **Genetic polymorphism of C-reactive protein (CRP) gene and a deep focus of infection determine maximal serum CRP level in Staphylococcus aureus bacteremia.** *Eur J Clin Microbiol Infect Dis* 2010;29(9):1131-1137.

II Rostila A, Puustinen A, Toljamo T, Vuopala K, Lindström I, Nyman TA, Oksa P, Vehmas T, Anttila SL. **Peroxiredoxins and tropomyosins as plasma biomarkers for lung cancer and asbestos exposure.** *Lung Cancer* 2012;77(2):450-459.

III Rostila A, Anttila SL, Lalowski M, Vuopala K, Toljamo T, Lindström I, Baumann M, Puustinen A. **ROS-regulating proteins peroxiredoxin 2 and thioredoxin, and glyceraldehyde-3-phosphate dehydrogenase are differentially abundant in induced sputum from smokers with lung cancer or asbestos exposure**. *European Journal of Cancer Prevention* 2020;29(3):238-247.

* authors have contributed equally to this study

ABBREVIATIONS

2D-DIGE	two-dimensional difference gel electrophoresis
2DE	two-dimensional gel electrophoresis
ALK	anaplastic lymphoma kinase
APCS	serum amyloid P-component
APOA1	apolipoprotein A1
APP	acute phase protein
APR	acute phase response
ASIP	agouti signaling protein
ASTA	astistaphylosin
BaP	benzo(a)pyrene
BRAF	v-Raf murine sarcoma viral oncogene homolog B
BRCA	breast cancer-associated gene
C5	component C5
cfRNA	cell-free RNA
CHIP	chemotaxis inhibitory protein
CRP	C-reactive protein
ctDNA	circulating tumor DNA
Eap	extracellular adherence protein
ECM	extracellular matrix protein
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbent assay
ESI	electronspray ionization
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GDR	genotype-driven recruitment
GWAS	genome-wide association studies
hCG	human corion gonadotropin
HDL	high-density lipoprotein
HER2	human epidermal growth factor receptor 2
HIV	human immunodeficiency virus

HPLC	high performance liquid chromatography
ICAM	intracellular adhesion molecule
IL	interleukin
KEAP1	kelch-like ECH-associated protein
KRAS	kirsten rat sarcoma viral oncogene homolog
LC	liquid chromatography
LD	linkage disequilibrium
LDL	low-density lipoprotein
LOH	loss of heterozygosity
LOX	lectin-like oxidized low-density lipoprotein receptor
LPS	lipopolysaccharide
LT	lymphotoxin
m/z	mass-to-charge
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
MET	MET proto-oncogene
miRNA	micro RNA
MS	mass spectrometry
MYC	myelocytomatosis
NRF2	transcription factor nuclear-factor-erythroid 2 p45-related factor 2
NSCLC	nonsmall cell lung cancer
PAH	polycyclic aromatic hydrocarbon
PCT	procalcitonin
PRDX	peroxiredoxin
PSA	prostate-specific antigen
PSM	phenol-soluble modulin
Q-TOF	quadrupole time-of-flight (mass spectrometer)
RBP4	retinol-binding protein 4
RET	RET proto-oncogene
RFLP	restriction fragment length polymorphism
RNS	reactive nitrogen species
ROS	reactive oxygen species

ROS1	proto-oncogene tyrosine-protein kinase ROS
S100A8	S100 calcium-binding protein A8
SA	*Staphylococcus aureus*
SAA	serum amyloid A
SAB	*Staphylococcus aureus* bacteraemia
SCLC	small cell lung cancer
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIRS	systemic inflammatory response syndrome
SNP	single nucleotide polymorphism
suPAR	soluble urokinase-type plasminogen activator receptor
TAA	techoid acid antibody
TFF	trefoil factor
TLR	toll-like receptor
TNF	tumor necrosis factor
TNM	tumor nodus metastasis
TP53	tumour protein 53
TPM	tropomyosin
TSST	toxic shock syndrome toxin
TTF-1	thyroid transcription factor 1
TTR	transthyretin
TXN	thioredoxin
TYR	tyrosinase
TYRP1	tyrosinase related protein 1
VLDL	very low-density lipoprotein
VNTR	variable number of tandem repeats
WCC	white blood cell count
WES	whole exome sequencing
WGS	whole genome sequencing

1. INTRODUCTION

Biomarkers are defined as measurable and reliable biomolecules that indicate the state of a biological process, or therapeutic response, typically nucleic acids, peptides, proteins, carbohydrates, lipids, or whole cells (Biomarkers Definitions Working Group, 2001). Clinical biomarkers can indicate a variety of health- or disease-related characteristics, such as genetic susceptibility, clinical disease, exposure to an environmental factor, or response to therapy (Chen *et al*., 2011), but a good clinical biomarker is specific, sensitive, easy to measure, minimally invasive to collect, and simple to interpret (Srivastava and Gopal-Srivastava, 2002). Good clinical biomarkers are hard to discover, as many lack sensitivity and specificity for a particular disease. For clinical purposes, biomarkers can be divided into three categories; predictive, diagnostic, and prognostic (Kisluk *et al*., 2014).

Predictive biomarkers are used for predicting a clinical outcome, monitoring the treatment response, or for evaluationg the disease risk. Diagnostic biomarkers are used for distinguishing an existing disease; hence they need to be specific and sensitive at an early stage and correlate directly with the manifestation of a disease. Such biomarkers are hard to obtain, especially for multifactorial diseases like cancers or infections. Prognostic biomarkers indicate survival, but a good prognostic biomarker can also aid in therapeutic decision making.

Advances made in molecular biology techniques, *i.e.* the “omics” referring to genomics, proteomics, transcriptomics, and metabolomics, have offered new possibilities in biomarker research. The development of mass spectrometry, and its coupled techniques, have made it possible to study a variety of sample specimen some with complex sample matrices (Hudler *et al*., 2014), such as the plasma proteome (Nanjappa *et al*., 2014). Together with the collection of high-quality biobanks (Olson *et al*., 2014), they enable the development of personalized medicine in which biomarkers play a key role (Vogenberg *et al*.,

2010). From a clinical viewpoint, it is essential to have measurable molecular biomarkers in daily use in order to treat the patients correctly. Such a biomarker is the C-reactive protein (CRP), which is used *e.g.* to distinguish between bacterial and viral infecions.

C-reactive protein (CRP) abundance of over 100 mg/l in blood typically indicates bacterial infection (Jeong *et al.*, 2012), and in severe cases, CRP levels can reach up to 10 000-fold above baseline (Pepys and Hirschfield, 2003). Upon the resolution of inflammation, the CRP levels decrease rapidly (Vigushin *et al*., 1993). Blood CRP levels are used in following *Staphylococcus aureus*-induced bacteraemia (SAB), since increased levels are known to associate with the severity of the disease (Mölkänen *et al*., 2016). Those with high CRP levels are prone to develop organ dysfunction, whereas rapidly decreasing CRP values in the course of the disease declare improvement. There are strong genetic components, *e. g.* single nucleotide polymoprhisms (SNPs), affecting the blood CRP values (Kushner et al., 2006; Kluft *et al*., 2003; Kathiresan *et al*., 2006), hence straightforward associations between the CRP values and the severity of SAB cannot be made.

Lung cancer is a deadly disease partially due to difficulties in early diagnosis (reviewed in Manser *et al*., 2013). Even with known risk factors, *e.g.* tobacco smoking and asbestos exposure, high-risk individuals are challenging to screen. Traditional imaging methods, such as with x-ray or computer tomography (CT), lack effectivity in detecting lung tumours early enough to reduce lung cancer mortality, thus predictive and diagnostic biomarkers are greatly needed. Lung cancer, typically for all cancers, harbours a variety of genetic aberrations, of which the most established kirsten rat sarcoma viral oncogene homolog (KRAS) and epidermal growth factor receptor (EGFR) mutations associate with lung adenocarcinomas (Pan *et al*., 2014). A variety of diagnostic biomarker proteins have been studied in lung cancer tissues and cell lines, of which napsin A and TTF1 are generally used in pathology practice to differentiate lung adenocarcinomas from squamous cell

carcinomas and from carcinomas of other organs metastatic to lungs (*Ma et al.*, 2015; Ao *et al*., 2014). Tissue biopsies, however, are invasive, and as such, not suitable for predicting cancer risk within healthy, asymptomatic individuals.

This thesis aimed at using mass spectrometry (MS) in identifying biomarkers in two high-mortality diseases, *Staphylococcus aureus* bacteraemia and lung cancer. Genomics MS was used to identify SNPs of the CRP gene associated with blood CRP levels in SAB, and with SAB severity. Proteomic approach was used to study proteins associated with lung cancer, smoking, and asbestos exposure from plasma and sputum. All studies aimed at finding either potential predictive, diagnostic or prognostic biomarkers that could further aid in risk evaluation, and treatment of SAB or lung cancer.

2. REVIEW OF THE LITERATURE

2.1 BIOMARKERS

Traditionally biomarkers have been described as biomeasurements, such as body mass index, blood pressure, or body temperature, clinically proven to indicate an individual´s state of health. These biomeasurements have long helped doctors to objectively evaluate risks and make decisions on suitable treatment options by relying on distinct measuments even before any disease symptoms have arisen. The development of molecular biology techniques has broadened the meaning of biomarkers from these tradional, and rather robust biomeasurements into biomolecules. Thus, biomarkers are now defined as measurable and reliable biomolecules that indicate the state of biological process, or therapeutic response, and can consist of *e.g.* nucleic acids, peptides, proteins, carbohydrates, lipids, or whole cells (Biomarkers Definitions Working Group, 2001). According to the International Programme on Chemical Safety led by the World Health Organization, a biomarker can be any substance, structure, or even a process that can be measured in the body or from its products, and influences or predicts an endpoint or a disease (World Health Organization, 2001).

2.1.1 *Guidelines for biomarker research*

Most candidate biomarkers, despite the effort used in finding them, never make it beyond their discovery, even with major technological advances in their examination from a single specimen. To overcome this problem, general biomarker discovery guidelines have been proposed that take into account study design (Mischak *et al.*, 2010), and organization of the field of biomarker research (Institute of Medicine, 2007) (**Figure 1**).

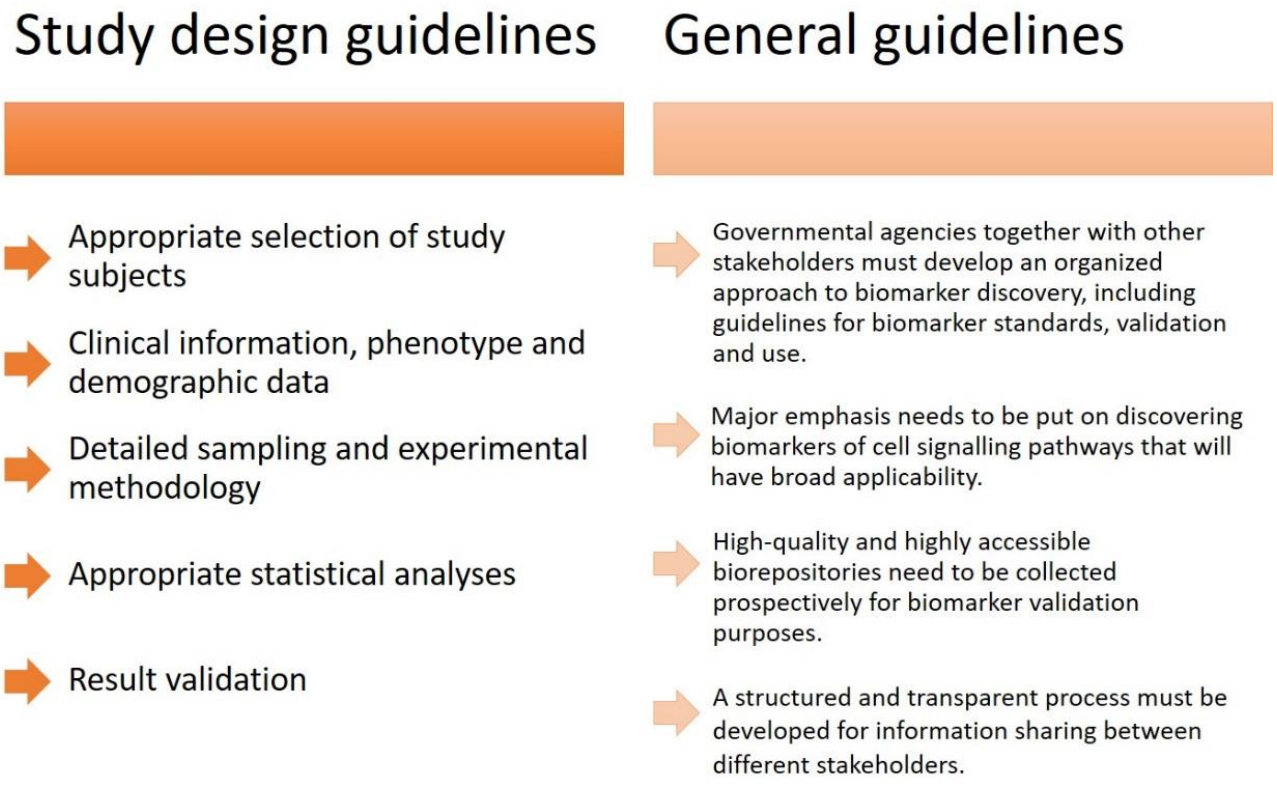


Figure 1 *Guidelines for improved biomarker discovery.*

2.1.2 *Different types of clinical biomarkers*

Biomarkers can be organized into different categories according to their clinical purpose (Kisluk *et al*. 2014) (**Figure 2**), and it is the categorization used throughout this thesis. Predictive biomarkers can either be used for predicting a clinical outcome, or for monitoring the treatment response. The T790M mutation found in the EGFR gene in lung adenocarcinomas (Pan *et al*., 2014) serves as an example of a predictive biomarker, since it arises during the targeted therapy and associates with resistance to 1st line tyrosinase kinase inhibitors. The term “predictive” refers to those biomarkers that can help in identifying the risks for developing a disease or indicate response to therapy, wheras diagnostic biomarkers are useful for the precise diagnosis of an existing disease. The diagnostic biomarkers should be the most specific and sensitive at an early stage, and correlate directly with the manifestation of a disease, but such biomarkers are difficult to obtain. For instance, procalcitonin (PCT) is a diagnostic biomarker in bacterial bloodstream infections, *i.e.* bacteraemias, even with a mean sensitivity of 77 %, and a mean specificity of 79 % (Wacker *et al.,* 2013). PCT is also used as a prognostic biomarker for deciphering the outcome of the overt disease, since high levels of PCT in blood indicate high mortality in bacteraemias, albeit with a sensitivity and specificity of around 80 % (Schneider *et al*., 2009, Lee *et al*., 2008). A good prognostic biomarker can also aid in therapeutic decision making.

In 1987, the National Research Counsil divided biomarkers into three groups, according to their effects (**Figure 3**). Exposure biomarkers are substances that can be found in the body or its secretions after exposure to a chemical. Such a biomarker can be *e.g.* carbon monoxide in expired air after a significant exposure has taken place (Lowry, 1995). The biomarker of effect would then be the carboxyhemoglobin found in blood, since carbon monoxide replaces hemoglobin-bound oxygen in red blood cells. Biomarker of susceptibility refers to a genetic factor that predisposes to a disease caused

by environmental exposure (National Research Council, 1987). Such biomarkers are *e.g.* genetic variants that predispose their carriers to UV-radiation-induced skin cancer, like the variants in genes tyrosinase (TYR) (Gudbjartsson *et al*., 2008), agouti signaling protein (ASIP), and tyrosinase related protein 1 (TYRP1) (Slingluff *et al*., 2000).

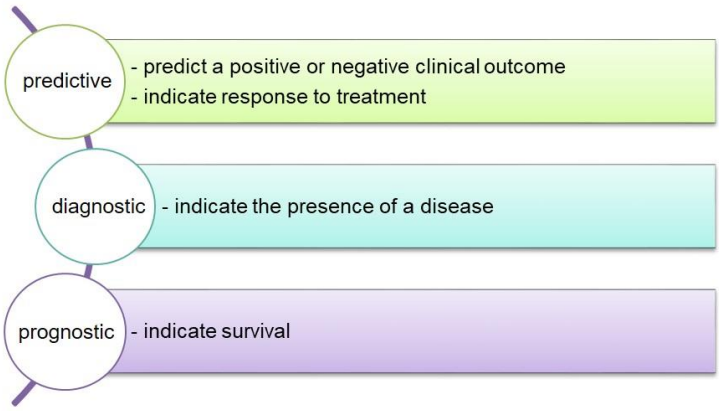


Figure 2 *Predictive, diagnostic, and prognostic biomarkers.*

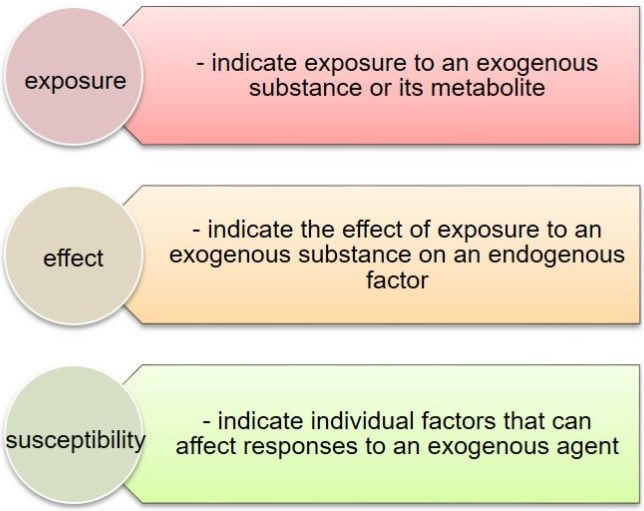


Figure 3 *Biomarkers of exposure, effect, and susceptibility.*

Another way of classifying biomarkers is by disease stage (Biomarkers Definitions Working Group, 2001; Chen *et al*., 2011) (**Figure 4**). Antecedent biomarkers are used for evaluating the risk of a clinical disease. These include *e.g.* mutations in the breast cancer 1 gene, which cause a lifetime risk of 65-80 % for breast cancer (Ford *et al*., 1994; Antoniou *et al*., 2003). Screening biomarkers help in recognizing an unsymptomatic disease, even on a population level. Such a biomarker is *e.g.* fecal hemoglobin that can be used to identify colorectal cancer (Zhu *et al*., 2010), or PSA in screening of prostate cancer (Catalona *et al*., 1991). CRP is used to identify between bacterial and viral infection (Jeong *et al.*, 2012), and to estimate the severity of a disease (Mölkänen *et al*., 2016). It can therefore be classified both as a diagnostic and a staging biomarker. Polymorphisms of the CRP gene are known to have an effect on survival in bacteraemias, hence the polymorphisms serve as prognostic biomarkers.

Figure 4 *Biomarkers according to disease stage.*

2.1.3 *Biomarkers in clinical use*

Clinical biomarkers can indicate a variety of health- or disease-related characteristics, such as genetic susceptibility, clinical disease, exposure to an environmental factor, or response to therapy (Chen *et al*., 2011). A good clinical biomarker is specific, sensitive, easy to measure, minimally invasive, and simple to interpret (Srivastava and Gopal-Srivastava, 2002). The biological characteristics monitored by biomarkers are often represented by pathogenic changes and embody a multitude of malignancies like cancer (Srivastava and Gopal-Srivastava, 2002), cardiovascular diseases (Vasan, 2006), metabolic disorders (Falahi *et al.,* 2015), and infectious diseases (Larsen and Petersen, 2017). In patient care, following therapeutic responses via biomarker tracking is essential especially in drug development, and therapy monitoring. Some examples of the biomarkers used in clinical decision making include glucose for type I diabetes (American Diabetes Association, 1997), phenylalanine for phenylketonuria (Scriver *et al*., 1994), C-reactive protein (CRP) for bacterial infection (Liu *et al.,* 2016), low-density lipoprotein (LDL) for cardiovascular risk (Albert, 2011), human chorionic gonadotropin (hCG) pregnancy (Marshall *et al*., 1968), prostate specific antigen (PSA) for prostate cancer (Kirby, 2016), and a variety of somatic and germ line genetic changes, like epidermal growth factor receptor gene (EGFR) mutations in lung cancer (Vincent *et al*., 2012), and heritable breast cancer 1 gene (BRCA1) (Martin and Weber, 2000) mutations (**Table 1**).

Table 1. *Examples of biomarkers in clinical use*

BIOMARKER NAME	BIOMARKER MOLECULE	IMPLICATION OF BIOMARKER
GLUCOSE	Carbohydrate	Plasma hyperglycemia, or manifestation in urine indicates diabetes mellitus type 1.
PHENYLALANINE	Amino acid	Increased plasma levels indicate phenylketonuria.
C-REACTIVE PROTEIN	Protein	Increased plasma levels indicate *e.g.* bacterial infection, risk for type 2 diabetes mellitus, or risk for cardiovascular diseases.
LOW-DENSITY LIPOPROTEIN	Lipoprotein	Elevated concentration in plasma indicates cardiovascular risk.
HUMAN CHORION GONADOTROPIN	Glycoprotein (hormone)	Manifestation in urine or plasma indicates pregnancy.
PROSTATE SPECIFIC ANTIGEN	Glycoprotein (enzyme)	Elevated concentration in plasma may indicate prostate cancer.
BREAST CANCER 1	DNA: single nucleotide mutation	Increased risk of breast and ovarian cancer.
EPIDERMAL GROWTH FACTOR RECEPTOR	DNA: *e.g.* changes in copy number, exon deletions.	Increased risk of NSCLC, especially adenocarcinoma. Resistance to chemotherapy and radiation. Good response to treatment with erlonitib or gefinitib.

2.1.4 *The use of biomarkers in personalized medicine*

Personalized medicine, sometimes referred to as precision medicine, is the individually tailor-made treatment that is based upon combining and analyzing biomarker data obtained from genomics, transcriptomics, proteomics, and metabolomics (**Figure 5**). The development of biotechnology, and accumulated knowledge on both heterogeneousness of diseases, and individual differences in responses to treatment, hold promise to medicine that does not only treat symptoms or cure diseases but prevents them by detecting early stage biomarkers through whole genome, proteome, trancriptome, and metabolome screening. This would, in time, lead to clinical practice that not only enables optimal therapy, or improves therapy monitoring, but reduces treatment times, and lowers the costs (Vogenberg *et al*., 2010).

Biobanks play an important role in supporting the development of personalized medicine, offering biological specimen for biomarker studies (Olson *et al*., 2014). Biomarker research has typically been phenotype-driven, but accumulated data from biobanks now enables genotype-driven recruitment (GDR) in genetic studies *i.e.* study group selection on the basis of genotype (McGuire and McGuire, 2008) instead of phenotype.

2.1.5 *Genomic biomarkers*

Virtually all medical issues are influenced by our genome. Some diseases posses a Mendelian genetic aetiology, making them easily available for genetic testing, genetic screening, and pedigree analysis. However, a growing proportion of our genetic disorders are common complex diseases caused by multiple predisposing genetic factors often in co-operation with our environment. Mapping our genome for these biomarkers for evaluation of

disease risk, or for precision therapy, creates the core of personalized genomics.

The human genome comprises of about 3 x 10^9 bases of which estimated 1 % is coding sequence (The Encode Project Consortium, 2007). These coding sequences make up about 19 000 protein-coding genes, 8000 pseudogenes, and 4000 non-coding RNA genes (Seal *et al.*, 2011). On approximation, 85 % of phenotype-affecting genetic mutations take place in the coding regions (Botstein and Risch, 2003), of which only 10 % are characterized (Rizzo and Buck, 2012).

Many worldwide projects, starting from the Human Genome Project initiated in 1990 (Lander *et al.*, 2001), have tried tackling the complexity of the genome, mapping the genome from different angles. For instance, the HapMap project concentrated on discovering the 3.5 million single nucleotide polymorphisms (International HapMap Consortium, 2007), SNPs, affecting many common diseases, such as cancer or heart diseases. The 1000 Genomes Project was aimed to deep sequence the human genome and discover structural variations from 2504 individuals from 26 different populations (1000 Genomes Project Consortium, 2010). Mapping the human genome and discovering the common variants has given rise to the genome-wide association studies (GWAS), and more recently to whole genome sequencing (WGS), whole exome sequencing (WES), and to static gene panel testing (Brittain *et al.*, 2017). WES in particular holds promise to uncover monogenic disease mutations, and genetic susceptibility to variants predisposing to *e.g.* cancer, thus making it an alluring tool for personalized genomics. **Figure 6** represents a possible layout of WES in personalized medicine.

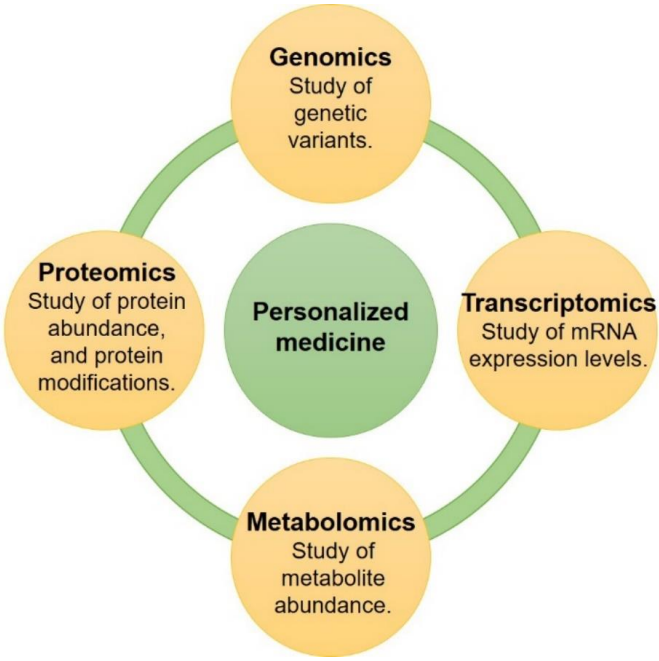


Figure 5 *Personalized medicine is comprised of determining biomarkers on the genome, transcriptome, proteome, and metabolome level.*

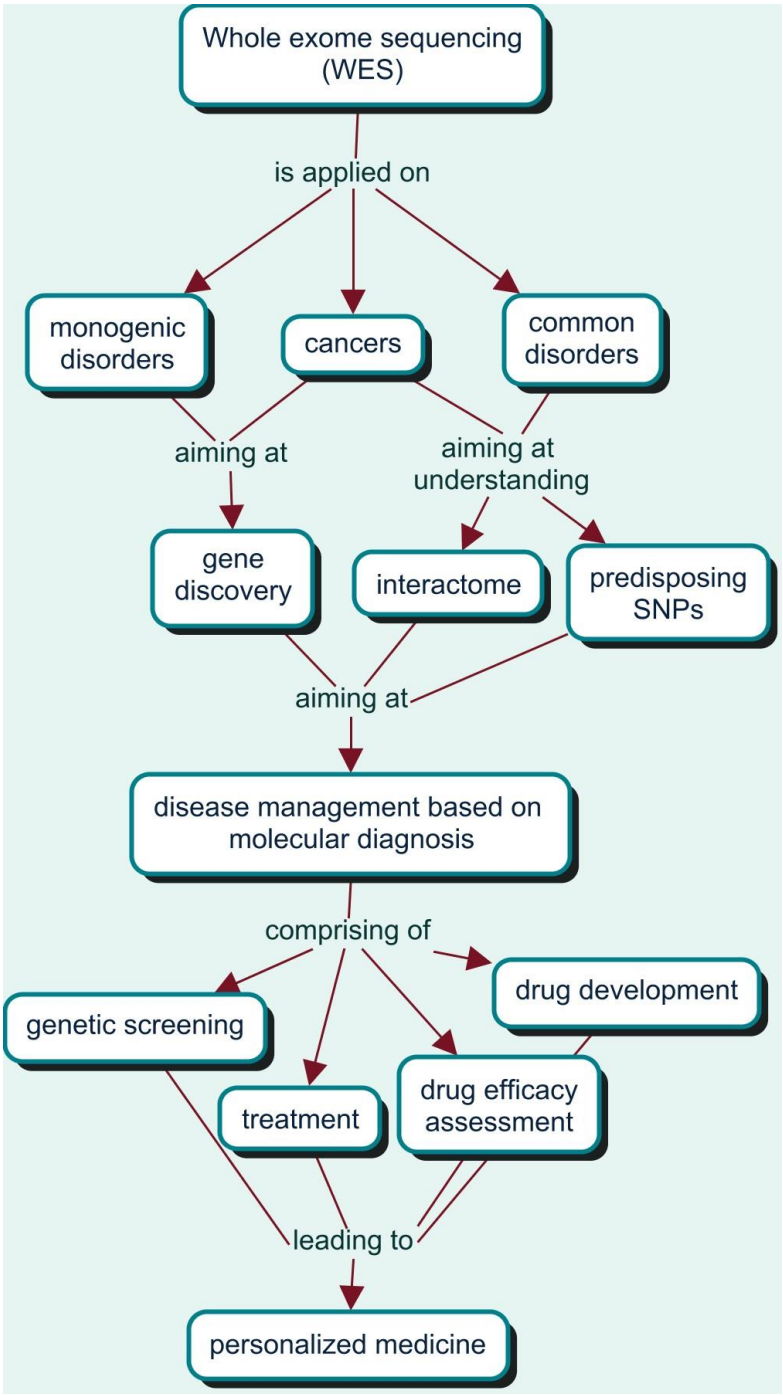


Figure 6 *From whole exome sequencing to personalized medicine.*

2.1.6 *Proteomic biomarkers*

Proteome is used to describe the protein consistency of *e.g.* a cell, tissue or biological fluid (such as plasma), at a given time point. Thus, the proteome represents more accurately than the genome (or the trancriptome) the actual phenotype of its source, as proteins are the functional end-point products of protein-coding genes (Hanash and Taguchi, 2010), and since variations on the DNA level are not always reflected on the protein level. Especially in genetically heterogeneous diseases like cancers, the proteome can offer more insight on the disease than the genome by unmasking the molecular mechanisms of the disease, since changes in the protein level offer pathway-specific knowledge on the disease (Latosinska *et al*., 2017).

The majority, 96 %, of molecular drug targets are proteins (Santos *et al*., 2017). Yet, most drugs are shown to be ineffective in 60 %, or in oncology in more than 75 %, of the patients of the disease population (Spear *et al*., 2001; Finley Austin and Babbiss, 2006). The technical development, such as high-resolution mass spectrometry in proteomics has rapidly led to the development of clinical proteomics aiming at analyzing the proteome of a variety of biological specimen (Latosinska *et al*., 2017). This in turn enables the advancement in personalized proteomics, *i.e.* proteomic profiling, in which individually designed drug therapy and monitoring of its biological responses is of importance. Personalized proteomics, especially in oncology, also aims at early recognition of the disease, or risk of disease on the proteome level, thereby improving prevention, and early detection of cancers. **Figure 7** represents a possible schema from protein biomarker discovery to personalized medicine.

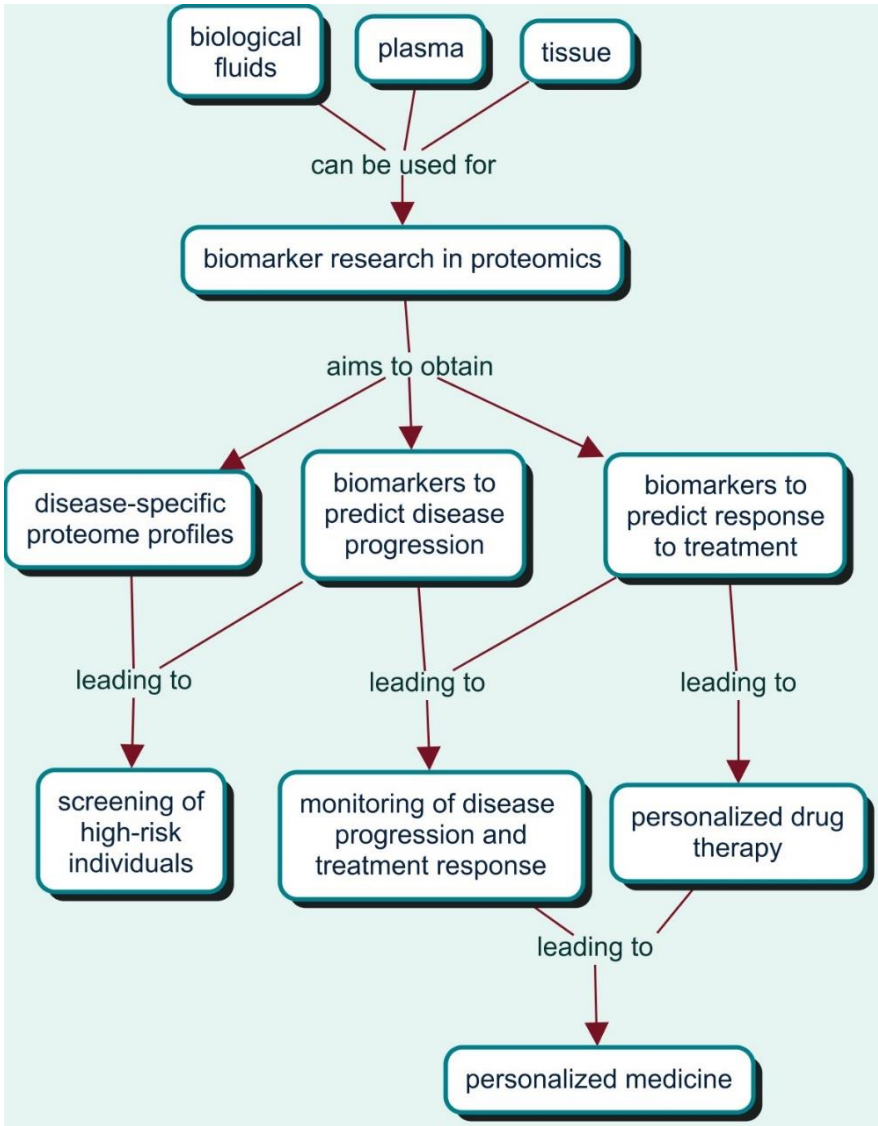


Figure 7 *From protein biomarker discovery to personalized medicine.*

2.1.7 *Sample matrices in biomarker research*

The human body and its secretions offer a multitude of different sample matrices appropriate for biomarker research. Especially liquid biopsies, such as blood, sputum, saliva or urine, offer a non-invasive sample matrix for

biomarker studies. For instance, in lung cancer, tissue biopsies can be hard to get, and the procedure can be harmful to the patient (Guibert *et al*., 2020). Collecting *e.g.* sputum that contains tumour derived cells, or proteins secreted by tumour cells, offers a non-invasive method for detecting changes in the genome or the proteome, or monitoring therapy response.

In proteomics, biomarkers have been studied from *e.g.* tissue biopsies (Chaurand *et al*., 2004), plasma (Geyer *et al*., 2017), urine (Rodríguez-Suárez *et al*., 2014), saliva (Xiao *et al*., 2016), bile (Barbhuiya *et al*., 2011), stool (Walsham and Sherwood, 2016), sputum (Suojalehto *et al.*, 2015), amniotic fluid (Tsangaris *et al*., 2006), cerebrospinal fluid (Khoonsari *et al*., 2016), bronchoalveolar lavage (Teirilä *et al*., 2014), follicular fluid (Oh *et al*., 2017), pancreatic fluid (Williams, 2013), and cell cultures (Geiger *et al*., 2012). Some of the most typical sample matrices for DNA or RNA isolation, *i.e.* the genomic biomarker discovery, are whole blood (Lin *et al*., 2009; Crowley *et al*., 2013; Klein *et al*., 2010), white blood cells (Alessandro *et al.*, 2020), tumour cells or biopsies (Ignatiadis *et al*., 2015; Moschini *et al*., 2016), urine (El Bali *et al*., 2014), and cell lines (Garnett *et al*., 2012). This thesis focused on plasma and sputum as sample matrices. Plasma was chosen for its easy accessibility in routine diagnostics, and sputum for its noninvasive sample collection from site of lung carcinogenesis.

2.1.7.1 Blood

Blood is a good source of candidate biomarkers, since it can be obtained in a minimally invasive way under standardized conditions in a clinical laboratory. Blood is a complex sample matrix withholding whole cells and a variety of biomolecules, such as proteins, lipids, carbohydrates, nucleic acids, and metabolic residues, all of which have the potential to serve as biomarkers. Proteomics typically focuses on studying the plasma proteome (Geyer *et al*., 2017), but genomics can use either a genome derived from

white blood cells (*e.g.* Alessandro *et al.*, 2020), or whole blood (*e.g.* Lin *et al*., 2009; Crowley *et al*., 2013).

The plasma proteome, currently known to consist of more than 10 000 proteins (Nanjappa *et al*., 2014), is described as the largest and most complex of the human proteomes. From a functional viewpoint, these proteins can be classified into eight categories: immunoglobulins, long and short distance receptors, temporary passengers, tissue leakage products, aberrant secretions, foreign proteins, and proteins secreted by solid tissues that have functions in the plasma (**Figure 8**) (Anderson and Anderson, 2002). Twenty-two proteins, including albumin, immunoglobulins, haptoglobin, and transferrin, make up about 99 % of the plasma proteome (Nanjappa *et al*., 2014), of which albumin alone counts for 55 % (Anderson and Anderson, 2002). The dynamic range of different proteins makes the plasma proteome extremely difficult to study, since for instance, albumin, transferrin, CRP, and α-fetoprotein are present in plasma in concentrations of around 40 mg/l, 3 mg/l, 0.01 mg/l, and 0.00001 mg/l, respectively. Depleting plasma of its most abundant proteins can help in enriching the less abundant proteins, thus improving the resolution of the assay, and decreasing the complexity of the sample. However, depleting plasma from *e.g.* albumin can lead to a loss of other molecules, due to albumin´s broad specificity as a carrier of for instance fatty acids (Ashbrook *et al*., 1975), amino acids (Peters, 1985), unconjugated bilirubin (Brodersen *et al*., 1979), thyroxine (Hoshikawa *et al*., 2004), and many proteins like haptoglobin, apolipoprotein AI, and transferrin (Gundry *et al*., 2007).

Compared to proteomics, blood is a simpler target for genomic approaches, since they are typically based on extracting DNA from white blood cells. However, accumulating amount of biomarker studies are focused on studying circulating cell-free RNA (cfRNA) or circulating tumour DNA (ctDNA) molecules from whole blood (Han *et al*., 2017).

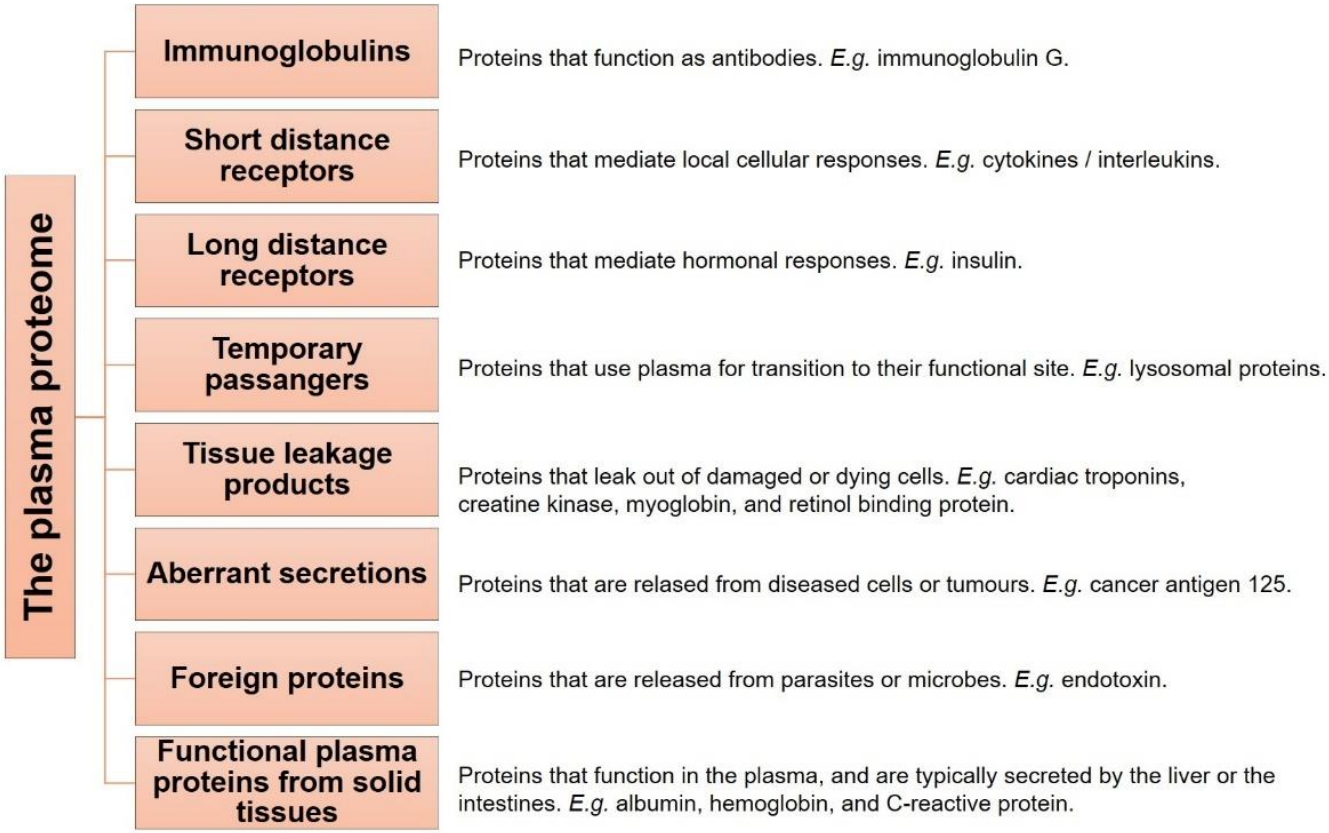


Figure 8 *Functional classes of the plasma proteome.*

2.1.7.2 Sputum

Sputum is a thick, mucous secretion of the lower airways that can consist of *e.g*. mucus, saliva, inflammatory cells, bacteria, proteins, DNA, and even blood. It can be spontaneuously expectorated, but a standardized technique in sputum sampling is induction of sputum by hypertonic saline (Pin *et al*., 1992; Djukanović *et al*., 2002), which improves sample yield and reproducibility. Induced sputum consists of two parts: a protein-rich fluid phase, and a cellular part comprised of *e.g.* epithelial, bacterial, red blood and inflammatory cells (Gharib *et al*., 2011).

Mucins, in particular MUC5AC and MUC5B (Holmen *et al*., 2004), are the major group of glycosylated proteins found in sputum (Thornton *et al*., 2008). Their abundance grows in airway inflammation (Matthews *et al*., 1963), since their gene expression is upregulated by toxins, pathogens, and inflammatory mediators (Rose and Voynow, 2006), such as tumour necrosis factor α (Lora *et al*., 2005), and interleukins 6 and 17 (Chen *et al*., 2003). Pollutants and oxidizing agents are also known to upregulate MUC5AC expression, especially tobacco smoke (Gensch *et al*., 2004), hydrogen peroxide (Takeyama *et al*., 2000), and oil fly ash (Longphre *et al*., 2000).

Biomarker studies on sputum in airway illnesses, such as chronic obstructive pulmonary disease (COPD) (Baraniuk *et al*., 2015), asthma (Gharib *et al*., 2011), cystic fibrosis (Sloane *et al*., 2005), and lung cancer (Cho and Sung, 2009), have been accumulating, since sputum collection is non-invasive, and offers an access to both the cellular incidences and soluble mediators directly from the site of interest. Hence, the consistency of sputum can predict disease risk, outcome or severity. The proteome is typically studied from the fluid phase of the sputum (*e.g.* Suojalehto *et al*., 2015), whereas the genome is extracted from the cellular phase (Cameron *et al*., 2017). For proteomic analyses, mucins set a challenge for sample homogenization, since they tangle up, forming netlike structures (Thornton *et al*., 2008). Dithiotreitol (DTT)

can be used to break the disulphide bonds of mucins, and to solubilize the consistency of the sample. This does, however, lead to reduced abundance of tumour necrosis factor α, leukotriene, and myeloperoxide, while increasing the levels of α-antitrypsin (Woolhouse *et al*, 2002).

2.1.8 *Techniques in biomarker research*

2.1.8.1 Mass spectrometry

Mass spectrometry enables identification of molecules based on their mass-to-charge (m/z) ratio. Molecules are first turned into positively or negatively charged ions in a process called ionization. These ions are then accelerated to gain kinetic energy and pushed into an electromagnet for deflection. The heavier, or the less charged the molecule, the less its projection will deflect. The detector records the flow of ions, and the chart recorder puts out the relative intensity of each ion with a unique m/z ratio.

For samples to enter the mass analyzer, they must be ionized. Due to their large size, proteins must also be fragmented into peptides with proteases, *e.g.* trypsin, for the mass spectrometrical analyses of proteomes to succeed. Volatile, stable samples are transformed to gas phase using gas chromatography (Gohlke and McLafferty, 1993). Large, less volatile samples, such as peptides, are first to go through liquid chromatography (LC) for purification and separation of peptides, after which they are converted into gas phase using *e.g.* electronspray ionization (ESI) (Aleksandrov *et al*., 1984; Yamashita and Fenn, 1984a; Yamashita and Fenn, 1984b). Another option for sample ionization is the soft ionization of biomolecules by matrix-assisted laser desorption/ionization (MALDI). In MALDI, a matrix is used to bind a sample to a sample surface, to ionize the sample, and to protect the sample from the high-energy laser in the desorption/ionization phase (Lai and Wang, 2017).

High analysis power of complex samples is typically achieved by combining mass analyzers, as in tandem mass spectrometry (MS/MS) (Jaiswal *et al*., 2015). The first MS acts as a filter that separates ions, letting only specific, pre-selected ions to go forward to the second MS. In the second MS, each individual ion is further fragmented, giving in-depth information on the molecules under identification. Peptides, for instance, can be fragmented at any bond yielding different fragments depending on the instrument used, although cleavage at the amino bond is the most typical (Wysockia *et al*., 2005). This type of fragmentation results in b or y ions depending on whether the charge is retained in the amino or carboxy terminal of the peptide, respectively.

Time of flight (TOF) mass spectrometry uses a specific type of mass analyzer based on detecting ion´s fly time to the detector, from which m/z ratio of the ion can be derived. MALDI is often, but not necessarily, used to transform the sample into ion form, resulting in MALDI-TOF mass spectrometry. TOF can also be coupled to ESI, where no matrix is needed. For ESI however, other options besides TOF are available as presented in **Figure 9**.

2.1.8.2 MS in genetics

MALDI-TOF was first introduced to studying genetic variants in the late 90´s (Haff and Smirnov, 1997; Rosset *et al*., 1998). Since then it has been used in high-throughput analysis of variants of DNA, mainly SNPs (Bray *et al*., 2001). MALDI-TOF can accurately detect oligonucleotide polymorphisms in a single reaction (Haff and Smirnov, 1997) even from large sample pools, *i.e.* multiplexes (Ross *et al*., 2000; Buetow *et al*., 2001). As a high-throughput genotyping technique, MALDI-TOF is also suitable for complex disease genetics (Werner *et al*., 2002), and since its emergence for genetic studies, it has been widely used for association studies in *e.g.* obesity, cancer, cardiovascular diseases, and diabetes (reviewed in Tost and Gut, 2005).

Beyond those, MALDI-TOF has successfully been implemented also in pharmacogenetics and in diagnostic testing.

Due to its accuracy, robustness and reproducibility compared to other methods (*e.g.* hybridization), MALDI-MS in genomics principally uses the primer extension reaction to detect variations in the genotype from the PCR-amplified DNA. After PCR in primer extension, the primer anneals to the 3´-end of the DNA, upstream and immediately adjacent to the target sequence. The primer is then followed by a terminating dideoxynucleotide in the single base nucleotide extension (Haff and Smirnov, 1997), or by deoxynucleotides and terminating dideoxynucleotides in the multiple base primer extension (Rodi *et al*., 2002). The single base nucleotide extension creates DNA strands with very small mass differences (*e.g.* only 9 Da between A and T) making it hard to resolve the reaction products in mass sprectrometry (Tost and Gut, 2005). The multiple base primer extension creates mass differences of at least 300 Da, allowing for a reliable discrimination between different genotypes, not to mention the possibility to multiplex the extension of four or more different DNA variants into single reaction. After the extension, the reaction products must be thoroughly cleaned for MALDI, since salts and reaction detergents can interfere with the MS. Several methods are available for the purification, most of them using ion exchange, such as resin, or magnetic beads.

2.1.8.3 MS in proteomics

Improved understanding of disease biology has led to a variety of study strategies beyond genomics, one being the analysis of proteins from biological specimen, *i.e.* proteomics. Mass spectrometry offers many different possibilities for both qualitative and quantitative proteomics. MS can be used to detect both steady-state and dynamic changes in the proteome, and it extends beyond protein identification to studying protein abundance, post-translational modifications, localization, and complex formation. These

advances in technology have led to targeted proteomics in the detection of proteins of interest, and to proteomic profiling, where the proteome of a particular specimen at a particular time is analyzed.

MALDI-TOF is well-suited for high-molecular weight protein analysis (up to 200 kDa), and MALDI-TOF is routinely used for microbial protein identification in clinical laboratories (reviewed in Patel, 2015). There it serves as an alternative for the biochemical methods shortening the microbe identification time from days to minutes. Multiple studies have also shown MALDI-TOF to perform at least as accurately as the traditional biochemical methods, and in rare occasions further testing is needed for the determination of microbial species. However, MALDI as an ion source in proteomics is often a worse option than ESI. According to Nadler *et al.* (2017) comparing the impact of MALDI and ESI ionization in different human cells lines of pancreatic origin, MALDI-MS produced worse results in identified peptide and protein uniqueness (18.6 %, 6.6 %), and sequence coverage (14.8 %) compared to ESI-MS (42.2 %, 19.9 %, and 21.7 % respectively). ESI also allows for more variation in the choice of mass analyzer (*e.g.* quadrupole or ion trap) and can easily be coupled to LC. Depending on the required mass accuracy, and study approach, different ESI instruments can be used for protein identification (**Figure 9**) (Han *et al.*, 2008).

There are two general workflow strategies for MS-based proteomics: bottom-up and top-down (**Figure 10**). Bottom-up-strategies are divided into two approaches depending on how and when the proteins are digested and separated into peptides. In the “sort-then-break”-approach proteins are first sorted, then digested and the peptides analyzed by peptide-mass-fingerprinting (PMF). The other alterative is to use the “break-then-sort”- or “shotgun”-approach in which proteins are first digested, and the created peptides sorted before entering MS. The shotgun approach, used in this thesis, has gained popularity in proteomics, since it is relatively simple. However, it creates a complex mixture of peptides resulting in various

problems with protein identification with MS. The top-down-strategy aims at identification of intact proteins, thus involves no digestion into peptides. From intact proteins it is more feasible to *e.g*. study post-translational modifcations. At any point, however, the already studied proteins can be individually digested into peptides and analyzed using PMF.

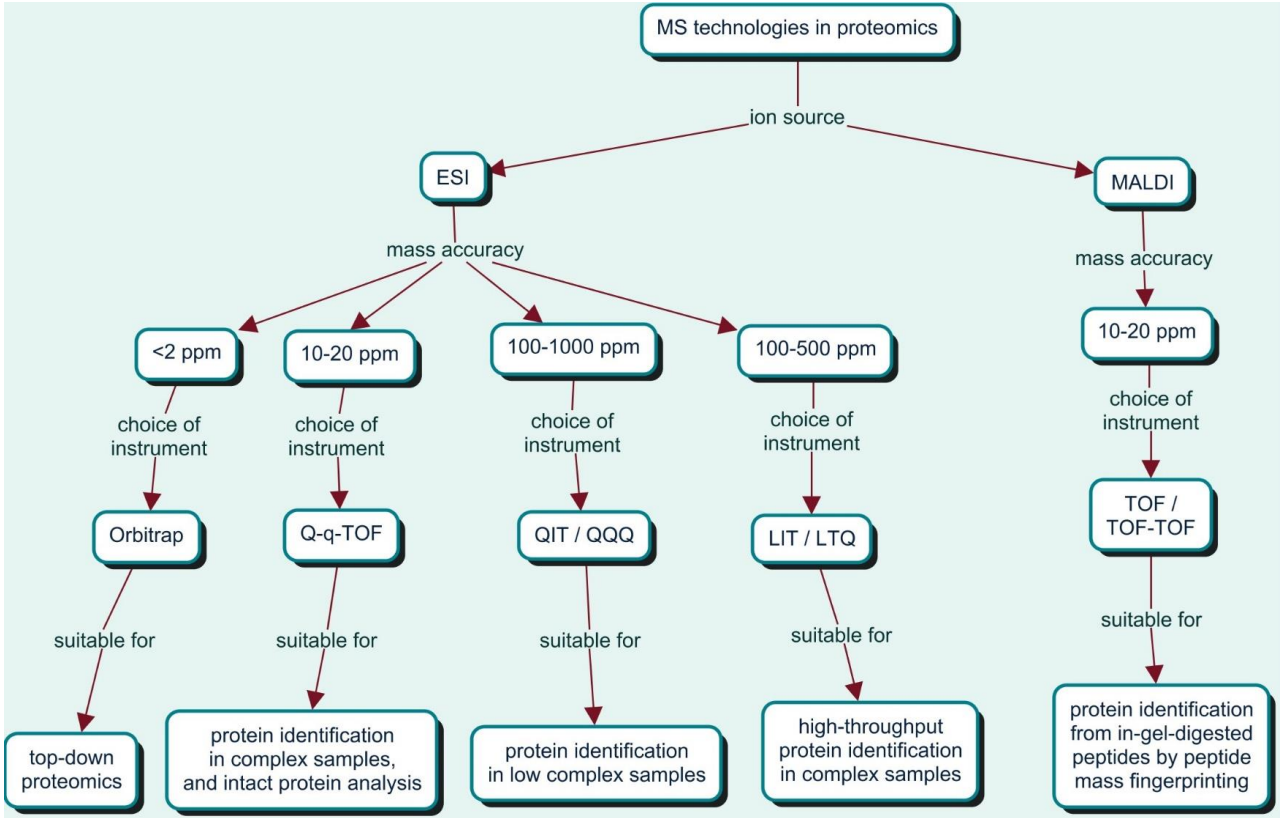


Figure 9 *Mass spectrometry technologies suitable for proteomics. Abbreviations: ESI, electrospray ionization; MALDI, matrix-assisted laser desorption/ionization; Q, quadrupole; TOF, time-of-flight; L, linear; IT, ion trap; TQ, trap quadrupole. Modified from Han et al., 2008.*

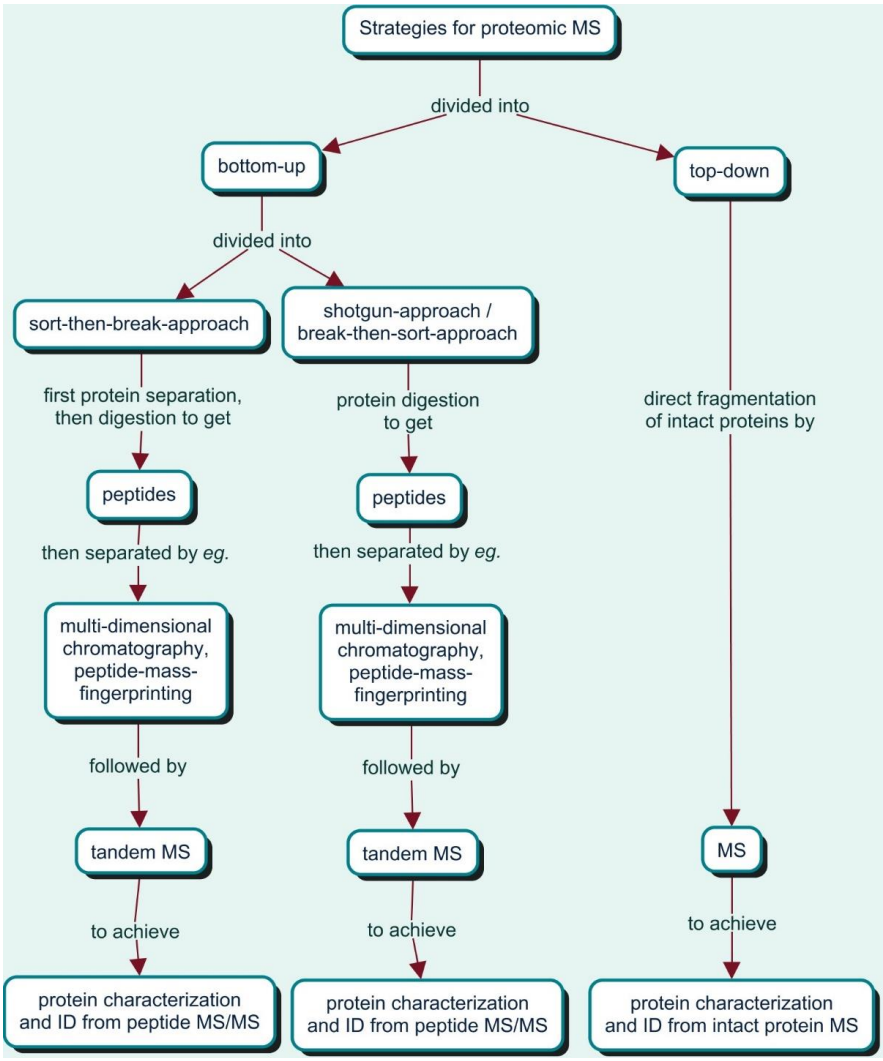


Figure 10 *Bottom-up- and top-down-strategies in mass spectrometry for proteomics.*

2.1.8.4 Gel-based proteomics

Gel-based proteomics is based on creating an electric current to a web-like matrix of the polyacrylamide gel, in which charged molecules migrate between the cathode and the anode, depending on the charge of the molecule. In sodium dodecyl sulfate polyacrylamide gel elctrophoresis (SDS-PAGE), negatively charged SDS molecules are used to bind proteins, and cover their intrinsic charges. Proteins are amphoteric in nature, *i.e*. their chemical structure contains both a negatively charged carboxy-group ($-COO^-$), and a positively charged amino-group ($-NH_3^+$). Once each amino acid has been covered with two negatively charged SDS molecules, all proteins travel inside the gel according to their size and their extrinsic charge (Reynolds and Tanford, 1970), *i.e.* from cathode to anode. SDS is a strong denaturating agent, and reduction of proteins to primary structures is further aided with heating the samples.

Gel-based proteomics can be done in one or two dimensions, depending on the resolution needed for protein separation. One-dimensional SDS-PAGE is solely based on the molecular weight of the proteins in the sample, the smaller proteins travelling a longer distance in the gel than the larger ones. Two-dimensional SDS-PAGE consists of two consecutive gel runs, and takes into account not only the molecular weight, but also the isoelectric point of the protein; hence separates the proteins according to two variables, and results in improved resolution compared to one-dimensional SDS-PAGE.

Isoelectric focusing, *i.e*. the first dimension of two-dimensional gel electrophoresis, is used to separate proteins according to their isoelectric point (pI), in which the net charge of the protein equals to zero. Each protein has a pH value in which its charges balance each other out, and the molecule gains polarity becoming a zwitter ion. Once a protein has reached its isoelectric point on the gel, it will no longer move in the electric current. SDS-PAGE, *i.e*. the second dimension of two-dimensional gel electrophoresis,

then separates the proteins according to their molecular weight. The distance travelled is proportional to the size of the molecule: the smaller the molecule, the longer the distance of travel at a given time (Shapiro *et al*., 1967).

Gel-based proteomics is a low-throughput way of analyzing proteins (Mäbert *et al*., 2014), and the development in MS technology has led to reduced use of the gel-based techniques. Gel-based proteomics offers, however, visualization of the sample proteome, and a way of detecting protein modifications (Rogowska-Wrzesinska *et al*., 2013), such as post-transcriptional modifications, alternative splicing, and truncation (Collier and Muddiman, 2012). Two-dimensional difference gel electrophoresis (2D-DIGE), which uses fluorescent protein labels, allows for simultaneous comparison of three protein samples on the same gel, and improved quantitation of proteins compared to the traditional SDS-PAGE. While MS alone is a high-throughput and effective tool for protein mapping in general, a study by Kim *et al*. (2014) showed its poor capacities in detecting altered gene products, leaving 66 % of the studied protein isoforms undetected.

2.1.8.5 Advantages and limitations of techniques in biomarkers research

Choosing the optimal method for biomarker search is a crucial point on the way to discovery. While MALDI-TOF is preferred in genomics MS, for proteomics options using ESI should be favored. For genomics, gel-based methods and sequencing, are other options in the search for genetic biomarkers, and can also be used for secondary validation. For protein separation and visualization, the gel-based techniques are beneficial in proteomics.

MALDI-TOF in genomics is ideal for SNP genotyping, suitable for GWAS as well as clinical diagnostics (Tost and Gut, 2005). Cost per sample is low, and

the results are considered highly reliable. However, depending on *e.g.* source and quality of DNA, MS might not be the best suitable approach. For instance, a study on DNA mutations in soft tissue sarcomas by Xu *et al.* (2016) compared MALDI-TOF to an ultra-deep sequencing assay. They found that DNA derived from formalin-fixed and paraffin-embedded samples resulted in surprisingly high false discovery rate of mutated DNA in MALDI-TOF. Hence, depending on sample material and study approach, secondary validation should be considered with MS-based genomics.

Study design, instrumentation, and study execution are crucial parts of any research planning, but even more so in proteomics (Drucker and Krapfenbaue, 2013). For instance, depending on the origin of the protein containing sample, the proteins can be difficult to solubilize in sample buffer, or cleaning the sample for MS might be a strenuous task. Sputum, for example, is a challenging sample matrix for MS-based studies, because it requires buffers with high salt and detergent concentrations in order to be soluble. Cleaning sputum from these substances results in low protein solubility and a thick, slime-like sample. Some of these problems can be overcome combining gel-based proteomics with the MS approach, as done in this thesis. However, multiple runs on gels can be expensive, and gel casting and handling requires technically skillful manual work in the laboratory. High abundant proteins easily monopolize the gels hampering the visualization of the entire proteome, especially the low abundant proteins. Thus, validations using other techniques, such as immunoassays, are pivotal.

MS in proteomics requires expensive instrumentation but results in low cost per sample. It is reliable, sensitive and precise, but requires high level of expertise, and is susceptible to breakdowns due to *e.g.* mishandling or sample impurities. However, MS with its coupled techniques is the best available approach to biomarker research.

2.2 *STAPHYLOCOCCUS AUREUS* BACTERAEMIA

2.2.1 *Staphylococcus aureus*

Staphylococcus aureus (SA) is a gram-positive bacterium found in the normal flora of the skin, and mucous membranes (Lowy, 1998). Despite its commensalistic nature, it is a major human pathogen prevailing as the second leading cause of bacteraemia (*Staphylococcus aureus* bacteraemia, SAB) (Laupland *et al*., 2013), and infective endocarditis in the world (Asgeirsson *et al*., 2018). On the contrary to bacterial infections in general, SA can infect its human host repeatedly and persistently throughout life (Grant and Hung, 2013). Despite the introduction of antibiotics to treat SAB, the case fatality rate (CFR) remains at 15-50 % (Van Hal *et al*., 2012). In Finland, 17 % of deaths caused by bloodstream infections are associated with positive blood culture for SA (Skogberg *et al*., 2012).

2.2.2 *Bacteraemia caused by Staphylococcus aureus*

Bacteraemia is a complex inflammatory process involving both humoral and cellular responses, caused by bacteria in the bloodstream. SA is the second most common bacterium found in positive blood cultures (Laupland *et al.* 2013), and therefore a frequent initiator of bacteraemia. The typical infection route for SA is by inoculation from skin to an open wound (Liu, 2009), most likely through nasal carriage, since 20 % of humans are persistently, and 30 % transiently colonized by SA in the nose (Wertheim *et al.*, 2004). The severity and outcome of SAB is significantly affected by underlying diseases, the presence of deep infection foci, and complications in the duration of SAB (Van Hal *et al*., 2012).

SA is a master in hiding from epithelial cells, endothelial cells, and macrophages (Liu 2009; Kubica *et al*., 2008), efficiently blocking the immune system from recognizing it. To aid this, SA expresses a clumping factor A

protein on its surface preventing opsonization, and destruction by host immunity (Foster, 2005). To avoid the neutrophil-secreted antimicrobial substances, such as reactive oxygen and nitrogen species (ROS, RNS), SA produces antioxidant enzymes like catalase and superoxidate dismutase. The Chemotaxis inhibitory protein (CHIP), and the Extracellular adherence protein (Eap) can block neutrophil recognition (De Haas *et al.*, 2004), and inhibit neutrophil binding to the Intercellular adhesion molecule 1 (ICAM-1) (Chavakis *et al.*, 2002) in order to prevent the leukocytes from entering the site on infection. Additionally, SA secreted enterotoxins, the Toxic Shock Syndrome Toxin (TSST) and Eap, impair the T cell receptor activation pathway of the host immunity (Llewelyn and Cohen, 2002). Some strains of SA have also tested positive for an enterotoxin Phenol soluble modulin (PSM), which can cause neutrophil cytolysis (Wang *et al.*, 2007).

2.2.3 *Epidemiology*

In the developed countries, the population incidence of SAB ranges from 10 to 30 per 100 000 person-years (Laupland *et al.* 2013). During 2004-2007 in Finland, about 13 % (n≈4000) of all positive blood cultures from bacteria were caused by SA (Skogberg *et al.*, 2012) This adds up to an annualized incidence of 20 per 100 000 person-years. Risk factors of SAB include *e.g.* infancy (Lyytikäinen *et al.*, 2005), old age, male sex (Skogberg *et al.*, 2012), ethnicity (Van Hal *et al.*, 2012), hemodialysis (Fitzgerald *et al.*, 2011), and HIV infection (Larsen *et al.*, 2012. There are also several conditions, such as diabetes (Zimakoff *et al.*, 1996), neutrophil dysfunction (Vanholder *et al.*, 1991), and iron overload (Boelaert *et al.*, 1990), that impair the immune system, and predispose to SA colonization.

2.2.4 *Diagnosing Staphylococcus aureus bacteraemia*

The presentation of SAB is characterized by typical symptoms of bacteraemia, such as high fever, hypotension, tachycardia, and tachypnea,

whereas severe cases present multiorgan failure, and death. Confirmation of the diagnosis can be obtained from blood culture positive for SA, although studies in bactereremias in general show that 40-60 % of severe bacteraemias or septic shocks remain microbiologically undetermined (De Prost *et al.*, 2013).

Protein biomarkers, such as CRP or PCT, are available for distinguishing between bacterial and viral infections (Jeong *et al.*, 2012), and to monitor the severity of the bacteraemia, but not to diagnose SAB in particular. Pierrakos and Vincent (2010) reported 170 protein biomarkers assessed for their usability in diagnostics and prognostics of bacteraemia, none of which offered either specificity or sensitivity for bacteraemias alone. Most of them were drawn from clinical studies, and offered prognostic, rather than diagnostic value.

2.2.5 *Acute phase response*

Acute phase response (APR) is a complex set of reactions put forward by release of cytokines (Beutler and Cerami, 1986), leading to pathological processes like inflammation, trauma, or infection. APR can generate responses like pyrexia, muscle protein depletion, hormone alteration, and leukocytosis in order for to maintain the homeostasis of the body. Manifestation of bacteria-derived antigens typically cause a strong APR (Alsemgeest, 1994). This is mostly due to the activation of the cells in the mononuclear phagocytic system of the innate immune system. These cells, such as monocytes and macrophages, secrete various cytokines (Van Miert, 1995) to recruit neutrophils to the site of inflammation. Cytokines, in turn, act as the main activators of transcription in genes involved in the immune defense.

2.2.6 *Biomarkers for bacteraemia*

Biomarkers for bacteraemias have been widely researched, due to difficulties in early diagnosis, severity evaluation, and outcome prediction in bacteraemias. Proteins involved in the APR are commonly determined from blood in the suspicion of bacteraemia, and genetic studies have revealed predisposing factors, influencing *e.g.* the blood levels of these proteins, for bacteraemia. Acute phase protein (APP) synthesis in the liver is cytokine-induced (Heinrich *et al.*, 1990; Le and Vilcek, 1989), a process that reaches its peak within 48 hours from the onset of the disease, causing the levels of APPs in blood to rise rapidly in the onset of the disease. Once the acute phase passes, and the disease resolves, APP levels in the blood descend steeply. Cytokines, CRP, PCT, and serum amyloid A (SAA) are the main contributors to APPs.

In recent years, association studies of microRNAs and bacteraemia have also emerged (*e.g.* Wang *et al*., 2014). Currently, good diagnostic biomarkers are not available for bacteraemias, but some biomarkers can help in the prognostic evaluation.

2.2.6.1 White blood cell count

White blood cells consist of granulocytes (neutrophils, eosinophils, basophils and mast cells), lymphocytes (T, B and NK cells), and monocytes, neutrophils representing the most abundant cell type. In bacterial infections, neutrophils of the innate immunity are recruited to migrate from blood into the site of infection in order to fight against the growing invasion of bacteria, mainly via phagocytosis. This requires activation of the bone marrow to produce vast amounts of neutrophils into the blood, to compensate for the increased neutrophil consumption. These changes in neutrophil count dramatically change the total white blood cell count (WCC), which reflects the status of the bacterial infection.

High WCC count does not solely apply to bacterial infections, since increased WCC is also detected in conditions like stress (Ogawa, 1993), hypersensitivity reactions (Denburg, 1992), metabolic diseases (Chabot-Richards and George, 2014), and tobacco smoking. In a study by Seigel *et al.* (2012), 17.4 % of patients with blood culture positive bacteraemia showed normal WCC. In the same study, 21 % of those, who developed septic shock, initially had normal WCC. In children with bacteraemia, WCC alone or together with body temperature, resulted in 57 % of false positive results concurring that WCC is not a sufficient measument for diagnostics or prognostics of bacteraemia (Jaffe and Fleisher, 1991). High WCC together with high CRP suggest infection by gram negative bacteria, yet essentially normal values were observed in those with gram positive bacteraemia (Vandijck *et al.*, 2007).

2.2.6.2 Cytokines

Cytokines have been proposed as biomarkers for bacteraemia, since they are the humoral mediators of adaptive immune response, and their production is tightly related to the onset of the disease. They are secreted by monocytes at the site on inflammation and induce the production of APPs in the liver (Laskin and Pendino, 1995). Cytokines can be divided into three groups according to their roles in acute inflammation (Van Miert, 1995): 1) negative or positive growth factors (interleukins IL-2/3/4/7/10/11/12, granulocyte colony stimulating factor), 2) proinflammatory (tumour necrosis factor TNFα/β, IL-1α/β, IL-6, interferon IFNα/γ, IL-8, macrophage inhibitory protein-1), and 3) anti-inflammatory (IL-1 receptor antagonists, soluble IL-1 receptors, TNFα binding protein, IL-1 binding protein).

Cytokines lack specificity for bacteraemias, since as acute phase proteins, their abundance in blood is not only correlated with antigens of bacterial origin (van Miert, 1995). Their levels also change rapidly and irregularly, which further hampers their usability in surveillance of a disease (Blackburn, 1994). However, according to clinical studies IL-10 associates with bacteremic shock

(Marchant *et al*., 1995), and IL-18 can differentiate between gram-negative and gram-positive bacteraemia (Oberholzer *et al*., 2001). Raised levels of IL-6 (Patel *et al*., 1994) and TNF (Calandra *et al*., 1990) have been linked to survival at 28 days from bacteraemia onset, and IL-8 has been observed in association with multiple organ failure in bacteraemia (El Maghraby *et al*., 2007).

2.2.6.3 C-reactive protein (CRP)

C-reactive protein (CRP) abundance of over 100 mg/l in blood typically indicates bacterial infection (Jeong et al., 2012). Levels of CRP in healthy individuals are normally under 10 mg/l, but these levels reach 350-400 mg/l within two days from the start of the disease (Pepys and Hirschfield, 2003). In severe cases, within 24 hours CRP levels can reach up to 10 000-fold above baseline. Upon the resolution of inflammation, the CRP levels decrease rapidly, with a reported half-life between four to nine hours (Clyne *et al*., 1999) and 19 hours (Vigushin et al., 1993). CRP is synthesized by hepatocytes (Clyne *et al*., 1999), and its production is induced by inflammatory stimuli, especially cytokine release, since the transcription of CRP gene is under the regulative control of IL-6 (Kolb-Bachofen, 1991). The release of IL-1β and TNFα, in turn activate IL-6 gene expression.

CRP activates both humoral and cell-mediated immunity of the adaptive immunune system. In humoral immunity, CRP binds to C1q protein of the complement system, whereafter the CRP-C1q-complex attaches to the Lectin-like oxidized LDL receptor 1 (LOX-1) on the surface of endothelial cells (Fujita *et al*., 2011). This activates the classical pathway of the complement system. In cell-mediated immunity, CRP binds directly to bacterial wall phosphocholine, resulting in bacterium opsonization (Marnell *et al*., 2005). The opsonized bacterium is recognized by leucocyte surface receptors FcγRI and FcγRII. This causes increased phagocytosis, and cytokine production, in particular IL-10, which suppresses IL-6 function in the liver, and

simultaneously transcription of CRP. **Figure 11** summarizes the production of CRP and its main impacts on the immune response.

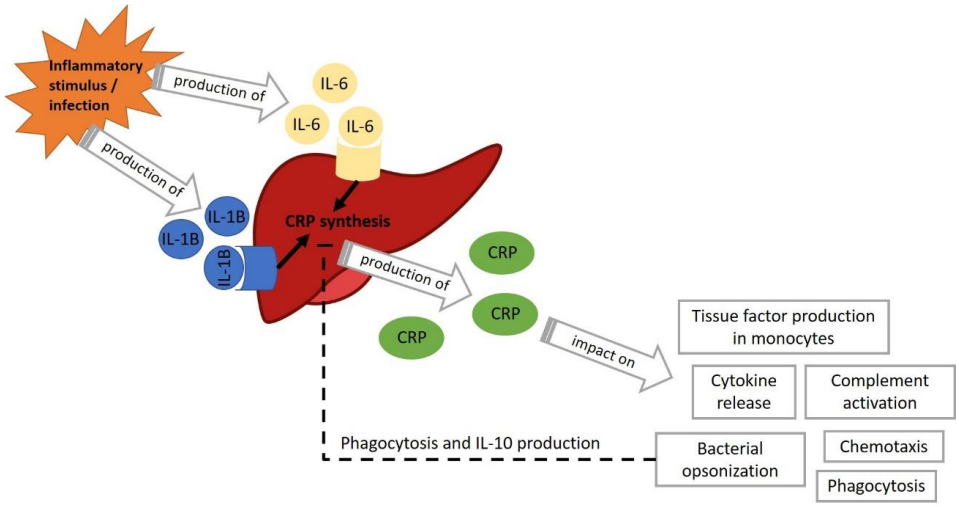


Figure 11 *CRP production and its effects on the immune system.*

The usefulness of CRP in predictive medicine is limited, due to lack of sensitivity and specificity (Clyne *et al*., 1999). In clinical biomarker studies, the value of CRP has been in evaluating the prognosis of the disease, rathen than to diagnose it (Lobo *et al*., 2003). Nonetheless, CRP protein abundance is used in clinical practice to determine risk factors for diseases that have infectious origins (Schmit and Vincent, 2008).

2.2.6.4 Procalcitonin (PCT)

Procalcitonin (PCT), the precursor of calcitonin, is produced in the thyroid gland in response to elevated blood calcium concentrations (Moya *et al.*, 1975). In healthy individuals, virtually all PCT is cleaved into calcitonin, leaving only trace amounts of uncleaved PCT in the blood. However, during severe systemic infections of bacterial origin, the levels of uncleaved PCT can rise up to 5000-fold compared to baseline (Nylen *et al.*, 1992; Benoist *et*

al., 1998; Gilbert, 2010). On the contrary to other acute phase proteins, PCT´s half-life is long, 25-30 hours (Meisner, 2002), resulting in persistent elevation of PCT levels even after the resolution of the acute phase.

PCT is a sensitive biomarker for detecting systemic bacterial infections (Harbarth *et al.*, 2001). Hyperprocalcitoninemia correlates with the severity of bacteraemia and predicts high mortality in bacteraemias of the critically ill. Local bacterial infections have a limited effect on PCT levels (Eberhard *et al.*, 1997), but high PCT levels do associate with other causes of inflammation, such as trauma (Benoist *et al.*, 1998), and surgery (Meisner *et al.*, 1998, Hensel *et al.*, 1998).

2.2.6.5 Serum amyloid A (SAA)

Serum amyloid A (SAA), like CRP, is able to distinguish between bacterial and viral infections. SAA too is produced by the liver, and its production increases during bacterial infection. SAA directly binds to the surface of several different bacteria acting as opsonin for neutrophils and macrophages (Hari-Dass *et al.*, 2005, Shah *et al.*, 2006). However, some studies have reported that SAA also plays a role in viral infection blocking viral entry into host cells (Lavie *et al.*, 2006).

In normal conditions, SAA is found from blood in trace amounts, but levels in blood rise up to 1000-fold compared to baseline within 24 hours from the inflammatory stimulus, such as release of IL-6 or IL-1β (Gabay and Kushner, 1999). Once released from hepatocytes, most of the SAA binds to high-density lipoprotein (HDL) (Coetzee *et al.*, 1986), partly or completely replacing apolipoprotein A1 (Apo-A1) from it (Banka *et al.*, 1995). This lipid-bound SAA modulates the cholesterol metabolism (Liang and Sipe, 1995), and apparently functions in an entirely different way than lipid-free SAA in the blood. According to studies *in vitro*, lipid-free SAA influnces the immunity system by attracting inflammatory cells, such as monocytes (Badolato *et al.*, 1994) and

T-lymphocytes (Xu *et al.*, 1995). SAA can also activate various types of cytokine relasing cells, such as neutrophils (Furlaneto and Campa, 2000), and act as opsonin for Gram-negative bacteria like *Pseudomomas aeruginosa* (Shah *et al.*, 2006).

Studies in neonatal bacteraemias, and early onset bacteraemias show that SAA is a better diagnostic biomarker than *e.g*. CRP, withholding a higher specificity and sensitivity (Cetinkaya *et al*., 2009, Clyne *et al*., 1999). However, there is evidence that hepatocytes are not the only source of SAA, but that tissues like epithelial (Urieli-Shoval *et al.*, 1998) and endothelial cells (Meek *et al.*, 1994), macrophages, and adipocytes, produce it too. Therefore, although tradionally known to be activated by the acute inflammatory stimuli, high levels of SAA are also associated with some chronic diseases, such as the Alzheimer´s disease (Liang *et al.*, 1997), atherosclerosis (Fyfe *et al.*, 1997), rheumatoid arthritis (O'Hara *et al.*, 2000), and cancer (Howard *et al*., 2003).

2.2.7 *Staphylococcus aureus-specific protein biomarkers*

2.2.7.1 Techoid acid antibody (TAA) and antistaphylosin (ASTA)

Two specific biomarkers, teichoid acid antibody (TAA) and antistaphylosin (ASTA), have been used to differentiate between complicated and uncomplicated SAB for decades (Larinkari, 1982). High TAA titers associate with staphylococcal endocarditis (Larinkari *et al.* 1977), deep wound infections, chronic osteomyelitis (Christensson *et al.* 1985) and septic arthritis. TAA is a more useful biomarker in chronic infections, when searching for the cause on infection, than in acute ones (Wise and Tosolini, 1992), since it takes up to two weeks for TAA to reach its peak values (Bayer and Guze, 1979). Though bearing high specificity for SAB, ASTA lacks sensivity, detecting only 32-62 % of deep infections (Larinkari and Valtonen, 1984).

Similar to TAA, it may take up to four weeks for ASTA to reach its peak values, (Christensson *et al.* 1985), limiting its use as an acute phase biomarker.

2.2.7.2 C-reactive protein (CRP)

Blood CRP levels are used in following SAB, since increased levels are known to associate with the severity of the disease (Mölkänen *et al*., 2016). For instance, those with high CRP levels are prone to develop organ dysfunction, whereas rapidly decreasing CRP values in the course of the disease declare improvement (Vigushin *et al*., 1993). There is, however, no clear-cut CRP value limit to aid clinical decision making in SAB (Clyne *et al*., 1999), but high CRP levels (over 100 mg/l) during the first days of infection raise a concern of complications, such as deep infection foci (Mölkänen *et al*., 2016). There are strong genetic components, SNPs in particular, affecting the blood CRP values (Kushner *et al*., 2006; Kluft *et al*., 2003; Kathiresan *et al*., 2006), hence straightforward associations between the CRP values and the severity of SAB cannot be made without knowledge of CRP values at baseline.

2.2.7.3 Soluble urokinase plasminogen activator receptor (suPAR)

There are many potential biomarkers that could be beneficial in diagnosing SAB. High levels of the soluble urokinase plasminogen activator receptor (suPAR) in blood predicts fatality in SAB (Mölkänen *et al.*, 2011). Similar results have been obtained for instance in *Streptococcus pneumoniae* bacteraemia (Wittenhagen *et al.* 2004), HIV (Ostrowski *et al.*, 2005), tuberculosis (Eugen-Olsen *et al.*, 2002), and coronary artery disease (Eugen-Olsen *et al.*, 2010), hence suPAR is neither a specific nor a diagnostic marker for SAB only.

2.2.8 *Genomic biomarkers in SAB*

Several genes have been studied as possible predictive or prognostic biomarkers in bacteraemia. Cytokines, in particular TNF, IL-1, and IL-6, and CRP are among the most important proteins in the pathogenesis of infection, hence their genetic polymorphisms are of interest also in bacteraemia. These genes are of particular interest, because their functions are tightly linked not only in the functional protein level, but also in the gene level; TNFα and IL-1 activate the transcription of IL-6, which in turn activates the transcription of CRP (Ferrari *et al*., 2003). Within the recent decade, association studies between micro-RNAs and bacteraemia have offered insights on new genetic predisposing factors, and possible biomarkers for bacteraemia.

2.2.8.1 Tumour necrosis factor (TNF) and interleukin 1 and 6 (IL-1, IL-6) gene polymorphisms

The tumour necrosis factor (TNF) gene consists of two gene subtypes, TNFα and TNFβ. TNFα is the functional gene that lies between TNFβ and the lymphotoxin β gene (LTβ), from where various polymorphisms and microsatellites have been identified (Ruuls *et al*., 1999). Two distinct polymorphisms of the TNF locus, the NcO1 restriction fragment length polymorphism (RFLP) of the TNFβ gene (Stüber *et al*., 1996), and SNP rs1800629 in the TNFα promoter region (McGuire *et al*., 1994), associate with infectious diseases, albeit with inconclusive results. In Sri Lankans, both mutations were shown to correlate with the manifestation of severe infectious diseases, among others malaria (Wattavidanage *et al.*, 1999), and another study on West African children, homozygocity for SNP rs1800629 A-allele associated with cerebral malaria (McGuire *et al.*, 1994). Stüber *et al.* (1996) found that the RFLP correlated with poor prognosis in severe bacteraemia, and with severe posttraumatic bacteraemia in blunt trauma patients (Majetschak *et al.*, 1999), but observed no influence of the rs1800629 on the survival (Stüber *et al.*, 1995-1996). Fang *et a*l (1999). found that the RFLP

associated with nonsurvival in those with severe bacteraemia. Nevertheless, in a study of more than 200 Caucasians with bacteraemia, no associations were found between either of the two polymorphisms and severe bacteraemia or septic shock (Gordon *et al.*, 2004).

Four SNPs of the IL-1 gene (rs1800587, rs16944, rs143634, rs1143627), a variable number tandem repeat (VNTR) of the IL-1 receptor antagonist (IL-1RN), and two SNPs of the IL-6 gene (rs1800796 and rs1800795) have been researched in association with bacteraemia, albeit with inconsistent results. The IL-1 gene SNPs rs1800587 (Gu *et al.*, 2010; Davis *et al.*, 2010), rs16944 and rs143627 (Gu *et al.*, 2010; Davis *et al.*, 2010; Shimada *et al.*, 2011) show no association with bacteraemias, although SNPs rs16944 and rs143627 influence the IL-1β levels *in vitro* under LPS stimulation (Wen *et al.*, 2006). The T allele of SNP rs143634 (Zhang *et al.*, 2005), and the VNTR polymorphism (Ma *et al.*, 2002) both associate with risk of severe bacteraemia or septic shock. *In vitro* the rs143634 T allele (Pociot *et al.*, 1992), and VNTR IL-1RN-2 allele (Santtila *et al.*, 1998) have elevated IL-1β levels under LPS stimulation compared to other alleles. IL-6 gene polymorphisms have been less studied than those in IL-1 gene. Lorente *et al.* (2016) found that those with rs1800795 CC genotype lower 30-day mortality in bacteraemia, and lower blood levels of IL-6 than other genotypes in the same locus. SNPs rs1800796 C allele is shown to increase levels of circulating IL-6, and simultaneously induce CRP gene expression (Ferrari *et al.*, 2003).

2.2.8.2 CRP gene polymorphisms

The CRP gene is located on chromosome 1, where its two exons encode for a 204 amino acid protein. The gene and its flanking areas include about 40 SNPs (Crawfod *et al.*, 2006), of which 6 were genotyped for his thesis. An estimated 35-40 % of interindividual variation in blood CRP levels is heritable (Pankow *et al.*, 2001). Obesity is a major contributor to elevated CRP levels,

which indicate the risk of developing the type 2 diabetes mellitus, and the metabolic syndrome (Laaksonen *et al*., 2004). Adipocytes are producers of IL-6, which stimulates CRP production in hepatocytes. Indeed, losing weight, and exercising lower the blood CRP value (Esposito *et al*., 2003), but the association between CRP and insulin resistance is independent of the body mass index. CRP gene SNP rs2794521 TT-genotype is associated with the type 2 diabetes mellitus but having the C-allele in the same locus improves training-induced insulin sensitivity significantly.

Aggregated CRP is present in virtually all atherosclerotic plaques, as it binds selectively to low-density lipoproteins (LDL), and very low-density lipoproteins (VLDL) (Ridker *et al*., 2007). After myocardial infarction, high peak levels of CRP at 48 hours predict increased mortality within six months of the event. Genetic studies in cardiovascular diseases have shown that polymorphisms in the CRP gene affect blood CRP levels, and predispose to atherothrombotic events (Lange *et al*., 2006). In periodontal inflammation patients, the TT genotype of the triallelic SNP rs1130864 in the 3´untranslated region of the CRP gene, caused 64 % higher CRP levels in blood compared to those with the CC genotype (Marsik *et al*., 2006). In another study, the same TT genotype was associated with high CRP levels at baseline, after training, and after coronary artery bypass graft, as compared to other genotypes (Brull *et al*., 2003). Homozygocity for SNP rs2794521 A-allele, and for SNP rs3093066 G-allele lead to high CRP levels in blood both at baseline and after exercise (Obisesan *et al*., 2004). In an Italian cohort, C-allele in the CRP gene SNP rs1800947 was associated with high blood CRP values, and was more often found in sufferers of myocardial infarction than in healthy controls (Balisteri *et al*., 2006). In an American study on healthy men, the C-allele had no correlation with artelial thrombosis (Zee and Ridker, 2002), nor did the British women´s health study report carrying the C-allele had any effect on blood pressure, pulse pressure, or hypertension (Davey *et al*., 2005). The A-allele in SNP rs2794521 is associated with high risk for coronary artery disease, myocardial infarction, and thromboembolic stroke (Miller *et al*., 2005). In a

study on Han Chinese, the A-allele raised the odds ratio for coronary heart disease to 6.8, compared to healthy controls (Chen *et al*., 2005).

To summarize, CRP gene polymorphisms evidently influence the blood CRP levels both in the healthy and in the diseased. SNPs that cause high blood CRP levels predispose to cardiovascular diseases, and lipid metabolism-related diseases, although results have not been indisputable. Some large-scale CRP tag SNP studies, such as the Cardiovascular Health Study (Lange *et al*., 2006), have, however, given evidence that CRP gene polymorphisms influence both blood CRP levels, and disease risk. **Figure 12** summarizes the CRP gene structure and represents the locations of both the referred and studied SNPs.

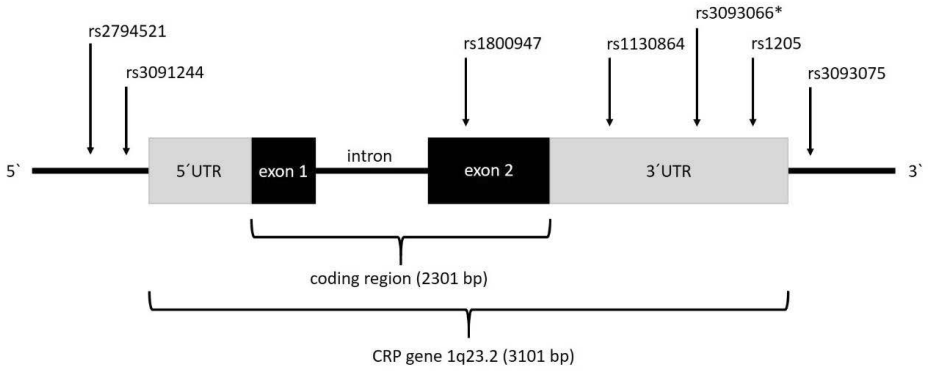


Figure 12 *The CRP gene. The rs-numbers guide to the SNPs studied or referred to in this thesis. UTR, untranslated region; bp, base pair. *not genotyped for this thesis.*

2.2.8.3 Association of micro-RNAs (miRNAs) with bacteraemia

Serum micro-RNAs (miRNAs) are emerging as the new biomarkers for bacteraemia. Vasulescu *et al.* (2009) identified the first miRNA in association with bacteraemia, miR-150. It was downregulated in those with bacteraemia, and in negative correlation with blood levels of TNFα, IL-10, and IL-18.

Furthermore, they observed that the ratio between miR-150 and IL-18 can be used as a prognostic biomarker for the severity of bacteraemia.

Other miRNAs, such as miR-15a, miR-16, miR-15b, miR-146a, and miR-223, have later been associated with bacteraemia. Wang *et al.* (2010) showed that blood levels of miR-146a and miR-223 were able to differentiate between bacteraemia and systemic inflammatory response syndrome (SIRS), being significantly lower in those with bacteraemia than in those with SIRS or healthy controls. Studies have shown associations with miR-15a and miR-16 in both neonatal and adult bacteraemia. In adults, high levels of miR-15a, and low levels of miR-16 are risk factors for death from bacteraemia (Wang *et al.*, 2012), whereas in neonates, high levels of miR-15a and miR-16 in the blood predict bacteraemia (Wang *et al.*, 2015). It was also shown in neonates that upregulated miR-15a and miR-16 downregulate the bacterial LPS/TLR4-signal transduction pathway, and hence work as prognostic biomarkers in neonatal bacteraemia.

To summarize, miR-146 and miR-223 can serve as diagnostic biomarkers for bacteraemia, whereas miR-150 is more of a prognostic biomarker. In adults miR-15a and miR-16 correlate with prognosis of bacteraemia, but in neonates these miRNAs can be used for diagnostic purposes.

2.2.9 *Treatment of Staphylococcus aureus bacteraemia*

Treatment of SAB is always based on a minimum of two-week administration of intravenous antibiotics, typically an antistaphylococcal (semi-synthetic) penicillin. Combinatory treatment, with *e.g.* fluoroquinolones, has led to good results in experimental studies, but clinical trials have given diversified results. In a Finnish study, Ruotsalainen *et al.* (2006) reported that adding fluoroquinolone to the standard treatment of SAB showed no improvement in patient outcome. Similar results have been obtained with adjunctive aminoglycosides.

2.3 LUNG CANCER

2.3.1 *Epidemiology of lung cancer*

Lung cancer is the leading cause of cancer-related mortality in the world, annually accounting for more than 1.5 million deaths (Fitzmaurice *et al*., 2013). In Finland in 2017, lung cancer accounted for 9.6 % and 6.1 % of all cancers in men and women, respectively (Suomen Syöpärekisteri, 2019). Only about 14 % of patients with lung cancer will survive for more than 5 years after the diagnosis. In Finnish men 22.0 % and in women 13.1 % of all cancer related deaths are due to lung cancer.

While cancers in general share many common risk factors, such as ageing (Zhang *et al*., 2017), irradiation (World Health Organization, 2008), unhealthy diet (World Health Organization, 2008; Renehan *et al*., 2008), and physical inactiveness (Ballard-Barbaoh *et al*., 2012) tobacco smoking is the predominant cause of lung-related carcinomas (IARC, 2004). Other risk factors include *e.g.* exposure to asbestos (Brown *et al*., 2012), radon (IARC, 2012), polycyclic hydrocarbons (Baan *et al*., 2009), and second-hand smoking (World Health Organization, 2008). Tobacco smoking is a factor in estimated 90% of lung cancers and causes all types of lung cancers. Smoking also plays a role in the formation of many other cancers, such as cancers of the oral and nasal cavity (Hecht, 2003), stomach (Brenner *et al*., 2009), bladder (Sanli *et al*., 2017), and kidney (Tahbaz *et al*., 2018). The younger a person starts to smoke, the more cigarettes they smoke per day, and the longer the smoking continues, the greater the risk of contracting lung cancer (Hakulinen and Pukkala, 1981). In Finland, second-hand smoking causes on average 10 to 50 new lung cancer cases per year (Pukkala and Rautalahti, 2013).

Tobacco smoking and exposure to asbestos contribute synergistically to the lung cancer risk (Selikoff *et al*., 1968; Vainio and Boffetta, 1994), meaning

that the combined effect of tobacco smoke and asbestos in risk for lung cancer is greater than the sum of these factors individually. Selikoff *et al*. (1968) showed that asbestos exposure alone increases the risk of lung cancer by fivefold, smoking alone by 10-fold, and both exposures by 50-fold, compared to the unexposed. A recent meta-analysis on the synergistic relationship between tobacco smoking and asbestos exposure in lung cancer by Klebe *et al*. (2020) claims that taking into account all lung cancer risk factors, a smoker´s actual average risk for lung cancer is 10 in the presence of tobacco smoke, 4 in the presence of asbestos exposure, and 38 in the presence of both risk factors. Thus, exposure to either risk factor alone is capable of causing lung cancer, but when both are present, they promote carcinogenesis in a multiplicative manner.

2.3.2 *Carcinogenicity of tobacco smoking*

Tobacco smoking is associated with 17 different cancers, and accounts for more than six million deaths annually (Secretan et al., 2009). Tobacco smoke contains more than 7000 different compounds (Rodgman and Perfetti, 2008), of which at least 60 are carcinogenic (Hecht, 2003). Polycyclic aromatic hydrocarbons (PAHs), nitrosamines and aromatic amines are among the strong carcinogens, of which benzo(a)pyrene (BaP) was the first to be detected (Cooper *et al.*, 1954). Oxidants, such as PAHs, form a high proportion of tobacco carcinogens (Church and Pryor, 1985). Whether in the gas or tar phase, oxidants react rapidly to form ROS and RNS, such as hydrogen peroxide and peroxynitrate (Pryor and Stone, 1993), which have a high reactive affinity for DNA, lipids, and proteins, and the capacity to provoke DNA damage, mitochondrial malfunction, cell membrane damage, cell proliferation and apoptosis (Sies, 1997). However, PAHs as such might not provoke DNA damage (Alexandrov *et al*., 2010). After entering the pulmonary cells, PAHs can act as procarcinogens that are later on transformed into carcinogens by metabolic enzymes involved in the CYP1A1/1B1 and epoxide hydrolase pathway, CYP peroxidase pathway, or the aldo-keto reductases

pathway. Changes caused to the integrity of the genetic material, and cell cycle, upset the normal growth and development of the cell, making the cell prone to tumourigenesis.

Continuous exposure to tobacco smoke affects both innate and adaptive immunity by provoking and inhibiting a multitude of responses, both pro-inflammatory and immunosuppressive in nature (Arnson *et al*., 2010). **Figure 13** presents some of these responses. Since alveolar macrophages serve as the first line of defence against pollutants in the airways, they are also the ones to engulf and withhold particles found in the tobacco smoke (Skold *et al*., 1996). These particles impair macrophage functions consequently leading to deficiencies in their bacterial killing abilities (Takeuchi *et al*., 2001). Nizri *et al*. (2009) showed that nicotine retains inhibitory functions, such as decrease in macrophage-released pro-inflammatory cytokines, in particular TNFα, IL-1β and IL-6, while other particles in tobacco smoke have converse effects (Bermudez *et al*., 2002). Macrophages with both pro- and anti-inflammatory functions are associated with *e.g.* pancreatic cancer (Helm *et al*., 2014).

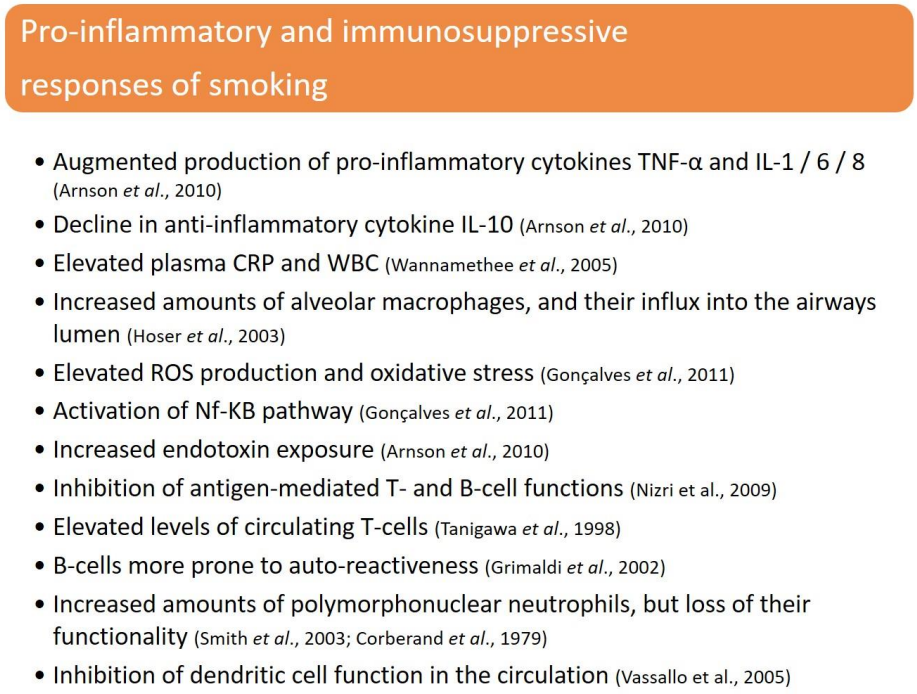


Figure 13 *Effects of tobacco smoke on the immune system.*

A variety of genetic alterations in cancers are associated with tobacco smoking. Base substitutions, and small insertions and deletions are more common in smokers than in non-smokers, and copy-number aberrations in lung adenocarcinomas from smokers are more frequent than in those from non-smokers (Alexandrov *et al.*, 2016). Tumour suppressor protein gene 53 (TP53) mutations are more common in lung cancers from smokers than from never-smokers (Husgafvel-Pursiainen *et al.*, 2000), and guanine to thymine transversions have been associated with smoking in multiple codons of the TP53 gene (Pfeifer *et al.*, 2002) experimentally shown prone to form DNA adducts with BaP metabolites found in tobacco smoke (Denissenko *et al.*, 1996). Proto-oncogene KRAS codon 12 mutatations in adenocarcinomas (Husgafvel-Pursiainen *et al.*, 1993), and loss of heterozygocity at

chromosomes 3p and 9q (Wistuba *et al.*, 1997) are also found more frequently in association with smoking.

2.3.3 *Carcinogenicity of asbestos*

Asbestos is a naturally occurring, solid silicate fiber known to act as a carcinogen by causing DNA damage, such as chromosomal aberrations, and by generating free radicals, such as hydrogen peroxide (Liu *et al*., 2013). Exposure to asbestos is typically occupational accounting for more than half of all occupational cancers, about 5-7 % of lung cancers, and withholding its place as the primary cause of malignant mesothelioma. Besides cancer, asbestos exposure can lead to an asbestos-induced pulmonary fibrosis called asbestosis, and pleural abnormalities, such as pleural plaques *i.e.* the fibrosis and calcification of the parietal pleura. Although the use of asbestos ceased in many countries decades ago, the long latency period of 15 to 40 years between the toxic exposure and the manifestation of the disease will result in occurrence of asbestos-related diseases long in the future. Also, countries like Canada, Russia, and China, continue to sustain their asbestos mines, and export asbestos to developing countries. In Finland, on estimation 200 000 people have been exposed to asbestos at work in the past, and 500 new cases of asbestos-related diseases emerge annually (Huuskonen and Rantanen, 2006), although asbestos has been banned in Finland since 1992 (Huuskonen *et al*., 1995).

Asbestos comes in two different forms. Amphibole fibers, *i.e.* crocidolite (blue asbestos), amosite (brown asbestos), tremolite, anthophyllite, and actinolite, are needle-shaped, sharp and rigid, whereas serpentine fibers, chrysotile (white asbestos), are curly, flexible and fragile (LaDou, 2004). Majority of the commercially used asbestos were chrysotile, crocidolite, and amosite fibers (Huuskonen and Rantanen, 2006), of which crocidolite is considered the most dangerous one (Heintz *et al*., 2010). Anthophyllite was mined in Finland until 1976, hence it was among the most used fiber types in Finland.

The pathogenicity of the fiber is dependent on its thinness, length, and biopersistence (Mossman et al., 1998). Fiber thinness is related to the ability of the fiber to penetrate beyond the ciliated cells in the airways: the thinner the fiber, the better it escapes the clearance of the cilia. Biopersistency refers to the solubility of the fiber; soluble serpentine fibers break down, and lose their shape within months, making them less harmful than the insoluble amphibole fibers that retain lungs for years (Boulanger *et al*., 2014). Long fibers penetrate and deposite beyond the ciliated airways better than short fibers. Schinwald *et al*. (2012) showed that fibers < 5 μm in length become completely phagocytosed by macrophages, and cleared from the lungs causing no inflammation, whereas fibers 5 μm in lenght are phagocytosed, but induce stress, and lead to inflammation. Fibers longer than 5 μm bring about a process called frustrated phagocytosis, where the fiber cannot fit inside the cell causing the cell to release its lysosomal content on cell surface. This results in inflammation. Fiber length of 5 μm serves as threshold value for the release of proinflammatory mediators, such as ROS and cytokines, from the mesothelial cells (Murphy *et al*., 2012), which, if prolonged, can lead to tumour growth (Mutsaers, 2002). Since tobacco smoking impairs the clearance capacity of the airways, pulmonary toxicity can be increased by the exposure to both asbestos and tobacco smoke (Nelson and Kelsey, 2002).

The magnitude and dosage of asbestos exposure plays a crucial role in triggering fiber-related inflammation. High dose over a short period of time activates neutrophil-mediated acute inflammation, whereas prolonged low dosages activate the macrophage-mediated chronic inflammation (Liu *et al*., 2013). Once macrophages phagocytose asbestos fibers, they provoke a fibrogenic response from the fibroblasts (Mossman *et al*., 2011). This response is mediated by *e.g.* cytokines, like TNFα and IL-1β, which promote collagen formation, and lead to asbestosis. Chronic inflammation is also known to evoke the formation of lung cancer (Balkwill and Mantovani, 2001), whose risk is directly correlated with the amount of asbestos fibers in the

lungs (Reid *et al.*, 2005). Asbestos exposed macrophages also release ROS, in particular hydrogen peroxide, after frustrated phagocytosis creating oxidative stress. In mice, Dostert *et al*. (2007) showed the ROS contribute to Nalp3 inflammasome activation, and IL-1β-mediated inflammation. Inflammasomes, of which Napl3 is the best characterized, are intracellular protein complexes that are activated *e.g.* upon stress or microbial molecules. In humans, activation of the Nalp3 inflammasome causes malignant mesothelioma (Wang *et al*., 2004), and promotes tumour growth and invasiveness (Krelin *et al*., 2007).

2.3.4 *Classification of lung carcinomas*

Classifying lung carcinomas into different subtypes is important from a therapeutical perspective, especially in targeted treatment approaches. Histologic and genetic changes observed in the tumor form the base of therapeutic decision making. Previously, lung tumours were separated to non-small cell lung cancers (NSCLC), and small cell lung cancers (SCLC). Non-small cell lung cancers were further divided into three subgroups: adenocarcinomas, squamous cell lung carcinomas, and large cell lung carcinomas. According to the latest criteria (Travis *et al*., 2015), adenocarcinomas, squamous cell lung carcinomas, large cell lung carcinomas, and SCLC all belong to tumors of the epithelia. The last two belong to the subgroup of neuroendocrine lung tumors. For clarity, this thesis uses NSCLC and SCLC when referring to different lung cancer subtypes.

2.3.4.1 Adenocarcinomas, squamous cell carcinomas and large cell carcinomas

Adenocarcinomas account for 40 %, squamous cell carcinomas for 20 %, and large cell carcinomas for 5 % of all lung cancers (Dela Cruz *et al*., 2011). Adenocarcinomas and large cell carcinomas typically grow in the periphery of the lung, *i.e.* in the smaller airways of the lung, whereas squamous cell

carcinomas often invade the central parts of the lung, the bronchi. Most typical lung cancer type in non-smokers and women is adenocarcinoma, whereas smokers typically manifest squamous cell carcinomas (Khuder, 2001). Large cell carcinomas are more common in men than in women.

2.3.4.2 Small-cell lung cancers

SCLCs account for approximately 15 % of lung cancers (American Cancer Society, 2015), and typically start growing from the bronchi. They metastasize more aggressively than NSCLCs (Matthews *et al*., 1973), and are more common in heavy smokers than non-smokers (Alexandrov *et al*., 2016). The five-year survival rate in SCLCs is only 7 % (Rudin and Poirier, 2017), and two-year survivals often develop a second primary tumour (Kawahara *et al*., 1998).

2.3.4.3 Protein markers for lung cancer diagnosis

Established genetic factors that maintain and drive lung tumourigenesis have enabled a more accurate lung cancer classification into subtypes according to their molecular composition. Specific protein expression, such as the thyroid transcription factor-1 (TTF-1) for adenocarcinoma (Travis *et al*., 2011), can help to distinguish lung carcinoma histotypes that lack otherwise histopathologically evident indications of carcinoma subtype.

Surfactants are surface-active proteins that reduce surface tension. This is important especially in the lungs, where water and gas form an interface, and airways need to stay inflated at all times. Pulmonary surfactants are produced by alveolar cells, and consist 90 % of lipids, and 10 % of proteins, mainly surfactant proteins SP A-D (Sunde *et al*., 2017). These proteins interact with other surfactants, such as TTF-1 and napsin A.

SP-A and TTF-1 differentiate AC from other NSCLCs, albeit TTF-1 is a more sensitive marker being abundant in 75-85 % of ACs (Bejarano *et al*., 1996). SCCs are typically negative for TTF-1, but in 80-90 % of SCLCs TTF-1 is abundant (Ordonez, 2000). A panel of cytokeratin 7 (CK7), TTF-1, napsin A, and mucin is most commonly used for diagnostic purposes of AC (Ao *et al*., 2014). In fine needle aspirations, napsin A and TTF-1 were found to discriminate ACs well from other types of NSCLCs, embodying sensitivity of 87 % and 74 %, and specificity of 96 % and 88 %, respectively (Fatima *et al*., 2011; Johnson *et al*., 2012). In a microarray study on lung tissue, TTF-1 was shown to be useful in segregating SQCs from other NSCLCs exhibiting 93 % sensitivity, and 92 % specificity (Brown *et al*., 2013). In a lung tissue study combination of TTF1, napsin A, and CK7 abundance was found a specific and sensitive diagnostic biomarker of AC (Ma *et al*., 2015). Combination of TTF-1, napsin, and protein p40, a marker for SCC, was able to distinguish between ACs and SCCs in a lung tissue microarray study (Ao *et al*., 2014). The triple protein panel was shown to embody a higher sensitivity and specificity for ACs than any of the three proteins individually. Li *et al*. (2013) compared the BAL proteome from individuals with ACs to those with benign lung disease and found that napsin A was abundant in the AC proteome, embodying a sensitivity of 84 %, and specificity of 67 %. Turner *et al*. (2012) compared the sensitivity and specificity of TTF-1 to napsin A in ACs, SQCs, SCLCs, and cancers of other organs, and found that napsin A was superior in detecting ACs from other lung carcinomas, and from carcinomas of any other organ excluding the kidney. Napsin A was also abundant in epithelial lining fluid (ELF) from those with AC as compared to those with other type of lung cancer (Uchida et al., 2017).

To summarize, it is evident that TTF-1 can be used as a diagnostic biomarker for lung ACs, but combining it with other proteins, such as SP-A or napsin A, offers improved specificity and sensitivity irrespective of sample type.

2.3.4.4 Staging of lung carcinomas

Cancer staging is an important part of cancer diagnosis and treatment. Its purpose is to aid therapeutic decision making by classifying the tumor by its anatomical extent. Together with classification into cancer subtypes, and details on *e.g.* patient age, gender and symptoms, staging helps to predict the prognosis, and to design cancer therapy.

The tumor-node-metastasis- (TNM) classification (IASLC, 2016) is based on the assessment of three components: T) the extent of the primary tumour (**Table 2**), N) the absence or presence and extent of regional lymph node metastasis (**Table 3**), and M) the absence or presence of distant metastasis (**Table 4**). The inclusion of numbers to these three components, T0, T1, T2, T3, T4, N0, N1, N2, N3, M0, and M1, give information on malignancy, zero being the most benign alternative in all cases. The combination of the TNM-classes gives cancer staging (**Table 5**).

Table 2. *Tumor classes in lung tumors*

Tumor class	Description
Tx	Primary tumour cannot be assessed, or tumour proven by the presence ofmalignant cells in sputum or bronchial washings but not visualized by imagingor bronchoscopy
T0	No evidence of primary tumour
Tis	Carcinoma *in situ*
T1	Tumour 3 cm or less in greatest dimension, surrounded by lung or visceralpleura, without bronchoscopic evidence of invasion more proximal than the lobar bronchus
T2	Tumour more than 3 cm but not more than 5 cm that does not involve any of the following features: involves main bronchus regardless of distance to the carina, but without involvement of the carina; Invades visceral pleura; Associated with atelectasis or obstructive pneumonitis that extends to the hilar region either involving part of or the entire lung
T3	Tumour more than 5 cm but not more than 7 cm in greatest dimension orone that directly invades any of the following: parietal pleura, chest wall (including superior sulcus tumours), phrenic nerve, parietal pericardium; or separate tumour nodule(s) in the same lobe as the primary
T4	Tumour more than 7 cm or of any size that invades any of the following: diaphragm, mediastinum, heart, great vessels, trachea, recurrent laryngealnerve, oesophagus, vertebral body, carina; separate tumour nodule(s) in adifferent ipsilateral lobe to that of the primary

Table 3. *Nodus classes in lung tumors*

Nodus class	Description
Nx	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in ipsilateral peribronchial and/or ipsilateral hilar lymph nodesand intrapulmonary nodes, including involvement by direct extension
N2	Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s)
N3	Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular lymph node(s)

Table 4. *Metastasis classes in lung tumors*

Metastasis class	Description
M0	No distant metastasis
M1	Distant metastasis

Table 5. *Lung cancer staging*

Stage	T	N	M
occult carcinoma	Tx	N0	M0
IA	T1	N0	M0
IB	T2	N0	M0
IIA	T2	N0	M0
IIB	T1-T3	N0, N1	M0
IIIA	T1-T4	N0-N2	M0
IIIB	T1-T4	N2, N3	M0
IIIC	T3, T4	N3	M0
IV	any T	any N	M1

2.3.5 ***Lung cancer screening***

Lung cancer comes with known, and often avoidable risk factors, but predicting lung cancer prevalence in high-risk groups is difficult mostly due to ineffective screening methods. Computer tomography (CT) scanning and x-raying offer traditional and visual approaches to lung cancer screening. In a large American study by Aberle *et al*. (2011), CT scanning was shown to be more efficient than x-raying in cancer screening, and able to reduce the death rate of lung cancer. However, multiple other studies conducted in Europe (Infante *et al*., 2015; Blanchon *et al*., 2007; Wille *et al*., 2016; Lopes *et al*., 2013; Horeweg *et al*., 2014) have not been able to show the benefits of CT scanning in the reduction of lung cancer mortality. Mansner *et al*. (2013) analyzed nine lung cancer screening trials with 453 965 individuals at high-risk of contracting lung cancer. None of the nine studies using chest x-ray, sputum cytology, a CT scan, or a combination of these, were able to show a statistically relevant reduction in lung cancer mortality.

Lung cancer screening is further hampered by indeterminate symptoms, and a relatively long lag time from the exposure to carcinogenic substances to the occurrence of the disease. Smokers can experience typical symptoms of lung cancer, such as cough, hoarseness, shortness of breath, and chest pain, on a daily basis, thus symptoms of cancer are easily overlooked. Also, in tobacco smoking, and asbestos exposure the incidence of lung cancer has a lag time of 10-50 years (Klebe *et al.*, 2020), making it difficult to follow the beginning of the disease. According to the Helsinki Criteria (Henderson and Leigh, 2011), lung cancer can be attributed to exposure to asbestos with a minimum latency period of 10 years. However, depending on smoking history, type of asbestos fibers, occupational history, magnitude of exposure, and genetic susceptibility, amendments can be applied.

2.3.6 *Biomarkers in lung cancer*

Low survival rate in all lung cancer subtypes can be explained by inadequancy in both screening and early diagnosis, and the growing resistance of tumours to anticancer drugs (Leon *et al.*, 2016). Predictive, diagnostic, and prognostic biomarkers could thus revolutionize lung cancer screening, detection, and treatment by offering information on the disease risk, severity of the disease, response to treatment, and survival.

2.3.7 *Genetic drivers of lung cancer*

Lung cancer has been extensively studied with regards to genetic aberrations. The best-known gene mutations in NSCLC occur in genes EFGR, KRAS, anaplastic lymphoma kinase (ALK), MET proto-oncogene (MET), human epidermal growth factor receptor 2 (HER2), proto-oncogene tyrosine-protein kinase ROS (ROS1), v-Raf murine sarcoma viral oncogene homolog B (BRAF), and RET proto-oncogene (RET), which account for more than half of known genetic aberrations in lung adenocarcinomas (Hirsch *et al.*, 2017). They are all molecular targets for cancer therapy.

Genetic aberrations in lung cancer caused by asbestos exposure are hard to distinguish, since tobacco smoking and asbestos exposure typically coincide (Nelson *et al*., 2002). Chromosome mutations have been detected in chromosomes 1 (Dopp *et al*., 1997; Lohani *et al*., 2002), 3 (Marsit *et al*., 2004), 9 (Dopp *et al*., 1997; Andujar *et al*., 2010) and 19 (Ruosaari *et al*., 2008), including chromosome breaks, deletions, and loss of heterozygosity (LOH). Asbestos-induced copy number variations have been detected in chromosomes 9 (Nymark *et al*., 2006; Nymark *et al*., 2009), 2 (Nymark *et al*., 2006; Kettunen *et al*., 2009), 5, 11, and 19 (Nymark *et al*., 2006). However, some of these changes, such as deletion in chromosome 9 (Tam *et al*., 2013), have also been reported in association with non-asbestos related lung cancer in smokers.

2.3.8 ***Proteomic markers of lung cancer***

2.3.8.1 Redox proteins

Reactive oxygen species (ROS), such as hydrogen peroxide (H_2O_2), are produced mainly in cellular respiration, but also as byproducts of protein folding, or by environmental exposure. Tobacco smoke (Carnevali *et al*., 2003), and asbestos exposure (Shukla *et al*., 2003) are important inducers of ROS generation and accumulation in cells. H_2O_2 is a key member of ROS, taking part in various signal transduction pathways, immune responses, and cell proliferation, also possessing a strong reactivity with biomolecules, *i.e.* DNA, lipids, and proteins. The effects of H_2O_2 in cells depend on exposure time, concentration, and cellular context, and maintaining the cellular ROS levels relatively low is important. Overburden of ROS causes oxidative stress, a hallmark of cancer, since the initiation of many tumours starts from ROS-inflicted DNA damage (**Figure 14**).

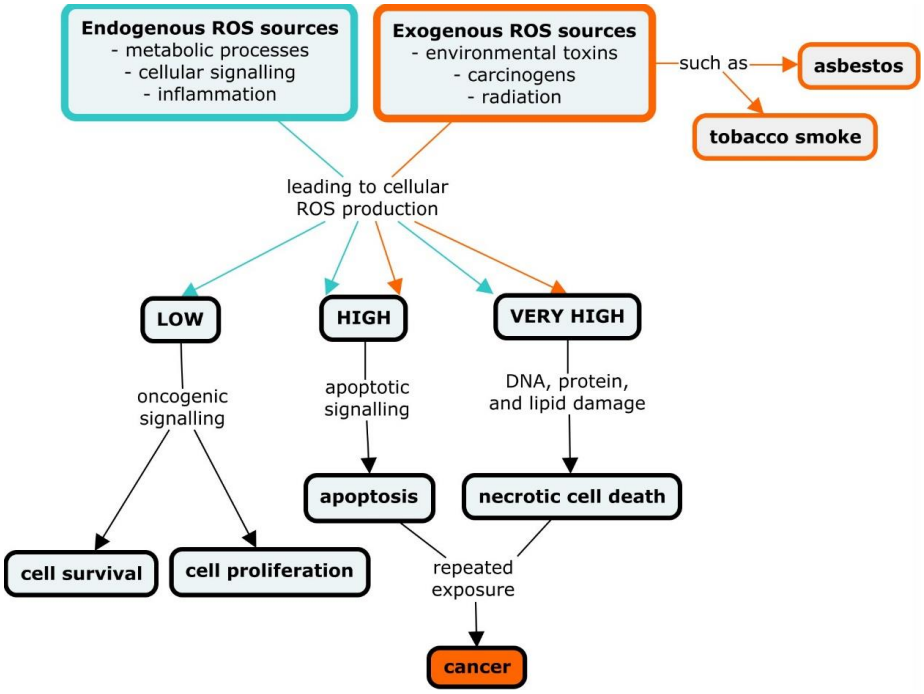


Figure 14 *Cellular effects of reactive oxygen species.*

The cellular levels of H_2O_2 are tightly regulated by redox proteins, such as peroxiredoxins (PRDX) and thioredoxin (TXN) (Netto and Antunes, 2016), all low molecular weight, ubiquitous antioxidant enzymes. The six different mammalian PRDXs are located in the cytosol (PRDX 1, 2, 4, 5, and 6), mitochondria (PRDX 3 and 5), the Golgi apparatus (PRDX4), and in peroxisomes (PRDX5) of cells, and they modulate the concentration of cellular H_2O_2 via redox-reactions mediated by their cysteine residues (Fujii and Ikeda, 2002). Thioredoxin participates in reduction of peroxiredoxins after H_2O_2 oxidation, but also possesses functions in cell signaling and cell proliferation.

Both PRDXs and TXN are under the regulative control of the transcription factor nuclear-factor-erythroid 2 p45-related factor 2 (NRF2), which is under the regulative control of the kelch-like ECH-associated protein (KEAP1)

(Lennicke *et al*., 2015). When ROS are limited, KEAP1 binds to NRF2 resulting in NRF2 degradation, and inhibition of PRDXs and TXN synthesis (**Figure 15**). In cells under high ROS burden, KEAP1-NRF2 signaling pathway is impaired by inactivation of KEAP1, which leads to promoted synthesis of PRDXs and TXN. Genetic alterations leading to loss of the regulative control by NRF2 and KEAP1 are common in lung cancer (Shibata *et al*., 2008; Singh *et al*., 2006; Wang *et al*., 2008). However, in some cancers, NRF2 abundancy is associated with resistance to chemotherapy, suggesting a cancer-promoting role for NRF2 (Wang *et al*., 2008).

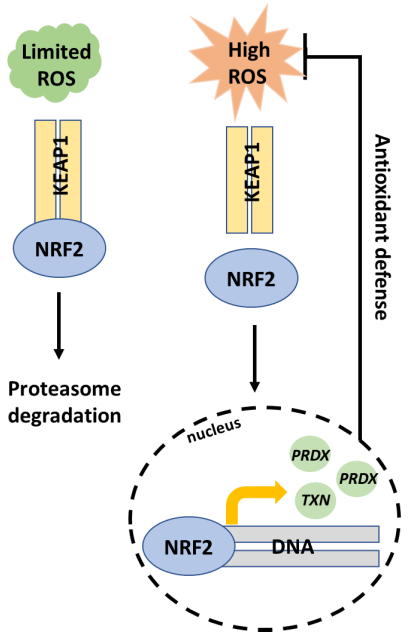


Figure 15 *ROS-dependent NRF2-KEAP1 signaling. Modified from Lennicke et al., 2015.*

The first larger study on PRDXs in lung cancer (Lehtonen *et al*., 2004) examined the abundance of all six PRDXs in lung cancer tissue, comprising different lung cancer subtypes. PRDX 1, 2, 4 and 6 were found to be highly abundant in lung carcinomas and possess differential abundance in lung cancer subtypes; SCLCs were predominantly PRDX negative, as ACs and

SQCs were PRDX positive in immunohistochemical staining. Mitochondrial PRDXs 3 and 5 are abundant in healthy lung tissue, thus no statistically significant alterations between tumour and healthy tissue in immunohistochemical staining were found in the study. However, in another study (Park *et al*., 2006) abundant PRDX3 and TXN were found in human lung cancer tissue, when comparing tumour tissue with its paired healthy tissue. PRDX1 has been found abundant in lung cancer tissue (Park *et al*., 2006), in plasma from individuals with lung cancer (Rostila *et al*., 2012), and lung cancer cell lines (Chang *et al*., 2001). An association with enhanced cell malignancy was found between abundant PRDX1 and PRDX4 in mouse xenografts, and human lung cancer cell lines (Jiang *et al*., 2014), and between abundant PRDX6 in mouse xenografts (Yun *et al*., 2014). Abundant PRDX4 in SQC tissue has been associated with poor patient survival, and high recurrence rate (Hwang *et al*., 2015), suggesting PRDX4 could be used as a prognostic biomarker.

Cancer cells can activate NRF2 synthesis by *e.g.* KEAP1 loss-of-function (Ohta *et al*., 2008), or NRF2 gain-of-function mutations (Ooi *et al*., 2013). Hypermethylation of KEAP1 has been observed to inhibit (Wang *et al*., 2008), and KRAS, MYC, and BRAF activation (DeNicola *et al*., 2011) to promote NRF2 transcription in lung cancer cell lines and tissues. A study on lung carcinogenesis in mice showed that abundant NRF2 promotes tumour growth after 24 weeks but protects from lung tumours before that (Satoh *et al*., 2013). In another mouse study (Kim *et al*., 2010), 10 % of lung carcinomas harboured malfunctioning NRF2. Merikallio *et al*. (2012) studied NRF2 abundancy in lung cancer tissues with immunohistochemical staining, and detected an association with abundant nuclear NRF2, and poor survival. Tao *et al*. (2018) studied the response of NRF2 between chemically and genetically induced lung cancer in mouse lung cancer models. They found that NRF2 protected cells from chemically induced cancer but promoted the growth of existing tumours irrespective of inducer.

To summarize, ROS-regulating proteins, such as PRDXs, TXN, NRF2, and KEAP1, play a key role in lung cancer. PRDX1, PRDX4, and NRF2 could serve as prognostic biomarkers, as they are associated with survival, and tumour progression. Other PRDXs and TXN could be predictive biomarkers, as they have been associated with presence of lung cancer.

2.3.9 *Treatment of lung carcinomas*

For NSCLC, surgery is the recommended option for treatment in stage I and II tumours (Vansteenkiste *et al*., 2014), often coupled to preoperative adjuvant chemotherapy (Hirsch *et al*., 2017). For stage I tumours, radiotherapy can be used if there are contraindications for surgery (Timmerman *et al*., 2010). Stage III inoperable tumours are treated with both radio- and chemotherapy (Curran *et al*., 2011). Advanced NSCLC is always beyond surgical treatment, leaving molecular target therapy as best practice. Data from the Lung Cancer Consortium suggests that up to 69 % of advanced NSCLCs could have a treatable molecular target (Tsao *et al*., 2016), especially among young never-smokers with adenocarcinomas. These targets include mutations in KRAS (25 %), EGFR (21 %), and ALK (7 %) driver oncogenes. In a Finnish study (Mäki-Nevala *et al.*, 2016), EGFR and KRAS mutations were present in 8 % and 26 % of lung adenocarcinomas, respectively. Other studies in Western populations have observed similar results (Boch *et al.*, 2013). In Finns, ALK fusion mutations have been reported in 2 % of NSCLCs (Tuononen *et al*., 2014).

SCLCs initially respond well to chemotherapy and targeted therapy but develop multidrug resistance soon after the start of therapy (Bunn *et al*., 2016). Treatment of SCLC is further hampered by cancer´s high metastatic rate, and manifestation of second primary tumours (Gazdar *et al*., 2017). As aggressive tumours, SCLCs are typically found at extensive disease state, beyond surgical treatment.

3. AIMS OF THE STUDY

The overall aim of this thesis was the search for potential biomarkers in lung cancer, and in *Staphylococcus aureus* bacteraemia with mass spectrometry, both disease groups of world-wide high mortality.

The specific study aims include:

- Identifying CRP-gene SNPs associated with clinical outcome, and blood CRP-levels in *Staphylococcus aureus* bacteraemia. (I)

- Identifying specific biomarkers for early diagnosis of lung cancer in high-risk individuals, such as tobacco smokers, and individuals significantly exposed to asbestos. (II, III)

- Comparing the use of different biological fluids, plasma and sputum, in biomarker discovery. (II, III)

4. MATERIALS AND METHODS

Three studies comprising this thesis can be divided into two different study approaches. The first study focused on genetics, more importantly SNPs of the CRP gene (**Figure 16**). The second two were proteomic studies on lung cancer using plasma and sputum as sample matrices (**Figure 17**).

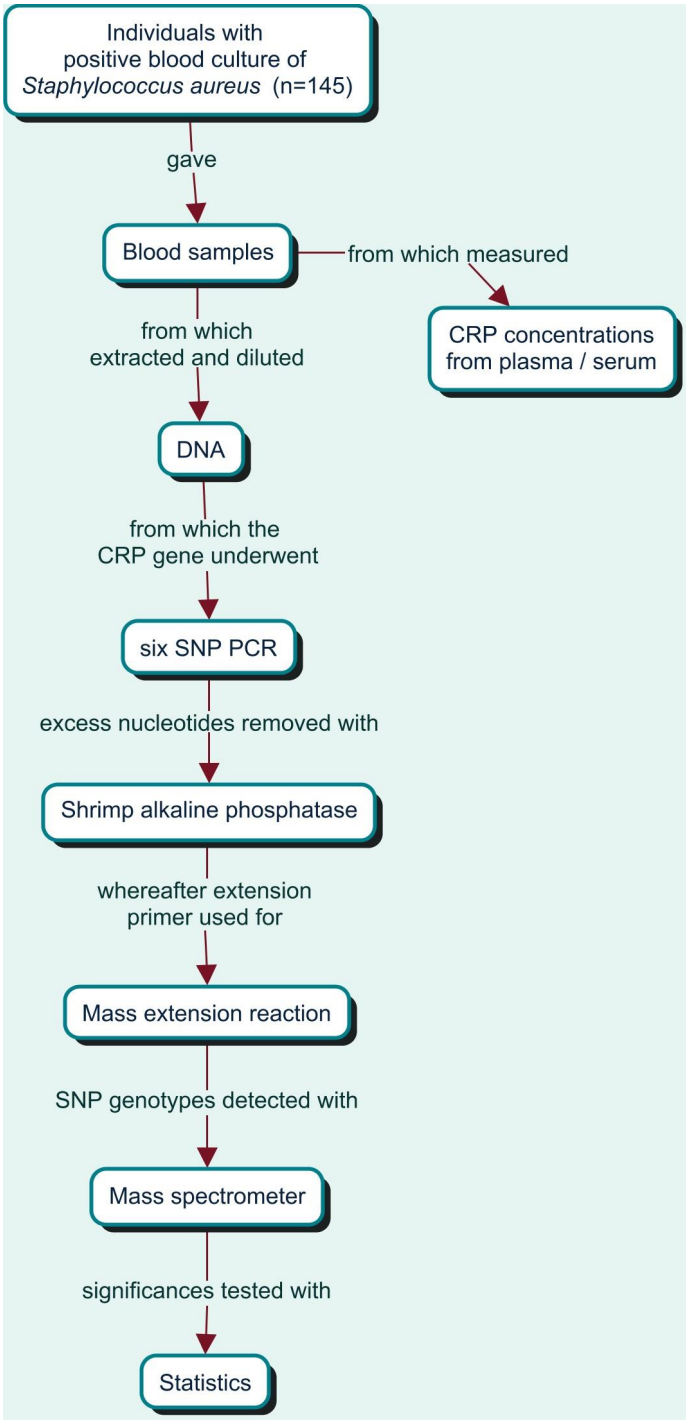


Figure 16 *Workflow of the CRP gene polymorphism study (I).*

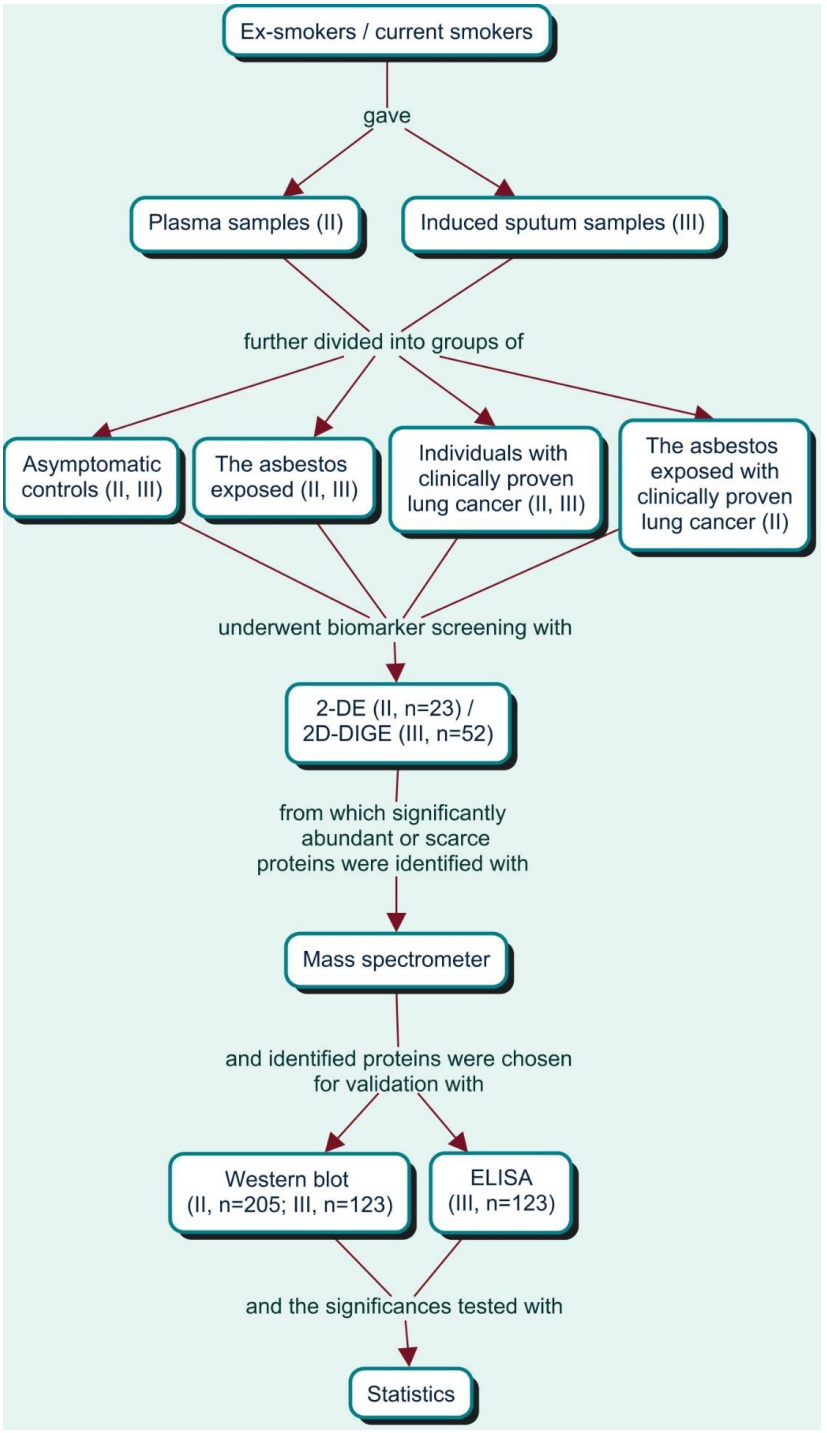


Figure 17 *Workflow of the proteomics studies (II, III).*

4.1 Study samples

4.1.1 *Blood samples from patients with Staphylococcus aureus bacteraemia (I)*

Blood samples were collected from patients with blood culture positive for *S. aureus* in the Helsinki University Hospital in 1999-2002. All samples were collected as a part of a five-university hospital trial, and seven tertiary care hospital trial aimed at examining the potential of adding two fluoroquinolones (trovafloxacin and levofloxacin) to the standard treatment of SAB in order to reduce the high mortality and complication rates. Permission for genetic testing of the blood samples was obtained from all individuals enrolled in the Helsinki University Hospital. The study protocol was approved by the ethics committees of all study sites.

There were several exclusion criteria in the SAB study: age < 18 years, imprisonment, pregnancy, breastfeeding, epilepsy, meningitis, neutropenia, another or polymicrobial bacteraemia, allergy to quinolone antibiotic, tendinitis in fluoroquinolone therapy, prior fluoroquinolone usage for more than five days before randomization, positive culture for S. aureus obtained only from a central intravenous catheter, methicillin-resistant *S. aureus* (MRSA), fluoroquinolone resistant *S. aureus*, and failure to give informed consent.

The severity of the SAB was classified with the McCabe and Jackson criteria (**Table 6)** (McCabe and Jackson, 1962). Other complications, such as endocarditis, deep infection foci, relapse of SAB, or infection of intravenous catheter, were monitored and defined by the latest available criteria (Ruotsalainen *et al*., 2005). During the three-month study period, mortality was surveyed at three time points: 1) first week, 2) at 28 days, and 3) at three months.

Table 6. *Criteria for McCabe classification*

McCabe classification	Explanation
1	no underlying disease
2	non-fatal
3	ultimately fatal (< 5 years)
4	rapidly fatal (< 1 year)

4.1.2 *Plasma and sputum samples in the lung disease studies (II, III)*

All plasma and sputa were from participants in Finnish health survey studies on tobacco smoking, and asbestos exposure. Control and lung cancer samples were from the Lapland central hospital, collected in Lapland as part of a health promotion campaign among tobacco smokers in 2003-2005. Asbestos-exposed samples came from the Finnish Institute of Occupational Health (FIOH), collected in a follow-up study on prevention and detection of asbestos-related disease program in 2006-2008. A pulmonary specialist at FIOH evaluated the extensiveness of asbestos exposure. Other information on *e.g.* smoking or work history was obtained by a personal interview at both study sites. An informed consent was obtained from all study participants, and the ethics committees of Lapland central hospital and FIOH approved the experimental protocols.

Both in plasma and in sputum, the general exclusion criteria included age less than 40 years, smoking duration of less than 10 years, and cancer (besides local basal cell carcinoma of the skin). Individuals with lung cancer bearing another cancer were omitted from the study. Controls and those with lung cancer were evaluated with an interview to check that they were not significantly exposed to asbestos. In neither study (II / III), no individual with less than ten packyears was approved to be part of the study population.

The studied individuals were divided into three groups of current and former smokers: 1) healthy (asymptomatic), 2) asbestos exposed, and 3) clinically proven lung cancer. In the plasma study, five individuals had both significant asbestos exposure and lung cancer.

4.2 Genetics (I)

4.2.1 *Blood CRP measurements*

Serum (or plasma) CRP levels, obtained from blood drawn from individuals with positive blood culture for SAB (n=145), were measured on the day of positive culture, at randomization, and every other day during the first week. The samples underwent automatic immunoturbidimetric analysis using analyzers 917 or Modular PP-analyzer (Hitachi Ltd, Tokyo, Japan), and Tina-quant CRP reagents (Roche Diagnostics, Tina-quant CRP). The normal serum / plasma CRP concentration was <10 mg/l for both methods.

4.2.2 *DNA processing*

DNA was extracted by phenol precipitation (n=145), and diluded twice with Tecan Genesis 150 (LabX, ON, Canada). The first dilution was done in TE-buffer, whereafter DNA concentrations were measured using the fluorescent absorbances obtained from the PicoGreen dsDNA Quantitation Kit (Molecular Probes, Carlsbad, USA). Samples were excited at 480nm, and the fluorescence emission intensity was measured at 520nm. These concentrations were used to carry out the second dilution in purified water, and to achieve the final concentration of 5ng/µl. DNA (10ng / 2µl) was dried using Hydra 96 (Robbins Scientific, CA, USA), and stored at room temperature.

4.2.3 *SNP Genotyping*

Six tag SNPs of the CRP gene were chosen to undergo genotyping: rs1130864, rs1205, rs1800947, rs2794521, rs3091244, and rs3093075. SNP genotyping was done using the Sequenom MassARRAY system with homogeneous MassEXTEND reaction (Sequenom Inc., CA, USA). DNA underwent PCR reaction with Peltier thermal cycler PTC-225 (MJ Research Inc., MA, USA), whereafter the PCR products were verified with agarose (2 %) gel electrophoresis. Shrimp alkaline phosphatase (SAP) was used to degrade all leftover nucleotides after PCR.

The mass extension reaction was carried out with Peltier thermal cycler PTC-225 using thermostable TERMIPol, or HOT TERMIPol extension enzymes (Solis Biodyne, Estonia) compatible with MALDI-TOF analysis of the primer extension products. The extension products were then treated with resin (SpectroCLEAN) in order to remove any excess salts from the samples, before transferred to the SprectoCHIP 601 microchip with Sequenom Robo Design (Sequenom Inc.).

Genotyping of the mass extension products was done with Sequenom Compact MALDI-TOF (Sequenom), and the results were analyzed with TYPER software (Sequenom Inc.). Yield-skew-ratio ≤ 0.2, and a success rate ≥ 92 % were considered as criteria for all genotypes approved in the study.

4.3 Proteomics (II, III)

4.3.1 *Plasma depletion of albumin and immunoglobulin G (II)*

Plasma samples (n=23) prepared for two-dimensional gel electrophoresis were depleted using the PROTIA immunoaffinity column (Sigma Aldrich, St. Louis, MO, USA) according to the manufacturer's instructions. In brief, a

plasma sample was adsorbed into the depletion medium that binds albumin and IgG, and the depleted sample was collected by centrifugation.

4.3.2 *Induction of sputum (III)*

Sputum was induced by hypertonic saline, fixed in Saccomanno fixative (2 % polyethyleneglycol / 50 % ethanol), and smear samples were prepared for papanicolaou staining. Fixed sputa were stored in -20 degrees prior to homogenization. Stained smear samples were screened for cytologic atopia by a cytotechnologist and evaluated by a pathologist.

4.3.3 *Sputum homogenization and concentration (III)*

Sputum was divided by centrifugation into a cellular part and a fluid phase. In this study, the fluid phase was homogenized with recurrent additions of protein reducing sputolysin (1 % dithiothreitol or DTT in 1 mol/l potassiumphosphate buffer solution, pH 7), and by filtrating the samples twice with nylon or PVDF filters (28-100 µm and 0.45 µm). Estimated protein concentrations of the filtered and centrifuged sputa were obtained from silver-stained gels after gel electrophoresis when comparing them to a known concentration of bovine serum albumin (BSA) on the same gel.

To maximize sputum protein concentration for biomarker screening with two-dimensional gel electrophoresis, and to remove any excess DTT, 52 sputum samples (1 ml) underwent a protein concentration protocol with ProteoSpin protein concentration spin columns (Norgen Biotek, Ontario, Canada). The protein concentrations were determined with 2-D Quant kit (GE Healthcare).

4.3.4 *Isoelectric focusing (IEF) (II, III)*

Concentrated sputum samples (III, n=52, 55 µg) underwent the 2-D Clean-Up protocol (GE Healthcare) to dispose of interfering impurities. Proteins were then denaturated in a sample buffer with urea (7 mol/l) and reduced with

thiourea (2 mol/l). CHAPS (4 %) was used as a detergent, and bromophenol blue (0.04 %) as a color marker. Sputum samples were then dyed with fluorescent dyes (Cy2, Cy3, and Cy5) that covalently bind to the ε-amino end of the lysine residues in proteins. These dyes enable the loading of three samples (two individual samples, and a pooled sample as an internal standard sample) on a single strip, improving gel-to-gel variability, spot matching, and quantitation of proteins. The dying was performed according to the manufacturer´s difference gel electrophoresis (DIGE) protocol (GE Healthcare).

For plasma, isoelectric focusing was carried out with the IPGphor II system (GE Healthcare). The 18cm immobilized non-linear strips (GE Healthcare) at three different pH intervals (3-6, 5-8, and 7-10) were hydrated over night with depleted plasma (II, n=23, 100 µg), and the rehydration solution (GE Healthcare) containing Reagent Type 4 (Sigma Aldrich).

Immobilized 18cm non-linear strips (pH 3-10, GE Healthcare) were used for concentrated sputum samples (55 µg). Prior to sample loading, the strips were rehydrated over night, and equilibrated in a solution containing urea (6 mol/l), sodium dodecyl sulfate (2 %), Tris-HCl (pH 8.8, 50 mmol/l), DTT (1 %), iodoacetamide (2 %), glycerol (30 %), and bromophenol blue (0.04 %). Samples were loaded on to the strips using the cup loading method.

4.3.5 ***SDS-PAGE (II, III)***

In the lung disease studies, depleted plasma (II, n=23, 100 µg), and concentrated sputum (III, n=52, 55 µg) samples underwent SDS-PAGE in 12 % polyacrylamide gels with the Ettan DALT*six* system (GE Healthcare). Sputum samples went through the two-dimensional difference gel electrophoresis (2D-DIGE), due to previous labelling with the fluorescent CyDyes.

Plasma gels were silver-stained, and spot detection was performed using the Image Master Platinum 7 software (GE Healthcare). Sputum gels were imaged, and spots were detected using the DeCyder 2D software 7.0 (GE Healthcare), whereafter the gels were silver-stained.

4.3.6 *Identification of proteins with high resolution tandem mass spectrometry (II, III)*

In the lung disease studies (II, III), differentially expressed proteins spots were chosen from the silver-stained 2D-gels for identification with mass spectrometry. Student´s t-tests were used both in Image Master Platinum (II, GE Healthcare) and DeCyder 2D 7.0 (III, GE Healthcare) softwares to evaluate the level of differentiation in expression of each protein spot between study groups. The cut-off values of $p<0.05$ and fold change $\geq \pm 1.5$ were used as signs of significant up- or down-regulation.

Chosen protein spots underwent in-gel digestion with trypsin (modified sequencing grade porcine trypsin, Promega, Madison, WI, USA). Plasma samples were analyzed with automated nanoflow capillary LC–MS/MS using CapLC system coupled to an electronspray ionization quadrupole time-of-flight (Q-TOF) mass spectrometer (Waters, Milford, MS, USA). Sputum samples were analyzed using automated EASY nanoLC 1000 (Proxeon; Thermo Fisher Scientific Inc., San Jose, California, USA) coupled to an electrospray ionization quadrupole-orbitrap mass spectrometer (Q Exactive, Thermo Fisher Scientific Inc.)

The mass fragment spectra obtained were analyzed with in-house Mascot v.2.1 (II, Matrix Science Ltd., London, UK) searched against all human entries in the NCBInr database, or Proteome Discoverer 1.4 (III, Thermo Fisher Scientific Inc.) searched against all human entries in the SwissProt database.

4.3.7 *Western blotting (II, III)*

In the lung disease plasma study (II), 40 undepleted plasma samples (80 µg) underwent WB validation of nine proteins identified with the mass spectrometer: transthyretin (TTR), apolipoprotein A-I (APOA1), retinol binding protein 4 (RBP4), serum amyloid P component (APCS), complement component 5 (C5), serum amyloid A component (SAA), peroxiredoxin 2 (PRDX2), tropomyosin 3 (TPM3), and tropomyosin 4 (TPM4). From this, another 205 undepleted plasma samples were WB validated against four proteins, SAA, PRDX2, TPM3, and TPM4 that got through from the first validation. In addition, PRDX1 was chosen for immunological validation due to its strong homoly with PRDX2. Gradient gels and blots were run on the Criterion system (Biorad). PVDF blots were visualized with chemiluminescent HRP-substrate ECL detection reagent (Perkin Elmer, Waltham, MA, USA). Imaging and analyzing was performed with Kodak X-OMAT 1000A processor (Eastman Kodak Company, Rochester, NY, USA) and ImageMaster Platinum 7.0 software (GE Healthcare), or with ImageQuant LAS 4000 Mini CCD camera, and ImageQuantTL software (GE Healthcare).

In the lung disease sputum study (III), 123 unconcentrated samples (15 µg) were cleaned from interfering reagents, such as salts and DTT, by acetone precipitation. The samples were dissolved, and gels were run according to the Bolt gel system protocol (Invitrogen by Thermo Fisher Scientific). Blotting was performed with Trans-Blot Turbo transfer system (Biorad) on PVDF membranes, and the chemiluminescent HRP-substrate ECL detection of the proteins was executed with the ChemiDoc Touch imaging system (Biorad). Blots were analyzed with the ImageLab 6.0 software (Biorad).

4.3.8 *The enzyme-linked immunosorbent assay (ELISA, III)*

123 unconcentrated sputum samples (15 µg) were tested for protein S100-A8 concentrations using the Bühlmann fcal ELISA kit (Bühlmann Laboratories AG, Schönenbuch, Switzerland) for calprotectin (heterodimer of S100-

A8/A9). Sputa were first precipitated with acetone, and then dissolved into ELISA sample buffer provided in the kit. Immunoassay was performed according to the manufacturer´s intructions. Results were measured in 450 nm with fluorescence spectrofotometer (Perkin Elmer). Standard curve was fitted, and protein concentrations calculated with MatLab software (MathWorks Inc., MA, USA).

4.4 Statistics

IBM SPSS Statistics software was used to conduct statistical testing in all studies: 1) version 14.0 in the CRP gene study (I, SPSS Inc., Chicago, IL, USA) 2) version 18.0 in the lund disease plasma study (II, IBM Corporation, NY, USA), 3) version 24.0 in the lung disease sputum study (III, IBM Corporation, NY, USA. GraphPad Prism 5 (GraphPad Software, La Jolla, CA, USA) was used in the lung disease plasma study (II) to execute pairwise testing.

4.4.1 *Haplotyping and Hardy-Weinberg equilibrium (I)*

Linkage disequilibrium, and haplotype blocks between SNPs was checked with Hapoloview 3.2 software (Broad Institute) using the R^2-model. Hardy-Weinberg equilibrium was calculated with SAS Learning Edition software (SAS Institute Inc., NC, USA).

4.4.2 *Pairwise testing (I-III)*

Analysis of normality was performed for continuous variables using the Kolmogorov-Smirnov test (I-III), Shapiro-Wilks test (I-III), and Levene´s test for homogeneity of variances (III). Wilcoxon-Mann-Whitney U-test was used to analyze the differences between the expression medians of sample groups, and Spearman's correlation to analyze the correlations between

variables in the lung disease studies (II, III). All tests were two-tailed, and statistical significance was determined with $p<0.05$.

The genotypic frequencies of CRP gene (I) for each SNP locus were tested against values expected from the Hardy-Weinberg proportion. The associations between the categorical variables were analyzed by 2-test, or Fisher's exact test, as appropriate. For skewed variables, heterogeneity of genotype groups was tested with Kruskal-Wallis H test.

4.4.3 *Regression analyses (I-III)*

Linear regression univariate analyses were first performed to analyze association between each explanatory variable and outcome variable. Logarithmic or square root transformation was used when necessary to convert the model into a linear form. Clinically meaningful explanatory variables showing association in univariate analysis with the outcome ($p<0.05$) were then included in the multivariate regression in order of strength of their correlation with the outcome variable. When a new explanatory variable was shown to contribute to the variance, influence on the variables already in the model and possible collinearity with them was assessed. For the multiple comparisons in the CRP gene study (I), the false discovery rate estimation method was applied to association analyses. In the lung disease sputum study (III), binomial logistic regression was used to explore the effects of different variables on group categories.

5. RESULTS

The following results were obtained in the three independent studies encompassing this thesis.

Study I)
145 individuals with SAB had blood CRP levels measured and DNA extracted to study the relationship of CRP SNPs and CRP levels in plasma during SAB. Six known SNPs of the CRP gene were chosen for the genotyping, rs3093075, rs1800947, rs2794521, rs1130864, rs1205, and the triallelic SNP rs3091244. Haplotypes generated from genotype data revealed strong linkage disequilibrium between SNPs rs3093075 and rs3091244_AT, rs3093075 and rs3091244_GT, rs1130864 and rs3091244_AT, and between rs1130864 and rs3091244_AG. Hence rs3093075 is in strong LD with the rs3091244 minor A/T-alleles, and rs1130864 with the rs3091244 minor A/T-alleles. Besides SNP rs3093075, all SNPs formed a haplotype block, thus are inherited together more commonly than would be expected by chance (**Figure 18**).

Blood CRP levels in SAB were observed to be influenced by SNPs in the CRP gene. Of the six CRP gene SNPs studied, SNP rs3091244 A-allele was associated with first week high maximal CRP levels, which also associated with the presence of a deep infection focus. The studied SNPs did not affect SAB mortality, degree of leukocytosis, time to defervescence, or number of deep infection foci.

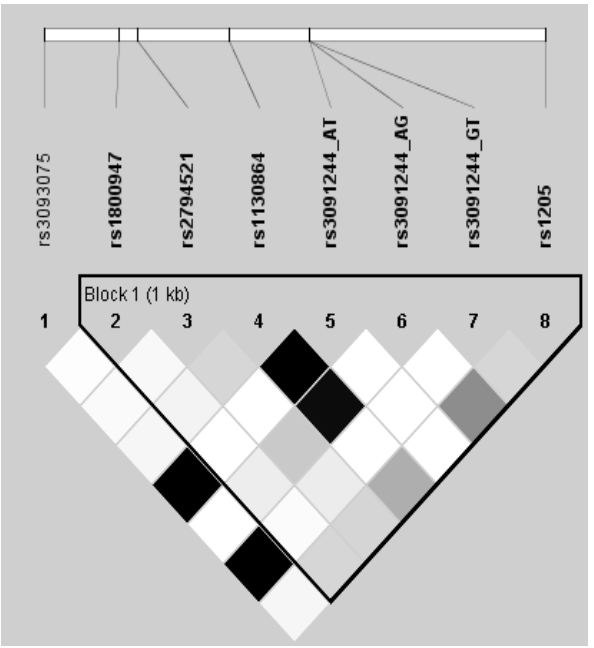


Figure 18 *CRP Haploview indicating linkage disequilibrium between the studied CRP SNPs. All SNPs, excluding SNP rs3093075, form a haplotype block. AT, minor allele T; AG, minor allele A; GT, minor allele T. Black, complete linkage disequilibrium; white, complete linkage equilibrium.*

Study II)

Plasma samples from 23 individuals were used to screen for differentially abundant proteins in four study groups of current and former smokers: 1) healthy, 2) asbestos exposed, 3) clinically proven lung cancer, and 4) asbestos exposed with clinically proven lung cancer. Screening was performed with 2-DE, and protein identification wih mass spectrometry. Altogether 36 differentially abundant protein spots revealed 28 distinctive proteins of which nine were chosen to undergo immunological validation in 40 plasma samples: TPM3, TPM4, C5, APCS, PRX2, TTR, APOA1, SAA, and RBP4. Additionally, PRDX1 was chosen for immunological validation due to its strong homology with PRDX2. Proteins C5, APCS, TTR, APOA1, and RBP4 did not show statistically significant abundance changes between the study groups (**Figure 19**) and were thus omitted from further studies.

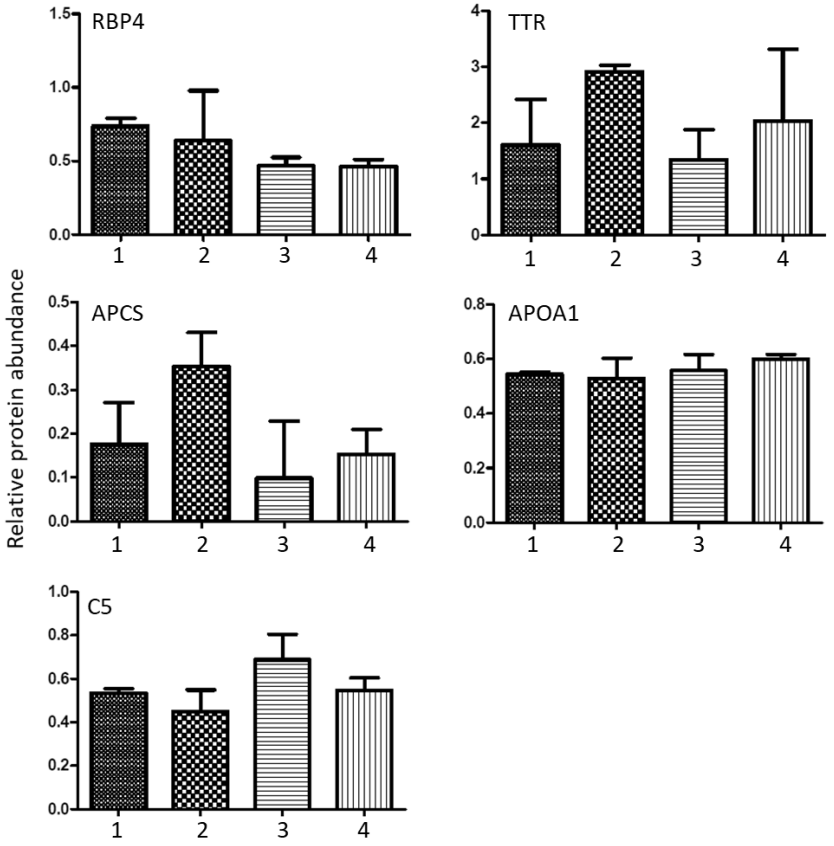


Figure 19 *Relative protein abundances of proteins not chosen for immunological validation in plasma samples of current and former smokers due to lack of statistical significance ($p<0.05$) between study groups. 1, asbestos exposed with clinically proven lung cancer; 2, asbestos exposed; 3, clinically proven lung cancer; 4, healthy.*

In the 40 plasma samples, five plasma proteins, PRDX2, PRDX1, SAA, TPM3, and TPM4 were found statistically significantly differentially abundant in lung cancer and / or asbestos exposure compared to healthy individuals. These results were further immunologically validated in 205 plasma samples. PRDX1 and SAA were associated with lung cancer (**Figure 20**). TPM3, TPM4, PRDX1 and PRDX2 were associated with asbestos exposure (**Figure 21**).

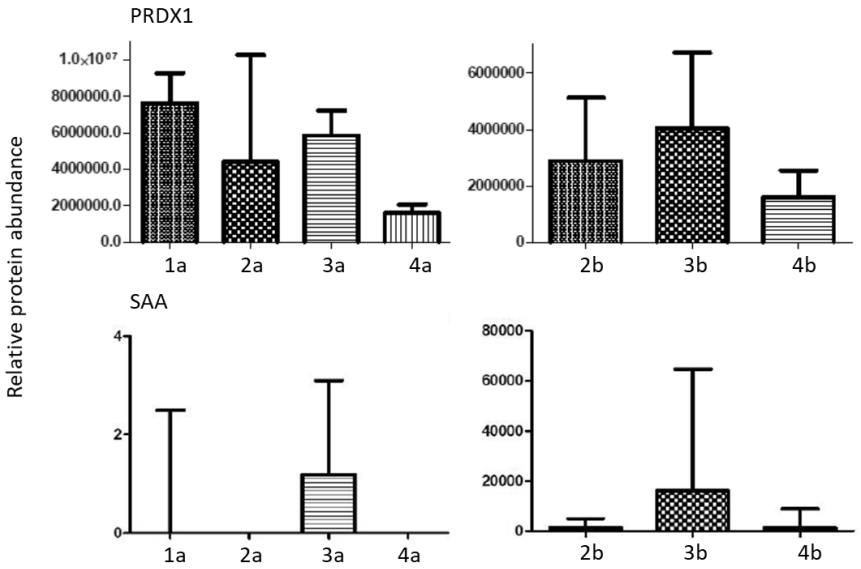


Figure 20 *PRDX1 and SAA were significantly ($p<0.05$) abundant in lung cancer both in the validation with 40 (a) and with 205 (b) plasma samples from current and former smokers. 1, asbestos exposed with clinically proven lung cancer; 2, asbestos exposed; 3, clinically proven lung cancer; 4, healthy.*

PRDX1 and PRDX2 were differentially abundant in lung cancer but correlated conversely with one another; PRDX1 abundance and PRDX2 scarcity was observed in lung cancer. The abundance of the studied proteins did not correlate with cancer type. TPM3 and TPM4 had a correlation with metastazing lung cancer. All studied proteins correlated with at least one smoking parameter. PRDX1 and PRDX2 were negatively correlated with smoking.

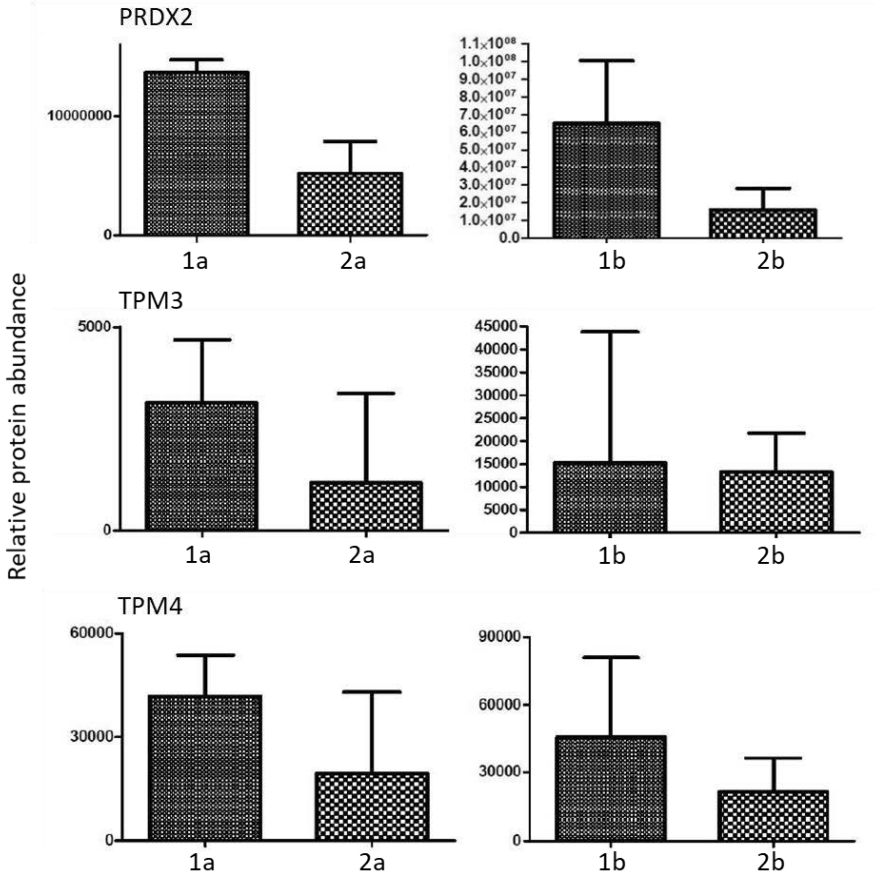
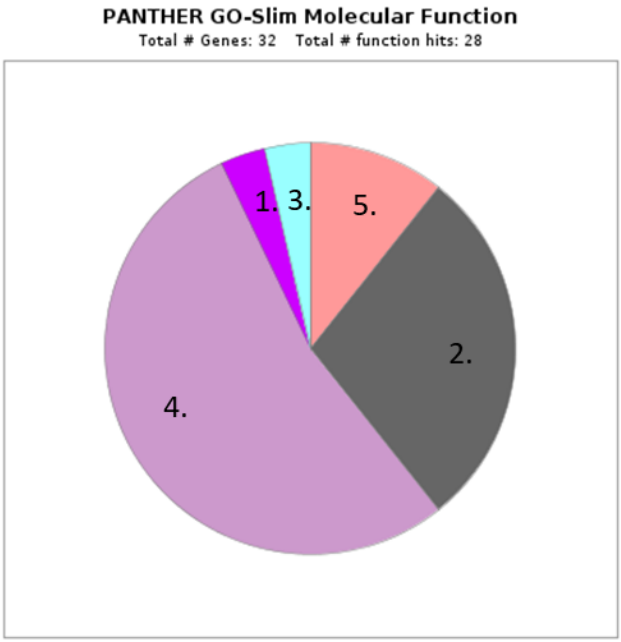


Figure 21 *PRDX2, TPM3, and TPM4 were significantly ($p<0.05$) abundant in asbestos exposure both in validation with 40 (a) and with 205 (b) plasma samples. 1, asbestos exposed; 2, non-exposed.*

Study III)

Screening for proteins of interest was performed using 2D-DIGE, and proteins were identified with mass spectrometry. Altogether 22 differentially abundant protein spots revealed 32 distinct proteins in 52 sputa divided into three study groups of current and former smokers: 1) healthy, 2) asbestos exposed, and 3) clinically proven lung cancer. Functional classification of the 32 differentially abundant proteins categorized the proteins into five main groups:

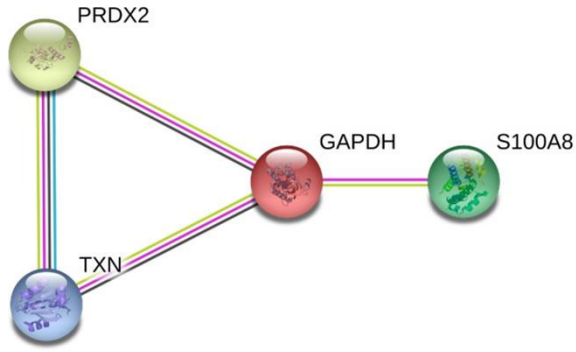
translation regulation, binding, structural activity, catalytic activity, and antioxidant activity (**Figure 22**). Five proteins, PRDX2, TXN, GAPDH, S100A8/9, and ECM1 were detected differentially abundant in lung cancer or asbestos exposure and chosen to undergo immunological validation in 123 sputa.



No.	Molecular function	No. of proteins
1	translation regulator activity (GO:0045182)	1
2	binding (GO:0005488)	8
3	structural molecule activity (GO:0005198)	1
4	catalytic activity (GO:0003824)	15
5	antioxidant activity (GO:0016209)	3

Figure 22 *Functional classification obtained from Panther gene list analysis tool. Data shows the molecular functions of 28/32 proteins identified from 52 sputa.*

PRDX2, TXN, GAPDH and S100A8/9 formed a biomarker panel for lung cancer, sharing a high confidence score in STRING pathway analysis with a PPI enrichment p-value of 0.00331 (**Figure 23**). They also correlated strongly with each other (**Table 7**). PRDX2, TXN, and GAPDH were found significantly abundant in lung cancer (**Figure 24**). Among the five proteins studied, only GAPDH and S100A8/9 were significantly more abundant in adenocarcinomas than in squamous cell carcinomas. Scarce TXN abundance was found in asbestos exposure. All studied proteins correlated with at least one smoking parameter, such as packyears, continued smoking, or increased smoking cessation age. PRDX2 and TXN were positively correlated with increasing number of packyears.



Node 1	Node 2	Co-expression	Experimentally determined interaction	Database annotated	Text-mining	Combined score
TXN	PRDX2	0.080	0.522	0.900	0.813	0.990
TXN	GAPDH	0.165	0.227	0	0.622	0.735
PRDX2	GAPDH	0.183	0.363	0	0.527	0.732
S100A8	GAPDH	0	0.336	0	0.446	0.616

Figure 23 *Protein interactions of PRDX2, TXN, GAPDH, and S100A8 from STRING database (v11.0). PRDX2, TXN, and GAPDH are co-expressed (purple line), and their linkage is determined both experimentally (pink line) and from textmining (green line). S100A8 connects with GAPDH by textmining, and experimentally. PRDX2 and TXN share a known interaction from curated databases (light blue line). The scores indicate the confidence in the interaction, and rank from 0 to 1, 1 indicating the highest confidence.*

Table 7. *Protein-protein correlations between PRDX2, TXN, GAPDH, ans S100A8/9*

	TXN	*GAPDH*	*S100A8*
PRDX2			
Spearman's correlation coefficient	0,471	0,610	0,267
Significance (2-tailed)	0,000	0,000	0,003
TXN			
Spearman's correlation coefficient		0,487	0,243
Significance (2-tailed)		0,000	0,007
GAPDH			
Spearman's correlation coefficient			0,335
Significance (2-tailed)			0,000

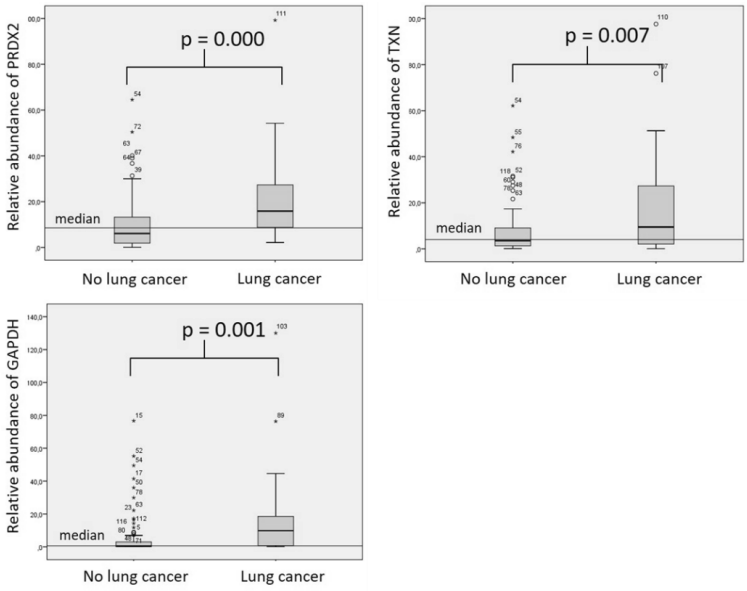


Figure 24 *Comparison of PRDX2, TXN, and GAPDH abundance between the lung cancer group and the non-cancer group from the immunological validation of 123 sputa. The plots present medians with interquartile range. The asymptomatic p-values were obtained from Mann-Whitney U pairwise 2-tailed testing.*

6. DISCUSSION

6.1 CRP gene SNP rs3091244 A-allele is a potential predictive biomarker for blood CRP abundance at baseline

Elevated blood CRP values are typical for bacterial infection, such as SAB, but increased levels are also associated with obesity, and risk of cardiovascular disease. CRP is a blood biomarker used in routine SAB diagnostics, despite its lack of sensitivity and specificity for SAB in particular. High blood CRP levels indicate bacteraemia and risk of complications, but information on CRP baseline levels are of importance when evaluating the severity of SAB.

SNPs of the CRP gene have been shown to influence blood CRP levels at baseline, but studies associating the gene polymorphisms to CRP levels in SAB are lacking. Of the six SNPs researched in study I, SNP rs3091244 A-allele in the gene´s promoter region was shown to affect blood CRP levels, which also correlated with the incidence of a deep infection focus during the first week of SAB. This result suggests that carriers of the A-allele may be more prone to suffer complicated SAB compared to those carrying the C- or the T-allele. As the SNP did not affect *e.g.* mortality, and the CRP levels were only affected for the first week, there must be underlying mechanisms allowing the A-allele carriers to overcome the disease quickly despite their initially bad prognosis.

A study on the effect of blood CRP levels on SAB mortality, Mölkänen *et al.* (2016) found that the mean CRP levels were the highest on the first day of postive blood culture in SAB and declined after that. These results differ from the study I presented here, since the SNP rs3091244 A-allele was only observed to affect the maximal CRP level during the first week of SAB.

Mölkänen *et al*. (2016) also linked the high maximal CRP levels to increased mortality in SAB, which was not detected in study I. Partly the same study population was used in both studies, but Mölkänen *et al*. (2016) did not research the connection between the genetics and the CRP levels or mortality. In accordance with our study, a large population study (Kathiresan *et al*., 2006), and a study on acute coronary syndrome (Suk Danik *et al*., 2006), showed that the SNP rs3091244 A-allele was associated with higher blood CRP levels at baseline.

Due to its location in the promoter of CRP gene, and from the epidemiological studies, it is evident that the SNP rs3091244 A-allele influences the CRP levels. Why the carriers of the A-allele have a delayed response to blood CRP abundance in SAB, is not known. The small study size (n=145), and a small group of patients (n=12) carrying the A-allele can influence the findings, despite statistical accomodations. Factors influencing the CRP levels, such as age, gender, and body mass index, were adjusted for in the study, thus did not affect the results. As studies have implicated, CRP abundance can be used as a diagnostic biomarker for bacteraemias, including SAB, but caution should be applied when using it as a prognostic biomarker. The SNP rs3091244 A-allele, however, acts as a predictive biomarker for CRP abundance in blood at baseline.

6.2 ROS-regulating enzymes, PRDX1, PRDX2, and TXN, are potential diagnostic biomarkers for lung cancer

ROS-regulation is important in maintaining the cell´s homeostasis (Lennicke *et al*., 2015). Redox proteins, such as PRDXs and TXN, react with ROS turning them into less reactive molecules, as when H_2O_2 is turned into water and oxygen. PRDXs and TXN are all regulated by the NRF2-KEAP1-pathway, known to harbor mutations in many cancers, including lung cancer. In the

case of defective NRF2-KEAP1-interaction, PRDXs and TXN are synthesized in excess (Menegon *et al*., 2016).

The role of PRDXs in cancer is complex. In human endothelial cells, the expression of PRDX2, but not PRDX1, is induced by H_2O_2 (Kim *et al*., 1997). In human HeLa cells, PRDX2 inhibits apoptosis via chaperone activity, thus promoting tumour development, and causing resistance in cancer therapy (Moon *et al*., 2005). Many studies have shown that PRDX1 (Park *et al*., 2006; Chang *et al*., 2001) and NRF2 (Satoh *et al*., 2013; Tao *et al*. 2018) are associated with tumour progression, whereas PRDX2 (Lehtonen *et al*., 2004) and TXN (Park *et al*., 2006) have been associated with presence of lung cancer. In mice, PRDX1 functions as tumour suppressor, inhibiting oncogenes such as KRAS, and MYC (Neumann and Fang, 2007; Egler *et al*., 2005), whereas PRDX2 inhibits the over-expression of cell cycle regulators such as TP53 (Han *et al*., 2005).

PRDXs are synthesized by most cells, and to be found in plasma, they need either to be excreted there or end up there from dying cells. In study II with plasma, PRDX1 was abundant in those with lung cancer, and in those exposed to asbestos. This can be explained by the continued activation of PRDX1 gene through loss of the regulative control of NRF2. The possible usability of PRDX1 as a predictive biomarker in lung cancer can be explained by the observation that compared to unsymptomatic healthy smokers, PRDX I was found abundant in the plasma of cancer-free asbestos exposed individuals, who are at risk high of contracting lung cancer. Study II suggests that the abundance of PRDX2 in plasma could inhibit tumour growth through its tumour suppressive functions, thus PRDX1 and PRDX2 respond differently to cellular increase of H_2O_2. This finding is supported by the observation that PRDX2 was abundant in those exposed to asbestos, thus it may exert a protective role in lung carcinogenesis. Also, in study III, TXN was correlated with asbestos exposure. Asbestos exposure is known to account for increased ROS burden in the cells (Shukla *et al*., 2003), thereby adding to

the cancer risk. However, whether the abundance of PRDX2 and TXN in those exposed to asbestos predicts cancer or protects from it, is not known.

Induced sputum is a secretion coughed up from the lower airways. Depending on sample processing, proteins found in sputum are either from its cell or liquid phase. The liquid phase proteins are released either locally by airway epithelial and inflammatory cells or through plasma exudation (Fu *et al*., 2012). In study III with sputum, PRDX2 and TXN abundance correlated with lung cancer, a result easily explained by the activation of NRF2-induced protein synthesis. Compared to results obtained in study II, where PRDX2 was scarce and TXN absent in plasma, it could be that PRDX2 and TXN are produced by cells in the lower airways of individuals with lung cancer. It is possible, however unlikely, that PRDX2 and TXN were not found abundant in the plasma of these individuals due to exudation into sputum or the airways. A more likely reason for the results is that plasma and sputum differ in their protein consistency as they are from different collection sites, thus they can express contrasting results.

It is evident that redox enzymes play a key role in lung carcinogenesis. Even though more studies are needed to understand how, why, and when they are abundant, according to studies II and III, PRDX1, PRDX2, and TXN could serve as diagnostic biomarkers for lung cancer depending on the sample matrix.

6.3 PRDX2, TXN, GAPDH and S100A8/9 form a potential diagnostic biomarker panel for lung cancer

Due to complex molecular mechanisms behind cancer formation, it is unlikely that one biomarker could reach a high sensitivity and specificity for lung cancer. In study III, PRDX2, TXN, GAPDH, and S100A8/9 formed a panel of proteins that associated with presence of lung cancer.

GAPDH is a ubiquitous enzyme involved in many necessary cellular functions (Baxi and Vishwanatha, 1995; Kim *et al*., 2002), but its gene overexpression has been linked to poor prognosis in NSCLC (Puzone *et al*., 2013), and with lung cancer-related EGFR gene mutations (Sasaki *et al*., 2007). In oxidative stress, cancer cells are at risk of becoming glucose deprived (Aykin-Burns *et al*., 2009), thus increasing GAPDH production allow them to maintain energy metabolism.

S100A8 gene mutations are involved in many different cancer types, such as oral cancer (Lunde *et al*., 2014) and NSCLC (Strazisar *et al*., 2009). Calprotectin, whose abundance was studied in study III, is a heterodimer of S100A8 and S100A9. It is associated with *e.g.* colorectal cancer (Turvill *et al*., 2016), and extended survival in NSCLC (Kawai *et al*., 2011). Calprotectin is known to interact with GAPDH, and together contribute to tumour formation (Bresnick *et al*., 2015; Coussens and Werb, 2002).

In study III, there was a strong correlation between PRDX2, TXN, GAPDH, and lung cancer, but likewise a strong correlation between the four proteins involved in inflammatory pathways S100A8/9, GAPDH, TXN, and PRDX2. GAPDH ans S100A8/9 associated with ACs rather than SCCs, and did not associate with smoking. This is in agreement with the fact that ACs are less common in smokers than SCCs. This in turn could indicate that depending on the level and type of stress experienced, different cellular pathways are activated that lead to different lung cancer types. Study II, alongside with other studies (Strazisar *et al*., 2009; Kawai *et al*., 2011), suggests that these proteins could form a diagnostic biomarker panel for lung cancer.

6.4 TPM3 and TPM4 could be diagnostic and prognostic biomarkers for lung cancer

Tropomyosins are actin-binding proteins involved in cytokinesis, cell division, cell motility, and apoptosis (Gunning *et al*., 2008). Cancer cells are characterized by the loss of tropomyosin-stabilized stress fibers (Tojkander *et al*., 2011), and of tropomyosins altogether (Gunning *et al*., 2008). Tropomyosin deficiency has been found in metastatic Lewis carcinoma (Takenaga *et al*., 1988), and apoptotic lung cancer cells (Li *et al*., 2011). In study II, scarce TMP3 and TPM4 abundance was found in association with lung cancer, and their abundance correlated negatively with metastasizing lung cancer. Scarcity of these proteins in blood, could then serve as potential diagnostic biomarkers for lung cancer, or as potential prognostic biomarkers for non-metastasizing lung cancer. Cell motility is a molecular mechanism fundamental to cell survival, and for cancer cells to spread, they need to be able to migrate to new locations (Stuelten *et al*., 2018). Loss of TPM3 and TPM4 can inhibit apoptosis, and hinder cell movement, both features of non-metastasizing cancer. However, setting cut-off values for protein scarcity in clinical use is difficult, especially with proteins needed for normal cell functions. Since both tropomyosins, notably TPM4, were abundant in those with significant exposure to asbestos, they could better serve as biomarkers for asbestos exposure.

6.5 Plasma and sputum are suitable for biomarker discovery, but yield different results

One of the aims of this study was to compare the usefulness of plasma and sputum for biomarker discovery in lung cancer. This was done using the same study population and similar methods in both lung cancer studies (II, III). The motive for the comparison was to find a noninvasive way of collecting samples from the origin of carcinogenesis in lung cancer biomarker studies, and later on for lung cancer diagnostics.

Plasma is an easily available sample, since plasma collection is a routine procedure in daily diagnostics. DNA extraction from blood is also a well-established method. However, due to the complexity of its proteome (Nanjappa *et al*., 2014), plasma might not be the best sample matrix in biomarker discovery. Genetic aberrations have been observed (Kettunen *et al*., 2006; Li *et al*., 2007; Katz *et al*., 2008) in DNA collected from the cellular component of the sputum (Cameron *et al*., 2017), but for proteomics the fluid phase of sputum must be used (Suojalehto *et al*., 2015). The facts that sputum proteome is much more simple than that of plasma, sputum originates in the lungs, and sputum can be collected by a noninvasive induction in a clinical laboratory, all favor the use of sputum in the search for potential biomarkers in lung cancer.

Despite its many advantages in proteomics, sputum is a challenging matrix to work with compared to plasma. After induction, the sample must be homogenized, filtered and cleaned for protein screening. Sample cleaning can cause the mucins to tangle up in the sample (Thornton *et al*., 2008), making it difficult to load on gel and impossible to use in mass spectrometry. Thus, testing and validating sample processing is laborious.

Interestingly in the proteomic studies (II, III), plasma and sputum yielded somewhat different results. In both studies redox proteins were observed in

association with lung cancer and/or asbestos exposure, although differences were found in direction of abundance change and protein type. SAA observed in plasma was understandably missing from sputum. Surprisingly TPMs were absent in sputum. On the contrary, GAPDH and S100A8/9 were proteins of interest in sputum, but not differentially abundant in plasma. Although the results differed, the validated proteins shared the underlying causes, such as inflammation and (oxidative) stress, for their expression. Thus, both plasma and sputum can be used for biomarker discovery, however, mass spectrometry alone might not be a suitable method with sputum samples.

7. CONCLUSIONS

The following conclusion can be drawn based on the three studies that comprise this thesis:

1) The SNP rs3091244 A-allele acts as a predictive biomarker for CRP abundance in blood at baseline. Carriers of the A-allele have high blood CRP levels during the first week of SAB. However, there is no association with mortality. Why the CRP levels were only affected for the first week, is not known. The results of the study indicate that since the A-allele affects the basal CRP levels in the blood, caution should be applied when treating SAB patients and using blood CRP as a predictive marker for SAB.

2) ROS-regulating proteins, here PRDX1 (plasma), PRDX2 (sputum), and TXN (sputum), are important in cancinogenesis and tumourigenesis, and have the highest potential to become predictive or diagnostic biomarkers for lung cancer, according to this thesis. PRDX2 and TXN interact with other cancer-related proteins, such as GAPDH (sputum) and S100A8/9 (sputum), thus their molecular mechanisms are complex, and most likely differ depending on the causing agent of cancer. SAA from plasma would be a good biomarker for lung cancer, but as a marker for acute phase inflammation, lacks specificity for lung cancer. To better understand these underlying mechanisms and interactions, more studies are needed.

3) Tropomyosins are vital for cells to grow and move. The loss of these proteins leads to abnormal cell survival by inhibition of apoptosis, *i.e.* cell immortality characteristic of cancer cells. However, TPM3 and TM4 might not be useful biomarkers in lung cancer, since they were scarce in the plasma of individuals with lung cancer. Their scarcity can be potential predictive biomarker for non-metastasizing lung cancer, since the loss of tropomyosins inhibits cell motility, thus cancer metastasion. As biomarkers, however, they best indicate exposure to asbestos.

4) Mass spectrometry, and its coupled methods (*e.g.* gel electrophoresis), are powerful tools in finding both genome and proteome alterations. These alterations can expose potential biomarkers for high-mortality diseases, such as SAB and lung cancer.

5) Sputum and plasma are both suitable for biomarker discovery, but both sample matrices present challenges. Sputum offers an interesting sample matrix from the origin of lung carcinogenesis, but more studies are needed in order to make it more easily available for high-throughput mass spectrometry.

8. ACKNOWLEDGEMENTS

These studies were carried out in National Public Health Institute (KTL for short, study I), Finnish Institute of Occupational Health (TTL for short, studies II and III), and in the Department of Biochemistry and Developmental Biology of the University of Helsinki (study III). I thank all these institutions for providing me with excellent research facilities.

I would like to express my deepest gratitude to my supervisors Docent Anne Puustinen, Docent Sisko Anttila, and Docent Marc Baumann for their support throughout this very long process. Anne and Sisko I thank first for taking me, a geneticist, on as a doctoral candidate to study proteomics, despite my limited knowledge on biochemistry. Anne, together with the excellent help of laboratory analyst Niina Ahonen, taught me everything I know about hands-on proteomics. Sisko I wish to thank for her great expertise in academic writing, and in lung pathology. Since funding was scarce, I had to do most of the work during holidays and days off from other work. There must have been times when the end of this project has seemed unreachable for all of us. I would like to thank Anne and Sisko for patience, for trusting me with this project and for making it possible to carry it on very independently, every now and then reminding me that they are still there to have my back. However, after TTL closed their reasearch facilities, the end of this project was closer than ever before. I owe my gratitude to Marc, who rescued me, a complete stranger, and let me work in his lab in order to finish my work. He has patiently been there for everything: listening, reassurance, guidance, calming me down, magic tricks, a lot of laughter, and for a shoulder to cry on. Throughout these last couple of years of the project, Marc´s help has been crucial in bringing the last manuscript to publication and this thesis to defence, and his support and encouragement has extended beyond anything one could expect from a supervisor.

The first part of this project started in KTL, where I was supervised by Professor Markus Perola. I wish to thank him for giving me my first chance to work in research, and together with Dr. Mervi Alanne for teaching me the core skills of conducting scientific work. The first publication was conducted with them, and the co-authors Dr. Tomi Mölkänen, Docent Asko Järvinen, and Docent Eeva Ruotsalainen, who I wish to thank for the collaboration.

I wish to thank Docent Katri Vuopala, Docent Tuula Toljamo, and Dr. Irmeli Lindström for a vast collection of plasma and sputum samples in the lung cancer studies. Katri Vuopala I also wish to thank for personally teaching me during my visit to the Lapland Central Hospital, and for offering to accommodate me at her home during my stay.

I wish to warmly thank the official pre-examiners Docent Steffen Ohlmeier and Dr. Petri Kouvonen for their thorough, detailed work in thesis revision. They pressured me into making important changes to the thesis, and I truly value that. Professor Outi Kuittinen is warmly thanked for accepting the role of official opponent at the defence.

I want to warmly acknowledge my thesis committee members, Docent Katarina Pelin and Docent Tuula Nyman, for listening to my concerns and for giving me an outside perspective on the progression of the work.

I am grateful for receiving HUS EVO funding for this thesis during 2009-2010. I also wish to warmly thank the head principal of Helsinki Normal Lyceum, Docent Tapio Lahtero for facilitating the completion of this project by enabling me to arrange my work schedule to get time for research. I have also been granted paid leave for manuscript and thesis writing.

Colleagues have been invaluable for this work to reach its destination. I want to thank Niina Ahonen, Jukka Sund and Pia Kinaret from TTL for creating an amicable working atmosphere, and for always being there to talk, listen and

to help. In the Biochemistry and Developmental Biology lab I was warmly welcomed by Teija Inkinen, Maciej Lalowski, Rabah Soliymani, Giulio Calza, Susanna Koskelainen, and Arvydas Dapkunas, who took me in to their lab group without hesitation. I am grateful for your help, encouragement and friendship. I also wish to thank lab technician Teija Inkinen for her help and guidance with the hands-on lab work in the third study. My fellow teachers in Norssi, Outi Houtsonen, Marjut Holopainen, Sara Lindholm, and Pauliina Toivonen, I want to thank for being there for me through the good and bad, and for never being too busy to listen. I thank them for the laughters, tears, hugs, mountains of chocolate, love, and endless support.

Mom and dad taught me the value of hard work, and they did their best to support my ambitions. Unfortunately, mom never got to see me graduate from the University, but as a teacher herself, she held education in high regard. I wish to thank my parents for letting me find my own way, to pursue my dreams and to become whatever I want to become. My brother, although no longer with us, I wish to thank for his love and support. He was a wonderful big brother, and always stood by me, despite our differing aspirations.

My “sisters” Anette, Justiina, and Josefiina, I thank for loving me, keeping my spirits high, and being there for me through thick and thin. My dear friends / family members, especially Heikki, Riitta, mummu, Sirpa, Mikaela, Liisa, Noora, Kimmo, Valentina, Clarice, and Elina, I thank for friendship, believing in me, and for good company.

Finally, I wish to thank my beautiful, smart, and resilient children Elisabet and Elias for making me feel special and loved every single day. Their smiles and hugs kept me going even when completing this work seemed like a hopeless task.

Espoo, August 3, 2020

Annina

9. REFERENCES

1. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature 2010;467(7319):1061-1073.
2. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD; National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365(5):395-409.
3. Albert MA. Biomarkers and heart disease. J Clin Sleep Med 2011;7(5 Suppl):S9-11.
4. Alessandro Leal, Nicole C T van Grieken, Doreen N Palsgrove, Jillian Phallen, Jamie E Medina, Carolyn Hruban, Mark A M Broeckaert, Valsamo Anagnostou, Vilmos Adleff, Daniel C Bruhm, Jenna V Canzoniero, Jacob Fiksel, Marianne Nordsmark, Fabienne A R M Warmerdam, Henk M W Verheul, Dick Johan van Spronsen, Laurens V Beerepoot, Maud M Geenen, Johanneke E A Portielje, Edwin P M Jansen, Johanna van Sandick, Elma Meershoek-Klein Kranenbarg, Hanneke W M van Laarhoven, Donald L van der Peet, Cornelis J H van de Velde, Marcel Verheij, Remond Fijneman, Robert B Scharpf, Gerrit A Meijer, Annemieke Cats, Victor E Velculescu. White Blood Cell and Cell-Free DNA Analyses for Detection of Residual Disease in Gastric Cancer. Nat Commun 2020;11(1):525.
5. Alexandrov K, Rojas M, Satarug S. The critical DNA damage by benzo(a)pyrene in lung tissues of smokers and approaches to preventing its formation. Toxicol Lett 2010;198(1):63-68.
6. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, Totoki Y, Fujimoto A, Nakagawa H, Shibata T, Campbell PJ, Vineis P, Phillips DH, Stratton MR. Mutational signatures associated with tobacco smoking in human cancer. Science 2016;354:618-622.

7. Aleksandrov ML, Gall LN, Krasnov VN, Nikolaev VI, Pavlenko VA, Shkurov VA. Ionextraction from solutions at atmospheric pressure – a method for mass spectrometricanalysis of bioorganic substances. Dokl Akad Nauk SSSR 1984;277:379-383.
8. Alsemgeest SP, Kalsbeek HC, Wensing T, Koeman JP, van Ederen AM, Gruys E. Concentrations of serum amyloid-A (SAA) and haptoglobin (HP) as parameters of inflammatory diseases in cattle. Vet Q 1994;16(1):21-23.
9. American Cancer Society. Global Cancer Facts & Figures 3rd edn (American Cancer Society, 2015).
10. American Diabetes Association. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20:1183-1197.
11. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002;1(11):845-867.
12. Andujar P, Wang J, Descatha A, Galateau-Sallé F, Abd-Alsamad I, Billon-Galland MA, Blons H, Clin B, Danel C, Housset B, Laurent-Puig P, Le Pimpec-Barthes F, Letourneux M, Monnet I, Régnard JF, Renier A, Zucman-Rossi J, Pairon JC, Jaurand MC. p16INK4A inactivation mechanisms in non-small-cell lung cancer patients occupationally exposed to asbestos. Lung Cancer 2010;67(1):23-30.
13. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tulinius H, Thorlacius S, Eerola H, Nevanlinna H, Syrjäkoski K, Kallioniemi OP, Thompson D, Evans C, Peto J, Lalloo F, Evans DG, Easton DF. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. Am J Hum Genet 2003;72:1117-1130.

14. Ao MH, Zhang H, Sakowski L, Sharma R, Illei PB, Gabrielson E, Askin F, Li QK. The utility of a novel triple marker (combination of TTF1, napsin A, and p40) in the subclassification of non-small cell lung cancer. Hum Pathol 2014;45(5):926-934.
15. Arnson Y, Shoenfeld Y, Amital H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. J Autoimmun 2010;34(3):258-265.
16. Asgeirsson H, Thalme A, Weiland O. Staphylococcus aureus bacteraemia and endocarditis - epidemiology and outcome: a review. Infect Dis (Lond) 2018;50(3):175-192.
17. Ashbrook JD, Spector AA, Santos EC, Fletcher JE. Long chain fatty acid binding to human plasma albumin. J Biol Chem 1975;250:2333-2338.
18. Aykin-Burns N, Ahmad IM, Zhu Y, Oberley LW, Spitz DR. Increased levels of superoxide and H2O2 mediate the differential susceptibility of cancer cells versus normal cells to glucose deprivation. Biochem J 2009;418(1):29-37.
19. Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V.; WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens - Part F: chemical agents and related occupations. Lancet Oncol 2009;10:1143-1144.
20. Badolato R, Wang JM, Murphy WJ, Lloyd AR, Michiel DF, Bausserman LL, Kelvin DJ, Oppenheim JJ. Serum amyloid A is a chemoattractant: induction of migration, adhesion, and tissue infiltration of monocytes and polymorphonuclear leukocytes. J Exp Med 1994;180(1):203-209.
21. Balistreri CR, Vasto S, Listì F, Grimaldi MP, Lio D, Colonna-Romano G, Caruso M, Caimi G, Hoffmann E, Caruso C, Candore G. Association between +1059G/C CRP polymorphism and acute myocardial infarction in a cohort of patients from Sicily: a pilot study. Ann N Y Acad Sci 2006;1067:276-281.
22. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? Lancet 2001;357(9255):539-545.

23. Ballard-Barbash R, Friedenreich CM, Courneya KS, Siddiqi SM, McTiernan A, Alfano CM. Physical activity, biomarkers, and disease outcomes in cancer survivors: a systematic review. J Natl Cancer Inst 2012;104:815-840.
24. Banka CL, Yuan T, de Beer MC, Kindy M, Curtiss LK, de Beer FC. Serum amyloid A (SAA): influence on HDL-mediated cellular cholesterol efflux. J Lipid Res 1995;36(5):1058-1065.
25. Baraniuk JN, Casado B, Pannell LK, McGarvey PB, Boschetto P, Luisetti M, Iadarola P. Protein networks in induced sputum from smokers and COPD patients. Int J Chron Obstruct Pulmon Dis 2015;10:1957-1975.
26. Barbhuiya MA, Sahasrabuddhe NA, Pinto SM, Muthusamy B, Singh TD, Nanjappa V, Keerthikumar S, Delanghe B, Harsha HC, Chaerkady R, Jalaj V, Gupta S, Shrivastav BR, Tiwari PK, Pandey A. Comprehensive proteomic analysis of human bile. Proteomics 2011;11(23):4443-4453.
27. Bayer AS, Guze LB. Staphylococcus aureus bacteremic syndromes: diagnostic and therapeutic update. Dis. Mond 1979;25:1-42.
28. Baxi MD, Vishwanatha JK. Uracil DNA-glycosylase/glyceraldehyde-3-phosphate dehydrogenase is an Ap4A binding protein. Biochemistry 1995;34(30):9700-9707.
29. Bejarano PA, Baughman RP, Biddinger PW, Miller MA, Fenoglio-Preiser C, al-Kafaji B, Di Lauro R, Whitsett JA. Surfactant proteins and thyroid transcription factor-1 in pulmonary and breast carcinomas. Mod Pathol 1996;9:445-452.
30. Benoist JF, Mimoz O, Assicot M, Edouard A. Serum procalcitonin, but not C-reactive protein, identifies sepsis in trauma patients. Clin Chem 1998;44:1778-1779.
31. Bermudez EA, Rifai N, Buring JE, Manson JE, Ridker PM. Relation between markers of systemic vascular inflammation and smoking in women. Am J Cardiol 2002;89:1117-1119.
32. Beutler B, Cerami A. Cachectin/tumor necrosis factor: an endogenous mediator of shock and inflammation. Immunol Res 1986;5(4):281-293.

33. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 2001;69(3):89-95.
34. Blackburn WD Jr. Validity of acute phase proteins as markers of disease activity. J Rheumatol Suppl 1994;42:9-13.
35. Blanchon T, Bréchot JM, Grenier PA, Ferretti GR, Lemarié E, Milleron B, Chagué D, Laurent F, Martinet Y, Beigelman-Aubry C, Blanchon F, Revel MP, Friard S, Rémy-Jardin M, Vasile M, Santelmo N, Lecalier A, Lefébure P, Moro-Sibilot D, Breton JL, Carette MF, Brambilla C, Fournel F, Kieffer A, Frija G, Flahault A; Dépiscan Group. Baseline results of the Depiscan study: a French randomized pilot trial of lung cancer screening comparing low dose CT scan (LDCT) and chest X-ray (CXR). Lung Cancer 2007;58(1):50-58.
36. Boch C, Kollmeier J, Roth A, Stephan-Falkenau S, Misch D, Grüning W, Bauer TT, Mairinger T. The frequency of EGFR and KRAS mutations in non-small cell lung cancer (NSCLC): routine screening data for central Europe from a cohort study. BMJ Open 2013;3(4).
37. Boelaert JR, Daneels RF, Schurgers ML, Matthys EG, Gordts BZ, Van Landuyt HW. Iron overload in haemodialysis patients increases the risk of bacteraemia: a prospective study. Nephrol Dial Transplant 1990;5:130-134.
38. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nat Genet 2003;33(Suppl):228-237.
39. Boulanger G, Andujar P, Pairon JC, Billon-Galland MA, Dion C, Dumortier P, Brochard P, Sobaszek A, Bartsch P, Paris C, Jaurand MC. Quantification of short and long asbestos fibers to assess asbestos exposure: a review of fiber size toxicity. Environ Health 2014;13:59.
40. Bray MS, Boerwinkle E, Doris PA. High-throughput multiplex SNP genotyping with MALDI-TOF mass spectrometry: practice, problems and promise. Hum Mutat 2001;17(4):296-304.

41. Brenner H, Rothenbacher D, Arndt V. Epidemiology of stomach cancer. Methods Mol Biol 2009;472:467-477.
42. Bresnick AR, Weber DJ, Zimmer DB. S100 proteins in cancer. Nat Rev Cancer 2015;15(2):96-109.
43. Brittain HK, Scott R, Thomas E. The rise of the genome and personalised medicine. Clin Med (Lond) 2017;17(6):545-551.
44. Brodersen R. Binding of bilirubin to albumin. CRC Crit Rev Clin Lab Sci 1979;11:305-399.
45. Brown AF, Sirohi D, Fukuoka J, Cagle PT, Policarpio-Nicolas M, Tacha D, Jagirdar J. Tissue-preserving antibody cocktails to differentiate primary squamous cell carcinoma, adenocarcinoma, and small cell carcinoma of lung. Arch Pathol Lab Med 2013;137(9):1274-1281.
46. Brown T, Darnton A, Fortunato L, Rushton L; British Occupational Cancer Burden Study Group. Occupational cancer in Britain. Respiratory cancer sites: larynx, lung and mesothelioma. Br J Cancer 2012;107 Suppl 1:S56-S70.
47. Brull DJ, Serrano N, Zito F, Jones L, Montgomery HE, Rumley A, Sharma P, Lowe GD, World MJ, Humphries SE, Hingorani AD. Human CRP gene polymorphism influences CRP levels: implications for the prediction and pathogenesis of coronary heart disease. Arterioscler Thromb Vasc Biol 2003;23(11):2063-2069.
48. Buetow KH, Edmonson M, MacDonald R, Clifford R, Yip P, Kelley J, Little DP, Strausberg R, Koester H, Cantor CR, Braun A. High-throughput development and characterization of a genome wide collection of gene-based single nucleotide polymorphism markers by chip- based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Proc Natl Acad Sci USA 2001;98:581-584.
49. Bunn PA Jr, Minna JD, Augustyn A, Gazdar AF, Ouadah Y, Krasnow MA, Berns A, Brambilla E, Rekhtman N, Massion PP, Niederst M, Peifer M, Yokota J, Govindan R, Poirier JT, Byers LA, Wynes MW, McFadden DG, MacPherson D, Hann CL, Farago AF, Dive C, Teicher BA, Peacock CD, Johnson JE, Cobb MH, Wendel HG, Spigel D, Sage J, Yang P, Pietanza

MC, Krug LM, Heymach J, Ujhazy P, Zhou C, Goto K, Dowlati A, Christensen CL, Park K, Einhorn LH, Edelman MJ, Giaccone G, Gerber DE, Salgia R, Owonikoko T, Malik S, Karachaliou N, Gandara DR, Slotman BJ, Blackhall F, Goss G, Thomas R, Rudin CM, Hirsch FR. Small cell lung cancer: can recent advances in biology and molecular biology be translated into improved outcomes? J Thorac Oncol 2016;11:453-474.

50. Calandra T, Baumgartner JD, Grau GE, Wu MM, Lambert PH, Schellekens J, Verhoef J, Glauser MP. Prognostic values of tumor necrosis factor/cachectin, interleukin-1, interferon-alpha, and interferon-gamma in the serum of patients with septic shock. Swiss-Dutch J5 Immunoglobulin Study Group. J Infect Dis 1990;161(5):982-987.
51. Cameron SJS, Lewis KE, Huws SA, Hegarty MJ, Lewis PD, Pachebat JA, Mur LAJ. A pilot study using metagenomic sequencing of the sputum microbiome suggests potential bacterial biomarkers for lung cancer. PLoS One 2017;12(5):e0177062.
52. Carnevali S, Petruzzelli S, Longoni B, Vanacore R, Barale R, Cipollini M, Scatena F, Paggiaro P, Celi A, Giuntini C. Cigarette smoke extract induces oxidative stress and apoptosis in human lung fibroblasts. Am J Physiol Lung Cell Mol Physiol 2003;284(6):L955-963.
53. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. N Engl J Med 1991; 324(17):1156-1161.
54. Cetinkaya M, Ozkan H, Köksal N, Celebi S, Hacimustafaoğlu M. Comparison of serum amyloid A concentrations with those of C-reactive protein and procalcitonin in diagnosis and follow-up of neonatal sepsis in premature infants. J Perinatol 2009;29(3):225-231.
55. Chabot-Richards DS, George TI. Leukocytosis. Int J Lab Hematol 2014;36(3):279-288.
56. Chang JW, Jeon HB, Lee JH, Yoo JS, Chun JS, Kim JH, Yoo YJ. Augmented expression of peroxiredoxin I in lung cancer. Biochem Biophys Res Commun 2001;289(2):507-512.

57. Chaurand P, Sanders ME, Jensen RA, Caprioli RM. Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. Am J Pathol 2004;165(4):1057-1068.
58. Chavakis T, Hussain M, Kanse SM, Peters G, Bretzel RG, Flock JI, Herrmann M, Preissner KT. Staphylococcus aureus extracellular adherence protein serves as anti-inflammatory factor by inhibiting the recruitment of host leukocytes. Nat Med 2002;8:687-693.
59. Chen J, Zhao J, Huang J, Su S, Qiang B, Gu D. −717A>G polymorphism of human C-reactive protein gene associated with coronary heart disease in ethnic Han Chinese: the Beijing atherosclerosis study. J Mol Med 2005;83:72-78.
60. Chen Y, Thai P, Zhao YH, Ho YS, DeSouza MM, Wu R. Stimulation of airway mucin gene expression by interleukin (IL)-17 through IL-6 paracrine/autocrine loop. J Biol Chem 2003;278:17036-17043.
61. Chen XH, Huang S, Kerr D. Biomarkers in clinical medicine. IARC Sci Publ 2011;(163):303-322.
62. Church DF, Pryor WA. Free radical chemistry of cigarette smoke and its toxicological implications. Environ Health Perspect 1985;64:111-126.
63. Christensson B, Espersen F, Hedström SA, Kronvall G. Serological assays against Staphylococcus aureus peptidoglycan, crude staphylococcal antigen and staphylolysin in the diagnosis of serious S. aureus infections. Scand J Infect Dis 1985;17(1):47-53.
64. Cho JY, Sung HJ. Proteomic approaches in lung cancer biomarker development. Expert Rev Proteomics 2009;6(1):27-42.
65. Clyne B, Olshaker JS. The C-reactive protein. J Emerg Med 1999;17(6):1019-1025.
66. Coetzee GA, Strachan AF, van der Westhuyzen DR, Hoppe HC, Jeenah MS, de Beer FC. Serum amyloid A-containing human high density lipoprotein 3. Density, size, and apolipoprotein composition. J Biol Chem 1986;261(21):9644-9651.
67. Collier TS, Muddiman DC. Analytical strategies for the global quantification of intact proteins. Amino Acids 2012;43:1109-1117.

68. Cooper RL, Lindsey AJ, Waller RE. The presence of 3,4–benzopyrene in cigarette smoke. Chem Ind 1954;46: 1418.
69. Corberand J, Nguyen F, Do AH, Dutau G, Laharrague P, Fontanilles AM, Gleizes B. Effect of tobacco smoking on the functions of polymorphonuclear leukocytes. Infect Immun 1979;23(3):577-581.
70. Coussens LM, Werb Z. Inflammation and cancer. Nature 2002;420(6917):860-867.
71. Crawford DC, Yi Q, Smith JD, Shephard C, Wong M, Witrak L, Livingston RJ, Rieder MJ, Nickerson DA. Allelic spectrum of the natural variation in CRP. Hum Genet 2006;119(5):496-504.
72. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. Nat Rev Clin Oncol 2013;10:472-484
73. Curran WJ Jr, Paulus R, Langer CJ, Komaki R, Lee JS, Hauser S, Movsas B, Wasserman T, Rosenthal SA, Gore E, Machtay M, Sause W, Cox JD. Sequential vs. concurrent chemoradiation for stage III non-small cell lung cancer: randomized phase III trial RTOG 9410. J Natl Cancer Inst 2011;103:1452-1460.
74. Davey Smith G, Lawlor DA, Harbord R, Timpson N, Rumley A, Lowe GD, Day IN, Ebrahim S. Association of C-reactive protein with blood pressure and hypertension: life course confounding and mendelian randomization tests of causality. Arterioscler Thromb Vasc Biol 2005;25:1051-1056.
75. Davis SM, Clark EA, Nelson LT, Silver RM. The association of innate immune response gene polymorphisms and puerperal group a streptococcal sepsis. Am J Obstet Gynecol 2010;202(3):e301-308.
76. De Haas CJ, Veldkamp KE, Peschel A, Weerkamp F, Van Wamel WJ, Heezius EC, Poppelier MJ, Van Kessel KP, van Strijp JA. Chemotaxis inhibitory protein of Staphylococcus aureus, a bacterial anti-inflammatory agent. J Exp Med 2004;199:687-695.
77. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. Clin Chest Med 2011;32(4):605-644.

78. Denburg JA. Basophil and mast cell lineages in vitro and in vivo. Blood 1992;79:846-860.
79. DeNicola GM, Karreth FA, Humpton TJ, Gopinathan A, Wei C, Frese K, Mangal D, Yu KH, Yeo CJ, Calhoun ES, Scrimieri F, Winter JM, Hruban RH, Iacobuzio-Donahue C, Kern SE, Blair IA, Tuveson DA. Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. Nature 2011;475(7354):106-109.
80. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. Science 1996;274:430-432.
81. De Prost N, Razazi K, Brun-Buisson C. Unrevealing culture-negative severe sepsis. Crit Care 2013;17(5):1001.
82. Djukanović R, Sterk PJ, Fahy JV, Hargreave FE. Standardised methodology of sputum induction and processing. Eur Respir J Suppl 2002;37:1s-2s.
83. Dopp E, Schuler M, Schiffman D, Eastmond DA. Induction of micronuclei, hyperdiploidy and chromosomal breakage affecting centric / pericentric regions of chromosomes 1 and 9 in human amniotic fluid cells after treatment with asbestos and ceramic fibers. Mutat Res 1997;377(1):77-87.
84. Dostert C, Pétrilli V, Van Bruggen R, Steele C, Mossman BT, Tschopp J. Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica. Science 2008;320(5876):674-677.
85. Eberhard OK, Haubitz M, Brunkhorst FM, Kliem V, Koch KM, Brunkhorst R. Usefulness of procalcitonin for differentiation between activity of systemic autoimmune and invasive bacterial infection. Arthritis Rheum 1997;40:1250-1256.
86. Egler RA, Fernandes E, Rothermund K, Sereika S, De Souza-Pinto N, Jaruga P, Dizdaroglu M, Prochownik EV. Regulation of reactive oxygen species, DNA damage, and c-Myc function by peroxiredoxin 1. Oncogene 2005;24:8038-8050.

87. El Bali L, Diman A, Bernard A, Roosens NH, De Keersmaecker SC. Comparative study of seven commercial kits for human DNA extraction from urine samples suitable for DNA biomarker-based public health studies. J Biomol Tech 2014;25(4):96-110.
88. El-Maghraby SM, Moneer MM, Ismail MM, Shalaby LM, El-Mahallawy HA. The diagnostic value of C-reactive protein, interleukin-8, and monocyte chemotactic protein in risk stratification of febrile neutropenic children with hematologic malignancies. J Pediatr Hematol Oncol 2007;29(3):131-136.
89. Esposito K, Pontillo A, Di Palo C, Giugliano G, Masella M, Marfella R, Giugliano D. Effect of weight loss and lifestyle changes on vascular inflammatory markers in obese women: a randomized trial. JAMA 2003;289(14):1799-1804.
90. Eugen-Olsen J, Gustafson P, Sidenius N, Fischer TK, Parner J, Aaby P, Gomes VF, Lisse I. The serum level of soluble urokinase receptor is elevated in tuberculosis patients and predicts mortality during treatment: a community study from Guinea-Bissau. Int J Tuberc Lung Dis 2002;6:686-692.
91. Eugen-Olsen J, Andersen O, Linneberg A, Ladelund S, Hansen TW, Langkilde A, Petersen J, Pielak T, Møller LN, Jeppesen J, Lyngbaek S, Fenger M, Olsen MH, Hildebrandt PR, Borch-Johnsen K, Jørgensen T, Haugaard SB. Circulating soluble urokinase plasminogen activator receptor predicts cancer, cardiovascular disease, diabetes and mortality in the general population. J Intern Med 2010;268(3):296-308.
92. Fang XM, Schröder S, Hoeft A, Stüber F. Comparison of two polymorphisms of the interleukin-1 gene family: interleukin-1 receptor antagonist polymorphism contributes to susceptibility to severe sepsis. Crit Care Med 1999;27(7):1330-1334.
93. Falahi E, Khalkhali Rad AH, Roosta S. What is the best biomarker for metabolic syndrome diagnosis? Diabetes Metab Syndr 2015;9(4):366-372.

94. Fatima N, Cohen C, Lawson D, Siddiqui MT. TTF-1 and Napsin A double stain: a useful marker for diagnosing lung adenocarcinoma on fine-needle aspiration cell blocks. Cancer Cytopathol 2011;119(2):127-133.
95. Ferrari SL, Ahn-Luong L, Garnero P, Humphries SE, Greenspan SL. Two promoter polymorphisms regulating interleukin-6 gene expression are associated with circulating levels of C-reactive protein and markers of bone resorption in postmenopausal women. J Clin Endocrinol Metab 2003;88(1):255-259.
96. Finley Austin MJ, Babiss L. Commentary: where and how could biomarkers be used in 2016? AAPS J 2006;8(1):E185-E189.
97. Fitzgerald SF, O'Gorman J, Morris-Downes MM, Crowley RK, Donlon S, Bajwa R, Smyth EG, Fitzpatrick F, Conlon PJ, Humphreys H. A 12-year review of Staphylococcus aureus bloodstream infections in haemodialysis patients: more work to be done. J Hosp Infect 2011;79:218-221.
98. Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, Allen C, Hansen G, Woodbrook R, Wolfe C, Hamadeh RR, Moore A, Werdecker A, Gessner BD, Te Ao B, McMahon B, Karimkhani C, Yu C, Cooke GS, Schwebel DC, Carpenter DO, Pereira DM, Nash D, Kazi DS, De Leo D, Plass D, Ukwaja KN, Thurston GD, Yun Jin K, Simard EP, Mills E, Park EK, Catalá-López F, deVeber G, Gotay C, Khan G, Hosgood HD 3rd, Santos IS, Leasher JL, Singh J, Leigh J, Jonas JB, Sanabria J, Beardsley J, Jacobsen KH, Takahashi K, Franklin RC, Ronfani L, Montico M, Naldi L, Tonelli M, Geleijnse J, Petzold M, Shrime MG, Younis M, Yonemoto N, Breitborde N, Yip P, Pourmalek F, Lotufo PA, Esteghamati A, Hankey GJ, Ali R, Lunevicius R, Malekzadeh R, Dellavalle R, Weintraub R, Lucas R, Hay R, Rojas-Rueda D, Westerman R, Sepanlou SG, Nolte S, Patten S, Weichenthal S, Abera SF, Fereshtehnejad SM, Shiue I, Driscoll T, Vasankari T, Alsharif U, Rahimi-Movaghar V, Vlassov VV, Marcenes WS, Mekonnen W, Melaku YA, Yano Y, Artaman A, Campos I, MacLachlan J, Mueller U, Kim D, Trillini M, Eshrati B, Williams HC, Shibuya K, Dandona R, Murthy K, Cowie B, Amare AT, Antonio CA, Castañeda-Orjuela C, van Gool CH, Violante F, Oh IH, Deribe K, Soreide

K, Knibbs L, Kereselidze M, Green M, Cardenas R, Roy N, Tillmann T, Li Y, Krueger H, Monasta L, Dey S, Sheikhbahaei S, Hafezi-Nejad N, Kumar GA, Sreeramareddy CT, Dandona L, Wang H, Vollset SE, Mokdad A, Salomon JA, Lozano R, Vos T, Forouzanfar M, Lopez A, Murray C, Naghavi M. The Global Burden of Cancer 2013. JAMA Oncology 2015;1(4): 505-527.

99. Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE. Risks of cancer in BRCA1-mutation carriers: Breast Cancer Linkage Consortium. Lancet 1994;343:692-695.

100. Foster TJ. Immune evasion by staphylococci. Nat Rev Microbiol 2005;3:948-958.

101. Fu YR, Yi ZJ, Guan SZ, Zhang SY, Li M. Proteomic analysis of sputum in patients with active pulmonary tuberculosis. Clin Microbiol Infect 2012;18(12):1241-1247.

102. Fujii J, Ikeda Y. Advances in our understanding of peroxiredoxin, a multifunctional, mammalian redox protein. Redox Rep 2002;7(3):123-130.

103. Fujita Y, Yamaguchi S, Kakino A, Iwamoto S, Yoshimoto R, Sawamura T. Lectin-like Oxidized LDL Receptor 1 Is Involved in CRP-Mediated Complement Activation. Clin Chem 2011;57(10):1398-1405.

104. Furlaneto CJ, Campa A. A novel function of serum amyloid A: a potent stimulus for the release of tumor necrosis factor-alpha, interleukin-1beta, and interleukin-8 by human blood neutrophil. Biochem Biophys Res Commun 2000;268(2):405-408.

105. Fyfe AI, Rothenberg LS, DeBeer FC, Cantor RM, Rotter JI, Lusis AJ. Association between serum amyloid A proteins and coronary artery disease: evidence from two distinct arteriosclerotic processes. Circulation 1997; 96(9):2914-2919.

106. Gabay C, Kushner I. Acute-phase proteins and other systemic responses to inflammation. N Engl J Med 1999;340(6):448-454.

107. Gang Liu G, Paul Cheresh P, David W. Kamp DW. Molecular Basis of Asbestos-Induced Lung Disease. Annu Rev Pathol 2013;8:161-187.

108. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, Greninger P, Thompson IR, Luo X, Soares J, Liu Q, Iorio F, Surdez D, Chen L, Milano RJ, Bignell GR, Tam AT, Davies H, Stevenson JA, Barthorpe S, Lutz SR, Kogera F, Lawrence K, McLaren-Douglas A, Mitropoulos X, Mironenko T, Thi H, Richardson L, Zhou W, Jewitt F, Zhang T, O'Brien P, Boisvert JL, Price S, Hur W, Yang W, Deng X, Butler A, Choi HG, Chang JW, Baselga J, Stamenkovic I, Engelman JA, Sharma SV, Delattre O, Saez-Rodriguez J, Gray NS, Settleman J, Futreal PA, Haber DA, Stratton MR, Ramaswamy S, McDermott U, Benes CH. Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 2012;483(7391):570-575.

109. Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. Nat Rev Cancer 2017;17(12):725-737.

110. Geiger T, Wehner A, Schaab C, Cox J, Mann M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. Mol Cell Proteomics 2012;11(3):M111.

111. Gensch E, Gallup M, Sucher A, Li D, Gebremichael A, Lemjabbar H, Mengistab A, Dasari V, Hotchkiss J, Harkema J, Basbaum C. Tobacco smoke control of mucin production in lung cells requires oxygen radicals AP-1 and JNK. J Biol Chem 2004;279:39085-39093.

112. Geyer PE, Holdt LM, Teupser D, Mann M. Revisiting biomarker discovery by plasma proteomics. Mol Syst Biol 2017;13(9):942.

113. Gharib SA, Nguyen EV, Lai Y, Plampin JD, Goodlett DR, Hallstrand TS. Induced sputum proteome in healthy subjects and asthmatic patients. J Allergy Clin Immunol 2011;128(6):1176-1184.

114. Gilbert DN. Use of plasma procalcitonin levels as an adjunct to clinical microbiology. J Clin Microbiol 2010;48(7):2325-2329.

115. Gohlke RS, McLafferty FW. Early gas chromatography/mass spectrometry. J Am Soc Mass Spectrom 1993;4(5):367-371.

116. Gonçalves RB, Coletta RD, Silvério KG, Benevides L, Casati MZ, da Silva JS, Nociti Jr. FH. Impact of smoking on inflammation: overview of molecular mechanisms. Inflamm Res 2011;60(5):409-424.

117. Gordon AC, Lagan AL, Aganna E, Cheung L, Peters CJ, McDermott MF, Millo JL, Welsh KI, Holloway P, Hitman GA, Piper RD, Garrard CS, Hinds CJ. TNF and TNFR polymorphisms in severe sepsis and septic shock: a prospective multicentre study. Genes Immun 2004;5(8):631-640.

118. Grant SS, Hung DT. Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. Virulence 2013;4(4):273-283.

119. Grimaldi CM, Cleary J, Dagtas AS, Moussai D, Diamond B. Estrogen alters thresholds for B cell apoptosis and activation. J Clin Invest 2002;109:1625e33.

120. Gu W, Zeng L, Zhou J, Jiang DP, Zhang L, Du DY, Hu P, Chen K, Liu Q, Wang ZG, Jiang JX. Clinical relevance of 13 cytokine gene polymorphisms in Chinese major trauma patients. Intensive Care Med 2010;36(7):1261-1265.

121. Gudbjartsson DF, Sulem P, Stacey SN, Goldstein AM, Rafnar T, Sigurgeirsson B, Benediktsdottir KR, Thorisdottir K, Ragnarsson R, Sveinsdottir SG, Magnusson V, Lindblom A, Kostulas K, Botella-Estrada R, Soriano V, Juberías P, Grasa M, Saez B, Andres R, Scherer D, Rudnai P, Gurzau E, Koppova K, Kiemeney LA, Jakobsdottir M, Steinberg S, Helgason A, Gretarsdottir S, Tucker MA, Mayordomo JI, Nagore E, Kumar R, Hansson J, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. Nat Genet 2008;40(7):886-891.

122. Guibert N, Pradines A, Favre G, Mazieres J. Current and future applications of liquid biopsy in nonsmall cell lung cancer from early to advanced stages. Eur Respir Rev 2020;29(155).

123. Gundry RL, Fu Q, Jelinek CA, Van Eyk JE, Cotter RJ. Investigation of an albumin-enriched fraction of human serum and its albuminome. Proteomics Clin Appl 2007;1(1):73-88.
124. Gunning P, O'Neill G, Hardeman E. Tropomyosin-based regulation of the actin cytoskeleton in time and space. Physiol Rev 2008;88:1-35.
125. Haff L, Smirnov I. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and MALDI-TOF mass spectrometry. Genome Res 1997;7:378-388.
126. Hakulinen T, Pukkala E. Future incidence of lung cancer: Forecasts based on hypothetical changes in the smoking habits of males. Int J Epidemiol 1981;10:233-240.
127. Han YH, Kim HS, Kim JM, Kim SK, Yu DY, Moon EY. Inhibitory role of peroxiredoxin II (Prx II) on cellular senescence. FEBS Lett 2005;579:4897-4902.
128. Han X, Aslanian A, Yates JR 3rd. Mass spectrometry for proteomics. Curr Opin Chem Biol 2008;12(5):483-490.
129. Han X, Wang J, Sun Y. Circulating Tumor DNA as Biomarkers for Cancer Detection. Genomics Proteomics Bioinformatics 2017;15(2):59-72.
130. Hanash S, Taguchi A. The grand challenge to decipher the cancer proteome. Nat Rev Cancer 2010;10(9):652-660.
131. Harbarth S, Holeckova K, Froidevaux C, Pittet D, Ricou B, Grau GE, Vadas L, Pugin J, Geneva Sepsis Network. Diagnostic value of procalcitonin, interleukin-6, and interleukin-8 in critically ill patients admitted with suspected sepsis. Am J Respir Crit Care Med 2001;164(3):396-402.
132. Hari-Dass R, Shah C, Meyer DJ, Raynes JG. Serum amyloid A protein binds to outer membrane protein A of gram-negative bacteria. J Biol Chem 2005;280(19):18562-18567.
133. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. Nat Rev Cancer 2003;3(10):733-744.

134. Heinrich PC, Castell JV, Andus T. Interleukin-6 and the acute phase response. Biochem J 1990;265(3):621-636.

135. Heintz NH, Janssen-Heininger YM, Mossman BT. Asbestos, lung cancers, and mesotheliomas: from molecular approaches to targeting tumor survival pathways. Am J Respir Cell Mol Biol 2010;42(2):133-139.

136. Helm O, Held-Feindt J, Grage-Griebenow E, Reiling N, Ungefroren H, Vogel I, Krüger U, Becker T, Ebsen M, Röcken C, Kabelitz D, Schäfer H, Sebens S. Tumor-associated macrophages exhibit pro- and anti-inflammatory properties by which they impact on pancreatic tumorigenesis. Int J Cancer 2014;135(4):843-861.

137. Henderson DW, Leigh J. (2011). Asbestos and carcinoma of the lung. In Asbestos: Risk Assessment, Epidemiology and Health Effects, 2nd ed. CRC Press/Taylor&Francis: FL, USA.

138. Hensel M, Volk T, Döcke WD, Kern F, Tschirna D, Egerer K, Konertz W, Kox WJ. Hyperprocalcitonemia in patients with noninfectious SIRS and pulmonary dysfunction associated with cardiopulmonary bypass. Anesthesiology 1998;89(1):93-104.

139. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJ Jr, Wu YL, Paz-Ares L. Lung cancer: current therapies and new targeted treatments. Lancet 2017;389(10066):299-311.

140. Holmen JM, Karlsson NG, Abdullah LH, Randell SH, Sheehan JK, Hansson GC, Davis CW. Mucins and their O-glycans from human bronchial epithelial cell cultures. Am J Physiol Lung Cell Mol Physiol 2004;287:L824-L834.

141. Horeweg N, van Rosmalen J, Heuvelmans MA, van der Aalst CM, Vliegenthart R, Scholten ET, ten Haaf K, Nackaerts K, Lammers JW, Weenink C, Groen HJ, van Ooijen P, de Jong PA, de Bock GH, Mali W, de Koning HJ, Oudkerk M. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol 2014;15(12):1332-1341.

142. Hoser G, Domagala-Kulawik J, Droszcz P, Droszcz W, Kawiak J. Lymphocyte subsets differences in smokers and nonsmokers with primary lung cancer: a flow cytometry analysis of bronchoalveolar lavage fluid cells. Med Sci Monit 2003;9:BR310-BR315.

143. Hoshikawa S, Mori K, Kaise N, Nakagawa Y, Ito S, Yoshida K. Artifactually elevated serum-free thyroxine levels measured by equilibrium dialysis in a pregnant woman with familial dysalbuminemic hyperthyroxinemia. Thyroid 2004;14:155-160.

144. Howard BA, Wang MZ, Campa MJ, Corro C, Fitzgerald MC, Patz EF Jr. Identification and validation of a potential lung cancer serum biomarker detected by matrix-assisted laser desorption/ionization-time of flight spectra analysis. Proteomics 2003;3:1720-1724.

145. Hudler P, Kocevar N, Komel R. Proteomic approaches in biomarker discovery: new perspectives in cancer diagnostics. Scientific World Journal 2014;2014:260348.

146. Husgafvel-Pursiainen K, Hackman P, Ridanpaa M, Anttila S, Karjalainen A, Partanen T, Taikina-Aho O, Heikkilä L, Vainio H. K-ras mutations in human adenocarcinoma of the lung: association with smoking and occupational exposure to asbestos. Int J Cancer 1993;53: 50-56.

147. Husgafvel-Pursiainen K, Boffetta P, Kannio A, Nyberg F, Pershagen G, Mukeria A, Constantinescu V, Fortes C, Benhamou S. p53 mutations and exposure to environmental tobacco smoke in a multicenter study on lung cancer. Cancer Res 2000;60:2906-2011.

148. Huuskonen MS, Karjalainen A, Tossavainen A, Rantanen J. Asbestos and cancer in Finland. Med Lav 1995;86(5):426-434.

149. Huuskonen MS, Rantanen J. Finnish Institute of Occupational Health (FIOH): prevention and detection of asbestos-related diseases, 1987-2005. Am J Ind Med 2006;49(3):215-20.

150. Hwang JA, Song JS, Yu DY, Kim HR, Park HJ, Park YS, Kim WS, Choi CM. Peroxiredoxin 4 as an independent prognostic marker for survival in patients with early-stage lung squamous cell carcinoma. Int J Clin Exp Pathol 2015;8(6):6627-6635.

151. IASLC, Rami-Porta R. (edit.). (2016). Staging Manual inThoracic Oncology, 2nd ed. Rx Press: FL, USA.

152. Ignatiadis M, Lee M, Jeffrey SS. Circulating tumor cells and circulating tumor DNA: challenges and opportunities on the path to clinical utility. Clin Cancer Res 2015:21:4786-4800.

153. Infante M, Cavuto S, Lutman FR, Passera E, Chiarenza M, Chiesa G, Brambilla G, Angeli E, Aranzulla G, Chiti A, Scorsetti M, Navarria P, Cavina R, Ciccarelli M, Roncalli M, Destro A, Bottoni E, Voulaz E, Errico V, Ferraroli G, Finocchiaro G, Toschi L, Santoro A, Alloisio M; DANTE Study Group. Long-Term Follow-up Results of the DANTE Trial, a Randomized Study of Lung Cancer Screening with Spiral Computed Tomography. Am J Respir Crit Care Med 2015;191(10):1166-1175.

154. Institute of Medicine (2007). Cancer Biomarkers: The Promises and Challenges of Improving Detection and Treatment. Washington, DC: The National Academies Press.

155. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851-862.

156. Jaffe DM, Fleisher GR. Temperature and total white blood cell count as indicators of bacteremia. Pediatrics 1991;87(5):670-674.

157. Jaiswal M, Washburn MP, Zybailov BL. Mass Spectrometry-Based Methods of Proteome Analysis. Rev Cell Biol Mol Med 2015.

158. Jeong S, Park Y, Cho Y, Kim HS. Diagnostic utilities of procalcitonin and C-reactive protein for the prediction of bacteremia determined by blood culture. Clin Chim Acta 2012;413:1731-1736.

159. Jiang H, Wu L, Mishra M, Chawsheen HA, Wei Q. Expression of peroxiredoxin 1 and 4 promotes human lung cancer malignancy. Am J Cancer Res 2014;4(5):445-460.

160. Johnson H, Cohen C, Fatima N, Duncan D, Siddiqui MT. Thyroid transcription factor 1 and Napsin A double stain: utilizing different vendor antibodies for diagnosing lung adenocarcinoma. Acta Cytol 2012;56(6):596-602.

161. Kathiresan S, Larson MG, Vasan RS, Guo CY, Gona P, Keaney JF Jr, Wilson PW, Newton-Cheh C, Musone SL, Camargo AL, Drake JA, Levy D, O'Donnell CJ, Hirschhorn JN, Benjamin EJ. Contribution of clinical correlates and 13 C-reactive protein gene polymorphisms to interindividual variability in serum C-reactive protein level. Circulation 2006;113(11):1415-1423.

162. Kawahara M, Ushijima S, Kamimori T, Kodama N, Ogawara M, Matsui K, Masuda N, Takada M, Sobue T, Furuse K. Second primary tumours in more than 2-year disease-free survivors of small-cell lung cancer in Japan: the role of smoking cessation. Br J Cancer 1998;78:409-412.

163. Kawai H, Minamiya Y, Takahashi N. Prognostic impact of S100A9 overexpression in non-small cell lung cancer. Tumour Biol 2011;32(4):641-646.

164. Kettunen E, Aavikko M, Nymark P, Ruosaari S, Wikman H, Vanhala E, Salmenkivi K, Pirinen R, Karjalainen A, Kuosma E, Anttila S. DNA copy number loss and allelic imbalance at 2p16 in lung cancer associated with asbestos exposure. Br J Cancer 2009;100(8):1336-1342.

165. Khoonsari PE, Häggmark A, Lönnberg M, Mikus M, Kilander L, Lannfelt L, Bergquist J, Ingelsson M, Nilsson P, Kultima K, Shevchenko G. Analysis of the Cerebrospinal Fluid Proteome in Alzheimer's Disease. PLoS One 2016;11(3):e0150672.

166. Khuder SA. Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. Lung Cancer 2001;31(2-3):139-148.

167. Kim AT, Sarafian TA, Shau H. Characterization of antioxidant properties of natural killer-enhancing factor-B and induction of its expression by hydrogen peroxide. Toxicol Appl Pharmacol 1997;147:135-142.

168. Kim HJ, Song EJ, Lee KJ. Proteomic analysis of protein phosphorylations in heat shock response and thermotolerance. J Biol Chem 2002;277(26):23193-23207.

169. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabuddhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang TC, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TS, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H, Pandey A. A draft map of the human proteome. Nature 2014;509(7502):575-581.

170. Kim YR, Oh JE, Kim MS, Kang MR, Park SW, Han JY, Eom HS, Yoo NJ, Lee SH. Oncogenic NRF2 mutations in squamous cell carcinomas of oesophagus and skin. J Pathol 2010;220:446-451.

171. Kirby, R. The role of PSA in detection and management of prostate cancer. Practitioner 2016;260(1792):17-21.

172. Kisluk J, Ciborowski M, Niemira M, Kretowski A, Niklinski J. Proteomics biomarkers for non-small cell lung cancer. J Pharm Biomed Anal 2014;101:40-49.

173. Klebe S, Leigh J, Henderson DW, Nurminen M. Asbestos, Smoking and Lung Cancer: An Update. Int J Environ Res Public Health 2019;17(1).

174. Klein RJ, Halldén C, Cronin AM, Ploner A, Wiklund F, Bjartell AS, Stattin P, Xu J, Scardino PT, Offit K, Vickers AJ, Grönberg H, Lilja H. Blood biomarker levels to aid discovery of cancer-related single-nucleotide polymorphisms: kallikreins and prostate cancer. Cancer Prev Res (Phila) 2010;3(5):611-619.

175. Kluft C, de Maat MP. Genetics of C-reactive protein: new possibilities and complications. Arterioscler Thromb Vasc Biol 2003;23(11):1956-1959.

176. Kolb-Bachofen V. A review on the biological properties of C-reactive protein. Immunology 1991;183:133-145.

177. Krelin Y, Voronov E, Dotan S, Elkabets M, Reich E, Fogel M, Huszar M, Iwakura Y, Segal S, Dinarello CA, Apte RN. Interleukin-1beta-driven inflammation promotes the development and invasiveness of chemical carcinogen-induced tumors. Cancer Res 2007;67(3):1062-1671.

178. Kubica M, Guzik K, Koziel J, Zarebski M, Richter W, Gajkowska B, Golda A, Maciag-Gudowska A, Brix K, Shaw L, Foster T, Potempa J. A potential new pathway for Staphylococcus aureus dissemination: the silent survival of S. aureus phagocytosed by human monocyte-derived macrophages. PLoS One 2008;3:e1409.

179. Kushner I, Rzewnicki D, Samols D. What does minor elevation of C-reactive protein signify? Am J Med 2006;119(2):166.e17-28.

180. Laaksonen DE, Niskanen L, Nyyssönen K, Punnonen K, Tuomainen TP, Valkonen VP, Salonen R, Salonen JT. C-reactive protein and the development of the metabolic syndrome and diabetes in middle-aged men. Diabetologia 2004;47(8):1403-1410.

181. LaDou J. The asbestos cancer epidemic. Environ Health Perspect 2004;112(3):285-290.

182. Lai YH, Wang YS. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry: Mechanistic Studies and Methods for Improving the Structural Identification of Carbohydrates. Mass Spectrom (Tokyo) 2017;6(Spec Iss 2):S0072.

183. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd

C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ,

Szustakowki J; International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921.

184. Lange LA, Carlson CS, Hindorff LA, Lange EM, Walston J, Durda JP, Cushman M, Bis JC, Zeng D, Lin D, Kuller LH, Nickerson DA, Psaty BM, Tracy RP, Reiner AP. Association of polymorphisms in the CRP gene with circulating C-reactive protein levels and cardiovascular events. JAMA 2006;296:2703-2711.

185. Larinkari U. Assay of teichoic acid antibodies and antistaphylolysin in the diagnosis of Staphylococcal osteomyelitis. Scand J Infect Dis 1982;14(2):123-126.

186. Larinkari U, Valmonen M, Sarvas M, Valtonen V. Teichoic Acid Antibody TestIts Use in Patients With Coagulase-Positive Staphylococcal Bacteremia. Arch Intern Med 1977;137(11):1522-1525.

187. Larinkari U, Valtonen VV. Comparison of anti-alpha-haemolysin and teichoic acid antibody tests in patients with Staphylococcus aureus endocarditis or bacteraemia. J Infect 1984;8(3):221-226.

188. Larsen MV, Harboe ZB, Ladelund S, Skov R, Gerstoft J, Pedersen C, Larsen CS, Obel N, Kronborg G, Benfield T. Major but differential decline in the incidence of Staphylococcus aureus bacteraemia in HIV-infected individuals from 1995 to 2007: a nationwide cohort study. HIV Med 2012;13:45-53.

189. Larsen FF, Petersen JA. Novel biomarkers for sepsis: A narrative review. Eur J Intern Med 2017;45:46-50.

190. Laskin DL, Pendino KJ. Macrophages and inflammatory mediators in tissue injury. Annu Rev Pharmacol Toxicol 1995;35:655-677.

191. Latosinska A, Frantzi M, Vlahou A, Merseburger AS, Mischak H. Clinical Proteomics for Precision Medicine: The Bladder Cancer Case. Proteomics Clin Appl 2018;12(2).

192. Laupland KB, Lyytikainen O, Sogaard M, Kennedy KJ, Knudsen JD, Ostergaard C, Galbraith JC, Valiquette L, Jacobsson G, Collignon P, Schonheyder HC, International Bacteremia Surveillance Collaborative.

The changing epidemiology of Staphylococcus aureus bloodstream infection: a multinational population-based surveillance study. Clin Microbiol Infect 2013;19:465-471.

193. Lavie M, Voisset C, Vu-Dac N, Zurawski V, Duverlie G, Wychowski C, Dubuisson J. Serum amyloid A has antiviral activity against hepatitis C virus by inhibiting virus entry in a cell culture system. Hepatology 2006;44(6):1626-1634.

194. Le JM, Vilcek J. Interleukin 6: a multifunctional cytokine regulating immune reactions and the acute phase protein response. Lab Invest 1989;61(6):588-602.

195. Lee CC, Chen SY, Tsai CL, Wu SC, Chiang WC, Wang JL, Sun HY, Chen SC, Chen WJ, Hsueh PR. Prognostic value of mortality in emergency department sepsis score, procalcitonin, and C-reactive protein in patients with sepsis at the emergency department. Shock 2008;29(3):322-327.

196. Lehtonen ST, Svensk AM, Soini Y, Pääkkö P, Hirvikoski P, Kang SW, Säily M, Kinnula VL. Peroxiredoxins, a novel protein family in lung cancer. Int J Cancer 2004;111(4):514-521.

197. Lennicke C, Rahn J, Lichtenfels R, Wessjohann LA, Seliger B. Hydrogen peroxide - production, fate and role in redox signaling of tumor cells. Cell Commun Signal 2015;13:39.

198. Leon G, MacDonagh L, Finn SP, Cuffe S, Barr MP. Cancer stem cells in drug resistant lung cancer: Targeting cell surface markers and signaling pathways. Pharmacol Ther 2016;158:71-90.

199. Li Y, Zhang B, Wang X, Yan H, Chen G, Zhang X. Proteomic analysis of apoptosis induction in human lung cancer cells by recombinant MVL. Amino Acids 2011;41(4):923-932.

200. Li QK, Shah P, Li Y, Aiyetan PO, Chen J, Yung R, Molena D, Gabrielson E, Askin F, Chan DW, Zhang H. Glycoproteomic analysis of bronchoalveolar lavage (BAL) fluid identifies tumor-associated glycoproteins from lung adenocarcinoma. J Proteome Res 2013;12(8):3689-3696.

201. Liang JS, Sipe JD. Recombinant human serum amyloid A (apoSAAp) binds cholesterol and modulates cholesterol flux. J Lipid Res 1995;36(1):37-46.

202. Liang JS, Sloane JA, Wells JM, Abraham CR, Fine RE, Sipe JD. Evidence for local production of acute phase response apolipoprotein serum amyloid A in Alzheimer's disease brain. Neurosci Lett 1997; 225(2):73-76.

203. Lin D, Hollander Z, Ng RT, Imai C, Ignaszewski A, Balshaw R, Freue GC, Wilson-McManus JE, Qasimi P, Meredith A, Mui A, Triche T, McMaster R, Keown PA, McManus BM; Biomarkers in Transplantation Team; NCE CECR Centre of Excellence for the Prevention of Organ Failure. Whole blood genomic biomarkers of acute cardiac allograft rejection. J Heart Lung Transplant 2009;28(9):927-935.

204. Liu GY. Molecular pathogenesis of Staphylococcus aureus infection. Pediatr Res 2009;65(5 Pt 2):71R-77R.

205. Liu Y, Hou J, Li Q, Chen K, Wang S-N, Wang J. Biomarkers for diagnosis of sepsis in patients with systemic inflammatory response syndrome: a systematic review and meta-analysis. Springerplus 2016;5(1):2091.

206. Llewelyn M, Cohen J. Superantigens: microbial agents that corrupt immunity. Lancet Infect Dis 2002;2:156-162.

207. Lobo SM, Lobo FR, Bota DP, Lopes-Ferreira F, Soliman HM, Mélot C, Vincent JL. C-reactive protein levels correlate with mortality and organ failure in critically ill patients. Chest 2003;123(6):2043-2049.

208. Lohani M, Dopp E, Becker HH, Seth K, Schiffmann D, Rahman Q. Smoking enhances asbestos-induced genotoxicity, relative involvement of chromosome 1: A study using multicolor fish with tandem labeling. Toxicol Lett 2002;136(1):55-63.

209. Longphre M, Li D, Li J, Matovinovic E, Gallup M, Samet JM, Basbaum CB. Lung mucin production is stimulated by the air pollutant residual oil fly ash. Tox Appl Pharmacol 2000;162:86-92.

210. Lopes Pegna A, Picozzi G, Falaschi F, Carrozzi L, Falchini M, Carozzi FM, Pistelli F, Comin C, Deliperi A, Grazzini M, Innocenti F, Maddau C, Vella A, Vaggelli L, Paci E, Mascalchi M; ITALUNG Study Research Group. Four-year results of low-dose CT screening and nodule management in the ITALUNG trial. J Thorac Oncol 2013;8(7):866-875.

211. Lora JM, Zhang DM, Liao SM, Burwell T, King AM, Barker PA, Singh L, Keaveney M, Morgenstern J, Gutiérrez-Ramos JC, Coyle AJ, Fraser CC. Tumor necrosis factor-α triggers mucus production in airway epithelium through an IκB kinase β-dependent mechanism. J Biol Chem 2005;280:36510-36517.

212. Lorente L, Martín MM, Pérez-Cejas A, Barrios Y, Solé-Violán J, Ferreres J, Labarta L, Díaz C, Jiménez A. Association between Interleukin-6 Promoter Polymorphism (-174 G/C), Serum Interleukin-6 Levels and Mortality in Severe Septic Patients. Int J Mol Sci 2016;17(11).

213. Lowry LK. Role of biomarkers of exposure in the assessment of health risks. Toxicol Lett 1995;77(1-3):31-38.

214. Lowy FD. Staphylococcus aureus infections. N Engl J Med 1998;339(8):520-532.

215. Lunde ML, Roman E, Warnakulasuriya S, Mehrotra R, Laranne J, Vasstrand EN, Ibrahim SO. Profiling of chromosomal changes in potentially malignant and malignant oral mucosal lesions from south and south-east Asia using array-comparative genomic hybridization. Cancer Genomics Proteomics 2014;11:127-140.

216. Lyytikainen O, Ruotsalainen E, Jarvinen A, Valtonen V, Ruutu P. Trends and outcome of nosocomial and community-acquired bloodstream infections due to Staphylococcus aureus in Finland, 1995-2001. Eur J Clin Microbiol Infect Dis 2005;24:399-404.

217. Ma P, Chen D, Pan J, Du B. Genomic polymorphism within interleukin-1 family cytokines influences the outcome of septic patients. Crit Care Med 2002;30(5):1046-1050.

218. Ma Y, Fan M, Dai L, Kang X, Liu Y, Sun Y, Yan W, Liang Z, Xiong H, Chen K. The expression of TTF-1 and Napsin A in early-stage lung

adenocarcinoma correlates with the results of surgical treatment. Tumour Biol 2015;36(10):8085-8092.

219. Majetschak M, Flohé S, Obertacke U, Schröder J, Staubach K, Nast-Kolb D, Schade FU, Stüber F. Relation of a TNF gene polymorphism to severe sepsis in trauma patients. Ann Surg 1999;230(2):207-214.

220. Manser R, Lethaby A, Irving LB, Stone C, Byrnes G, Abramson MJ, Campbell D. Screening for lung cancer. Cochrane Database Syst Rev 2013;6:CD001991.

221. Marchant A, Alegre ML, Hakim A, Piérard G, Marécaux G, Friedman G, De Groote D, Kahn RJ, Vincent JL, Goldman M. Clinical and biological significance of interleukin-10 plasma levels in patients with septic shock. J Clin Immunol 1995;15(5):266-273.

222. Marnell L, Mold C, Du Clos TW. C-reactive protein: ligands, receptors and role in inflammation. Clin Immunol 2005;117(2):104-111.

223. Marshall JR, Hammond CB, Ross GT, Jacobson A, Rayford P, Odell WD. Plasma and urinary chorionic gonadotropin during early human pregnancy. Obstet Gynecol 1968;32(6):760-764.

224. Marsik C, Sunder-Plassmann R, Jilma B, Kovar FM, Mannhalter C, Wagner O, Rumpold H, Endler G. The C-reactive protein (+)1444C/T alteration modulates the inflammation and coagulation response in human endotoxemia. Clin Chem 2006;52(10):1952-1957.

225. Marsit CJ, Hasegawa M, Hirao T, Kim DH, Aldape K, Hinds PW, Wiencke JK, Nelson H, Kelsey KT. Loss of heterozygocity of chromosome 3p21 is associated with mutant Tp53 and better patient survival in non-small-cell lung cancer. Cancer Res 2004;64(23):8702-8707.

226. Martin AM, Weber BL. Genetic and hormonal risk factors in breast cancer. J Natl Cancer Inst 2000;92(14):1126-1135.

227. Matthews LW, Spector S, Lemm J, Potter JL. Studies on pulmonary secretions: I. The over-all chemical composition of pulmonary secretions from patients with cystic fibrosis, bronchiectasis, and laryngectomy. Am Rev Respir Dis 1963;88:199-204.

228. Matthews MJ, Kanhouwa S, Pickren J, Robinette D. Frequency of residual and metastatic tumor in patients undergoing curative surgical resection for lung cancer. Cancer Chemother Rep 1973;3(4):63-67.
229. McCabe WR, Jackson GG. Gram negative bacteraemia 1. Etiology and ecology. Arch Intern Med 1962;110:847-855.
230. McGuire W, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D. Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. Nature 1994;371(6497):508-510.
231. McGuire SE, McGuire AL. Don't throw the baby out with the bathwater: enabling a bottom-up approach in genome-wide association studies. Genome Res 2008;18(11):1683-1685.
232. Meek RL, Urieli-Shoval S, Benditt EP. Expression of apolipoprotein serum amyloid A mRNA in human atherosclerotic lesions and cultured vascular cells: implications for serum amyloid A function. Proc Natl Acad Sci U S A 1994; 91(8):3186-3190.
233. Meisner M, Tschaikowsky K, Hutzler A, Schick C, Schüttler J. Postoperative plasma concentrations of procalcitonin after different types of surgery. Intensive Care Med 1998;24(7):680-684.
234. Meisner M. Pathobiochemistry and clinical use of procalcitonin. Clin Chim Acta 2002;323(1-2):17-29.
235. Menegon S, Columbano A, Giordano S. The Dual Roles of NRF2 in Cancer. Trends Mol Med 2016;22(7):578-593.
236. Merikallio H, Pääkkö P, Kinnula VL, Harju T, Soini Y. Nuclear factor erythroid-derived 2-like 2 (Nrf2) and DJ1 are prognostic factors in lung cancer. Hum Pathol 2012;43(4):577-584.
237. Mischak H, Allmaier G, Apweiler R, Attwood T, Baumann M, Benigni A, Bennett SE, Bischoff R, Bongcam-Rudloff E, Capasso G, Coon JJ, D'Haese P, Dominiczak AF, Dakna M, Dihazi H, Ehrich JH, Fernandez-Llama P, Fliser D, Frokiaer J, Garin J, Girolami M, Hancock WS, Haubitz M, Hochstrasser D, Holman RR, Ioannidis JP, Jankowski J, Julian BA, Klein JB, Kolch W, Luider T, Massy Z, Mattes WB, Molina F, Monsarrat B, Novak J, Peter K, Rossing P, Sánchez-Carbayo M, Schanstra JP,

Semmes OJ, Spasovski G, Theodorescu D, Thongboonkerd V, Vanholder R, Veenstra TD, Weissinger E, Yamamoto T, Vlahou A. Recommendations for biomarker identification and qualification in clinical proteomics. Sci Transl Med 2010;2(46):46ps42.

238. Miller DT, Zee RY, Suk Danik J, Kozlowski P, Chasman DI, Lazarus R, Cook NR, Ridker PM, Kwiatkowski DJ. Association of common CRP gene variants with CRP levels and cardiovascular events. Ann Hum Genet 2005;69:623-638.

239. Moon JC, Hah YS, Kim WY, Jung BG, Jang HH, Lee JR, Kim SY, Lee YM, Jeon MG, Kim CW, Cho MJ, Lee SY. Oxidative stress-dependent structural and functional switching of a human 2-Cys peroxiredoxin isotype II that enhances HeLa cell resistance to H2O2-induced cell death. J Biol Chem 2005;280:28775-28784.

240. Moschini M, Spahn M, Mattei A, Cheville J, Karnes RJ. Incorporation of tissue-based genomic biomarkers into localized prostate cancer clinics. BMC Med 2016;14:67.

241. Mossman BT, Churg A. Mechanisms in the pathogenesis of asbestosis and silicosis. Am J Respir Crit Care Med 1998;157(5 Pt 1):1666-1680.

242. Mossman BT, Lippmann M, Hesterberg TW, Kelsey KT, Barchowsky A, Bonner JC. Pulmonary endpoints (lung carcinoma and asbestosis) following inhalation exposure to asbestos. J Toxicol Environ Health 2011;14:76-121.

243. Moya F, Nieto A, Candela JL. Calcitonin biosynthesis: evidence for a precursor. Eur J Biochem 1975;55:407-413.

244. Murphy FA, Schinwald A, Poland CA, Donaldson K. The mechanism of pleural inflammation by long carbon nanotubes: interaction of long fibres with macrophages stimulates them to amplify pro-inflammatory responses in mesothelial cells. Part Fibre Toxicol 2012;9:8.

245. Mutsaers SE. Mesothelial cells: their structure, function and role in serosal repair. Respirology 2002;7:171-191.

246. Mäbert K, Cojoc M, Peitzsch C, Kurth I, Souchelnytskyi S, Dubrovska A. Cancer biomarker discovery: current status and future perspectives. Int J Radiat Biol 2014;90(8):659-677.

247. Mäki-Nevala S, Sarhadi VK, Rönty M, Kettunen E, Husgafvel-Pursiainen K, Wolff H, Knuuttila A, Knuutila S. Hot spot mutations in Finnish non-small cell lung cancers. Lung Cancer 2016;99:102-110.

248. Mölkänen T, Ruotsalainen E, Thorball CW, Järvinen A. Elevated soluble urokinase plasminogen activator receptor (suPAR) predicts mortality in Staphylococcus aureus bacteremia. Eur J Clin Microbiol Infect Dis 2011;30(11):1417-1424.

249. Mölkänen T, Ruotsalainen E, Rintala EM, Järvinen A. Predictive Value of C-Reactive Protein (CRP) in Identifying Fatal Outcome and Deep Infections in Staphylococcus aureus Bacteremia. PLoS One 2016;11(5):e0155644.

250. Nadler WM, Waidelich D, Kerner A, Hanke S, Berg R, Trumpp A, Rösli C. MALDI versus ESI: The Impact of the Ion Source on Peptide Identification. J Proteome Res 2017;16(3):1207-1215.

251. Nanjappa V, Thomas JK, Marimuthu A, Muthusamy B, Radhakrishnan A, Sharma R, Ahmad Khan A, Balakrishnan L, Sahasrabuddhe NA, Kumar S, Jhaveri BN, Sheth KV, Kumar Khatana R, Shaw PG, Srikanth SM, Mathur PP, Shankar S, Nagaraja D, Christopher R, Mathivanan S, Raju R, Sirdeshmukh R, Chatterjee A, Simpson RJ, Harsha HC, Pandey A, Prasad TS. Plasma Proteome Database as a resource for proteomics research: 2014 update. Nucleic Acids Res 2014;42(Database issue):D959-965.

252. National Research Council, Committee on Biological Markers. Biological markers in environmental health research. Environ Health Perspect 1987;77:3-9.

253. Nelson H, Kelsey K. The molecular epidemiology of asbestos and tobacco in lung cancer. Oncogene 2002;21(48):7284-7288.

254. Netto LE, Antunes F. The Roles of Peroxiredoxin and Thioredoxin in Hydrogen Peroxide Sensing and in Signal Transduction. Mol Cells 2016;39(1):65-71.

255. Neumann CA, Fang Q. Are peroxiredoxins tumor suppressors? Curr Opin Pharmacol 2007;7:375-380.

256. Nizri E, Irony-Tur-Sinai M, Lory O, Orr-Urtreger A, Lavi E, Brenner T. Activation of the cholinergic anti-inflammatory system by nicotine attenuates neuroinflammation via suppression of Th1 and Th17 responses. J Immunol 2009;183:6681-6688.

257. Nylen ES, O´Neil W, Jordan MH, Snider RH, Moore CF, Lewis M, Silva OL, Becker KL. Serum procalcitonin as index of inhalation injury in burns. Horm Metab Res 1992;24:439-443.

258. Nymark P, Kettunen E, Aavikko M, Ruosaari S, Kuosma E, Vanhala E, Salmenkivi K, Pirinen R, Karjalainen A, Knuutila S, Wikman H, Anttila S. Molecular alterations at 9q33.1 and polyploidy in asbestos-related lung cancer. Clin Cancer Res 2009;15(2):468-475.

259. Nymark P, Wikman H, Ruosaari S, Hollmén J, Vanhala E, Karjalainen A, Anttila S, Knuutila S. Identification of specific gene copy number changes in asbestos-related lung cancer. Cancer Res 2006;66(11):5737-5743.

260. Oberholzer A, Steckholzer U, Kurimoto M, Trentz O, Ertel W. Interleukin-18 plasma levels are increased in patients with sepsis compared to severely injured patients. Shock 2001;16(6):411-414.

261. Obisesan TO, Leeuwenburgh C, Phillips T, Ferrell RE, Phares DA, Prior SJ, Hagberg JM. C-reactive protein genotypes affect baseline, but not exercise training-induced changes, in C-reactive protein levels. Arterioscler Thromb Vasc Biol 2004;24(10):1874-1879.

262. Ogawa M. Differentiation and proliferation of hematopoietic stem cells. Blood 1993;81:2844-2853.

263. Oh JW, Kim SK, Cho KC, Kim MS, Suh CS, Lee JR, Kim KP. Proteomic analysis of human follicular fluid in poor ovarian responders during in vitro fertilization. Proteomics 2017;17(6).

264. O'Hara R, Murphy EP, Whitehead AS, FitzGerald O, Bresnihan B. Acute-phase serum amyloid A production by rheumatoid arthritis synovial tissue. Arthritis Res 2000;2(2):142-144.

265. Ohta T, Iijima K, Miyamoto M, Nakahara I, Tanaka H, Ohtsuji M, Suzuki T, Kobayashi A, Yokota J, Sakiyama T, Shibata T, Yamamoto M, Hirohashi S. Loss of Keap1 function activates Nrf2 and provides advantages for lung cancer cell growth. Cancer Res 2008;68(5):1303-1309.

266. Olson JE, Bielinski SJ, Ryu E, Winkler EM, Takahashi PY, Pathak J, Cerhana JR. Biobanks and personalized medicine. Clin Genet 2014;86(1):50-55.

267. Ooi A, Dykema K, Ansari A, Petillo D, Snider J, Kahnoski R, Anema J, Craig D, Carpten J, Teh BT, Furge KA. CUL3 and NRF2 mutations confer an NRF2 activation phenotype in a sporadic form of papillary renal cell carcinoma. Cancer Res 2013;73(7):2044-2051.

268. Ordonez NG. Value of thyroid transcription factor-1 immunostaining in distinguishing small cell lung carcinomas from other small cell carcinomas. Am J Surg Pathol 2000;24:1217-1223.

269. Ostrowski SR, Piironen T, Hoyer-Hansen G, Gerstoft J, Pedersen BK, Ullum H. High plasma levels of intact and cleaved soluble urokinase receptor reflect immune activation and are independent predictors of mortality in HIV-1-infected patients. J Acquir Immune Defic Syndr 2005;39:23-31.

270. Pan Q, Wang Y, Chen J, Xu G, Chen B, Pan J, Huang K. Investigation of the epidermal growth factor receptor mutation rate in non-small cell lung cancer patients and the analysis of associated risk factors using logistic regression. Oncol Lett. 2014;8(2):813-818.

271. Pankow JS, Folsom AR, Cushman M, Borecki IB, Hopkins PN, Eckfeldt JH, Tracy RP. Familial and genetic determinants of systemic markers of inflammation: the NHLBI family heart study. Atherosclerosis 2001;154(3):681-689.

272. Park JH, Kim YS, Lee HL, Shim JY, Lee KS, Oh YJ, Shin SS, Choi YH, Park KJ, Park RW, Hwang SC. Expression of peroxiredoxin and thioredoxin in human lung cancer and paired normal lung. Respirology 2006;11(3):269-275.

273. Patel RT, Deen KI, Youngs D, Warwick J, Keighley MR. Interleukin 6 is a prognostic indicator of outcome in severe intra-abdominal sepsis. Br J Surg 1994;81(9):1306-1308.

274. Patel, RT. MALDI-TOF MS for the Diagnosis of Infectious Diseases. Clin Chem 2015;61(1):100-111.

275. Pepys MB, Hirschfield GM. C-reactive protein: a critical update. J Clin Invest 2003;111(12):1805-1812.

276. Peters T Jr. Serum albumin. Adv Protein Chem 1985;37:161-245.

277. Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. Oncogene 2002;21:7435-7451.

278. Pierrakos C, Vincent JL. Sepsis biomarkers: a review. Crit Care 2010;14(1):R15.

279. Pin I, Gibson PG, Kolendowicz R, Girgis-Gabardo A, Denburg JA, Hargreave FE, Dolovich J. Use of induced sputum cell counts to investigate airway inflammation in asthma. Thorax 1992 ;47(1):25-29.

280. Pociot F, Molvig J, Wogensen L, Worsaae H, Nerup J. A TaqI polymorphism in the human interleukin-1 beta (IL-1 beta) gene correlates with IL-1 beta secretion in vitro. Eur J Clin Invest 1992;22(6):396-402.

281. Pryor WA, Stone K. Oxidants in cigarette smoke: radicals, hydrogen peroxide, peroxynitrite and peroxynitrite. Ann N Y Acad Sci 1993;686:12-27.

282. Pukkala E, Rautalahti M. Cancer in Finland. Cancer Society of Finland Publication No. 86 (2013).

283. Puzone R, Savarino G, Salvi S, Dal Bello MG, Barletta G, Genova C, Rijavec E, Sini C, Esposito AI, Ratto GB, Truini M, Grossi F, Pfeffer U. Glyceraldehyde-3-phosphate dehydrogenase gene over expression

correlates with poor prognosis in non small cell lung cancer patients. Mol Cancer 2013;12(1):97.

284. Reid A, de Klerk N, Ambrosini GL, Olsen N, Pang SC, Berry G, Musk AW. The effect of asbestosis on lung cancer risk beyond the dose related effect of asbestos alone. Occup Environ Med 2005;62(12):885-889.

285. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 2008;371:569-578.

286. Reynolds JA, Tanford C. Possible implications for the state of proteins in biological membranes. Proc Natl Acad Sci U S A 1970;66(3):1002-1007.

287. Ridker PM, Rifai N, Rose L, Buring JE, Cook NR. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. N Engl J Med 2007;347:1557-1565.

288. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. Cancer Prev Res 2012;5:887-900.

289. Rodgman A., Perfetti TA. (2008). The chemical components of tobacco and tobacco smoke. CRC Press; FL, USA.

290. Rodi CP, Darnhofer-Patel B, Stanssens P, Zabeau M, van den Boom D. A strategy for the rapid discovery of disease markers using the MassARRAY system. BioTechniques 2002;Suppl 62-66:68-69.

291. Rodríguez-Suárez E, Siwy J, Zürbig P, Mischak H. Urine as a source for clinical proteome analysis: from discovery to clinical application. Biochim Biophys Acta 2014;1844(5):884-898.

292. Rogowska-Wrzesinska A, Le Bihan MC, Thaysen-Andersen M, Roepstorff P. 2D gels still have a niche in proteomics. J Proteomics 2013;88:4-13.

293. Rose MC, Voynow JA. Respiratory tract mucin genes and mucin glycoproteins in health and disease. Physiol Rev 2006;86:245-278.

294. Ross P, Hall L, Smirnov I, Haff L. High level multiplex genotyping by MALDI-TOF mass spectrometry. Nat Biotechnol 1998;16:1347-1351.

295. Ross P, Hall L, Haff LA. Quantitative approach to single-nucleotide polymorphism analysis using MALDI-TOF mass spectrometry. Biotechniques 2000;29:620-629.

296. Rudin CM, Poirier JT. Small-cell lung cancer in 2016: shining light on novel targets and therapies. Nat Rev Clin Oncol 2017;14:75-76.

297. Ruosaari ST, Nymark PE, Aavikko MM, Kettunen E, Knuutila S, Hollmén J, Norppa H, Anttila SL. Aberrations of chromosome 19 in asbestos-associated lung cancer and in asbestos-induced micronuclei of bronchial epithelial cells in vitro. Carcinogenesis 2008;29(5):913-917.

298. Ruotsalainen E, Järvinen A, Koivula I, Kauma H, Rintala E, Lumio J, Kotilainen P, Vaara M, Nikoskelainen J, Valtonen V; Finlevo Study Group. Levofloxacin does not decrease mortality in Staphylococcus aureus bacteraemia when added to the standard treatment: a prospective and randomized clinical trial of 381 patients. J Intern Med 2006;259(2):179-190.

299. Ruuls SR, Sedgwick JD. Unlinking tumor necrosis factor biology from the major histocompatibility complex: lessons from human genetics and animal models. Am J Hum Genet 1999;65(2):294-301.

300. Sanli O, Dobruch J, Knowles MA, Burger M, Alemozaffar M, Nielsen ME, Lotan Y. Bladder cancer. Nat Rev Dis Primers 2017;3:17022.

301. Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, Karlsson A, Al-Lazikani B, Hersey A, Oprea TI, Overington JP. A comprehensive map of molecular drug targets. Nat Rev Drug Discov 2017;16(1):19-34.

302. Santtila S, Savinainen K, Hurme M. Presence of the IL-1RA allele 2 (IL1RN*2) is associated with enhanced IL-1beta production in vitro. Scand J Immunol 1998;47(3):195-198.

303. Sasaki H, Okuda K, Endo K, Kawano O, Yukiue H, Yokoyama T, Yano M, Fujii Y. CCND1 messenger RNA expression is correlated with EGFR mutation status in lung cancer. Clin Lung Cancer 2007;8(8):493-496.

304. Satoh H, Moriguchi T, Takai J, Ebina M, Yamamoto M. Nrf2 prevents initiation but accelerates progression through the Kras signaling pathway during lung carcinogenesis. Cancer Res 2013;73(13):4158-4168.

305. Schinwald A, Murphy FA, Prina-Mello A, Poland CA, Byrne F, Movia D, Glass JR, Dickerson JC, Schultz DA, Jeffree CE, Macnee W, Donaldson K. The threshold length for fiber-induced acute pleural inflammation: shedding light on the early events in asbestos-induced mesothelioma. Toxicol Sci 2012;128(2):461-470.

306. Schneider CP, Yilmaz Y, Kleespies A, Jauch KW, Hartl WH. Accuracy of procalcitonin for outcome prediction in unselected postoperative critically ill patients. Shock 2009;31(6):568-573.

307. Scriver CR, Eisensmith RC, Woo SL, Kaufman S. The hyperphenylalaninemias of man and mouse. Annu Rev Genet 1994;28:141-165.

308. Seal RL, Gordon SM, Lush MJ, Wright MW, Bruford EA. Genenames.org: the HGNC resources in 2011. Nucleic Acids Res 2011;39:D514-D519.

309. Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V. WHO International Agency for Research on Cancer Monograph Working Group, A review of human carcinogens—Part E: Tobacco, areca nut, alcohol, coal smoke, and salted fish. Lancet Oncol 2009;10:1033-1034.

310. Selikoff I, Hammond EC, Churg J. Asbestos exposure, smoking, and neoplasia. JAMA 1968;204:106-112.

311. Seigel TA, Cocchi MN, Salciccioli J, Shapiro NI, Howell M, Tang A, Donnino MW. Inadequacy of temperature and white blood cell count in predicting bacteremia in patients with suspected infection. J Emerg Med 2012;42(3):254-259.

312. Schmit X, Vincent JL. The time course of blood C-reactive protein concentrations in relation to the response to initial antimicrobial therapy in patients with sepsis. Infection 2008;36(3):213-219.

313. Shah C, Hari-Dass R, Raynes JG. Serum amyloid A is an innate immune opsonin for Gram-negative bacteria. Blood 2006;108(5):1751-1757.

314. Shapiro AL, Viñuela E, Maizel JV Jr. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. Biochem Biophys Res Commun 1967;28(5):815-820.

315. Shibata T, Ohta T, Tong KI, Kokubu A, Odogawa R, Tsuta K, Asamura H, Yamamoto M, Hirohashi S. Cancer related mutations in NRF2 impair its recognition by Keap1-Cul3 E3 ligase and promote malignancy. Proc Natl Acad Sci U S A 2008;105(36):13568-13573.

316. Shimada T, Oda S, Sadahiro T, Nakamura M, Hirayama Y, Watanabe E, Abe R, Nakada T-a, Tateishi Y, Otani S, Hirasawa H, Tokuhisa T, Uno H. Outcome prediction in sepsis combined use of genetic polymorphisms - A study in Japanese population. Cytokine 2011;54(1):79-84.

317. Shukla A, Gulumian M, Hei TK, Kamp D, Rahman Q, Mossman BT. Multiple roles of oxidants in the pathogenesis of asbestos-induced diseases. Free Radic Biol Med 2003;34(9):1117-1129.

318. Sies H. Oxidative stress: oxidants and antioxidants. Exp Physiol 1997;82:291-295.

319. Singh A, Misra V, Thimmulappa RK, Lee H, Ames S, Hoque MO, Herman JG, Baylin SB, Sidransky D, Gabrielson E, Brock MV, Biswal S. Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer. PLoS Med 2006;3(10):e420.

320. Skogberg K, Lyytikäinen O, Ollgren J, Nuorti JP, Ruutu P. Population-based burden of bloodstream infections in Finland. Clin Microbiol Infect 2012;18(6):E170-176.

321. Skold CM, Lundahl J, Hallden G, Hallgren M, Eklund A. Chronic smoke exposure alters the phenotype pattern and the metabolic response in human alveolar macrophages. Clin Exp Immunol 1996;106:108-113.

322. Slingluff CL Jr, Colella TA, Thompson L, Graham DD, Skipper JCA, Caldwell J, Brinckerhoff L, Kittlesen DJ, Deacon DH, Oei C, Harthun NL,

Huczko EL, Hunt DF, Darrow TL, Engelhard VH. Melanomas with concordant loss of multiple melanocytic differentiation proteins: immune escape that may be overcome by targeting unique or undefined antigens. Cancer Immunol Immunother 2000;48,661-672.

323. Sloane AJ , Lindner RA, Prasad SS, Sebastian LT, Pedersen SK, Robinson M, Bye PT, Nielson DW, Harry JL. Proteomic Analysis of Sputum from Adults and Children with Cystic Fibrosis and from Control Subjects. Am J Respir Crit Care Med 2005;172(11):1416-1426.

324. Smith MR, Kinmonth AL, Luben RN, Bingham S, Day NE, Wareham NJ, Welch A, Khaw KT. Smoking status and differential white cell count in men and women in the EPIC-Norfolk population. Atherosclerosis 2003;169(2):331-337.

325. Spear BB, Heath-Chiozzi ME, Huff J. Clinical application of pharmacogenetics. Trends Mol Med 2001;7(5):201-204.

326. Srivastava S, Gopal-Srivastava R. Biomarkers in cancer screening: a public health perspective. J Nutr 2002;132(8 Suppl):2471S-2475S.

327. Strazisar M, Rott T, Glavac D. Frequent polymorphic variations but rare tumour specific mutations of the S100A2 on 1q21 in non-small cell lung cancer. Lung Cancer 2009;63:354-359.

328. Stuelten CH, Parent CA, Montell DJ. Cell motility in cancer invasion and metastasis: insights from simple model organisms. Nat Rev Cancer 2018;18(5):296-312.

329. Stüber F, Udalova IA, Book M, Drutskaya LN, Kuprash DV, Turetskaya RL, Schade FU, Nedospasov SA. - 308 tumor necrosis factor (TNF) polymorphism is not associated with survival in severe sepsis and is unrelated to lipopolysaccharide inducibility of the human TNF promoter. J Inflamm 1995-1996; 46(1):42-50.

330. Stüber F, Petersen M, Bokelmann F, Schade U. A genomic polymorphism within the tumor necrosis factor locus influences plasma tumor necrosis factor-alpha concentrations and outcome of patients with severe sepsis. Crit Care Med 1996;24(3):381-384.

331. Suk Danik J, Chasman DI, Cannon CP, Miller DT, Zee RY, Kozlowski P, Kwiatkowski DJ, Ridker PM. Influence of genetic variation in the C-reactive protein gene on the inflammatory response during and after acute coronary ischemia. Ann Hum Genet 2006;70(Pt 6):705-716.

332. Sunde M, Pham CLL, Kwan AH. Molecular Characteristics and Biological Functions of Surface-Active and Surfactant Proteins. Annu Rev Biochem 2017;86:585-608.

333. Suojalehto H, Kinaret P, Kilpeläinen M, Toskala E, Ahonen N, Wolff H, Alenius H, Puustinen A. Level of Fatty Acid Binding Protein 5 (FABP5) Is Increased in Sputum of Allergic Asthmatics and Links to Airway Remodeling and Inflammation. PLoS One 2015;10(5).

334. Suomen Syöpärekisteri. (2019). Syöpä. https://syoparekisteri.fi/assets/files/2019/10/sy%C3%B6p%C3%A42017_raportti.pdf [accessed 12.3.2020]

335. Tahbaz R, Schmid M, Merseburger AS. Prevention of kidney cancer incidence and recurrence: lifestyle, medication and nutrition. Curr Opin Urol 2018;28(1):62-79.

336. Takenaga K, Nakamura Y, Sakiyama S. Differential expression of a tropomyosin isoform in low- and high-metastatic Lewis lung carcinoma cells. Mol Cell Biol 1988;8:3934-3937.

337. Takeuchi M, Nagai S, Nakajima A, Shinya M, Tsukano C, Asada H, et al. Inhibition of lung natural killer cell activity by smoking: the role of alveolar macrophages. Respiration 2001;68:262-267.

338. Takeyama K, Dabbagh K, Jeong Shim J, Dao-Pick T, Ueki IF, Nadel JA. Oxidative stress causes mucin synthesis via transactivation of epidermal growth factor receptor: role of neutrophils. J Immunol 2000;164:1546-1552.

339. Tam KW, Zhang W, Soh J, Stastny V, Chen M, Sun H, Thu K, Rios JJ, Yang C, Marconett CN, Selamat SA, Laird-Offringa IA, Taguchi A, Hanash S, Shames D, Ma X, Zhang MQ, Lam WL, Gazdar A. CDKN2A/p16 inactivation mechanisms and their relationship to smoke

exposure and molecular features in non-small-cell lung cancer. J Thorac Oncol 2013;8(11):1378-1388.

340. Tanigawa T, Araki S, Nakata A, Kitamura F, Yasumoto M, Sakurai S, Kiuchi T. Increase in memory (CD4+CD29+ and CD4+CD45RO+) T and naive (CD4+CD45RA+) T-cell subpopulations in smokers. Arch Environ Health 1998;53(6):378-383.

341. Tao S, Rojo de la Vega M, Chapman E, Ooi A, Zhang DD. The effects of NRF2 modulation on the initiation and progression of chemically and genetically induced lung cancer. Mol Carcinog 2018;57(2):182-192.

342. Teirilä L, Karvala K, Ahonen N, Riska H, Pietinalho A, Tuominen P, Piirilä P, Puustinen A, Wolff H. Proteomic changes of alveolar lining fluid in illnesses associated with exposure to inhaled non-infectious microbial particles. PLoS One 2014;9(7):e102624.

343. The Encode Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007;447:779-814.

344. Thornton DJ, Rousseau K, McGuckin MA. Structure and function of the polymeric mucins in airways mucus. Annu Rev Physiol 2008;70:459-486.

345. Timmerman R, Paulus R, Galvin J, Michalski J, Straube W, Bradley J, Fakiris A, Bezjak A, Videtic G, Johnstone D, Fowler J, Gore E, Choy H. Stereotactic body radiation therapy for inoperable early stage lung cancer. JAMA 2010;303:1070-1076.

346. Tojkander S, Gateva G, Schevzov G, Hotulainen P, Naumanen P, Martin C, Gunning, PW, Lappalainen P. A molecular pathway for myosin II recruitment to stress fibers. Curr Biol 2011;21:539-550.

347. Tost J, Gut IG. Genotyping single nucleotide polymorphisms by MALDI mass spectrometry in clinical applications. Clin Biochem 2005;38(4):335-350.

348. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, Beer DG, Powell CA, Riely GJ, Van Schil PE, Garg K, Austin JH, Asamura H, Rusch VW, Hirsch FR, Scagliotti G, Mitsudomi T, Huber

RM, Ishikawa Y, Jett J, Sanchez-Cespedes M, Sculier JP, Takahashi T, Tsuboi M, Vansteenkiste J, Wistuba I, Yang PC, Aberle D, Brambilla C, Flieder D, Franklin W, Gazdar A, Gould M, Hasleton P, Henderson D, Johnson B, Johnson D, Kerr K, Kuriyama K, Lee JS, Miller VA, Petersen I, Roggli V, Rosell R, Saijo N, Thunnissen E, Tsao M, Yankelewitz D. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011;6(2):244-285.

349. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS, Wistuba I, WHO Panel. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol 2015;10(9):1243-1260.

350. Tsangaris GT, Kolialexi A, Karamessinis PM, Anagnostopoulos AK, Antsaklis A, Fountoulakis M, Mavrou A. The normal human amniotic fluid supernatant proteome. In Vivo 2006;20(4):479-490.

351. Tsao AS, Scagliotti GV, Bunn PA Jr, Carbone DP, Warren GW, Bai C, de Koning HJ, Yousaf-Khan AU, McWilliams A, Tsao MS, Adusumilli PS, Rami-Porta R, Asamura H, Van Schil PE, Darling GE, Ramalingam SS, Gomez DR, Rosenzweig KE, Zimmermann S, Peters S, Ignatius Ou SH, Reungwetwattana T, Jänne PA, Mok TS, Wakelee HA, Pirker R, Mazières J, Brahmer JR, Zhou Y, Herbst RS, Papadimitrakopoulou VA, Redman MW, Wynes MW, Gandara DR, Kelly RJ, Hirsch FR, Pass HI. Scientific advances in lung cancer 2015. J Thorac Oncol 2016;11:613-638.

352. Tuononen K, Kero M, Mäki-Nevala S, Sarhadi VK, Tikkanen M, Wirtanen T, Rönty M, Knuuttila A, Knuutila S. ALK fusion and its association with other driver gene mutations in Finnish non-small cell lung cancer patients. Genes Chromosomes Cancer 2014;53(11):895-901.

353. Turner BM, Cagle PT, Sainz IM, Fukuoka J, Shen SS, Jagirdar J. Napsin A, a new marker for lung adenocarcinoma, is complementary and

more sensitive and specific than thyroid transcription factor 1 in the differential diagnosis of primary pulmonary carcinoma: evaluation of 1674 cases by tissue microarray. Arch Pathol Lab Med 2012;136(2):163-171.

354. Turvill J, Aghahoseini A, Sivarajasingham N, Abbas K, Choudhry M, Polyzois K, Lasithiotakis K, Volanaki D, Kim B, Langlands F, Andrew H, Roos J, Mellen S, Turnock D, Jones A. Faecal calprotectin in patients with suspected colorectal cancer: a diagnostic accuracy study. Br J Gen Pract 2016;66(648):e499-506.

355. Uchida A, Samukawa T, Kumamoto T, Ohshige M, Hatanaka K, Nakamura Y, Mizuno K, Higashimoto I, Sato M, Inoue H. Napsin A levels in epithelial lining fluid as a diagnostic biomarker of primary lung adenocarcinoma. BMC Pulm Med 2017;17(1):195.

356. Urieli-Shoval S, Cohen P, Eisenberg S, Matzner Y. Widespread expression of serum amyloid A in histologically normal human tissues. Predominant localization to the epithelium. J Histochem Cytochem 1998;46(12):1377-1384.

357. Vainio H, Boffetta P. Mechanisms of the combined effect of asbestos and smoking in the etiology of lung cancer. Scand J Work Environ 1994;20(4):235-242.

358. Vandijck DM, Hoste EA, Blot SI, Depuydt PO, Peleman RA, Decruyenaere JM. Dynamics of C-reactive protein and white blood cell count in critically ill patients with nosocomial Gram positive vs. Gram negative bacteremia: a historical cohort study. BMC Infect Dis 2007;7:106.

359. Van Hal SJ, Jensen SO, Vaska VL, Espedido BA, Paterson DL, Gosbell IB. Predictors of mortality in Staphylococcus aureus bacteremia. Clin Microbiol Rev 2012;25:362-386.

360. Vanholder R, Ringoir S, Dhondt A, Hakim R. Phagocytosis in uremic and hemodialysis patients: a prospective and cross sectional study. Kidney Int 1991;39:320-327.

361. Vansteenkiste J, Crinò L, Dooms C, Douillard JY, Faivre-Finn C, Lim E, Rocco G, Senan S, Van Schil P, Veronesi G, Stahel R, Peters S, Felip E. Panel Members. 2nd ESMO Consensus Conference on Lung Cancer:

early-stage non-small-cell lung cancer consensus on diagnosis, treatment and follow-up. Ann Oncol 2014;25:1462-1474.

362. Van Miert AS. Pro-inflammatory cytokines in a ruminant model: pathophysiological, pharmacological, and therapeutic aspects. Vet Q 1995;17(2):41-50.

363. Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. Circulation 2006;113:2335-2362.

364. Vasilescu C, Rossi S, Shimizu M, Tudor S, Veronese A, Ferracin M, Nicoloso MS, Barbarotto E, Popa M, Stanciulea O, Fernandez MH, Tulbure D, Bueso-Ramos CE, Negrini M, Calin GA. MicroRNA fingerprints identify miR-150 as a plasma prognostic marker in patients with sepsis. PLoS One 2009;4(10):e7405.

365. Vassallo R, Tamada K, Lau JS, Kroening PR, Chen L. Cigarette smoke extract suppresses human dendritic cell function leading to preferential induction of Th-2 priming. J Immunol 2005;175:2684-2691.

366. Vigushin DM, Pepys MB, Hawkins PN. Metabolic and scintigraphic studies of radioiodinated human C-reactive protein in health and disease. J Clin Invest 1993;91(4):1351-1357.

367. Vincent MD, Kuruvilla MS, Leighl NB, Kamel-Reid S. Biomarkers that currently affect clinical practice: EGFR, ALK, MET, KRAS. Curr Oncol 2012;19(Suppl 1):S33-44.

368. Vogenberg FR, Barash CI, Pursel M. Personalized medicine: part 1: evolution and development into theranostics. Pharmacy and Therapeutics 2010;35(10):560-562, 565-567, 576.

369. Wacker C, Prkno A, Brunkhorst FM, Schlattmann P. Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis. Lancet Infect Dis 2013;13(5):426-435.

370. Walsham NE, Sherwood RA. Fecal calprotectin in inflammatory bowel disease. Clin Exp Gastroenterol 2016;9:21-29.

371. Wang H, Zhang P, Chen W, Feng D, Jia Y, Xie LX. Evidence for serum miR-15a and miR-16 levels as biomarkers that distinguish sepsis from systemic inflammatory

372. response syndrome in human subjects. Clin Chem Lab Med 2012;50,1423-1428.

373. Wang HJ, Wang BZ, Zhang PJ, Deng J, Zhao ZR, Zhang X, Xiao K, Feng D, Jia YH, Liu YN, Xie LX. Identification of four novel serum protein biomarkers in sepsis patients encoded by target genes of sepsis-related miRNAs. Clin Sci (Lond) 2014;126(12):857-867.

374. Wang JF, Yu ML, Yu G, Bian JJ, Deng XM, Wan XJ, Zhu KM. Serum miR-146a and miR-223 as potential new biomarkers for sepsis. Biochem Biophys Res Commun 2010;394(1):184-188.

375. Wang R, An J, Ji F, Jiao H, Sun H, Zhou D. Hypermethylation of the Keap1 gene in human lung cancer cell lines and lung cancer tissues. Biochem Biophys Res Commun 2008;373(1):151-154.

376. Wang R, Braughton KR, Kretschmer D, Bach TH, Queck SY, Li M, Kennedy AD, Dorward DW, Klebanoff SJ, Peschel A, DeLeo FR, Otto M. Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. Nat Med 2007;13:1510-1514.

377. Wang X, Wang X, Liu X, Wang X, Xu J, Hou S, Zhang X, Ding Y. miR-15a/16 are upreuglated in the serum of neonatal sepsis patients and inhibit the LPS-induced inflammatory pathway. Int J Clin Exp Med 2015;8(4):5683-5690.

378. Wang Y, Faux SP, Hallden G, Kirn DH, Houghton CE, Lemoine NR, Patrick G. Interleukin-1beta and tumour necrosis factor-alpha promote the transformation of human immortalised mesothelial cells by erionite. Int J Oncol 2004;25(1):173-178.

379. Wannamethee SG, Lowe GD, Shaper AG, Rumley A, Lennon L, Whincup PH. Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. Eur Heart J 2005;26(17):1765-1773.

380. Wattavidanage J, Carter R, Perera KL, Munasingha A, Bandara S, McGuinness D, Wickramasinghe AR, Alles HK, Mendis KN, Premawansa S. TNFalpha*2 marks high risk of severe disease during Plasmodium

falciparum malaria and other infections in Sri Lankans. Clin Exp Immunol 1999;115(2):350-355.

381. Wen AQ, Wang J, Feng K, Zhu PF, Wang ZG, Jiang JX. Effects of haplotypes in the interleukin 1beta promoter on lipopolysaccharide-induced interleukin 1beta expression. Shock 2006;26(1):25-30.

382. Werner M, Sych M, Herbon N, Illig T, König IR, Wjst M. Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. Hum Mutat 2002;20(1):57-64.

383. Wertheim HF, Vos MC, Ott A, van Belkum A, Voss A, Kluytmans JA, van Keulen PH, Vandenbroucke-Grauls CM, Meester MH, Verbrugh HA. Risk and outcome of nosocomial Staphylococcus aureus bacteraemia in nasal carriers versus non-carriers. Lancet 2004;364(9435):703-705.

384. Wille MM, Dirksen A, Ashraf H, Saghir Z, Bach KS, Brodersen J, Clementsen PF, Hansen H, Larsen KR, Mortensen J, Rasmussen JF, Seersholm N, Skov BG, Thomsen LH, Tønnesen P, Pedersen JH. Results of the Randomized Danish Lung Cancer Screening Trial with Focus on High-Risk Profiling. Am J Respir Crit Care Med 2016;193(5):542-551.

385. Williams JA. Proteomics as a systems approach to pancreatitis. Pancreas 2013;42(6):905-911.

386. Wise KA, Tosolini FA. Detection of teichoic acid antibodies in Staphylococcus aureus infections. Pathology 1992;24(2):102-108.

387. Wistuba II, Lam S, Behrens C, Virmani AK, Fong KM, LeRiche J, Samet JM, Srivastava S, Minna JD, Gazdar AF. Molecular damage in the bronchial epithelium of current and former smokers. J Natl Cancer Inst 1997;89: 1366-1373.

388. Wittenhagen P, Kronborg G, Weis N, Nielsen H, Obel N, Pedersen SS, Eugen-Olsen J. The plasma level of soluble urokinase receptor is elevated in patients with Streptococcus pneumoniae bacteraemia and predicts mortality. Clin Microbiol Infect 2004;10:409-415.

389. Woolhouse IS, Bayley DL, Stockley RA. Effect of sputum processing with dithiothreitol on the detection of inflammatory mediators in chronic bronchitis and bronchiectasis. Thorax 2002;57(8):667-671.

390. World Health Organization; International Programme on Chemical Safety. (2001). Biomarkers In Risk Assessment: Validity And Validation. http://www.inchem.org/documents/ehc/ehc/ehc222.htm [accessed 11.3.2020]

391. World Health Organization. (2008) The Global Burden of Disease. Update. Geneva: World Health Organization.

392. Wysockia VH, Resingb KA, Zhanga Q, Chenga G. Mass spectrometry of peptides and proteins. Methods 2005;35:211–222.

393. Xiao H, Zhang Y, Kim Y, Kim S, Kim JJ, Kim KM, Yoshizawa J, Fan LY, Cao CX, Wong DT. Differential Proteomic Analysis of Human Saliva using Tandem Mass Tags Quantification for Gastric Cancer Detection. Sci Rep 2016;6:22165.

394. Xu L, Badolato R, Murphy WJ, Longo DL, Anver M, Hale S, Oppenheim JJ, Wang JM. A novel biologic function of serum amyloid A. Induction of T lymphocyte migration and adhesion. J Immunol 1995; 155(3):1184-1190.

395. Xu L, Wilson RA, Laetsch TW, Oliver D, Spunt SL, Hawkins DS, Skapek SX. Potential pitfalls of mass spectrometry to uncover mutations in childhood soft tissue sarcoma: A report from the Children's Oncology Group. Sci Rep 2016;6:33429.

396. Yamashita M, Fenn JB. Electrospray ion source. Another variation on the free-jet theme. J Phys Chem 1984a;88:4451-4459.

397. Yamashita M, Fenn JB. Negative ion production with the electrospray ion source. J Phys Chem 1984b;88:4671-4675.

398. Yun HM, Park KR, Lee HP, Lee DH, Jo M, Shin DH, Yoon DY, Han SB, Hong JT. PRDX6 promotes lung tumor progression via its GPx and iPLA2 activities. Free Radic Biol Med 2014;69:367-376.

399. Zee RY, Ridker PM. Polymorphism in the human C-reactive protein (CRP) gene, plasma concentrations of CRP, and the risk of future arterial thrombosis. Atherosclerosis 2002;162:217-219.

400. Zimakoff J, Bangsgaard Pedersen F, Bergen L, Baago-Nielsen J, Daldorph B, Espersen F, Gahrn Hansen B, Hoiby N, Jepsen OB, Joffe P,

Kolmos HJ, Klausen M, Kristoffersen K, Ladefoged J, Olesen-Larsen S, Rosdahl VT, Scheibel J, Storm B, Tofte-Jensen P. Staphylococcus aureus carriage and infections among patients in four haemo- and peritoneal-dialysis centres in Denmark. The Danish Study Group of Peritonitis in Dialysis (DASPID). J Hosp Infect 1996;33:289-300.

401. Zhang DL, Zheng HM, Yu BJ, Jiang ZW, Li JS. Association of polymorphisms of IL and CD14 genes with acute severe pancreatitis and septic shock. World J Gastroenterol 2005;11(28):4409-4413.

402. Zhang X, Meng X, Chen Y, Leng SX, Zhang H. The Biology of Aging and Cancer: Frailty, Inflammation, and Immunity. Cancer J 2017;23(4):201-205.

403. Zhu MM, Xu XT, Nie F, Tong JL, Xiao SD, Ran ZH. Comparison of immunochemical and guaiac-based fecal occult blood test in screening and surveillance for advanced colorectal neoplasms: a meta-analysis. J Dig Dis 2010;11:148-160.