Institute of Molecular Pathology
Faculty of Military Health Sciences
University of Defence

Identification of potential intraamniotic infection and inflammation biomarkers in amniotic fluid from preterm birth patients

Dissertation Thesis

Vojtěch Tambor, MSc

Supervisors:

Juraj Lenčo, PhD
Ramkumar Menon, PhD

Hradec Králové 2011

Ústav molekulární patologie
Fakulta vojenského zdravotnictví
Univerzita obrany

Identifikace potenciálních biomarkerů intraamniální infekce a zánětu v plodové vodě u pacientek s předčasným porodem

Disertační práce

Mgr. Vojtěch Tambor

Školitelé:

Juraj Lenčo, PhD
Ramkumar Menon, PhD

Hradec Králové 2011

Prohlašuji, že jsem tuto práci vypracoval samostatně. Veškerá literatura a další zdroje, z nichž jsem při zpracování této práce čerpal a na které se odkazuji, jsou uvedeny v seznamu použité literatury a v práci jsou řádně citovány.

V Hradci Králové, 11.10.2011

Mgr. Vojtěch Tambor

Acknowledgements

I wish to express my appreciation to all the women who agreed with the participation in our studies and to all others who have contributed to this thesis in different ways. In particular I would like to thank:

Juraj Lenco, my main supervisor, friend and colleague, for generously sharing his knowledge of research, for his ideas and suggestions. Thank you also for all the patience with the sometimes not-so-bright-as-you-might-have-wanted student.

Marian Kacerovsky, my clinical supervisor, friend and colleague, for actually enabling this whole project and thus my PhD thesis. Without his enthusiasm, passion and devotion to the hard work, he keeps on delivering, this would never have been possible.

Marek Link, our mass spectrometry wizard and LC magician, my friend and colleague, for endless patient hours spent tutoring, explaining and troubleshooting.

Jiri Stulik, head of our department, for providing me the very possibility to join his great team, for keeping the excellent spirit in the institute and for fighting the windmills in the name of the whole group.

Ramkumar Menon, University of Texas Medical Branch, Galveston, USA, my expert supervisor, for a very fruitful collaboration, discussion and support.

The "girls & boys" at the office, Alena, Jirka, Klarka, Lucka for being helpful, positive and for keeping up the great spirit.

All my workmates and lab technicians, as well as to Ivana Musilova and Daniel Lesko, clinicians at the Department of Obstetrics and Gynecology, for aiding in the recruitment of patients and finally to researchers at the Department of Immunology, for performing the cathelicidin quantitation experiments.

Most importantly, however, I would like to thank my family, for all the endless support and love. Without them, I would have never made this.

Grant acknowledgements

This work was supported by the following grants:

Ministry of Defence, Czech Republic, Grant No. MO0FVZ0000501
Ministry of Health, Czech Republic, Grant No. NS/10382-3/2009
Ministry of Education, Czech Republic, Grant No. ME10025
Czech Science Foundation, Grant No. 301/09/P241

Table of Contents

Summary 8
Souhrn 9
Abbreviations used in this thesis 10
Papers in this thesis 12
1. Introduction 13
1.1. Characterization and classification of preterm birth 13
1.2. Risk factors 15
1.2.1. Non-infectious risk factors 16
1.2.2. Intraamniotic infection 17
1.3. Available diagnostic tools 22
1.4. Proteomics in biomarker discovery 22
1.4.1. Discovery phase 24
1.4.2. Verification and validation phase 26
2. Aims of the work 29
3. Materials and methods 30
3.1. Amniotic fluid sample collection, classification and preparation 30
3.1.1. Amniotic fluid samples from PPROM patients 30
3.1.2. Amniotic fluid samples from sPTL patients with intact membranes 31
3.2. Genaration of pooled samples for discovery phases 32
3.3. AF sample processing for SRM analyses 32
3.4. Immunoafinity depletion of high abundance AF proteins 32
3.5. Development and application of the CysTRAQ method 33
3.6. Basic pH RP peptide fractionation 33
3.7. LC-MS/MS analysis 33
3.8. Data analysis 34
3.9. SRM assay design and experiments 35
3.9.1. Peptide selection 35
3.9.2. Transition selection 35
3.9.3. Peptide identification confirmation 35
3.9.4. LC-SRM analysis of AF samples 36
3.10. ELISA measurements 36
4. Results and discussion 37

4.1. Specific aim 1 - Development and optimization of an iTRAQ compatible peptide fractionation strategy based on cysteinyl peptide enrichment 37
4.2. Specific aim 2 - Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from PPROM patients using multidimensional comparative proteomics 40
4.3. Specific aim 3 - Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from patients with sPTL with intact membranes using multidimensional comparative proteomics 61
4.4. Specific aim 4 - Verification of differential abundance of selected proteins using both antibody based as well as proteomic targeted techniques. 65
4.4.1. Antibody-based verification experiments 65
4.4.2. SRM-based verification experiments 67
5. Conclusions 74
6. References 76
Outputs 88
Publications related to the thesis 88
Other publications and outputs 90
Awards 91
External tables 92
Attached papers I. and II.

Summary

In this work, we took advantage of modern proteomics in order to characterize intraamniotic infection and inflammation related changes in the proteome of amniotic fluid from preterm birth patients. Proteins with altered levels could subsequently serve as potential biomarkers for timely recognition of intraamniotic infection and inflammation.

Due to the extremely high complexity and high dynamic range of amniotic fluid proteome, we first developed a technique - CysTRAQ - which enables proteome complexity reduction based on cysteinyl peptide capturing and features a multiplexed protein quantitation across four samples. The developed method was subsequently applied into the comparative proteomic analyses of amniotic fluid.

Our study included patients with both principal phenotypes of spontaneous preterm birth - spontaneous preterm labour with intact membranes as well as preterm premature rupture of membranes - in order to provide a comprehensive insight into the proteomic background of the infectious and inflammatory processes occurring in amniotic fluid. By employing our CysTRAQ approach in combination with additional fractionation and separation techniques, we managed to describe a remarkable number of the amniotic fluid proteins. Owing to the quantitation feature of CysTRAQ, we were also able to quantify the differences between samples, where intraamniotic infection and inflammation was confirmed or ruled out, respectively.

In our results, the preterm premature rupture of membranes cohort showed a considerably higher degree of proteome dysregulation with regard to the presence of intraamniotic infection and inflammation compared to the spontaneous preterm labour patients. In both cohorts, we observed major changes in antimicrobial peptides, protease inhibitors and acute inflammatory phase signaling molecules. In the preterm premature rupture of membranes cohort, we described a dysregulation of a complex web of proteases as well as of their respective inhibitors. In the spontaneous preterm labour cohort, on the other hand, proteins related to neutrophil degranulation were among the most obviously dysregulated. Noteworthy, we observed profound changes in proteins related to neutrophils extracellular traps. Proteins constituting these structures were found to be dysregulated in both preterm premature rupture of membranes as well as in spontaneous preterm labour cohort and include histone proteins, neutrophil defensins and azurocidin. According to our knowledge, this work is the first one to suggest the presence of neutrophil extracellular traps in amniotic fluid during intraamniotic infection and inflammation.

Souhrn

V předkládané disertační práci jsme využili moderní proteomické technologie k charakterizaci rozdílů způsobených přítomností intraamniální infekce a zánětu v plodové vodě pacientek s předčasným porodem a k nalezení potenciálních biomarkerů tohoto patologického stavu.

Vzhledem k velmi vysoké komplexitě a širokému dynamickému rozsahu proteomu plodové vody jsme nejdříve vyvinuli techniku CysTRAQ, která umožňuje snížení komplexity proteomu a současnou relativní kvantifikaci proteinů napříč čtyřmi vzorky. Tato metoda byla následně aplikována při komparativní proteomické analýze plodových vod.

V naší studii byly zahrnuty oba hlavní fenotypy spontánního předčasného porodu - spontánní předčasný porod s intaktními membránami a předčasný odtok plodové vody. Zahrnutí obou fenotypů nám umožnilo získat zajímavý pohled na změny v plodové vodě pacientek s předčasným porodem způsobené infekcí a zánětem. S využitím metody CysTRAQ v kombinaci s dalšími frakcionačními a separačními technikami se nám podařilo popsat vysoký počet proteinů plodové vody. Díky kvantifikační složce metody CysTRAQ jsme popsali kvantitativní rozdíly mezi vzorky, u nichž byla potvrzena přítomnost infekce a zánětu a těmi, u nichž byl nález negativní.

Naše výsledky ukázaly, že změny v kohortě pacientek s předčasným odtokem plodové vody byly mnohem výraznější než u pacientek se spontánním předčasným porodem. V obou kohortách jsme pozorovali změny v hladinách antimikrobiálních peptidů, proteázových inhibitorů a signalizačních molekul akutní fáze zánětu. Ve skupině pacientek s předčasným odtokem plodové vody jsme popsali dysregulaci celé sítě proteáz a jejich inhibitorů. Oproti tomu, nejvýraznější změny u pacientek se spontánním předčasným porodem byly pozorovány u proteinů spojených s degranulací neutrofilů. Dále jsme nalezli řadu dysregulovaných proteinů, které jsou součástí neutrofilových extracelulárních pastí. Proteiny tvořící tyto sítě byly zvýšeny v obou studovaných kohortách a zahrnují histonové proteiny, neutrofil defensiny a azurocidin. Podle našich informací je tato práce první, která navrhuje přítomnost neutrofilových extracelulárních pastí v plodové vodě během intraamniální infekce a zánětu.

Abbreviations used in this thesis

1D	One dimensional
ACN	Acetonitrile
AF	Amniotic fluid
amu	Atomic mass unit
ANPEP	Aminopeptidase N
AZU1	Azurocidin
BMI	Body mass index
BSA	Bovine serum albumin
CLCA1	Calcium-activated chloride channel regulator 1
CRP	C-reactive protein
CysTRAQ	Strategy enabling sample fractionation and quantitation
Da	Dalton
ELA2	Neutrophil elastase
ELISA	Enzyme-linked immunosorbent assay
FA	Formic acid
FDR	False discovery rate
FIRS	Fetal inflammatory response syndrome
HCA	Histological chorioamnionitis
HPLC	High performance liquid chromatography
ICAM1	Intracellular adhesion molecule 1
IL	Interleukin
ILEU	Leukocyte elastase inhibitor
ITGAM	Integrin alpha-M
iTRAQ	Isobaric tags for relative and absolute quantitation
LC	Liquid chromatography
LPS	Lipopolysaccharide
LTA4H	Leukotriene A4 hydrolase
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption/ionization
MIAC	Microbial invasion of the amniotic cavity
MIDAS	MRM-initiated Detection and Sequencing
MMP	Matrix metalloproteinase
MPO	Myeloperoxidase

MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NETs	Neutrophil extracellular traps
NFκB	Nuclear factor κB
PAMP	Pathogen-associated molecular patterns
PANTHER	Protein ANalysis THrough Evolutionary Relationships
PCR	Polymerase chain reaction
PGLYRP-1	Peptidoglycan recognition protein 1
PMA	Phorbol myristate acetate
PPROM	Preterm premature rupture of membranes
PRR	Pattern recognition receptor
PRTN3	Proteinase 3; Myeloblastin
Q	Quadrupole
ROS	Reactive oxygen species
RP	Reverse phase
SCX	Strong cation exchange
SELDI	Surface Enhanced Laser Desorption Ionization
SLPI	Antileukoproteinase
sPTL	Spontaneous preterm labour
SRM	Selected reaction monitoring
TFA	Trifluoroacetic acid
TIMP	Tissue inhibitor of metalloproteinases
TLR	Toll-like receptor
TNF	Tumor necrosis factor
TOF	Time of Flight
VDBP	Vitamin D binding protein

Papers in this thesis

The thesis is based on the following articles and manuscripts, which will be referred to throughout the text. Both are attached at the end of this work.

Paper I.
Tambor V, Fucikova A, Lenco J, Kacerovsky M, Rehacek V, Stulik J and Pudil R
Application of proteomics in biomarker discovery: a primer for the clinician.
Physiological Research, 2010;59(4):471-97.

Paper II.
Tambor V, Hunter CL, Seymour SL, Kacerovsky M, Stulik J and Lenco J
CysTRAQ - A combination of iTRAQ and Enrichment of Cysteinyl Peptides for Uncovering and Quantifying Hidden Proteomes
Journal of Proteomics, 2011

1. Introduction

1.1. Characterization and classification of preterm birth

Preterm birth remains one of the main causes of perinatal morbidity and mortality and is therefore an enduring healthcare problem worldwide. According to the definition by World Health Organization, preterm birth is a delivery occurring at less than 37 weeks of gestation [1]. Although the global rate was estimated to fluctuate around 9.6% in 2005, there are evident geographic variations. While the rate in Europe was 6.2%, the data for Asia and Africa were markedly worse, showing rates of 9.1 and 11.9%, respectively. The assumption, that preterm birth is a problem of the developing countries only is promptly proven wrong, regarding the fact that the North American incidence was reported to be 10.6% and has been rising steadily for the past three decades [2-4].

Preterm births are not distributed evenly during pregnancy, with more than 60% occurring during 34-36 weeks of gestational age (near term), about 20% at 32-33 weeks (moderately preterm), around 15% at 28-31 weeks (severely preterm) and less than 5% before 28 weeks (extremely preterm) [5, 6] (Figure 1).

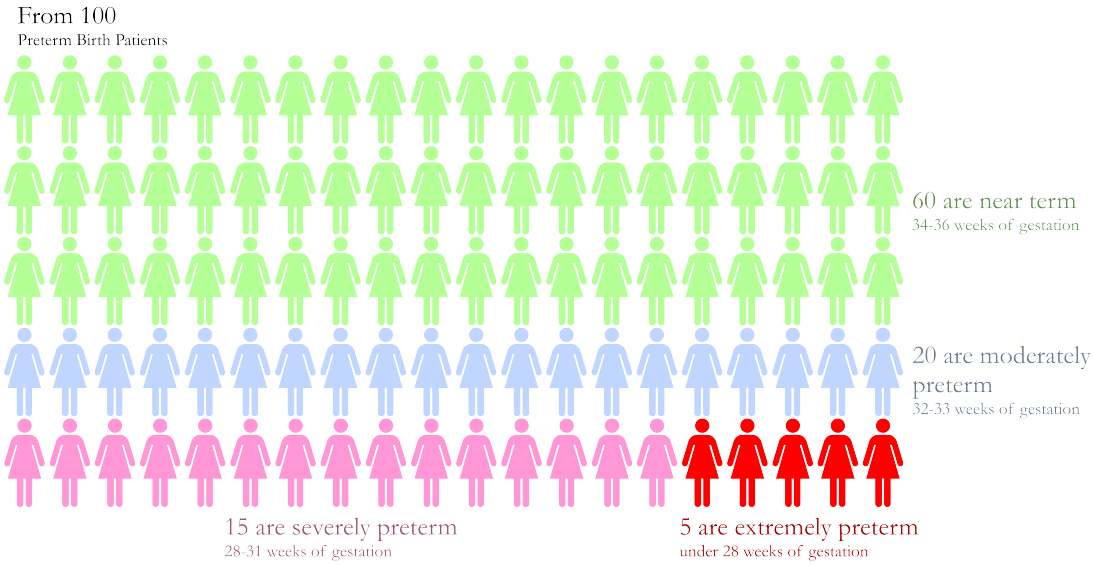


Figure 1. Distribution of preterm births according to weeks. The vast majority of preterm births occur near term (~60%). The most rare preterm deliveries, on the other hand, take place prior to 28th week of gestation.

Another point of view for the categorization may be based on principal clinical conditions leading to preterm birth. These include 1) iatrogenic preterm labour due to maternal or fetal indications, 2) spontaneous preterm labour (sPTL) with intact

membranes and 3) preterm premature rupture of membranes (PPROM). Iatrogenic deliveries account for about 30% of all preterm births, roughly 45% are due to idiopathic sPTL with intact membranes and finally around 25% follow PPROM [7]. Births caused by sPTL and by PPROM are commonly referred to as spontaneous preterm births (Figure 2). Although sPTL with intact membranes and PPROM are two distinct etiologies, there is no clear cut between them with regard to risk or causal factors, a majority of which play a role in both pathological states.

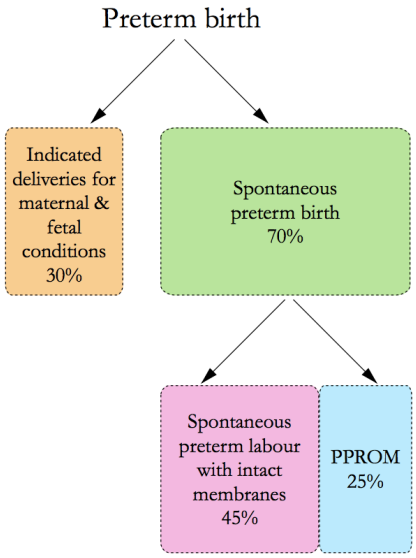


Figure 2. Individual precursors leading to preterm birth. Three principal causes lead to preterm birth - medical indications, sPTL with intact membranes and PPROM. The latter two are commonly referred to as spontaneous preterm birth. The percentage shown in the figure is regarded as percentage of all preterm births.

sPTL is defined as regular contractions prior 37 weeks of gestation. Although the pathophysiology is not completely understood, it has been suggested that sPTL might be the result of early activation of labour by pathologic processes. Even the influence of the fetus in the timing of labour has been suggested, with the fetal-adrenal axis and fetal cortisol playing key roles [8-10]. Importantly, the initiation has been shown to be tightly associated with the activation of inflammatory pathways [11].

PPROM, on the other hand, is manifested by spontaneous rupture of the membranes and leakage of amniotic fluid (AF) at least two hours before the onset of regular uterine activity in gestational age below 37 weeks of gestation [12]. The cause of this rupture is mostly unknown and the risk factors are similar to those for sPTL, but an important role of subclinical intraamniotic infection and smoking has been suggested [13]. Although most of the PPROM patients deliver spontaneously within days, a small part can continue with the pregnancy for several weeks. Unfortunately, not only can

intraamniotic infection be the underlying cause of PPROM, but is also its frequent complication. As the membranes form a natural barrier against microbial invasion, their rupture opens a gateway for ascending infection [14].

1.2. Risk factors

An exact mechanism responsible for sPTL or PPROM cannot be identified in most cases. Preterm birth is thus regarded as a syndrome, which may be initiated by multiple mechanisms, including intraamniotic infection and inflammation, placental ischemia or hemorrhage, uterine overdistension and stress [11, 14]. Moreover, a number of risk factors is thought to interact to cause a shift from uterine quiescence to sPTL or PPROM [15]. There are many maternal and fetal characteristics, which have been associated with spontaneous preterm birth and which can be categorized into the following groups; 1) demographic factors; 2) previous pregnancy history; 3) current pregnancy findings; 4) nutritional factors and 5) associated biomarkers (Figure 3).

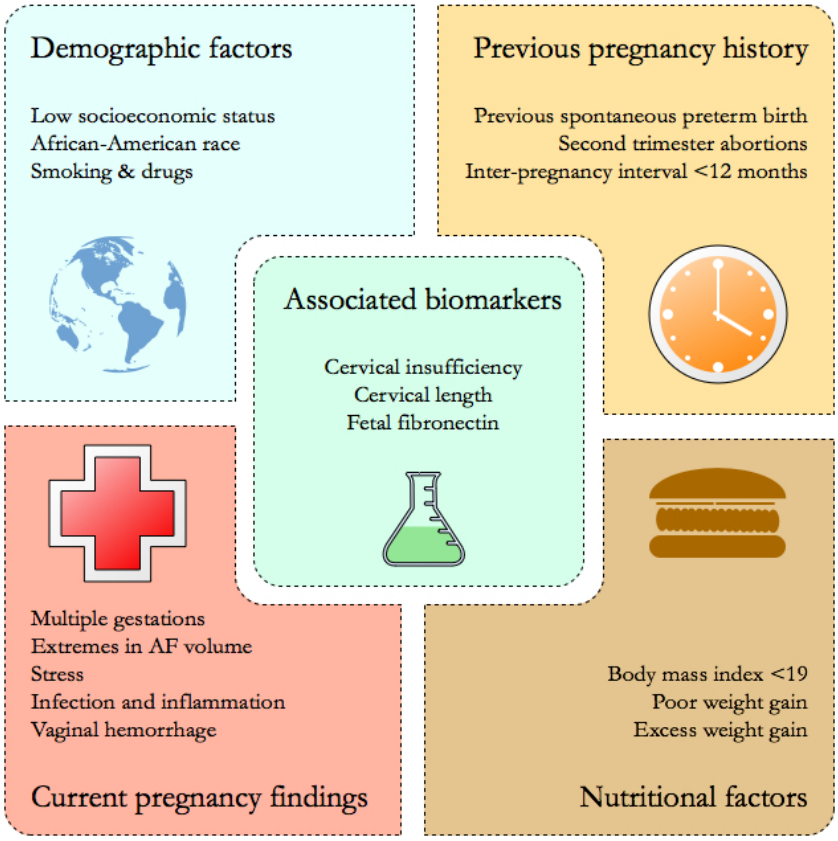


Figure 3. Risk factors of preterm birth. Although preterm birth is regarded as a syndrome rather than a disease caused by a particular symptom, several risk factors associated with preterm birth have been identified and are outlined in this figure. Importantly, it is rare that preterm birth occurs only due to one risk factor - mostly, several risk factors interact.

1.2.1. Non-infectious risk factors

African-American and Afro-Caribbean mothers are reported to be at higher risk of preterm delivery and PPROM compared to other races, as the frequency in black women is significantly higher (16-18%) compared to white women (5-9%) [16-18]. Up to date, the reason for this disproportion has not been clearly elucidated, although it has been suggested recently, that a specific combination of multiple risk factors occurring in a particular race could result in unique inflammatory response patterns and might thus contribute to this racial disparity in preterm birth frequency [19]. Asian and Hispanic women generally have low preterm birth rates and Indian mothers, although delivering low birth weight newborns, were not shown to have increased preterm birth rates [12]. A relationship evaluation of several perinatal morbidities revealed strong correlation of preterm birth to several socioeconomic measures, particularly to education, occupational status and area-based (neighborhood) measures [20]. Interestingly, many US immigrant groups show positive correlation between the preterm birth rate and the time spent living in the USA. Whether this is due to high rate of lacking health insurance, absence of a supportive social safety net or due to other factors, is unknown. Substance abuse, often tightly linked to low socioeconomic status, has been reported to be another risk factor for preterm birth. Although it has been shown that tobacco use increases the likelihood of preterm birth approximatel 2-fold, a substantial portion of smoking women continues also during pregnancy [21, 22]. Cigarette smoke contains a large number of chemicals, including carbon monoxide and nicotine. Both are potent vasoconstrictors and are linked to placental damage, decreased placental blood flow and could thus lead to induction of preterm labour. Smoking is also related to oxidative stress and apoptosis induction, which can lead to increased likelihood of preterm birth [23-25]. Even though alcoholism has been associated to serious defects in pregnancy as well as to preterm birth, mild to moderate alcohol consumption is not regarded as a risk for preterm birth. On the other hand, the abuse of drugs, i.e. cocaine or heroin, has been shown to correlate with increased risk of preterm delivery [26].

Women with previous preterm birth or low birth weight during previous pregnancy are well know to be at risk, as well as patients with repeated second trimester abortions [27, 28]. Low interval between current and previous pregnancy also leads to higher risk, which is increased ~2 fold if the time gap is 6 months or less. This may be explained by the fact that the uterus needs certain amount of time in order to fully recover to normal state. Similarly, the inflammatory process, which is partially considered

physiological in normal pregnancy, needs to be resolved during the reconditioning process of the uterus. Even the depletion of maternal minerals, vitamins and aminoacids may be a contributing factor, as these supplies need certain amount time to regenerate.

Nutritional factors, usually described using body mass index (BMI) or nutritional intake are also regarded as risk factors, and women with BMI<19 are threatened by high risk of spontaneous preterm birth. This might be caused by decreased blood flow and reduced uterine perfusion in skinny mothers [29]. These women also consume less food and have thus lower vitamins and nutrients intake. Interestingly, the other extreme - high BMI (obesity) - was shown to have protective effects and decreases the risk [30].

Although multiple pregnancies account for only 2-3% newborns, they are associated with considerable risk and constitute up to 20% of all preterm deliveries. Nearly half of the twins and the majority of triplets are born upon PPROM or sPTL. The associated uterine overdistention is regarded to be the major underlying cause, in particular in PPROM [11]. Women experiencing considerable stress during pregnancy, both psychological and social, have ~2 fold higher risk of preterm birth. Central role of corticotropin releasing hormone has been proposed in stress induced preterm labour, as women experiencing high levels of stress have high levels of inflammatory markers, including C-reactive protein (CRP). This suggests that stress might induce preterm birth, with the inflammatory pathway as the effector [31-33].

1.2.2. Intraamniotic infection

Another risk factor tightly associated with this pathway is intraamniotic infection, a frequent and important cause of preterm birth [34]. It is the only pathological process for which a firm causal link to preterm birth is known [35, 36]. Studies suggest, that infection accounts for 25-40% of all preterm births. [37, 38]. Under normal conditions, amniotic fluid is a sterile environment and less than 1% of women at term will have microbial invasion of the amniotic cavity (MIAC). In pathological states, the proportion of MIAC positive patients rises considerably and is dependent on gestational age and clinical presentation. To illustrate, the frequency of MIAC in sPTL patients, who deliver preterm, is 22%. On the other hand, more than 32% PPROM women are positive upon admission and by the onset of labour, positive AF culture is found in 75% cases [14, 38]. An interesting observation is that the lower the gestational age at sPTL or PPROM, the higher the proportion of positive AF cultures [39]. The microorganisms most frequently detected in amniotic cavity are *Mycoplasma hominis, Ureaplasma urealyticum, Streptococcus*

agalactiae, Escherichia coli and *Gardnerella vaginalis*. Moreover, a wide variety of pathogens implicated in MIAC are difficult-to-cultivate or uncultivable species [40]. The above-described prevalence of MIAC positive cases is described based on classical microbiological techniques, e.g. cultivation. A positive culture can be obtained only if the test is performed under conditions, which enable the growth of a particular organism. Therefore, a negative culture result does not necessarily exclude the presence of microorganisms. Whereas positive culture indicates MIAC (true-positive result), a negative culture could be the consequence of the absence of microbes (true-negative result) or due to non-suitable conditions for the growth of a particular species (false-negative result). Given the fact that only 1% of the microbes can be detected by cultivation techniques, the above reported numbers regarding MIAC positive patients could be considerably underrated [41]. It has been demonstrated, that the rate of MIAC is higher if polymerase chain reaction (PCR) detection of conserved prokaryotic 16S ribosomal DNA is used [42]. These results are also in concordance with the clinical context, as patients with positive PCR for *U. urealyticum* but negative AF culture have similar adverse outcomes as women with positive AF culture, whereas patients with both negative AF culture and negative PCR had significantly better results [43]. Similarly, patients with positive PCR, but negative AF culture had comparable degree of inflammation as patients with positive AF culture [44]. This suggests, that PCR detection might be a more reliable indicator of microbial presence compared to classical cultivation methods.

While the sterility of AF during physiological pregnancy has been confirmed even using molecular biology techniques, the situation in fetal membranes is far less evident. As expected, bacterial infiltration in membranes is found in most MIAC positive patients, both in sPTL as well as in PPROM women. However, fluorescence *in situ* hybridization with probe against 16S ribosomal DNA detected presence of microbes in a striking 70% of women delivering at term. The bare presence of microorganisms might therefore not be sufficient to cause neither MIAC and subsequent sPTL or PPROM, nor a maternal or fetal immune system response [45]. Moreover, major causative agents of MIAC, genital mycoplasmas, are very often of low virulence. This accounts for rather chronic nature of the infection and explains the frequent absence of clinical symptoms [34]. The complexity of this problem is further illustrated by the fact that patients diagnosed with MIAC in the second trimester due to invasion of *M. hominis* delivered healthy neonates in term, with no complications. In contrary, a similar cohort of women

with MIAC caused by *U. urealyticum* delivered preterm prior to 30th week of gestation, with neonatal sepsis or neonatal death [46]. The chronicity aspect of *Mycoplasma*-caused MIAC may be explained by the fact that this particular microbial strain does not cause prostaglandin secretion upon invasion and thus does not cause contractions of the uterus [47]. This goes well with the findings that PPROM patients are more frequently diagnosed with *Mycoplasmas,* compared to patients with sPTL [48, 49].

A parallel can be drawn between intraamniotic infection and intrauterine inflammation, both of which are often associated. Women with sPTL and PPROM have elevated inflammatory markers in AF several weeks earlier, during the second trimester screening, compared to women who delivered at term. This points out to the fact, that rather than an acute condition, intrauterine inflammation (histological chorioamnionitis, HCA), very much like intraamniotic infection, is often a chronic process [50, 51].

Microorganisms can gain access to the amniotic cavity by numerous ways; 1) ascending from the vagina; 2) retrograde invasion through the fallopian tubes; 3) transplacental infection; 4) accidental introduction due to invasive procedures, i.e. amniocentesis, chorionic villi sampling, with the most common one being the ascending route (Figure 4) [34].

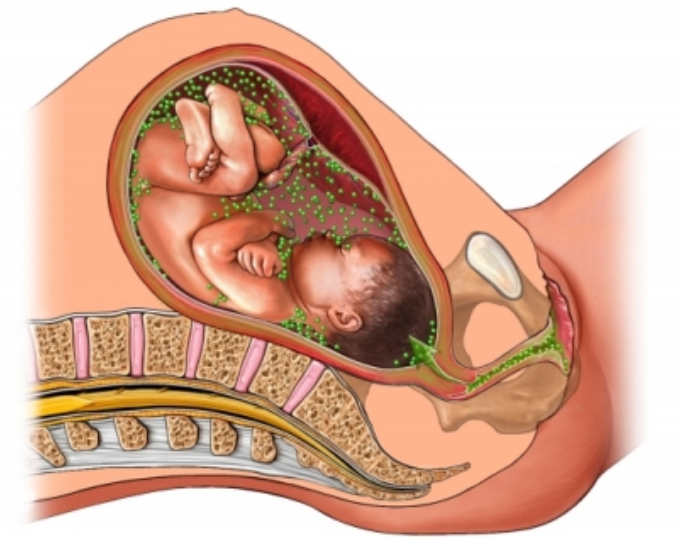


Figure 4. Ascending route of infection. Microbes can gain access to the uterus *via* the ascending route from vagina. Despite several barriers, both physical and chemical, obstruct the advancement of bacteria, these can fail and the microbes can gain access to the fetal membranes, AF or in the worst cases invade the fetus. (© 2009 Nucleus Medical Art, Inc., reprint permission requested).

The immune system has developed several defensive strategies to cope with pathogens. Epithelial surfaces, i.e. of the fetal membranes, represent the first barrier. Unfortunately, not only can the microorganisms invade through even minute injuries in these epithelia, but there is also evidence that bacteria can cross intact fetal membranes [52, 53]. The membranes therefore need to be more than a plain physical barrier. Indeed, most epithelia produce natural antimicrobial peptides like alpha- and beta-defensins, protegrin or cathelicidin, which can kill bacteria by interaction with their membrane [54-56]. Another component of the immune system has been designed to prevent tissue infiltration and microbial proliferation by triggering an inflammatory process, which is activated i.e. in response to detection of microbes by pattern recognition receptors (PRRs) [57]. Being a part of the innate immunity, these receptors are designed to identify pathogen-associated molecular patterns (PAMPs) such as those present on surface of microorganisms [58]. The molecules recognized include lipopolysaccharides (LPS), bacterial carbohydrates present in the cell wall of Gram-negative bacteria or peptidoglycans and lipotechoic acid from Gram-positive bacteria [58, 59]. PRRs can be classified based on their localization into: 1) secreted PRRs, including the complement receptors and pentraxin proteins such as serum amyloid and CRP, which opsonize pathogens to be eliminated by the complement system; 2) cytoplasmic PRRs - NOD-like receptors and RNA helicases, which seem to mediate intracellular recognition of viral RNA; and 3) membrane-bound or transmembrane PRRs, such as toll-like receptors (TLRs), mannose receptors and scavenger receptors [57]. Eleven TLRs have been discovered in mammals. To illustrate, TLR-2 recognizes peptidoglycans and lipoproteins (Gram-positive bacteria, *Mycoplasmas* and funghi), TLR-3 detects double-stranded viral RNA, the specific ligands for TLR-4 are lipopolysaccharides (Gram-negative bacteria) and heat shock proteins [60]. Upon ligand detection by TLR, the nuclear factor κB (NFκB) is activated, which in turn leads to elevated production of cytokines, chemokines and antimicrobial peptides [58]. As TLRs are critical in detection of microorganisms, defective signaling through these PRRs should prevent infection-induced preterm birth. Indeed, a mouse strain with TLR-4 mutation has lower rate of preterm birth upon administration of LPS or dead bacteria than the wild-type mice [61]. During pregnancy, TLR-2 and -4 are expressed in the amniotic epithelium and this expression rises during labour both at term and preterm, regardless of the membrane rupture status. This fact suggests, that the innate immune system plays a key role in both preterm as well as in physiological term labour [62].

Another part of the immune system, which plays a role in preterm birth are inflammatory mediators. The key molecules involved in this process include the proinflammatory interleukin (IL)-1β, tumor necrosis factor (TNF)-α and IL-8, but other pro- and anti-inflammatory cytokines or chemokines, such as prostaglandins, platelet-activating factor and mediators, can also play a role.

Microbes following the ascending infection route may reach the decidua, infiltrate the fetal membranes and elicit a local inflammatory response, which in turn causes the release of proinflammatory agents (prostaglandins, leukotrienes, reactive oxygen species (ROS), nitric oxide, etc.). If this reaction itself does not cause preterm birth, the microorganisms can invade the amniotic cavity, where they also stimulate production of inflammatory mediators by intraamniotic immune system cells. Finally, the microorganism can gain access to the fetus and cause a systemic inflammatory response, followed by production of cytokines as well as by activation of neutrophils and monocytes [63, 64]. Due to the similarity with systemic inflammatory response syndrome, the resulting condition is termed fetal inflammatory response syndrome (FIRS) [65]. This stage is considered the most advanced and results in very high rates of neonatal morbidities, such as neonatal sepsis, pneumonia, respiratory distress syndrome, bronchopulmonary dysplasia or necrotizing enterocolitis. These may result in multiple organ damage or failure and ultimately, FIRS may even result in the death of the fetus [66, 67].

Not only are the neonates disadvantaged by being born preterm, but the associated intraamniotic infection and subsequent intrauterine inflammation, both local and systemic, further deteriorate their status. Besides the resulting conditions requiring acute intensive care upon birth, the newborns are threatened also by lifelong adverse sequelæ (e.g. neurodevelopmental impairments, respiratory and gastrointestinal complications), which have both direct and long-term consequences on quality of life and health care costs [12]. Therefore, timely recognition of ongoing intraamniotic infection and inflammation is crucial for reducing the associated risks. Unfortunately, the rather silent course of the infectious process makes the recognition of the patients at risk truly difficult. In turn, these patients thus often remain without appropriate clinical management.

1.3. Available diagnostic tools

Early detection of ongoing MIAC and HCA remains an obstetrical challenge and the available options are limited.

PPROM patients may choose to have amniocentesis performed for AF cultivation and for PCR detection of MIAC. This, however, is time consuming, expensive, not routinely performed and thus only optional for the patient. Moreover, a number of healthcare facilities does not even enable these tests or the national policies regulating the management do not offer this option. A protein detection-based bedside test would thus greatly facilitate such tests. HCA is diagnosed either by evaluating the "golden trias" in AF (glucose, lactate and leukocytes) or assessment of IL-6 and matrix metalloproteinase (MMP)-8 levels [68].

The situation in sPTL patients with intact membranes is even more complicated, as AF sampling is not generally performed in these cases. Maternal status and blood are then the only sources for monitoring inflammatory signs (CRP, leukocytes, fever etc.). Unfortunately, by the time these become elevated in the mother, the fetus is often already endangered [69-71].

This illustrates that the routinely used clinical examinations and markers often suffer from major drawbacks, including low sensitivity, low specificity, long sampling to result time or unacceptable risk for the patient. Researches have therefore been attempting to find new robust and sensitive markers for presence of MIAC, ongoing HCA and even FIRS.

1.4. Proteomics in biomarker discovery

A common research concept, frequently used in protein biomarker discovery, is the involvement of antibody-based techniques for confirming or disproving a particular hypothesis. Besides these hypothesis-driven experiments, proteomic approach has recently gained substantial attention in biomedical research, including the novel biomarker identification field. In contrary to biochemical and molecular biology methods, no particular hypothesis is needed prior to global proteomic analysis. As proteomic methods enable unbiased characterization of the protein composition of a biological system and its dynamics, new and undescribed phenomena, which occur in context with a particular pathologic state, can be found using this approach. In novel biomarker research, quantitative proteomic methods are of particular interest, as these

enable precise assessment of even minute quantitative alterations of protein levels across the studied samples. For closer details, please follow an extensive introduction into proteomics with focus on biomarker-related research, which can be found in Paper I.

From the proteomic point of view, amniotic fluid analysis presents an analytically challenging task. Due to the fact that AF is derived from plasma, it also inherits its two major stumbling blocks - extremely high complexity and dynamic range of the proteins. Similarly to plasma, amniotic fluid is expected to be a mixture of several thousand proteins, which can be present at concentrations spanning across up to 10 orders of magnitude [72]. In comparison to maternal plasma, however, these changes in protein levels are not diluted by the large maternal blood volume and are thus more likely to be detected using proteomic techniques in AF. Moreover, AF is contained in an enclosed compartment and thus represents a very interesting environment. It surrounds the fetus not only from the outside, but also passes through several cavities, including the fetal digestive and respiratory tract. Changes occurring in the fetal organism might therefore be reflected in the AF composition, proteomic in particular.

In response to infection and inflammation, a wide range of proteins can be specifically secreted, released or shed from pathologically affected fetal cells and tissues. In turn, these could serve as potential biomarkers suitable for diagnosis and early detection. Most of these proteins, however, are expected to be of low abundance and thus hard to detect, as the AF proteome is dominated by a few high abundance proteins, e.g. albumin. Finding such hidden proteins can very well turn into looking for a needle in a haystack [73].

In several projects, proteomics has been applied into intraamniotic infection and inflammation biomarker research. One of the pioneering piece of work in this regard was conducted by Buhimschi et al. [74]. The authors compared profiles of low-molecular segment of the amniotic fluid proteome of intraamniotic infection and inflammation negative and positive samples using direct mass spectrometric (MS) profiling on a Surface Enhanced Laser Desorption Ionization - Time of Flight (SELDI-TOF) instrument. As a result, the authors proposed detection of neutrophil defensins-1 and -2, and calgranulins A and C as putative markers. Although this approach benefits from its high-throughput nature and requires a very low amount of sample, it inherently cannot discover low-abundance molecules. In addition, conventional SELDI-TOF instruments do not provide any direct means for peak identification and the potential of SELDI-TOF in biomarker discovery process is nowadays regarded as rather limited.

In contrast, the combination of high-performance liquid chromatography (HPLC) and tandem mass spectrometry (MS/MS), also known as shotgun proteomics, allows characterization of proteomes with greater depth and dynamic range, but most importantly provides direct means for large-scale protein identification. Along with the ability to use a plethora of quantitation methods, the shotgun strategy has become highly popular in proteomic research. In a typical shotgun experiment, proteins are digested into peptides with a sequence-specific protease, usually trypsin. The resulting peptide mixture is simplified using HPLC, using either one dimensional (1D) reversed-phase (RP) separation or alternatively 2D fractionation approach. Finally, the separated peptides are analyzed on a tandem mass spectrometer. The acquired peptide-level data are in turn used for database searching and for "reconstructing" original protein-level evidence. Using this approach, up to several hundred proteins can be identified and quantified in a proteomic experiment.

Due to the large number of fractions generated in order to reach the low abundance proteins, a typical multidimensional shotgun experiment may take several days or even weeks to perform. As such, it is therefore unsuitable for analyzing each clinical sample from a large cohort individually. A two-stage strategy can therefore be advantageously used in biomarker proteomic projects [75, 76]. In the first phase of the novel biomarker identification process - the discovery phase - representative surrogate samples for each group are analyzed and compared using quantitative proteomics. This poses strict requirements on proper selection and rigorous stratification of the samples into groups to be compared. The most promising candidates from the deep quantitative characterization of the AF proteome are subsequently taken further forward into the second step of the strategy - the verification phase.

1.4.1. Discovery phase

Due to the fact that the majority of potential protein biomarkers are supposed to be of low abundance and thus hidden among high abundance ballast, substantial demands are placed on fractionation and separation techniques since even the latest proteomic technologies are not able to comprehensively interrogate such a complex AF protein digest, and moreover, cannot cover more than 4 orders of concentration range [77]. As a consequence, technologies reducing the sample complexity, and thus increasing the likelihood of uncovering low-abundant proteins with promising diagnostic potential, are required to address these challenges [78-80]. Such technologies may involve

multiple rounds of fractionation both at the protein level and peptide level after digestion. Immunoaffinity depletion, often used as a first step for complexity reduction in plasma or serum proteomics, can be employed also in AF proteome research, as the majority of high abundance proteins are identical [81]. By this means, the most abundant proteins with no diagnostic potential are removed, but as these represent about 90% of the total protein mass, this step results in significant sample simplification. In quantitative proteomic projects, the choice of suitable quantitation strategy is crucial. An overview of available choices can be found in Paper I. and II. According to published data and also our experience, the isobaric tags for relative and absolute quantitation (iTRAQ) technology is very well suitable for clinical proteomics projects [82-84]. Using this method, up to four, or even eight samples in the latest version, respectively, can be analyzed, quantified and compared in a single analysis. Due to the fact that upon iTRAQ labeling, the samples can be mixed and processed as a single one, this technology is ideal for subsequent peptide fractionation and separation strategies for proteome complexity reduction. Researchers have exploited a range of peptide properties, which could be used for peptide fractionation. A detailed overview can be found in Paper I. and II.

Due to the reactivity, simplicity of reaction and ease of reaction reversibility, cysteine is the favourite amino acid for this strategy. An *in-silico* digest of the human proteome revealed that merely 15 % of all human tryptic peptides with molecular mass spanning from 800 to 3000 Daltons (Da) contain at least one cysteine in their sequence. As a consequence, specific enrichment of these peptides results in considerable reduction of sample complexity and therefore similar number of MS/MS events during a proteomic analysis may lead to increased number of identified proteins. On the other hand, due to the fact that 97% of all human proteins contain at least one tryptic peptide with cysteine, the major part of the proteome is still represented in such simplified mixture [85]. Moreover, a recent work combining fractionation based on cysteines and multidimensional separation in a concentration-annotated yeast standard proteome study has clearly shown increased detection of low abundance proteins [86]. All these facts and features make cysteine an ideal target for enrichment strategies to improve the proteome coverage.

However, even after fractionation, the complexity of the AF digest is still high, and additional separation steps might therefore result in increased likelihood of detecting low abundant species. A popular concept, described in Paper I., uses 2D peptide fractionation in order to reach these low abundance analytes. Although strong cation

exchange (SCX) is commonly used in the first dimension to fractionate peptides, several reports suggest that orthogonality of SCX and RP modes is not ideal [87]. It is known that separation in SCX is directed by the peptide charge. Since the tryptic peptides are mostly 2+ and 3+ charged, the peptides cluster in a narrow retention time windows. In addition, the majority of desirable peptides elute from the column early in the analysis, leaving a portion of separation space relatively devoid of peaks. Therefore, the peak capacity of the 2D HPLC based on SCX-RP combination may be lower than expected. Recently, an alternative replacing SCX with RP at basic pH in the first dimension was described and suggests overall better performance. The RP columns have ~2-fold higher peak capacity, compared to SCX [88]. Therefore, the eluting peptides are not distributed across multiple fractions, are highly concentrated and thus more probable to detect. Also, the peptides elute evenly and not in "charge batches" as in SCX. The peptides show better recovery from the column and the mobile phases used are compatible with subsequent MS analysis, requiring just lowering organic solvent concentration and not desalting, as opposed to SCX fractionated peptides. Individual peptide fractions generated in the first dimension are separated by a routine acidic pH RP prior to MS and MS/MS analysis, upon which the acquired data are evaluated, as closer described in Paper I. and II.

The set of identified and quantified proteins is subsequently evaluated using statistical methods and literature search in order to fish the most promising potential biomarker proteins. This ensures that only high-quality marker candidates from the discovery phase are taken further forward into the verification phase and ultimately to preclinical validation.

1.4.2. Verification and validation phase

Until recently, antibody-based assays (i.e. enzyme-linked immunosorbent assay, ELISA) would have been introduced at this point as the method of choice for the verification and validation phase in proteomics. However, as more and more novel marker candidates were reported, it became obvious that the lack of high quality antibodies would hinder the verification of a large portion of these potential biomarkers. This induced a paradigm shift, which has been apparent in this area since the introduction of targeted HPLC-Selected Reaction Monitoring (SRM, also called Multiple Reaction Monitoring, MRM) technology into proteomics, allowing the verification phase to be carried out without the need for a specific antibody. For more details on SRM,

please follow Paper I. Briefly, SRM analyses are generally performed on mass spectrometers capable of working in triple quadrupole mode. These instruments are set to select only a specific precursor peptide in the first quadrupole (Q1), which is then fragmented in the collision cell and only a specific fragment is allowed to go through the third quadrupole (Q3) and reach the detector (Figure 5). The combination of Q1 and Q3 filter masses is known as a transition and is unique and selective for a particular analyte. This configuration allows high sensitivity and high specificity targeted detection and more importantly quantitation of chosen molecules. Moreover, as the instrument alternates between different transitions in few milliseconds, tens to hundreds of different molecules may be analyzed in a single HPLC run. By adding internal standards containing heavily labeled amino acids, latest SRM instruments are capable of accurate and precise absolute quantitation over four orders of magnitude [89].

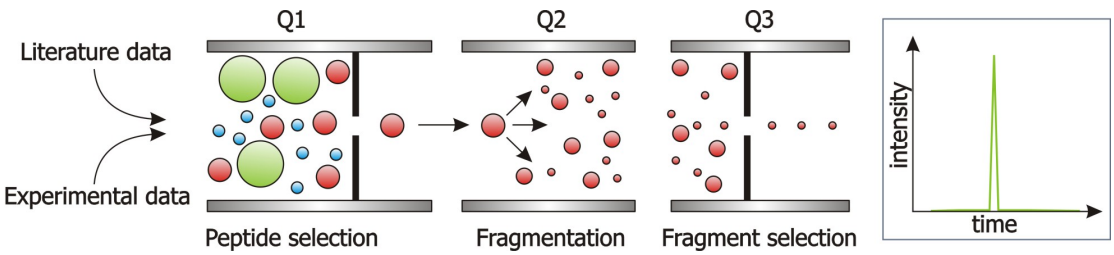


Figure 5. Schematic representation of the principle of SRM. A triple quadrupole mass spectrometer is set to select a specific precursor mass in Q1, fragment it in Q2 and finally selects a specific fragment in Q3. As the peptide elutes from the column, the response intensity is plotted over time and the peptide amount can be correlated to the area under the chromatographic peak.

Hence, when considering optimal strategy for discovery phase candidate verification, antibody-based ELISA are a good choice if a suitable pair of antibodies or a ready-made kit is available. On the other hand, if no high quality antibodies are at disposal, SRM assays represent a compelling alternative. The development step is the most critical part, but once successfully completed and optimized, these assays can be used in a fairly high-throughput manner at very low operating costs.

Recently, a novel concept for designing and optimizing SRM assays has been introduced. First, synthetic analogues of peptides of interest are generated in a high-throughput and relatively cheap fashion [90, 91]. These are then used in the first step of the assay development. As the peptides are available in relatively high abundance, they enable easy selection of optimal transitions, sequence confirmation, retention time

monitoring and collision energy optimization. The optimized method is subsequently used in the final assay in order to detect the desired peptide in the sample.

The concept of synthetic peptides brings along yet another important feature. If stable-isotope labeled heavy amino acids are used in peptide production and precise amount of the peptide of interest is determined, this heavy labeled counterpart of the intrinsic peptide can be spiked into the sample. Due to identical chemical composition, both natural and heavy labeled peptides co-migrate during all steps of the experiment, but owing to a specific mass difference, they may be readily distinguished by the mass spectrometer. The comparison of peak areas of individual peptides finally enables precise absolute quantitation of the protein of interest [89, 90].

2. Aims of the work

The major goal of this thesis was to identify novel putative biomarkers of intraamniotic infection and inflammation in amniotic fluid of preterm birth patients using proteomics. The studied cohorts included both sPTL patients with intact membranes as well as patients with PPROM.

Due to the complexity of amniotic fluid, the project included development a novel strategy for sample complexity reduction, which would be compatible with protein quantitation. This strategy was subsequently applied into the discovery phase of two large-scale proteomics analyses of both studied cohorts. Finally, several proteins, which were found to be dysregulated, were verified using complementary methods in the verification phase.

The specific aims were:

1. Development and optimization of an iTRAQ compatible peptide fractionation strategy based on cysteinyl peptide enrichment.

2. Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from PPROM patients using multidimensional comparative proteomics.

3. Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from patients with sPTL with intact membranes using multidimensional comparative proteomics.

4. Verification of differential abundance of selected proteins using both antibody based as well as proteomic targeted techniques.

3. Materials and methods

3.1. Amniotic fluid sample collection, classification and preparation

3.1.1. Amniotic fluid samples from PPROM patients

The sample collection within the study was approved by the Ethical Committee of the University Hospital in Hradec Kralove on March 19th 2008 (No. 200804 SO1P). Pregnant women with gestational age between 24-36 weeks, who were admitted to the Department of Obstetrics and Gynecology University Hospital Hradec Kralove since May 2008 with a diagnosis of PPROM were involved into the project. Only women with the following criteria were enrolled: singleton pregnancy, sonographically estimated fetal weight between the 10th-90th percentiles for gestational age, and the absence of fetal structural malformations or chromosomal abnormalities. Exclusion criteria were significant vaginal bleeding and signs of fetal hypoxia. Amniotic fluid samples (3-5 ml) were obtained on admission before administration of corticosteroids, antibiotics, or tocolytics, by ultrasound-guided transabdominal amniocentesis after signing written informed consent. Samples were supplemented with Complete EDTA free protease inhibitors (Roche, Basel, Switzerland) and centrifuged at 300x *g* to remove cells and debris. The supernatants were filtered through a 0.22 μm syringe-driven filter (TPP, Trasadingen, Switzerland) and stored at -80°C until use. The samples from the placenta, the fetal membranes, and the umbilical cord were obtained at delivery.

PPROM was defined as fetal membrane rupture with leakage of AF that precedes the onset of uterine contraction by at least 2 hours at <37 weeks of gestation and was diagnosed by sterile speculum examination confirming the pooling of AF in the vagina in association with the positive test for the presence of the insulinlike growth factor-binding protein (ACTIM PROM test, Medix Biochemica, Kauniainen, Finland) in the vaginal fluid. The presence of HCA was determined by histological examination of the placenta, the fetal membranes and umbilical cord was performed in all cases. The degree of polymorphonuclear leukocyte infiltration was assessed according to criteria given by Salafia et al. [92]. Intraamniotic infection was allocated as a presence of MIAC, which was defined as a positive result of PCR analysis for genital mycoplasmas (*U. urealyticum, Ureaplasma parvum, M. hominis*) and/or positive cultivation results of any bacteria in the amniotic fluid, except *Staphylococcus epidermidis*, which was considered a skin contaminant.

Only patients with confirmed MIAC and HCA were considered as positive. Patients with both MIAC and HCA ruled out were considered as negative.

3.1.2. Amniotic fluid samples from sPTL patients with intact membranes

The sample collection within the study was approved by the TriStar Nashville Institutional Review Board. Pregnant women admitted at the Centennial Women's Hospital in Nashville, Tennessee, USA, for sPTL with intact membranes (defined as the presence of regular uterine contractions at a minimum frequency of 2 contractions per 10 minutes combined with documented cervical change) between 22 and 36 weeks were enrolled to the study after signing written informed consent. Race was identified by self-report from a set of provided choices and determined by the race of the mother and father of the fetus and their parents and grandparents. Subjects of mixed races were excluded from the study. Only Caucasians of non-Hispanic origin were included [93]. Gestational age was determined by last menstrual period dating and corroborated by early ultrasound. Subjects with oligo- or polyhydramnios, multiple gestation, preterm premature rupture of the membranes, preeclampsia, placental previa, fetal anomalies, or medical/surgical complications of pregnancy or surgeries during pregnancies were excluded. Amniotic fluid samples were collected using a 22-gauge needle by transvaginal amniocentesis during labor after complete dilatation before rupture of the membranes and delivery. Samples were not collected from subjects who presented with spontaneous rupture of the membranes either at preterm or term. The samples were collected from the forebag from which maximum cytokine concentrations would be expected [94, 95]. Amniotic fluid was centrifuged immediately for 10 minutes at 2000x *g* to remove cellular and particulate matter. Aliquots of amniotic fluid were stored at -70°C until analysis.

Demographic data collected included maternal age, socioeconomic status (education, yearly income, and marital status), and a complete medical and obstetric history. Histologic examination of the placenta and umbilical cord was performed. HCA was defined as a dense polymorphonuclear leukocyte/neutrophil infiltration of the amniochorionic membrane excluding decidua, and funisitis was defined as inflammation in one or more of the umbilical cord vessels (vasculitis) with or without inflammation in the Wharton's jelly.

MIAC was defined by the presence of bacteria in the amniotic fluid detected by amplification of microbial 16s ribosomal DNA by PCR (TaqMan assay, Foster City, CA) [96, 97]. PCR data were correlated with the presence of histologic evidence of

chorioamnionitis to rule out contamination of vaginal colonizers during collection of samples. Only patients with confirmed MIAC and HCA were considered as positive. Patients with both MIAC and HCA ruled out were considered as negative.

3.2. Genaration of pooled samples for discovery phases

In order to create a pooled sample for the discovery phase, protein concentration was determined in each AF sample using the bicinchoninic acid protein assay kit (Pierce, Rockford, IL). Equal protein amount from each sample was taken and used to create a pooled positive and a pooled negative sample, both in duplicates (Figure 6). The resulting samples contained 2 mg of total protein each. The volume was adjusted to 4 ml using the immunoaffinity depletion buffer A and all four samples were concentrated using Amicon Ultra centrifugal filters (Millipore, Bedford, MA) with 3 kDa molecular weight cutoff to reach protein concentration roughly equal to human plasma. The retenates were collected, adjusted to 200 µl and stored at -80°C until immunoaffinity depletion.

3.3. AF sample processing for SRM analyses

One milligram of total protein from each AF sample was adjusted to 4 ml using the immunoaffinity depletion buffer A and concentrated using the Amicon Ultra filters with 3 kDa molecular weight cutoff to roughly 150 µl, retenates were collected, adjusted to 200 µl and stored at -80°C until immunoaffinity depletion.

3.4. Immunoafinity depletion of high abundance AF proteins

The top 14 high abundance proteins, which constitute up to 90% of the total protein mass of human amniotic fluid, were removed using the MARS Hu-14 immunoaffinity column (Agilent, Palo Alto, CA) on an Alliance 2695 HPLC system (Waters, Milford, MA) according to manufacturer's instruction. The flow-through fraction (~3 ml) containing low abundance proteins was collected between 5th and 22nd minute and immediately frozen at -80°C. In order to exchange the buffer for downstream applications, the immunoaffinity depletion buffer was exchanged for water using the 3 kDa cutoff Amicon Ultra filters by 3 subsequent concentration steps, so that after the final spin, the original buffer concentration was below 1% of the original amount. The retenates were collected and stored at -80°C.

3.5. Development and application of the CysTRAQ method

These steps are closer described in Paper II.

3.6. Basic pH RP peptide fractionation

Desalted lyophilized cysteinyl and non-cysteinyl peptide fractions were dissolved in 200 µl 20 mM ammonium formamate. The separation and fractionation was performed on the Alliance 2695 HPLC system. Due to higher peptide content in the non-cysteinyl peptides, 100 µl was injected from this sample as opposed to 200 µl from the cysteinyl peptide fraction, both of which were loaded onto a Gemini C_{18} 150 x 2 mm column (Phenomenex, Torrance, CA) filled with 3 µm, 110 Å particles. The peptides were separated by a linear gradient, from 5% acetonitrile (ACN), 20 mM ammonium formamate to 55% ACN, 20 mM ammonium formamate in 62 min. The eluting peptides were collected between 20-60 min of the gradient and in total, 18 fractions were collected *per* sample. Right after collection, 2.5 µl of 100% formic acid (FA) was added to each sample to adjust the pH to slightly acidic, the samples were dried in a SpeedVac and stored at -80°C until analysis.

3.7. LC-MS/MS analysis

Individual fractions from the first dimension fractionation were redissolved in 40 µl 5% ACN, 0.1% trifluoroacetic acid (TFA), following nanoLC peptide separation, which was performed on an UltiMate3000 HPLC system (Dionex, Sunnyvale, CA). First, a UV only run was performed by injecting 5 µl from each fraction in order to provide insight on the total peptide content *per* fraction. Based on the UV trace, the final peptide load was normalized to be constant across all the fractions. Peptides from each fraction were desalted on a µ-Precolumn 300 µm x 5 mm filled with C_{18} PepMap, 5 µm, 100 Å particles (Dionex). The peptides were then eluted on an analytical NanoEase column 100 µm x 150 µm filled with Atlantis C_{18}, 3 µm, 100 Å particles (Waters). Peptides were separated by a linear gradient formed by 5% ACN, 0.1% TFA and 80% ACN, 0.1% TFA, from 0-90% of 80% ACN, 0.1% TFA in 90 min at a flow rate of 360 nl/min. The Probot fraction collector (Dionex) was set to collect fractions every 8 s for 60 min onto a blank OptiTOF LC-MALDI sample plate (AB SCIEX, Foster City, CA). The eluate was mixed 1:4 post-column with 3 mg/ml α-cyano-4-hydroxycinnamic acid matrix (LaserBio Labs, Sophia-Antipolis, France) in 70% ACN, 0.1% TFA. The MALDI analysis was

performed on a 4800 MALDI-TOF/TOF instrument (AB SCIEX). MS spectra were acquired across the mass range of 800-4000 mass-to-charge ratio (m/z) using 625 laser shots *per* spectrum. A maximum of 12 precursors were chosen for fragmentation in each MS spectrum. For each precursor, collision induced dissociation MS/MS spectra were acquired with a total accumulation of 3000 laser shots, starting with the weakest precursor.

3.8. Data analysis

Peptide identification and quantitation was conducted using the ProteinPilot 2.0.1 software (AB SCIEX) using the Paragon database search algorithm and the integrated false discovery rate (FDR) analysis function [98, 99]. The software used only unique peptide sequences as evidence for protein identification. The data were searched against UniProtKB/Swiss-Prot database containing proteins entries. The samples were described using the following parameters in the Paragon method: Sample Type - iTRAQ 4plex (Peptide Labeled); Cys Alkylation - MMTS; Digestion - trypsin; Special Factors - no selection; Species - *Homo sapiens*. The processing was specified as follows: Quantitate - On; Bias Correction - On; Background Correction - On; ID Focus - Biological Modifications; Search Effort - Thorough; Detected Protein Threshold - 0.05 (10.0%). For FDR determination, data were searched against concatenated databases by *in silico* on-the-fly reversal for decoy sequences automatically by the software. Only proteins at 5% FDR and distinct peptides at 5% FDR were used for further analysis of the amniotic fluid data. For quantitation, the ProteinPilot software excluded peptides with confidence <1%, peptides without iTRAQ modification, and spectra shared between different proteins - i.e. where a spectrum is also claimed by a different protein but with the same, similar, or unrelated peptide sequence having reasonable confidence as well. All the remaining peptides contributed to protein quantitation in the ProteinPilot software. Intensities of iTRAQ reporter ions were corrected using isotope correction factors supplied with the iTRAQ kit. Automatic normalization of quantitative data (bias correction) was further performed for each iTRAQ pair to correct any experimental or systematic bias. The background correction function in ProteinPilot software was also used. This function uses an algorithm to estimate and subtract out the background contribution of many low-level coeluting peptides in complex mixtures, which tends to attenuate the extremity of ratios, making them less accurate.

3.9. SRM assay design and experiments

3.9.1. Peptide selection

The sequence for each protein was imported into the Skyline software (MacCoss Lab, University of Washington, Seattle, WA) [100]. The software enables automatic peptide selection based on user-defined criteria, which are outlined in Table 1.

Enzyme	Trypsin
Missed cleavages	Not allowed
Peptide length	6-18 amino acids
Potential ragged ends	Excluded
Cysteine containing peptides	Excluded
Methionine containing peptides	Excluded
NXT/NXS consensus containing peptides	Excluded
NG consensus containing peptides	Excluded
RP/KP consensus containing peptides	Excluded

Table 1. Skyline peptide selection criteria. The following criteria were used for peptide selection. NXT - Asparagine - X - Threonine consensus sequence. NXS - Asparagine - X - Serine consensus sequence. X stands for any amino acid. NG - Asparagine-Glycine sequence. RP - Arginine-Proline. KP-Lysine-Proline.

For each predicted peptide, we also considered its cleavage probability using the PeptideCutter tool (http://web.expasy.org/peptide_cutter) and finally assessed its uniqueness by searching the peptide sequence against an *in-silico* trypsin digest of the human proteome directly in the Skyline software.

3.9.2. Transition selection

In the Skyline software, we selected the transitions using the following criteria. Precursor charges 2+ and 3+ were allowed. We manually deleted all triply charged precursors bellow 350 m/z and for singly charged fragments, we selected only *y*-type ions. Product ions were selected in a range from (m/z > precursor)- 1 to last ion -1 (with maximum at 1400 m/z). If the peptide contained proline, the first proline *y*-ion was automatically added to the transition list. The selected transitions were exported into a transition list with predicted collision energy and declustering potential optimized for the 4000 QTRAP instrument (AB SCIEX) and a minimum dwell time of 10 ms was set for each transition.

3.9.3. Peptide identification confirmation

For MRM-initiated Detection and Sequencing (MIDAS) experiments [101], the enhanced product ion scan (full MS/MS spectrum acquisition) was triggered if the

transition intensity was higher than 1000 counts *per* second. The resulting MIDAS triggered MS/MS spectra were evaluated using the MASCOT (Matrix Science, Boston, MA) search engine in the ProteinPilot software.

3.9.4. LC-SRM analysis of AF samples

Chromatographic separations were performed on a Tempo nano MDLC system (AB SCIEX). After loading the sample (1 μg), peptides were pre-concentrated on a trapping column, filled with 3 μm Atlantis dC18 particles (Waters), 100 μm x 25 mm at a flow rate 2 μl/min for ten minutes. Gradient from 5% to 40% of 98% ACN, 0.1% FA in 60 min at a flow rate 360 nl/min was used to resolve peptides on an PicoTip Emmiter (New Objective, Woburn, MA) filled in house with 3 μm Atlantis dC18, 75 μm x 100 mm. Peptides were analyzed on a 4000 QTRAP hybrid mass spectrometer equipped with a nanoelectrospray. A spray voltage of 2400 V was used with a source temperature 175°C. The mass spectrometer was operated in SRM mode with both quadrupoles filtering with 0.7 unit mass resolution. For all SRM analyses, 10 ms dwell time was used for each transition. MIDAS based enhanced product ion spectra were recorded with Q1 filter set to 1.0 unit resolution. Enhanced product ion spectra were recorded at a scan speed 4000 atomic mass units *per* second between 250 to 1400 m/z, with enabled dynamic fill time and Q0 trapping.

3.10. ELISA measurements

Cathelicidin AF levels were determined using Human LL-37 ELISA kit (Hycult Biotech, Uden, The Netherlands) according to manufacturer's instructions. The limit of detection of the kit was 0.14 ng/ml. Samples of amniotic fluid were diluted 1:4 using phosphate buffer saline. Absorbance values were read at 450 nm using Multiskan RC ELISA reader (Thermo Fisher Scientific, Waltham, MA). These measurements were perfomed by Dr. Andrys, Department of Clinical Immunology, University Hospital Hradec Kralove.

Myeloperoxidase (MPO) AF levels were determined using Human MPO ELISA kit (R&D Systems, Minneapolis, MN) according to manufacturer's instructions. The limit of detection of the MPO kit was 1.56 ng/ml. Samples of amniotic fluid were diluted 1:10 using dilution buffer supplied with the kit. Absorbance values were read at 450 nm using Paradigm ELISA reader (Beckman Coulter, Indianapolis, IN).

4. Results and discussion

4.1. Specific aim 1 - Development and optimization of an iTRAQ compatible peptide fractionation strategy based on cysteinyl peptide enrichment

We have successfully developed, optimized and applied a method, which enables iTRAQ reagent quantitation of peptides fractionated based on presence of a cysteine residue, thus CysTRAQ. For the first time, we have shown, that iTRAQ quantitation is fully compatible with cysteinyl peptide enrichment and is not influenced by the fractionation process. This technique is closely described in Paper II.

In the initial phase of the project we used radioactively ^{35}S-labeled cysteine to generate radioactive *Francisella tularensis* proteome. Upon trypsin digestion, the peptides were iTRAQ labeled and subsequently used to optimize and fine tune the cysteinyl peptides capturing on Thiopropyl-Sepharose 6B beads, which was based on a workflow described by Liu et al. [79]. Based on the liquid scintigraphy data, we managed to capture 78% of the original radioactivity detected in the loaded sample using the optimized resulting protocol. The reason why we did not enrich the entire radioactivity remains unclear. We experimentally excluded the ability of the bacterium to convert cysteine to methionine and thus we hypothesize that other sulfur substances without thiol group are responsible for the residual radioactivity in the unbound fraction, e.g. peptides derived from cysteine acylated lipoproteins [102]. Our assumption is supported by the LC-MS/MS results, which show almost total selectivity of the protocol towards enriched cysteinyl peptides. Importantly, we found that only ~5% of the radioactivity remained bound to the beads, showing excellent peptide recovery (>90%). This is a key improvement compared to another quantitative proteomics workflow targeting cysteines, Isotope-Coded Affinity Tags, as efficient elution of the labeled peptides from the enrichment streptavidin column was one of the drawbacks of this approach [103].

Once we had a robust assay based on cysteinyl peptide fractionation at disposal, we sought to find out whether the fractionation would influence the iTRAQ quantitation. We used a standard bovine serum albumin (BSA) digest, which was iTRAQ labeled in three ratios and subjected to the fractionation protocol. The results show, that the quantitative data are consistent across both cysteinyl and non-cysteinyl fractions and are in agreement with the original ratios. Thus, we conclude that the iTRAQ quantitation is compatible with the cysteinyl peptide enrichment step and is not influenced by the fractionation.

The final step involved application of the CysTRAQ workflow on AF samples in order to evaluate the performance of the method in complex sample analysis. For a proof-of-concept purpose, we included two representative AF samples from PPROM patients, which were confirmed to be both infection and inflammation positive as well as two samples with both parameters negative. The samples were processed as described in Paper II. and subsequently subjected to the CysTRAQ workflow, whereby the unfractionated sample as well as both cysteinyl and non-cysteinyl peptide fractions were included for analysis.

The combined analysis of unfractionated sample and of the cysteinyl peptides fraction resulted in 20% increase in protein identifications compared to unfractionated sample only. Similarly, when a combined analysis of unfractionated sample and non-cysteinyl peptides fraction was performed, the total number of identified proteins increased by 40%. Ultimately, when both cysteinyl and non-cysteinyl peptide fractions were analyzed together, 60% more proteins were identified compared to analyzing unfractionated sample only. In this experiment, only 29 proteins were unique for the unfractionated sample, compared to 846 identified in the cysteinyl and non-cysteinyl peptide fractions, showing low benefit of analyzing all three fractions. Based on our data, if the instrument time is limited, the analysis of the fraction depleted of cysteinyl peptides is the most advantageous, as the highest number of proteins *per* fraction was identified here. On the other hand, if the goal is to identify a higher number of proteins and thus increase the likelihood of detecting low abundant proteins, both fractions can be analyzed.

The inspection of peptides identified in this study revealed exceptional selectivity of the sepharose beads towards cysteinyl peptides. In the non-cysteinyl fraction, only 3.2% of the peptides contained cysteine in their structure. More importantly, the peptides detected in the cysteinyl peptides fraction were enriched with outstanding specificity, as 98.6% of them did indeed contain a cysteine residue. We hypothesize that this specificity is enabled by covalent binding of the cysteinyl peptides to the solid support. Being truly unique among cysteine targeting approaches, the covalently bound peptides can withhold much more stringent washing out of non-specifically bound peptides. The exceptional specificity towards cysteinyl peptides is unmatched by any of the previous works [79, 86]. In fact, the presence of cysteine in the sequence could be an additional criterion for confident peptide identification and thus for FDR rate reduction. Besides the advantage with regards to washing, the formation of disulfide bond is a simple oxidation reaction.

The elution of the peptides of interest from the beads is therefore based on plain reduction. We support this by our radioactive peptides-based experiments, where 93.6% of the total radioactivity was recovered from the resin. The enrichment step does not modify the peptides with any kind of tag to mediate the capturing of cysteinyl peptides. Thus, we can omit potential mass increment associated issues during MS analysis, ionization efficiency modification or evaluation software custom modifications, which in turn simplifies the whole workflow.

We also analyzed the AF results with regard to protein quantitation. Our previous BSA experiments suggested, that the iTRAQ quantitation was not influenced by the fractionation step. To verify if this conclusion was valid also in the case of a complex sample analysis, we performed a parametric Pearson test with calculated protein ratios from individual fractions. The comparison results of quantitative data from unfractionated sample vs. both cysteinyl and non-cysteinyl peptide fractions (correlation coefficient r=0.837; p<0.001) as well as a comparison of cysteinyl vs. non-cysteinyl peptide fractions (r=0.795; p<0.001) are in good agreement with our BSA results and show, that the iTRAQ quantitation is consistent across all fractions. The workflow reproducibility was also assessed and showed excellent performance in two subsequent enrichment experiments, as the quantitation correlation was 0.89 and 0.91 (p<0.001 for both) for two cysteinyl and two non-cysteinyl peptide fractions, respectively. The quantitative information is not incorporated in a cysteine targeting tag, thus the quantitation is not limited to the cysteinyl peptides fraction, providing additional versatility. As the quantitation is based on the iTRAQ amine labeling strategy, both cysteinyl and non-cysteinyl fractions can be readily analyzed and profit from the high multiplex advantages of iTRAQ MS/MS quantitation. We show, that the non-cysteinyl peptide group in fact contributed with the highest number of uniquely identified proteins. Owing to the high quantitation correlation between both generated fractions, the analysis of the non-cysteinyl peptides fraction is a perfect complementary for the cysteinyl peptides analysis. The ability of both analyzing and quantifying the non-cysteinyl fractions is unique for CysTRAQ and presents a major advantage, as supported by our results.

4.2. Specific aim 2 - Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from PPROM patients using multidimensional comparative proteomics

Amniotic fluid represents a very appealing source of potential biomarkers of a wide range of pathological disorders. Unfortunately, the extreme complexity, similar to plasma or serum, hinders easy and straightforward discovery of these highly promising molecules. Due to the fact that the AF proteome is dominated by a few high abundance proteins, the detection of low abundance molecules remains an analytical challenge. Our experience shows, that merely 80 proteins can be identified from an AF analysis using 1D-RP LC-MALDI workflow. The majority of proteins detected in this simple experiment belong to the high abundance proteins and for most of them the diagnostic potential is poor. In order to get around this problem and to identify a larger number of proteins in order to reach the low abundance proteome, with the ultimate goal to identify potential biomarkers of intraamniotic infection and inflammation, we used a two-stage strategy.

In the discovery phase of the project, we selected representative surrogate samples from our PPROM AF sample set, accepting only samples from patients diagnosed and described with high confidence. Nineteen samples from patients with confirmed MIAC and with confirmed HCA were selected and used as a positive group. As a control, we included 19 AF samples from PPROM patients, where both MIAC and HCA were ruled out. Samples in both groups were matched for gestational age, maternal age, parity and smoking status in order to eliminate potential bias (Table 2).

Upon generating pooled positive and negative samples and immunoaffinity depletion, we applied the CysTRAQ protocol. Both cysteinyl and non-cysteinyl peptide fractions were subjected to basic pH RP fractionation and all resulting fractions were analyzed using the LC-MALDI approach. This multidimensional quantitative proteomics workflow lead to identification of 9 422 distinct peptides, which were identified using more than 12 000 MS/MS spectra. Based on these peptides, 851 AF proteins were identified (all data presented at 5% FDR rate). The iTRAQ labeling enabled simultaneous relative quantitation of abundance changes across individual samples. We used the iTRAQ 4-plex version and thus we were able to analyze two replicates in a single analysis as shown in Figure 6. For all proteins, we annotated the matching protein class using the Protein ANalysis THrough Evolutionary Relationships (PANTHER) database [104]. The

complete list of these proteins is included in Table 8. Due to its extensive size, this table can be found in the "External tables" section at the end of this work.

	The presence of both MIAC and HCA (n=19)	The absence of both MIAC and HCA (n=19)	*p*-value
Maternal age (years)	29.5±6.6	28.1±4.3	0.46
Nulliparous	10 (53%)	11 (58%)	1.00
Prepregnancy BMI	19.9 (17.0 - 33.0)	22.0 (17.6 - 33.2)	0.12
Gestational age at sampling (days)	215 (177-243)	224 (168-244)	0.13
Gestational age at delivery (days)	217 (179-243)	224 (168-244)	0.23
CRP level at admission (mg/l)	6.6 (0.0 - 59.0)	6.02 (0 - 33.0)	0.84
White blood count at admission ($x10^9$/l)	13.0 (0 - .0)	13.6 (7.0 - 19.0)	0.61
Birth weight (grams)	1572±401	1782±529	0.18
PPROM to delivery interval (hours)	51 (10 -120)	21 (8 - 244)	0.39
Spontaneous delivery	14 (74%)	14 (74%)	1.00
Cesarean section	5 (26%)	5 (26%)	1.00
Apgar score in 5 minutes	9 (0 - 10)	9 (0 -10)	0.69
Apgar score in 10 minutes	10 (6 -10)	10 (7 -10)	0.54

Table 2. Maternal and newborn characteristics based on the presence and absence of both MIAC and HCA - PPROM cohort. Continuous variables were compared using parametric t-test (presented as mean ± SD) or a nonparametric Mann-Whitney U test, presented as median (range). Categorical variables were compared using Fisher exact test and presented as number (%).

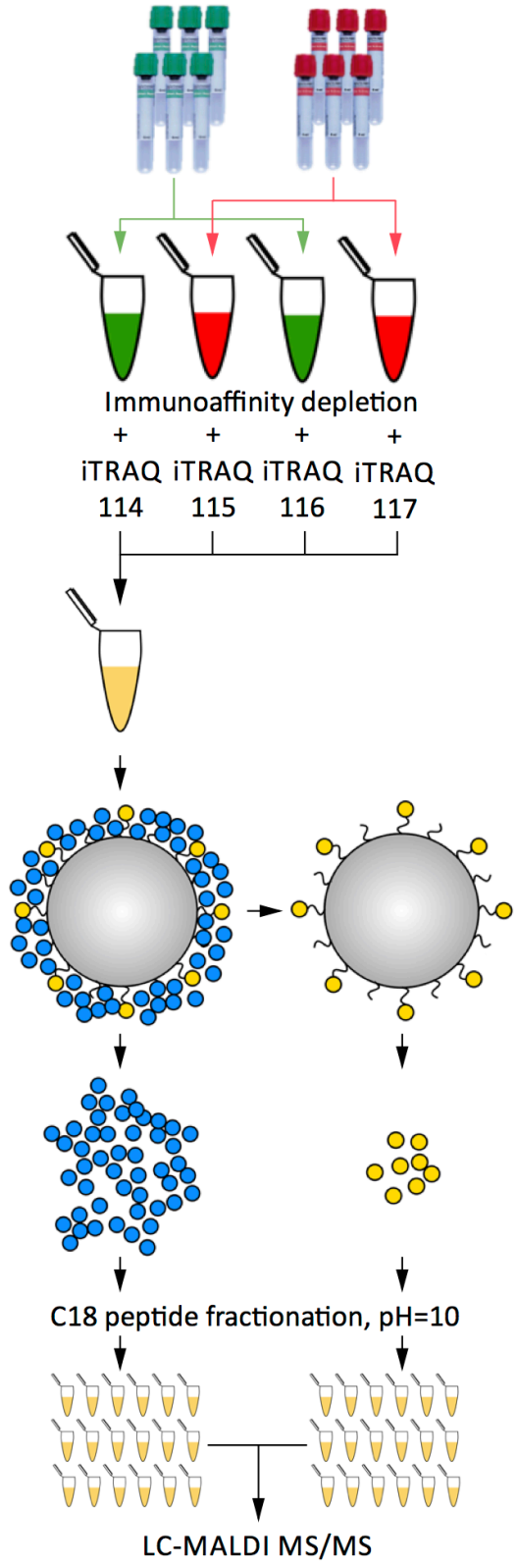


Figure 6. Discovery phase workflow. MIAC and HCA positive and negative samples were pooled in duplicates. These four samples were immunodepleted and the CysTRAQ protocol was applied. Both cysteinyl and non-cysteinyl fractions were subjected to basic pH RP fractionation and each generated fraction was subjected to the LC-MALDI workflow.

The discovery phase results should be considered as a snapshot of the processes occurring in the AF based on the presence of infection and inflammation, which in turn lead to quantitative and qualitative protein abundance changes. Therefore, rather than being final and definite, the results along with other information available should guide the scientist to select the most promising candidates suitable for verification. Thus, any mathematical or statistical criteria used to filter the results are actually artificial. In Table 3, we provide an overview of how these criteria change the resulting "protein of interest" count.

	5% FDR - complete dataset	$p<0.05$ in one replicate	$p<0.05$ in both replicates	$p<0.01$ in one replicate	$p<0.01$ in both replicates	$p<0.01$ in both replicates, 2-fold change
PPROM cohort	851	299	146	133	99	46
sPTL cohort	846	72	27	23	9	3

Table 3. Filtering criteria based on *p*-value and relative protein ratio change, with corresponding matching protein counts.

As mentioned in the introduction, proteomics enables unbiased analyses, from which specific hypotheses are created based on the obtained results. In our PPROM dataset, we chose the following initial criterion for the hypothesis derivation and subsequent candidate selection: $p \leq 0.01$ in both replicates and an at least 2-fold abundance change, either upwards and downwards. This resulted in a subset of 46 proteins (included in Table 9, External tables). However, we will discuss also proteins, which did not meet these conditions, but were found to have a potential role in the pathophysiology of intraamniotic infection and inflammation and thus we think are worth mentioning.

In concordance with previously published data, we detected protein S100-A8 and -A12 (calgranulins A and C) and neutrophil defensin 1 significantly elevated in the positive group [74]. In addition, our study revealed high levels of several other S100 proteins, including protein S100-A9 (calgranulin B), protein S100-A11 (calgizzarin) and protein S100-P.

Calgranulins A and B are calcium-binding proteins with direct antimicrobial activity against bacteria and fungi. Both proteins aggregate in form of a complex and together play a key role in innate resistance to invasion by pathogenic bacteria. The complex promotes myeloid cell function by binding to and activating TLR-4, regulates macrophage and neutrophil migration, infiltration and accumulation as well as macrophage cytokine production [105]. Through TLR activation, genes that are under the control of NFκB are upregulated. Both act as pro-inflammatory mediators in acute

and chronic inflammation and up-regulate the production and release of IL-8 and cell-surface expression of intercellular adhesion molecule 1 (ICAM1). We were able to detect IL-8 and confirm high upregulation in both replicates (12 and 20 fold, respectively). On the other hand, we confidently detected and quantified ICAM1, but the levels did not differ in both groups. This is most probably caused by the fact that ICAM1 is a membrane bound protein and its levels in AF might not reflect the cellular ones [106]. In contrast, IL-8 is a secreted protein and increased production of this cytokine might lead to a corresponding rise in AF IL-8 level [107].

Calgranulin C shares the majority of properties with calgranulin A and B, including antimicrobial effects, immune system response and calcium binding ability. Calgranulin C, however, requires zinc ions for proper oligomerization and its antimicrobial activity is then thus enhanced, unlike calgranulin A and C, in which zinc inhibits their antimicrobial effect [108, 109]. A protein interaction analysis using STRING protein interaction database (http://www.string-db.org, [110]) revealed close relationship of the three calgranulins, whereas no direct link with the remaining two S100 family representatives, protein S100-A11 (calgizzarin) and protein S100-P was found. Interestingly, calgizzarin is directly influenced by IL-8, which bridges the link with the above-described calgranulins (Figure 7).

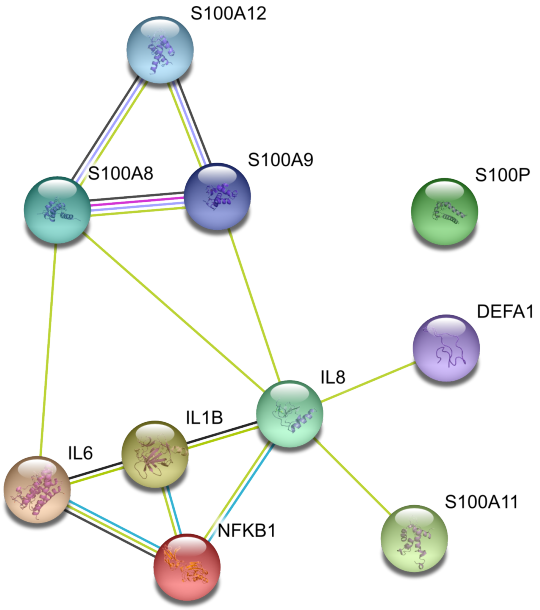


Figure 7. S100 proteins interaction network. The proteins were shown to be interacting with IL-8 and further on with the NFκB network. Interestingly, protein S100 A11 (calgizzarin) was shown to be related to the network as well, but indirectly, through IL-8. Althought the protein S100P was shown to be also dysregulated in our analysis and the name suggests similar involvement as for the remaining S100 proteins, the STRING analysis suggests a different type of regulation, as this protein had no partner in the network.

IL-8 expression is also enhanced by neutrophil defensin-1, another protein mentioned in several intraamniotic infection and inflammation studies [111-113]. We were able to confirm these findings and detected a ~4-fold higher levels of this protein in the positive group. Neutrophil defensin-1 is a part of the alpha-defensins family of cytotoxic and antimicrobial peptides involved in host defence. The protein is highly abundant in microbicidal azurophilic granules of neutrophils and is also found in the epithelia of mucosal surfaces such as those of the intestine, respiratory tract, urinary tract, and vagina, possessing fungicidal, antiviral and antimicrobial activity against Gram-negative and Gram-positive bacteria. Defensins are thought to kill microbes by permeabilizing their plasma membrane. Three major theories explain this phenomenon - the barrel model, the torroidal model and the carpet model (Figure 8) [114].

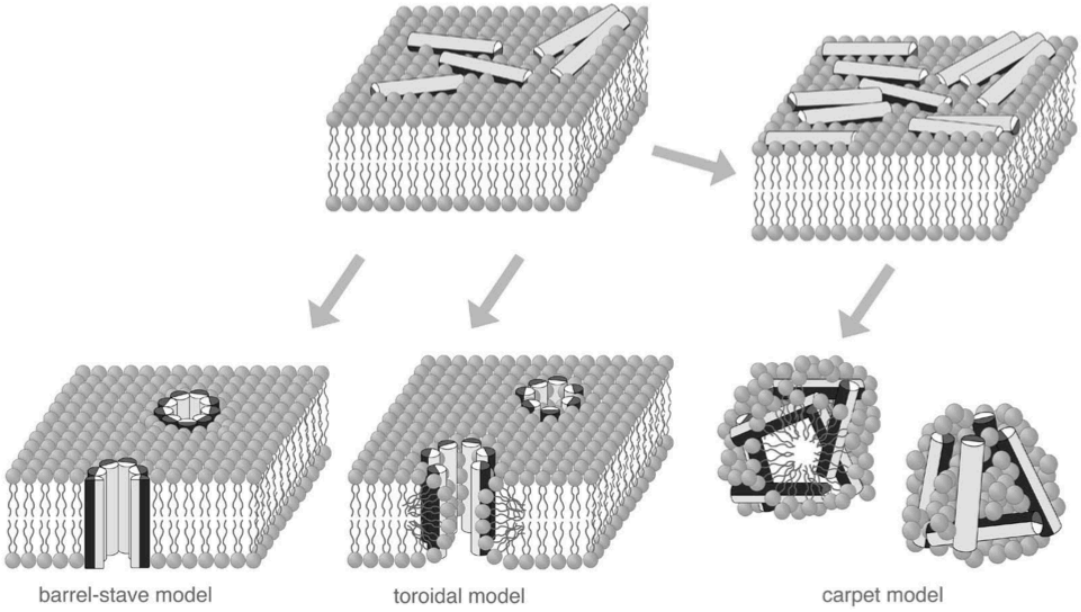


Figure 8. Three models explaining possible mechanism of action of cationic antimicrobial peptides. Hydrophobic parts are depicted in black, hydrophilic regions are show in grey. Figure kindly provided by prof. Vilcinskas et al. [115].

Members of the defensin family are highly similar in protein sequence - neutrophil defensin-1 differs from the defensins-2 and -3 by only one amino acid (Figure 9).

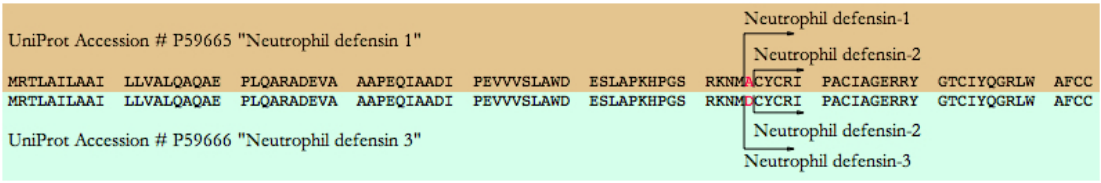


Figure 9. Sequence comparisons of Neutrophil defensin-1 and -3. Primary sequences of both precursors differ only in one amino acid (alanine vs. aspartic acid). Moreover, these precursors undergo processing into final acting peptides.

Another group of proteins, which showed markedly higher levels in the positive group, are matrix degradation proteins. The physicochemical properties of intracellular matrix form a barrier for immune system cells, which tend to infiltrate the inflammation epicenter and use proteolytic enzymes - proteases - to break the matrix proteins. This process is enhanced by cytokines and enables the immune system cells to escape the blood vessel system and invade the extravascular space. The most important proteases used for this purpose include serine proteases, cysteine proteases and MMPs. A substantial number of proteins from each group were detected in our study and are outlined in Table 4.

Most out of 21 serine proteases identified in our work did not show marked differences, but we did observed increased levels of neutrophil elastase (ELA2), myeloblastin (proteinase 3, PRTN3) and azurocidin (AZU1) in the positive patient group. Just like the above-described defensins, these serine proteases are located in azurophilic granules of neutrophils, from which they are released during the inflammatory process. The vast and often devastating effects of serine proteases need to be regulated using specific inhibitors, not to cause collateral damage. More than 20 serine protease inhibitors, including α_2-macroglobulin, α_1-antitrypsin and antileukoproteinase were identified, but the majority of them do not show dysregulation or have just slightly lower levels in the positive group, with several exceptions described below.

	Protein name	Change	Ref.
Serine proteases & inhibitors	Azurocidin	↑	[116]
	Neutrophil elastase	↑	[117-119]
	Proteinase 3	↑	-
	Antileukoproteinase	-	[120, 121]
	Leukocyte elastase inhibitor	↑	[116]
	α_1-antitrypsin	-	[122]
	α_2-macroglobulin	-	-
Cysteine and aspartic proteases & inhibitors	Cathepsin B	↑	-
	Cathepsin D (aspartic)	↑	-
	Cathepsin H	↑	-
	Cathepsin L	↑	-
	Cystatin A, B, C, M, SA	-	-
MMPs & inhibitors	Aminopeptidase N	↑	-
	Matrix metalloproteinase-2	-	[123]
	Matrix metalloproteinase-7	-	-
	Matrix metalloproteinase-8	↑	[67, 68]
	Matrix metalloproteinase-9	↑	[124, 125]
	TIMP-1	-	[126]
	TIMP-2	-	[123]

Table 4. Selected proteases and protease inhibitors detected in the analysis. We detected representatives from all major proteases groups and a large portion of these showed dysregulation.

During intraamniotic infection and inflammation, the AF ELA2 levels were shown to rise in several studies [117, 119]. This protein has also been suggested as a potential marker of histologic chorioamnionitis, both in sPTL patients with intact membranes as well as in PPROM cases [118]. There are two major inhibitors, which modulate its activity, antileukoproteinase (SLPI) and leukocyte elastase inhibitor (ILEU), both of which were identified in our study. The role of SLPI is well described in the literature and suggests that SLPI is crucial for preventing ELA2-caused fetal membranes rupture, as a drop in SLPI levels causes rupture of the membranes, both at term as well as during preterm birth [118, 120]. In PPROM, the SLPI levels drop relatively to the amount of ELA2, which results in increased activity of this enzyme and finally in membrane rupture [118, 120]. In agreement with this fact, we observed a relative drop of SLPI compared to ELA2 in the PPROM cohort. Interestingly, we confidently identified ILEU and observed a two fold increased level of this protein [127, 128]. In contrast to SLPI, no information is available on ILEU and its role during intraamniotic infection and

inflammation in PPROM patients. The only study relating increased ILEU to AF and ongoing infection was another proteomic study in sPTL patients with intact membranes carried out by Gravett et al. [116].

However, it turned out, that ILEU inhibits not only ELA2, but also PRTN3, another matrix protease released by the neutrophils and performing similar tasks as ELA2 - degrading elastin, fibronectin, laminin, vitronectin etc. In analogy to ELA2, also PRTN3 was found to be markedly elevated in the positive group. Although no studies were carried out to evaluate the performance of PRTN3 in detecting chorioamnionitis, we speculate that due to the similar function, binding partners and specificity, next to ELA2 this protein might be another potential candidate for detection of chorioamnionitis, especially regarding the exceptional performance of ELA2, which even enables stratification of the patients based on the grade of infection [117, 119].

AZU1 was another protein categorized as serine protease representative, which was detected to be highly upregulated in the positive patient group. As indicated by its name, AZU1 is an azurophil granule antimicrobial glycoprotein. It is an important multifunctional inflammatory mediator with chemotactic and antibacterial activity. Although this protein is a member of the serine protease gene family, it is not a true serine proteinase by function, because the active site serine and histidine residues are replaced. Upon release from azurophil neutrophilic granules, it acts as a chemoattractant and activator of monocytes and macrophages. The functional consequence is enhancement of cytokine release and bacterial phagocytosis, allowing for a more efficient bacterial clearance. In addition, AZU1 activates endothelial cells, which leads to vascular leakage and edema formation. The cytotoxic action is limited to many species of Gram-negative bacteria, whereby this specificity may be explained by a strong affinity of the very basic *N*-terminal portion of the protein towards the negatively charged lipopolysaccharides that are unique to the Gram-negative bacterial outer envelope [129]. Interestingly, the genes encoding for *azu1*, *ela2* and *prtn3* are in a cluster located at chromosome 19. All 3 genes are expressed coordinately and their protein products are packaged together into azurophil granules during neutrophil differentiation [130, 131]. Similarly to ILEU, there is only one study mentioning AF AZU1 to intraamniotic infection and inflammation, carried out in sPTL patients with intact membranes [116]. Our results in PPROM cohort are thus in agreement with these findings.

We identified several cathepsin family members, namely cathepsin B, H and L uniformly elevated. These members of cysteine proteases take advantage of cysteine to

initiate protein degradation, which occurs in the highly conserved catalytic core of the enzyme. Cathepsins differ in substrate specificity, i.e. cathepsins L and S are elastin-specific. Similarly to serine proteases, cysteine proteases are also regulated by specific inhibitors, cystatins. We successfully identified several members of this inhibitor family (cystatin A, B, C, M, SA), but none showed significantly altered levels. According to our knowledge, this is the first time where the role of increased AF levels of cysteine proteases, cathepsins in particular, is suggested in intraamniotic infection and inflammation.

The last representatives of matrix degrading enzymes detected in our experiments are MMPs, a family of zinc-dependent enzymes that hydrolyze a wide range of extracellular matrix constituents [132]. They are also known to be involved in the cleavage of cell surface receptors, the release of apoptotic ligands and chemokine/cytokine activation and/or inactivation [133]. Thus, MMPs are thought to play a major role in cell proliferation, migration, adhesion, differentiation, angiogenesis, apoptosis, and host defense. The family is divided into subfamilies that are based on substrate specificity.

The interstitial collagenases (MMP-1 and MMP-8) degrade collagen types I and III. The levels of MMP-1, which is produced by numerous cell types, as well as of MMP-8, which is produced exclusively by neutrophils, was found to be increased in pregnancies complicated by infection related sPTL as well as PPROM [67, 134]. MMP-8 was even suggested as a potential target for a bedside test for detection of intraamniotic infection and inflammation in AF [68]. We were able to confirm these findings as we observed a significant increase in AF MMP-8 levels in the positive patient group.

The gelatinases (MMP-2 and MMP-9) are ubiquitously expressed and degrade denatured interstitial collagen, i.e. gelatin, and molecules that are associated with epithelial basement membranes, including collagen types IV and V, laminin, and fibronectin [132]. According to published data, MMP-2 is expressed constitutively and its levels do not change with the presence of MIAC [123]. The levels of MMP-9, on the other hand, were shown to be tightly associated with the presence of intraamniotic infection and inflammation, both *in vivo* [124] as well as *in vitro* [125] evaluations. We detected both gelatinases and detected aberrant expression of MMP-9, which was present at levels comparable with MMP-8. In concordance with previous findings, we show that MMP-2 levels do not change between both groups. In analogy to serine and cysteine proteases, also MMPs have specific inhibitors regulating their catalytic activity - tissue

inhibitors of metalloproteinases (TIMPs). We identified TIMP-1 and -2 in our analysis, but observed no alterations in their levels [123, 126].

During MMP results analysis, yet another interesting enzyme caught our attention. Aminopeptidase N (ANPEP) does not resemble the routine metalloproteinase name, despite being a member of this family, as the large extracellular carboxyterminal domain contains a consensus sequence characteristic for members of the zinc-binding metalloproteinase superfamily. ANPEP is membrane-anchored protein located in the small-intestinal and renal microvillar membrane. In the small intestine ANPEP plays a role in the final digestion of peptides generated from hydrolysis of proteins by gastric and pancreatic proteases. Its function in proximal tubular epithelial cells and other cell types is less clear. It is thought to be involved in the metabolism of regulatory peptides originating in diverse cell types, including small intestinal and renal tubular epithelial cells, macrophages and granulocytes. ANPEP is used as a biomarker of kidney damage, as high levels of this protein in urine are found in various kidney pathologies. Even though no specific information was found on the role of ANPEP during intraamniotic infection and inflammation, we speculate that increased levels of this protein might be caused by severe inflammation in the fetal kidneys. The fetal urine might thus be the source of this peptidase in AF.

Similarly to ANPEP, neither leukotriene A4 hydrolase (LTA4H) is a typical MMP representative. It does however contain typical catalytic zinc binding site and the sequence comparison with several other zinc hydrolases/MMPs showed large common sequence regions [135, 136]. Even the peptidase mechanism action was shown to be identical to that of ANPEP [137]. The biological role of the peptidase activity of LTA4H is not known but we speculate that the aminopeptidase activity may be involved in the processing of peptides related to inflammation and host-defense, especially regarding the sequence and functional similarities with ANPEP as well as other MMPs. Another reason, which leads us to this assumption, is the much better described role of LTA4H in leukotriene metabolism, where it acts as a hydrolase. Based on cell type, an unstable precursor, epoxide leukotriene A4 (LTA4), is enzymatically converted into final acting leukotrienes. In cells, which produce LTA4H, i.e. neutrophils or monocytes, LTA4 is converted to the dihydroxy acid leukotriene B4 (LTB4), which is a powerful chemoattractant for neutrophils acting at B-leukotriene (BLT)-1 and -2 receptors on the plasma membrane of these cells [138]. In cells that express leukotriene C4 (LTC4) synthase, such as mast cells and eosinophils, LTA4 is conjugated with glutathione to

form the first of the cysteinyl-leukotrienes, LTC4 (Figure 10). Outside the cell, LTC4 can be converted to LTD4 and LTE4, which retain biological activity. The cysteinyl-leukotrienes act at their cell-surface Cys-leukotriene(CysLT)-1 and -2 receptors on target cells to increase permeability of small blood vessels, to enhance secretion of mucus and to recruit leukocytes to sites of inflammation [139].

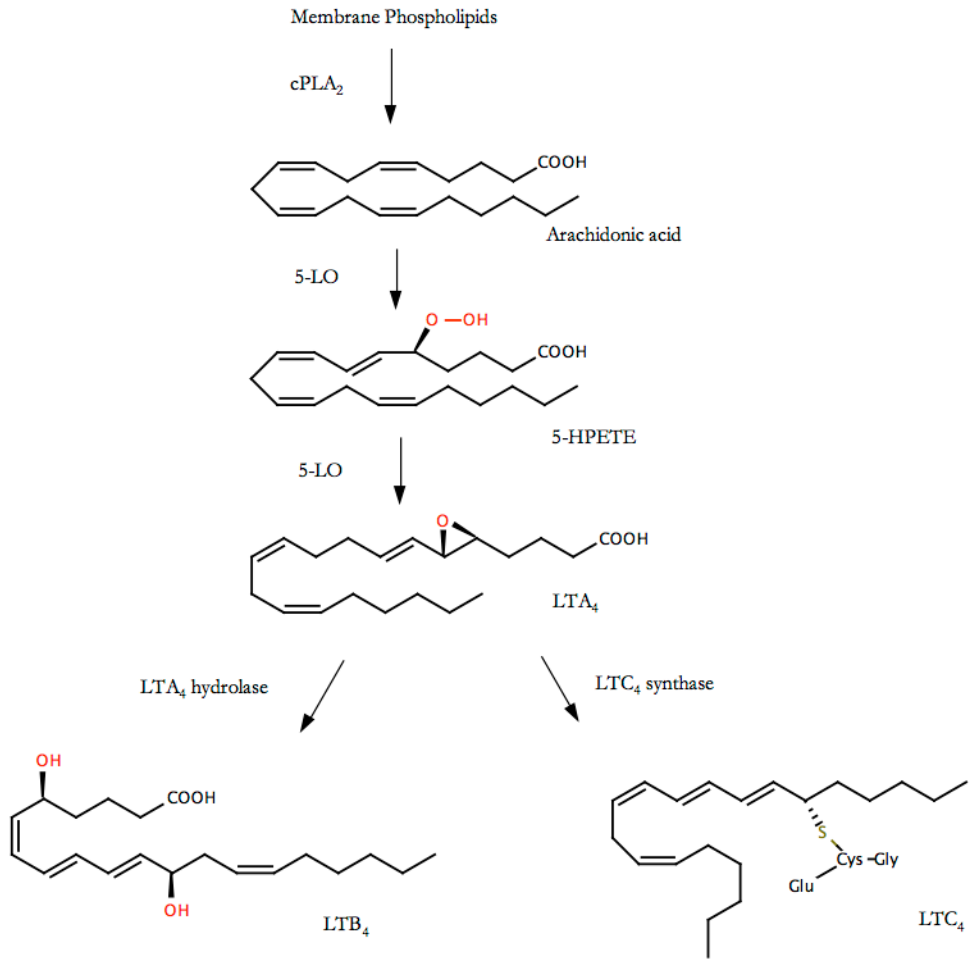


Figure 10. Leukotriene metabolism. Cytosolic phospholipase A2 ($cPLA_2$) cleaves phospholipids and produces arachidonic acid. 5-lipooxygenase (5-LO) then processess arachidonic acid into 5-hydroperoxyeicosatetraenoic acid (5-HPETE) and further into LTA_4, which is processed based on the respective cell type into either LTB_4 or LTC_4.

Calcium-activated chloride channel regulator 1 (CLCA1) was confidently identified and showed a two-fold increase in the positive patient group. Even though the name suggests its principal role in ion transfer, several pathway analysis tools indicated links to peptidase activity. Indeed, it has been recently described that this protein appears to be a membrane anchored metal-dependent hydrolase, possibly protease. A metallohydrolase structural domain was predicted, unexpectedly, in the CLCA sequences

and confirmed by a comparison modeling analysis using MMP-11 as template [140]. The authors even doubted the protein role as a chloride channel. This was supported by a study, where the authors did not detect any transmembrane domain in the protein sequence [141], but showed that the protein contained membrane anchor domains and domains for protein-protein interaction. In another study, CLCA1 was found to be a membrane bound epithelial protein and the authors identified integrin alpha as a binding partner at the surface of leukocytes and suggested that CLCA1 might mediate leukocyte adhesion [142]. We detected remarkably high levels of integrin alpha-M (ITGAM) in our experiments, which corresponds to these previous findings. This also suggests, that both CLCA1 and ITGAM might be released from both epithelial cells and leukocytes and detected in AF, signaling occurring inflammatory process. Although it has been suggested that CLCA1 might be both membrane bound as well as actively secreted, the source of CLCA1 in AF remains unclear [140, 141]. The protein was detected in both bronchoalveolar lavage fluid and lung tissue epithelia taken from asthmatic patients [141, 143]. With respect to the findings that CLCA1 expression is limited and was found to be restricted to small intestine, colon and airways epithelium [144-146], we speculate that the increased AF level of this protein might originate in the fetal digestive and/or respiratory tract and could thus be a result of FIRS. Our assumption is supported by the fact, that we detect elevation in both circulating immune system cells specific protein, ITGAM, as well as in its binding partner, CLCA1, specifically produced in gastrointestinal and lung epithelia.

We observed a ~6-fold increase in resistin, a 12.5 kDa cysteine-rich protein, which was suggested to play principal role in insulin resistance and obesity as well as during infection and inflammation. This cytokine is produced by leukocytes and adipocytes and increases the expression of several proinflammatory agents, including IL-1, IL-6 and TNF-α in an NFκB-mediated fashion [147]. It has been also shown, that resistin, similarly to calgranulins, upregulates adhesion molecules ICAM1 and vascular cell-adhesion molecule-1, both of which play a role in pathways of leukocyte attraction and recruitment to sites of infection [148]. Resistin itself can be upregulated by interleukins and also by PRR-mediated recognition of microbial antigens such as LPS [149]. Interestingly, resistin has been shown to compete with LPS at TLR-4 receptor, suggesting a certain degree of autocrine and paracrine signaling [150]. During pregnancy, resistin levels in maternal blood rise with gestational age, most probably due to its involvement in fetal growth [151]. A recent study by Kusanovic et al. assessing AF

resistin levels with respect to intraamniotic infection and inflammation showed significantly elevated levels of this cytokine in patients with inflammation compared to matched negative controls. Moreover, patients who had MIAC and histological signs of inflammation were presented with an additional resistin increase compared to patients with histological inflammation but were negative for MIAC. The resistin dysregulation was observed in both clinical phenotypes of preterm birth, sPTL as well as PPROM [152].

Several proteins discussed above possess antimicrobial properties and are thus capable of killing microorganisms by either forming pores in their cell membrane or by specific interaction with intracellular components of the microbe [115]. After that, antimicrobial peptides exert various effects on bacterial functions *via* changes in cell membrane synthesis or composition, inhibition of various enzymes, inhibition of protein or nucleic acid synthesis, binding to DNA or by pore formation and membrane destabilization [153].

A well known, ~20 kDa protein cathelicidin was confidently identified and showed ~6-fold increase in patients with positive intraamniotic findings. Cathelicidin serves a critical role in mammalian innate immune defense against invasive bacterial infection and is the only member of the cathelicidin family found in humans. This family shares sequence homology with cystatins, cysteine protease inhibitors, which also possess antimicrobial properties [154]. It has been even suggested, that cathelicidin originates in the cystatin proteins superfamily [155]. It is produced in epithelial cells, macrophages and most importantly neutrophils upon stimulation by bacteria, viruses and, interestingly, 1,25-dihydroxyvitamin D - an active form of vitamin D. Cathelicidin is secreted in high amounts in tissues exposed to environmental microbes, particularly in those with squamous epithelia (mouth, tongue, cervix, vagina, esophagus etc.) or in derived fluids [156]. Decreased production of cathelicidin has been linked to diseases, where the common denominator is enhanced susceptibility of infection [157, 158]. A similar effect was observed in an animal model, as cathelicidin-deficient mice were more prone to skin and urinary tract infection [159, 160]. The antimicrobial effect was tested and confirmed experimentally in body fluids, including AF or urine [159, 161]. Similar findings also lead to elucidation of the "antimicrobial" effect of vitamin D. Since the 19th century, liver oil, eggs or sunlight were used in tuberculosis therapy. Even after the isolation and identification of vitamin D from cod liver, which enabled therapeutic employment of vitamin D, the exact mechanism of action in fighting infection remained unexplained

[162]. It was not known until the 1980, where vitamin D was found to boost antimicrobial activity of human monocytes against *Mycobacterium tuberculosis* [163]. The proposed mechanism of action is triggered by TLR2/1 activation, which leads to production of 25-hydroxyvitamin D-1 α -hydroxylase, which in turn converts circulating inactive 25-hydroxyvitamin D into active 1,25-dihydroxyvitamin D. This active form finally binds to vitamin D receptor, a transcription factor from the steroid receptor family and activates cathelicidin gene transcription [164].

The association between cathelicidin and vitamin D may be also regarded from another point of view. While vitamin D promotes antimicrobial agent production, it also has anti-inflammatory effects [165]. Even the "executing" component of the antimicrobial effect, cathelicidin, was shown to have anti-inflammatory influence [166]. A similar behavior can be observed in cystatins, which block cathepsins in order to limit their pro-inflammatory and destructive effects and thus exert anti-inflammatory function, but at the same time they possess antimicrobial properties. Is there a link between these functional similarities and evolutional homology? Several studies have shown, that maternal vitamin D deficiency is associated with a range of pregnancy related morbidities and adverse neonatal outcome [167-171]. It can be speculated, that low levels of vitamin D may result in induction of non-infectious inflammation and/or in impaired production of antimicrobials, which in turn leads to reduced ability to face microbial invasion. Given the fact that infection and/or inflammation are one of the key components of causes leading to preterm birth, low vitamin D levels might be associated with increased risk of preterm labour [172].

Moreover, African-Americans were shown to be at significantly higher risk for vitamin D insufficiency or even deficiency compared to other races. According to Center for Disease Control and Prevention, Atlanta, GA, nearly half of black US women in productive age are deficient in vitamin D [173]. This goes very well with the fact that African-American women are threatened by preterm birth far more often than Caucasian, Asian or Hispanic mothers. One possible explanation suggests, that high melanin content absorbs UVB radiation in the skin and thus diminishes vitamin D synthesis [174].

Two possible explanations of the causes of increased cathelicidin levels in our study may be deducted. The first alternative suggests that microbial components are detected by a range of TLRs, which execute a specific response, based on the TLR subtype. TLR 2/1 would then signal through the above described pathway and increase

the amount of cathelicidin to be produced and secreted in order to cope with the microbial invasion.

The other point of view goes back and suggests vitamin D insufficiency as one of the causative agents of the very infection and inflammation. In our study, we detected very high levels of AF vitamin D binding protein (VDBP) in general. In fact, this protein was the third most abundant protein in AF based on number of detected peptides. VDBP acts as a carrier protein for vitamin D and its metabolites and serves for their transport to target cells and tissues. The positive patient group had slightly, but significantly decreased levels of VDBP compared to negative patients. Given the high concentration of this carrier protein in AF, we speculate that even a minute drop in its concentration might result in decreased availability of vitamin D and thus in increased susceptibility to infection as well as inflammation. The observed rise in cathelicidin concentration in the positive group, on the other hand, could be caused by induction of cathelicidin production by other cytokine, i.e. IL-6. Another powerful mediator, which promotes the production and release of cathelicidin from neutrophils, is LTB4. As described above, we detected significantly increased levels of LTA4H, the enzyme responsible for LTB4 production.

Dermcidin, another member of peptides with microbicidal effects identified in our work showed substantial drop in concentration in the positive group. This trend was a rare direction of change, regarding the fact that the majority of proteins involved in microbe killing were upregulated. This peptide is secreted in large quantities in sweat and is a principal constituent of the skin chemical barrier [115, 175]. It has been also detected in breath condensate and in bronchoalveolar lavage fluid in studies focused at asthma [176, 177]. It was shown that dermcidin is just a precursor protein, which gives rise to a multitude of active antibiotical peptides upon secretion and proteolytic processing [178]. These generated peptides are both cationic and anionic and are capable of killing a wide range of microbes, both Gram-postive and -negative, including *Staphylococcus aureus, E. coli, Enterococcus faecalis, Candida albicans, S. epidermidis* or *Pseudomonas putida.* The mechanism by which dermcidin far less clear compared to i.e. cathelicidin, which is known to interact with the bacterial cell wall forming transmembrane pores which results in disruption of the membrane. Although dermcidin showed affinity to the bacterial cell envelope, albeit a rather weak one, and was shown to be able to kill Gram-negative bacteria by membrane depolarization, this was not achieved *via* membrane pore formation [179]. The authors observed a similarly weak affinity to components of the

Gram-positive cell structures. In contrast, cathelicidin binds strongly to these structures in a dose-dependent fashion. This suggested that the mode of action is different from that of cathelicidin and most likely from other pore-forming antimicrobial peptides. This was confirmed also by the observation that while pore-forming peptides kill the microbes within several first minutes, the kinetics observed during *in vitro* interaction of dermcidin with *S. aureus* indicated that the killing is much slower and takes at least one hour. Interestingly, dermcidin inhibited bacterial macromolecular biosynthesis, but did not bind directly to bacterial DNA or RNA. Unfortunately, the authors were not able to fully explain the mechanism of action [180]. Another study showed that cathepsin D, an aspartic protease, participates in dermcidin precursor processing into final acting peptides [181]. We detected about two-fold increase in this protein compared to negative controls. Whether the surprising drop in dermcidin in the positive patient group is a consequence of the immune system response or a cause of the actual MIAC and subsequent intraamniotic infection and inflammation, remains to be elucidated.

Mentioning histone proteins in the last section of the antimicrobial acting proteins might seem odd. It is well known, that histones are core components of nucleosomes and play a key role in eukaryotic chromosome stabilization, DNA replication and repair, transcription regulation etc. [182]. While histone H1 keeps the nucleosome structure condensed and compact, the core histones (H2A, H2B, H3 and H4) aggregate in an octamer and form the very nucleosome [183]. We detected all five types in our analysis and more importantly, all of these proteins showed high upregulation in the positive group. In fact, these proteins showed some of the most profound changes in the whole dataset. Surprisingly, it turned out that aside their role in the nucleus, histones as well as their fragments possess broad antimicrobial activity. This function seems to be highly conserved, as the same effect was confirmed across all vertebrates, including fish, frog, chicken and mammals [184-186]. Histone H1 was detected in a wide range of mammalian tissues and demonstrated a broad-spectrum activity against both Gram-positive and -negative species, including drug resistant ones, such as methicillin-resistant *S. aureus* (MRSA) [187]. Other studies proved antifungal and even antiviral activities, suggesting that histone H1 inhibits virus attachment to the cell. All these effects were observed in both native as well as in recombinant forms of the protein and were without any hemolytic or cytotoxic side effects [188, 189]. Histone H2A antimicrobials are the best described histone family mainly due to well known fragments found in frogs and fish (buforins, parasins or hipposins). However, even full-

length antimicrobial H2A was isolated from human placenta as well as other sources [190]. Similarly to H1, also this second type of histones exerts wide-spectrum antimicrobial activity against both Gram-positive and -negative species. Histone H2B was isolated from human placenta along with H2A and from intestinal epithelia along with H1 and was described to show anti-*E. coli* activity. In general H2B has similar broad-spectrum activity as previous two mentioned histones, but in addition seems to possess antiparasitic properties [191]. Compared to H2A, histone H2B has significantly higher affinity towards LPS and is thus suggested to inhibit the endotoxin activity of LPS, as shown in a study assessing the bactericidal activity against pathogenic bacteria in AF [190]. Histone H2B also relies on a different strategy when interacting with microbes as opposed to H1 or H2A. The latter two use, at least partially, certain endogenous proteases, which cleave the full-length protein into various acting fragment peptides. A study of H2B action against an *E. coli* strain defective in outer membrane proteinase T gene showed impaired penetration into the bacterial membrane, which apparently abolished the antimicrobial function of H2B. Based on these findings, it has been suggested that rather than using an endogenous protease, histone H2B curiously takes advantage of the pathogen and seems to be cleaved by an exogenous protease [192].

While the above-mentioned histones are well described for their antimicrobial effects, the role of histone H3 and H4 in this area is far less obvious. A study assessing not the antimicrobial activity itself, but LPS binding ability, suggested their role in host defence [193]. Recently, histone H4 was described to be a principal antimicrobial component of human sebatocytes [194]. These pieces of evidence thus suggest, that also H3 and H4 might possess means for bacterial killing, similar to remaining histone proteins.

Except the nucleus, there is, however, a very interesting structure where histones H3 and H4 are found along with remaining histone family members, as well as with several others above described proteins. Next to secretion of antimicrobial agents and engulfment of microbes, neutrophils were recently found to possess yet another means for killing pathogens - formation of neutrophil extracellular traps (NETs). NETs are actively produced upon neutrophil activation by IL-8, LPS or phorbol myristate acetate (PMA), but not by naïve cells, in a fashion, which is distinct from apoptosis or necrosis [195, 196]. Several arguments support the theory that NETs are actively "constructed"; it was demonstrated that NETs are made by motile cells; neutrophils start to produce NETs as early as 10 min from activation, which is much faster than apoptosis; cytokines

and antigens, which promote life of neutrophils also activate NET formation; this process is not accompanied by DNA fragmentation. With regard to these findings, NETs are not a result of a leakage caused by cell rupture. On the other hand, as neutrophils are terminally differentiated cells and usually die within several hours after entering the circulation, the NET formation might be an early event in the programmed neutrophil death [197]. Evidence points out to the fact, that several constituents of NETs are in fact required for the trap formation. In particular, ELA2 and MPO were shown to be critical in the production. Although the exact mechanism is unknown, the authors speculate that upon neutrophil activation, ELA2 is released from the granules, migrates to the nucleus, cleaves histones and promotes chromatin decondensation. MPO binds to the chromatin structures later in the process and promotes further decondensation [198]. This was supported by the findings that ELA2 knockout mice were unable to create NETs in lungs upon infection.

The composition analysis of these trapping arrangements revealed that rather than being based on classical cytoplasmic structural proteins, like actin or tubulin, these unique structures employ DNA and granular proteins as their building blocks. The DNA strings forming the web are decorated by proteins originating in neutrophil azurophilic granules (i.e. ELA2, AZU1, MPO), cytoplasm (i.e. S100 proteins) or nucleus (histone proteins) [197, 199, 200]. Proteins reported to be involved in NETs are outlined in Table 5. The constructed traps bind Gram-negative as well as -positive bacteria as well as fungi, preventing them from spreading. Additionally, high local concentration of antimicrobials and degrading enzymes, which are incorporated into the webs, enables disarmament and direct killing of the pathogens. In contrast to plain release of these degrading enzymes, the collateral damage is much lower if these proteins are attached to the produced scaffold. A particular theory justifying the existence and necessity of NETs suggests that these structures enable efficient capturing for subsequent destruction of microorganisms too large to be engulfed and phagocytized by the neutrophils [200].

Protein	Cellular localization	Detected	Dysregulated	Reference
Azurocidin	Azurophilic granules	Yes	↑	[200]
Permeability-increasing protein	Azurophilic granules	Yes	↓	[197, 200]
Catalase	Peroxisome	Yes	↑	[200]
Cathelicidin antimicrobial peptide	Azurophilic granules	Yes	↑	[201]
Cathepsin G	Azurophilic granules	No	-	[197, 200]
Histone H1	Nucleus	Yes	↑	[197, 200]
Histone H2A	Nucleus	Yes	↑	[197, 200]
Histone H2B	Nucleus	Yes	↑	[197, 200]
Histone H3	Nucleus	Yes	↑	[197, 200]
Histone H4	Nucleus	Yes	↑	[197, 200]
Lactotransferrin	Granules	Yes	↑	[197, 200]
Lysozyme C	Azurophilic granules	Yes	↑	[200]
Matrix metalloproteinase-9	Granules	Yes	↑	[197, 199]
Myeloblastin	Azurophilic granules	Yes	↑	[200]
Myeloid cell nuclear differentiation antigen	Nucleus	Yes	↑	[200]
Myeloperoxidase	Azurophilic granules	Yes	↑	[197, 200]
Neutrophil defensin 1	Azurophilic granules	Yes	↑	[200]
Neutrophil elastase	Azurophilic granules	Yes	↑	[197, 200]
Pentraxin 3	Granules	Yes	↑	[199]
Peptidoglycan recognition protein 1	Granules	Yes	↑	[202]
Plastin-2	Cytoplasm	Yes	↑	[200]
Protein S100-A12	Cytoplasm	Yes	↑	[200]
Protein S100-A8	Cytoplasm	Yes	↑	[200]
Protein S100-A9	Cytoplasm	Yes	↑	[200]

Table 5. Overview of NETs components. Most of the proteins involved in NETs formation were identified in our study and in concordance with the NETs theory, the majority of them were present in higher levels compared to negative patient group.

In addition, not only do these chromatin/protein complexes aid in microbe trapping and immobilizing, but also the very presence of DNA seems to be critical for proper functioning of the attached proteins. This was demonstrated by protease-free DNase treatment, which did not affect any of the protein constituents, but lead to a dramatic reduction of microbe killing capability. Similar effect was observed, when an H2A-H2B-DNA complex antibody was added to the activated neutrophils [197]. The authors suggested that along other well-known antimicrobials involved in NETs, histones are of particular importance in the bactericidal effect of these traps, once again pointing out on the exceptional role of these proteins.

In our analysis, we detected the far majority of protein so far described to be a part of NETs. More importantly, all of our detected proteins showed strong dysregulation, mostly upregulation. The question, whether these changes in AF proteins are caused by simple neutrophil degranulation, NET formation, or both, however, remains to be addressed. The presence of NETs *in vivo* has been confirmed in a host-pathogen interaction model, spontaneous appendicitis, atopic asthmatic airways or bacterial pneumonia [197, 203, 204]. The only evidence found for pregnancy-related disorders, however, suggests possible role of NETs in the pathophysiology of preeclampsia [205, 206]. The authors found broad infiltration of the intravillous spaces by NETs in preeclamptic placentae. Interestingly, in this case the release of traps was stimulated by syncytiotrophoblast microparticles and the results were comparable to neutrophil activation by IL-8 or PMA.

Our findings for the first time suggest the presence of these supracellular trapping structures in AF of patients affected by intraamniotic infection and inflammation.

4.3. Specific aim 3 - Identification of novel potential intraamniotic infection and inflammation biomarkers in amniotic fluid from patients with sPTL with intact membranes using multidimensional comparative proteomics

The strategy as well as the workflow used in the analysis of AF samples form sPTL patients with intact membranes was identical with that used in PPROM patients AF analysis.

In the discovery phase of the project, we selected representative surrogate samples from the sPTL AF sample set, accepting only samples from patients diagnosed and described with high confidence. Thirty-one samples from patients with confirmed MIAC and with confirmed HCA were selected and used as a positive group. As a control, we included 26 AF samples from patients, where both MIAC and HCA were ruled out. Samples in both groups were matched for gestational age, maternal age, parity and smoking status in order to eliminate potential bias (Table 6).

	The presence of both MIAC and HCA (n=31)	The absence of both MIAC and HCA (n=26)	*p*-value
Maternal age (years)	26.8±6.1	26.4±6.2	0.78
Nulliparous	16 (52%)	18 (69%)	0.29
Smoking in pregnancy	7 (23%)	6 (26%)	1.00
Birth weight (grams)	2099±774	2238±720	0.49
Apgar score in 1 minutes	8 (2-9)	8 (1-9)	0.91
Apgar score in 5 minutes	9 (7-9)	8 (2-9)	0.64
Clinical chorioamnionitis	31 (100%)	4 (15%)	<0.0001
Apgar score in 10 minutes	7 (23%)	0 (0%)	0.01

Table 6. Maternal and newborn characteristics based on the presence and absence of both MIAC and HCA - sPTL cohort. Continuous variables were compared using parametric t-test (presented as mean ± SD) or a nonparametric Mann-Whitney U test, presented as median (range). Categorical variables were compared using Fisher exact test and presented as number (%).

Upon generating pooled positive and negative samples and immunoaffinity depletion, we applied the CysTRAQ protocol. Both cysteinyl and non-cysteinyl peptide fractions were subjected to basic pH RP fractionation and all resulting fractions were analyzed using the LC-MALDI approach (Figure 6). This multidimensional quantitative proteomics workflow lead to identification of 12 777 distinct peptides, which were identified using more than 15 000 MS/MS spectra. Based on these peptides, 846 AF proteins were identified (all data presented at 5% FDR rate). The iTRAQ labeling enabled simultaneous relative quantitation of abundance changes across individual samples. We used the iTRAQ 4-plex version and thus we were able to analyze two

replicates in a single analysis as shown in. For all proteins, we annotated the matching protein class using the PANTHER database. The complete list of these proteins is included in Table 8, section External tables.

The total number of identified proteins was comparable between both cohorts, which was expected due to the fact that we used the same workflow. What was more surprising was the number of dysregulated proteins at various levels of statistical significance (Table 3). The underlying cause for the much more profound proteome dysregulation in the PPROM cohort is unclear. We speculate, that one of the reasons for this striking difference may be the fact that a much wider range of factors and causes may result in membrane rupture. sPTL, on the other hand, is initiated by the onset of contractions, and thus also the number of pathways involved could be lower, resulting in a much more specific dysregulation.

Similarly to the PPROM cohort, we used analogous criteria for dysregulated proteins filtering and subsequent candidate selection: $p \leq 0.01$ in both replicates and an at least 2-fold abundance change, either upwards and downwards. This, however, resulted in a list of just three candidates and we thus decided to include also several additional proteins which did not meet this initial criterion, but showed significant dysregulation in both replicates ($p \leq 0.05$; included in Table 9, External tables).

We identified SLPI in both sPTL and PPROM cohorts. While we did not observe any change in SLPI levels in PPROM patients, we detected a significant rise in its levels in sPTL patients. This is in agreement with published data, as increased levels of SLPI are required during increased protease activity, especially that of ELA2, as SLPI is the natural inhibitor of this enzyme. SLPI was reported to be increased in intact membrane patients, both preterm as well as term and is thus apparently required to prevent membrane rupture [120].

Carboxypeptidase M was identified to be significantly increased in the positive patient group. This protein is supposed to play a crucial role in peptide hormone and growth factor activity regulation due to its specific *C*-terminus proteolytic activity. Interestingly, neutrophil defensins, which are known to be involved in intraamniotic infection and inflammation, need to be processed from their inactive form in order to gain their antimicrobial activity. Sequence specificity analysis of carboxypeptidase M revealed, that the defensins contain target sequences, which are cleavable by this enzyme [207]. Whether the increased levels of carboxypeptidase M are associated with increased defensin levels is sPTL AF however unknown. Importantly, carboxypeptidase M was

found to be produced in neutrophils upon activation and could thus share the production location with defensins [208].

Carboxypeptidase A1, along with bile salt-activated lipase were one of the most highly dysregulated proteins from the whole sPTL dataset. Both proteins are produced in and secreted from the pancreas. Upon secretion, bile salt-activated lipase aids in digestion of fats, while carboxypeptidase A1 is an exopeptidase and catalyzes the release of the *C*-terminal amino acid [209, 210]. While no specific information is available on the role of these two enzymes during intraamniotic infection and inflammation, carboxypeptidase A1 was recently shown to be implicated in both acute and chronic pancreatitis [211]. Could the increased levels of bile salt-activated lipase coincide with those of carboxypeptidase A1? And if so, could the high levels of both proteins in the positive patient group be due to fetal pancreas inflammatory response? Interestingly, both proteins were detected in a recent AF proteomic study with regards to searching for Down syndrome and both of these proteins were found to be dysregulated [76]. The authors even confirmed decreased levels of both proteins using SRM in AF [75].

Peptidoglycan recognition protein 1 (PGLYRP-1) was another dysregulated molecule, which caught our attention. Being a part of the innate immunity, PGLYRP-1 is a specific PRR, which was found to be highly conserved in evolution and has been detected across vertebrates, as well as in invertebrates [212, 213]. Individual PGLYRP family members (PGLYRP1-4) have various functions. To illustrate, while PGLYRP-1 is present in granulocytes and likely participates in killing phagocytized bacteria, PGLYRP-2 is constitutively produced and secreted into the bloodstream from liver [214, 215]. The remaining two PGLYRPs (-3 and -4) share antimicrobial properties with PGLYRP-1, but the spectrum of individual representatives against individual microbes changes considerably [216]. Unfortunately, we found no data on the role of PGLYRPs during intraamniotic infection. Due to the considerably larger amount of peptidoglycan in Gram-positive bacteria, we speculate whether the PGLYRP-1 level would be higher in AF samples invaded by Gram-positive species. Our hypothesis is supported by antimicrobial PGLYPRs effect against Gram-negative microbes, as the authors observed a bacteriostatic effects of these proteins, but no bacterial killing. The response to Gram-negative species is thus different [216].

While the above discussed changes in protein levels seem to be unique for sPTL, we did expect a certain amount of dysregulations to be shared between both studied cohorts. Indeed, we observed altered levels of proteins discussed in the PPROM cohort

to be changed also in the sPTL patients. These commonly dysregulated proteins include two histone proteins, neutrophil defensins, protein S100-A11, resistin etc. and are summarized in Table 9, External tables.

4.4. Specific aim 4 - Verification of differential abundance of selected proteins using both antibody based as well as proteomic targeted techniques.

In order to verify findings from the discovery phase, we chose several promising candidate proteins suitable for verification. These targeted experiments were carried out both using classical ELISA-based assays, as well as using targeted SRM proteomic technique.

4.4.1. Antibody-based verification experiments

Our discovery phase data showed that cathelicidin levels in the positive PPROM patient group were about six times higher compared to negative patient group. To verify these changes, we evaluated cathelicidin levels in PPROM AF samples using ELISA. We assessed the two PPROM patient groups used for the discovery phase as a verification cohort (negative MIAC and HCA, n=19; positive MIAC and HCA, n=19).

Women with presence of MIAC and HCA had a higher median AF cathelicidin level compared with those with both parameters negative (the absence of MIAC and HCA: median 1.4 ng/ml, interquartile range (IQR) 0.8-2.4 vs. the presence of MIAC and HCA: median 3.6 ng/ml, IQR 2.0-102.2, p=0.0003; Figure 11).

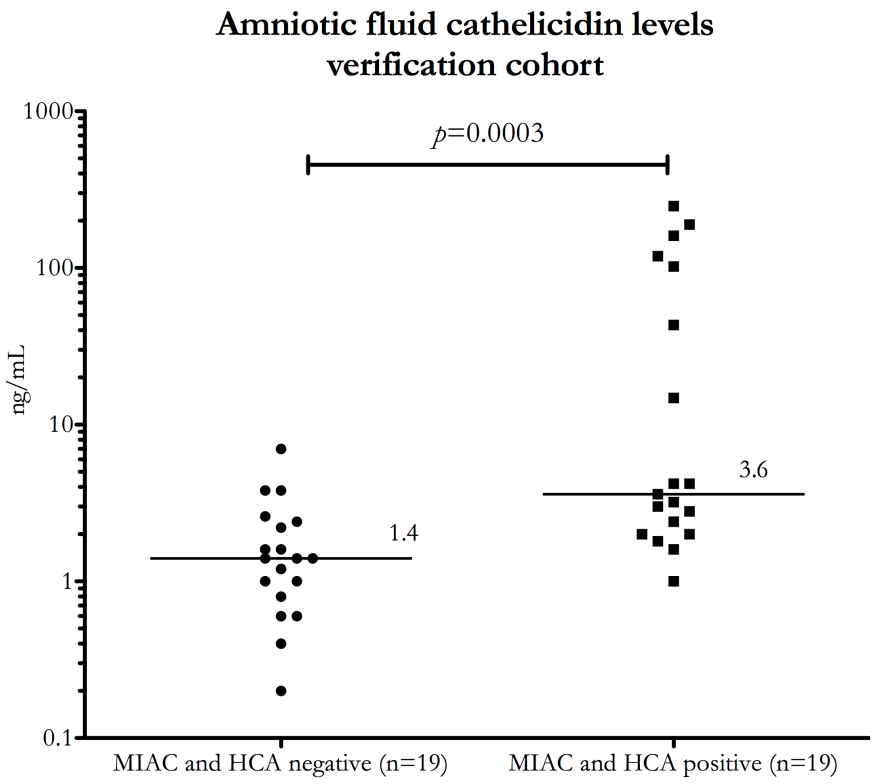


Figure 11. AF cathelicidin levels in women with PPROM according to the absence and presence of MIAC and HCA, respectively. Women with the presence of MIAC and HCA had a higher cathelicidin level than women without MIAC and HCA. Horizontal bars indicate median values.

MPO was another protein, which was selected for verification. The discovery phase experiments indicated significantly increased MPO levels in the positive PPROM patient group. MPO is produced by macrophages, neutrophils and dendritic cells and belongs to the first line of defense against invading microorganisms. MPO is well known as important intracellular tool for killing microbes during the respiratory burst. In addition, neutrophils release MPO after activation together with other granule enzymes extracellularly. Moreover, similarly to cathelicidin, MPO was shown to be involved in the formation of NETs. Therefore, increased levels of MPO in the amniotic fluid may indicate the presence of ongoing intraamniotic infection and inflammation. Currently, MPO has been suggested as a predictor for myocardial infarction due to its association with atherogenesis, plaque destabilization and thrombosis [217]. Surprisingly, although several earlier studies pointed out altered levels of this protein in MIAC positive patients, its diagnostic potential hasn't been validated yet.

To verify the observed changes in our dataset, we assessed MPO levels in PPROM AF samples using ELISA. The sample set used in this experiment was identical

with that used in cathelicidin evaluation (negative MIAC and HCA, n=19; positive MIAC and HCA, n=19).

Women with presence of MIAC and HCA had a higher median AF MPO level compared with those with both parameters negative (the absence of MIAC and HCA: median 73.4 ng/ml, IQR 50.0-163.8 vs. the presence of MIAC and HCA: median 430.4 ng/ml, IQR 85.8-2000.0, p=0.018; Figure 12).

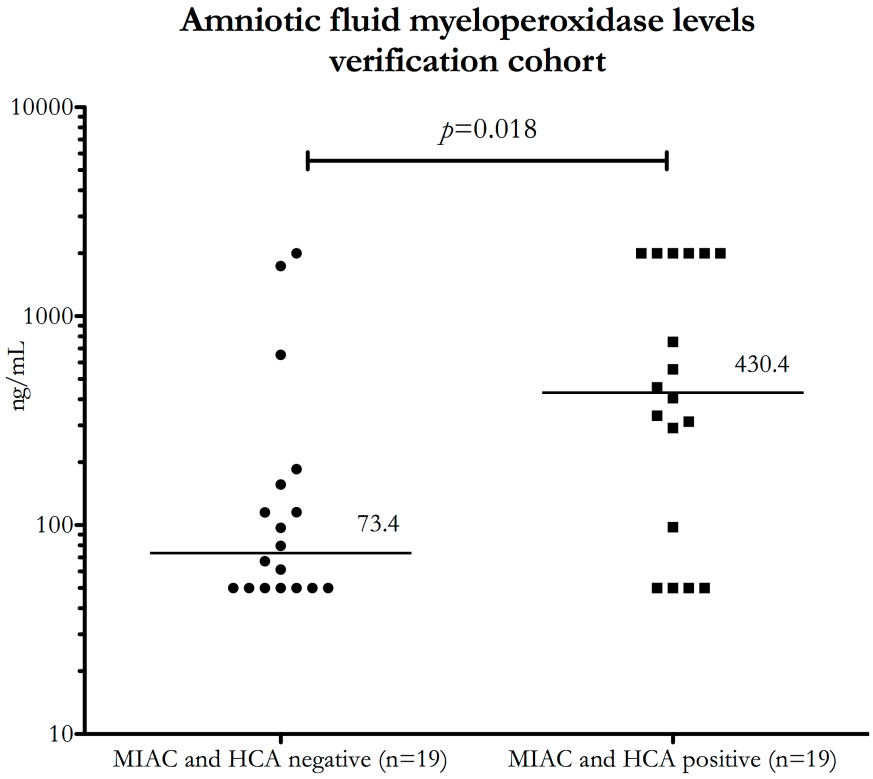


Figure 12. AF MPO levels in women with PPROM according to the absence and presence of MIAC and HCA, respectively. Women with the presence of MIAC and HCA had a higher MPO level than women without MIAC and HCA. Horizontal bars indicate median values.

4.4.2. SRM-based verification experiments

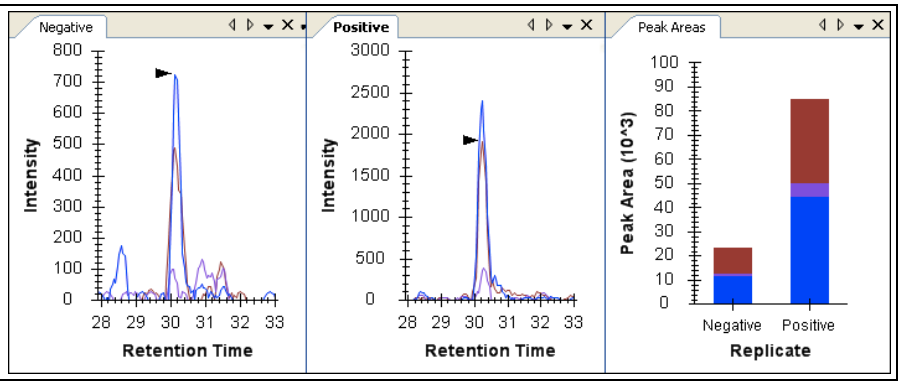
Based on our initial findings from discovery phase of the PPROM cohort, we selected 20 proteins for subsequent verification using SRM. Using the Skyline software, we selected suitable peptides and subsequently also their fragments, as described above. Thanks to the ion-trapping feature of the 4000 QTRAP mass spectrometer, we were able to acquire full MS/MS spectra of selected chromatographic peaks and thus were able to confirm the peak identity using the MIDAS workflow [101]. The knowledge of the elution time of the peptide of interest thus provided another level of certainty.

Once we were confident with the used transitions and knew in which retention time window these peptides are to be expected to elute, we assessed these 20 proteins in a pooled negative and a pooled positive AF sample - these samples were identical to those used in the initial step of the discovery phase. Our SRM data are in good agreement with the iTRAQ data (Table 7). As expected, the majority of SRM quantitation values were more profound compared to iTRAQ values. This confirms our previous findings and experience, that iTRAQ quantitation smoothens the quantitation results [218].

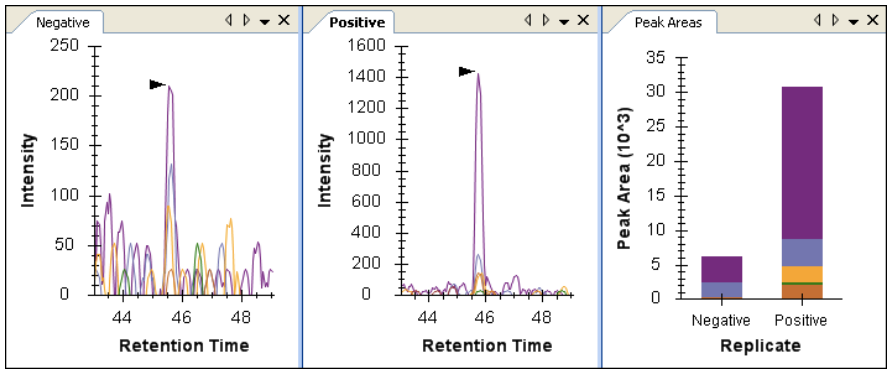
Due to the fact, that we were able to assess MPO using both SRM and ELISA, we can roughly estimate the lower limit of detection in the SRM experiments. We were able to detect MPO in both positive and negative pooled sample using SRM. As the median level of MPO in MIAC and HCA negative group of the PPROM cohort was 73.4 ng/ml, we estimate that we can reach the tens of ng/ml protein levels in AF. We also attempted to detect cathelicidin using SRM in the PPROM AF samples, but were not successful, most probably to the low levels of cathelicidin, which are in the single ng/ml range.

This demonstrates that although the SRM is a truly powerful method in terms of throughput and sensitivity, it also has its limits, which are determined mainly by the used mass spectrometer. We thus conclude that although we have the assays ready for assessing the presented proteins in individual AF sample, we will postpone these measurements and will perform them using a novel type of SRM mass spectrometer, which will be available in the upcoming months. This will increase the limit of detection for individual analytes and thus will enable more precise quantitation proteins of interest.

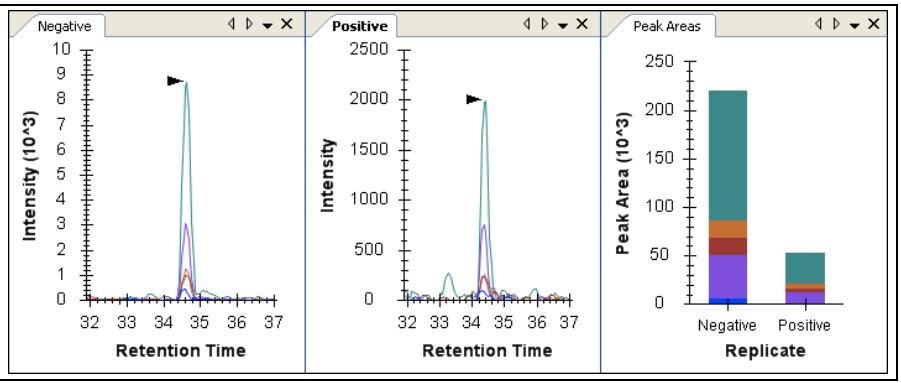
Protein:	
6-phosphogluconate dehydrogenase	
SRM Peptide:	
TIFQGIAAK	
iTRAQ quant:	3.78
SRM quant:	3.61



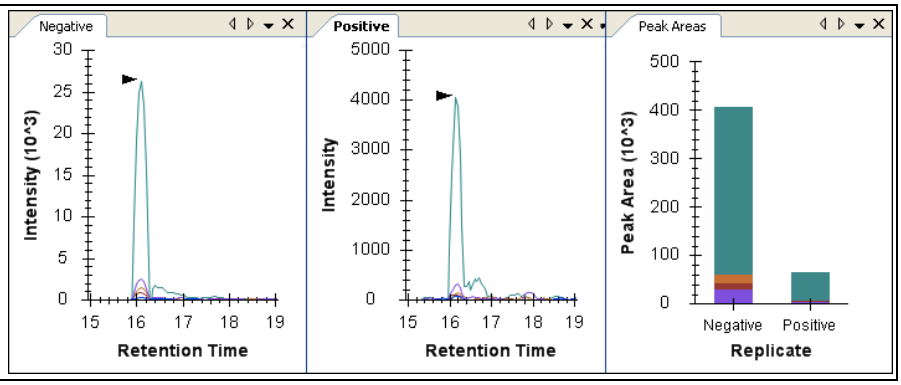
Protein:	
Chitinase-3-like protein1	
SRM Peptide:	
TLLSVGGWNFGSQR	
iTRAQ quant:	3.37
SRM quant:	4.86



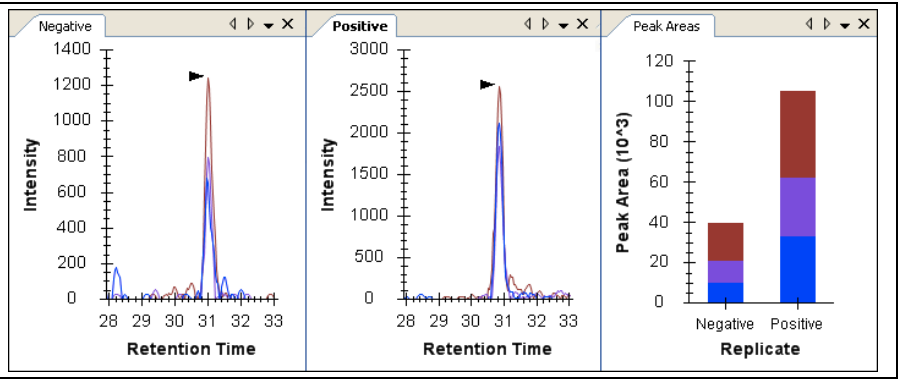
Protein:	
Decorin	
SRM Peptide:	
VSPGAFTPLVK	
iTRAQ quant:	0.38
SRM quant:	0.24



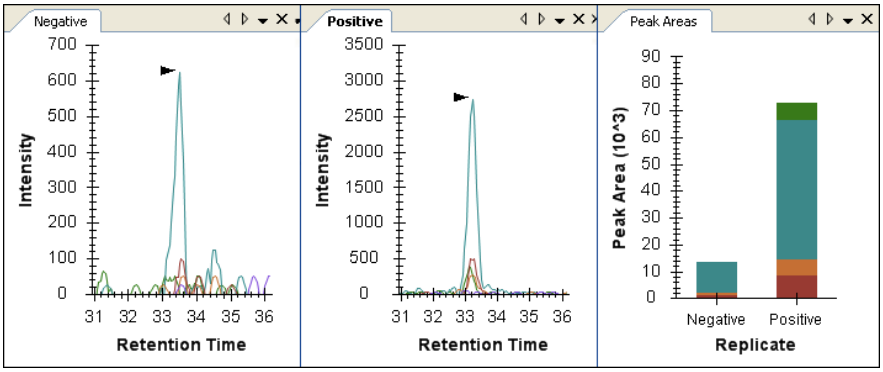
Protein:	
Dermcidin	
SRM Peptide:	
ENAGEDPGLAR	
iTRAQ quant:	0.26
SRM quant:	0.17



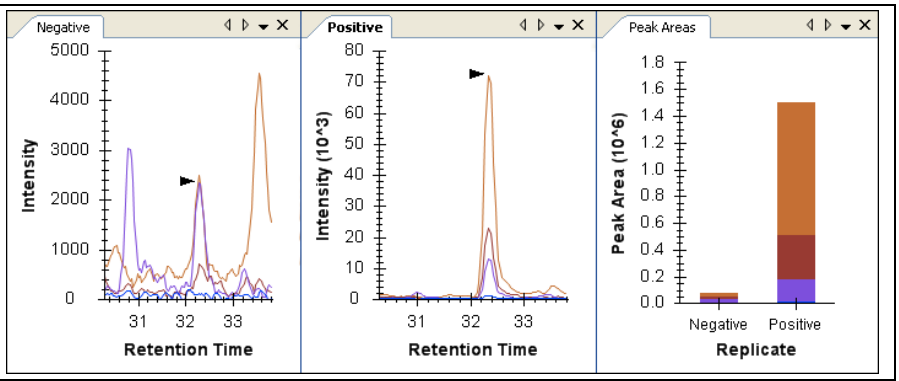
Protein:	
Fructose-bisphosphate aldolase A	
SRM Peptide:	
QLLLTADDR	
iTRAQ quant:	2.11
SRM quant:	2.63



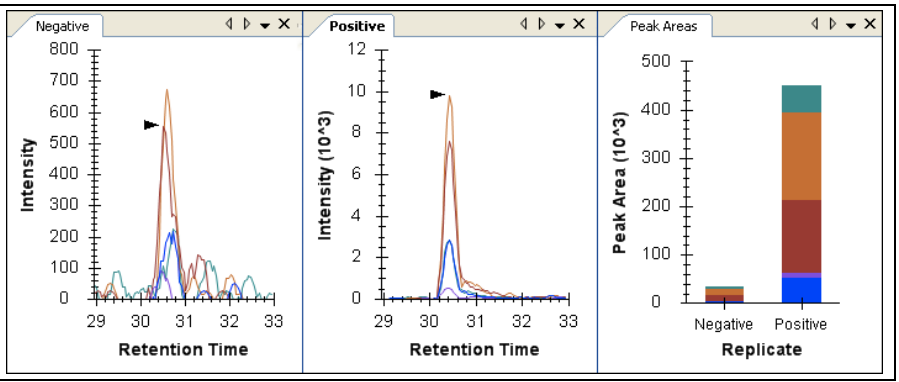
Protein:	
Glyceraldehyde-3-P dehydrogenase	
SRM Peptide:	
GALQNIIPASTGAAK	
iTRAQ quant:	2.12
SRM quant:	5.34



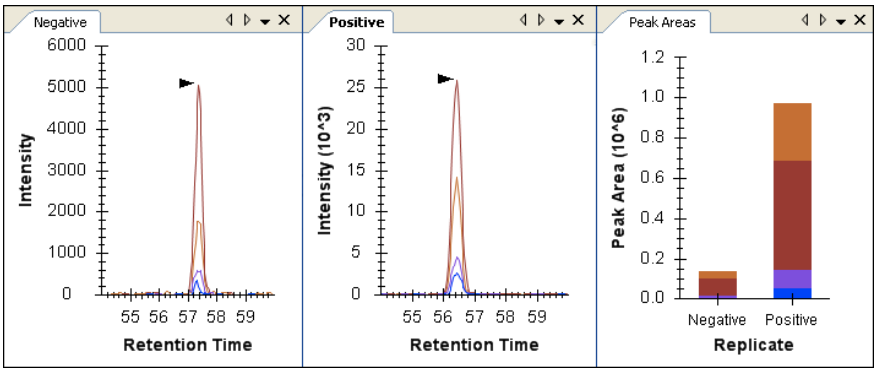
Protein:	
Histone H2A	
SRM Peptide:	
AGLQFPVGR	
iTRAQ quant:	5.34
SRM quant:	19.05



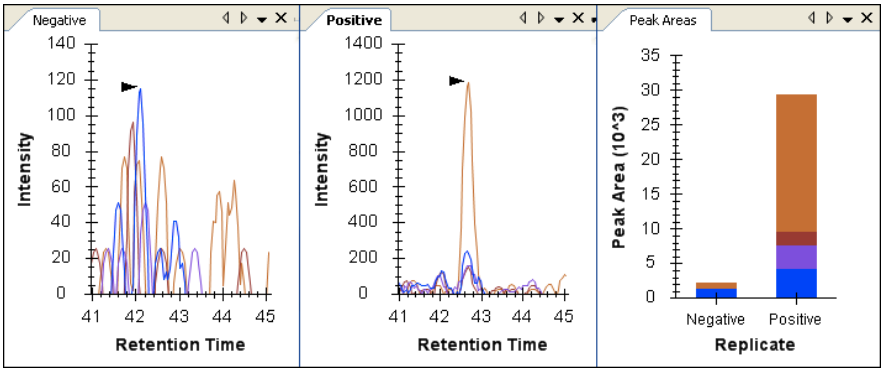
Protein:	
Histone H4	
SRM Peptide:	
ISGLIYEETR	
iTRAQ quant:	8.73
SRM quant:	12.93



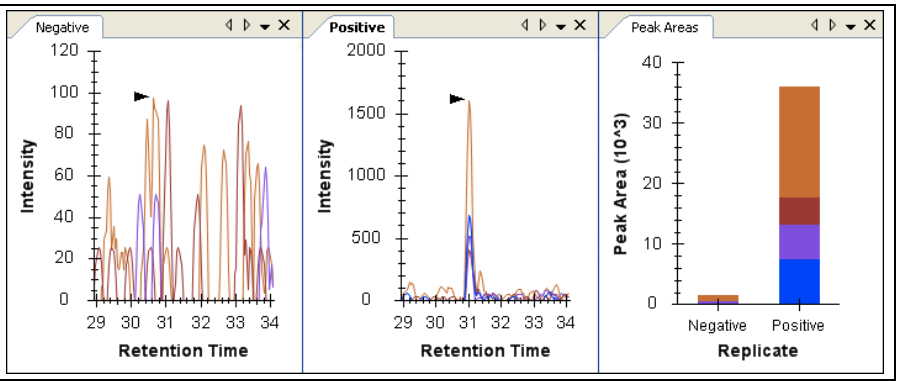
Protein:	
Lactotransferrin	
SRM Peptide:	
EDAIWNLLR	
iTRAQ quant:	2.37
SRM quant:	7.08



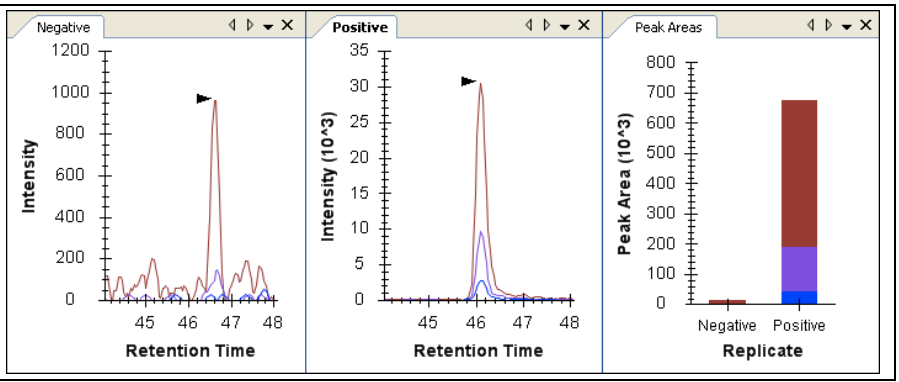
Protein:	
Leukotriene A4 hydrolase	
SRM Peptide:	
WEDAIPLALK	
iTRAQ quant:	3.90
SRM quant:	13.00



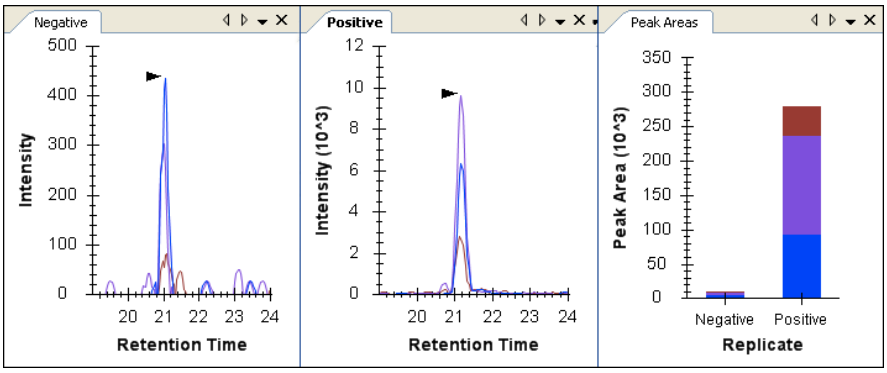
Protein:	
Matrix metalloproteinase-9	
SRM Peptide:	
FQTFEGDLK	
iTRAQ quant:	4.68
SRM quant:	23.68



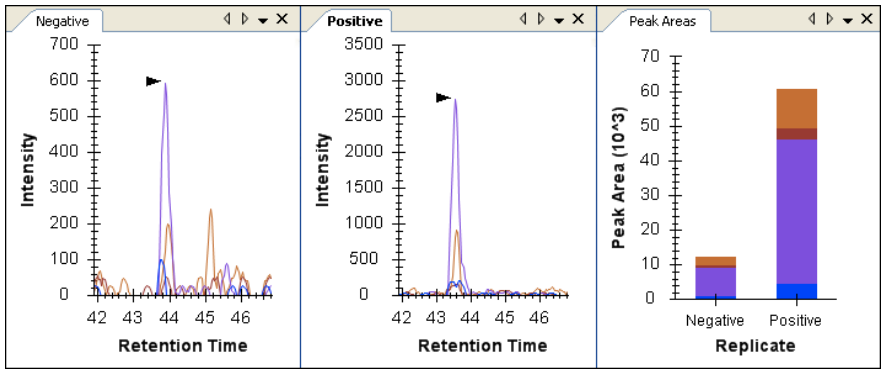
Protein:	
Myeloperoxidase	
SRM Peptide:	
VVLEGGIDPILR	
iTRAQ quant:	5.61
SRM quant:	42.56



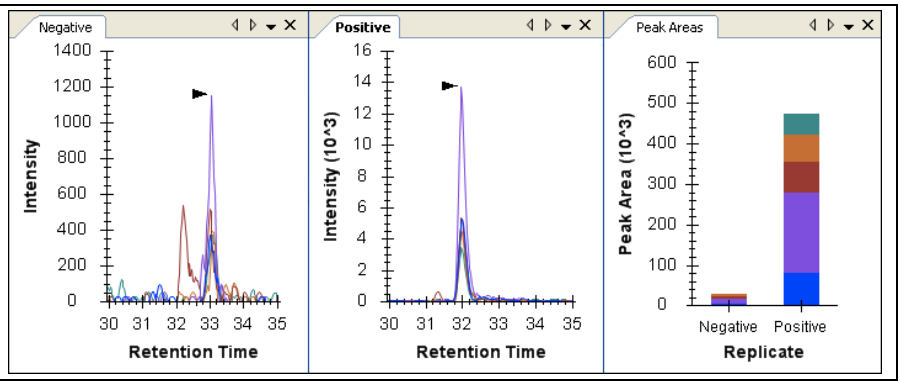
Protein:	
Neutrophil gelatinase-associated lipocalin	
SRM Peptide:	
ELTSELK	
iTRAQ quant:	4.91
SRM quant:	27.91



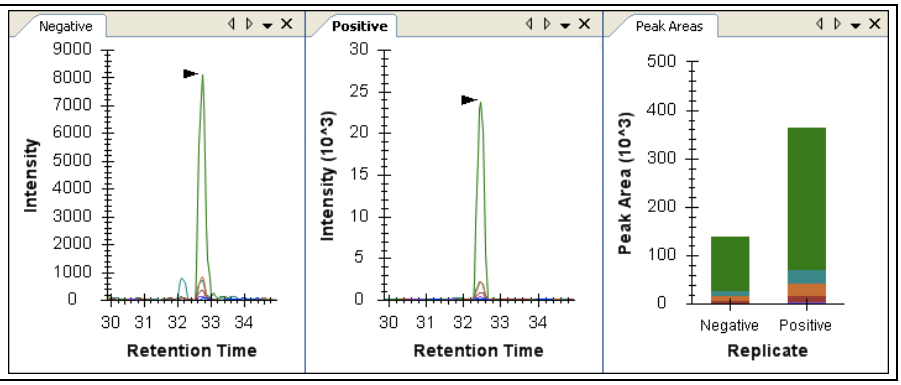
Protein:	
Phosphoglycerate kinase 1	
SRM Peptide:	
YSLEPVAVELK	
iTRAQ quant:	2.35
SRM quant:	4.86



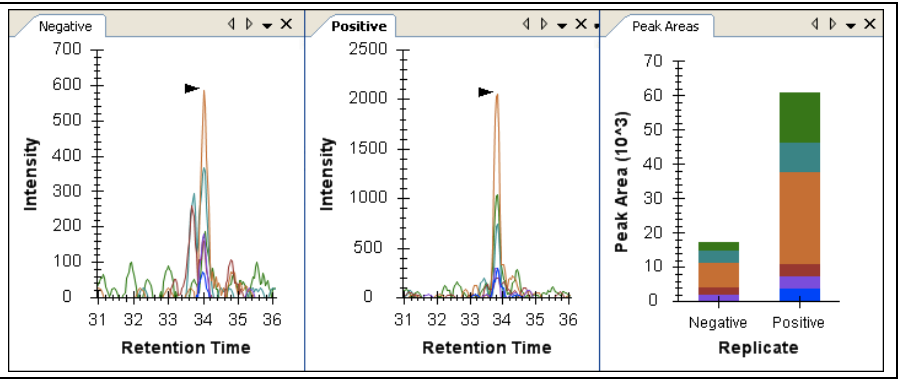
Protein:	
Plastin-2	
SRM Peptide:	
QFVTATDVVR	
iTRAQ quant:	4.28
SRM quant:	15.83



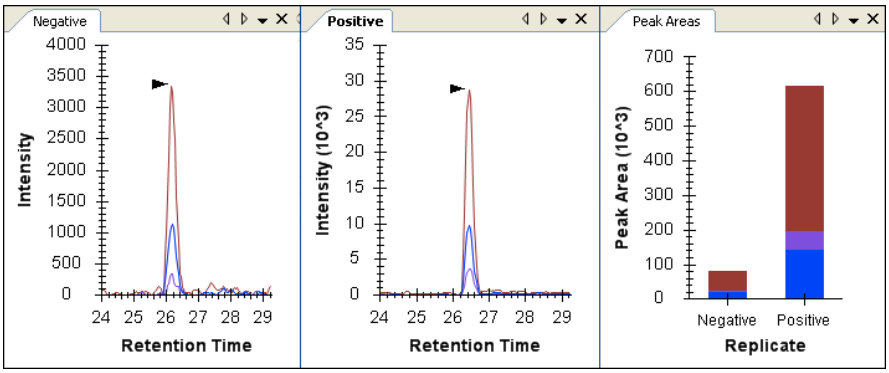
Protein:	
Profilin-1	
SRM Peptide:	
DSPSVWAAVPGK	
iTRAQ quant:	3.40
SRM quant:	2.61



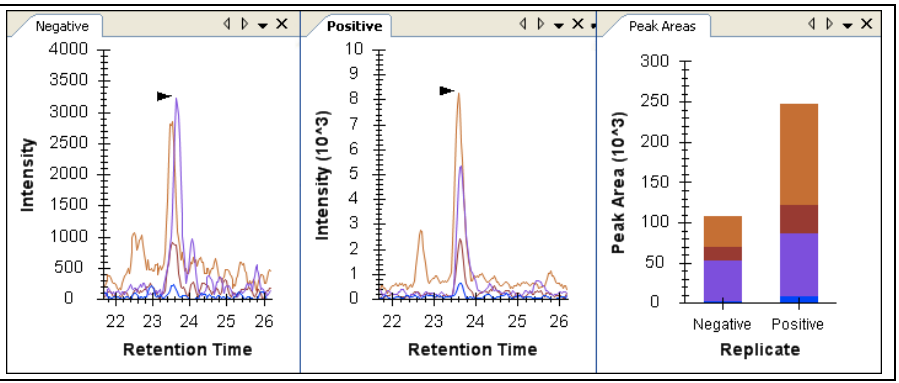
Protein:	
Protein S100-A8	
SRM Peptide:	
ALNSIIDVYHK	
iTRAQ quant:	3.03
SRM quant:	3.50



Protein:	
Protein S100-A9	
SRM Peptide:	
LGHPDTLNQGEFK	
iTRAQ quant:	3.50
SRM quant:	7.48



Protein:	
Protein S100-A11	
SRM Peptide:	
DGYNYTLSK	
iTRAQ quant:	2.75
SRM quant:	2.30



Protein:	
Vimentin	
SRM Peptide:	
ILLAELEQLK	
iTRAQ quant:	3.12
SRM quant:	4.38

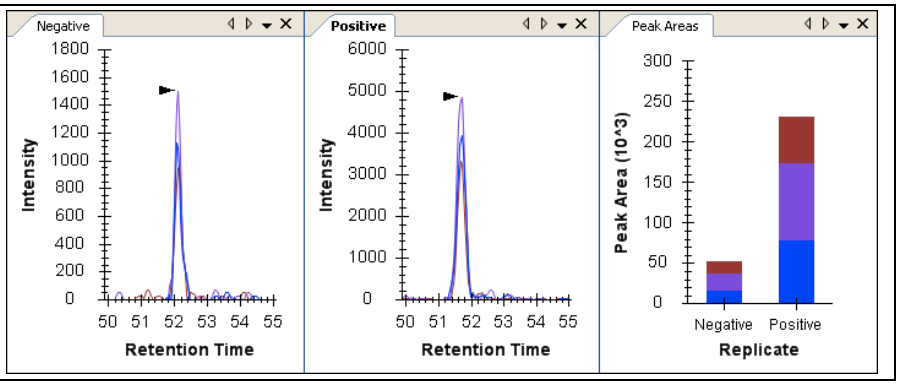


Table 7. SRM assays for 20 selected proteins shown to be dysregulated in PPROM AF samples. Except protein name, the used SRM peptide sequence is shown, along with discovery phase quantitation data (iTRAQ quant) as well as with targeted SRM quantitation data (SRM quant). The quantitation data are presented as an x-fold change in the positive vs. negative sample. LC-SRM chromatograms for the negative and positive sample are shown as well as a column representation of the chromatogram quantitation. The chromatogram intensity is shown in counts *per* second.

5. Conclusions

CysTRAQ:

- We developed and optimized a method, which enables sample complexity reduction and multiplexed quantitation across four samples
- This technique is based on cysteinyl peptide capturing in combination with iTRAQ quantitation. We thus call this approach CysTRAQ
- For the first time we demonstrated, that iTRAQ labeling is compatible with cysteinyl peptide enrichment.
- In addition, the quantitation is not influenced by the fractionation step.
- The application of CysTRAQ enables analysis and quantitation of both cysteinyl and non-cysteinyl fractions.
- If both fractions are analyzed, 60% more proteins can be identified compared to unfractionated sample.

Comparative proteomic analysis of PPROM AF samples:

- We applied CysTRAQ into multidimensional AF analysis with the goal to describe changes associated with the presence of MIAC and HCA in PPROM patients.
- Using our workflow, we identified 851 proteins, which were also quantified owing to the quantitation feature of CysTRAQ.
- A considerable portion of the proteins showed significant dysregulation.
- We speculate, that several of these changes might be caused by the presence of NETs in AF during intraamniotic infection and inflammation

Comparative proteomic analysis of sPTL AF samples

- We applied CysTRAQ into multidimensional AF analysis with the goal to describe changes associated with the presence of MIAC and HCA in sPTL patients.
- We identified and quantified 846 proteins in sPTL AF samples.
- Compared to the PPROM cohort, far less proteins showed significant dysregulation and moreover, the relative changes in protein levels were considerably lower.
- Similarly to the PPROM patients, several proteins indicate the presence of NETs also in AF of the sPTL cohort.

Verification of differential abundance of selected proteins

- Based on our findings from the PPROM discovery phase, we verified dysregulation of two proteins using ELISA.
- In both cathelicidin and MPO, we were able to confirm our discovery phase results and demonstrated considerably higher levels of both proteins in the positive MIAC and HCA PPROM patient group.
- We developed SRM assays for 20 AF proteins based on our PPROM discovery phase results.
- Using SRM, we were able to confirm dysregulation of these proteins in the PPROM cohort. Our SRM quantitation results are in good agreement with the discovery phase iTRAQ data.

6. References

1. *WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976.* Acta Obstet Gynecol Scand, 1977. **56**(3): p. 247-53.
2. Beck, S., et al., *The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity.* Bull World Health Organ, 2010. **88**(1): p. 31-8.
3. Goldenberg, R.L. and D.J. Rouse, *Prevention of premature birth.* N Engl J Med, 1998. **339**(5): p. 313-20.
4. MacDorman, M.F., et al., *Annual summary of vital statistics--2001.* Pediatrics, 2002. **110**(6): p. 1037-52.
5. Lumley, J., *The epidemiology of preterm birth.* Baillieres Clin Obstet Gynaecol, 1993. **7**(3): p. 477-98.
6. Moutquin, J.M., *Classification and heterogeneity of preterm birth.* BJOG, 2003. **110 Suppl 20**: p. 30-3.
7. Tucker, J.M., et al., *Etiologies of preterm birth in an indigent population: is prevention a logical expectation?* Obstet Gynecol, 1991. **77**(3): p. 343-7.
8. Liggins, G.C., et al., *Parturition in the sheep.* Ciba Found Symp, 1977(47): p. 5-30.
9. Anderson, A.B., K.M. Laurence, and A.C. Turnbull, *The relationship in anencephaly between the size of the adrenal cortex and the length of gestation.* J Obstet Gynaecol Br Commonw, 1969. **76**(3): p. 196-9.
10. Sfakianaki, A.K. and E.R. Norwitz, *Mechanisms of progesterone action in inhibiting prematurity.* J Matern Fetal Neonatal Med, 2006. **19**(12): p. 763-72.
11. Romero, R., et al., *The preterm parturition syndrome.* BJOG, 2006. **113 Suppl 3**: p. 17-42.
12. Goldenberg, R.L., et al., *Epidemiology and causes of preterm birth.* Lancet, 2008. **371**(9606): p. 75-84.
13. Mercer, B.M., et al., *The Preterm Prediction Study: prediction of preterm premature rupture of membranes through clinical findings and ancillary testing. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network.* Am J Obstet Gynecol, 2000. **183**(3): p. 738-45.
14. Romero, R., et al., *Intraamniotic infection and the onset of labor in preterm premature rupture of the membranes.* Am J Obstet Gynecol, 1988. **159**(3): p. 661-6.
15. Mercer, B.M., et al., *The preterm prediction study: a clinical risk assessment system.* Am J Obstet Gynecol, 1996. **174**(6): p. 1885-93; discussion 1893-5.
16. Goldenberg, R.L., et al., *Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women.* Am J Obstet Gynecol, 1996. **175**(5): p. 1317-24.
17. Menon, R., *Spontaneous preterm birth, a clinical dilemma: etiologic, pathophysiologic and genetic heterogeneities and racial disparity.* Acta Obstet Gynecol Scand, 2008. **87**(6): p. 590-600.
18. Fiscella, K., *Race, perinatal outcome, and amniotic infection.* Obstet Gynecol Surv, 1996. **51**(1): p. 60-6.
19. Menon, R., et al., *An overview of racial disparities in preterm birth rates: caused by infection or inflammatory response?* Acta Obstet Gynecol Scand, 2011.
20. Blumenshine, P., et al., *Socioeconomic disparities in adverse birth outcomes: a systematic review.* Am J Prev Med, 2010. **39**(3): p. 263-72.
21. Andres, R.L. and M.C. Day, *Perinatal complications associated with maternal tobacco use.* Semin Neonatol, 2000. **5**(3): p. 231-41.

22. Ebrahim, S.H., et al., *Trends in pregnancy-related smoking rates in the United States, 1987-1996.* JAMA, 2000. **283**(3): p. 361-6.
23. Bermudez, E.A., et al., *Relation between markers of systemic vascular inflammation and smoking in women.* Am J Cardiol, 2002. **89**(9): p. 1117-9.
24. Tracy, R.P., et al., *Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects.* Arterioscler Thromb Vasc Biol, 1997. **17**(10): p. 2167-76.
25. Menon, R., et al., *Cigarette smoke induces oxidative stress and apoptosis in normal term fetal membranes.* Placenta, 2011. **32**(4): p. 317-22.
26. Gouin, K., K. Murphy, and P.S. Shah, *Effects of cocaine use during pregnancy on low birthweight and preterm birth: systematic review and metaanalyses.* Am J Obstet Gynecol, 2011. **204**(4): p. 340 e1-12.
27. Berkowitz, G.S. and E. Papiernik, *Epidemiology of preterm birth.* Epidemiol Rev, 1993. **15**(2): p. 414-43.
28. Conde-Agudelo, A., A. Rosas-Bermudez, and A.C. Kafury-Goeta, *Birth spacing and risk of adverse perinatal outcomes: a meta-analysis.* JAMA, 2006. **295**(15): p. 1809-23.
29. Neggers, Y. and R.L. Goldenberg, *Some thoughts on body mass index, micronutrient intakes and pregnancy outcome.* J Nutr, 2003. **133**(5 Suppl 2): p. 1737S-1740S.
30. Hendler, I., et al., *The Preterm Prediction Study: association between maternal body mass index and spontaneous and indicated preterm birth.* Am J Obstet Gynecol, 2005. **192**(3): p. 882-6.
31. Copper, R.L., et al., *The preterm prediction study: maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network.* Am J Obstet Gynecol, 1996. **175**(5): p. 1286-92.
32. Wadhwa, P.D., et al., *Stress and preterm birth: neuroendocrine, immune/inflammatory, and vascular mechanisms.* Matern Child Health J, 2001. **5**(2): p. 119-25.
33. Sheldon, J., et al., *C-reactive protein and its cytokine mediators in intensive-care patients.* Clin Chem, 1993. **39**(1): p. 147-50.
34. Goldenberg, R.L., J.C. Hauth, and W.W. Andrews, *Intrauterine infection and preterm delivery.* N Engl J Med, 2000. **342**(20): p. 1500-7.
35. Minkoff, H., *Prematurity: infection as an etiologic factor.* Obstet Gynecol, 1983. **62**(2): p. 137-44.
36. Romero, R., et al., *Infection in the pathogenesis of preterm labor.* Semin Perinatol, 1988. **12**(4): p. 262-79.
37. Romero, R., et al., *The relationship between acute inflammatory lesions of the preterm placenta and amniotic fluid microbiology.* Am J Obstet Gynecol, 1992. **166**(5): p. 1382-8.
38. Goncalves, L.F., T. Chaiworapongsa, and R. Romero, *Intrauterine infection and prematurity.* Ment Retard Dev Disabil Res Rev, 2002. **8**(1): p. 3-13.
39. Watts, D.H., et al., *The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor.* Obstet Gynecol, 1992. **79**(3): p. 351-7.
40. Han, Y.W., et al., *Uncultivated bacteria as etiologic agents of intra-amniotic inflammation leading to preterm birth.* J Clin Microbiol, 2009. **47**(1): p. 38-47.
41. Relman, D.A., *The search for unrecognized pathogens.* Science, 1999. **284**(5418): p. 1308-10.
42. Gardella, C., et al., *Identification and sequencing of bacterial rDNAs in culture-negative amniotic fluid from women in premature labor.* Am J Perinatol, 2004. **21**(6): p. 319-23.
43. Yoon, B.H., et al., *Clinical implications of detection of Ureaplasma urealyticum in the amniotic cavity with the polymerase chain reaction.* Am J Obstet Gynecol, 2000. **183**(5): p. 1130-7.

44. Yoon, B.H., et al., *The clinical significance of detecting Ureaplasma urealyticum by the polymerase chain reaction in the amniotic fluid of patients with preterm labor.* Am J Obstet Gynecol, 2003. **189**(4): p. 919-24.
45. Steel, J.H., et al., *Bacteria and inflammatory cells in fetal membranes do not always cause preterm labor.* Pediatr Res, 2005. **57**(3): p. 404-11.
46. Cassell, G.H., et al., *Isolation of Mycoplasma hominis and Ureaplasma urealyticum from amniotic fluid at 16-20 weeks of gestation: potential effect on outcome of pregnancy.* Sex Transm Dis, 1983. **10**(4 Suppl): p. 294-302.
47. Menon, R., et al., *Diversity in cytokine response to bacteria associated with preterm birth by fetal membranes.* American journal of obstetrics and gynecology, 2009. **201**(3): p. 306 e1-6.
48. Jacobsson, B., et al., *Microbial invasion and cytokine response in amniotic fluid in a Swedish population of women with preterm prelabor rupture of membranes.* Acta obstetricia et gynecologica Scandinavica, 2003. **82**(5): p. 423-31.
49. Jacobsson, B., et al., *Microbial invasion and cytokine response in amniotic fluid in a Swedish population of women in preterm labor.* Acta obstetricia et gynecologica Scandinavica, 2003. **82**(2): p. 120-8.
50. Wenstrom, K.D., et al., *Elevated second-trimester amniotic fluid interleukin-6 levels predict preterm delivery.* Am J Obstet Gynecol, 1998. **178**(3): p. 546-50.
51. Yoon, B.H., et al., *An elevated amniotic fluid matrix metalloproteinase-8 level at the time of mid-trimester genetic amniocentesis is a risk factor for spontaneous preterm delivery.* Am J Obstet Gynecol, 2001. **185**(5): p. 1162-7.
52. Romero, R., et al., *Infection and labor: the detection of endotoxin in amniotic fluid.* Am J Obstet Gynecol, 1987. **157**(4 Pt 1): p. 815-9.
53. Galask, R.P., et al., *Bacterial attachment to the chorioamniotic membranes.* Am J Obstet Gynecol, 1984. **148**(7): p. 915-28.
54. Shai, Y., *Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides.* Biochim Biophys Acta, 1999. **1462**(1-2): p. 55-70.
55. Yang, L., et al., *Crystallization of antimicrobial pores in membranes: magainin and protegrin.* Biophys J, 2000. **79**(4): p. 2002-9.
56. Zasloff, M., *Antimicrobial peptides of multicellular organisms.* Nature, 2002. **415**(6870): p. 389-95.
57. Hargreaves, D.C. and R. Medzhitov, *Innate sensors of microbial infection.* J Clin Immunol, 2005. **25**(6): p. 503-10.
58. Janeway, C., *Immunobiology : the immune system in health and disease.* 6th ed2005, New York: Garland Science. xxiii, 823 p.
59. Wang, J.E., et al., *Peptidoglycan and lipoteichoic acid in gram-positive bacterial sepsis: receptors, signal transduction, biological effects, and synergism.* Shock, 2003. **20**(5): p. 402-14.
60. Doyle, S.L. and L.A. O'Neill, *Toll-like receptors: from the discovery of NFkappaB to new insights into transcriptional regulations in innate immunity.* Biochem Pharmacol, 2006. **72**(9): p. 1102-13.
61. Wang, H. and E. Hirsch, *Bacterially-induced preterm labor and regulation of prostaglandin-metabolizing enzyme expression in mice: the role of toll-like receptor 4.* Biol Reprod, 2003. **69**(6): p. 1957-63.
62. Kim, Y.M., et al., *Toll-like receptor-2 and -4 in the chorioamniotic membranes in spontaneous labor at term and in preterm parturition that are associated with chorioamnionitis.* Am J Obstet Gynecol, 2004. **191**(4): p. 1346-55.

63. Romero, R., et al., *Further observations on the fetal inflammatory response syndrome: a potential homeostatic role for the soluble receptors of tumor necrosis factor alpha.* Am J Obstet Gynecol, 2000. **183**(5): p. 1070-7.
64. Berry, S.M., et al., *Premature parturition is characterized by in utero activation of the fetal immune system.* Am J Obstet Gynecol, 1995. **173**(4): p. 1315-20.
65. Gomez, R., et al., *The fetal inflammatory response syndrome.* American journal of obstetrics and gynecology, 1998. **179**(1): p. 194-202.
66. Miralles, R., et al., *Relationship between antenatal inflammation and antenatal infection identified by detection of microbial genes by polymerase chain reaction.* Pediatric research, 2005. **57**(4): p. 570-7.
67. Lee, S.E., et al., *The intensity of the fetal inflammatory response in intraamniotic inflammation with and without microbial invasion of the amniotic cavity.* American journal of obstetrics and gynecology, 2007. **197**(3): p. 294 e1-6.
68. Kim, K.W., et al., *A rapid matrix metalloproteinase-8 bedside test for the detection of intraamniotic inflammation in women with preterm premature rupture of membranes.* American journal of obstetrics and gynecology, 2007. **197**(3): p. 292 e1-5.
69. Greig, P.C., J.M. Ernest, and L. Teot, *Low amniotic fluid glucose levels are a specific but not a sensitive marker for subclinical intrauterine infections in patients in preterm labor with intact membranes.* American journal of obstetrics and gynecology, 1994. **171**(2): p. 365-70; discussion 370-1.
70. Larsen, J.W., Jr., et al., *Significance of neutrophils and bacteria in the amniotic fluid of patients in labor.* Obstetrics and gynecology, 1976. **47**(2): p. 143-7.
71. Bobitt, J.R. and W.J. Ledger, *Amniotic fluid analysis. Its role in maternal neonatal infection.* Obstetrics and gynecology, 1978. **51**(1): p. 56-62.
72. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects.* Molecular & cellular proteomics : MCP, 2002. **1**(11): p. 845-67.
73. Picotti, P., R. Aebersold, and B. Domon, *The implications of proteolytic background for shotgun proteomics.* Molecular & cellular proteomics : MCP, 2007. **6**(9): p. 1589-98.
74. Buhimschi, I.A., R. Christner, and C.S. Buhimschi, *Proteomic biomarker analysis of amniotic fluid for identification of intra-amniotic inflammation.* BJOG : an international journal of obstetrics and gynaecology, 2005. **112**(2): p. 173-81.
75. Cho, C.K., et al., *Verification of a biomarker discovery approach for detection of Down syndrome in amniotic fluid via multiplex selected reaction monitoring (SRM) assay.* Journal of proteomics, 2011. **74**(10): p. 2052-9.
76. Cho, C.K., C.R. Smith, and E.P. Diamandis, *Amniotic fluid proteome analysis from Down syndrome pregnancies for biomarker discovery.* Journal of proteome research, 2010. **9**(7): p. 3574-82.
77. Picotti, P., R. Aebersold, and B. Domon, *The implications of proteolytic background for shotgun proteomics.* Mol Cell Proteomics, 2007. **6**(9): p. 1589-98.
78. Liu, T., et al., *The use of a quantitative cysteinyl-peptide enrichment technology for high-throughput quantitative proteomics.* Methods Mol Biol, 2007. **359**: p. 107-24.
79. Liu, T., et al., *Improved proteome coverage by using high efficiency cysteinyl peptide enrichment: the human mammary epithelial cell proteome.* Proteomics, 2005. **5**(5): p. 1263-73.
80. Liu, T., et al., *High dynamic range characterization of the trauma patient plasma proteome.* Mol Cell Proteomics, 2006. **5**(10): p. 1899-913.
81. Cho, C.K., et al., *Proteomics analysis of human amniotic fluid.* Molecular & cellular proteomics : MCP, 2007. **6**(8): p. 1406-15.
82. Voisin, S.N., et al., *Identification of novel molecular targets for endometrial cancer using a drill-down LC-MS/MS approach with iTRAQ.* PloS one, 2011. **6**(1): p. e16352.

83. Overgaard, A.J., et al., *Quantitative iTRAQ-Based Proteomic Identification of Candidate Biomarkers for Diabetic Nephropathy in Plasma of Type 1 Diabetic Patients.* Clinical proteomics, 2010. **6**(4): p. 105-114.
84. Chen, Y.T., et al., *Discovery of novel bladder cancer biomarkers by comparative urine proteomics using iTRAQ technology.* Journal of proteome research, 2010. **9**(11): p. 5803-15.
85. Giron, P., L. Dayon, and J.C. Sanchez, *Cysteine tagging for MS-based proteomics.* Mass spectrometry reviews, 2010.
86. Lin, D., et al., *Cysteinyl peptide capture for shotgun proteomics: global assessment of chemoselective fractionation.* J Proteome Res, 2010. **9**(10): p. 5461-72.
87. Peng, J., et al., *Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.* Journal of proteome research, 2003. **2**(1): p. 43-50.
88. Slonecker, P.J., et al., *Informational orthogonality of two-dimensional chromatographic separations.* Analytical chemistry, 1996. **68**(4): p. 682-9.
89. Kuzyk, M.A., et al., *Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma.* Molecular & cellular proteomics : MCP, 2009. **8**(8): p. 1860-77.
90. Pan, S., et al., *Mass spectrometry based targeted protein quantification: methods and applications.* Journal of proteome research, 2009. **8**(2): p. 787-97.
91. Picotti, P., et al., *High-throughput generation of selected reaction-monitoring assays for proteins and proteomes.* Nature methods, 2010. **7**(1): p. 43-6.
92. Salafia, C.M., C. Weigl, and L. Silberman, *The prevalence and distribution of acute placental inflammation in uncomplicated term pregnancies.* Obstet Gynecol, 1989. **73**(3 Pt 1): p. 383-9.
93. Rosenberg, N.A., et al., *Genetic structure of human populations.* Science, 2002. **298**(5602): p. 2381-5.
94. MacDonald, P.C. and M.L. Casey, *The accumulation of prostaglandins (PG) in amniotic fluid is an aftereffect of labor and not indicative of a role for PGE2 or PGF2 alpha in the initiation of human parturition.* The Journal of clinical endocrinology and metabolism, 1993. **76**(5): p. 1332-9.
95. Romero, R., et al., *Topographic differences in amniotic fluid concentrations of prostanoids in women in spontaneous labor at term.* Prostaglandins, leukotrienes, and essential fatty acids, 1994. **50**(2): p. 97-104.
96. Gardella, C., et al., *Identification and sequencing of bacterial rDNAs in culture-negative amniotic fluid from women in premature labor.* American journal of perinatology, 2004. **21**(6): p. 319-23.
97. Hitti, J., et al., *Broad-spectrum bacterial rDNA polymerase chain reaction assay for detecting amniotic fluid infection among women in premature labor.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 1997. **24**(6): p. 1228-32.
98. Shilov, I.V., et al., *The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra.* Molecular & cellular proteomics : MCP, 2007. **6**(9): p. 1638-55.
99. Tang, W.H., I.V. Shilov, and S.L. Seymour, *Nonlinear fitting method for determining local false discovery rates from decoy database searches.* Journal of proteome research, 2008. **7**(9): p. 3661-7.
100. MacLean, B., et al., *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.* Bioinformatics, 2010. **26**(7): p. 966-8.

101. Anderson, L. and C.L. Hunter, *Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins.* Molecular & cellular proteomics : MCP, 2006. **5**(4): p. 573-88.
102. Kovacs-Simon, A., R.W. Titball, and S.L. Michell, *Lipoproteins of bacterial pathogens.* Infect Immun, 2011. **79**(2): p. 548-61.
103. Gygi, S.P., et al., *Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags.* J Proteome Res, 2002. **1**(1): p. 47-54.
104. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome research, 2003. **13**(9): p. 2129-41.
105. Vogl, T., et al., *Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock.* Nature medicine, 2007. **13**(9): p. 1042-9.
106. Sumagin, R. and I.H. Sarelius, *Intercellular adhesion molecule-1 enrichment near tricellular endothelial junctions is preferentially associated with leukocyte transmigration and signals for reorganization of these junctions to accommodate leukocyte passage.* Journal of immunology, 2010. **184**(9): p. 5242-52.
107. Baggiolini, M. and I. Clark-Lewis, *Interleukin-8, a chemotactic and inflammatory cytokine.* FEBS letters, 1992. **307**(1): p. 97-101.
108. Cole, A.M., et al., *Calcitermin, a novel antimicrobial peptide isolated from human airway secretions.* FEBS letters, 2001. **504**(1-2): p. 5-10.
109. Moroz, O.V., et al., *Both Ca2+ and Zn2+ are essential for S100A12 protein oligomerization and function.* BMC biochemistry, 2009. **10**: p. 11.
110. Szklarczyk, D., et al., *The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored.* Nucleic acids research, 2011. **39**(Database issue): p. D561-8.
111. Amenomori, M., et al., *Differential effects of human neutrophil peptide-1 on growth factor and interleukin-8 production by human lung fibroblasts and epithelial cells.* Experimental lung research, 2010. **36**(7): p. 411-9.
112. Cobo, T., et al., *Clinical and inflammatory markers in amniotic fluid as predictors of adverse outcomes in preterm premature rupture of membranes.* American journal of obstetrics and gynecology, 2011.
113. Erez, O., et al., *Differential expression pattern of genes encoding for anti-microbial peptides in the fetal membranes of patients with spontaneous preterm labor and intact membranes and those with preterm prelabor rupture of the membranes.* The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians, 2009. **22**(12): p. 1103-15.
114. Lehrer, R.I., A.K. Lichtenstein, and T. Ganz, *Defensins: antimicrobial and cytotoxic peptides of mammalian cells.* Annual review of immunology, 1993. **11**: p. 105-28.
115. Wiesner, J. and A. Vilcinskas, *Antimicrobial peptides: the ancient arm of the human immune system.* Virulence, 2010. **1**(5): p. 440-64.
116. Gravett, M.G., et al., *Diagnosis of intra-amniotic infection by proteomic profiling and identification of novel biomarkers.* JAMA : the journal of the American Medical Association, 2004. **292**(4): p. 462-9.
117. Daoud, G.A., et al., *Reduced nitric oxide in amniotic fluid of patients with chorioamnionitis.* Fetal diagnosis and therapy, 2006. **21**(1): p. 77-80.
118. Helmig, B.R., et al., *Neutrophil elastase and secretory leukocyte protease inhibitor in prelabor rupture of membranes, parturition and intra-amniotic infection.* The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians, 2002. **12**(4): p. 237-46.

119. Kidokoro, K., et al., *Amniotic fluid neutrophil elastase and lactate dehydrogenase: association with histologic chorioamnionitis.* Acta obstetricia et gynecologica Scandinavica, 2006. **85**(6): p. 669-74.
120. Malamitsi-Puchner, A., et al., *Investigation of midtrimester amniotic fluid factors as potential predictors of term and preterm deliveries.* Mediators of inflammation, 2006. **2006**(4): p. 94381.
121. Romero, R., et al., *Isobaric labeling and tandem mass spectrometry: a novel approach for profiling and quantifying proteins differentially expressed in amniotic fluid in preterm labor with and without intra-amniotic infection/inflammation.* The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians, 2010. **23**(4): p. 261-80.
122. O'Brien, W.F., et al., *Amniotic fluid alpha 1-antitrypsin concentration in premature rupture of the membranes.* American journal of obstetrics and gynecology, 1990. **162**(3): p. 756-9.
123. Maymon, E., et al., *A role for the 72 kDa gelatinase (MMP-2) and its inhibitor (TIMP-2) in human parturition, premature rupture of membranes and intraamniotic infection.* Journal of perinatal medicine, 2001. **29**(4): p. 308-16.
124. Di Ferdinando, A., et al., *Expression of matrix metalloproteinase-9 (MMP-9) in human midpregnancy amniotic fluid and risk of preterm labor.* Clinical and experimental obstetrics & gynecology, 2010. **37**(3): p. 193-6.
125. Zaga-Clavellina, V., et al., *In vitro secretion and activity profiles of matrix metalloproteinases, MMP-9 and MMP-2, in human term extra-placental membranes after exposure to Escherichia coli.* Reproductive biology and endocrinology : RB&E, 2011. **9**: p. 13.
126. Locksmith, G.J., et al., *Amniotic fluid concentrations of matrix metalloproteinase 9 and tissue inhibitor of metalloproteinase 1 during pregnancy and labor.* American journal of obstetrics and gynecology, 2001. **184**(2): p. 159-64.
127. Remold-O'Donnell, E., *A fast-acting elastase inhibitor in human monocytes.* The Journal of experimental medicine, 1985. **162**(6): p. 2142-55.
128. Cooley, J., et al., *The serpin MNEI inhibits elastase-like and chymotrypsin-like serine proteases through efficient reactions at two active sites.* Biochemistry, 2001. **40**(51): p. 15762-70.
129. Morgan, J.G., et al., *Human neutrophil granule cationic protein CAP37 is a specific macrophage chemotaxin that shares homology with inflammatory proteinases.* Advances in experimental medicine and biology, 1991. **305**: p. 89-96.
130. *Entrez Gene: AZU1 azurocidin 1 (cationic antimicrobial protein 37).* 2011; Available from: http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=566.
131. Miyasaki, K.T. and A.L. Bodeau, *Human neutrophil azurocidin synergizes with leukocyte elastase and cathepsin G in the killing of Capnocytophaga sputigena.* Infection and immunity, 1992. **60**(11): p. 4973-5.
132. Ennis, B.W. and L.M. Matrisian, *Matrix degrading metalloproteinases.* Journal of neuro-oncology, 1994. **18**(2): p. 105-9.
133. Van Lint, P. and C. Libert, *Chemokine and cytokine processing by matrix metalloproteinases and its effect on leukocyte migration and inflammation.* Journal of leukocyte biology, 2007. **82**(6): p. 1375-81.
134. Bryant-Greenwood, G.D. and S.Y. Yamamoto, *Control of peripartal collagenolysis in the human chorion-decidua.* American journal of obstetrics and gynecology, 1995. **172**(1 Pt 1): p. 63-70.

135. Haeggstrom, J.Z., et al., *Leukotriene A4 hydrolase: a zinc metalloenzyme.* Biochemical and biophysical research communications, 1990. **172**(3): p. 965-70.
136. Minami, M., et al., *Leukotriene A4 hydrolase is a zinc-containing aminopeptidase.* Biochemical and biophysical research communications, 1990. **173**(2): p. 620-6.
137. Blomster, M., et al., *Evidence for a catalytic role of tyrosine 383 in the peptidase reaction of leukotriene A4 hydrolase.* European journal of biochemistry / FEBS, 1995. **231**(3): p. 528-34.
138. Tager, A.M. and A.D. Luster, *BLT1 and BLT2: the leukotriene B(4) receptors.* Prostaglandins, leukotrienes, and essential fatty acids, 2003. **69**(2-3): p. 123-34.
139. Lynch, K.R., et al., *Characterization of the human cysteinyl leukotriene CysLT1 receptor.* Nature, 1999. **399**(6738): p. 789-93.
140. Pawlowski, K., et al., *Novel conserved hydrolase domain in the CLCA family of alleged calcium-activated chloride channels.* Proteins, 2006. **63**(3): p. 424-39.
141. Gibson, A., et al., *hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels.* The Journal of biological chemistry, 2005. **280**(29): p. 27205-12.
142. Furuya, M., et al., *Lymphatic endothelial murine chloride channel calcium-activated 1 is a ligand for leukocyte LFA-1 and Mac-1.* Journal of immunology, 2010. **185**(10): p. 5769-77.
143. Toda, M., et al., *A calcium-activated chloride channel (HCLCA1) is strongly related to IL-9 expression and mucus production in bronchial epithelium of patients with asthma.* The Journal of allergy and clinical immunology, 2002. **109**(2): p. 246-50.
144. Hoshino, M., et al., *Increased expression of the human Ca2+-activated Cl- channel 1 (CaCC1) gene in the asthmatic airway.* American journal of respiratory and critical care medicine, 2002. **165**(8): p. 1132-6.
145. Gruber, A.D., et al., *Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of Ca2+-activated Cl- channel proteins.* Genomics, 1998. **54**(2): p. 200-14.
146. Agnel, M., T. Vermat, and J.M. Culouscou, *Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea.* FEBS letters, 1999. **455**(3): p. 295-301.
147. Silswal, N., et al., *Human resistin stimulates the pro-inflammatory cytokines TNF-alpha and IL-12 in macrophages by NF-kappaB-dependent pathway.* Biochemical and biophysical research communications, 2005. **334**(4): p. 1092-101.
148. Verma, S., et al., *Resistin promotes endothelial cell activation: further evidence of adipokine-endothelial interaction.* Circulation, 2003. **108**(6): p. 736-40.
149. Lu, S.C., et al., *Lipopolysaccharide increases resistin gene expression in vivo and in vitro.* FEBS letters, 2002. **530**(1-3): p. 158-62.
150. Tarkowski, A., et al., *Resistin competes with lipopolysaccharide for binding to toll-like receptor 4.* Journal of cellular and molecular medicine, 2010. **14**(6B): p. 1419-31.
151. Nien, J.K., et al., *Resistin: a hormone which induces insulin resistance is increased in normal pregnancy.* Journal of perinatal medicine, 2007. **35**(6): p. 513-21.
152. Kusanovic, J.P., et al., *Resistin in amniotic fluid and its association with intra-amniotic infection and inflammation.* The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians, 2008. **21**(12): p. 902-16.
153. Brogden, K.A., et al., *The nervous system and innate immunity: the neuropeptide connection.* Nature immunology, 2005. **6**(6): p. 558-64.

154. Zanetti, M., R. Gennaro, and D. Romeo, *Cathelicidins: a novel protein family with a common proregion and a variable C-terminal antimicrobial domain.* FEBS letters, 1995. **374**(1): p. 1-5.
155. Zhu, S., *Did cathelicidins, a family of multifunctional host-defense peptides, arise from a cysteine protease inhibitor?* Trends in microbiology, 2008. **16**(8): p. 353-60.
156. Frohm Nilsson, M., et al., *The human cationic antimicrobial protein (hCAP18), a peptide antibiotic, is widely expressed in human squamous epithelia and colocalizes with interleukin-6.* Infection and immunity, 1999. **67**(5): p. 2561-6.
157. Putsep, K., et al., *Deficiency of antibacterial peptides in patients with morbus Kostmann: an observation study.* Lancet, 2002. **360**(9340): p. 1144-9.
158. Yang, Y.H., et al., *Expression of LL-37/hCAP-18 gene in human leukemia cells.* Leukemia research, 2003. **27**(10): p. 947-50.
159. Chromek, M., et al., *The antimicrobial peptide cathelicidin protects the urinary tract against invasive bacterial infection.* Nature medicine, 2006. **12**(6): p. 636-41.
160. Nizet, V., et al., *Innate antimicrobial peptide protects the skin from invasive bacterial infection.* Nature, 2001. **414**(6862): p. 454-7.
161. Yoshio, H., et al., *Antimicrobial polypeptides of human vernix caseosa and amniotic fluid: implications for newborn innate defense.* Pediatric research, 2003. **53**(2): p. 211-6.
162. Martineau, A.R., et al., *IFN-gamma- and TNF-independent vitamin D-inducible human suppression of mycobacteria: the role of cathelicidin LL-37.* Journal of immunology, 2007. **178**(11): p. 7190-8.
163. Rook, G.A., et al., *Vitamin D3, gamma interferon, and control of proliferation of Mycobacterium tuberculosis by human monocytes.* Immunology, 1986. **57**(1): p. 159-63.
164. Liu, P.T., et al., *Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response.* Science, 2006. **311**(5768): p. 1770-3.
165. Adorini, L. and G. Penna, *Control of autoimmune diseases by the vitamin D endocrine system.* Nature clinical practice. Rheumatology, 2008. **4**(8): p. 404-12.
166. Amatngalim, G.D., et al., *Cathelicidin peptide LL-37 modulates TREM-1 expression and inflammatory responses to microbial compounds.* Inflammation, 2011. **34**(5): p. 412-25.
167. Stene, L.C. and G. Joner, *Use of cod liver oil during the first year of life is associated with lower risk of childhood-onset type 1 diabetes: a large, population-based, case-control study.* The American journal of clinical nutrition, 2003. **78**(6): p. 1128-34.
168. Camargo, C.A., Jr., et al., *Maternal intake of vitamin D during pregnancy and risk of recurrent wheeze in children at 3 y of age.* The American journal of clinical nutrition, 2007. **85**(3): p. 788-95.
169. McGrath, J., *Hypothesis: is low prenatal vitamin D a risk-modifying factor for schizophrenia?* Schizophrenia research, 1999. **40**(3): p. 173-7.
170. McGrath, J., et al., *Vitamin D supplementation during the first year of life and risk of schizophrenia: a Finnish birth cohort study.* Schizophrenia research, 2004. **67**(2-3): p. 237-45.
171. Shand, A.W., et al., *Maternal vitamin D status in pregnancy and adverse pregnancy outcomes in a group at high risk for pre-eclampsia.* BJOG : an international journal of obstetrics and gynaecology, 2010. **117**(13): p. 1593-8.
172. Arora, C.P. and C.J. Hobel, *Vitamin D - a novel role in pregnancy.* Biopolymers and Cell, 2010. **26**(2): p. 8.
173. *Differences in infant mortality between blacks and whites--United States, 1980-1991.* MMWR. Morbidity and mortality weekly report, 1994. **43**(16): p. 288-9.
174. Clemens, T.L., et al., *Increased skin pigment reduces the capacity of skin to synthesise vitamin D3.* Lancet, 1982. **1**(8263): p. 74-6.
175. Park, J.H., et al., *An antimicrobial protein, lactoferrin exists in the sweat: proteomic analysis of sweat.* Experimental dermatology, 2011. **20**(4): p. 369-71.

176. Bloemen, K., et al., *A new approach to study exhaled proteins as potential biomarkers for asthma.* Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology, 2011. **41**(3): p. 346-56.
177. Kurova, V.S., et al., *Proteomics of exhaled breath: methodological nuances and pitfalls.* Clinical chemistry and laboratory medicine : CCLM / FESCC, 2009. **47**(6): p. 706-12.
178. Rieg, S., et al., *Generation of multiple stable dermcidin-derived antimicrobial peptides in sweat of different body sites.* The Journal of investigative dermatology, 2006. **126**(2): p. 354-65.
179. Steffen, H., et al., *Naturally processed dermcidin-derived peptides do not permeabilize bacterial membranes and kill microorganisms irrespective of their charge.* Antimicrobial agents and chemotherapy, 2006. **50**(8): p. 2608-20.
180. Senyurek, I., et al., *Dermcidin-derived peptides show a different mode of action than the cathelicidin LL-37 against Staphylococcus aureus.* Antimicrobial agents and chemotherapy, 2009. **53**(6): p. 2499-509.
181. Baechle, D., et al., *Cathepsin D is present in human eccrine sweat and involved in the postsecretory processing of the antimicrobial peptide DCD-1L.* The Journal of biological chemistry, 2006. **281**(9): p. 5406-15.
182. Wolffe, A.P. and D. Guschin, *Review: chromatin structural features and targets that regulate transcription.* Journal of structural biology, 2000. **129**(2-3): p. 102-22.
183. Khorasanizadeh, S., *The nucleosome: from genomic organization to genomic regulation.* Cell, 2004. **116**(2): p. 259-72.
184. Hiemstra, P.S., et al., *Antimicrobial proteins of murine macrophages.* Infection and immunity, 1993. **61**(7): p. 3038-46.
185. Fernandes, J.M., et al., *Anti-microbial properties of histone H2A from skin secretions of rainbow trout, Oncorhynchus mykiss.* The Biochemical journal, 2002. **368**(Pt 2): p. 611-20.
186. Kawasaki, H., et al., *A protein with antimicrobial activity in the skin of Schlegel's green tree frog Rhacophorus schlegelii (Rhacophoridae) identified as histone H2B.* Biochemical and biophysical research communications, 2003. **312**(4): p. 1082-6.
187. Jacobsen, F., et al., *Activity of histone H1.2 in infected burn wounds.* The Journal of antimicrobial chemotherapy, 2005. **55**(5): p. 735-41.
188. Chinchar, V.G., et al., *Inactivation of viruses infecting ectothermic animals by amphibian and piscine antimicrobial peptides.* Virology, 2004. **323**(2): p. 268-75.
189. Tamura, M., et al., *Inhibition of attachment of virions of Norwalk virus to mammalian cells by soluble histone molecules.* Archives of virology, 2003. **148**(9): p. 1659-70.
190. Kim, H.S., et al., *Endotoxin-neutralizing antimicrobial proteins of the human placenta.* Journal of immunology, 2002. **168**(5): p. 2356-64.
191. Noga, E.J., Z. Fan, and U. Silphaduang, *Histone-like proteins from fish are lethal to the parasitic dinoflagellate Amyloodinium ocellatum.* Parasitology, 2001. **123**(Pt 1): p. 57-65.
192. Kawasaki, H. and S. Iwamuro, *Potential roles of histones in host defense as antimicrobial agents.* Infectious disorders drug targets, 2008. **8**(3): p. 195-205.
193. Augusto, L.A., et al., *Histones: a novel class of lipopolysaccharide-binding molecules.* Biochemistry, 2003. **42**(13): p. 3929-38.
194. Lee, D.Y., et al., *Histone H4 is a major component of the antimicrobial action of human sebocytes.* The Journal of investigative dermatology, 2009. **129**(10): p. 2489-96.
195. Fadeel, B., et al., *Involvement of caspases in neutrophil apoptosis: regulation by reactive oxygen species.* Blood, 1998. **92**(12): p. 4808-18.
196. Fuchs, T.A., et al., *Novel cell death program leads to neutrophil extracellular traps.* The Journal of cell biology, 2007. **176**(2): p. 231-41.

197. Brinkmann, V., et al., *Neutrophil extracellular traps kill bacteria.* Science, 2004. **303**(5663): p. 1532-5.
198. Papayannopoulos, V., et al., *Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps.* The Journal of cell biology, 2010. **191**(3): p. 677-91.
199. Jaillon, S., et al., *The humoral pattern recognition receptor PTX3 is stored in neutrophil granules and localizes in extracellular traps.* The Journal of experimental medicine, 2007. **204**(4): p. 793-804.
200. Urban, C.F., et al., *Neutrophil extracellular traps capture and kill Candida albicans yeast and hyphal forms.* Cellular microbiology, 2006. **8**(4): p. 668-76.
201. Lauth, X., et al., *M1 protein allows Group A streptococcal survival in phagocyte extracellular traps through cathelicidin inhibition.* Journal of innate immunity, 2009. **1**(3): p. 202-14.
202. Cho, J.H., et al., *Human peptidoglycan recognition protein S is an effector of neutrophil-mediated innate immunity.* Blood, 2005. **106**(7): p. 2551-8.
203. Dworski, R., et al., *Eosinophil and neutrophil extracellular DNA traps in human allergic asthmatic airways.* The Journal of allergy and clinical immunology, 2011. **127**(5): p. 1260-6.
204. Yamada, M., et al., *Interferon-gamma production by neutrophils during bacterial pneumonia in mice.* American journal of respiratory and critical care medicine, 2011. **183**(10): p. 1391-401.
205. Gupta, A., et al., *Occurrence of neutrophil extracellular DNA traps (NETs) in pre-eclampsia: a link with elevated levels of cell-free DNA?* Annals of the New York Academy of Sciences, 2006. **1075**: p. 118-22.
206. Gupta, A.K., et al., *Neutrophil NETs: a novel contributor to preeclampsia-associated placental hypoxia?* Seminars in immunopathology, 2007. **29**(2): p. 163-7.
207. Deiteren, K., et al., *Carboxypeptidase M: Multiple alliances and unknown partners.* Clinica chimica acta; international journal of clinical chemistry, 2009. **399**(1-2): p. 24-39.
208. Marquez-Curtis, L., et al., *Carboxypeptidase M expressed by human bone marrow cells cleaves the C-terminal lysine of stromal cell-derived factor-1alpha: another player in hematopoietic stem/progenitor cell mobilization?* Stem cells, 2008. **26**(5): p. 1211-20.
209. Kumar, B.V., et al., *Structure of the human pancreatic cholesterol esterase gene.* Biochemistry, 1992. **31**(26): p. 6077-81.
210. Fernandez, D., et al., *Structural and Functional Analysis of the Complex between Citrate and the Zinc Peptidase Carboxypeptidase A.* Enzyme research, 2011. **2011**: p. 128676.
211. Kemik, O., et al., *Serum procarboxypeptidase a and carboxypeptidase a levels in pancreatic disease.* Human & experimental toxicology, 2011.
212. Yoshida, H., K. Kinoshita, and M. Ashida, *Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, Bombyx mori.* The Journal of biological chemistry, 1996. **271**(23): p. 13854-60.
213. Kang, D., et al., *A peptidoglycan recognition protein in innate immunity conserved from insects to humans.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(17): p. 10078-82.
214. Liu, C., et al., *Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth.* The Journal of biological chemistry, 2000. **275**(32): p. 24490-9.
215. Gelius, E., et al., *A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity.* Biochemical and biophysical research communications, 2003. **306**(4): p. 988-94.
216. Lu, X., et al., *Peptidoglycan recognition proteins are a new class of human bactericidal proteins.* The Journal of biological chemistry, 2006. **281**(9): p. 5895-907.

217. Dominguez-Rodriguez, A. and P. Abreu-Gonzalez, *Current role of myeloperoxidase in routine clinical practice.* Expert review of cardiovascular therapy, 2011. **9**(2): p. 223-30.
218. Dieppedale, J., et al., *Identification of a putative chaperone involved in stress resistance and virulence in Francisella tularensis.* Infection and immunity, 2011. **79**(4): p. 1428-39.

Outputs

Publications related to the thesis

Tambor V, Hunter CL, Seymour SL, Kacerovsky M, Stulik J and Lenco J

CysTRAQ - A combination of iTRAQ and Enrichment of Cysteinyl Peptides for Uncovering and Quantifying Hidden Proteomes

Journal of Proteomics, 2011

Kacerovsky M, Andrys C, Hornychova H, Pliskova L, Lancz K, Musilova I, Drahosova M, Bolehovska R, Tambor V, Jacobsson B

Amniotic fluid soluble Toll-like receptor 4 in pregnancies complicated by preterm prelabor rupture of the membranes

Journal of Maternal-Fetal and Neonatal Medicine, 2011

Kacerovsky M, Drahosova M, Andrys C, Hornychova H, Tambor V, Lenco J, Tosner J, Krejsek J.

Amniotic fluid concentrations of soluble scavenger receptor for hemoglobin (sCD163) in pregnancy complicated by preterm premature rupture of the membranes and histologic chorioamnionitis.

Journal of Maternal-Fetal and Neonatal Medicine, 2011;24(8):995-1001.

Kacerovsky M, Pliskova L, Bolehovska R, Musilova I, Hornychova H, Tambor V, Jacobsson B.

The microbial load with genital mycoplasmas correlates with the degree of histologic chorioamnionitis in preterm PROM.

American Journal of Obstetrics and Gynecology, 2011

Menon R, Merialdi M, Torloni R, Massinen T, Voltolini Ch, Kacerovsky M, Tambor V. et al.

Biomarkers of Spontaneous Preterm Birth: An Overview of The Literature in the Last Four Decades.

Reproductive Sciences, 2011

Kacerovsky M, Drahosova M, Hornychova H, Tambor V, Flidrova E, Musilova I, Tosner J, Andrys C.

Umbilical cord blood concentration of soluble scavenger receptor for hemoglobin, but not pentraxin 3, is of value for early postpartum identification of the presence of histological chorioamnionitis.

Journal of Maternal-Fetal and Neonatal Medicine, 2011

Musilova I, Kacerovsky M, Tambor V, Tosner J.

Proteomics and biomarkers for detection of preterm labor: a systematic review

Ceska Gynekologie, 2011;76(1)37-45

Tambor V, Fucikova A, Lenco J, Kacerovsky M, Rehacek V, Stulik J and Pudil R

Application of proteomics in biomarker discovery: a primer for the clinician.

Physiological Research, 2010;59(4):471-97.

Andrys C, Drahosova M, Hornychova H, Tambor V, Musilova I, Tosner J, Flidrova E, Kacerovsky M.

Umbilical cord blood concentrations of IL-6, IL-8, and MMP-8 in pregnancy complicated by preterm premature rupture of the membranes and histological chorioamnionitis.

Neuroendocrinology Letters, 2010;31(6):857-63.

Other publications and outputs

Lenco J and Tambor V

Quantitative Mass Spectrometric Approaches.

In Proteomics, Glycomics and Antigenicity of BSL3 and BSL4 Agents, First Edition.

Jiri Stulik, Rudolf Toman, Patrick Butaye, Robert G. Ulrich.

Wiley-VCH Verlag GmbH & Co. KGaA. 2011

Bazhin AV, Tambor V, Dikov B, Philippov PP, Schadendorf D, Eichmüller SB.

cGMP-phosphodiesterase 6, transducin and Wnt5a/Frizzled-2-signaling control cGMP and Ca(2+) homeostasis in melanoma cells.

Cellular and Molecular Life Sciences, 2010;67(5):817-28.

Lenco J, Link M, Tambor V, Zaková J, Cerveny L, Stulik J.

iTRAQ quantitative analysis of Francisella tularensis ssp. holarctica live vaccine strain and Francisella tularensis ssp. tularensis SCHU S4 response to different temperatures and stationary phases of growth.

Proteomics, 2009;9(10):2875-82.

Kacerovsky M, Tambor V, Lenco J, Tosner J.

Proteomics and biomarkers for detection of ovarian cancer.

Ceska Gynekologie, 2009;74(3):163-70.

Awards

1st place - Oral presentation, PhD students conference, Faculty of Military Health Sciences, University of Defence, Hradec Kralove, October 2011
Deep proteome characterization as a tool for identification of novel intraamniotic infection and inflammation biomarkers in preterm birth patients

1st place - Oral presentation, 5th Central and Eastern European Proteomic Conference, Prague, September 2011
Deep proteome characterization as a tool for identification of novel intraamniotic infection and inflammation biomarkers in PPROM patients

1st place - Oral presentation, PhD student forum, 6th Annual Meeting, Bridges in Life Sciences, RECOOP Consortium, Bratislava, April 2011
Deep proteome characterization as a tool for identification of novel intraamniotic infection and inflammation biomarkers in preterm birth patients

1st place - Oral presentation, PhD students conference, Faculty of Military Health Sciences, University of Defence, Hradec Kralove, October 2010
CysTRAQ: nová strategie pro zkoumání a kvantifikaci skrytých proteomů

1st place - Oral presentation, PhD student forum, 5th Annual Meeting, Bridges in Life Sciences, RECOOP Consortium, Lviv, Ukraine, April 2010
CysTRAQ: a novel strategy for uncovering and quantifying hidden proteomes

External tables

Table 8. Proteins identified in the discovery phases of the PPROM and sPTL cohorts. Proteins are reported at 5% FDR. UniProt Accession number is a unique protein identifier in the UniProt database (http://www.uniprot.org). The "Peptides" column shows the number of peptides detected for a particular protein. "Pos vs. Neg" columns show an x-fold change of a particular protein in the positive group compared to negative patient group. These data are presented in two replicates, with corresponding *p*-values. The "PANTHER Protein Class" is an entry from the PANTHER database for a particular protein. "N.D." indicates that the particular protein was not detected in one of the projects. "N.A." is entered if no PANTHER Protein Class was annotated for a particular protein in the database. A dash instead of a *p*-value means that no value could be computed due to low number of peptides.

Table 9. Dysregulated protein comparison between both PPROM and sPTL projects. This table is derived from Table 8. In order for a protein to be listed in this table, the *p*-value had to be 0.05 or less in at least one replicate in either cohort. Also, the relative change had to be at least 1.5-fold (1.50 for proteins with increased levels or 0.66 for proteins with decreased levels) in at least one replicate in either cohort. The columns are closer described in the caption of Table 8.

EXTERNAL TABLES
TABLE 8

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P61604	10 kDa heat shock protein, mitochondrial	1	**1.66**	-	**0.89**	-			N.D.			chaperonin
P31946	14-3-3 protein beta/alpha	9	**1.14**	-	**1.40**	-			N.D.			chaperone
P62258	14-3-3 protein epsilon	9	**1.57**	6.3E-05	**1.47**	2.5E-02	12	**0.90**	9.2E-01	**0.86**	8.4E-01	chaperone
P61981	14-3-3 protein gamma	9	**1.70**	2.2E-01	**1.41**	2.5E-01	11	**1.51**	3.2E-01	**1.63**	4.6E-01	chaperone
P31947	14-3-3 protein sigma			N.D.			16	**1.16**	8.2E-01	**0.99**	9.9E-01	chaperone
P27348	14-3-3 protein theta	8	**6.41**	-	**1.08**	-	10	**1.17**	6.3E-01	**0.93**	8.4E-01	chaperone
P63104	14-3-3 protein zeta/delta	14	**1.66**	1.8E-05	**1.58**	6.0E-03	24	**0.90**	7.0E-01	**0.90**	7.3E-01	chaperone
P25398	40S ribosomal protein S12	1	**3.18**	-	**2.18**	-			N.D.			ribosomal protein
P62857	40S ribosomal protein S28	1	**1.42**	-	**1.49**	-			N.D.			N.A.
Q9BRK5	45 kDa calcium-binding protein	2	**0.82**	2.4E-01	**0.79**	4.4E-01	2	**0.94**	5.7E-01	**0.83**	5.8E-01	calmodulin
P52209	6-phosphogluconate dehydrogenase, decarboxylating	11	**3.43**	6.8E-08	**4.13**	1.6E-06	10	**1.20**	5.8E-01	**1.21**	6.1E-01	dehydrogenase
O95336	6-phosphogluconolactonase	2	**1.90**	1.2E-01	**1.68**	1.1E-02	4	**1.18**	1.3E-01	**0.99**	8.6E-01	hydrolase
P08253	72 kDa type IV collagenase	20	**1.03**	4.2E-01	**1.02**	8.4E-01	21	**0.83**	6.5E-01	**0.82**	6.8E-01	metalloprotease
P11021	78 kDa glucose-regulated protein	6	**1.40**	7.5E-04	**1.31**	2.8E-01	18	**1.10**	5.8E-01	**1.13**	4.9E-01	Hsp70 family chaperone
Q9UHI8	A disintegrin and metalloproteinase with thrombospondin motifs 1	11	**0.97**	6.9E-01	**0.91**	5.6E-01	7	**0.80**	6.0E-01	**0.92**	8.6E-01	metalloprotease
O95450	A disintegrin and metalloproteinase with thrombospondin motifs 2	1	**0.91**	-	**0.95**	-			N.D.			metalloprotease
Q9UNA0	A disintegrin and metalloproteinase with thrombospondin motifs 5	2	**0.88**	2.7E-02	**0.81**	5.6E-02			N.D.			metalloprotease
Q13085	Acetyl-CoA carboxylase 1			N.D.			1	**1.16**	-	**1.02**	-	ligase
Q13510	Acid ceramidase	1	**1.16**	-	**1.31**	-			N.D.			N.A.
O15143	Actin-related protein 2/3 complex subunit 1B	1	**3.20**	-	**3.37**	-			N.D.			actin family cytoskeletal protein
O15144	Actin-related protein 2/3 complex subunit 2	1	**5.19**	-	**3.23**	-	4	**0.94**	8.0E-01	**0.97**	9.3E-01	N.A.
O15145	Actin-related protein 2/3 complex subunit 3			N.D.			2	**1.06**	9.7E-01	**0.96**	9.8E-01	actin family cytoskeletal protein
P59998	Actin-related protein 2/3 complex subunit 4	4	**3.64**	1.0E-02	**3.98**	6.9E-04	5	**1.15**	5.0E-01	**1.16**	6.8E-01	N.A.
O15511	Actin-related protein 2/3 complex subunit 5			N.D.			1	**1.36**	-	**1.11**	-	actin family cytoskeletal protein
P61158	Actin-related protein 3	4	**3.79**	1.6E-04	**3.42**	5.1E-03	2	**1.48**	4.1E-02	**1.05**	8.7E-01	actin and actin related protein
P62736	Actin, aortic smooth muscle	33	**2.38**	1.4E-01	**2.55**	1.7E-02			N.D.			actin and actin related protein
P60709	Actin, cytoplasmic 1	70	**3.47**	4.2E-12	**3.47**	3.8E-09	88	**1.49**	3.1E-07	**1.36**	4.7E-05	actin and actin related protein
P63261	Actin, cytoplasmic 2	64	**2.09**	7.9E-03	**2.12**	3.6E-03			N.D.			actin and actin related protein
P63267	Actin, gamma-enteric smooth muscle			N.D.			38	**0.82**	2.1E-01	**0.76**	1.3E-01	actin and actin related protein
P36896	Activin receptor type-1B	1	**0.71**	-	**0.88**	-			N.D.			protein kinase
P07108	Acyl-CoA-binding protein	2	**0.97**	7.9E-01	**1.27**	2.3E-01	6	**0.63**	4.2E-01	**0.68**	4.9E-01	transfer/carrier protein
P13798	Acylamino-acid-releasing enzyme			N.D.			2	**0.86**	-	**0.82**	-	serine protease
Q6UY14	ADAMTS-like protein 4	1	**0.95**	4.2E-01	**0.78**	3.1E-01	2	**1.08**	4.0E-01	**1.30**	1.7E-01	metalloprotease
P55263	Adenosine kinase	1	**1.49**	-	**1.14**	-			N.D.			carbohydrate kinase
P23526	Adenosylhomocysteinase			N.D.			1	**1.01**	1.0E+00	**1.00**	-	hydrolase
O60503	Adenylate cyclase type 9	1	**3.65**	-	**4.97**	-			N.D.			adenylate cyclase
P54819	Adenylate kinase 2, mitochondrial			N.D.			1	**1.02**	-	**1.12**	-	nucleotide kinase
P00568	Adenylate kinase isoenzyme 1			N.D.			2	**1.09**	-	**1.17**	-	nucleotide kinase
P30520	Adenylosuccinate synthetase isozyme 2	2	**1.57**	4.7E-01	**2.03**	6.8E-02	1	**0.98**	9.5E-01	**1.23**	5.2E-01	ligase
Q01518	Adenylyl cyclase-associated protein 1	7	**4.41**	8.3E-04	**3.56**	2.3E-03	16	**1.26**	1.2E-01	**1.24**	2.2E-01	actin family cytoskeletal protein
Q9H6B4	Adipocyte adhesion molecule			N.D.			2	**1.05**	7.9E-01	**1.26**	4.6E-01	receptor
Q15848	Adiponectin			N.D.			2	**0.60**	1.4E-02	**0.72**	8.7E-02	peptide hormone
Q10588	ADP-ribosyl cyclase 2	1	**1.72**	2.5E-02	**1.41**	7.6E-02	3	**1.33**	1.5E-01	**1.04**	7.5E-01	cyclase
P84077	ADP-ribosylation factor 1			N.D.			1	**0.14**	-	**0.15**	-	hydrolase
P43652	Afamin	72	**0.90**	8.3E-05	**0.90**	1.1E-01	113	**0.93**	6.2E-01	**0.92**	5.6E-01	transfer/carrier protein
P16112	Aggrecan core protein	1	**0.93**	6.0E-01	**0.74**	3.8E-01	2	**0.88**	5.7E-01	**0.97**	8.8E-01	extracellular matrix glycoprotein
O00468	Agrin	7	**0.95**	3.5E-01	**0.93**	2.2E-01	6	**0.93**	5.5E-01	**0.97**	7.9E-01	N.A.
P14550	Alcohol dehydrogenase [NADP+]	1	**1.24**	-	**0.89**	-	3	**0.94**	7.5E-01	**1.22**	4.1E-01	reductase
P15121	Aldose reductase	1	**1.74**	-	**0.83**	-			N.D.			reductase
P05186	Alkaline phosphatase, tissue-nonspecific isozyme	1	**2.87**	-	**3.34**	-			N.D.			nucleotide phosphatase
P55008	Allograft inflammatory factor 1	1	**6.54**	-	**4.63**	-	2	**1.13**	9.5E-01	**0.90**	9.7E-01	N.A.
P02763	Alpha-1-acid glycoprotein 1	2	**0.75**	1.8E-01	**0.65**	4.0E-01			N.D.			N.A.
P19652	Alpha-1-acid glycoprotein 2	1	**0.42**	-	**0.41**	-			N.D.			N.A.
P01011	Alpha-1-antichymotrypsin	44	**0.70**	3.2E-09	**0.79**	9.1E-02	79	**0.72**	2.6E-01	**0.79**	4.4E-01	serine protease inhibitor
P01009	Alpha-1-antitrypsin	183	**0.62**	0.0E+00	**0.67**	8.1E-13	39	**0.97**	9.2E-01	**0.95**	8.7E-01	serine protease inhibitor
P26572	Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase	2	**0.98**	5.4E-01	**0.93**	1.0E-01	1	**0.56**	7.0E-01	**0.61**	8.6E-01	glycosyltransferase
P04217	Alpha-1B-glycoprotein	69	**0.95**	1.8E-01	**0.91**	6.1E-02	84	**0.82**	1.2E-01	**0.89**	3.9E-01	immunoglobulin receptor superfamily
P08697	Alpha-2-antiplasmin	26	**0.86**	1.8E-02	**0.66**	5.6E-03	38	**0.91**	8.1E-01	**0.88**	7.5E-01	serine protease inhibitor
P02765	Alpha-2-HS-glycoprotein	137	**0.94**	2.4E-02	**0.99**	8.9E-01	159	**0.83**	2.1E-01	**0.90**	4.9E-01	extracellular matrix glycoprotein
A8K2U0	Alpha-2-macroglobulin-like protein 1	17	**0.91**	1.3E-03	**0.88**	3.8E-01	28	**0.90**	6.8E-01	**0.96**	8.6E-01	cytokine
P12814	Alpha-actinin-1	19	**3.13**	4.8E-04	**2.87**	2.4E-03	29	**0.93**	7.7E-01	**0.98**	9.4E-01	N.A.
O43707	Alpha-actinin-4	19	**1.76**	1.1E-08	**1.63**	3.2E-03	31	**1.16**	7.3E-01	**1.09**	8.6E-01	N.A.
P04745	Alpha-amylase 1	6	**0.79**	1.7E-01	**0.64**	3.9E-02			N.D.			N.A.
P19961	Alpha-amylase 2B			N.D.			5	**1.15**	7.2E-01	**1.15**	6.9E-01	amylase
P06733	Alpha-enolase	20	**3.14**	1.6E-07	**3.03**	3.1E-06	31	**1.42**	2.5E-01	**1.31**	4.3E-01	lyase
P02771	Alpha-fetoprotein	31	**1.65**	2.2E-11	**1.64**	1.2E-05	26	**0.93**	7.2E-01	**0.98**	9.3E-01	transfer/carrier protein
P00709	Alpha-lactalbumin			N.D.			2	**0.40**	4.1E-01	**0.43**	4.7E-01	hydrolase
Q16706	Alpha-mannosidase 2	1	**0.98**	-	**0.66**	-			N.D.			glycosidase
Q9NSC7	Alpha-N-acetylgalactosaminide alpha-2,6-sialyltransferase 1	1	**0.91**	-	**1.11**	-			N.D.			glycosyltransferase
P19801	Amiloride-sensitive amine oxidase [copper-containing]	30	**0.67**	1.6E-11	**0.74**	3.1E-04	34	**1.22**	2.5E-01	**1.30**	1.2E-01	oxidase
P15144	Aminopeptidase N	14	**2.48**	6.8E-09	**2.29**	1.0E-03	10	**0.88**	4.9E-01	**0.86**	3.6E-01	metalloprotease
P05067	Amyloid beta A4 protein	1	**0.99**	-	**1.14**	-	1	**1.18**	-	**1.22**	-	signaling molecule
P03950	Angiogenin	1	**1.01**	9.7E-01	**0.77**	4.2E-01	4	**0.66**	4.7E-01	**0.59**	3.3E-01	endoribonuclease
Q9UKU9	Angiopoietin-related protein 2	6	**0.93**	3.0E-01	**0.80**	1.2E-02	3	**0.73**	1.9E-01	**0.55**	2.2E-01	signaling molecule
Q9BY76	Angiopoietin-related protein 4			N.D.			1	**1.13**	-	**1.04**	-	signaling molecule
P01019	Angiotensinogen	64	**1.01**	7.6E-01	**0.94**	6.8E-01	56	**0.76**	4.6E-01	**0.73**	2.8E-01	serine protease inhibitor
Q3KP44	Ankyrin repeat domain-containing protein 55			N.D.			1	**19.79**	-	**25.72**	-	N.A.
P04083	Annexin A1	11	**0.86**	2.1E-02	**1.01**	9.5E-01	17	**1.19**	5.8E-01	**1.21**	5.1E-01	transfer/carrier protein
P07355	Annexin A2	12	**0.93**	5.4E-01	**0.93**	6.5E-01	16	**1.04**	8.1E-01	**0.95**	7.4E-01	transfer/carrier protein
P12429	Annexin A3	3	**1.96**	2.5E-03	**1.96**	8.7E-02			N.D.			transfer/carrier protein
P09525	Annexin A4	1	**1.29**	3.5E-02	**0.91**	9.3E-02	1	**0.95**	7.5E-01	**1.28**	3.3E-01	transfer/carrier protein
P08758	Annexin A5	1	**1.45**	-	**1.41**	-	2	**1.25**	5.1E-01	**0.88**	6.8E-01	transfer/carrier protein
P08133	Annexin A6			N.D.			1	**1.67**	-	**1.34**	-	transfer/carrier protein
Q5VT79	Annexin A8-like protein 2	1	**0.90**	-	**0.82**	-			N.D.			transfer/carrier protein
O95994	Anterior gradient protein 2 homolog			N.D.			2	**0.67**	4.5E-01	**0.73**	1.0E-01	surfactant
Q9H6X2	Anthrax toxin receptor 1	1	**0.95**	7.0E-01	**0.89**	3.6E-01	2	**1.37**	2.8E-01	**1.33**	2.8E-01	receptor
P03973	Antileukoproteinase	10	**0.89**	5.5E-02	**0.94**	7.7E-01	28	**2.06**	2.3E-04	**1.93**	4.4E-04	serine protease inhibitor
P01008	Antithrombin-III	88	**0.94**	4.4E-02	**0.84**	4.2E-03	101	**0.80**	2.5E-01	**0.80**	2.0E-01	serine protease inhibitor
P63010	AP-2 complex subunit beta	1	**1.69**	-	**2.33**	-			N.D.			membrane traffic protein
P02647	Apolipoprotein A-I	7	**1.04**	8.2E-01	**0.86**	4.4E-01	21	**0.96**	9.2E-01	**0.91**	8.5E-01	transporter
P02652	Apolipoprotein A-II	16	**0.94**	2.4E-01	**1.08**	7.9E-01	14	**1.00**	1.0E+00	**1.01**	9.9E-01	transporter
P06727	Apolipoprotein A-IV	44	**0.76**	7.7E-07	**0.80**	1.1E-02	73	**0.91**	5.2E-01	**0.83**	2.8E-01	transporter
P04114	Apolipoprotein B-100	18	**1.44**	6.8E-09	**0.99**	9.4E-01	112	**0.95**	7.7E-01	**0.92**	6.6E-01	apolipoprotein
Q0VD83	Apolipoprotein B-100 receptor	1	**2.41**	-	**3.15**	-			N.D.			receptor
P02656	Apolipoprotein C-III			N.D.			7	**1.68**	2.2E-01	**1.37**	5.2E-01	N.A.
P05090	Apolipoprotein D	1	**1.70**	-	**1.06**	-			N.D.			apolipoprotein
P02649	Apolipoprotein E	7	**1.10**	2.7E-01	**1.34**	3.4E-01	15	**1.28**	4.4E-01	**1.30**	4.0E-01	transporter
O14791	Apolipoprotein L1			N.D.			2	**1.13**	5.1E-01	**1.13**	5.5E-01	transporter
O95445	Apolipoprotein M	2	**1.29**	-	**1.20**	-	3	**1.74**	6.4E-02	**1.51**	1.2E-01	N.A.
P08519	Apolipoprotein(a)			N.D.			5	**1.44**	5.1E-01	**1.68**	4.8E-01	peptide hormone
Q9ULZ3	Apoptosis-associated speck-like protein containing a CARD	1	**3.72**	-	**4.27**	-	2	**1.06**	7.7E-01	**0.80**	5.5E-01	cysteine protease
P05089	Arginase-1	1	**8.77**	-	**1.23**	-			N.D.			hydrolase

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P17174	Aspartate aminotransferase, cytoplasmic	1	**1.34**	-	**1.10**	-	1	**0.83**	3.0E-01	**0.73**	2.9E-01	transaminase
Q15121	Astrocytic phosphoprotein PEA-15			N.D.			1	**1.45**	-	**0.87**	-	N.A.
O94911	ATP-binding cassette sub-family A member 8	1	**1.17**	-	**1.30**	-			N.D.			ATP-binding cassette (ABC) transporter
P53396	ATP-citrate synthase			N.D.			1	**0.64**	-	**0.80**	-	transferase
O75882	Attractin	8	**0.90**	1.6E-01	**1.01**	9.6E-01	19	**1.26**	4.8E-01	**1.31**	4.7E-01	N.A.
P20160	Azurocidin	3	**6.57**	5.5E-04	**6.70**	2.5E-02	1	**1.30**	2.1E-01	**1.26**	2.2E-01	serine protease
P17213	Bactericidal permeability-increasing protein			N.D.			3	**0.79**	4.4E-01	**0.67**	1.0E-01	transfer/carrier protein
Q8N4F0	Bactericidal/permeability-increasing protein-like 1	9	**0.86**	1.5E-02	**0.50**	1.1E-02	8	**0.68**	5.0E-01	**0.57**	4.7E-01	carbohydrate transporter
O95817	BAG family molecular chaperone regulator 3			N.D.			1	**1.44**	-	**1.00**	-	chaperone
P02730	Band 3 anion transport protein			N.D.			2	**0.68**	5.8E-01	**0.94**	8.6E-01	transporter
P50895	Basal cell adhesion molecule	1	**1.11**	5.8E-01	**0.70**	6.8E-02			N.D.			receptor
P98160	Basement membrane heparan sulfate proteoglycan core protein	59	**1.01**	8.3E-01	**1.00**	1.0E+00	68	**0.97**	8.0E-01	**0.98**	8.6E-01	N.A.
P15291	Beta-1,4-galactosyltransferase 1	4	**0.93**	1.3E-01	**0.71**	4.7E-01	3	**1.21**	4.9E-01	**1.24**	4.8E-01	glycosyltransferase
P02749	Beta-2-glycoprotein 1	46	**1.05**	1.1E-01	**0.94**	5.8E-01	80	**1.04**	8.6E-01	**0.99**	9.6E-01	apolipoprotein
P61769	Beta-2-microglobulin	37	**0.91**	1.4E-01	**1.05**	5.4E-01	39	**0.94**	6.9E-01	**1.21**	3.0E-01	major histocompatibility complex
Q562R1	Beta-actin-like protein 2	19	**2.54**	-	**3.31**	-			N.D.			actin and actin related protein
Q96KN2	Beta-Ala-His dipeptidase	3	**0.94**	1.7E-01	**0.75**	2.7E-01	6	**0.84**	5.4E-01	**0.88**	8.3E-01	metalloprotease
P60022	Beta-defensin 1	1	**0.84**	-	**0.94**	-			N.D.			antibacterial response protein
P06865	Beta-hexosaminidase subunit alpha			N.D.			3	**0.38**	2.7E-02	**0.34**	4.4E-02	glycosidase
P07686	Beta-hexosaminidase subunit beta	1	**1.60**	1.6E-01	**1.42**	3.5E-01	4	**0.98**	9.8E-01	**1.02**	9.7E-01	glycosidase
O00462	Beta-mannosidase	1	**0.72**	-	**0.75**	-			N.D.			galactosidase
P08118	Beta-microseminoprotein	5	**0.89**	5.2E-01	**1.31**	7.2E-01	3	**1.25**	8.8E-01	**1.20**	8.9E-01	peptide hormone
P21810	Biglycan			N.D.			1	**0.90**	-	**0.62**	-	receptor
P19835	Bile salt-activated lipase			N.D.			2	**2.07**	1.1E-01	**2.44**	7.4E-02	peptide hormone
Q8NFC6	Biorientation of chromosomes in cell division protein 1-like			N.D.			1	**1.53**	7.4E-01	**2.02**	5.6E-01	N.A.
P43251	Biotinidase	9	**0.90**	1.1E-01	**0.74**	1.3E-01	4	**0.56**	3.7E-01	**0.58**	4.6E-01	hydrolase
P07738	Bisphosphoglycerate mutase	1	**0.93**	-	**0.30**	-	6	**1.18**	5.6E-01	**1.18**	3.7E-01	mutase
Q13867	Bleomycin hydrolase	3	**0.81**	6.4E-02	**0.97**	8.9E-01	2	**0.81**	2.7E-01	**0.88**	2.2E-01	cysteine protease
P13727	Bone marrow proteoglycan	37	**0.66**	7.4E-12	**0.61**	4.1E-17	39	**0.96**	6.7E-01	**1.03**	7.2E-01	extracellular matrix structural protein
P13497	Bone morphogenetic protein 1	2	**0.79**	2.4E-01	**0.96**	9.1E-01	4	**0.98**	8.0E-01	**0.95**	7.9E-01	transporter
P80723	Brain acid soluble protein 1	3	**4.39**	7.0E-03	**6.77**	1.7E-02	6	**1.41**	1.8E-01	**1.37**	1.5E-01	N.A.
Q96CX2	BTB/POZ domain-containing protein KCTD12			N.D.			2	**1.21**	2.0E-01	**1.35**	3.0E-01	enzyme modulator
Q8WVV5	Butyrophilin subfamily 2 member A2	4	**0.74**	2.6E-01	**0.72**	3.0E-01			N.D.			ubiquitin-protein ligase
Q16627	C-C motif chemokine 14			N.D.			2	**0.24**	1.1E-01	**0.20**	7.7E-02	chemokine
P55774	C-C motif chemokine 18	1	**1.55**	-	**2.00**	-			N.D.			chemokine
P02741	C-reactive protein	3	**1.25**	7.2E-02	**1.81**	4.4E-01	3	**1.88**	5.6E-01	**2.24**	4.7E-01	antibacterial response protein
Q9UBG0	C-type mannose receptor 2	2	**1.04**	8.6E-01	**1.24**	6.4E-01			N.D.			receptor
P04003	C4b-binding protein alpha chain			N.D.			10	**0.94**	8.7E-01	**1.10**	7.6E-01	apolipoprotein
P12830	Cadherin-1	9	**0.99**	8.0E-01	**0.70**	6.9E-02	9	**0.72**	4.6E-01	**0.76**	6.0E-01	G-protein coupled receptor
P55287	Cadherin-11	2	**0.93**	3.7E-01	**0.93**	7.6E-01	1	**1.39**	9.2E-01	**1.23**	-	G-protein coupled receptor
P55290	Cadherin-13	3	**0.98**	8.5E-01	**1.00**	9.8E-01	8	**1.07**	3.5E-01	**1.10**	3.6E-01	G-protein coupled receptor
P55291	Cadherin-15			N.D.			2	**1.05**	8.0E-01	**0.91**	3.7E-01	G-protein coupled receptor
P19022	Cadherin-2	1	**0.60**	-	**0.53**	-	1	**0.59**	-	**0.29**	-	G-protein coupled receptor
P33151	Cadherin-5	3	**1.17**	2.6E-02	**1.06**	7.6E-01	2	**0.93**	7.2E-01	**0.80**	2.2E-01	G-protein coupled receptor
Q9BYE9	Cadherin-related family member 2	2	**2.39**	1.0E-01	**1.46**	2.7E-01			N.D.			G-protein coupled receptor
A8K7I4	Calcium-activated chloride channel regulator 1	3	**2.09**	1.5E-01	**2.24**	1.1E-01			N.D.			ion channel
Q13938	Calcyphosin			N.D.			4	**0.89**	7.7E-01	**0.86**	6.1E-01	signaling molecule
Q05682	Caldesmon			N.D.			1	**0.89**	-	**1.19**	-	non-motor actin binding protein
P62158	Calmodulin	9	**1.47**	1.3E-03	**1.91**	4.0E-03	8	**0.71**	3.9E-02	**0.76**	1.6E-01	calmodulin
O94983	Calmodulin-binding transcription activator 2			N.D.			1	**0.75**	9.0E-01	**0.68**	9.0E-01	transcription factor
P27482	Calmodulin-like protein 3	4	**0.95**	5.3E-01	**1.06**	7.5E-01	10	**0.78**	1.4E-01	**0.80**	2.5E-01	calmodulin
Q9NZT1	Calmodulin-like protein 5			N.D.			6	**0.97**	9.4E-01	**1.10**	8.7E-01	calmodulin
P04632	Calpain small subunit 1	3	**1.86**	3.1E-01	**0.82**	3.3E-01			N.D.			cysteine protease
P07384	Calpain-1 catalytic subunit	2	**1.22**	3.8E-01	**0.90**	4.3E-01	4	**1.00**	9.8E-01	**1.02**	9.1E-01	cysteine protease
P17655	Calpain-2 catalytic subunit	1	**1.05**	-	**1.10**	-	1	**0.81**	7.1E-01	**1.27**	6.6E-01	cysteine protease
P20810	Calpastatin	1	**1.25**	3.5E-01	**1.05**	1.3E-01	6	**0.97**	9.0E-01	**0.86**	6.2E-01	cysteine protease inhibitor
P27797	Calreticulin	10	**1.49**	1.5E-08	**1.57**	1.7E-05	13	**1.02**	9.6E-01	**1.08**	8.7E-01	calcium-binding protein
P22676	Calretinin			N.D.			1	**0.49**	-	**0.55**	-	calcium-binding protein
O94985	Calsyntenin-1	6	**0.83**	1.2E-02	**0.86**	8.9E-02	5	**1.10**	7.7E-01	**1.07**	8.4E-01	cell adhesion molecule
O43852	Calumenin	2	**1.06**	-	**0.83**	-	4	**0.72**	7.5E-01	**0.61**	5.8E-01	calmodulin
Q08499	cAMP-specific 3',5'-cyclic phosphodiesterase 4D	1	**0.91**	-	**1.02**	-			N.D.			phosphodiesterase
P30622	CAP-Gly domain-containing linker protein 1			N.D.			1	**1.73**	3.0E-01	**1.83**	2.8E-01	non-motor microtubule binding protein
P00915	Carbonic anhydrase 1	9	**0.62**	1.8E-03	**0.57**	2.2E-04	48	**0.59**	5.6E-02	**0.65**	5.9E-02	dehydratase
P00918	Carbonic anhydrase 2	2	**1.11**	3.8E-01	**1.15**	4.3E-01	17	**1.44**	1.3E-01	**1.35**	1.7E-01	dehydratase
P07451	Carbonic anhydrase 3			N.D.			5	**0.85**	7.9E-01	**0.81**	6.9E-01	dehydratase
O00748	Carboxylesterase 2			N.D.			2	**0.83**	5.5E-01	**0.83**	4.6E-01	peptide hormone
P15085	Carboxypeptidase A1			N.D.			2	**7.53**	1.6E-01	**9.62**	1.5E-01	metalloprotease
Q96IY4	Carboxypeptidase B2	7	**0.72**	8.7E-03	**0.67**	5.5E-03	9	**0.80**	1.2E-01	**0.71**	1.4E-02	metalloprotease
P14384	Carboxypeptidase M	7	**0.75**	4.9E-04	**0.78**	2.1E-02	14	**0.78**	6.8E-03	**0.78**	3.6E-03	metalloprotease
P15169	Carboxypeptidase N catalytic chain			N.D.			3	**1.55**	2.3E-01	**1.65**	2.1E-01	metalloprotease
P22792	Carboxypeptidase N subunit 2	2	**1.03**	6.1E-01	**1.07**	9.3E-01	7	**0.93**	8.6E-01	**0.99**	9.8E-01	receptor
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	3	**0.89**	1.7E-01	**0.93**	6.1E-01	4	**0.41**	9.8E-02	**0.42**	1.3E-01	immunoglobulin cell adhesion molecule
Q9NQ79	Cartilage acidic protein 1	2	**1.15**	3.4E-01	**1.06**	6.7E-01	5	**0.81**	5.8E-01	**0.96**	8.8E-01	annexin
O75339	Cartilage intermediate layer protein 1			N.D.			1	**2.09**	-	**2.30**	-	N.A.
Q8IUL8	Cartilage intermediate layer protein 2	6	**0.83**	5.3E-03	**1.02**	8.3E-01	3	**1.03**	9.0E-01	**1.29**	3.8E-01	N.A.
P49747	Cartilage oligomeric matrix protein	13	**0.89**	2.0E-02	**0.89**	4.8E-01	18	**0.87**	7.3E-01	**0.86**	7.5E-01	signaling molecule
P31944	Caspase-14			N.D.			2	**0.50**	1.7E-01	**0.45**	5.7E-02	cysteine protease
P04040	Catalase	12	**1.52**	4.2E-03	**1.40**	1.2E-02	26	**1.12**	7.3E-01	**1.17**	5.8E-01	peroxidase
P49913	Cathelicidin antimicrobial peptide	4	**5.93**	2.0E-03	**6.96**	4.0E-03	10	**0.83**	2.8E-01	**0.80**	3.1E-01	protease inhibitor
P07858	Cathepsin B	8	**1.89**	1.4E-04	**1.80**	1.0E-03	7	**0.94**	9.2E-01	**0.88**	8.2E-01	cysteine protease
P07339	Cathepsin D	3	**2.10**	9.4E-04	**1.89**	1.7E-01	5	**0.92**	8.9E-01	**1.04**	9.4E-01	aspartic protease
P08311	Cathepsin G			N.D.			3	**0.99**	9.8E-01	**0.85**	6.9E-01	serine protease
P09668	Cathepsin H	4	**1.73**	3.6E-03	**1.87**	8.2E-02	6	**0.73**	9.3E-02	**0.76**	6.4E-02	cysteine protease
P07711	Cathepsin L1	1	**2.00**	-	**2.18**	-			N.D.			cysteine protease
P25774	Cathepsin S			N.D.			4	**0.97**	9.2E-01	**1.04**	9.4E-01	cysteine protease
Q6YHK3	CD109 antigen	4	**1.03**	8.3E-01	**0.88**	6.7E-01	1	**0.64**	-	**3.23**	-	cytokine
Q13740	CD166 antigen	2	**0.98**	9.1E-01	**0.67**	1.4E-01	2	**0.59**	5.8E-01	**0.64**	5.7E-01	receptor
Q8N6Q3	CD177 antigen			N.D.			2	**2.01**	7.4E-01	**1.77**	8.5E-01	N.A.
P16070	CD44 antigen	4	**1.15**	7.6E-02	**1.17**	6.3E-01	7	**1.08**	9.4E-01	**1.13**	9.1E-01	receptor
P13987	CD59 glycoprotein	17	**0.83**	1.9E-04	**0.88**	5.5E-01	18	**0.89**	7.3E-01	**0.84**	5.9E-01	membrane-bound signaling molecule
P21926	CD9 antigen	2	**0.91**	7.1E-01	**0.90**	8.8E-01	2	**1.21**	9.5E-01	**1.12**	9.7E-01	membrane-bound signaling molecule
Q9BY67	Cell adhesion molecule 1	1	**1.00**	-	**0.95**	-	2	**0.80**	4.0E-01	**0.93**	7.6E-01	receptor
Q8NFZ8	Cell adhesion molecule 4	3	**0.80**	3.2E-01	**1.12**	5.0E-01	3	**0.93**	7.2E-01	**0.97**	9.1E-01	receptor
P60953	Cell division control protein 42 homolog	3	**1.43**	1.0E-01	**1.38**	1.7E-01			N.D.			small GTPase
P43121	Cell surface glycoprotein MUC18			N.D.			1	**1.21**	8.4E-01	**1.09**	9.3E-01	receptor
P29373	Cellular retinoic acid-binding protein 2	1	**0.83**	-	**0.93**	-	6	**0.86**	3.7E-01	**0.91**	6.5E-01	transfer/carrier protein
O75503	Ceroid-lipofuscinosis neuronal protein 5	1	**0.92**	8.7E-01	**0.85**	6.1E-01			N.D.			N.A.
P00450	Ceruloplasmin	239	**0.84**	4.8E-15	**0.83**	6.2E-04	330	**0.79**	7.7E-02	**0.82**	1.3E-01	transporter
O43633	Charged multivesicular body protein 2a			N.D.			1	**1.58**	-	**0.78**	-	N.A.
P36222	Chitinase-3-like protein 1	13	**3.46**	1.0E-10	**3.28**	6.8E-07	11	**0.92**	6.4E-01	**0.80**	4.0E-01	glycosidase
Q13231	Chitotriosidase-1	2	**6.63**	6.5E-02	**5.46**	3.3E-02	3	**1.40**	2.3E-01	**1.69**	1.6E-01	glycosidase
O00299	Chloride intracellular channel protein 1	3	**1.93**	2.9E-02	**2.30**	3.8E-03	3	**1.49**	4.0E-02	**1.12**	6.3E-01	N.A.

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P11597	Cholesteryl ester transfer protein	1	**1.36**	-	**0.64**	-			N.D.			transfer/carrier protein
Q8WWI5	Choline transporter-like protein 1			N.D.			1	**1.34**	-	**1.77**	-	transporter
P06276	Cholinesterase			N.D.			2	**0.78**	9.2E-01	**0.92**	9.7E-01	signaling molecule
Q6UVK1	Chondroitin sulfate proteoglycan 4	4	**1.05**	8.2E-01	**1.18**	5.0E-01	5	**1.05**	6.7E-01	**0.98**	9.0E-01	N.A.
Q9H2X0	Chordin			N.D.			1	**1.12**	-	**1.21**	-	signaling molecule
Q9BU40	Chordin-like protein 1	1	**0.71**	2.3E-01	**0.84**	6.7E-01	4	**0.79**	8.3E-01	**0.85**	8.7E-01	signaling molecule
Q6NT52	Choriogonadotropin subunit beta variant 2			N.D.			2	**0.74**	8.2E-01	**0.58**	6.6E-01	peptide hormone
P01243	Chorionic somatomammotropin hormone	10	**0.73**	4.0E-02	**0.76**	2.2E-01	20	**0.87**	7.5E-01	**0.89**	7.9E-01	growth factor
Q13185	Chromobox protein homolog 3			N.D.			1	**1.18**	-	**0.91**	-	chromatin/chromatin-binding protein
P10645	Chromogranin-A	1	**1.12**	-	**0.65**	-	1	**0.35**	-	**0.46**	-	peptide hormone
P09497	Clathrin light chain B			N.D.			1	**0.96**	-	**0.73**	-	vesicle coat protein
P10909	Clusterin	29	**0.73**	4.3E-07	**0.82**	5.7E-02	42	**1.01**	9.6E-01	**0.98**	9.5E-01	N.A.
Q6UXG3	CMRF35-like molecule 9	1	**0.66**	-	**0.94**	-			N.D.			immunoglobulin receptor superfamily
Q14019	Coactosin-like protein	3	**2.16**	7.0E-02	**2.02**	2.0E-01	8	**0.98**	9.9E-01	**0.89**	9.0E-01	N.A.
P00740	Coagulation factor IX	3	**0.68**	4.9E-02	**0.58**	8.9E-02	3	**0.75**	3.0E-01	**0.90**	4.4E-01	serine protease
P12259	Coagulation factor V			N.D.			2	**1.24**	8.5E-01	**0.83**	8.8E-01	transporter
P00742	Coagulation factor X	4	**0.83**	5.9E-02	**0.81**	4.7E-01	5	**1.33**	5.8E-01	**1.31**	6.4E-01	serine protease
P00748	Coagulation factor XII	26	**0.85**	7.5E-03	**0.80**	7.8E-02	32	**1.07**	6.5E-01	**1.14**	4.1E-01	serine protease
P00488	Coagulation factor XIII A chain	1	**2.06**	-	**1.61**	-			N.D.			acyltransferase
P05160	Coagulation factor XIII B chain	3	**0.98**	9.1E-01	**1.03**	8.8E-01	14	**0.83**	5.9E-01	**0.94**	8.5E-01	apolipoprotein
P23528	Cofilin-1	4	**3.49**	1.2E-02	**3.31**	1.7E-02	14	**1.32**	5.1E-01	**1.30**	5.5E-01	non-motor actin binding protein
Q6ZUS5	Coiled-coil domain-containing protein 121			N.D.			1	**1.36**	-	**0.78**	-	N.A.
P02452	Collagen alpha-1(I) chain	85	**0.78**	4.1E-08	**0.93**	3.1E-01	127	**0.71**	5.1E-02	**0.71**	7.7E-02	peptide hormone
P02458	Collagen alpha-1(II) chain	20	**0.83**	3.0E-02	**0.90**	3.3E-01	28	**0.78**	6.0E-01	**0.76**	4.6E-01	surfactant
P02461	Collagen alpha-1(III) chain	105	**0.74**	1.0E-10	**0.78**	3.0E-04	107	**0.75**	2.2E-01	**0.75**	2.7E-01	surfactant
P20908	Collagen alpha-1(V) chain	16	**0.89**	6.4E-02	**1.03**	9.0E-01	15	**0.90**	7.5E-01	**0.91**	7.5E-01	peptide hormone
P12109	Collagen alpha-1(VI) chain	20	**0.89**	4.1E-02	**1.01**	9.5E-01	18	**0.95**	8.4E-01	**0.90**	6.9E-01	peptide hormone
Q03692	Collagen alpha-1(X) chain	2	**0.00**	-	**0.00**	-			N.D.			peptide hormone
P12107	Collagen alpha-1(XI) chain	3	**0.67**	9.9E-02	**0.64**	4.7E-02			N.D.			peptide hormone
Q99715	Collagen alpha-1(XII) chain	5	**1.23**	3.9E-02	**1.11**	2.0E-01	5	**0.78**	7.4E-02	**0.86**	2.6E-01	peptide hormone
P39059	Collagen alpha-1(XV) chain	6	**0.90**	8.0E-02	**0.74**	5.5E-03	4	**0.78**	1.5E-01	**0.79**	3.3E-01	peptide hormone
Q9UMD9	Collagen alpha-1(XVII) chain	1	**1.65**	-	**0.90**	-	1	**0.51**	-	**0.41**	-	peptide hormone
P39060	Collagen alpha-1(XVIII) chain	4	**0.82**	4.6E-01	**0.75**	7.8E-02	7	**0.85**	7.2E-01	**0.82**	6.4E-01	peptide hormone
P08123	Collagen alpha-2(I) chain	52	**0.79**	8.6E-07	**1.08**	6.0E-01	108	**0.94**	7.3E-01	**0.91**	6.1E-01	surfactant
P08572	Collagen alpha-2(IV) chain	13	**0.74**	2.1E-04	**0.79**	8.2E-02	14	**0.84**	2.9E-01	**0.81**	2.9E-01	peptide hormone
P05997	Collagen alpha-2(V) chain	2	**0.92**	6.9E-01	**0.82**	4.4E-01	3	**0.89**	1.6E-01	**0.90**	4.5E-01	surfactant
P12110	Collagen alpha-2(VI) chain	3	**0.87**	1.1E-01	**0.90**	3.1E-01			N.D.			peptide hormone
P13942	Collagen alpha-2(XI) chain	1	**0.57**	5.3E-02	**0.68**	1.0E-01			N.D.			peptide hormone
Q01955	Collagen alpha-3(IV) chain	2	**0.77**	6.1E-01	**0.67**	2.4E-01			N.D.			peptide hormone
P25940	Collagen alpha-3(V) chain	1	**0.71**	4.4E-02	**1.17**	3.7E-01	3	**0.79**	6.6E-01	**0.78**	6.1E-01	peptide hormone
P12111	Collagen alpha-3(VI) chain	42	**1.03**	4.7E-01	**1.15**	2.7E-01	45	**0.99**	9.7E-01	**0.97**	8.9E-01	surfactant
P29400	Collagen alpha-5(IV) chain			N.D.			2	**1.57**	7.2E-02	**2.00**	3.2E-01	peptide hormone
Q5KU26	Collectin-12	4	**0.90**	2.2E-01	**0.96**	7.0E-01	3	**0.91**	4.7E-01	**0.81**	2.1E-01	transporter
P02745	Complement C1q subcomponent subunit A	1	**1.42**	-	**1.50**	-	3	**1.24**	2.1E-02	**1.12**	4.2E-01	peptide hormone
P02746	Complement C1q subcomponent subunit B			N.D.			4	**1.06**	9.1E-01	**0.98**	9.8E-01	peptide hormone
P02747	Complement C1q subcomponent subunit C	1	**1.52**	-	**1.19**	-	3	**1.25**	7.7E-01	**1.21**	7.8E-01	peptide hormone
P00736	Complement C1r subcomponent	9	**1.10**	3.9E-02	**1.13**	3.8E-01	16	**1.22**	4.4E-01	**1.26**	5.4E-01	serine protease
Q9NZP8	Complement C1r subcomponent-like protein	5	**1.02**	9.4E-01	**1.46**	6.6E-01	10	**0.84**	6.2E-01	**0.77**	3.8E-01	serine protease
P09871	Complement C1s subcomponent	3	**1.00**	9.9E-01	**0.83**	2.6E-01	8	**0.77**	5.4E-01	**0.76**	5.1E-01	serine protease
P06681	Complement C2	23	**0.94**	2.5E-01	**0.99**	8.8E-01	32	**0.96**	8.6E-01	**0.96**	8.8E-01	N.A.
P01024	Complement C3	52	**1.54**	3.9E-12	**1.62**	1.3E-05	111	**1.08**	5.8E-01	**1.06**	7.4E-01	cytokine
P0C0L4	Complement C4-A	102	**0.99**	-	**0.74**	-			N.D.			cytokine
P0C0L5	Complement C4-B	103	**1.04**	7.9E-01	**0.79**	4.0E-01	160	**0.00**	-	**0.00**	-	N.A.
P01031	Complement C5	31	**0.91**	1.6E-02	**0.86**	5.3E-02	23	**1.79**	1.3E-01	**1.56**	4.2E-01	cytokine
Q9NPY3	Complement component C1q receptor	2	**1.83**	-	**1.49**	-			N.D.			transcription factor
P13671	Complement component C6	23	**0.87**	1.7E-03	**0.87**	2.0E-01	30	**1.01**	9.8E-01	**0.95**	8.6E-01	apolipoprotein
P10643	Complement component C7	17	**0.86**	1.7E-02	**0.74**	1.0E-03	34	**0.80**	1.9E-01	**0.76**	1.3E-01	apolipoprotein
P07357	Complement component C8 alpha chain	15	**0.73**	3.6E-04	**0.79**	8.8E-03	19	**0.88**	5.3E-01	**0.91**	5.9E-01	apolipoprotein
P07358	Complement component C8 beta chain	20	**0.73**	1.9E-05	**0.71**	7.7E-05	17	**0.65**	2.4E-03	**0.62**	3.9E-03	receptor
P07360	Complement component C8 gamma chain	7	**0.76**	4.6E-04	**0.70**	5.2E-04	11	**0.77**	2.2E-01	**0.72**	2.0E-01	serine protease inhibitor
P02748	Complement component C9	28	**0.81**	6.7E-04	**0.80**	6.9E-02	42	**0.86**	5.6E-01	**0.79**	3.9E-01	receptor
P08174	Complement decay-accelerating factor	9	**0.92**	2.3E-01	**0.97**	7.5E-01	12	**1.14**	5.1E-01	**1.19**	4.0E-01	apolipoprotein
P00751	Complement factor B	97	**0.92**	7.9E-03	**1.01**	8.9E-01	133	**1.05**	7.3E-01	**1.07**	6.3E-01	apolipoprotein
P00746	Complement factor D	11	**1.14**	2.9E-01	**1.19**	5.9E-01	14	**1.16**	8.1E-01	**1.04**	9.3E-01	serine protease
P08603	Complement factor H	71	**1.01**	7.9E-01	**1.05**	4.6E-01	148	**1.08**	5.7E-01	**1.07**	5.6E-01	apolipoprotein
Q03591	Complement factor H-related protein 1	12	**0.91**	2.6E-01	**0.78**	7.8E-04	22	**1.14**	8.2E-01	**1.07**	9.3E-01	apolipoprotein
P36980	Complement factor H-related protein 2	12	**0.92**	3.8E-01	**0.91**	1.6E-04	17	**1.07**	8.7E-01	**1.06**	9.1E-01	apolipoprotein
Q02985	Complement factor H-related protein 3			N.D.			6	**1.21**	1.2E-01	**1.06**	3.5E-01	apolipoprotein
P05156	Complement factor I	32	**0.79**	1.2E-06	**0.86**	1.9E-01	52	**0.91**	5.9E-01	**0.95**	8.1E-01	serine protease
P20023	Complement receptor type 2			N.D.			1	**1.00**	9.6E-01	**0.85**	5.7E-01	apolipoprotein
P29279	Connective tissue growth factor	8	**0.66**	8.9E-04	**0.82**	2.0E-01	5	**0.74**	6.8E-01	**0.80**	7.8E-01	growth factor
O14618	Copper chaperone for superoxide dismutase			N.D.			1	**0.85**	-	**0.87**	-	cation transporter
Q9P0M6	Core histone macro-H2A.2	1	**2.30**	-	**1.45**	-			N.D.			histone
Q15517	Corneodesmosin	2	**0.95**	7.6E-01	**0.88**	4.9E-01	3	**0.78**	3.0E-01	**0.86**	8.9E-01	N.A.
P35321	Cornifin-A	24	**0.90**	2.3E-02	**1.11**	3.9E-02	30	**1.11**	1.1E-01	**1.02**	6.0E-01	structural protein
P22528	Cornifin-B	15	**0.84**	6.2E-01	**1.09**	8.3E-01	22	**1.24**	2.2E-03	**1.06**	3.9E-01	structural protein
Q9UBG3	Cornulin	15	**1.00**	9.7E-01	**1.07**	3.9E-01	25	**0.96**	8.5E-01	**0.82**	3.4E-01	signaling molecule
P31146	Coronin-1A	8	**3.72**	1.9E-04	**4.64**	1.9E-04	5	**1.21**	2.7E-01	**1.33**	5.6E-01	non-motor actin binding protein
Q9ULV4	Coronin-1C	1	**2.00**	-	**2.44**	-	1	**1.07**	7.1E-01	**1.04**	5.1E-01	non-motor actin binding protein
P08185	Corticosteroid-binding globulin	18	**0.97**	5.9E-01	**0.77**	1.0E-01	20	**0.63**	2.6E-01	**0.61**	7.2E-02	serine protease inhibitor
P12277	Creatine kinase B-type			N.D.			3	**1.28**	6.9E-01	**1.28**	8.3E-01	amino acid kinase
P01040	Cystatin-A	2	**0.93**	3.0E-01	**0.36**	1.5E-01	6	**0.26**	6.5E-02	**0.33**	1.6E-01	cysteine protease inhibitor
P04080	Cystatin-B	30	**0.88**	4.6E-02	**0.93**	3.7E-01	42	**0.98**	9.6E-01	**1.00**	9.9E-01	cysteine protease inhibitor
P01034	Cystatin-C	9	**0.80**	7.3E-03	**1.03**	9.1E-01	15	**0.72**	4.2E-01	**0.75**	4.8E-01	cysteine protease inhibitor
Q15828	Cystatin-M	6	**1.12**	1.3E-01	**1.00**	1.0E+00	12	**1.08**	6.0E-01	**1.06**	5.2E-01	cysteine protease inhibitor
P09228	Cystatin-SA	5	**1.13**	3.0E-01	**1.10**	2.1E-01	5	**0.39**	4.5E-03	**0.38**	1.2E-02	cysteine protease inhibitor
P21291	Cysteine and glycine-rich protein 1			N.D.			3	**1.31**	6.9E-01	**1.21**	7.8E-01	actin family cytoskeletal protein
P50238	Cysteine-rich protein 1	1	**1.11**	2.0E-01	**0.86**	1.3E-01	3	**1.19**	6.2E-01	**1.10**	8.3E-01	homeobox transcription factor
P54108	Cysteine-rich secretory protein 3	7	**1.07**	1.6E-01	**1.05**	7.7E-01	16	**0.99**	9.8E-01	**0.98**	9.6E-01	defense/immunity protein
P32320	Cytidine deaminase	2	**3.69**	7.4E-03	**3.14**	1.5E-01			N.D.			deaminase
Q13409	Cytoplasmic dynein 1 intermediate chain 2			N.D.			1	**0.81**	-	**0.57**	-	microtubule family cytoskeletal protein
Q96KP4	Cytosolic non-specific dipeptidase			N.D.			2	**0.85**	8.2E-01	**0.86**	7.5E-01	metalloprotease
P30046	D-dopachrome decarboxylase			N.D.			2	**0.95**	8.6E-01	**1.22**	4.6E-01	cytokine
P07585	Decorin	7	**0.39**	3.3E-06	**0.36**	7.0E-05	4	**1.12**	8.4E-01	**1.20**	7.4E-01	receptor
Q9UGM3	Deleted in malignant brain tumors 1 protein	24	**1.53**	8.8E-06	**1.39**	1.1E-02	14	**1.13**	3.5E-01	**1.23**	2.1E-01	receptor
P13716	Delta-aminolevulinic acid dehydratase	1	**1.31**	-	**1.02**	-			N.D.			dehydratase
Q6P3S1	DENN domain-containing protein 1B	2	**0.91**	-	**0.74**	-			N.D.			N.A.
Q6ZUT9	DENN domain-containing protein 5B	1	**1.74**	-	**0.05**	-			N.D.			ion channel
P24855	Deoxyribonuclease-1	1	**0.95**	-	**0.83**	-			N.D.			endodeoxyribonuclease
Q07507	Dermatopontin	3	**0.75**	4.2E-02	**0.66**	1.7E-01	1	**0.98**	9.4E-01	**1.05**	7.7E-01	extracellular matrix protein
P81605	Dermcidin	22	**0.24**	3.3E-12	**0.27**	5.0E-12	30	**0.60**	3.1E-01	**0.49**	1.2E-01	N.A.
Q08554	Desmocollin-1			N.D.			1	**0.88**	-	**0.87**	-	cell junction protein

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
Q02487	Desmocollin-2	7	**1.00**	9.9E-01	**1.00**	1.0E+00	9	**1.01**	9.5E-01	**1.00**	9.9E-01	cell junction protein
Q14574	Desmocollin-3	1	**0.78**	3.6E-01	**0.67**	3.5E-02	2	**0.77**	5.5E-01	**0.73**	2.5E-01	cell junction protein
Q02413	Desmoglein-1	2	**1.00**	9.9E-01	**1.29**	6.1E-01	2	**1.22**	7.6E-01	**1.42**	6.5E-01	cell junction protein
Q14126	Desmoglein-2	3	**0.89**	1.2E-01	**1.03**	8.3E-01	4	**0.96**	8.8E-01	**0.90**	3.3E-01	cell junction protein
P32926	Desmoglein-3	5	**0.99**	9.0E-01	**0.98**	9.0E-01	5	**1.40**	6.9E-01	**1.47**	6.7E-01	cell junction protein
P15924	Desmoplakin	2	**0.91**	6.2E-01	**0.94**	4.3E-01	6	**0.99**	9.3E-01	**1.00**	1.0E+00	intermediate filament binding protein
P60981	Destrin			N.D.			5	**0.92**	6.5E-01	**0.85**	2.0E-01	non-motor actin binding protein
Q01459	Di-N-acetylchitobiase			N.D.			5	**0.85**	3.5E-01	**0.81**	3.1E-01	glycosidase
O94907	Dickkopf-related protein 1	3	**0.97**	8.4E-01	**0.98**	8.4E-01	3	**0.78**	6.7E-02	**0.80**	2.0E-01	signaling molecule
Q9UBP4	Dickkopf-related protein 3	3	**0.92**	4.5E-01	**0.68**	1.6E-01	3	**0.81**	6.4E-01	**0.82**	3.2E-01	N.A.
Q16555	Dihydropyrimidinase-related protein 2	1	**1.34**	-	**1.38**	-			N.D.			hydrolase
P16444	Dipeptidase 1	1	**1.72**	-	**1.02**	-			N.D.			metalloprotease
P53634	Dipeptidyl peptidase 1	3	**1.15**	3.5E-01	**1.30**	2.6E-02	3	**0.96**	8.7E-01	**0.79**	3.1E-01	cysteine protease
Q9NY33	Dipeptidyl peptidase 3			N.D.			2	**0.96**	8.9E-01	**0.87**	7.1E-01	metalloprotease
P27487	Dipeptidyl peptidase 4			N.D.			1	**0.74**	1.7E-01	**0.76**	4.5E-02	serine protease
O14672	Disintegrin and metalloproteinase domain-containing protein 10	2	**0.74**	5.1E-01	**0.44**	1.6E-01	1	**1.17**	-	**1.02**	-	metalloprotease
O43184	Disintegrin and metalloproteinase domain-containing protein 12	1	**0.96**	-	**1.32**	-	7	**0.57**	1.8E-02	**0.60**	7.0E-02	metalloprotease
Q9P0K1	Disintegrin and metalloproteinase domain-containing protein 22	1	**0.94**	4.7E-01	**1.37**	8.3E-01			N.D.			metalloprotease
Q13443	Disintegrin and metalloproteinase domain-containing protein 9	2	**0.90**	4.1E-01	**1.28**	5.6E-02	1	**1.25**	9.0E-01	**1.17**	9.2E-01	metalloprotease
Q8TDJ6	DmX-like protein 2			N.D.			2	**1.92**	6.1E-04	**1.97**	5.6E-03	N.A.
P27695	DNA-(apurinic or apyrimidinic site) lyase			N.D.			3	**0.90**	8.6E-01	**1.00**	9.9E-01	endodeoxyribonuclease
O75937	DnaJ homolog subfamily C member 8	1	**1.53**	-	**1.39**	-			N.D.			chaperone
Q8TD84	Down syndrome cell adhesion molecule-like protein 1			N.D.			2	**1.01**	9.6E-01	**0.35**	5.5E-01	immunoglobulin receptor superfamily
Q96FJ2	Dynein light chain 2, cytoplasmic	1	**1.05**	-	**1.29**	-			N.D.			enzyme modulator
Q14118	Dystroglycan	8	**0.85**	2.9E-02	**0.78**	8.9E-02	8	**1.15**	7.2E-01	**0.92**	8.5E-01	receptor
Q13822	Ectonucleotide phosphodiesterase family member 2	13	**1.08**	2.1E-01	**1.00**	9.9E-01	15	**0.88**	8.2E-01	**0.88**	8.2E-01	nucleotide phosphatase
Q96C19	EF-hand domain-containing protein D2	2	**8.20**	-	**5.99**	-	5	**0.46**	3.0E-02	**0.56**	8.5E-02	calcium-binding protein
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	25	**0.87**	1.1E-02	**0.91**	3.2E-01	28	**0.67**	4.8E-02	**0.71**	1.4E-01	membrane-bound signaling molecule
O95967	EGF-containing fibulin-like extracellular matrix protein 2	5	**0.90**	5.2E-01	**0.75**	9.9E-02	1	**1.05**	6.8E-01	**1.07**	3.2E-01	membrane-bound signaling molecule
P19957	Elafin	3	**1.47**	5.8E-03	**1.73**	3.5E-01	4	**2.17**	3.7E-01	**1.80**	5.6E-01	serine protease inhibitor
P26641	Elongation factor 1-gamma			N.D.			1	**1.44**	2.2E-01	**1.71**	2.6E-01	anion channel
P13639	Elongation factor 2	1	**1.22**	5.9E-02	**1.12**	6.9E-01	6	**1.12**	5.7E-01	**1.12**	5.2E-01	nucleotidyltransferase
Q9BXX0	EMILIN-2			N.D.			3	**1.09**	4.6E-01	**1.18**	2.2E-01	extracellular matrix glycoprotein
P17813	Endoglin	2	**0.72**	1.3E-02	**0.83**	4.1E-01	5	**0.80**	8.9E-02	**0.88**	3.2E-01	TGF-beta receptor
O94919	Endonuclease domain-containing 1 protein	3	**0.84**	4.3E-01	**0.76**	2.0E-01	3	**1.78**	3.4E-01	**1.65**	4.6E-01	N.A.
P30040	Endoplasmic reticulum resident protein 29			N.D.			1	**1.62**	4.8E-01	**1.71**	6.9E-01	membrane traffic protein
Q9BS26	Endoplasmic reticulum resident protein 44	1	**1.11**	-	**0.87**	-	2	**1.17**	5.5E-02	**1.31**	7.8E-03	isomerase
P14625	Endoplasmin	4	**1.23**	6.3E-01	**0.85**	4.6E-01	4	**0.99**	9.9E-01	**0.95**	9.3E-01	Hsp90 family chaperone
Q9HCU0	Endosialin	11	**1.06**	5.7E-01	**1.36**	4.5E-03	10	**1.09**	6.4E-01	**1.22**	3.3E-01	transcription factor
Q96AP7	Endothelial cell-selective adhesion molecule	3	**1.13**	2.7E-02	**1.41**	2.4E-01	2	**1.22**	4.8E-01	**1.19**	1.6E-01	receptor
Q9UNN8	Endothelial protein C receptor	3	**0.94**	8.8E-02	**0.85**	2.9E-01	3	**0.80**	8.0E-01	**0.85**	6.8E-01	receptor
Q92556	Engulfment and cell motility protein 1	1	**0.97**	-	**1.06**	-			N.D.			signaling molecule
Q92817	Envoplakin			N.D.			2	**0.81**	7.4E-01	**0.71**	7.4E-01	intermediate filament binding protein
Q05315	Eosinophil lysophospholipase	1	**0.94**	-	**0.92**	-			N.D.			signaling molecule
P21709	Ephrin type-A receptor 1	1	**1.32**	-	**1.04**	-			N.D.			protein kinase
P29323	Ephrin type-B receptor 2	1	**0.95**	-	**0.97**	-	1	**1.26**	5.7E-01	**1.43**	4.1E-01	protein kinase
P54760	Ephrin type-B receptor 4	2	**0.84**	3.8E-01	**0.86**	2.0E-01	4	**1.17**	2.9E-01	**0.97**	8.5E-01	protein kinase
P20827	Ephrin-A1	1	**0.97**	5.8E-01	**1.00**	9.8E-01	3	**1.20**	8.6E-01	**1.30**	7.4E-01	membrane-bound signaling molecule
P98172	Ephrin-B1			N.D.			3	**1.20**	3.3E-01	**1.10**	3.8E-01	membrane-bound signaling molecule
Q9H6S3	Epidermal growth factor receptor kinase substrate 8-like protein 2	3	**1.34**	2.6E-01	**1.54**	2.6E-01	2	**1.02**	9.2E-01	**1.32**	1.9E-01	transmembrane receptor
P61916	Epididymal secretory protein E1	6	**1.25**	1.3E-01	**1.28**	1.6E-01	9	**0.79**	7.6E-01	**0.81**	7.9E-01	N.A.
Q99645	Epiphycan	3	**0.87**	1.2E-01	**0.80**	1.4E-01	2	**0.87**	4.2E-01	**1.02**	9.5E-01	receptor
Q96HE7	ERO1-like protein alpha	4	**1.34**	3.0E-03	**0.92**	7.3E-02	2	**0.96**	8.5E-01	**1.00**	1.0E+00	oxidoreductase
P27105	Erythrocyte band 7 integral membrane protein	1	**1.14**	6.6E-01	**1.05**	3.5E-01			N.D.			cytoskeletal protein
Q6IS14	Eukaryotic translation initiation factor 5A-1-like			N.D.			3	**1.28**	3.4E-02	**1.44**	8.4E-03	translation factor
Q9GZZ8	Extracellular glycoprotein lacritin	1	**1.05**	8.1E-01	**0.92**	9.9E-02			N.D.			N.A.
Q16610	Extracellular matrix protein 1	18	**0.96**	5.8E-01	**1.00**	9.7E-01	24	**1.11**	6.8E-01	**1.03**	9.1E-01	N.A.
Q8IWU5	Extracellular sulfatase Sulf-2	1	**0.50**	-	**0.52**	-	1	**1.14**	-	**1.16**	-	hydrolase
P08294	Extracellular superoxide dismutase [Cu-Zn]	5	**0.82**	1.7E-01	**0.95**	7.2E-01	7	**1.23**	1.9E-01	**1.22**	2.3E-01	oxidoreductase
P15311	Ezrin			N.D.			20	**0.94**	8.8E-01	**0.94**	9.0E-01	actin family cytoskeletal protein
P52907	F-actin-capping protein subunit alpha-1			N.D.			2	**1.09**	-	**1.42**	-	non-motor actin binding protein
P47756	F-actin-capping protein subunit beta	2	**3.56**	-	**2.04**	-	3	**1.05**	7.9E-01	**0.98**	9.4E-01	non-motor actin binding protein
Q16658	Fascin			N.D.			2	**1.65**	2.7E-01	**1.55**	3.2E-01	non-motor actin binding protein
Q01469	Fatty acid-binding protein, epidermal	13	**1.05**	1.8E-01	**0.99**	9.6E-01	27	**0.82**	6.9E-01	**0.79**	6.2E-01	transfer/carrier protein
P07148	Fatty acid-binding protein, liver			N.D.			1	**0.43**	-	**0.52**	-	transfer/carrier protein
Q5SYB0	FERM and PDZ domain-containing protein 1			N.D.			1	**0.95**	8.1E-01	**0.87**	3.8E-01	N.A.
P02794	Ferritin heavy chain	2	**3.21**	1.8E-02	**2.02**	7.5E-02			N.D.			storage protein
P02792	Ferritin light chain	2	**1.35**	9.4E-02	**0.85**	3.0E-01	1	**0.96**	-	**0.63**	-	storage protein
Q9UGM5	Fetuin-B	10	**0.76**	5.0E-04	**0.81**	6.4E-03	12	**0.83**	2.3E-01	**0.87**	4.2E-01	extracellular matrix glycoprotein
P35555	Fibrillin-1	15	**1.02**	7.3E-01	**1.10**	3.0E-01	8	**1.06**	6.0E-01	**1.03**	8.5E-01	signaling molecule
P35556	Fibrillin-2	3	**1.02**	8.6E-01	**1.14**	2.6E-01			N.D.			signaling molecule
P02671	Fibrinogen alpha chain	29	**1.04**	7.5E-01	**1.02**	9.2E-01	106	**0.82**	2.4E-01	**0.78**	1.3E-01	signaling molecule
P02675	Fibrinogen beta chain	7	**1.75**	6.3E-05	**1.77**	2.4E-03	63	**0.52**	2.2E-02	**0.50**	1.1E-02	signaling molecule
P02679	Fibrinogen gamma chain	3	**1.99**	5.4E-03	**2.30**	8.5E-03	31	**0.64**	3.3E-02	**0.65**	2.7E-02	signaling molecule
Q9BYJ0	Fibroblast growth factor-binding protein 2	1	**0.00**	-	**0.00**	-			N.D.			N.A.
Q14314	Fibroleukin	1	**0.77**	-	**1.31**	-	1	**1.61**	-	**1.31**	-	signaling molecule
Q06828	Fibromodulin	2	**0.90**	5.7E-01	**1.01**	9.3E-01	3	**0.94**	6.8E-01	**0.82**	5.5E-02	receptor
P02751	Fibronectin	305	**0.90**	2.3E-09	**0.76**	4.7E-13	355	**0.81**	1.5E-02	**0.85**	4.9E-02	signaling molecule
P23142	Fibulin-1	22	**0.95**	3.7E-01	**0.88**	3.7E-01	23	**1.26**	3.9E-01	**1.19**	5.4E-01	membrane-bound signaling molecule
P98095	Fibulin-2	9	**1.08**	1.2E-01	**1.13**	3.7E-01	3	**0.97**	9.7E-01	**0.90**	8.4E-01	membrane-bound signaling molecule
Q9UBX5	Fibulin-5	9	**0.80**	7.0E-02	**0.80**	1.6E-01	13	**0.86**	3.2E-01	**0.84**	3.3E-01	membrane-bound signaling molecule
P20930	Filaggrin	3	**0.62**	2.1E-01	**0.49**	2.9E-03	9	**0.70**	1.9E-02	**0.71**	2.5E-02	cytoskeletal protein
Q5D862	Filaggrin-2			N.D.			1	**1.05**	-	**0.95**	9.3E-01	cytoskeletal protein
P21333	Filamin-A	11	**2.21**	1.5E-04	**1.99**	2.9E-03	22	**1.00**	9.8E-01	**0.97**	8.7E-01	N.A.
O75369	Filamin-B	1	**1.29**	5.1E-01	**1.21**	5.6E-02	4	**1.44**	2.8E-02	**1.28**	3.2E-01	N.A.
P30043	Flavin reductase	2	**0.88**	6.0E-01	**0.73**	4.8E-01	7	**0.83**	5.1E-01	**1.00**	9.9E-01	reductase
P15328	Folate receptor alpha	4	**0.89**	2.1E-01	**0.70**	1.4E-01	3	**0.80**	7.5E-01	**0.86**	8.6E-01	transporter
P41439	Folate receptor gamma	3	**3.26**	3.0E-01	**4.20**	2.8E-01			N.D.			transporter
Q12841	Follistatin-related protein 1	14	**0.90**	4.0E-01	**1.12**	5.0E-01	22	**0.87**	6.5E-01	**0.84**	5.9E-01	enzyme modulator
O95633	Follistatin-related protein 3	8	**0.72**	4.3E-05	**0.67**	1.3E-02	9	**0.73**	2.3E-01	**0.71**	1.4E-01	enzyme modulator
Q5SZK8	FRAS1-related extracellular matrix protein 2	5	**0.98**	8.1E-01	**0.95**	4.9E-01	3	**1.00**	9.8E-01	**1.09**	7.0E-01	cation transporter
Q14332	Frizzled-2			N.D.			2	**0.94**	-	**0.98**	-	signaling molecule
Q9ULV1	Frizzled-4	1	**0.87**	-	**0.87**	-			N.D.			signaling molecule
P09467	Fructose-1,6-bisphosphatase 1			N.D.			2	**0.99**	9.8E-01	**0.83**	1.0E-01	carbohydrate phosphatase
P04075	Fructose-bisphosphate aldolase A	24	**2.11**	4.1E-12	**2.08**	5.4E-06	29	**1.17**	6.7E-01	**1.06**	8.6E-01	aldolase
P09972	Fructose-bisphosphate aldolase C			N.D.			7	**1.23**	4.0E-01	**1.25**	4.6E-01	aldolase
Q86SQ4	G-protein coupled receptor 126	1	**0.79**	-	**0.86**	-	1	**0.85**	-	**1.11**	-	G-protein coupled receptor
Q9NZH0	G-protein coupled receptor family C group 5 member B	1	**0.80**	-	**0.82**	-	1	**0.84**	1.7E-01	**0.76**	3.3E-02	G-protein coupled receptor
P09382	Galectin-1	6	**1.29**	2.6E-04	**1.20**	8.2E-02	9	**1.37**	4.6E-01	**1.18**	7.5E-01	signaling molecule
P17931	Galectin-3	7	**1.12**	1.7E-02	**1.37**	1.6E-02	8	**1.11**	5.7E-01	**1.10**	1.8E-01	signaling molecule
Q08380	Galectin-3-binding protein	17	**0.89**	3.8E-01	**1.00**	9.9E-01	24	**1.23**	3.5E-01	**1.33**	2.1E-01	receptor
P09104	Gamma-enolase	3	**1.14**	-	**1.09**	-			N.D.			lyase

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
Q92820	Gamma-glutamyl hydrolase			N.D.			1	**0.66**	7.2E-02	**0.57**	5.5E-02	cysteine protease
O75223	Gamma-glutamylcyclotransferase	2	**1.02**	8.7E-01	**0.93**	8.5E-03			N.D.			N.A.
P17900	Ganglioside GM2 activator	5	**1.15**	3.5E-02	**1.10**	6.4E-01	7	**0.80**	6.7E-01	**0.76**	6.7E-01	transfer/carrier protein
P07098	Gastric triacylglycerol lipase			N.D.			3	**0.73**	4.8E-01	**0.75**	4.3E-01	lipase
P20142	Gastricsin	1	**2.88**	-	**2.27**	-	3	**0.99**	9.4E-01	**0.84**	2.4E-01	aspartic protease
Q9NS71	Gastrokine-1	6	**1.05**	2.3E-01	**1.16**	2.3E-02	5	**0.87**	6.4E-01	**0.90**	6.9E-01	N.A.
P06396	Gelsolin	49	**1.03**	2.9E-01	**1.05**	7.3E-01	83	**0.98**	9.1E-01	**0.97**	8.9E-01	non-motor actin binding protein
O60234	Glia maturation factor gamma	2	**4.45**	2.3E-01	**5.31**	2.0E-01	2	**1.58**	1.2E-01	**2.04**	2.8E-02	signaling molecule
P07093	Glia-derived nexin	21	**0.57**	1.5E-06	**0.64**	7.1E-03	16	**1.32**	4.2E-01	**1.22**	5.7E-01	serine protease inhibitor
P11413	Glucose-6-phosphate 1-dehydrogenase	2	**3.09**	2.4E-02	**3.02**	1.8E-02	6	**1.54**	2.1E-01	**1.56**	1.8E-01	dehydrogenase
P06744	Glucose-6-phosphate isomerase	9	**4.20**	2.0E-06	**4.30**	3.1E-05	12	**0.79**	2.9E-01	**0.92**	7.2E-01	isomerase
P14314	Glucosidase 2 subunit beta	4	**1.35**	4.6E-02	**1.23**	4.7E-01	5	**1.14**	8.2E-01	**1.32**	6.4E-01	transferase
P15104	Glutamine synthetase	1	**0.87**	-	**0.86**	-			N.D.			ligase
Q07075	Glutamyl aminopeptidase			N.D.			2	**0.91**	6.6E-01	**0.76**	2.2E-01	metalloprotease
P35754	Glutaredoxin-1	3	**1.34**	3.2E-01	**1.51**	1.6E-01	6	**0.42**	2.4E-01	**0.44**	1.4E-01	oxidoreductase
P22352	Glutathione peroxidase 3	4	**0.81**	7.0E-02	**1.08**	8.7E-01	4	**0.69**	5.4E-01	**0.75**	6.3E-01	peroxidase
P78417	Glutathione S-transferase omega-1	5	**1.76**	1.5E-03	**1.42**	2.2E-01	5	**1.20**	1.9E-01	**1.22**	3.6E-01	anion channel
P09211	Glutathione S-transferase P	10	**1.55**	8.2E-05	**1.72**	1.8E-02	16	**1.34**	3.3E-01	**1.36**	3.4E-01	transferase
P48637	Glutathione synthetase	1	**1.49**	-	**1.74**	-	2	**1.15**	5.5E-01	**0.90**	4.9E-01	ligase
P04406	Glyceraldehyde-3-phosphate dehydrogenase	9	**2.16**	1.2E-05	**2.07**	9.3E-05	13	**1.16**	6.4E-01	**1.17**	7.2E-01	dehydrogenase
P09466	Glycodelin	14	**1.31**	1.7E-04	**1.37**	1.6E-06	22	**1.70**	1.1E-02	**1.72**	7.4E-03	transfer/carrier protein
P06737	Glycogen phosphorylase, liver form	8	**3.49**	1.3E-05	**3.55**	1.9E-04	3	**1.31**	2.3E-01	**1.37**	1.6E-01	phosphorylase
P46976	Glycogenin-1	1	**3.44**	6.5E-02	**4.05**	7.9E-03			N.D.			glycosyltransferase
P01215	Glycoprotein hormones alpha chain	3	**0.75**	6.9E-02	**0.64**	2.6E-02	3	**0.37**	2.2E-03	**0.37**	4.7E-04	peptide hormone
Q9HC38	Glyoxalase domain-containing protein 4	1	**1.65**	2.1E-01	**1.43**	1.1E-02	3	**0.81**	8.5E-01	**0.86**	9.0E-01	lyase
P35052	Glypican-1			N.D.			2	**1.33**	7.5E-01	**1.45**	7.4E-01	extracellular matrix glycoprotein
O75487	Glypican-4			N.D.			2	**0.80**	6.9E-01	**0.92**	8.7E-01	extracellular matrix glycoprotein
Q8NBJ4	Golgi membrane protein 1	4	**0.90**	3.3E-01	**0.86**	4.4E-01	7	**0.66**	4.5E-01	**0.83**	6.0E-01	N.A.
Q9H4G4	Golgi-associated plant pathogenesis-related protein 1	1	**1.62**	-	**1.35**	-			N.D.			defense/immunity protein
P28799	Granulins	4	**1.48**	9.6E-03	**1.67**	2.8E-02	3	**1.02**	8.7E-01	**0.94**	4.8E-01	cytokine
P54826	Growth arrest-specific protein 1	1	**1.18**	2.0E-01	**1.02**	5.3E-01	1	**1.07**	7.8E-01	**1.09**	7.9E-01	N.A.
Q14393	Growth arrest-specific protein 6	3	**0.72**	3.4E-02	**0.65**	1.2E-01			N.D.			signaling molecule
P62993	Growth factor receptor-bound protein 2	1	**1.46**	-	**1.46**	-	1	**1.44**	2.7E-02	**1.44**	9.3E-02	transmembrane receptor
P08107	Heat shock 70 kDa protein 1A/1B	8	**2.15**	2.8E-06	**1.79**	3.5E-05	25	**1.17**	7.6E-01	**1.03**	9.5E-01	N.A.
P34932	Heat shock 70 kDa protein 4			N.D.			2	**1.27**	4.3E-01	**1.46**	3.7E-01	Hsp70 family chaperone
P11142	Heat shock cognate 71 kDa protein	6	**2.15**	2.7E-02	**2.29**	6.8E-02	20	**1.15**	8.3E-01	**1.06**	9.1E-01	Hsp70 family chaperone
P04792	Heat shock protein beta-1	5	**0.95**	1.5E-01	**0.89**	2.1E-01	12	**0.92**	8.4E-01	**1.00**	9.9E-01	structural protein
O14558	Heat shock protein beta-6			N.D.			1	**1.18**	-	**1.04**	-	structural protein
P07900	Heat shock protein HSP 90-alpha	10	**2.23**	3.0E-04	**2.50**	1.5E-04	7	**0.98**	9.4E-01	**0.99**	9.7E-01	Hsp90 family chaperone
P08238	Heat shock protein HSP 90-beta	5	**0.00**	-	**0.00**	-	6	**0.78**	7.5E-01	**0.96**	6.8E-01	Hsp90 family chaperone
P14317	Hematopoietic lineage cell-specific protein	1	**3.20**	-	**5.91**	-	1	**1.04**	7.8E-01	**0.91**	1.7E-01	basic helix-loop-helix transcription factor
Q9Y5Z4	Heme-binding protein 2	2	**1.22**	4.6E-01	**1.45**	7.5E-02	7	**1.00**	9.9E-01	**1.04**	7.2E-01	N.A.
Q96RW7	Hemicentin-1	4	**0.95**	5.0E-01	**1.11**	7.4E-01	2	**1.19**	7.8E-01	**1.00**	1.0E+00	N.A.
Q8NDA2	Hemicentin-2	1	**0.89**	-	**0.78**	-			N.D.			N.A.
P69905	Hemoglobin subunit alpha	51	**0.75**	6.7E-06	**0.57**	5.5E-07	199	**1.30**	6.2E-03	**1.22**	2.4E-02	transfer/carrier protein
P68871	Hemoglobin subunit beta	59	**0.82**	2.8E-03	**0.52**	2.0E-06	262	**0.93**	4.7E-01	**0.87**	1.6E-01	N.A.
P02042	Hemoglobin subunit delta	28	**1.15**	6.0E-01	**0.68**	7.6E-01	127	**1.04**	8.7E-01	**0.92**	8.0E-01	transfer/carrier protein
P69892	Hemoglobin subunit gamma-2	9	**0.38**	5.6E-02	**0.28**	1.4E-02	23	**1.68**	3.3E-01	**1.83**	2.0E-01	transfer/carrier protein
P02008	Hemoglobin subunit zeta	1	**1.80**	8.4E-02	**0.57**	1.4E-01	17	**1.00**	1.0E+00	**1.66**	1.2E-01	transfer/carrier protein
Q6ZVN8	Hemojuvelin	1	**0.82**	1.3E-01	**0.79**	2.5E-02			N.D.			N.A.
P02790	Hemopexin	159	**0.77**	2.3E-24	**0.77**	4.0E-05	266	**0.84**	1.9E-01	**0.88**	3.3E-01	transfer/carrier protein
P05546	Heparin cofactor 2	22	**0.86**	2.0E-02	**0.69**	9.6E-03	21	**0.84**	5.6E-01	**0.88**	6.7E-01	serine protease inhibitor
Q8TDQ0	Hepatitis A virus cellular receptor 2	1	**0.76**	8.2E-02	**1.07**	7.5E-01			N.D.			immunoglobulin receptor superfamily
P26927	Hepatocyte growth factor-like protein	7	**0.84**	1.4E-01	**0.90**	3.6E-01	16	**0.85**	2.6E-01	**0.82**	2.5E-01	growth factor
P51858	Hepatoma-derived growth factor	3	**2.13**	9.5E-02	**1.47**	2.7E-01	1	**2.39**	-	**2.77**	-	transcription cofactor
Q14103	Heterogeneous nuclear ribonucleoprotein D0	2	**2.74**	2.5E-01	**3.38**	3.0E-01			N.D.			replication origin binding protein
P52597	Heterogeneous nuclear ribonucleoprotein F			N.D.			2	**1.42**	1.5E-01	**1.36**	4.5E-01	ribosomal protein
P61978	Heterogeneous nuclear ribonucleoprotein K			N.D.			3	**0.84**	7.6E-01	**0.60**	2.6E-01	mRNA processing factor
O60506	Heterogeneous nuclear ribonucleoprotein Q	1	**1.02**	9.4E-01	**0.90**	7.9E-01			N.D.			mRNA processing factor
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	4	**3.47**	1.8E-01	**4.19**	1.6E-01	5	**1.27**	7.5E-01	**1.27**	7.8E-01	replication origin binding protein
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2			N.D.			1	**0.98**	6.7E-02	**0.84**	4.2E-01	mRNA processing factor
P52790	Hexokinase-3	4	**3.05**	4.0E-03	**3.38**	8.4E-03			N.D.			carbohydrate kinase
P09429	High mobility group protein B1			N.D.			3	**1.02**	9.7E-01	**0.98**	9.2E-01	HMG box transcription factor
P04196	Histidine-rich glycoprotein	26	**0.76**	6.4E-08	**0.83**	2.4E-01	36	**0.92**	4.1E-01	**0.82**	4.4E-02	N.A.
P16403	Histone H1.2	2	**4.49**	9.0E-02	**4.18**	2.0E-02	2	**1.18**	-	**1.00**	-	histone
P16401	Histone H1.5			N.D.			2	**1.15**	7.9E-01	**1.18**	1.4E-01	histone
Q9BTM1	Histone H2A.J	2	**6.49**	6.7E-02	**4.18**	2.2E-01			N.D.			histone
Q71UI9	Histone H2A.V			N.D.			1	**2.84**	-	**3.51**	-	histone
Q99880	Histone H2B type 1-L	13	**7.16**	2.1E-06	**6.88**	2.2E-05	3	**4.12**	2.9E-01	**2.32**	1.5E-01	histone
Q16695	Histone H3.1t			N.D.			5	**2.53**	-	**4.34**	-	histone
Q71DI3	Histone H3.2	5	**8.17**	1.8E-05	**8.26**	2.2E-06			N.D.			N.A.
P62805	Histone H4	12	**8.43**	2.6E-09	**9.03**	1.2E-05	10	**2.10**	2.7E-02	**1.98**	1.8E-01	N.A.
P16188	HLA class I histocompatibility antigen, A-30 alpha chain	3	**0.89**	-	**0.54**	-			N.D.			N.A.
P30501	HLA class I histocompatibility antigen, Cw-2 alpha chain			N.D.			1	**3.78**	-	**2.44**	-	N.A.
P04233	HLA class II histocompatibility antigen gamma chain	1	**1.08**	-	**0.92**	-			N.D.			major histocompatibility complex
Q9BPY8	Homeodomain-only protein			N.D.			2	**1.00**	-	**1.01**	9.7E-01	N.A.
Q14520	Hyaluronan-binding protein 2	8	**0.88**	1.1E-03	**0.66**	1.0E-03	11	**0.74**	4.1E-01	**0.67**	2.9E-01	growth factor
P00492	Hypoxanthine-guanine phosphoribosyltransferase	1	**2.03**	-	**2.36**	-			N.D.			glycosyltransferase
Q9Y4L1	Hypoxia up-regulated protein 1	1	**1.06**	-	**0.50**	-			N.D.			Hsp70 family chaperone
P01857	Ig gamma-1 chain C region			N.D.			3	**1.30**	5.3E-01	**1.43**	5.1E-01	immunoglobulin
P01860	Ig gamma-3 chain C region	2	**0.93**	6.8E-01	**1.19**	1.5E-01	3	**1.30**	2.1E-01	**1.22**	1.4E-01	immunoglobulin
P01834	Ig kappa chain C region	1	**2.17**	-	**1.78**	-			N.D.			immunoglobulin
P0CG06	Ig lambda-3 chain C regions	1	**1.77**	1.2E-02	**2.03**	1.3E-01			N.D.			N.A.
P01871	Ig mu chain C region			N.D.			3	**1.22**	1.0E-01	**1.51**	4.3E-03	immunoglobulin
Q9Y6R7	IgGFc-binding protein	34	**2.31**	1.8E-16	**2.28**	1.6E-13	36	**1.09**	4.3E-01	**1.16**	2.4E-01	signaling molecule
O14498	Immunoglobulin superfamily containing leucine-rich repeat protein	8	**0.96**	5.2E-01	**0.80**	7.8E-02	6	**0.89**	5.7E-01	**0.77**	2.9E-01	receptor
Q8N436	Inactive carboxypeptidase-like protein X2	1	**0.94**	-	**1.11**	-			N.D.			metalloprotease
P05019	Insulin-like growth factor I			N.D.			2	**1.24**	5.9E-01	**1.21**	2.4E-01	growth factor
P01344	Insulin-like growth factor II	4	**1.04**	9.2E-01	**0.96**	9.7E-01	6	**1.29**	7.3E-01	**1.11**	9.1E-01	growth factor
P08833	Insulin-like growth factor-binding protein 1	111	**0.83**	6.4E-09	**0.99**	7.7E-01	143	**0.85**	9.3E-02	**0.87**	6.4E-02	N.A.
P18065	Insulin-like growth factor-binding protein 2	24	**0.89**	5.1E-02	**0.83**	1.5E-02	37	**0.98**	8.4E-01	**0.92**	5.1E-01	N.A.
P17936	Insulin-like growth factor-binding protein 3	31	**0.89**	1.8E-02	**0.82**	1.6E-02	31	**0.88**	4.9E-01	**0.91**	6.3E-01	N.A.
P22692	Insulin-like growth factor-binding protein 4	26	**0.82**	2.3E-03	**0.83**	8.0E-02	34	**1.06**	7.1E-01	**1.05**	7.8E-01	N.A.
P24593	Insulin-like growth factor-binding protein 5	10	**1.07**	4.5E-01	**1.09**	5.1E-01	18	**1.21**	3.9E-01	**1.18**	2.5E-01	N.A.
P24592	Insulin-like growth factor-binding protein 6	10	**0.95**	6.1E-01	**0.88**	1.8E-02	8	**0.90**	6.6E-01	**0.90**	6.6E-01	N.A.
Q16270	Insulin-like growth factor-binding protein 7	6	**0.84**	6.1E-03	**0.84**	5.8E-02	10	**0.81**	6.2E-01	**0.70**	4.5E-01	N.A.
P35858	Insulin-like growth factor-binding protein complex acid labile subunit	21	**0.80**	3.9E-04	**0.91**	3.8E-01	21	**0.83**	3.0E-01	**0.86**	4.6E-01	receptor
Q8WX77	Insulin-like growth factor-binding protein-like 1	1	**0.61**	-	**0.80**	-			N.D.			N.A.
Q9Y287	Integral membrane protein 2B	3	**0.75**	1.4E-01	**0.72**	4.7E-01	3	**1.08**	9.4E-01	**1.06**	9.4E-01	N.A.
P98153	Integral membrane protein DGCR2/IDD			N.D.			1	**0.92**	-	**0.87**	-	N.A.
P11215	Integrin alpha-M	1	**7.55**	-	**11.43**	-			N.D.			cell adhesion molecule
P05107	Integrin beta-2	1	**3.12**	-	**2.48**	-	1	**1.06**	8.1E-01	**1.13**	3.9E-01	receptor

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P19827	Inter-alpha-trypsin inhibitor heavy chain H1	22	**0.76**	4.3E-04	**0.68**	2.3E-03	44	**1.10**	7.7E-01	**0.97**	9.1E-01	serine protease inhibitor
P19823	Inter-alpha-trypsin inhibitor heavy chain H2	26	**0.77**	7.0E-05	**0.81**	2.4E-04	62	**1.13**	3.2E-01	**1.15**	3.0E-01	serine protease inhibitor
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	5	**0.99**	8.7E-01	**0.86**	1.1E-01	18	**1.06**	7.6E-01	**0.97**	8.7E-01	serine protease inhibitor
Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	58	**0.85**	3.6E-04	**0.88**	8.1E-02	78	**0.89**	4.4E-01	**0.94**	6.9E-01	serine protease inhibitor
Q86UX2	Inter-alpha-trypsin inhibitor heavy chain H5	2	**0.72**	-	**0.65**	-	2	**1.03**	9.4E-01	**1.09**	6.7E-01	serine protease inhibitor
P05362	Intercellular adhesion molecule 1	12	**0.83**	6.7E-06	**0.85**	1.7E-01	18	**0.96**	8.9E-01	**1.00**	9.9E-01	signaling molecule
P13598	Intercellular adhesion molecule 2			N.D.			2	**1.42**	4.4E-01	**1.40**	4.3E-01	signaling molecule
P32942	Intercellular adhesion molecule 3	1	**2.20**	-	**1.82**	-			N.D.			signaling molecule
Q9NPH3	Interleukin-1 receptor accessory protein	6	**0.89**	2.4E-01	**0.77**	1.9E-01	2	**0.94**	9.4E-01	**0.81**	7.6E-01	type I cytokine receptor
P18510	Interleukin-1 receptor antagonist protein	6	**1.09**	1.3E-01	**0.94**	8.4E-01	10	**1.06**	7.1E-01	**1.02**	9.3E-01	interleukin superfamily
P27930	Interleukin-1 receptor type 2	6	**0.97**	7.4E-01	**1.05**	7.2E-01	2	**0.97**	9.5E-01	**1.20**	-	type I cytokine receptor
Q01638	Interleukin-1 receptor-like 1	14	**0.72**	1.4E-05	**0.73**	1.1E-02	11	**0.85**	6.3E-01	**0.77**	4.6E-01	type I cytokine receptor
O95998	Interleukin-18-binding protein	1	**0.86**	-	**1.04**	-			N.D.			N.A.
P40189	Interleukin-6 receptor subunit beta	1	**0.97**	-	**1.20**	-	2	**1.22**	4.9E-01	**1.45**	2.7E-01	signaling molecule
P10145	Interleukin-8	1	**12.27**	-	**20.85**	-	1	**5.23**	-	**3.06**	-	chemokine
P07476	Involucrin	21	**0.97**	8.0E-02	**0.98**	7.9E-01	32	**0.96**	8.4E-01	**1.01**	9.6E-01	extracellular matrix protein
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic	2	**1.39**	2.9E-01	**1.94**	1.5E-02	6	**0.93**	8.4E-01	**0.96**	9.3E-01	dehydrogenase
P53990	IST1 homolog	1	**1.05**	-	**1.06**	-			N.D.			N.A.
Q6H9L7	Isthmin-2	10	**0.73**	6.9E-03	**0.72**	4.4E-03	12	**0.95**	8.8E-01	**1.07**	8.3E-01	membrane-bound signaling molecule
Q9Y624	Junctional adhesion molecule A	4	**0.78**	1.2E-02	**0.79**	6.5E-02	6	**0.99**	9.4E-01	**0.95**	6.2E-01	receptor
Q9BX67	Junctional adhesion molecule C	1	**0.81**	-	**0.59**	-			N.D.			receptor
O43240	Kallikrein-10	1	**1.02**	-	**1.06**	-			N.D.			serine protease
Q9UBX7	Kallikrein-11	1	**0.89**	3.2E-01	**0.49**	3.6E-01	3	**1.34**	8.4E-01	**1.28**	9.0E-01	serine protease
Q9UKR3	Kallikrein-13	1	**0.74**	-	**0.84**	-	1	**0.97**	7.7E-01	**1.05**	3.0E-01	serine protease
Q92876	Kallikrein-6			N.D.			2	**1.32**	2.8E-01	**1.45**	3.0E-01	serine protease
P29622	Kallistatin	14	**0.85**	7.4E-03	**0.72**	3.0E-03	21	**0.57**	1.6E-01	**0.59**	2.3E-01	serine protease inhibitor
P13645	Keratin, type I cytoskeletal 10	4	**0.73**	2.0E-01	**0.72**	1.3E-02	12	**0.32**	3.4E-02	**2.38**	2.2E-02	structural protein
P13646	Keratin, type I cytoskeletal 13	15	**1.24**	8.7E-03	**1.23**	2.1E-01	41	**1.84**	2.5E-02	**1.83**	5.6E-02	structural protein
P02533	Keratin, type I cytoskeletal 14			N.D.			5	**1.37**	5.8E-02	**1.60**	2.1E-02	structural protein
Q04695	Keratin, type I cytoskeletal 17	8	**1.94**	2.7E-01	**7.37**	4.0E-01	4	**0.70**	-	**0.65**	-	structural protein
P05783	Keratin, type I cytoskeletal 18			N.D.			6	**1.02**	8.8E-01	**1.01**	9.7E-01	structural protein
P08727	Keratin, type I cytoskeletal 19	10	**1.27**	1.4E-01	**1.19**	4.1E-01	12	**1.42**	4.5E-01	**1.33**	4.9E-01	structural protein
P35527	Keratin, type I cytoskeletal 9	1	**0.97**	-	**1.23**	-	1	**0.89**	1.8E-01	**9.04**	6.9E-02	structural protein
P04264	Keratin, type II cytoskeletal 1	4	**0.91**	7.5E-01	**1.08**	6.7E-01	19	**0.46**	5.7E-03	**4.09**	6.2E-04	structural protein
P35908	Keratin, type II cytoskeletal 2 epidermal			N.D.			9	**0.27**	1.0E-01	**3.89**	2.8E-03	structural protein
P19013	Keratin, type II cytoskeletal 4	10	**1.02**	8.5E-01	**1.04**	7.0E-01	24	**1.59**	3.5E-01	**1.56**	4.7E-01	structural protein
P13647	Keratin, type II cytoskeletal 5	13	**1.71**	2.5E-01	**1.17**	8.0E-01	31	**0.93**	7.7E-01	**1.11**	7.0E-01	structural protein
P02538	Keratin, type II cytoskeletal 6A	16	**1.80**	4.9E-02	**1.69**	5.8E-02	44	**1.54**	4.1E-02	**1.49**	6.4E-02	structural protein
P04259	Keratin, type II cytoskeletal 6B			N.D.			41	**0.83**	-	**1.06**	-	structural protein
P48668	Keratin, type II cytoskeletal 6C	14	**2.98**	-	**2.22**	-			N.D.			structural protein
P08729	Keratin, type II cytoskeletal 7	5	**1.48**	4.2E-02	**1.54**	5.3E-02	8	**1.53**	5.2E-01	**1.27**	7.4E-01	structural protein
Q8N1N4	Keratin, type II cytoskeletal 78			N.D.			5	**1.23**	4.2E-02	**1.21**	1.4E-01	structural protein
P05787	Keratin, type II cytoskeletal 8	10	**1.42**	7.2E-03	**1.52**	5.8E-03	18	**1.00**	1.0E+00	**0.98**	9.7E-01	structural protein
P01042	Kininogen-1	76	**0.91**	1.1E-02	**1.09**	3.9E-01	135	**1.07**	6.8E-01	**1.08**	6.2E-01	cysteine protease inhibitor
O43278	Kunitz-type protease inhibitor 1	5	**0.88**	4.3E-01	**0.94**	4.9E-01	4	**0.86**	6.5E-01	**0.88**	7.2E-01	serine protease inhibitor
P00338	L-lactate dehydrogenase A chain	13	**2.16**	7.7E-08	**1.78**	8.0E-05	11	**0.89**	6.7E-01	**0.86**	6.6E-01	dehydrogenase
P07195	L-lactate dehydrogenase B chain	3	**1.75**	4.3E-02	**1.08**	7.2E-01	6	**1.05**	9.3E-01	**1.07**	9.0E-01	dehydrogenase
P14151	L-selectin	3	**1.36**	1.2E-01	**1.50**	5.7E-03	3	**0.97**	7.4E-01	**0.88**	2.5E-01	apolipoprotein
P22079	Lactoperoxidase	8	**0.84**	6.5E-03	**0.89**	6.0E-02	9	**0.85**	7.1E-02	**0.89**	1.2E-01	peroxidase
P02788	Lactotransferrin	62	**2.46**	9.4E-38	**2.28**	3.4E-21	60	**0.76**	4.4E-02	**0.75**	5.7E-02	transfer/carrier protein
Q04760	Lactoylglutathione lyase	1	**1.42**	-	**0.99**	-			N.D.			lyase
P02545	Lamin-A/C	2	**1.31**	1.0E-01	**1.25**	1.0E-01	7	**1.20**	6.4E-01	**1.30**	4.5E-01	structural protein
P20700	Lamin-B1	2	**3.14**	-	**3.14**	-			N.D.			structural protein
Q16787	Laminin subunit alpha-3	1	**0.84**	6.3E-01	**1.03**	9.4E-01	3	**1.05**	8.7E-01	**0.94**	7.1E-01	N.A.
O15230	Laminin subunit alpha-5	1	**0.95**	3.7E-01	**1.25**	2.8E-01	1	**1.33**	1.6E-01	**0.86**	2.1E-01	N.A.
P07942	Laminin subunit beta-1	1	**1.11**	-	**1.19**	-	3	**1.22**	7.7E-01	**1.09**	8.0E-01	N.A.
Q13751	Laminin subunit beta-3	1	**1.05**	-	**0.70**	-			N.D.			N.A.
P11047	Laminin subunit gamma-1	3	**1.34**	8.1E-03	**1.24**	6.9E-02	3	**1.04**	7.4E-01	**1.35**	6.2E-03	N.A.
Q14766	Latent-transforming growth factor beta-binding protein 1	11	**0.94**	4.4E-01	**1.07**	5.7E-01	18	**0.95**	6.0E-01	**1.01**	9.2E-01	signaling molecule
Q14767	Latent-transforming growth factor beta-binding protein 2	5	**0.86**	2.3E-01	**1.18**	5.3E-01	5	**1.31**	1.4E-01	**1.25**	2.9E-01	signaling molecule
Q8N2S1	Latent-transforming growth factor beta-binding protein 4	8	**0.88**	8.1E-02	**0.80**	4.6E-01	6	**0.97**	8.5E-01	**0.95**	6.1E-01	signaling molecule
Q9BS40	Latexin			N.D.			1	**0.99**	-	**0.94**	-	N.A.
O94910	Latrophilin-1	2	**0.80**	2.7E-01	**0.52**	4.0E-01	1	**1.04**	9.9E-01	**1.06**	9.8E-01	G-protein coupled receptor
Q6UX15	Layilin			N.D.			1	**1.26**	-	**1.22**	-	N.A.
O00292	Left-right determination factor 2	1	**0.34**	-	**0.37**	-			N.D.			growth factor
P02750	Leucine-rich alpha-2-glycoprotein	19	**1.00**	9.2E-01	**1.00**	9.7E-01	20	**1.20**	6.2E-01	**1.12**	8.0E-01	receptor
P30740	Leukocyte elastase inhibitor	11	**2.18**	6.8E-06	**1.69**	1.1E-03	15	**1.38**	9.3E-02	**1.44**	9.1E-02	serine protease inhibitor
Q8N6C8	Leukocyte immunoglobulin-like receptor subfamily A member 3	4	**2.77**	2.7E-03	**3.14**	5.1E-05			N.D.			membrane-bound signaling molecule
Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1	3	**1.02**	8.8E-01	**0.97**	7.2E-01	3	**1.00**	9.4E-01	**1.00**	1.0E+00	immunoglobulin receptor superfamily
P09960	Leukotriene A-4 hydrolase	8	**3.70**	1.0E-05	**4.09**	2.2E-04	6	**1.87**	4.2E-01	**1.91**	5.3E-01	metalloprotease
Q14847	LIM and SH3 domain protein 1	3	**1.64**	3.8E-01	**1.98**	2.7E-01	2	**1.49**	1.8E-01	**1.39**	2.2E-01	non-motor actin binding protein
Q13449	Limbic system-associated membrane protein	1	**0.88**	-	**0.69**	-			N.D.			immunoglobulin cell adhesion molecule
P31025	Lipocalin-1	11	**0.81**	1.8E-01	**0.53**	2.1E-03	20	**0.33**	1.7E-03	**0.37**	1.7E-03	transfer/carrier protein
Q6UWW0	Lipocalin-15	3	**0.88**	3.3E-01	**0.88**	6.0E-01	3	**0.92**	8.7E-01	**0.83**	5.4E-01	transfer/carrier protein
Q86X29	Lipolysis-stimulated lipoprotein receptor	1	**0.89**	-	**0.80**	-	3	**1.00**	1.0E+00	**0.90**	6.2E-01	N.A.
P18428	Lipopolysaccharide-binding protein	7	**1.03**	5.3E-01	**0.97**	7.7E-01	5	**0.88**	8.5E-01	**1.00**	1.0E+00	transfer/carrier protein
P48304	Lithostathine-1-beta	1	**1.03**	-	**0.74**	-			N.D.			growth factor
Q8TDL5	Long palate, lung and nasal epithelium carcinoma-associated protein 1	2	**1.30**	6.6E-02	**1.16**	6.7E-02	5	**1.09**	5.4E-01	**1.24**	7.6E-01	carbohydrate transporter
P59827	Long palate, lung and nasal epithelium carcinoma-associated protein 4	2	**0.79**	2.1E-01	**0.61**	1.2E-01	1	**0.71**	4.4E-02	**0.60**	1.8E-01	carbohydrate transporter
P08637	Low affinity immunoglobulin gamma Fc region receptor III-A	2	**3.38**	1.3E-02	**3.61**	2.5E-04	3	**1.74**	1.7E-01	**1.57**	2.0E-01	immunoglobulin receptor superfamily
P01130	Low-density lipoprotein receptor	1	**0.70**	-	**0.75**	-	2	**0.92**	3.3E-01	**1.02**	9.2E-01	receptor
P98164	Low-density lipoprotein receptor-related protein 2	1	**0.68**	-	**0.85**	-			N.D.			receptor
P51884	Lumican	91	**0.95**	2.4E-02	**1.05**	7.3E-02	85	**0.98**	7.9E-01	**0.98**	7.9E-01	receptor
O95274	Ly6/PLAUR domain-containing protein 3	6	**0.80**	2.4E-02	**0.95**	9.1E-01	7	**0.72**	6.2E-01	**0.65**	5.5E-01	N.A.
Q9Y5Y7	Lymphatic vessel endothelial hyaluronic acid receptor 1			N.D.			2	**1.62**	7.8E-01	**1.19**	9.4E-01	N.A.
Q14210	Lymphocyte antigen 6D	1	**0.77**	4.6E-01	**0.87**	2.4E-01			N.D.			N.A.
P33241	Lymphocyte-specific protein 1	2	**3.38**	8.0E-02	**2.97**	2.0E-01	3	**0.90**	6.9E-01	**0.88**	8.4E-01	signaling molecule
Q9Y2K7	Lysine-specific demethylase 2A	1	**0.94**	-	**0.80**	-			N.D.			damaged DNA-binding protein
P13473	Lysosome-associated membrane glycoprotein 2	2	**1.35**	4.3E-01	**1.13**	7.5E-01	2	**0.85**	9.2E-01	**0.80**	9.0E-01	membrane trafficking regulatory protein
P61626	Lysozyme C	14	**1.38**	4.6E-08	**1.26**	9.6E-04	27	**0.83**	4.4E-01	**0.84**	4.6E-01	hydrolase
Q08397	Lysyl oxidase homolog 1	1	**0.79**	3.3E-01	**1.04**	6.1E-01			N.D.			receptor
P09603	Macrophage colony-stimulating factor 1	2	**0.84**	6.6E-01	**0.83**	3.8E-01	2	**1.31**	3.8E-01	**1.08**	4.8E-01	cytokine
P07333	Macrophage colony-stimulating factor 1 receptor	3	**1.26**	3.5E-01	**1.11**	9.8E-02	5	**1.14**	3.7E-01	**1.38**	1.3E-01	protein kinase
P40121	Macrophage-capping protein	5	**2.64**	1.4E-03	**2.56**	1.6E-03	8	**1.44**	2.8E-01	**1.17**	7.0E-01	non-motor actin binding protein
P04156	Major prion protein	3	**0.91**	1.8E-01	**0.74**	1.7E-01	4	**1.09**	4.2E-01	**1.13**	2.0E-01	N.A.
P40925	Malate dehydrogenase, cytoplasmic	3	**1.62**	9.4E-03	**1.51**	1.3E-01	9	**1.52**	3.7E-01	**1.48**	3.9E-01	dehydrogenase
O43451	Maltase-glucoamylase, intestinal	2	**1.05**	7.1E-01	**1.00**	9.8E-01	2	**1.12**	2.9E-01	**0.78**	4.4E-01	glucosidase
P48740	Mannan-binding lectin serine protease 1	1	**0.57**	-	**0.78**	-	5	**1.33**	5.2E-01	**1.29**	3.7E-01	serine protease
O00187	Mannan-binding lectin serine protease 2	1	**0.89**	1.4E-02	**1.02**	6.8E-01	3	**0.88**	5.2E-01	**0.87**	4.7E-01	serine protease
P11226	Mannose-binding protein C			N.D.			1	**0.26**	-	**0.40**	-	defense/immunity protein
P33908	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	3	**1.01**	9.1E-01	**0.73**	1.7E-01	3	**0.50**	2.1E-01	**0.31**	1.1E-01	hydrolase
O00339	Matrilin-2	1	**0.74**	-	**0.75**	-			N.D.			extracellular matrix glycoprotein

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P09237	Matrilysin	1	**1.22**	7.4E-01	**1.08**	9.3E-01			N.D.			metalloprotease
P08493	Matrix Gla protein	1	**0.92**	-	**1.06**	-			N.D.			structural protein
P14780	Matrix metalloproteinase-9	14	**4.60**	2.4E-06	**4.75**	1.2E-05	25	**1.13**	5.4E-01	**1.01**	9.7E-01	metalloprotease
Q9NR99	Matrix-remodeling-associated protein 5	2	**1.65**	-	**1.11**	-			N.D.			immunoglobulin receptor superfamily
Q9BRK3	Matrix-remodeling-associated protein 8	3	**0.84**	1.0E-01	**1.11**	6.9E-01	4	**0.71**	8.0E-02	**0.75**	9.6E-03	receptor
Q5JRA6	Melanoma inhibitory activity protein 3	1	**1.75**	-	**0.90**	-			N.D.			growth factor
Q16819	Meprin A subunit alpha	2	**1.23**	8.6E-01	**1.18**	8.8E-01			N.D.			transporter
Q13421	Mesothelin	19	**1.06**	5.0E-01	**0.94**	8.3E-01	16	**0.81**	6.3E-01	**0.86**	5.8E-01	extracellular matrix glycoprotein
P01033	Metalloproteinase inhibitor 1	13	**1.07**	2.3E-01	**1.09**	3.8E-01	12	**0.91**	8.4E-01	**0.92**	8.0E-01	metalloprotease inhibitor
P16035	Metalloproteinase inhibitor 2	10	**0.76**	8.9E-03	**1.27**	3.6E-01	12	**1.01**	9.8E-01	**1.01**	9.9E-01	metalloprotease inhibitor
P80297	Metallothionein-1X	1	**1.58**	-	**2.39**	-			N.D.			N.A.
Q15726	Metastasis-suppressor KiSS-1	1	**0.63**	-	**0.60**	-	3	**1.30**	7.8E-01	**0.87**	9.1E-01	N.A.
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	1	**0.88**	-	**0.76**	-			N.D.			ligase
Q13361	Microfibrillar-associated protein 5	2	**0.55**	3.9E-01	**0.70**	7.1E-02	2	**0.93**	7.2E-01	**1.19**	5.9E-01	extracellular matrix glycoprotein
P27816	Microtubule-associated protein 4	1	**1.63**	-	**2.25**	-			N.D.			non-motor microtubule binding protein
Q9NU22	Midasin			N.D.			1	**6.43**	3.8E-01	**6.22**	3.8E-01	chaperone
P21741	Midkine	1	**1.10**	-	**1.01**	-			N.D.			cytokine
P20774	Mimecan	5	**1.05**	7.4E-01	**1.02**	8.7E-01	6	**1.01**	8.1E-01	**0.91**	3.4E-01	receptor
Q16584	Mitogen-activated protein kinase kinase kinase 11			N.D.			1	**0.28**	-	**0.30**	-	protein kinase
P26038	Moesin	17	**3.97**	2.6E-05	**3.27**	1.4E-06	30	**1.56**	2.1E-01	**1.40**	3.6E-01	actin family cytoskeletal protein
P08571	Monocyte differentiation antigen CD14	18	**1.15**	4.3E-02	**1.12**	4.7E-01	12	**0.96**	9.2E-01	**0.91**	8.1E-01	receptor
P15941	Mucin-1	3	**1.20**	2.3E-02	**0.68**	1.4E-01	3	**1.07**	9.2E-01	**1.10**	9.0E-01	non-motor actin binding protein
Q8WXI7	Mucin-16	7	**0.98**	7.7E-01	**0.87**	1.2E-01	6	**1.23**	2.2E-02	**1.39**	1.9E-02	N.A.
P98088	Mucin-5AC (Fragments)	22	**0.80**	6.9E-09	**0.77**	1.0E-03	28	**1.06**	6.2E-01	**1.07**	5.8E-01	signaling molecule
Q9HC84	Mucin-5B	44	**0.86**	7.8E-07	**0.82**	4.8E-04	77	**1.33**	1.2E-02	**1.42**	3.4E-03	signaling molecule
Q8TAX7	Mucin-7	4	**0.92**	5.5E-01	**0.94**	9.5E-01	6	**0.44**	1.6E-01	**0.41**	1.2E-01	N.A.
P33527	Multidrug resistance-associated protein 1	1	**0.00**	-	**0.00**	-			N.D.			ATP-binding cassette (ABC) transporter
Q13201	Multimerin-1	1	**0.96**	5.1E-01	**1.24**	8.0E-01	2	**1.47**	1.3E-01	**1.97**	6.8E-02	extracellular matrix glycoprotein
P24158	Myeloblastin	3	**5.14**	6.7E-02	**2.81**	1.7E-01	3	**0.73**	6.1E-01	**0.65**	3.3E-01	serine protease
P41218	Myeloid cell nuclear differentiation antigen	1	**6.81**	-	**5.25**	-	5	**0.96**	8.9E-01	**1.03**	9.2E-01	transcription factor
P05164	Myeloperoxidase	34	**5.11**	3.5E-16	**6.11**	1.5E-14	30	**1.18**	3.7E-01	**1.18**	3.6E-01	peroxidase
P02144	Myoglobin			N.D.			7	**0.64**	4.2E-01	**0.64**	3.2E-01	transfer/carrier protein
P60660	Myosin light polypeptide 6	5	**1.93**	6.4E-03	**1.94**	1.6E-02	5	**1.62**	2.4E-01	**1.77**	3.5E-01	actin family cytoskeletal protein
P19105	Myosin regulatory light chain 12A	1	**2.99**	-	**2.53**	-			N.D.			actin family cytoskeletal protein
P35579	Myosin-9	12	**2.26**	3.3E-04	**2.01**	3.4E-04	9	**1.25**	6.4E-01	**1.28**	6.4E-01	G-protein modulator
P29966	Myristoylated alanine-rich C-kinase substrate	1	**1.39**	2.7E-01	**0.85**	7.6E-01			N.D.			structural protein
Q9UJJ9	N-acetylglucosamine-1-phosphotransferase subunit gamma	2	**1.22**	6.0E-01	**0.93**	8.3E-01	2	**0.89**	2.2E-01	**0.89**	4.2E-01	transferase
P15586	N-acetylglucosamine-6-sulfatase	3	**1.32**	7.8E-02	**1.24**	2.2E-01	3	**0.92**	5.5E-01	**0.88**	4.9E-01	hydrolase
O43505	N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase			N.D.			2	**0.87**	8.2E-01	**0.87**	8.1E-01	glycosyltransferase
Q96PD5	N-acetylmuramoyl-L-alanine amidase	14	**0.85**	1.4E-01	**0.85**	3.6E-01	22	**0.98**	9.5E-01	**0.89**	7.1E-01	signaling molecule
O95865	N(G),N(G)-dimethylarginine dimethylaminohydrolase 2	1	**0.85**	1.7E-01	**1.37**	1.6E-01			N.D.			N.A.
O96009	Napsin-A	3	**1.41**	1.8E-02	**1.34**	3.2E-01	3	**0.74**	6.7E-02	**0.69**	4.5E-02	aspartic protease
Q15843	NEDD8	1	**1.62**	-	**2.94**	-	2	**1.89**	4.7E-01	**1.68**	3.2E-01	ribosomal protein
Q9UMX5	Neudesin			N.D.			3	**0.97**	8.9E-01	**1.01**	9.8E-01	signaling molecule
Q9UJJ8	Neurabin-1			N.D.			1	**1.31**	2.6E-01	**1.27**	1.6E-01	N.A.
P13591	Neural cell adhesion molecule 1	0	**0.00**	-	**0.00**	-			N.D.			immunoglobulin receptor superfamily
Q9ULB1	Neurexin-1-alpha	2	**0.88**	1.4E-01	**0.61**	2.2E-01	1	**0.23**	-	**0.29**	-	transporter
Q09666	Neuroblast differentiation-associated protein AHNAK	4	**1.07**	4.2E-01	**0.84**	5.3E-02	10	**1.33**	1.4E-01	**1.17**	3.8E-01	N.A.
P41271	Neuroblastoma suppressor of tumorigenicity 1	1	**1.21**	-	**0.73**	-			N.D.			N.A.
P21359	Neurofibromin	1	**0.00**	-	**0.00**	-			N.D.			signaling molecule
Q14697	Neutral alpha-glucosidase AB	1	**1.50**	-	**1.10**	-			N.D.			glucosidase
P22894	Neutrophil collagenase	9	**4.14**	1.5E-02	**4.60**	1.4E-02	11	**1.23**	3.4E-01	**1.35**	3.2E-01	metalloprotease
P59665	Neutrophil defensin 1	8	**4.28**	1.8E-08	**4.30**	4.9E-08			N.D.			N.A.
P59666	Neutrophil defensin 3			N.D.			10	**1.89**	9.5E-03	**1.84**	5.1E-02	N.A.
P08246	Neutrophil elastase	1	**5.17**	-	**5.17**	-			N.D.			serine protease
P80188	Neutrophil gelatinase-associated lipocalin	20	**4.54**	3.8E-18	**5.28**	9.3E-13	44	**1.09**	7.6E-01	**1.15**	5.8E-01	transfer/carrier protein
Q96TA1	Niban-like protein 1	1	**1.29**	-	**0.92**	-	1	**0.87**	-	**0.98**	-	N.A.
P43490	Nicotinamide phosphoribosyltransferase	3	**6.94**	6.0E-03	**4.80**	7.4E-02	10	**1.37**	6.1E-02	**1.39**	1.2E-01	cytokine
Q6XQN6	Nicotinate phosphoribosyltransferase	1	**4.70**	-	**0.94**	-			N.D.			glycosyltransferase
Q15274	Nicotinate-nucleotide pyrophosphorylase [carboxylating]	1	**0.91**	-	**1.14**	-			N.D.			N.A.
P14543	Nidogen-1	9	**0.96**	7.6E-01	**0.90**	7.4E-01	9	**1.16**	7.9E-01	**0.98**	9.8E-01	receptor
Q14112	Nidogen-2	6	**1.10**	2.1E-01	**1.04**	7.6E-01	7	**0.90**	8.1E-01	**0.94**	8.7E-01	receptor
P05204	Non-histone chromosomal protein HMG-17			N.D.			4	**1.11**	9.4E-01	**0.86**	9.3E-01	N.A.
P10153	Non-secretory ribonuclease	2	**1.58**	1.4E-01	**1.55**	2.7E-01	3	**1.02**	8.7E-01	**1.13**	6.0E-01	endoribonuclease
Q6ZVX7	Non-specific cytotoxic cell receptor protein 1 homolog	3	**0.95**	6.9E-01	**0.87**	1.5E-01	3	**1.06**	7.2E-01	**0.97**	8.9E-01	N.A.
P61970	Nuclear transport factor 2	2	**1.20**	3.5E-01	**1.35**	2.5E-01	3	**0.87**	5.4E-01	**1.03**	9.6E-01	N.A.
Q02818	Nucleobindin-1	8	**1.19**	1.5E-02	**1.07**	4.6E-01	10	**0.75**	3.7E-02	**0.77**	7.8E-02	nucleic acid binding
P80303	Nucleobindin-2	1	**1.35**	-	**0.87**	-	1	**0.86**	-	**0.84**	-	nucleic acid binding
P19338	Nucleolin			N.D.			1	**0.21**	-	**0.25**	-	replication origin binding protein
P06748	Nucleophosmin	1	**2.38**	-	**1.83**	-			N.D.			chaperone
P22392	Nucleoside diphosphate kinase B	3	**1.91**	5.0E-03	**1.62**	1.3E-01	4	**0.73**	4.8E-01	**0.77**	4.8E-01	N.A.
P55209	Nucleosome assembly protein 1-like 1			N.D.			1	**1.08**	-	**1.34**	-	phosphatase inhibitor
Q99733	Nucleosome assembly protein 1-like 4	1	**1.83**	-	**1.47**	-			N.D.			phosphatase inhibitor
Q9NTK5	Obg-like ATPase 1			N.D.			1	**1.21**	-	**1.11**	-	G-protein
Q6UX06	Olfactomedin-4			N.D.			3	**0.78**	6.2E-01	**0.65**	3.4E-01	structural protein
Q92882	Osteoclast-stimulating factor 1	2	**6.68**	5.3E-02	**6.99**	9.0E-03			N.D.			N.A.
P10451	Osteopontin	12	**1.31**	3.7E-02	**1.42**	3.1E-01	16	**0.83**	7.3E-01	**0.80**	4.4E-01	cytokine
Q86UD1	Out at first protein homolog			N.D.			1	**0.82**	-	**0.52**	-	N.A.
P00995	Pancreatic secretory trypsin inhibitor	3	**0.89**	5.4E-01	**1.10**	5.2E-01	2	**1.62**	8.8E-02	**1.76**	1.5E-01	serine protease inhibitor
O95428	Papilin	2	**1.16**	2.1E-01	**1.07**	7.3E-01	2	**1.63**	3.3E-01	**1.52**	1.1E-01	metalloprotease
Q13219	Pappalysin-1	9	**0.71**	2.3E-02	**0.79**	1.0E-02	15	**1.18**	1.9E-01	**1.09**	4.2E-01	apolipoprotein
Q9BXP8	Pappalysin-2	27	**0.71**	1.2E-07	**0.69**	1.0E-03	28	**1.00**	9.9E-01	**1.00**	9.9E-01	apolipoprotein
P26022	Pentraxin-related protein PTX3	1	**3.08**	-	**2.82**	-	1	**1.59**	4.5E-01	**1.19**	-	antibacterial response protein
O43692	Peptidase inhibitor 15	2	**0.76**	5.0E-01	**1.40**	7.2E-01			N.D.			defense/immunity protein
Q6UXB8	Peptidase inhibitor 16	9	**0.90**	3.4E-01	**1.15**	3.0E-01	9	**1.26**	1.7E-01	**1.27**	2.9E-01	defense/immunity protein
O75594	Peptidoglycan recognition protein 1	5	**4.33**	6.7E-07	**4.34**	1.0E-06	5	**1.49**	4.7E-02	**1.49**	3.5E-02	signaling molecule
P19021	Peptidyl-glycine alpha-amidating monooxygenase	1	**1.04**	-	**0.96**	-	3	**0.73**	5.3E-02	**0.79**	4.7E-01	oxygenase
P62937	Peptidyl-prolyl cis-trans isomerase A	11	**2.19**	1.9E-04	**1.99**	3.6E-02	21	**0.91**	8.8E-01	**0.92**	8.9E-01	N.A.
P23284	Peptidyl-prolyl cis-trans isomerase B			N.D.			3	**1.00**	1.0E+00	**1.21**	9.0E-01	isomerase
P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A	2	**1.69**	2.0E-01	**1.45**	6.0E-01	3	**1.18**	4.5E-02	**0.99**	8.7E-01	isomerase
O60664	Perilipin-3			N.D.			2	**1.41**	9.7E-02	**1.05**	8.2E-01	transfer/carrier protein
Q15063	Periostin	4	**1.14**	1.7E-01	**1.18**	1.1E-01	8	**0.90**	5.7E-01	**0.88**	5.2E-01	signaling molecule
O60437	Periplakin	9	**0.85**	1.9E-01	**0.77**	1.4E-01	19	**0.84**	4.2E-01	**0.85**	4.3E-01	intermediate filament binding protein
Q92626	Peroxidasin homolog	0	**1.13**	-	**0.68**	-			N.D.			peroxidase
Q06830	Peroxiredoxin-1	6	**1.16**	2.5E-02	**1.34**	1.5E-03	10	**1.35**	7.2E-01	**1.27**	8.4E-01	peroxidase
P32119	Peroxiredoxin-2	5	**0.67**	1.1E-02	**0.71**	3.2E-02	17	**1.28**	5.6E-01	**1.25**	6.8E-01	peroxidase
P30044	Peroxiredoxin-5, mitochondrial	2	**2.34**	2.5E-02	**1.73**	1.5E-01	3	**0.71**	3.8E-01	**0.59**	2.8E-01	peroxidase
P30041	Peroxiredoxin-6	6	**1.02**	5.4E-01	**1.08**	4.7E-01	13	**0.87**	5.3E-01	**0.78**	4.0E-01	peroxidase
Q9BYK8	Peroxisomal proliferator-activated receptor A-interacting complex			N.D.			1	**0.00**	-	**0.00**	-	DNA helicase
P04180	Phosphatidylcholine-sterol acyltransferase	1	**1.30**	-	**1.21**	-	1	**1.11**	3.5E-01	**1.04**	9.3E-01	acyltransferase
P30086	Phosphatidylethanolamine-binding protein 1	5	**1.18**	1.6E-01	**1.15**	7.8E-01	8	**1.04**	8.7E-01	**1.10**	6.9E-01	transfer/carrier protein
Q96S96	Phosphatidylethanolamine-binding protein 4			N.D.			6	**1.05**	8.4E-01	**1.19**	5.0E-01	transfer/carrier protein

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P80108	Phosphatidylinositol-glycan-specific phospholipase D			N.D.			5	**1.14**	4.9E-01	**1.03**	8.3E-01	lipase
P36871	Phosphoglucomutase-1	1	**2.90**	4.2E-02	**2.13**	4.0E-03	3	**1.26**	7.0E-02	**1.07**	8.4E-01	mutase
P00558	Phosphoglycerate kinase 1	14	**2.32**	3.0E-08	**2.39**	2.3E-04	26	**1.07**	8.4E-01	**1.09**	8.1E-01	carbohydrate kinase
P18669	Phosphoglycerate mutase 1	10	**2.43**	1.5E-04	**2.48**	2.0E-03	15	**0.95**	9.0E-01	**1.11**	7.7E-01	mutase
Q96FE7	Phosphoinositide-3-kinase-interacting protein 1	2	**0.85**	2.5E-01	**0.46**	2.7E-01	4	**0.37**	1.4E-01	**0.33**	2.1E-01	growth factor
P55058	Phospholipid transfer protein	8	**1.30**	1.1E-05	**0.74**	1.3E-01	7	**0.70**	3.9E-01	**0.54**	1.1E-01	transfer/carrier protein
Q96FC7	Phytanoyl-CoA hydroxylase-interacting protein-like	1	**1.03**	-	**0.56**	-			N.D.			N.A.
P36955	Pigment epithelium-derived factor	31	**0.95**	2.9E-01	**0.95**	6.8E-01	38	**1.00**	9.9E-01	**1.01**	9.7E-01	serine protease inhibitor
P03952	Plasma kallikrein	5	**0.98**	8.3E-01	**0.89**	1.7E-01	5	**1.06**	8.8E-01	**0.99**	9.8E-01	serine protease
P05155	Plasma protease C1 inhibitor	24	**0.86**	4.8E-04	**0.90**	4.2E-01	15	**0.88**	7.1E-01	**0.92**	8.3E-01	serine protease inhibitor
P05154	Plasma serine protease inhibitor	8	**0.85**	2.1E-03	**0.71**	1.1E-03	7	**0.76**	8.8E-02	**0.71**	6.8E-02	serine protease inhibitor
P00747	Plasminogen	113	**0.70**	1.6E-20	**0.77**	1.6E-04	151	**1.04**	7.4E-01	**1.01**	9.6E-01	peptide hormone
P05121	Plasminogen activator inhibitor 1	17	**0.87**	4.4E-02	**0.83**	2.7E-01	18	**0.60**	1.0E-01	**0.57**	5.7E-02	serine protease inhibitor
P05120	Plasminogen activator inhibitor 2	5	**0.95**	8.2E-01	**0.87**	3.4E-01	5	**0.81**	3.9E-01	**0.71**	2.2E-01	serine protease inhibitor
P13796	Plastin-2	35	**4.39**	1.6E-21	**4.17**	2.6E-15	40	**1.08**	7.0E-01	**1.00**	9.9E-01	non-motor actin binding protein
P13797	Plastin-3			N.D.			11	**0.93**	5.9E-01	**0.90**	3.1E-01	non-motor actin binding protein
P02775	Platelet basic protein	1	**0.58**	-	**0.27**	-	5	**1.24**	6.5E-02	**1.22**	8.5E-02	chemokine
P10720	Platelet factor 4 variant			N.D.			2	**0.99**	8.3E-01	**0.95**	4.9E-01	chemokine
P07359	Platelet glycoprotein Ib alpha chain	1	**1.41**	-	**0.91**	-	2	**1.16**	1.4E-01	**1.15**	4.7E-01	receptor
Q15149	Plectin	1	**1.53**	-	**1.47**	-	4	**1.35**	1.8E-01	**1.53**	2.3E-01	N.A.
O15031	Plexin-B2	1	**1.09**	7.0E-01	**0.80**	-			N.D.			protein kinase
Q9Y4D7	Plexin-D1	1	**0.86**	7.2E-01	**0.86**	8.0E-01			N.D.			protein kinase
Q9NZ53	Podocalyxin-like protein 2	1	**0.90**	-	**1.42**	-			N.D.			N.A.
P15151	Poliovirus receptor	2	**0.81**	1.3E-01	**1.10**	4.7E-01	3	**1.19**	3.9E-01	**1.18**	2.5E-01	receptor
Q92692	Poliovirus receptor-related protein 2	3	**0.89**	2.5E-01	**1.04**	9.0E-01	1	**1.23**	7.9E-01	**1.18**	6.5E-01	receptor
Q9NQS3	Poliovirus receptor-related protein 3	1	**1.08**	-	**1.07**	-			N.D.			receptor
Q96NY8	Poliovirus receptor-related protein 4			N.D.			1	**1.02**	-	**1.50**	-	receptor
Q15365	Poly(rC)-binding protein 1			N.D.			1	**1.15**	-	**1.01**	-	mRNA processing factor
P21128	Poly(U)-specific endoribonuclease	1	**0.83**	-	**0.68**	-	1	**0.41**	-	**0.53**	-	serine protease
P01833	Polymeric immunoglobulin receptor	33	**1.09**	2.2E-02	**1.06**	5.6E-01	56	**0.96**	8.6E-01	**0.94**	7.9E-01	immunoglobulin receptor superfamily
Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	1	**1.06**	-	**0.87**	-	3	**1.02**	8.5E-01	**1.01**	9.2E-01	glycosyltransferase
Q8NCL4	Polypeptide N-acetylgalactosaminyltransferase 6	1	**1.11**	-	**1.12**	-	1	**1.35**	8.1E-01	**1.49**	8.0E-01	glycosyltransferase
P0CG48	Polyubiquitin-C	3	**1.06**	7.0E-01	**1.00**	1.0E+00	6	**0.25**	4.9E-04	**0.26**	1.4E-03	N.A.
P0CG38	POTE ankyrin domain family member I	20	**1.86**	-	**2.54**	-			N.D.			N.A.
Q5VWM3	PRAME family member 18			N.D.			3	**0.98**	-	**0.80**	-	N.A.
P20742	Pregnancy zone protein	24	**1.01**	9.2E-01	**0.83**	7.9E-02	48	**0.72**	8.6E-02	**0.73**	1.1E-01	cytokine
P11464	Pregnancy-specific beta-1-glycoprotein 1	18	**0.67**	3.0E-02	**0.82**	2.6E-01	24	**1.00**	8.7E-01	**0.94**	8.6E-01	N.A.
Q9UQ72	Pregnancy-specific beta-1-glycoprotein 11	8	**0.76**	1.4E-01	**0.65**	1.3E-01	11	**0.81**	7.7E-02	**0.73**	1.6E-01	N.A.
P11465	Pregnancy-specific beta-1-glycoprotein 2	10	**0.96**	8.9E-01	**1.02**	6.3E-01	17	**0.97**	6.8E-01	**0.84**	3.4E-01	immunoglobulin cell adhesion molecule
Q16557	Pregnancy-specific beta-1-glycoprotein 3	15	**0.85**	-	**0.79**	-			N.D.			immunoglobulin cell adhesion molecule
Q00888	Pregnancy-specific beta-1-glycoprotein 4	12	**0.48**	9.4E-03	**0.73**	4.2E-02	15	**1.33**	4.2E-01	**1.43**	2.9E-01	N.A.
Q15238	Pregnancy-specific beta-1-glycoprotein 5	4	**0.88**	-	**0.83**	-			N.D.			N.A.
Q00887	Pregnancy-specific beta-1-glycoprotein 9	12	**0.57**	1.8E-03	**1.08**	7.7E-01	15	**1.06**	5.3E-01	**1.01**	9.2E-01	immunoglobulin cell adhesion molecule
Q9UKY0	Prion-like protein doppel	2	**0.83**	3.6E-01	**0.82**	3.8E-01			N.D.			N.A.
P01133	Pro-epidermal growth factor	1	**0.52**	-	**1.12**	-			N.D.			growth factor
O14511	Pro-neuregulin-2, membrane-bound isoform			N.D.			1	**5.27**	-	**4.62**	-	growth factor
P07602	Proactivator polypeptide	20	**1.57**	2.3E-08	**2.04**	1.5E-07	14	**0.97**	8.7E-01	**0.85**	3.3E-01	enzyme modulator
Q15113	Procollagen C-endopeptidase enhancer 1	10	**0.85**	3.5E-02	**0.88**	2.9E-01	10	**1.05**	8.5E-01	**1.04**	9.1E-01	transporter
P01210	Proenkephalin-A			N.D.			1	**0.73**	-	**0.63**	-	neuropeptide
P07737	Profilin-1	11	**3.37**	3.3E-08	**3.43**	3.7E-09	22	**1.01**	9.8E-01	**0.98**	9.6E-01	N.A.
Q8WUM4	Programmed cell death 6-interacting protein	3	**1.11**	9.6E-02	**1.10**	2.8E-01	7	**1.27**	4.8E-01	**1.39**	1.6E-01	transmembrane receptor
P01236	Prolactin	21	**1.06**	5.8E-01	**0.92**	3.8E-01	16	**1.09**	8.0E-01	**1.10**	8.1E-01	growth factor
P12273	Prolactin-inducible protein	4	**0.74**	2.6E-02	**0.62**	6.3E-02	1	**0.55**	6.6E-01	**0.41**	5.0E-01	N.A.
Q16378	Proline-rich protein 4	6	**0.67**	7.5E-02	**0.63**	3.3E-02	8	**0.80**	5.0E-02	**0.76**	9.2E-03	N.A.
O43490	Prominin-1	2	**0.99**	5.3E-01	**0.99**	9.9E-01	3	**0.29**	1.3E-01	**0.54**	1.3E-01	membrane traffic protein
P27918	Properdin	2	**1.72**	5.2E-01	**1.87**	2.9E-01	3	**1.11**	1.3E-01	**1.17**	2.1E-01	N.A.
P41222	Prostaglandin-H2 D-isomerase	37	**0.92**	1.3E-03	**1.04**	4.9E-01	36	**1.08**	5.8E-01	**1.06**	6.8E-01	isomerase
Q16651	Prostasin			N.D.			1	**0.50**	3.7E-01	**0.39**	2.3E-01	serine protease
P15309	Prostatic acid phosphatase	2	**0.92**	6.8E-01	**0.46**	1.2E-01			N.D.			phosphatase
Q06323	Proteasome activator complex subunit 1	4	**2.41**	9.3E-02	**1.43**	4.7E-01	2	**0.88**	6.9E-01	**0.84**	6.4E-01	N.A.
Q9UL46	Proteasome activator complex subunit 2	2	**2.68**	7.9E-03	**2.09**	2.9E-01	1	**1.67**	1.6E-01	**1.45**	7.9E-02	N.A.
P25786	Proteasome subunit alpha type-1	1	**2.09**	-	**2.26**	-			N.D.			protease
P25789	Proteasome subunit alpha type-4			N.D.			1	**0.91**	-	**0.88**	-	protease
P28066	Proteasome subunit alpha type-5	1	**1.70**	-	**2.11**	-	2	**0.99**	1.0E+00	**0.90**	9.5E-01	protease
P60900	Proteasome subunit alpha type-6	1	**1.67**	-	**2.07**	-	2	**1.17**	4.5E-01	**1.34**	4.0E-01	protease
P20618	Proteasome subunit beta type-1	2	**1.59**	3.1E-01	**1.60**	2.6E-01	2	**1.28**	3.2E-01	**1.16**	4.3E-01	protease
P28070	Proteasome subunit beta type-4			N.D.			2	**1.31**	-	**1.28**	-	protease
P02760	Protein AMBP	162	**0.92**	3.6E-04	**0.94**	1.0E-01	141	**0.90**	4.4E-01	**0.90**	2.9E-01	serine protease inhibitor
Q9UKY7	Protein CDV3 homolog			N.D.			2	**1.72**	3.2E-01	**1.22**	5.5E-01	N.A.
P80370	Protein delta homolog 1	27	**0.73**	1.9E-06	**0.79**	2.9E-03	19	**0.72**	6.4E-02	**0.73**	6.8E-02	transcription factor
P07237	Protein disulfide-isomerase	13	**1.33**	5.4E-03	**1.38**	1.4E-01	18	**1.19**	3.7E-01	**1.18**	4.4E-01	isomerase
P30101	Protein disulfide-isomerase A3	7	**1.23**	1.5E-02	**1.18**	2.2E-01	13	**1.11**	7.9E-01	**1.12**	5.8E-01	isomerase
P13667	Protein disulfide-isomerase A4	2	**1.08**	6.5E-01	**1.06**	6.2E-01			N.D.			isomerase
Q15084	Protein disulfide-isomerase A6	4	**1.10**	4.2E-01	**1.01**	9.4E-01	6	**1.04**	8.9E-01	**1.16**	5.8E-01	isomerase
Q99497	Protein DJ-1	5	**1.63**	8.0E-03	**1.92**	1.4E-01	7	**1.45**	4.4E-01	**1.52**	3.9E-01	transcription factor
Q6UWH4	Protein FAM198B	2	**0.73**	4.3E-01	**1.03**	8.2E-01			N.D.			N.A.
Q92520	Protein FAM3C	3	**1.00**	9.8E-01	**1.11**	1.8E-01	6	**1.32**	6.8E-01	**1.27**	7.7E-01	N.A.
Q9NUQ9	Protein FAM49B	1	**2.70**	-	**2.95**	-	2	**0.87**	3.8E-01	**0.79**	1.2E-01	N.A.
Q6P988	Protein notum homolog	12	**0.74**	5.0E-06	**0.64**	2.4E-02	17	**1.02**	9.2E-01	**1.11**	5.9E-01	glycosidase
P48745	Protein NOV homolog	10	**0.82**	1.9E-03	**0.94**	5.1E-01	4	**1.08**	4.8E-01	**1.09**	7.7E-01	growth factor
Q9NP55	Protein Plunc	5	**1.33**	1.7E-01	**0.29**	3.0E-03	2	**0.28**	1.8E-01	**0.16**	1.8E-01	carbohydrate transporter
P60903	Protein S100-A10			N.D.			3	**1.35**	8.8E-01	**1.69**	7.8E-01	calmodulin
P31949	Protein S100-A11	4	**2.60**	8.4E-04	**2.90**	6.3E-04	6	**1.86**	3.9E-01	**1.93**	3.7E-01	signaling molecule
P80511	Protein S100-A12	5	**2.55**	1.3E-01	**2.08**	2.5E-01	3	**1.17**	8.4E-01	**1.56**	5.5E-01	signaling molecule
P29034	Protein S100-A2	1	**0.83**	-	**1.25**	-	4	**1.14**	9.3E-01	**1.08**	9.6E-01	growth factor
P26447	Protein S100-A4	1	**1.30**	-	**1.35**	-	7	**1.55**	5.9E-01	**1.27**	8.2E-01	growth factor
P06703	Protein S100-A6			N.D.			5	**1.07**	9.7E-01	**0.93**	9.4E-01	growth factor
P31151	Protein S100-A7	4	**0.88**	2.2E-01	**1.00**	9.9E-01	10	**0.87**	7.4E-01	**0.94**	8.8E-01	signaling molecule
P05109	Protein S100-A8	25	**3.17**	1.1E-18	**2.89**	1.7E-09	35	**1.17**	6.5E-01	**1.05**	9.1E-01	signaling molecule
P06702	Protein S100-A9	44	**3.86**	7.5E-33	**3.14**	1.5E-27	48	**1.18**	2.1E-01	**1.16**	3.6E-01	signaling molecule
P25815	Protein S100-P	2	**3.09**	2.4E-01	**3.47**	1.8E-01	2	**0.68**	8.1E-01	**0.62**	7.9E-01	signaling molecule
Q01105	Protein SET	2	**3.23**	6.5E-02	**3.49**	9.2E-02	2	**1.04**	4.7E-01	**1.05**	3.3E-01	phosphatase inhibitor
Q8N114	Protein shisa-5	2	**0.86**	4.1E-01	**0.93**	7.3E-01	4	**1.28**	1.6E-01	**1.17**	3.9E-01	zinc finger transcription factor
Q9C0D5	Protein TANC1	1	**0.00**	-	**0.00**	-			N.D.			N.A.
Q9GZM5	Protein YIPF3	2	**0.79**	1.3E-01	**0.76**	6.8E-02	2	**0.97**	8.3E-01	**0.80**	3.7E-01	N.A.
Q9UK55	Protein Z-dependent protease inhibitor	3	**0.99**	9.1E-01	**0.79**	2.7E-02	3	**0.80**	4.9E-01	**0.87**	4.1E-01	serine protease inhibitor
P21980	Protein-glutamine gamma-glutamyltransferase 2	2	**0.87**	2.8E-01	**1.04**	6.6E-01	5	**1.30**	7.5E-01	**1.29**	7.8E-01	acyltransferase
Q08188	Protein-glutamine gamma-glutamyltransferase E	2	**0.80**	5.0E-02	**0.80**	8.0E-02	1	**1.20**	6.5E-01	**1.51**	6.2E-01	acyltransferase
P28300	Protein-lysine 6-oxidase	1	**0.78**	-	**0.83**	-	1	**0.70**	8.2E-01	**0.54**	6.5E-01	receptor
P00734	Prothrombin	69	**0.71**	1.8E-19	**0.71**	1.9E-06	81	**0.94**	6.0E-01	**0.97**	7.8E-01	serine protease
Q8IWL2	Pulmonary surfactant-associated protein A1	10	**1.39**	3.1E-04	**1.39**	1.9E-03	10	**0.59**	6.0E-02	**0.60**	2.4E-01	N.A.
P07988	Pulmonary surfactant-associated protein B	14	**1.34**	1.4E-05	**1.19**	3.9E-01	9	**0.73**	5.1E-01	**0.69**	3.1E-01	surfactant

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P35247	Pulmonary surfactant-associated protein D	6	**0.93**	5.9E-01	**0.77**	2.4E-01	9	**0.90**	2.1E-01	**0.89**	2.8E-01	surfactant
P00491	Purine nucleoside phosphorylase	4	**1.98**	2.0E-04	**1.47**	8.5E-03	5	**0.72**	5.3E-01	**0.84**	6.9E-01	phosphorylase
P55786	Puromycin-sensitive aminopeptidase	2	**1.18**	1.8E-02	**0.79**	4.8E-01	6	**1.07**	3.6E-01	**1.09**	5.2E-01	metalloprotease
Q5VTE0	Putative elongation factor 1-alpha-like 3	2	**1.60**	2.0E-01	**1.40**	9.3E-02	5	**0.67**	1.0E-01	**0.61**	1.8E-01	N.A.
A6NGU5	Putative gamma-glutamyltranspeptidase 3			N.D.			1	**1.53**	-	**1.61**	-	acyltransferase
P0C7M2	Putative heterogeneous nuclear ribonucleoprotein A1-like 3	4	**3.96**	-	**2.79**	-	6	**1.58**	1.2E-01	**1.40**	3.7E-01	replication origin binding protein
Q6DRA6	Putative histone H2B type 2-D	2	**1.47**	-	**1.28**	-			N.D.			N.A.
Q13046	Putative pregnancy-specific beta-1-glycoprotein 7	14	**0.72**	4.4E-01	**0.82**	8.5E-01	20	**1.22**	9.3E-01	**1.18**	9.4E-01	N.A.
Q8NFI4	Putative protein FAM10A5	1	**1.15**	-	**1.08**	-			N.D.			chaperone
Q92928	Putative Ras-related protein Rab-1C			N.D.			2	**0.93**	-	**1.00**	-	small GTPase
Q9H7F4	Putative transmembrane protein 185B	1	**0.97**	-	**1.07**	-			N.D.			N.A.
A6NL28	Putative tropomyosin alpha-3 chain-like protein			N.D.			2	**1.42**	-	**1.56**	-	actin binding motor protein
P14618	Pyruvate kinase isozymes M1/M2	11	**2.28**	3.3E-06	**2.18**	4.1E-06	11	**1.07**	4.1E-01	**0.97**	8.0E-01	carbohydrate kinase
Q53FA7	Quinone oxidoreductase PIG3	1	**0.48**	-	**0.74**	-			N.D.			dehydrogenase
P31150	Rab GDP dissociation inhibitor alpha	3	**1.46**	5.1E-01	**1.27**	6.3E-01	6	**1.19**	2.5E-01	**1.32**	3.9E-02	acyltransferase
P50395	Rab GDP dissociation inhibitor beta	8	**2.34**	4.0E-05	**2.01**	4.9E-02	16	**0.70**	3.3E-01	**0.67**	2.1E-01	acyltransferase
P35241	Radixin	9	**1.50**	-	**1.11**	-			N.D.			actin family cytoskeletal protein
P46940	Ras GTPase-activating-like protein IQGAP1	3	**3.05**	3.7E-02	**1.98**	9.8E-02	2	**0.79**	6.6E-01	**0.99**	8.9E-01	G-protein modulator
P61026	Ras-related protein Rab-10	1	**1.47**	-	**0.99**	-			N.D.			small GTPase
P51149	Ras-related protein Rab-7a	1	**2.45**	-	**2.20**	-			N.D.			small GTPase
P11234	Ras-related protein Ral-B	1	**0.80**	-	**1.14**	-			N.D.			small GTPase
P10586	Receptor-type tyrosine-protein phosphatase F	1	**1.05**	-	**0.92**	-			N.D.			receptor
P23470	Receptor-type tyrosine-protein phosphatase gamma	2	**1.30**	-	**0.82**	-	1	**0.79**	-	**0.50**	-	receptor
Q13332	Receptor-type tyrosine-protein phosphatase S	1	**0.83**	8.0E-01	**0.84**	6.4E-01			N.D.			receptor
Q9NZ71	Regulator of telomere elongation helicase 1			N.D.			1	**0.00**	-	**0.00**	-	DNA helicase
Q9HD89	Resistin	2	**5.73**	1.2E-01	**7.99**	1.1E-01	1	**1.82**	1.3E-01	**1.84**	3.8E-02	N.A.
Q15293	Reticulocalbin-1			N.D.			2	**0.24**	-	**0.32**	-	calmodulin
Q96D15	Reticulocalbin-3			N.D.			3	**0.78**	1.8E-01	**0.66**	2.4E-01	calmodulin
P00352	Retinal dehydrogenase 1			N.D.			5	**0.68**	5.0E-01	**0.71**	4.1E-01	dehydrogenase
Q99969	Retinoic acid receptor responder protein 2	1	**0.76**	-	**0.64**	-			N.D.			N.A.
Q8NFJ5	Retinoic acid-induced protein 3	7	**0.82**	2.0E-02	**0.54**	1.3E-02	8	**0.50**	2.7E-01	**0.54**	1.7E-01	G-protein coupled receptor
P09455	Retinol-binding protein 1	1	**0.00**	-	**0.00**	-			N.D.			transfer/carrier protein
P02753	Retinol-binding protein 4	27	**0.99**	9.0E-01	**1.11**	3.7E-01	53	**1.10**	6.4E-01	**1.16**	5.5E-01	transfer/carrier protein
P52565	Rho GDP-dissociation inhibitor 1	5	**2.58**	5.2E-06	**2.71**	1.2E-05	7	**1.07**	8.0E-01	**1.08**	7.4E-01	signaling molecule
P52566	Rho GDP-dissociation inhibitor 2	2	**7.45**	1.1E-02	**8.02**	1.4E-02	10	**0.66**	6.8E-01	**0.64**	6.9E-01	signaling molecule
P34096	Ribonuclease 4	1	**0.87**	6.4E-01	**0.65**	4.7E-01	4	**0.77**	2.8E-02	**0.82**	1.9E-01	endoribonuclease
P13489	Ribonuclease inhibitor	3	**1.73**	7.2E-03	**1.88**	5.8E-02	2	**0.98**	9.7E-01	**0.98**	9.4E-01	enzyme modulator
P07998	Ribonuclease pancreatic	20	**0.90**	7.4E-02	**0.99**	9.0E-01	22	**0.91**	4.8E-01	**0.84**	2.9E-01	endoribonuclease
O00584	Ribonuclease T2			N.D.			6	**1.14**	8.4E-01	**1.17**	8.1E-01	endoribonuclease
P51812	Ribosomal protein S6 kinase alpha-3	1	**0.00**	-	**0.00**	-			N.D.			protein kinase
Q6ZP01	RNA-binding protein 44			N.D.			1	**0.72**	5.6E-01	**0.53**	5.3E-01	mRNA processing factor
Q8WZ75	Roundabout homolog 4	4	**0.93**	6.2E-01	**1.03**	7.9E-01	5	**1.23**	1.5E-01	**1.10**	7.3E-01	immunoglobulin receptor superfamily
P10768	S-formylglutathione hydrolase	1	**1.82**	-	**2.16**	-	2	**0.92**	6.7E-01	**0.91**	2.0E-01	esterase
P63208	S-phase kinase-associated protein 1			N.D.			3	**1.13**	6.3E-01	**1.18**	6.5E-01	ubiquitin-protein ligase
Q6ZMJ2	Scavenger receptor class A member 5			N.D.			1	**0.79**	-	**1.05**	-	transporter
A1L4H1	Scavenger receptor cysteine-rich domain-containing protein	3	**0.96**	5.4E-01	**0.97**	8.4E-01	3	**1.20**	1.8E-01	**1.03**	8.1E-01	receptor
Q86VB7	Scavenger receptor cysteine-rich type 1 protein M130	12	**1.24**	9.1E-04	**1.49**	6.2E-02	13	**1.28**	4.9E-01	**1.25**	5.3E-01	receptor
O95171	Sciellin			N.D.			1	**1.26**	-	**1.17**	-	N.A.
Q8WVN6	Secreted and transmembrane protein 1			N.D.			2	**1.03**	8.6E-01	**1.03**	8.9E-01	N.A.
Q92765	Secreted frizzled-related protein 3	4	**0.87**	1.1E-01	**0.95**	2.7E-01	2	**1.30**	3.5E-01	**1.26**	8.1E-02	signaling molecule
O95969	Secretoglobin family 1D member 2	3	**0.90**	4.7E-01	**0.82**	2.0E-01			N.D.			transfer/carrier protein
Q96QR1	Secretoglobin family 3A member 1	1	**1.26**	-	**0.82**	-			N.D.			N.A.
P13521	Secretogranin-2			N.D.			1	**0.34**	-	**0.45**	-	neuropeptide
Q13228	Selenium-binding protein 1	3	**0.83**	2.1E-01	**0.84**	3.3E-01	8	**1.01**	9.8E-01	**0.97**	9.3E-01	defense/immunity protein
P49908	Selenoprotein P	3	**0.90**	1.7E-02	**0.65**	3.6E-01	4	**0.79**	8.7E-01	**0.80**	8.9E-01	extracellular matrix glycoprotein
Q14563	Semaphorin-3A	1	**1.23**	7.2E-01	**1.11**	7.9E-01			N.D.			membrane-bound signaling molecule
Q9NPR2	Semaphorin-4B			N.D.			2	**1.32**	-	**1.24**	-	membrane-bound signaling molecule
O75326	Semaphorin-7A	4	**1.04**	8.8E-01	**0.95**	7.8E-01	1	**1.29**	6.8E-01	**1.13**	8.6E-01	membrane-bound signaling molecule
O95084	Serine protease 23			N.D.			2	**0.86**	6.3E-01	**0.86**	6.9E-01	N.A.
Q92743	Serine protease HTRA1	1	**0.85**	9.3E-02	**0.73**	2.1E-02	2	**1.18**	6.4E-01	**0.99**	7.4E-01	serine protease
Q9NQ38	Serine protease inhibitor Kazal-type 5	17	**0.78**	3.7E-04	**0.92**	6.8E-01	35	**1.19**	3.4E-01	**1.19**	3.6E-01	enzyme modulator
P58062	Serine protease inhibitor Kazal-type 7	1	**0.82**	-	**0.68**	-	2	**1.32**	3.3E-01	**1.04**	7.4E-01	serine protease inhibitor
Q9UPZ9	Serine/threonine-protein kinase ICK	1	**0.00**	-	**0.00**	-			N.D.			protein kinase
P02787	Serotransferrin	116	**1.08**	1.0E-02	**1.25**	7.5E-04	23	**1.79**	9.9E-02	**1.74**	1.4E-01	transfer/carrier protein
Q96P63	Serpin B12	2	**0.54**	-	**0.80**	-			N.D.			serine protease inhibitor
Q9UIV8	Serpin B13	1	**0.99**	9.7E-01	**0.90**	3.9E-01	3	**0.67**	5.1E-01	**0.58**	5.8E-01	serine protease inhibitor
P29508	Serpin B3	10	**1.03**	5.8E-01	**0.71**	1.4E-02	25	**1.12**	6.4E-01	**1.15**	6.0E-01	serine protease inhibitor
P48594	Serpin B4			N.D.			17	**1.38**	8.4E-02	**1.35**	2.3E-01	serine protease inhibitor
P36952	Serpin B5	1	**0.84**	-	**0.38**	-	1	**1.42**	-	**1.04**	-	serine protease inhibitor
P02768	Serum albumin	53	**1.45**	1.9E-12	**1.46**	6.4E-04	19	**1.76**	1.3E-02	**1.79**	1.7E-02	transfer/carrier protein
P35542	Serum amyloid A-4 protein			N.D.			2	**1.03**	6.0E-01	**1.03**	6.9E-01	transporter
P02743	Serum amyloid P-component	2	**1.56**	1.0E-01	**1.52**	4.2E-01	6	**0.75**	7.5E-01	**0.81**	8.2E-01	antibacterial response protein
P27169	Serum paraoxonase/arylesterase 1	2	**0.68**	3.8E-01	**0.69**	2.0E-01	5	**0.86**	5.2E-01	**1.07**	7.1E-01	peroxidase
P04278	Sex hormone-binding globulin	13	**0.95**	7.3E-01	**0.96**	7.2E-01	20	**0.74**	4.9E-01	**0.78**	6.3E-01	N.A.
O75368	SH3 domain-binding glutamic acid-rich-like protein	3	**2.08**	1.3E-02	**1.88**	1.4E-02	7	**0.75**	8.1E-02	**0.70**	7.8E-02	N.A.
Q9UJC5	SH3 domain-binding glutamic acid-rich-like protein 2	1	**0.75**	-	**0.93**	-			N.D.			N.A.
Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	2	**1.57**	1.1E-02	**1.52**	1.0E-01	6	**1.20**	8.7E-01	**1.46**	4.9E-01	N.A.
Q9NQ36	Signal peptide, CUB and EGF-like domain-containing protein 2	5	**0.89**	4.8E-03	**0.93**	4.4E-01	3	**0.96**	9.3E-01	**1.24**	5.4E-01	extracellular matrix glycoprotein
O00241	Signal-regulatory protein beta-1	2	**0.99**	9.6E-01	**1.18**	7.3E-01			N.D.			chemokine
P35326	Small proline-rich protein 2A	6	**1.01**	-	**0.80**	-			N.D.			structural protein
P35325	Small proline-rich protein 2B			N.D.			8	**0.99**	-	**0.67**	-	structural protein
Q9UBC9	Small proline-rich protein 3	40	**1.04**	5.6E-01	**1.17**	4.0E-01	85	**0.78**	3.8E-01	**0.72**	1.7E-01	structural protein
P11166	Solute carrier family 2, facilitated glucose transporter member 1	1	**1.11**	7.4E-01	**1.02**	4.8E-01			N.D.			carbohydrate transporter
Q8TDB8	Solute carrier family 2, facilitated glucose transporter member 14	2	**1.02**	-	**0.88**	-			N.D.			carbohydrate transporter
Q00796	Sorbitol dehydrogenase			N.D.			1	**1.23**	-	**0.93**	-	dehydrogenase
P09486	SPARC	21	**0.61**	1.9E-10	**0.71**	5.5E-02	21	**0.90**	7.5E-01	**0.83**	6.0E-01	growth factor
Q14515	SPARC-like protein 1	6	**0.80**	9.6E-03	**0.74**	3.5E-02	9	**0.92**	8.1E-01	**0.91**	7.9E-01	growth factor
Q13813	Spectrin alpha chain, brain	3	**1.46**	3.0E-02	**1.42**	2.8E-01	2	**0.61**	4.8E-01	**0.59**	5.2E-01	N.A.
Q9HCB6	Spondin-1	2	**1.02**	9.2E-01	**1.03**	8.8E-01	4	**1.07**	2.9E-01	**0.95**	6.6E-01	extracellular matrix glycoprotein
Q9BUD6	Spondin-2	5	**0.67**	1.5E-01	**0.70**	8.0E-02	6	**0.80**	4.3E-01	**0.85**	5.7E-01	extracellular matrix glycoprotein
P52823	Stanniocalcin-1	3	**1.26**	1.7E-01	**1.09**	7.7E-01			N.D.			transferase
O76061	Stanniocalcin-2	2	**0.94**	4.0E-01	**0.78**	3.0E-01	3	**0.77**	5.6E-01	**0.79**	1.1E-01	peptide hormone
Q9P2P6	StAR-related lipid transfer protein 9			N.D.			1	**1.13**	7.8E-01	**0.75**	5.4E-01	N.A.
P16949	Stathmin	1	**0.00**	-	**0.00**	-			N.D.			N.A.
P31948	Stress-induced-phosphoprotein 1	1	**1.58**	3.4E-01	**2.38**	4.1E-01	2	**2.80**	4.8E-02	**2.26**	2.6E-02	chaperone
P09238	Stromelysin-2			N.D.			2	**0.80**	5.9E-01	**0.77**	5.8E-01	metalloprotease
O00391	Sulfhydryl oxidase 1	18	**0.84**	5.0E-03	**0.81**	1.2E-01	30	**0.96**	8.8E-01	**0.93**	8.0E-01	oxidase
P00441	Superoxide dismutase [Cu-Zn]	2	**1.09**	7.1E-01	**1.16**	3.5E-01	5	**1.68**	4.5E-01	**2.03**	3.5E-01	oxidoreductase
P04179	Superoxide dismutase [Mn], mitochondrial			N.D.			4	**1.48**	5.4E-01	**1.89**	2.9E-01	oxidoreductase
Q6UWP8	Suprabasin	1	**0.67**	-	**0.89**	-	7	**0.92**	7.0E-01	**0.97**	8.7E-01	N.A.
Q9UGT4	Sushi domain-containing protein 2	2	**0.66**	1.9E-01	**0.60**	3.3E-01	1	**0.38**	1.0E-01	**0.41**	2.3E-02	N.A.
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	1	**1.88**	-	**2.85**	-			N.D.			dehydrogenase

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P18827	Syndecan-1	1	**1.45**	-	**0.96**	-			N.D.			membrane-bound signaling molecule
P31431	Syndecan-4	1	**0.99**	-	**0.74**	-			N.D.			membrane-bound signaling molecule
O15400	Syntaxin-7			N.D.			2	**1.04**	6.8E-01	**1.04**	4.5E-01	SNARE protein
O00560	Syntenin-1	3	**1.00**	9.8E-01	**0.85**	3.0E-02	5	**0.80**	4.4E-01	**0.64**	1.4E-01	membrane trafficking regulatory protein
Q9UHF0	Tachykinin-3	2	**0.65**	4.0E-01	**0.87**	6.0E-01	2	**1.62**	6.9E-01	**1.29**	8.2E-01	neuropeptide
Q9Y490	Talin-1	2	**1.03**	9.1E-01	**1.05**	8.8E-01	7	**1.09**	7.3E-01	**1.09**	6.5E-01	actin family cytoskeletal protein
Q7Z7G0	Target of Nesh-SH3	5	**0.90**	2.8E-01	**1.12**	5.6E-01	7	**0.99**	9.8E-01	**1.17**	8.5E-01	N.A.
P24821	Tenascin	5	**1.12**	3.1E-01	**1.03**	8.2E-01	4	**0.70**	3.0E-01	**0.60**	1.4E-01	signaling molecule
P22105	Tenascin-X	14	**1.17**	6.7E-02	**1.07**	4.7E-01	11	**0.95**	8.6E-01	**0.93**	8.0E-01	signaling molecule
P05452	Tetranectin	9	**1.03**	8.1E-01	**0.80**	8.4E-02	16	**0.74**	5.3E-01	**0.72**	4.2E-01	extracellular matrix structural protein
P37173	TGF-beta receptor type-2	1	**0.90**	3.3E-01	**1.04**	8.8E-01			N.D.			protein kinase
P10599	Thioredoxin	6	**1.13**	6.3E-01	**1.02**	9.3E-01	9	**0.94**	9.0E-01	**0.96**	9.4E-01	oxidoreductase
Q9BRA2	Thioredoxin domain-containing protein 17	2	**1.32**	1.9E-01	**1.38**	7.6E-02	2	**1.15**	1.9E-02	**1.06**	5.1E-01	transporter
Q8NBS9	Thioredoxin domain-containing protein 5	2	**1.09**	6.7E-01	**1.19**	1.8E-01	6	**1.16**	5.9E-01	**1.21**	2.3E-01	isomerase
P30048	Thioredoxin-dependent peroxide reductase, mitochondrial			N.D.			2	**1.27**	6.1E-02	**1.36**	9.9E-02	peroxidase
P07996	Thrombospondin-1	30	**0.74**	5.4E-07	**0.76**	7.3E-03	38	**0.97**	8.5E-01	**0.96**	8.3E-01	signaling molecule
P35442	Thrombospondin-2	2	**0.85**	6.3E-01	**1.03**	8.5E-01			N.D.			signaling molecule
P35443	Thrombospondin-4	3	**0.86**	-	**1.29**	-			N.D.			signaling molecule
P04216	Thy-1 membrane glycoprotein	3	**0.95**	5.2E-01	**0.96**	7.6E-01	4	**1.12**	8.2E-01	**1.12**	8.6E-01	membrane-bound signaling molecule
P19971	Thymidine phosphorylase	2	**1.72**	1.6E-01	**1.41**	3.5E-01			N.D.			glycosyltransferase
P63313	Thymosin beta-10	3	**1.32**	4.4E-01	**1.84**	9.4E-02	4	**1.18**	9.3E-01	**1.24**	9.2E-01	N.A.
P62328	Thymosin beta-4	9	**3.05**	1.1E-03	**4.63**	1.3E-01	16	**0.91**	9.1E-01	**1.14**	8.6E-01	N.A.
P05543	Thyroxine-binding globulin	19	**0.95**	5.1E-01	**0.79**	2.5E-02	20	**0.58**	3.1E-02	**0.59**	9.1E-02	serine protease inhibitor
P10646	Tissue factor pathway inhibitor	1	**0.99**	-	**0.76**	-			N.D.			serine protease inhibitor
Q8WZ42	Titin			N.D.			2	**1.14**	7.9E-01	**1.33**	5.8E-01	N.A.
P37837	Transaldolase	3	**3.81**	1.3E-02	**4.63**	7.9E-03	12	**1.13**	6.7E-01	**1.11**	7.8E-01	transaldolase
P20061	Transcobalamin-1	6	**1.00**	9.5E-01	**0.79**	3.8E-01	7	**0.61**	3.2E-01	**0.56**	1.8E-01	cation transporter
Q15582	Transforming growth factor-beta-induced protein ig-h3	20	**1.15**	6.1E-04	**1.26**	5.3E-02	19	**0.96**	8.7E-01	**0.99**	9.8E-01	signaling molecule
Q01995	Transgelin	3	**0.58**	2.7E-02	**0.53**	6.2E-03	16	**0.69**	2.9E-01	**0.69**	2.7E-01	non-motor actin binding protein
P37802	Transgelin-2	5	**2.04**	4.3E-04	**1.94**	9.1E-05	16	**1.18**	2.8E-01	**1.21**	2.2E-01	N.A.
P55072	Transitional endoplasmic reticulum ATPase	1	**2.16**	-	**1.87**	-	1	**2.18**	4.0E-01	**2.60**	1.0E-01	transmembrane receptor
P29401	Transketolase	7	**3.21**	1.8E-03	**4.13**	1.3E-03	17	**1.33**	5.1E-01	**1.55**	2.1E-01	transketolase
P13693	Translationally-controlled tumor protein			N.D.			2	**1.36**	5.8E-01	**1.19**	7.4E-01	non-motor microtubule binding protein
Q9UL52	Transmembrane protease serine 11E	1	**1.11**	-	**1.12**	-			N.D.			receptor
Q24JP5	Transmembrane protein 132A			N.D.			1	**1.00**	1.0E+00	**0.83**	8.7E-01	N.A.
P02766	Transthyretin	5	**0.94**	5.1E-01	**0.89**	4.3E-01	7	**1.38**	4.4E-01	**1.30**	4.8E-01	transporter
P04155	Trefoil factor 1	2	**0.97**	8.5E-01	**1.19**	5.6E-01	2	**1.07**	8.5E-01	**1.16**	6.5E-01	cytokine
Q03403	Trefoil factor 2	1	**2.03**	-	**2.93**	-	2	**0.50**	3.4E-01	**0.88**	4.8E-01	cytokine
Q07654	Trefoil factor 3	16	**1.00**	9.8E-01	**1.02**	9.2E-01	28	**1.00**	9.9E-01	**0.96**	9.2E-01	cytokine
Q07283	Trichohyalin	0	**0.80**	3.1E-03	**0.80**	2.8E-01			N.D.			N.A.
P60174	Triosephosphate isomerase	12	**1.89**	1.7E-06	**2.42**	3.0E-07	26	**1.48**	6.2E-04	**1.47**	4.8E-02	isomerase
O14773	Tripeptidyl-peptidase 1	1	**1.43**	-	**1.53**	-			N.D.			serine protease
Q9NYL9	Tropomodulin-3			N.D.			2	**1.46**	2.6E-01	**1.42**	1.9E-02	non-motor actin binding protein
P06753	Tropomyosin alpha-3 chain	5	**3.65**	5.6E-02	**7.85**	6.4E-02	9	**0.99**	9.5E-01	**1.04**	9.0E-01	actin binding motor protein
P67936	Tropomyosin alpha-4 chain	6	**1.54**	2.3E-01	**2.39**	9.1E-02	12	**1.07**	7.4E-01	**1.07**	8.8E-01	actin binding motor protein
P07477	Trypsin-1			N.D.			6	**0.00**	-	**0.00**	-	serine protease
P35030	Trypsin-3			N.D.			9	**1.32**	1.1E-02	**1.26**	2.0E-02	serine protease
P23381	Tryptophanyl-tRNA synthetase, cytoplasmic	3	**1.57**	1.1E-02	**1.57**	9.1E-02	2	**1.11**	7.1E-01	**0.77**	2.0E-01	aminoacyl-tRNA synthetase
Q9BQE3	Tubulin alpha-1C chain			N.D.			7	**0.93**	4.1E-01	**0.98**	9.0E-01	tubulin
P68366	Tubulin alpha-4A chain	2	**2.74**	4.6E-02	**2.22**	1.9E-01			N.D.			tubulin
P07437	Tubulin beta chain	4	**1.97**	-	**1.73**	-	6	**1.24**	1.3E-01	**1.12**	8.7E-01	N.A.
P68371	Tubulin beta-2C chain	4	**1.16**	-	**1.09**	-			N.D.			tubulin
Q9BW30	Tubulin polymerization-promoting protein family member 3	1	**1.12**	-	**1.27**	-			N.D.			non-motor microtubule binding protein
O75347	Tubulin-specific chaperone A	1	**1.19**	-	**1.00**	-	5	**0.93**	9.2E-01	**0.92**	8.7E-01	chaperonin
Q9NP84	Tumor necrosis factor receptor superfamily member 12A	2	**0.84**	6.9E-01	**0.59**	1.4E-01			N.D.			N.A.
P08138	Tumor necrosis factor receptor superfamily member 16	3	**0.93**	5.8E-01	**0.87**	4.3E-01	1	**0.53**	3.4E-01	**0.53**	3.5E-01	tumor necrosis factor receptor
P20333	Tumor necrosis factor receptor superfamily member 1B	1	**1.31**	1.7E-01	**3.69**	2.2E-01			N.D.			tumor necrosis factor receptor
O43399	Tumor protein D54	1	**1.82**	-	**0.86**	-			N.D.			N.A.
P09758	Tumor-associated calcium signal transducer 2			N.D.			1	**1.13**	-	**0.91**	-	receptor
Q9GZX9	Twisted gastrulation protein homolog 1	1	**0.86**	-	**1.04**	-	2	**1.10**	3.7E-01	**1.03**	8.6E-01	N.A.
P30530	Tyrosine-protein kinase receptor UFO	9	**0.98**	8.2E-01	**0.96**	5.1E-01	7	**1.16**	5.6E-01	**1.18**	3.9E-01	protein kinase
Q13308	Tyrosine-protein kinase-like 7	1	**1.25**	-	**0.98**	-			N.D.			protein kinase
P29350	Tyrosine-protein phosphatase non-receptor type 6			N.D.			1	**1.66**	-	**1.76**	-	receptor
P78324	Tyrosine-protein phosphatase non-receptor type substrate 1	4	**1.74**	1.0E-02	**2.02**	5.2E-03	2	**1.28**	3.5E-01	**1.19**	6.0E-01	chemokine
P54578	Ubiquitin carboxyl-terminal hydrolase 14	1	**1.61**	-	**1.21**	-	1	**1.10**	-	**1.11**	-	cysteine protease
Q96FW1	Ubiquitin thioesterase OTUB1	1	**1.87**	8.8E-02	**1.05**	8.9E-01	2	**0.72**	3.5E-01	**0.60**	3.8E-01	hydrolase
P61086	Ubiquitin-conjugating enzyme E2 K			N.D.			1	**0.99**	-	**0.54**	-	transfer/carrier protein
P68036	Ubiquitin-conjugating enzyme E2 L3			N.D.			2	**1.29**	4.8E-01	**1.26**	4.1E-01	ligase
Q13404	Ubiquitin-conjugating enzyme E2 variant 1			N.D.			4	**1.01**	8.8E-01	**1.01**	9.4E-01	N.A.
Q15819	Ubiquitin-conjugating enzyme E2 variant 2	1	**1.29**	-	**1.12**	-			N.D.			transfer/carrier protein
P22314	Ubiquitin-like modifier-activating enzyme 1	2	**2.05**	1.0E-02	**1.64**	2.5E-01	9	**1.16**	1.6E-01	**1.23**	1.9E-01	transfer/carrier protein
P05161	Ubiquitin-like protein ISG15	1	**1.04**	-	**0.90**	-	1	**1.09**	-	**1.22**	-	ribosomal protein
Q8NFL0	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7	0	**0.69**	-	**0.84**	-			N.D.			glycosyltransferase
Q9BY64	UDP-glucuronosyltransferase 2B28	1	**0.75**	-	**0.71**	-			N.D.			glycosyltransferase
Q5T011	Uncharacterized protein KIAA0467	1	**1.07**	1.7E-01	**1.01**	9.9E-01	2	**0.62**	6.9E-01	**0.85**	8.9E-01	N.A.
Q14146	Unhealthy ribosome biogenesis protein 2 homolog	1	**0.96**	-	**0.81**	-			N.D.			N.A.
Q9GZP4	UPF0424 protein C1orf128			N.D.			2	**1.24**	1.0E-01	**1.05**	6.6E-01	N.A.
Q969H8	UPF0556 protein C19orf10	3	**1.21**	6.4E-03	**1.37**	1.4E-02	3	**1.13**	2.7E-01	**1.14**	2.7E-02	N.A.
Q6P5S2	UPF0762 protein C6orf58	7	**0.90**	5.5E-02	**1.27**	4.3E-01	5	**1.40**	6.3E-01	**1.23**	8.6E-01	N.A.
Q6UX73	UPF0764 protein C16orf89			N.D.			1	**0.55**	3.7E-01	**0.50**	1.4E-02	N.A.
Q03405	Urokinase plasminogen activator surface receptor	1	**1.35**	-	**0.91**	-			N.D.			receptor
P07911	Uromodulin	4	**0.75**	2.5E-02	**0.75**	1.8E-03	6	**0.89**	6.7E-01	**0.83**	5.3E-01	receptor
P11684	Uteroglobin	8	**0.69**	9.1E-03	**0.91**	7.7E-01	8	**1.02**	9.5E-01	**1.03**	9.3E-01	cytokine
Q16851	UTP--glucose-1-phosphate uridylyltransferase	3	**2.12**	4.9E-03	**2.26**	4.5E-03	4	**1.40**	1.7E-02	**1.19**	4.7E-01	nucleotidyltransferase
P46939	Utrophin			N.D.			1	**1.29**	-	**0.75**	-	N.A.
P54727	UV excision repair protein RAD23 homolog B			N.D.			3	**0.87**	9.3E-01	**0.68**	7.0E-01	damaged DNA-binding protein
Q9Y279	V-set and immunoglobulin domain-containing protein 4	4	**0.89**	5.5E-01	**0.98**	8.8E-01	2	**1.22**	4.5E-01	**1.25**	3.3E-01	N.A.
P38606	V-type proton ATPase catalytic subunit A	1	**1.78**	-	**1.62**	-			N.D.			ATP synthase
Q7Z7G8	Vacuolar protein sorting-associated protein 13B			N.D.			2	**1.04**	6.6E-01	**1.13**	1.5E-01	N.A.
O75351	Vacuolar protein sorting-associated protein 4B	2	**1.08**	1.8E-01	**0.93**	5.7E-01			N.D.			hydrolase
P19320	Vascular cell adhesion protein 1	5	**1.01**	9.1E-01	**0.87**	3.8E-01	4	**1.55**	7.0E-01	**1.53**	6.5E-01	immunoglobulin receptor superfamily
P17948	Vascular endothelial growth factor receptor 1	23	**0.65**	1.2E-08	**0.70**	2.4E-02	31	**0.93**	8.4E-01	**0.90**	7.8E-01	protein kinase
P50552	Vasodilator-stimulated phosphoprotein	2	**5.46**	1.6E-02	**5.00**	3.9E-02	2	**1.72**	4.6E-02	**1.82**	1.5E-01	N.A.
Q6EMK4	Vasorin	6	**1.09**	3.3E-01	**1.21**	1.4E-01	8	**1.06**	2.9E-01	**1.03**	7.6E-01	receptor
P13611	Versican core protein	8	**1.28**	3.2E-02	**1.34**	6.8E-02	3	**1.83**	1.5E-01	**1.89**	4.3E-02	extracellular matrix glycoprotein
Q12907	Vesicular integral-membrane protein VIP36	9	**0.94**	1.9E-01	**0.85**	1.8E-01	12	**0.88**	5.9E-01	**0.86**	5.7E-01	membrane traffic protein
P08670	Vimentin	27	**3.15**	1.7E-12	**3.09**	2.1E-10	29	**1.54**	2.1E-01	**1.42**	3.9E-01	structural protein
P18206	Vinculin	6	**1.50**	4.9E-03	**1.37**	1.9E-01	14	**0.80**	3.3E-01	**0.76**	2.4E-01	non-motor actin binding protein
P02774	Vitamin D-binding protein	253	**0.83**	1.9E-27	**0.89**	1.5E-02	335	**0.81**	4.8E-02	**0.86**	1.2E-01	transfer/carrier protein
P07225	Vitamin K-dependent protein S	8	**0.88**	1.1E-01	**0.80**	7.0E-02	12	**0.82**	5.4E-01	**0.87**	7.1E-01	N.A.
P22891	Vitamin K-dependent protein Z	1	**0.00**	-	**0.00**	-	2	**1.05**	7.7E-01	**1.13**	4.1E-01	serine protease
Q7Z5L0	Vitelline membrane outer layer protein 1 homolog	1	**0.84**	-	**0.87**	-	1	**0.90**	3.9E-01	**0.78**	3.1E-01	N.A.

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	**Pos vs. Neg Replicate 1**	*p*-value	**Pos vs. Neg Replicate 2**	*p*-value	Peptides	**Pos vs. Neg Replicate 1**	*p*-value	**Pos vs. Neg Replicate 2**	*p*-value	
P04004	Vitronectin	51	**0.74**	1.7E-11	**0.70**	3.4E-06	80	**1.07**	6.7E-01	**1.05**	7.7E-01	cell adhesion molecule
O00555	Voltage-dependent P/Q-type calcium channel subunit alpha-1A			N.D.			1	**0.91**	8.0E-01	**0.83**	7.1E-01	calcium channel
P04275	von Willebrand factor			N.D.			3	**1.48**	4.7E-01	**1.50**	4.0E-01	signaling molecule
Q14508	WAP four-disulfide core domain protein 2	13	**0.92**	1.8E-01	**0.92**	4.7E-01	10	**1.28**	2.8E-01	**1.20**	4.4E-01	serine protease inhibitor
O75083	WD repeat-containing protein 1	3	**1.68**	1.7E-01	**1.82**	1.1E-01	3	**1.00**	9.9E-01	**0.97**	9.6E-01	non-motor actin binding protein
P13010	X-ray repair cross-complementing protein 5	1	**2.82**	-	**2.59**	-			N.D.			DNA helicase
P12955	Xaa-Pro dipeptidase	1	**1.49**	-	**0.86**	-	1	**0.77**	-	**1.03**	-	transcription factor
P15822	Zinc finger protein 40			N.D.			1	**1.62**	-	**1.78**	-	zinc finger transcription factor
P25311	Zinc-alpha-2-glycoprotein	36	**0.79**	5.2E-07	**0.71**	2.5E-03	31	**0.99**	9.8E-01	**0.91**	8.3E-01	immunoglobulin receptor superfamily
Q96DA0	Zymogen granule protein 16 homolog B	10	**1.07**	5.4E-01	**0.67**	1.1E-01	14	**0.57**	5.7E-01	**0.54**	3.7E-01	N.A.
Q15942	Zyxin			N.D.			2	**1.28**	5.4E-01	**1.38**	4.7E-01	kinase modulator

EXTERNAL TABLES
TABLE 9

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P62258	14-3-3 protein epsilon	9	**1.57**	6.3E-05	**1.47**	2.5E-02	12	**0.90**	9.2E-01	**0.86**	8.4E-01	chaperone
P63104	14-3-3 protein zeta/delta	14	**1.66**	1.8E-05	**1.58**	6.0E-03	24	**0.90**	7.0E-01	**0.90**	7.3E-01	chaperone
P52209	6-phosphogluconate dehydrogenase, decarboxylating	11	**3.43**	6.8E-08	**4.13**	1.6E-06	10	**1.20**	5.8E-01	**1.21**	6.1E-01	dehydrogenase
O95336	6-phosphogluconolactonase	2	**1.90**	1.2E-01	**1.68**	1.1E-02	4	**1.18**	1.3E-01	**0.99**	8.6E-01	hydrolase
P59998	Actin-related protein 2/3 complex subunit 4	4	**3.64**	1.0E-02	**3.98**	6.9E-04	5	**1.15**	5.0E-01	**1.16**	6.8E-01	N.A.
P61158	Actin-related protein 3	4	**3.79**	1.6E-04	**3.42**	5.1E-03	2	**1.48**	4.1E-02	**1.05**	8.7E-01	actin and actin related protein
P62736	Actin, aortic smooth muscle	33	**2.38**	1.4E-01	**2.55**	1.7E-02			N.D.			actin and actin related protein
P60709	Actin, cytoplasmic 1	70	**3.47**	4.2E-12	**3.47**	3.8E-09	88	**1.49**	3.1E-07	**1.36**	4.7E-05	actin and actin related protein
P63261	Actin, cytoplasmic 2	64	**2.09**	7.9E-03	**2.12**	3.6E-03			N.D.			actin and actin related protein
Q01518	Adenylyl cyclase-associated protein 1	7	**4.41**	8.3E-04	**3.56**	2.3E-03	16	**1.26**	1.2E-01	**1.24**	2.2E-01	actin family cytoskeletal protein
Q15848	Adiponectin			N.D.			2	**0.60**	1.4E-02	**0.72**	8.7E-02	peptide hormone
Q10588	ADP-ribosyl cyclase 2	1	**1.72**	2.5E-02	**1.41**	7.6E-02	3	**1.33**	1.5E-01	**1.04**	7.5E-01	cyclase
P01009	Alpha-1-antitrypsin	183	**0.62**	0.0E+00	**0.67**	8.1E-13	39	**0.97**	9.2E-01	**0.95**	8.7E-01	serine protease inhibitor
P08697	Alpha-2-antiplasmin	26	**0.86**	1.8E-02	**0.66**	5.6E-03	38	**0.91**	8.1E-01	**0.88**	7.5E-01	serine protease inhibitor
P12814	Alpha-actinin-1	19	**3.13**	4.8E-04	**2.87**	2.4E-03	29	**0.93**	7.7E-01	**0.98**	9.4E-01	N.A.
O43707	Alpha-actinin-4	19	**1.76**	1.1E-08	**1.63**	3.2E-03	31	**1.16**	7.3E-01	**1.09**	8.6E-01	N.A.
P04745	Alpha-amylase 1	6	**0.79**	1.7E-01	**0.64**	3.9E-02			N.D.			N.A.
P06733	Alpha-enolase	20	**3.14**	1.6E-07	**3.03**	3.1E-06	31	**1.42**	2.5E-01	**1.31**	4.3E-01	lyase
P02771	Alpha-fetoprotein	31	**1.65**	2.2E-11	**1.64**	1.2E-05	26	**0.93**	7.2E-01	**0.98**	9.3E-01	transfer/carrier protein
P15144	Aminopeptidase N	14	**2.48**	6.8E-09	**2.29**	1.0E-03	10	**0.88**	4.9E-01	**0.86**	3.6E-01	metalloprotease
Q9UKU9	Angiopoietin-related protein 2	6	**0.93**	3.0E-01	**0.80**	1.2E-02	3	**0.73**	1.9E-01	**0.55**	2.2E-01	signaling molecule
P12429	Annexin A3	3	**1.96**	2.5E-03	**1.96**	8.7E-02			N.D.			transfer/carrier protein
P03973	Antileukoproteinase	10	**0.89**	5.5E-02	**0.94**	7.7E-01	28	**2.06**	2.3E-04	**1.93**	4.4E-04	serine protease inhibitor
P20160	Azurocidin	3	**6.57**	5.5E-04	**6.70**	2.5E-02	1	**1.30**	2.1E-01	**1.26**	2.2E-01	serine protease
Q8N4F0	Bactericidal/permeability-increasing protein-like 1	9	**0.86**	1.5E-02	**0.50**	1.1E-02	8	**0.68**	5.0E-01	**0.57**	4.7E-01	carbohydrate transporter
P06865	Beta-hexosaminidase subunit alpha			N.D.			3	**0.38**	2.7E-02	**0.34**	4.4E-02	glycosidase
P13727	Bone marrow proteoglycan	37	**0.66**	7.4E-12	**0.61**	4.1E-17	39	**0.96**	6.7E-01	**1.03**	7.2E-01	extracellular matrix structural protein
P80723	Brain acid soluble protein 1	3	**4.39**	7.0E-03	**6.77**	1.7E-02	6	**1.41**	1.8E-01	**1.37**	1.5E-01	N.A.
P62158	Calmodulin	9	**1.47**	1.3E-03	**1.91**	4.0E-03	8	**0.71**	3.9E-02	**0.76**	1.6E-01	calmodulin
P27797	Calreticulin	10	**1.49**	1.5E-08	**1.57**	1.7E-05	13	**1.02**	9.6E-01	**1.08**	8.7E-01	calcium-binding protein
P00915	Carbonic anhydrase 1	9	**0.62**	1.8E-03	**0.57**	2.2E-04	48	**0.59**	5.6E-02	**0.65**	5.9E-02	dehydratase
P04040	Catalase	12	**1.52**	4.2E-03	**1.40**	1.2E-02	26	**1.12**	7.3E-01	**1.17**	5.8E-01	peroxidase
P49913	Cathelicidin antimicrobial peptide	4	**5.93**	2.0E-03	**6.96**	4.0E-03	10	**0.83**	2.8E-01	**0.80**	3.1E-01	protease inhibitor
P07858	Cathepsin B	8	**1.89**	1.4E-04	**1.80**	1.0E-03	7	**0.94**	9.2E-01	**0.88**	8.2E-01	cysteine protease
P07339	Cathepsin D	3	**2.10**	9.4E-04	**1.89**	1.7E-01	5	**0.92**	8.9E-01	**1.04**	9.4E-01	aspartic protease
P09668	Cathepsin H	4	**1.73**	3.6E-03	**1.87**	8.2E-02	6	**0.73**	9.3E-02	**0.76**	6.4E-02	cysteine protease
P36222	Chitinase-3-like protein 1	13	**3.46**	1.0E-10	**3.28**	6.8E-07	11	**0.92**	6.4E-01	**0.80**	4.0E-01	glycosidase
Q13231	Chitotriosidase-1	2	**6.63**	6.5E-02	**5.46**	3.3E-02	3	**1.40**	2.3E-01	**1.69**	1.6E-01	glycosidase
O00299	Chloride intracellular channel protein 1	3	**1.93**	2.9E-02	**2.30**	3.8E-03	3	**1.49**	4.0E-02	**1.12**	6.3E-01	N.A.
P00740	Coagulation factor IX	3	**0.68**	4.9E-02	**0.58**	8.9E-02	3	**0.75**	3.0E-01	**0.90**	4.4E-01	serine protease
P23528	Cofilin-1	4	**3.49**	1.2E-02	**3.31**	1.7E-02	14	**1.32**	5.1E-01	**1.30**	5.5E-01	non-motor actin binding protein
P12107	Collagen alpha-1(XI) chain	3	**0.67**	9.9E-02	**0.64**	4.7E-02			N.D.			peptide hormone
P01024	Complement C3	52	**1.54**	3.9E-12	**1.62**	1.3E-05	111	**1.08**	5.8E-01	**1.06**	7.4E-01	cytokine
P01031	Complement C5	31	**0.91**	1.6E-02	**0.86**	5.3E-02	23	**1.79**	1.3E-01	**1.56**	4.2E-01	cytokine
P07358	Complement component C8 beta chain	20	**0.73**	1.9E-05	**0.71**	7.7E-05	17	**0.65**	2.4E-03	**0.62**	3.9E-03	receptor
P29279	Connective tissue growth factor	8	**0.66**	8.9E-04	**0.82**	2.0E-01	5	**0.74**	6.8E-01	**0.80**	7.8E-01	growth factor
P31146	Coronin-1A	8	**3.72**	1.9E-04	**4.64**	1.9E-04	5	**1.21**	2.7E-01	**1.33**	5.6E-01	non-motor actin binding protein
P09228	Cystatin-SA	5	**1.13**	3.0E-01	**1.10**	2.1E-01	5	**0.39**	4.5E-03	**0.38**	1.2E-02	cysteine protease inhibitor
P32320	Cytidine deaminase	2	**3.69**	7.4E-03	**3.14**	1.5E-01			N.D.			deaminase
P07585	Decorin	7	**0.39**	3.3E-06	**0.36**	7.0E-05	4	**1.12**	8.4E-01	**1.20**	7.4E-01	receptor
Q9UGM3	Deleted in malignant brain tumors 1 protein	24	**1.53**	8.8E-06	**1.39**	1.1E-02	14	**1.13**	3.5E-01	**1.23**	2.1E-01	receptor
Q07507	Dermatopontin	3	**0.75**	4.2E-02	**0.66**	1.7E-01	1	**0.98**	9.4E-01	**1.05**	7.7E-01	extracellular matrix protein
P81605	Dermcidin	22	**0.24**	3.3E-12	**0.27**	5.0E-12	30	**0.60**	3.1E-01	**0.49**	1.2E-01	N.A.
O43184	Disintegrin and metalloproteinase domain-containing protein 12	1	**0.96**		**1.32**		7	**0.57**	1.8E-02	**0.60**	7.0E-02	metalloprotease
Q8TDJ6	DmX-like protein 2			N.D.			2	**1.92**	6.1E-04	**1.97**	5.6E-03	N.A.
Q96C19	EF-hand domain-containing protein D2	2	**8.20**		**5.99**		5	**0.46**	3.0E-02	**0.56**	8.5E-02	calcium-binding protein
P19957	Elafin	3	**1.47**	5.8E-03	**1.73**	3.5E-01	4	**2.17**	3.7E-01	**1.80**	5.6E-01	serine protease inhibitor
P02794	Ferritin heavy chain	2	**3.21**	1.8E-02	**2.02**	7.5E-02			N.D.			storage protein
P02675	Fibrinogen beta chain	7	**1.75**	6.3E-05	**1.77**	2.4E-03	63	**0.52**	2.2E-02	**0.50**	1.1E-02	signaling molecule
P02679	Fibrinogen gamma chain	3	**1.99**	5.4E-03	**2.30**	8.5E-03	31	**0.64**	3.3E-02	**0.65**	2.7E-02	signaling molecule
P20930	Filaggrin	3	**0.62**	2.1E-01	**0.49**	2.9E-03	9	**0.70**	1.9E-02	**0.71**	2.5E-02	cytoskeletal protein
P21333	Filamin-A	11	**2.21**	1.5E-04	**1.99**	2.9E-03	22	**1.00**	9.8E-01	**0.97**	8.7E-01	N.A.
P04075	Fructose-bisphosphate aldolase A	24	**2.11**	4.1E-12	**2.08**	5.4E-06	29	**1.17**	6.7E-01	**1.06**	8.6E-01	aldolase
O60234	Glia maturation factor gamma	2	**4.45**	2.3E-01	**5.31**	2.0E-01	2	**1.58**	1.2E-01	**2.04**	2.8E-02	signaling molecule
P07093	Glia-derived nexin	21	**0.57**	1.5E-06	**0.64**	7.1E-03	16	**1.32**	4.2E-01	**1.22**	5.7E-01	serine protease inhibitor
P11413	Glucose-6-phosphate 1-dehydrogenase	2	**3.09**	2.4E-02	**3.02**	1.8E-02	6	**1.54**	2.1E-01	**1.56**	1.8E-01	dehydrogenase
P06744	Glucose-6-phosphate isomerase	9	**4.20**	2.0E-06	**4.30**	3.1E-05	12	**0.79**	2.9E-01	**0.92**	7.2E-01	isomerase
P78417	Glutathione S-transferase omega-1	5	**1.76**	1.5E-03	**1.42**	2.2E-01	5	**1.20**	1.9E-01	**1.22**	3.6E-01	anion channel
P09211	Glutathione S-transferase P	10	**1.55**	8.2E-05	**1.72**	1.8E-02	16	**1.34**	3.3E-01	**1.36**	3.4E-01	transferase
P04406	Glyceraldehyde-3-phosphate dehydrogenase	9	**2.16**	1.2E-05	**2.07**	9.3E-05	13	**1.16**	6.4E-01	**1.17**	7.2E-01	dehydrogenase
P09466	Glycodelin	14	**1.31**	1.7E-04	**1.37**	1.6E-06	22	**1.70**	1.1E-02	**1.72**	7.4E-03	transfer/carrier protein
P06737	Glycogen phosphorylase, liver form	8	**3.49**	1.3E-05	**3.55**	1.9E-04	3	**1.31**	2.3E-01	**1.37**	1.6E-01	phosphorylase
P46976	Glycogenin-1	1	**3.44**	6.5E-02	**4.05**	7.9E-03			N.D.			glycosyltransferase
P01215	Glycoprotein hormones alpha chain	3	**0.75**	6.9E-02	**0.64**	2.6E-02	3	**0.37**	2.2E-03	**0.37**	4.7E-04	peptide hormone
Q9HC38	Glyoxalase domain-containing protein 4	1	**1.65**	2.1E-01	**1.43**	1.1E-02	3	**0.81**	8.5E-01	**0.86**	9.0E-01	lyase
P28799	Granulins	4	**1.48**	9.6E-03	**1.67**	2.8E-02	3	**1.02**	8.7E-01	**0.94**	4.8E-01	cytokine
Q14393	Growth arrest-specific protein 6	3	**0.72**	3.4E-02	**0.65**	1.2E-01			N.D.			signaling molecule
P08107	Heat shock 70 kDa protein 1A/1B	8	**2.15**	2.8E-06	**1.79**	3.5E-05	25	**1.17**	7.6E-01	**1.03**	9.5E-01	N.A.
P11142	Heat shock cognate 71 kDa protein	6	**2.15**	2.7E-02	**2.29**	6.8E-02	20	**1.15**	8.3E-01	**1.06**	9.1E-01	Hsp70 family chaperone
P07900	Heat shock protein HSP 90-alpha	10	**2.23**	3.0E-04	**2.50**	1.5E-04	7	**0.98**	9.4E-01	**0.99**	9.7E-01	Hsp90 family chaperone
P69905	Hemoglobin subunit alpha	51	**0.75**	6.7E-06	**0.57**	5.5E-07	199	**1.30**	6.2E-03	**1.22**	2.4E-02	transfer/carrier protein
P68871	Hemoglobin subunit beta	59	**0.82**	2.8E-03	**0.52**	2.0E-06	262	**0.93**	4.7E-01	**0.87**	1.6E-01	N.A.
P69892	Hemoglobin subunit gamma-2	9	**0.38**	5.6E-02	**0.28**	1.4E-02	23	**1.68**	3.3E-01	**1.83**	2.0E-01	transfer/carrier protein
P52790	Hexokinase-3	4	**3.05**	4.0E-03	**3.38**	8.4E-03			N.D.			carbohydrate kinase
P16403	Histone H1.2	2	**4.49**	9.0E-02	**4.18**	2.0E-02	2	**1.18**		**1.00**		histone
Q99880	Histone H2B type 1-L	13	**7.16**	2.1E-06	**6.88**	2.2E-05	3	**4.12**	2.9E-01	**2.32**	1.5E-01	histone
Q71DI3	Histone H3.2	5	**8.17**	1.8E-05	**8.26**	2.2E-06			N.D.			N.A.
P62805	Histone H4	12	**8.43**	2.6E-09	**9.03**	1.2E-05	10	**2.10**	2.7E-02	**1.98**	1.8E-01	N.A.
Q14520	Hyaluronan-binding protein 2	8	**0.88**	1.1E-03	**0.66**	1.0E-03	11	**0.74**	4.1E-01	**0.67**	2.9E-01	growth factor
P0CG06	Ig lambda-3 chain C regions	1	**1.77**	1.2E-02	**2.03**	1.3E-01			N.D.			N.A.
P01871	Ig mu chain C region			N.D.			3	**1.22**	1.0E-01	**1.51**	4.3E-03	immunoglobulin
Q9Y6R7	IgGFc-binding protein	34	**2.31**	1.8E-16	**2.28**	1.6E-13	36	**1.09**	4.3E-01	**1.16**	2.4E-01	signaling molecule
O75874	Isocitrate dehydrogenase [NADP] cytoplasmic	2	**1.39**	2.9E-01	**1.94**	1.5E-02	6	**0.93**	8.4E-01	**0.96**	9.3E-01	dehydrogenase
P29622	Kallistatin	14	**0.85**	7.4E-03	**0.72**	3.0E-03	21	**0.57**	1.6E-01	**0.59**	2.3E-01	serine protease inhibitor
P13645	Keratin, type I cytoskeletal 10	4	**0.73**	2.0E-01	**0.72**	1.3E-02	12	**0.32**	3.4E-02	**2.38**	2.2E-02	structural protein
P13646	Keratin, type I cytoskeletal 13	15	**1.24**	8.7E-03	**1.23**	2.1E-01	41	**1.84**	2.5E-02	**1.83**	5.6E-02	structural protein
P02533	Keratin, type I cytoskeletal 14			N.D.			5	**1.37**	5.8E-02	**1.60**	2.1E-02	structural protein
P04264	Keratin, type II cytoskeletal 1	4	**0.91**	7.5E-01	**1.08**	6.7E-01	19	**0.46**	5.7E-03	**4.09**	6.2E-04	structural protein
P35908	Keratin, type II cytoskeletal 2 epidermal			N.D.			9	**0.27**	1.0E-01	**3.89**	2.8E-03	structural protein
P02538	Keratin, type II cytoskeletal 6A	16	**1.80**	4.9E-02	**1.69**	5.8E-02	44	**1.54**	4.1E-02	**1.49**	6.4E-02	structural protein
P08729	Keratin, type II cytoskeletal 7	5	**1.48**	4.2E-02	**1.54**	5.3E-02	8	**1.53**	5.2E-01	**1.27**	7.4E-01	structural protein
P05787	Keratin, type II cytoskeletal 8	10	**1.42**	7.2E-03	**1.52**	5.8E-03	18	**1.00**	1.0E+00	**0.98**	9.7E-01	structural protein

UniProt Accession #	Protein name	PPROM cohort					sPTL cohort					PANTHER Protein Class
		Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	Peptides	Pos vs. Neg Replicate 1	*p*-value	Pos vs. Neg Replicate 2	*p*-value	
P00338	L-lactate dehydrogenase A chain	13	**2.16**	7.7E-08	**1.78**	8.0E-05	11	**0.89**	6.7E-01	**0.86**	6.6E-01	dehydrogenase
P07195	L-lactate dehydrogenase B chain	3	**1.75**	4.3E-02	**1.08**	7.2E-01	6	**1.05**	9.3E-01	**1.07**	9.0E-01	dehydrogenase
P14151	L-selectin	3	**1.36**	1.2E-01	**1.50**	5.7E-03	3	**0.97**	7.4E-01	**0.88**	2.5E-01	apolipoprotein
P02788	Lactotransferrin	62	**2.46**	9.4E-38	**2.28**	3.4E-21	60	**0.76**	4.4E-02	**0.75**	5.7E-02	transfer/carrier protein
P30740	Leukocyte elastase inhibitor	11	**2.18**	6.8E-06	**1.69**	1.1E-03	15	**1.38**	9.3E-02	**1.44**	9.1E-02	serine protease inhibitor
Q8N6C8	Leukocyte immunoglobulin-like receptor subfamily A member 3	4	**2.77**	2.7E-03	**3.14**	5.1E-05			N.D.			membrane-bound signaling molecule
P09960	Leukotriene A-4 hydrolase	8	**3.70**	1.0E-05	**4.09**	2.2E-04	6	**1.87**	4.2E-01	**1.91**	5.3E-01	metalloprotease
P31025	Lipocalin-1	11	**0.81**	1.8E-01	**0.53**	2.1E-03	20	**0.33**	1.7E-03	**0.37**	1.7E-03	transfer/carrier protein
P59827	Long palate, lung and nasal epithelium carcinoma-associated protein 4	2	**0.79**	2.1E-01	**0.61**	1.2E-01	1	**0.71**	4.4E-02	**0.60**	1.8E-01	carbohydrate transporter
P08637	Low affinity immunoglobulin gamma Fc region receptor III-A	2	**3.38**	1.3E-02	**3.61**	2.5E-04	3	**1.74**	1.7E-01	**1.57**	2.0E-01	immunoglobulin receptor superfamily
O95274	Ly6/PLAUR domain-containing protein 3	6	**0.80**	2.4E-02	**0.95**	9.1E-01	7	**0.72**	6.2E-01	**0.65**	5.5E-01	N.A.
P40121	Macrophage-capping protein	5	**2.64**	1.4E-03	**2.56**	1.6E-03	8	**1.44**	2.8E-01	**1.17**	7.0E-01	non-motor actin binding protein
P40925	Malate dehydrogenase, cytoplasmic	3	**1.62**	9.4E-03	**1.51**	1.3E-01	9	**1.52**	3.7E-01	**1.48**	3.9E-01	dehydrogenase
P14780	Matrix metalloproteinase-9	14	**4.60**	2.4E-06	**4.75**	1.2E-05	25	**1.13**	5.4E-01	**1.01**	9.7E-01	metalloprotease
P26038	Moesin	17	**3.97**	2.6E-05	**3.27**	1.4E-06	30	**1.56**	2.1E-01	**1.40**	3.6E-01	actin family cytoskeletal protein
P05164	Myeloperoxidase	34	**5.11**	3.5E-16	**6.11**	1.5E-14	30	**1.18**	3.7E-01	**1.18**	3.6E-01	peroxidase
P60660	Myosin light polypeptide 6	5	**1.93**	6.4E-03	**1.94**	1.6E-02	5	**1.62**	2.4E-01	**1.77**	3.5E-01	actin family cytoskeletal protein
P35579	Myosin-9	12	**2.26**	3.3E-04	**2.01**	3.4E-04	9	**1.25**	6.4E-01	**1.28**	6.4E-01	G-protein modulator
P22894	Neutrophil collagenase	9	**4.14**	1.5E-02	**4.60**	1.4E-02	11	**1.23**	3.4E-01	**1.35**	3.2E-01	metalloprotease
P59665	Neutrophil defensin 1	8	**4.28**	1.8E-08	**4.30**	4.9E-08			N.D.			N.A.
P59666	Neutrophil defensin 3			N.D.			10	**1.89**	9.5E-03	**1.84**	5.1E-02	N.A.
P80188	Neutrophil gelatinase-associated lipocalin	20	**4.54**	3.8E-18	**5.28**	9.3E-13	44	**1.09**	7.6E-01	**1.15**	5.8E-01	transfer/carrier protein
P43490	Nicotinamide phosphoribosyltransferase	3	**6.94**	6.0E-03	**4.80**	7.4E-02	10	**1.37**	6.1E-02	**1.39**	1.2E-01	cytokine
P22392	Nucleoside diphosphate kinase B	3	**1.91**	5.0E-03	**1.62**	1.3E-01	4	**0.73**	4.8E-01	**0.77**	4.8E-01	N.A.
Q92882	Osteoclast-stimulating factor 1	2	**6.68**	5.3E-02	**6.99**	9.0E-03			N.D.			N.A.
O75594	Peptidoglycan recognition protein 1	5	**4.33**	6.7E-07	**4.34**	1.0E-06	5	**1.49**	4.7E-02	**1.49**	3.5E-02	signaling molecule
P62937	Peptidyl-prolyl cis-trans isomerase A	11	**2.19**	1.9E-04	**1.99**	3.6E-02	21	**0.91**	8.8E-01	**0.92**	8.9E-01	N.A.
P62942	Peptidyl-prolyl cis-trans isomerase FKBP1A	2	**1.69**	2.0E-01	**1.45**	6.0E-01	3	**1.18**	4.5E-02	**0.99**	8.7E-01	isomerase
P30044	Peroxiredoxin-5, mitochondrial	2	**2.34**	2.5E-02	**1.73**	1.5E-01	3	**0.71**	3.8E-01	**0.59**	2.8E-01	peroxidase
P36871	Phosphoglucomutase-1	1	**2.90**	4.2E-02	**2.13**	4.0E-03	3	**1.26**	7.0E-02	**1.07**	8.4E-01	mutase
P00558	Phosphoglycerate kinase 1	14	**2.32**	3.0E-08	**2.39**	2.3E-04	26	**1.07**	8.4E-01	**1.09**	8.1E-01	carbohydrate kinase
P18669	Phosphoglycerate mutase 1	10	**2.43**	1.5E-04	**2.48**	2.0E-03	15	**0.95**	9.0E-01	**1.11**	7.7E-01	mutase
P55058	Phospholipid transfer protein	8	**1.30**	1.1E-05	**0.74**	1.3E-01	7	**0.70**	3.9E-01	**0.54**	1.1E-01	transfer/carrier protein
P05121	Plasminogen activator inhibitor 1	17	**0.87**	4.4E-02	**0.83**	2.7E-01	18	**0.60**	1.0E-01	**0.57**	5.7E-02	serine protease inhibitor
P13796	Plastin-2	35	**4.39**	1.6E-21	**4.17**	2.6E-15	40	**1.08**	7.0E-01	**1.00**	9.9E-01	non-motor actin binding protein
P0CG48	Polyubiquitin-C	3	**1.06**	7.0E-01	**1.00**	1.0E+00	6	**0.25**	4.9E-04	**0.26**	1.4E-03	N.A.
Q00888	Pregnancy-specific beta-1-glycoprotein 4	12	**0.48**	9.4E-03	**0.73**	4.2E-02	15	**1.33**	4.2E-01	**1.43**	2.9E-01	N.A.
Q00887	Pregnancy-specific beta-1-glycoprotein 9	12	**0.57**	1.8E-03	**1.08**	7.7E-01	15	**1.06**	5.3E-01	**1.01**	9.2E-01	immunoglobulin cell adhesion molecule
P07602	Proactivator polypeptide	20	**1.57**	2.3E-08	**2.04**	1.5E-07	14	**0.97**	8.7E-01	**0.85**	3.3E-01	enzyme modulator
P07737	Profilin-1	11	**3.37**	3.3E-08	**3.43**	3.7E-09	22	**1.01**	9.8E-01	**0.98**	9.6E-01	N.A.
P12273	Prolactin-inducible protein	4	**0.74**	2.6E-02	**0.62**	6.3E-02	1	**0.55**	6.6E-01	**0.41**	5.0E-01	N.A.
Q16378	Proline-rich protein 4	6	**0.67**	7.5E-02	**0.63**	3.3E-02	8	**0.80**	5.0E-02	**0.76**	9.2E-03	N.A.
Q9UL46	Proteasome activator complex subunit 2	2	**2.68**	7.9E-03	**2.09**	2.9E-01	1	**1.67**	1.6E-01	**1.45**	7.9E-02	N.A.
Q99497	Protein DJ-1	5	**1.63**	8.0E-03	**1.92**	1.4E-01	7	**1.45**	4.4E-01	**1.52**	3.9E-01	transcription factor
Q6P988	Protein notum homolog	12	**0.74**	5.0E-06	**0.64**	2.4E-02	17	**1.02**	9.2E-01	**1.11**	5.9E-01	glycosidase
Q9NP55	Protein Plunc	5	**1.33**	1.7E-01	**0.29**	3.0E-03	2	**0.28**	1.8E-01	**0.16**	1.8E-01	carbohydrate transporter
P31949	Protein S100-A11	4	**2.60**	8.4E-04	**2.90**	6.3E-04	6	**1.86**	3.9E-01	**1.93**	3.7E-01	signaling molecule
P05109	Protein S100-A8	25	**3.17**	1.1E-18	**2.89**	1.7E-09	35	**1.17**	6.5E-01	**1.05**	9.1E-01	signaling molecule
P06702	Protein S100-A9	44	**3.86**	7.5E-33	**3.14**	1.5E-27	48	**1.18**	2.1E-01	**1.16**	3.6E-01	signaling molecule
Q8IWL2	Pulmonary surfactant-associated protein A1	10	**1.39**	3.1E-04	**1.39**	1.9E-03	10	**0.59**	6.0E-02	**0.60**	2.4E-01	N.A.
P00491	Purine nucleoside phosphorylase	4	**1.98**	2.0E-04	**1.47**	8.5E-03	5	**0.72**	5.3E-01	**0.84**	6.9E-01	phosphorylase
P14618	Pyruvate kinase isozymes M1/M2	11	**2.28**	3.3E-06	**2.18**	4.1E-06	11	**1.07**	4.1E-01	**0.97**	8.0E-01	carbohydrate kinase
P50395	Rab GDP dissociation inhibitor beta	8	**2.34**	4.0E-05	**2.01**	4.9E-02	16	**0.70**	3.3E-01	**0.67**	2.1E-01	acyltransferase
P46940	Ras GTPase-activating-like protein IQGAP1	3	**3.05**	3.7E-02	**1.98**	9.8E-02	2	**0.79**	6.6E-01	**0.99**	8.9E-01	G-protein modulator
Q9HD89	Resistin	2	**5.73**	1.2E-01	**7.99**	1.1E-01	1	**1.82**	1.3E-01	**1.84**	3.8E-02	N.A.
Q8NFJ5	Retinoic acid-induced protein 3	7	**0.82**	2.0E-02	**0.54**	1.3E-02	8	**0.50**	2.7E-01	**0.54**	1.7E-01	G-protein coupled receptor
P52565	Rho GDP-dissociation inhibitor 1	5	**2.58**	5.2E-06	**2.71**	1.2E-05	7	**1.07**	8.0E-01	**1.08**	7.4E-01	signaling molecule
P52566	Rho GDP-dissociation inhibitor 2	2	**7.45**	1.1E-02	**8.02**	1.4E-02	10	**0.66**	6.8E-01	**0.64**	6.9E-01	signaling molecule
P34096	Ribonuclease 4	1	**0.87**	6.4E-01	**0.65**	4.7E-01	4	**0.77**	2.8E-02	**0.82**	1.9E-01	endoribonuclease
P13489	Ribonuclease inhibitor	3	**1.73**	7.2E-03	**1.88**	5.8E-02	2	**0.98**	9.7E-01	**0.98**	9.4E-01	enzyme modulator
P49908	Selenoprotein P	3	**0.90**	1.7E-02	**0.65**	3.6E-01	4	**0.79**	8.7E-01	**0.80**	8.9E-01	extracellular matrix glycoprotein
P02787	Serotransferrin	116	**1.08**	1.0E-02	**1.25**	7.5E-04	23	**1.79**	9.9E-02	**1.74**	1.4E-01	transfer/carrier protein
P02768	Serum albumin	53	**1.45**	1.9E-12	**1.46**	6.4E-04	19	**1.76**	1.3E-02	**1.79**	1.7E-02	transfer/carrier protein
O75368	SH3 domain-binding glutamic acid-rich-like protein	3	**2.08**	1.3E-02	**1.88**	1.4E-02	7	**0.75**	8.1E-02	**0.70**	7.8E-02	N.A.
Q9H299	SH3 domain-binding glutamic acid-rich-like protein 3	2	**1.57**	1.1E-02	**1.52**	1.0E-01	6	**1.20**	8.7E-01	**1.46**	4.9E-01	N.A.
P09486	SPARC	21	**0.61**	1.9E-10	**0.71**	5.5E-02	21	**0.90**	7.5E-01	**0.83**	6.0E-01	growth factor
Q13813	Spectrin alpha chain, brain	3	**1.46**	3.0E-02	**1.42**	2.8E-01	2	**0.61**	4.8E-01	**0.59**	5.2E-01	N.A.
P31948	Stress-induced-phosphoprotein 1	1	**1.58**	3.4E-01	**2.38**	4.1E-01	2	**2.80**	4.8E-02	**2.26**	2.6E-02	chaperone
Q9UGT4	Sushi domain-containing protein 2	2	**0.66**	1.9E-01	**0.60**	3.3E-01	1	**0.38**	1.0E-01	**0.41**	2.3E-02	N.A.
O00560	Syntenin-1	3	**1.00**	9.8E-01	**0.85**	3.0E-02	5	**0.80**	4.4E-01	**0.64**	1.4E-01	membrane trafficking regulatory protein
P62328	Thymosin beta-4	9	**3.05**	1.1E-03	**4.63**	1.3E-01	16	**0.91**	9.1E-01	**1.14**	8.6E-01	N.A.
P05543	Thyroxine-binding globulin	19	**0.95**	5.1E-01	**0.79**	2.5E-02	20	**0.58**	3.1E-02	**0.59**	9.1E-02	serine protease inhibitor
P37837	Transaldolase	3	**3.81**	1.3E-02	**4.63**	7.9E-03	12	**1.13**	6.7E-01	**1.11**	7.8E-01	transaldolase
Q01995	Transgelin	3	**0.58**	2.7E-02	**0.53**	6.2E-03	16	**0.69**	2.9E-01	**0.69**	2.7E-01	non-motor actin binding protein
P37802	Transgelin-2	5	**2.04**	4.3E-04	**1.94**	9.1E-05	16	**1.18**	2.8E-01	**1.21**	2.2E-01	N.A.
P29401	Transketolase	7	**3.21**	1.8E-03	**4.13**	1.3E-03	17	**1.33**	5.1E-01	**1.55**	2.1E-01	transketolase
P60174	Triosephosphate isomerase	12	**1.89**	1.7E-06	**2.42**	3.0E-07	26	**1.48**	6.2E-04	**1.47**	4.8E-02	isomerase
P23381	Tryptophanyl-tRNA synthetase, cytoplasmic	3	**1.57**	1.1E-02	**1.57**	9.1E-02	2	**1.11**	7.1E-01	**0.77**	2.0E-01	aminoacyl-tRNA synthetase
P68366	Tubulin alpha-4A chain	2	**2.74**	4.6E-02	**2.22**	1.9E-01			N.D.			tubulin
P78324	Tyrosine-protein phosphatase non-receptor type substrate 1	4	**1.74**	1.0E-02	**2.02**	5.2E-03	2	**1.28**	3.5E-01	**1.19**	6.0E-01	chemokine
P22314	Ubiquitin-like modifier-activating enzyme 1	2	**2.05**	1.0E-02	**1.64**	2.5E-01	9	**1.16**	1.6E-01	**1.23**	1.9E-01	transfer/carrier protein
Q6UX73	UPF0764 protein C16orf89			N.D.			1	**0.55**	3.7E-01	**0.50**	1.4E-02	N.A.
Q16851	UTP--glucose-1-phosphate uridylyltransferase	3	**2.12**	4.9E-03	**2.26**	4.5E-03	4	**1.40**	1.7E-02	**1.19**	4.7E-01	nucleotidyltransferase
P17948	Vascular endothelial growth factor receptor 1	23	**0.65**	1.2E-08	**0.70**	2.4E-02	31	**0.93**	8.4E-01	**0.90**	7.8E-01	protein kinase
P50552	Vasodilator-stimulated phosphoprotein	2	**5.46**	1.6E-02	**5.00**	3.9E-02	2	**1.72**	4.6E-02	**1.82**	1.5E-01	N.A.
P13611	Versican core protein	8	**1.28**	3.2E-02	**1.34**	6.8E-02	3	**1.83**	1.5E-01	**1.89**	4.3E-02	extracellular matrix glycoprotein
P08670	Vimentin	27	**3.15**	1.7E-12	**3.09**	2.1E-10	29	**1.54**	2.1E-01	**1.42**	3.9E-01	structural protein
P18206	Vinculin	6	**1.50**	4.9E-03	**1.37**	1.9E-01	14	**0.80**	3.3E-01	**0.76**	2.4E-01	non-motor actin binding protein

ATTACHED PAPERS
PAPER I.

Physiol. Res. 59: 471-497, 2010

REVIEW

Application of Proteomics in Biomarker Discovery: a Primer for the Clinician

V. TAMBOR[1*], A. FUČÍKOVÁ[1*], J. LENČO[1], M. KACEROVSKÝ[2,3], V. ŘEHÁČEK[2,4], J. STULÍK[1], R. PUDIL[2,5]

**These authors contributed equally to this work.*

[1]Institute of Molecular Pathology, Faculty of Military Health Sciences, University of Defence, Hradec Králové, Czech Republic, [2]Charles University in Prague, Faculty of Medicine in Hradec Králové, Czech Republic, [3]Department of Obstetrics and Gynecology, University Hospital Hradec Králové, Czech Republic, [4]Transfusion Department, University Hospital Hradec Králové, Czech Republic, [5]Department of Internal Medicine, University Hospital Hradec Králové, Czech Republic

Received February 3, 2009
Accepted October 5, 2009
On-line November 20, 2009

Summary

Ever since proteomics was proven to be capable of characterizing a large number of differences in both protein quality and quantity, it has been applied in various areas of biomedicine, ranging from the deciphering molecular pathogenesis of diseases to the characterization of novel drug targets and the discovery of potential diagnostic biomarkers. Indeed, the biomarker discovery in human plasma is clearly one of the areas with enormous potential. However, without proper planning and implementation of specific techniques, the efforts and expectations may very easily be hampered. Numerous earlier projects aimed at clinical proteomics, characterized by exaggerated enthusiasm, often underestimated some principal obstacles of plasma biomarker discovery. Consequently, ambiguous and insignificant results soon led to a more critical view in this field. In this article, we critically review the current state of proteomic approaches for biomarker discovery and validation, in order to provide basic information and guidelines for both clinicians and researchers. These need to be closely considered prior to initiation of a project aimed at plasma biomarker discovery. We also present a short overview of recent applications of clinical proteomics in biomarker discovery.

Key words

Proteomics • Biomarker • Plasma • Mass spectrometry

Corresponding author

J. Lenčo, Institute of Molecular Pathology, Faculty of Military Health Sciences, University of Defence, Třebešská 1575, 500 01 Hradec Králové, Czech Republic. Fax: +420 435 513 018. E-mail: lenco@pmfhk.cz

Introduction

During the past decade, several groundbreaking discoveries in life science were made. The completion of sequencing the human genome certainly belongs to the key tasks successfully completed, representing a true milestone in biomedicine (Collins *et al.* 2004). Indeed, this has provided an important knowledge base, thus enabling rapid development in life science-oriented research, in such areas as prenatal and postnatal diagnostics, gene therapy, discovery of new drug targets, and development of personalized therapies (Workman 2003, Lau and Leung 2005, Young *et al.* 2006, Rosa *et al.* 2008). The accomplishment of the complete genome also brings along a new, even more challenging task for scientists: the characterization of the human proteome.

The term "proteome" was used first in 1994 and describes a set of all proteins expressed by a given genome (Wasinger *et al.* 1995). A more accurate definition,

emphasizing its dynamic nature, further specifies the proteome as a set of proteins in a given time and space, as its composition may vary from tissue to tissue or even from cell to cell. Furthermore, the structure of a proteome is dependent on a wide range of internal and external factors such as environment, age, sex, diseases, etc., which is in sharp contrast with the nature of the genome.

A protein, the basic unit of a proteome, is a molecule composed of single amino acids, further forming secondary, tertiary, and quaternary three-dimensional structures. Although the amino acid sequence is defined by the appropriate gene, the genetic information itself cannot provide the complete information about a protein. In contrast to the stable, rigid, single- dimensional genomic information based on a combination of four nucleotides, the information encoded in proteins is not exclusively limited to the amino acid sequence. Specific properties of proteins like various conformation states, posttranslational modifications, and alternative splicing demonstrate the multidimensionality, high variability, and dynamic nature of the proteomic information. This explains the high number of unique protein molecules, far exceeding the number of respective genes, particularly in eukaryotes.

Proteomics, the main tool for proteome research, is a relatively new and extremely dynamically evolving branch of science, focused on the evaluation of gene expression at proteome level. Due to the specific properties of proteins mentioned above, current proteomics deals with different issues, such as protein identification, quantification, characterization of posttranslational modification, structure and function elucidation and description of possible interactions. The rapid development of proteomics was made possible by progress in analytical instrumentation, especially in mass spectrometry (MS) with the introduction of new, cutting-edge types of mass spectrometers and improvements of soft ionization techniques. No less important are the advances in technologies and methodologies dealing with protein or peptide separation and sample complexity reduction, mainly in liquid chromatography and electrophoretic techniques. Bioinformatics is the third important foundation for advances in proteomics, as the ability to collect, store, process and visualize vast amount of data is crucial in extensive proteomics studies.

Although genomic research dominated the area of biomedical research in the past decades, proteomics is increasingly gaining ground in leading scientific workgroups and in clinical research labs. One of the reasons driving this platform change is the fact that a protein pattern of a biological sample is much more accurately up to reflecting the current physiological state of an organism than is the genome, and thus holds great promise in biomedicine.

Biomarkers

Timely recognition of an ongoing pathological process is a crucial factor that influences a patient's chances for successful treatment (Etzioni *et al.* 2003, Zhang *et al.* 2007b). To accelerate and facilitate the determination of diagnosis, current medicine strongly relies on the specialized assessment of certain molecules, where the concentration of these molecules in a biological sample more or less correlates with the occurrence of a given disease. Determination of the concentration change of such biomarkers may allow screening of high-risk individuals and detect disease at early, still well curable stages, as well as facilitate the prognosis prediction and monitoring of treatment response. The ultimate goal of implementing these biomarkers in routine clinical tests is the reduction of morbidity and mortality. Unfortunately, even with these tools, it is not always easy to realize the full potential of well-established markers (Andriole *et al.* 2009, Schroder *et al.* 2009).

Requirements of an ideal biomarker

According to the Food and Drug Administration (FDA), a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biologic or pathogenic processes or pharmacological responses to a therapeutic intervention. It may be also defined as an *in vivo* derived molecule present at levels deviating significantly from the average in association with specific conditions of health (Atkinson *et al.* 2001, Zhang *et al.* 2007b). From a biochemical point of view a biomarker is often a protein, the presence or quantitative characteristics of which are measured mostly using methods based on monoclonal antibodies. An ideal biomarker should enable unbiased diagnosis determination, particularly in patients without specific symptoms. It should therefore fulfill several criteria, particularly high specificity towards the given disease and high sensitivity. A correlation of the biomarker level and the disease stage is also desirable (Guo *et al.* 2007). Ease of use, standardization, and clarity and readability of the results

for the clinician are all factors that further affect the biomarker performance in the clinical setting. Unfortunately, many of these requirements are not met by most of the potential and even approved and used biomarkers (Anderson 2005). In theory, every disease may be uncovered and characterized by its unique biomarker. To see this biomarker as a single molecule, however, is just one alternative. Rather than as a unique protein, a biomarker should be regarded as a panel of up- and down-regulated proteins or proteins with altered posttranslational modifications, which differ in diseased and normal state (Etzioni *et al.* 2003, Rifai *et al.* 2006).

These facts along with the diagnostic potential of proteins and advances in proteomics technologies recently caused a significant increase of interest in biomarker research. These indicators hold great promise in early detection screening, disease progression monitoring, or in therapy efficiency evaluation, as new, more sensitive and specific markers are yet to be found (Etzioni *et al.* 2003, Veenstra *et al.* 2005, Hu *et al.* 2006, Hanash *et al.* 2008). To illustrate, we present some of recent studies dealing with biomarker discovery, which deserve particular attention because of clinical relevance or biological/methodical approach. These studies are summarized in Table 1.

Biomarker sources

One of the key issues in biomarker research is the accessibility of the source of biological matrix. Among a wide variety of available body fluids, blood is considered the most promising. Other fluids (urine, amniotic fluid, saliva, cerebrospinal fluid, nipple aspirate fluid, synovial fluid, etc.) cannot offer a protein profile as representative as that of blood, and availability of these samples may be very restricted. Blood as a source of biomarkers is easily accessible; its collection is minimally invasive, low risk, and cheap. The processing of crude blood to plasma is a routine task in clinical labs.

Blood

The most important advantage of blood is its contact with virtually all cells of the organism. Due to specific secretion, shedding from the surface, or non-specific leakage, tissue-related proteins are released into the blood stream (Zhang *et al.* 2007a). Therefore, pathologically affected cells with deregulated proteomes may create a specific "barcode" by disease-related proteins released into circulating blood. Besides the proteins originating from affected cells, the barcode is also represented by molecules resulting from organism response to the disease (Bijian *et al.* 2009). Therefore, this barcode includes high-abundance proteins, which can be readily analyzed using conventional techniques. Doubts have emerged, however, on whether these markers would be up to fulfill the criteria required for validation and pass all phases of testing. Except for intact proteins, the barcode also includes protein fragments due to proteases/peptidases deregulation. These are advantageously analyzed using MS profiling (Villanueva *et al.* 2006, Hashiguchi *et al.* 2009). However, the most interesting proteins originate from pathologically affected cells. Unfortunately, owing to the large blood volume, the final concentration of these diagnostically interesting proteins drops to about nanogram per milliliter (ng/ml) levels or even less (Anderson and Anderson 2002, Anderson *et al.* 2004b). To successfully analyze these compounds, sophisticated methods and specific procedures need to be implemented.

Because changes in the plasma proteome are not solely caused by pathological processes, the preanalytical phase is a crucial part of the biomarker discovery workflow. Factors like age, circadian rhythms, stress, medication usage, physical activity, pregnancy etc., may also significantly influence the plasma protein profile. Therefore, all the preanalytical steps – patient preliminary, blood collection, sample transport and storage – need to be strictly standardized and monitored, in order to prevent the occurrence of random and disease-unrelated changes in the plasma proteome. Even minor deviations in the pre-analytical phase may lead to false conclusions of the analysis (Rai *et al.* 2005, Banks 2008, Govorukhina *et al.* 2009). To prevent such deviations, i.e., in blood collection, specialized products like the BD P100 blood collection set (BD Diagnostics, USA) have been developed for proteomic purposes, standardizing the collection procedure. Another crucial aspect, namely, the number of cases and controls enrolled for a study, should also be carefully considered, as an insufficient number of patients may easily lead to false results. For higher credibility, it is advantageous to include patients from multiple clinical centers. In this case, however, strict requirements on standardized sample processing need to be closely monitored, as variations in preanalytical steps may lead even to contradictory results (Fiedler *et al.* 2009).

Table 1. Overview of recent clinical applications of protemics in biomarker discovery projects.

Research area and usefulness	Proteomic platform and validation method	Candidate markers	Ref.
Membranous nephropathy – diagnostic biomarkers	SDS-PAGE of glomeruli protein extract and Western blotting using human sera	Autoantibodies against phospholipase A_2 receptor	(Beck *et al.* 2009)
Hepatocellular carcinoma (HCC) – diagnostic biomarkers	2D-PAGE of HepG2 cells extract and Western blotting using human sera ELISA	11 immunoreactive protein spots were reactive only with HCC sera, among them HSP60 and HSP70	Looi *et al.* (2008)
Chronic inflammatory demyelinating polyneuropathy – diagnostic biomarkers	2D-PAGE (DIGE) analysis of human CSF Nephelometry	Transferrin, α-1 acid glycoprotein 1, apolipoprotein A IV, haptoglobin, transthyretin, retinol binding protein, proapolipoprotein, integrin β 8	Tumani *et al.* (2009)
Lung adenocarcinoma – biomarkers for cancer development and progression	WGA lectin affinity chromatography, 2D-PAGE (DIGE) analysis of human sera Western blot	Adiponectin, ceruloplasmin, cyclin H, proto-oncogene protein kinase Fyn, vanin-2 (GPI-anchored 80-kDa glycoprotein), additional 34 proteins	Hongsachart *et al.* (2009)
Colorectal cancer (CRC) – diagnostic biomarkers	2D-PAGE (DIGE) analysis of human tissue samples Western blot	From 51 tissue protein spots associated with development of CRC, S100A8 and S100A9 were found to be elevated in patients' plasma	Kim *et al.* (2009)
Hepatocellular carcinoma (HCC) – diagnostic biomarkers	SELDI-TOF profiling of human sera ELISA	Peak at *m/z* 13 391 identified as cystatin C, additional 10 peak signatures	Zinkin *et al.* (2008)
Renal cell carcinoma – diagnostic biomarkers	SELDI-TOF profiling of human sera	Peak at *m/z* 8937 identified as eukaryotic initiation factor 2B δ subunit, additional 24 peak signatures	Xu *et al.* (2009)
Melanoma – prognostic biomarkers in early-stage patients	MALDI-TOF profiling of human sera Unspecified immunoassay	Peak at *m/z* 11 680, identified as serum amyloid A, correlating with poor survival	Findeisen *et al.* (2009)
Pancreatic cancer – diagnostic biomarker	MALDI-TOF profiling of human sera ELISA	Three peak signatures at *m/z* 3194, 4055, 5959, and platelet factor 4 represented by peak at *m/z* 7767 and its doubly charged variant at *m/z* 3884	Fiedler *et al.* (2009)
Hepatocellular carcinoma with HCV etiology – diagnostic biomarkers	MALDI-TOF profiling of human sera	Complement C3 peptide, complement C4a peptide and additional four peak signatures	Goldman *et al.* (2007)
Breast cancer – diagnostic and prognostic biomarkers	MALDI-TOF profiling of *N*-glycans released from human plasma glycoproteins	Eight glycan signatures characteristic for breast cancer	Kyselova *et al.* (2008)

Hepatocellular carcinoma (HCC) with HCV etiology – diagnostic biomarkers	MALDI-TOF profiling of *N*-glycans released from human plasma glycoproteins	Three glycan signatures at *m/z* of 2473, 3242 and 4052	Goldman *et al.* (2009)
Chronic allograft dysfunction (CAD) – diagnostic biomarker	LC-MSMS (label free) analysis of human urine peptides LC-MSMS based on Extracted Ion Chromatogram	Uromodulin peptide SGSVIDQSRVLNLGPITR Kininogen peptide DLIATMMPPISPAPIQSDDDW IPDIQI, ions at *m/z* 645.59 and at *m/z* 642.61	Quintana *et al.* (2009)
Rheumatoid arthritis (RA) – diagnostic biomarker	LC-MSMS (label free) analysis of human plasma peptides	Peptides from 25 proteins found differently abundant in patients with RA, peptides derived from thymosin β4 found among the most elevated	Wei *et al.* (2008)
Dilated cardiomyopathy (DC) – diagnostic biomarkers	2D-LC-MSMS (label-free) analysis of mouse tissue Western blot	From 593 mouse tissue proteins associated with development of DC, RTN4 protein found to be elevated in patients' plasma	Gramolini *et al.* (2008)
Breast cancer – diagnostic biomarkers	LC-MSMS (label-free) analysis of mouse tissue MRM, ELISA and Western Blot	Osteopontin and fibulin-2 confirmed as circulating potential markers in mouse model	Whiteaker *et al.* (2007)
Pancreatic cancer – diagnostic biomarkers	2D-LC-MSMS (SILAP) analysis of human sera ELISA	ICAM-1 and BCAM were selected for validation from 121 proteins elevated by factor 1.5 in serum	Yu *et al.* (2009)
Preterm birth (PTB) – screening biomarkers for women at risk	2D-LC-MSMS (SILAP) analysis of human cell lines supernatant MRM for validation in cervicovaginal fluid	From 15 candidates identified in cell line supernatants mixture, desmoplakin isoform 1, stratifin, thrombospondin 1 were confirmed significantly elevated in PTB	Shah *et al.* (2009)
Pancreatic cancer (PC) – early diagnostic biomarkers	LC-MSMS (stable isotope labeling of cystein residues using D0/D3 acrylamide) analysis of mouse plasma ELISA	Five proteins discriminating between patients with PC and healthy individuals up to 13 months prior to development of clinical symptoms	Faca *et al.* (2008)
Endometrial cancer – diagnostic biomarkers	2D-LC-MSMS (iTRAQ) analysis of human endometrial tissue MRM	From nine markers, pyruvate kinase and polymeric immunoglobulin receptor were chosen for subsequent verification and absolute quantification	DeSouza *et al.* (2008, 2009)
Cardiovascular injury biomarkers – previously known markers or marker candidates	MRM, ELISA	CRP, MRP14, MPO, cTnT, cTnI, and NT-proBNP were absolutely quantified in plasma using internal standard.	Keshishian *et al.* (2009)

Although the enormous complexity of blood as a factor reflecting the state of the whole organism may be regarded as an advantage, it may be also seen as a disadvantage from the analytical point of view. Indeed, blood plasma is an extremely rich mixture of proteins and peptides as well as proteins originating from microorganisms. Moreover, proteins may be represented in a number of various forms due to their posttranslational modifications or alternative splicing, which further greatly increases the diversity of the plasma proteome (Anderson and Anderson 2002). Although more than 9,000 plasma proteins have been identified so far, as reported by the HUPO consortium, this was achieved in a collaborative project of 35 laboratories (States *et al.* 2006). Unfortunately, this number of identified proteins is extremely hard to achieve in single-laboratory settings. To illustrate, a more recent work led through very extensive fractionation of serum to the identification of 4,396 proteins in one study (Tucholska *et al.* 2009). The wide concentration range of plasma protein is another limiting factor, as the estimated concentration span exceeds 10 orders of magnitude (Anderson and Anderson 2002). This exceeds the dynamic range of any current analytical instrument or method. The questing for biomarkers thus presents a real challenge for plasma-based proteomics research, as these molecules are hidden among 20 very high-abundance proteins, representing ~ 99 % of total plasma protein (Veenstra *et al.* 2005).

Addressing the problem of high-abundance proteins

In present proteomic research, several methods have been introduced in order to solve some of the pitfalls associated with plasma analysis. One of the key points, often implemented as the first step of proteomic sample workflow, is the removal of ballast high-abundance proteins with no diagnostic potential using immunoaffinity depletion (Tam *et al.* 2004, Echan *et al.* 2005, Huang and Fang 2008). This approach takes advantage of immobilized polyclonal antibodies to remove a portion of high-abundance proteins. These antibodies are designed to bind defined proteins and their isoforms, allowing the removal of up to ~ 95 % of total plasma protein, which results in significant reduction of complexity and dynamic range (Fig. 1). This may lead, in turn, to a higher number of identified proteins, improved sequence coverage, and more accurate protein quantification (Chromy *et al.* 2004, Tam *et al.* 2004, Huang *et al.* 2005b). The depletion step is subsequently included in the validation phase as well, as it enables adequate sample loading (Kim *et al.* 2009). This approach, however, brings along certain disadvantages, as some of the high-abundance proteins, albumin in particular, are known to act as carrier molecules for other proteins, possibly with diagnostic potential. Thus, by removing the carrier proteins, these potentially interesting molecules may be lost as well (Huang *et al.* 2005a, Liu *et al.* 2006).

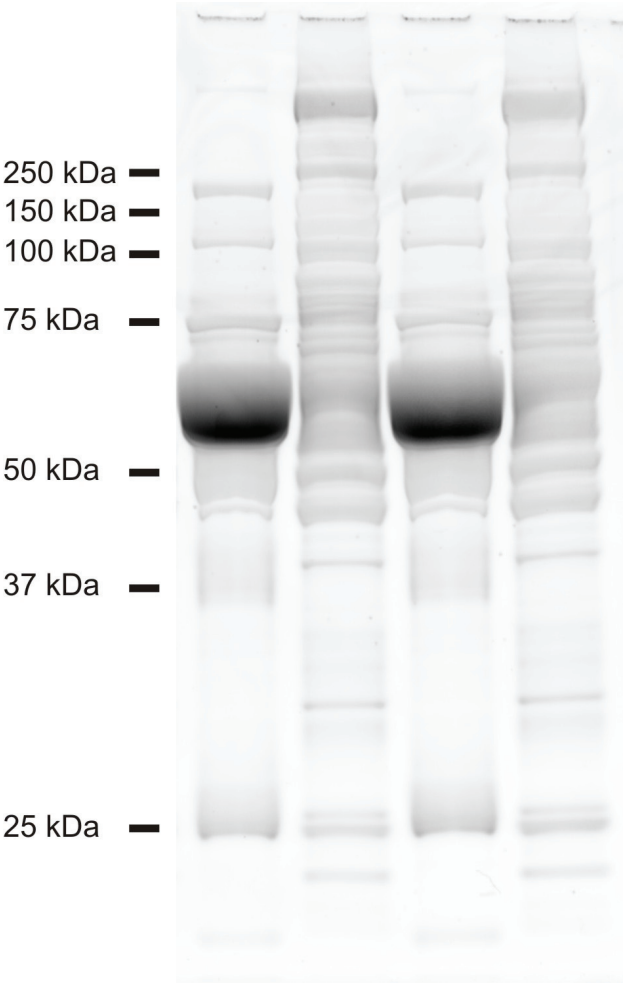


Fig. 1. SDS-PAGE gel of plasma samples processed by immunoaffinity depletion on a MARS Hu-14 column (Agilent). The first and third lane was loaded with the bound fraction, i.e. a fraction containing depleted high abundance proteins. The second and fourth lane present a plasma sample depleted from high abundance proteins.

Peptide libraries present an alternative solution for dynamic range reduction. Instead of removing a portion of high-abundance proteins, the peptide libraries equilibrate concentration of plasma proteins to a similar level. Microscopic beads are covered with a library of hexapeptides prepared using combinatorial synthesis

from common amino acids (Thulasiraman *et al.* 2005, Righetti *et al.* 2006, Righetti and Boschetti 2007, Sennels *et al.* 2007). This results in millions of bead populations, each population carrying a unique peptide sequence. Based on probability, the majority of plasma proteins is supposed to find a binding partner. After the binding capacity of a particular bead population is saturated, the remaining portion of the given protein cannot bind any more and is washed out. The proteins are then eluted from the beads and further analyzed. However, due to the nature of this method, the differences in protein concentration are smoothed among individual samples after saturating the capacity, and only low-abundance proteins that are not up to saturate the beads may be quantified among more samples without employing a method based on stable isotope labeling (Roux-Dalvai *et al.* 2008).

Mining the plasma glycoproteins

The glycosylation of proteins is known to be aberrant in different disease states, especially in cancer (Spiro 2002, Brooks *et al.* 2008). In addition, most of the proteins localized at the surface or secreted by cells are glycosylated. Therefore, disease-related glycoproteins, either actively secreted, or passively shed or leaked from the cells due to cellular damage or death, are likely to occur in the blood stream. Unsurprisingly, numerous clinically used protein markers are glycosylated, such as PSA, CA125, and CEA (Kui Wong *et al.* 2003, Comegys *et al.* 2004, Ludwig and Weinstein 2005, Tajiri *et al.* 2008). Hence, glycoproteomics has been attracting considerable attention in the biomarker discovery field because suitable technologies and methods for glycoproteomic analysis have emerged. With respect to techniques used for this purpose, two approaches can be identified. Lectin affinity chromatography is capable of enriching glycosylated proteins from complex matrices by interaction with various types of lectin without destroying the glycan part, leaving it available for analysis (Mechref *et al.* 2008). During the second alternative approach, covalent capturing of glycoproteins/ glycopeptides, the oxidized glycan moiety is covalently bound to hydrazide solid phase support. The protein/ peptide backbone can be released by enzymatic cleavage from the glycan part and analyzed. The glycan part cannot be recovered from the hydrazide resin and is therefore unavailable for analysis using this approach (Tian *et al.* 2007).

Proximal fluids and tissue

Proximal fluids as a source of biomarkers present a compelling alternative to blood. Although proximal fluids are not as representative as blood, their expedience increases if the nidus of a disease is in close contact with the particular body fluid, i.e., urine may be a prospective source of kidney diseases biomarkers (Quintana *et al.* 2009), or cerebrospinal fluid for central nervous system diseases (Tumani *et al.* 2009). The anticipated biomarker molecules are present in a significantly higher concentration than in body fluids. Moreover, if a disease-specific marker is found in tissue, targeted approaches may be introduced to assess its presence in body fluids as described further in this article (Schiess *et al.* 2009). Unfortunately, the accessibility of tissue specimens or some of the proximal fluids is mostly more complicated compared to collection of blood and may present a level of risk for the patient. Analysis of both tissue and proximal fluids from an animal model of the respective disease may be an option, as these are much more easily obtainable and results from these studies may be then translated to human analogues of the disease (Whiteaker *et al.* 2007, Gramolini *et al.* 2008).

Multistage strategies

Direct analysis of human plasma is just one of many possible ways to seek for new markers. The major obstacles to direct biomarker discovery in plasma – enormous complexity and high concentration range – recently caused several new strategies to emerge (Schiess *et al.* 2009). These are generally divided into multiple parts. First, diseased and control biological samples with anticipated concentrations of potential markers higher than in plasma are compared. These might include model cell lines, affected tissue samples, and proximal fluids (Kulasingam and Diamandis 2008). Apart from the fact that potential markers are present in higher amounts in these sources than in plasma, the overall protein concentration range in cells is lower than in blood, and proteomic analysis of these sources results in higher proteome coverage. Even an animal model may be used, as the diseased and control animals are precisely defined and their genetic backgrounds are alike in all respects (Whiteaker *et al.* 2007, Gramolini *et al.* 2008). Along with the ability to grow a human cancer tissue in the animal host, subsequent analysis allows differentiation of cancer tissue-released proteins from host response proteins (Bijian *et al.* 2009). These are all ways of identifying more specific and sensitive potential

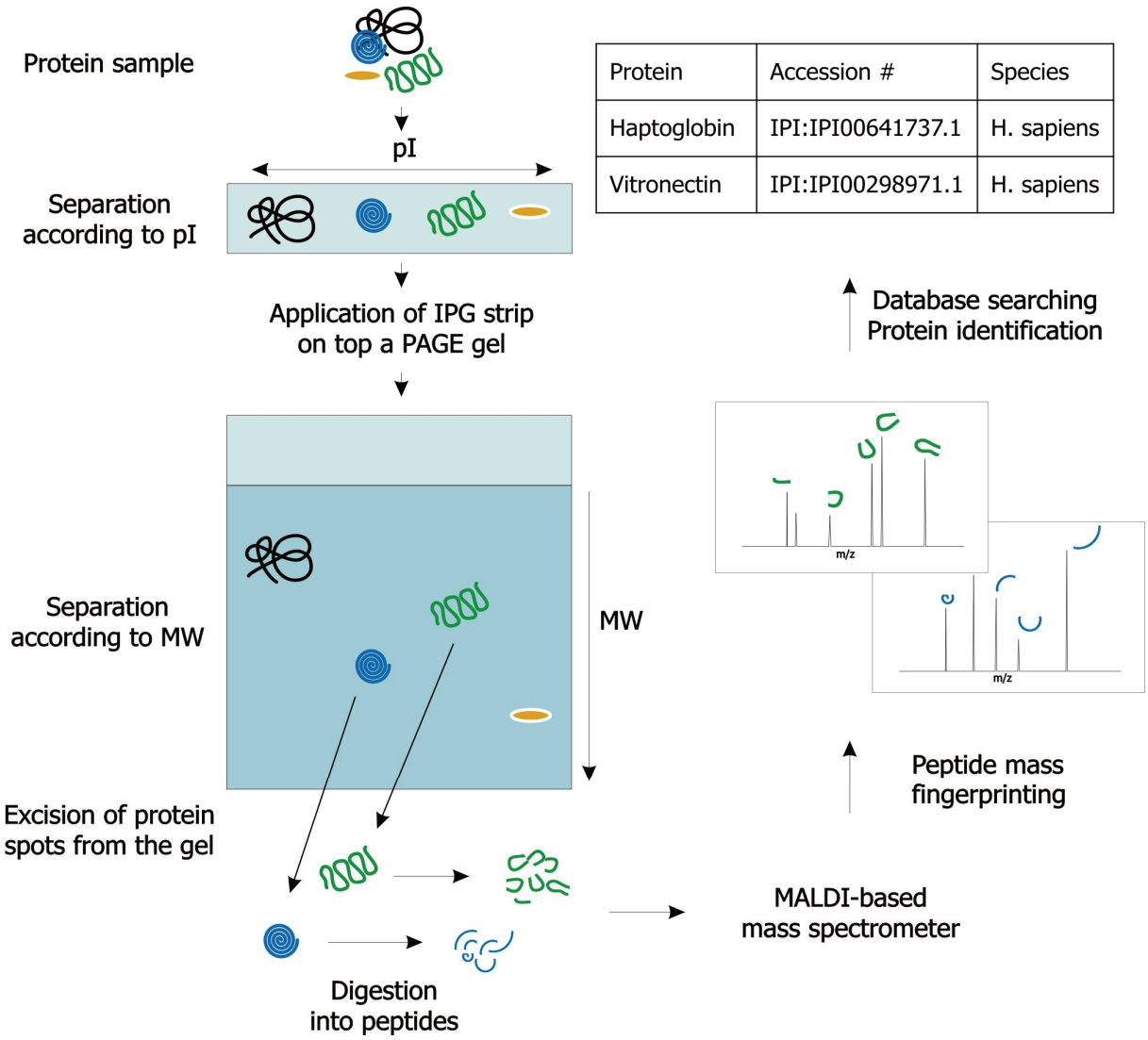


Fig. 2A. 2D-PAGE workflow: A complex protein sample is applied onto an IPG strip and the proteins are separated according to their pI. Then, the strip is placed on top a SDS-PAGE gel and the proteins are separated according to their molecular weight (MW) in second, perpendicular dimension. After gel staining, protein spots of interest may be cut out, digested into peptides and identified mostly by means of PMF approach.

biomarkers. These candidate markers are subsequently targeted in plasma and if their presence is confirmed, then they are simultaneously precisely quantified using targeted mass spectrometry, as described in respective section of this article.

An interesting approach based on identification of glycosylated cell surface proteins was published recently (Wollscheid *et al.* 2009). By this means, it is possible to precisely describe the cell surface proteome. Subsequently, these proteins are targeted using targeted proteomics in human plasma as the cell surface proteins are released into the blood stream upon cellular death or damage. The major drawback of this workflow is the requirement of a suitable representative tissue sample or cellular model of a disease, which is not always available.

Proteomic approaches for plasma analysis

Currently, there are three primary approaches available in biomarker discovery projects (Fig. 2A-C). Each of these methods offers unique advantages but also suffers from specific and often substantial drawbacks. Therefore, one should keep in mind that none of these techniques is ideal and a thorough discussion is crucial prior to selecting the definitive approach. Even though these three methods are fundamentally distinct, a common denominator for all three is the application of mass spectrometry. Therefore, we present a brief description of this key technique.

In principle, mass spectrometry as an analytical technique enables accurate measurements of molecular

weights of individual components in a given sample. A mass spectrometer comprises three major parts: ion source, analyzer, and detector. During a typical MS experiment, sample molecules are ionized and converted into gas phase in the ion source, separated according to their mass to charge ratio in the analyzer, and finally detected by the detector. As for individual segments, the most commonly used ion sources in proteomics are Matrix-Assisted Laser Desorption/Ionization (MALDI) and Electrospray Ionization (ESI). In a MALDI source, analyte molecules are ionized from solid state by a pulsed laser beam, whereas an ESI source ionizes dissolved molecules by spraying them in an extremely fine beam directly into a mass analyzer. For individual types, the commonly used analyzers in proteomics are time-of-flight (TOF), quadrupole (Q), ion trap (IT), Fourier transform ion cyclotron resonance (FT-ICR), and Orbitrap. These might be used either singly or in a tandem configuration. In tandem mass spectrometry (MSMS), multiple steps of mass analysis can be performed with individual analyzers separated in space or in a single analyzer with steps separated in time. In MSMS separated in space, analyzers are physically separated, but are tightly connected in order to maintain vacuum. This configuration is used in the following instruments: Q-TOF, TOF-TOF, Triple Quadrupole, etc. MSMS in time, on the other hand, can be performed with ions trapped in the same place, with individual analysis steps carried out over time. Ion traps or FT-ICRs can be used for this purpose.

Two-dimensional polyacrylamide gel electrophoresis

The very first method employed in comprehensive proteomic experiments was the two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (Fig. 2A). The proteins are separated in a gel matrix based on two independent physicochemical properties of each protein: isoelectric point (pI) and molecular weight (MW) (O'Farrell 1975, Gorg *et al.* 2004, Carrette *et al.* 2006). By a combination of these two features, a high-resolution separation of proteins may be readily achieved.

The protein mixture is separated using isoelectric focusing (IEF) according to the pI of the proteins in the first dimension. The IEF is carried out on commercial gel strips with an immobilized pH gradient (IPG) strips (Bjellqvist *et al.* 1982). The IPG strips containing focused proteins are incubated with sodium dodecyl sulphate (SDS), a detergent that covers the

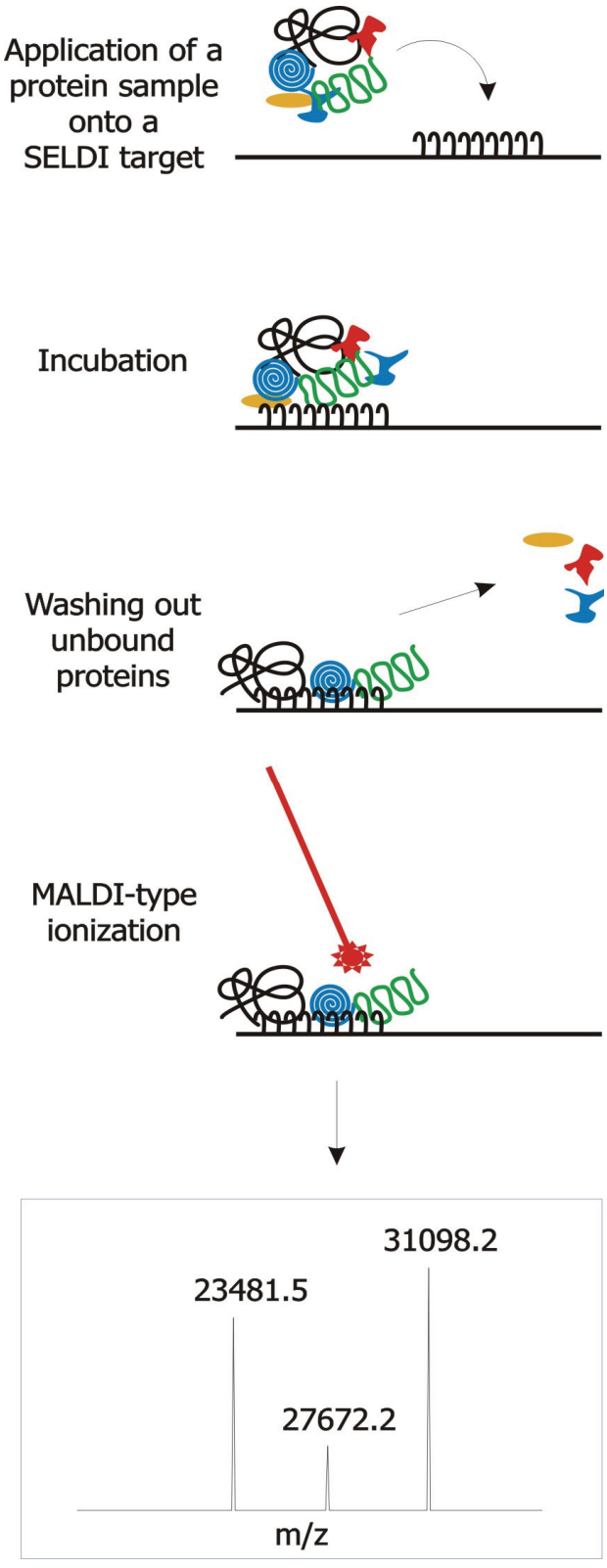


Fig. 2B. Protein profiling workflow: Crude sample is applied onto a SELDI target modified by a specific chromatographic surface. After incubation the unbound fraction is washed away. The SELDI chip is directly analyzed using a SELDI-TOF mass spectrometer. A protein profile is obtained, each protein being represented by a peak with a corresponding m/z value. Note that information on protein identity is missing and cannot be obtained by this type of analysis.

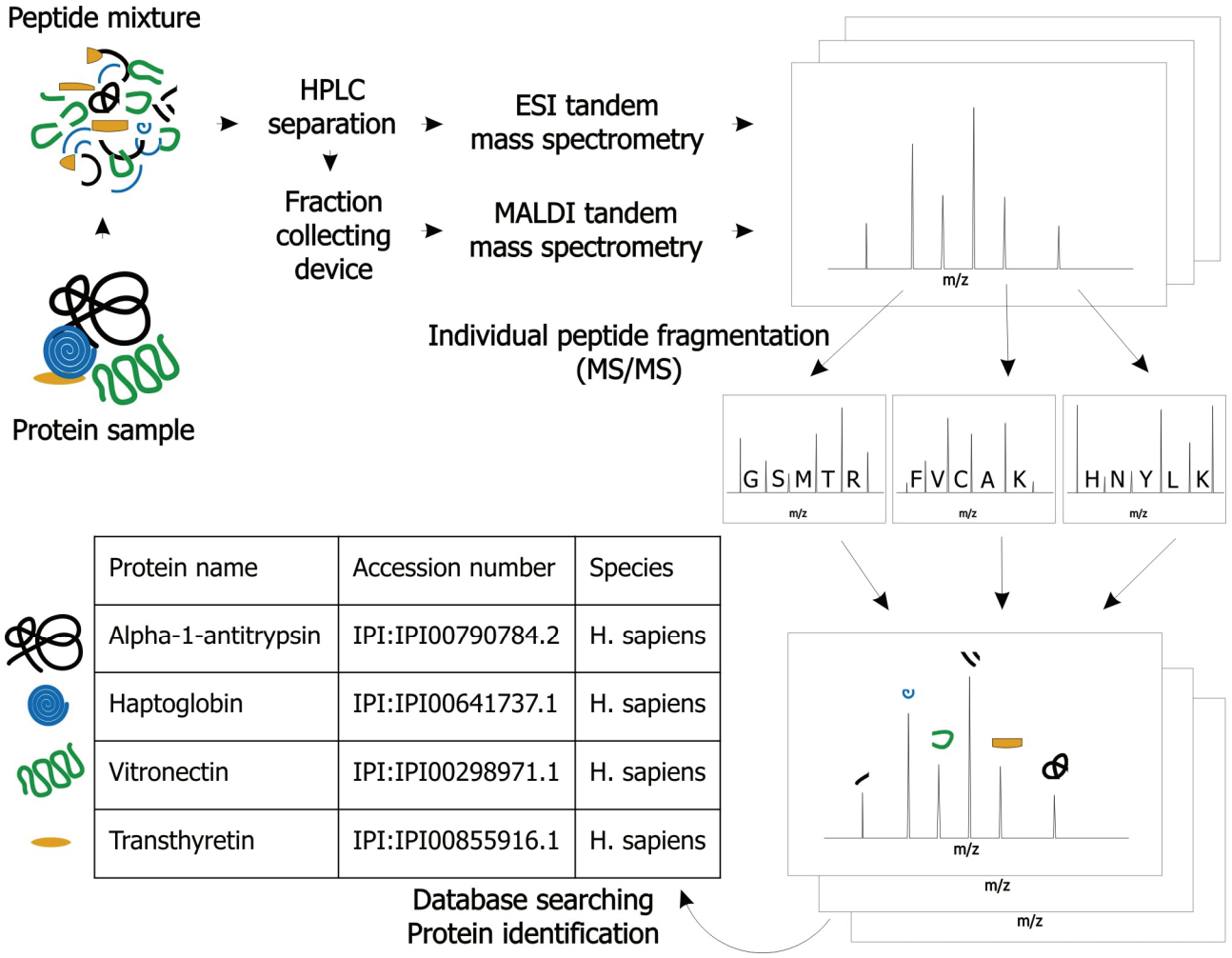


Fig. 2C. Shotgun proteomics workflow: A complex protein sample is digested by a sequence specific protease into peptides. This mixture of peptide may be optionally fractionated and separated. The separated peptides are subjected to MS analysis. First, the MS spectra are acquired and selected peptides from these spectra are fragmented. Resulting MSMS spectra are used for peptide identification. A list of identified peptides is then used in order to identify individual protein components of original sample.

proteins with a negative charge. After incubation, the IPG strips are placed on top of the SDS polyacrylamide slab gels, and the proteins are separated based on their MW in the second, perpendicular dimension (Laemmli 1970). This results in a two-dimensional protein map, where the proteins can be visualized using various approaches.

Conventional staining protocols include Coomassie™ Blue G-250 and R-250 dyes (Neuhoff *et al.* 1988, Candiano *et al.* 2004), or a color reaction based on silver ions reduction of ionic to metallic silver onto the protein surface (Rabilloud *et al.* 1994, Chevallet *et al.* 2008). Increasingly popular fluorescent dyes, e.g. Sypro™ Ruby (Berggren *et al.* 2000) and Deep Purple™ formerly known as Lightning Fast, (Mackintosh *et al.* 2003) offer ameliorated sensitivity and linearity for quantification compared to classic staining agents. The Differential Gel Electrophoresis (DIGE) employs three fluorescent dyes (Cy2, Cy3, and Cy5) for covalent protein labeling prior to 2D-PAGE. Due to their identical physicochemical properties in regard of pI and MW, labeled proteins are run on the same gel simultaneously. However, due to different excitation and emission wavelengths of the dyes, a unique 2D protein map can be acquired for each protein sample loaded on the gel (Unlu *et al.* 1997). When choosing the appropriate staining protocol, factors like sensitivity, dynamic linearity, and compatibility with MS analysis should be taken into consideration (Miller *et al.* 2006, Berth *et al.* 2007) (Table 2). Subsequently, the stained gels are digitalized and evaluated by means of specialized software enabling quantification of proteins via comparison of the intensity of stained spots (Berth *et al.* 2007). This final step is crucial, as any variance in image processing may lead to false results, mostly in quantification (Stessl *et al.* 2009).

Table 2. The most frequent staining methods (based on Miller *et al.* 2006).

Staining method	Principle of detection	Sensitivity	Linearity for quantification	MS compatibility
Coomassie™ Blue G-250 colloidal	Absorption	++	++	+
Silver staining	Absorption	+++	+	–/+
Sypro™ Ruby	Fluorescence	+++	+++	+
CyDyes - DIGE staining	Fluorescence	++++	++++	+

The separated protein spots are identified on a mass spectrometer, mostly using the peptide mass fingerprinting method (PMF) (Shevchenko *et al.* 1996, Henzel *et al.* 2003). A gel piece containing an isolated protein is excised and enzymatically digested by trypsin or any other sequence specific protease, resulting in a mixture of peptides. A MS spectrum is acquired, each peptide being represented by its mass-to-charge ratio (*m/z*) value. The recorded *m/z* values are compared with theoretical values and in case of a match, a protein is assigned to a spectrum with certain probability, according to the Mowse scoring algorithm (Pappin *et al.* 1993). The theoretical *m/z* values are obtained by *in silico* translation of DNA sequences of genes into proteins, from which theoretical proteolytic peptide masses are computed. If a spectrum fails to provide sufficient data for confidential protein identification, a tandem spectrometer may be used, as this type of instrument enables direct acquisition of a peptide sequence (Thiede *et al.* 2005).

Several hundreds to a few thousands of protein spots may be separated on a single 2D-PAGE gel. This approach is one of the most suitable for separating isoforms of identical proteins. Also, the expenditure for the required equipment and chemicals is relatively low. However, the main drawbacks of 2D-PAGE include reproducibility issues, time and labor intensiveness of the process, and imperfect separation of protein in both pI and MW extremities and of hydrophobic proteins. A partial solution to the reproducibility and dynamic range problems may be achieved using the DIGE approach, solving also problems regarding the low dynamic range of conventional staining methods.

Although the 2D-PAGE method has been applied to numerous projects for biomarker discovery, the proteins with altered concentration belong mostly to the group of high-abundance proteins (Tumani *et al.* 2009). However, if specific fraction or enrichment methods are employed during the sample processing workflow, even tissue-derived proteins may be detected using this approach (Hongsachart *et al.* 2009). Nonetheless, gel-based techniques may bring substantial results in a very specific field of biomarkers, namely autoantibodies that act as markers. In autoimmune diseases or in cancers autoantibodies are often found to be targeted against own cellular proteins (Bazhin *et al.* 2009). In this case, the strategy of searching for biomarker is far different from those described above, as the marker itself is an immunoglobulin and the task is to determine against which antigen it is targeted. The strategy is to perform Western blotting of affected tissue proteins by using imunoglobulins from the sera of patients. Although protein-antibody arrays currently dominate this area, conventional gel-based proteomic methods can still bring significant results (Looi *et al.* 2008, Beck *et al.* 2009).

Proteomic profiling (Fig. 2B)

Direct MS analysis of a sample may provide rapid insight into its protein profile. An instrument based on MALDI-TOF in linear configuration is ideal for this purpose, as it enables an acquisition of wide *m/z* range. By this approach, protein profiles of samples may be quickly compared, resulting in a list of differentially concentrated protein peaks (Fig. 3). However, due to the complexity of biological samples, the majority of low-abundance proteins remain undetected. This issue is partially solved by sample prefractionation on a carrier, covered by various chromatographic surfaces. These bind only the desired subset of proteins and the corresponding protein profile is then acquired using a mass spectrometer directly from these carriers. This approach is also known as Surface-Enhanced Laser Desorption/Ionization time-of-flight mass spectrometry (SELDI-TOF MS) (Dattelbaum and Iyer 2006, Poon 2007). Currently, a variety of chemical and biochemical surfaces is at disposal, enabling analysis of a wide range of protein subgroups. Analogous analyses may be also performed

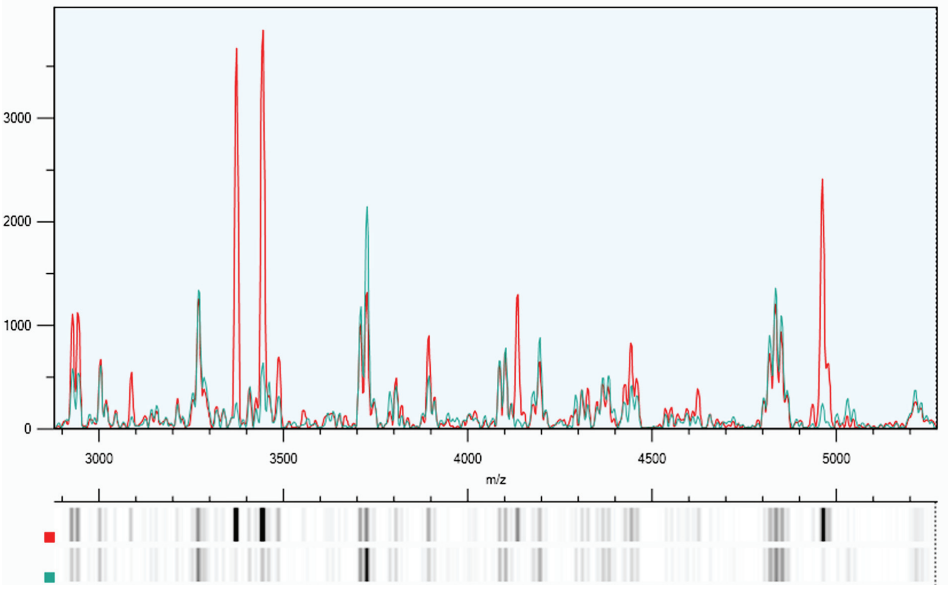


Fig. 3. Representative MS profiling spectrum. Two samples obtained from infection free (shown in green) and infected (shown in red) amniotic fluid were acquired on a MALDI-TOF mass spectrometer in linear mode and compared. Several markedly altered peaks were detected. Except the spectra, an alternative gel-like view is also shown.

on a MALDI-TOF instrument, but the sample prefractionation has to be performed separately, i.e., using magnetic beads modified by various chromatographic surfaces, similar to those on SELDI carriers, or using column devices filled with chromatographic phases. This configuration enhances sensitivity, as the surface of beads is higher compared to those of SELDI targets. Due to the poor analytical performance of SELDI-TOF instruments, researchers experienced in mass spectrometry prefer alternatives based on MALDI-TOF technology for biomarker discovery applications (Villanueva *et al.* 2004, Callesen *et al.* 2009).

Compared to 2D-PAGE, a SELDI-TOF analysis requires a much lower amount of sample, which may be in addition applied directly onto the target, without extensive preparation. Also, this technique is remarkably fast and high-throughput. Nevertheless, the SELDI-TOF approach suffers from some major drawbacks, including low spectra resolution and low accuracy. In addition, concerns about reproducibility discourage MS profiling from becoming a routine proteomic tool prior to addressing standardization of preanalytic and analytic factors (Banks 2008, Bruegel *et al.* 2009, Callesen *et al.* 2009). Furthermore, the absence of means for precise protein identification in SELDI-TOF limits the information about a biomarker candidate protein to just its *m/z* value in most cases. Although publications presenting just these limited data on candidate markers keep emerging, proteins/peptides defined just by *m/z* are worthless for diagnostic applications because their unknown identity hinders further validation by independent orthogonal methods. Limited or no options for this validation step further increase controversy and skepticism currently associated with this approach. Even though the SELDI-TOF technique or profiling based on MALDI-TOF instrument shows some disease-related changes in plasma, these occur mainly among the higher-abundance proteins (Hu *et al.* 2006, Findeisen *et al.* 2009). Due to their low specificity, however, these would unlikely pass the validation for a disease-specific biomarker. On the other hand, as the profiling approach focuses on low *m/z* segment, disease-specific low-molecular weight fragments may be detected in plasma as certain pathologies are characterized by profound deregulation in proteases/peptidases activities (Villanueva *et al.* 2006, Goldman *et al.* 2007, Hashiguchi *et al.* 2009). Another area where the profiling strategy can be advantageously employed involves analysis of glycans from glycoproteins. As already mentioned, the glycosylation pattern of proteins is known to be aberrant in different diseases. One of the methods shown to be able to uncover disease-specific changes in glycosylation is MALDI profiling of *N*-glycan moieties released from plasma/serum glycoproteins. This method has been proved to be well reproducible (Wada *et al.* 2007). To date, sera from various cancer patients have been tested using this approach (Kyselova *et al.* 2008, Goldman *et al.* 2009). Although this approach seems to be very promising, as it has been shown to be able to distinguish individual cancer stages (Kyselova *et al.* 2008), concerns have been raised on how to identify the parent glycoproteins, allowing further validation.

Shotgun proteomics (LC-MSMS) (Fig. 2C)

The combination of liquid chromatography (LC) and MS allows detection of proteomes with greater depth, dynamic range, and enhanced accuracy of quantification than when using one-dimensional profiling techniques that record all ions in a single mass spectrum. The shotgun approach is closely linked to advances and progress in MSMS. A tandem mass spectrometer is an instrument capable of isolating a precursor ion, fragmenting it, and detecting resulting fragments (Domon and Aebersold 2006).

During a typical shotgun experiment, a protein mixture of various complexity is cleaved by a sequence-specific protease first. The most commonly used protease in proteomics is trypsin that cleaves a typical protein into several tens of peptides. Therefore, in case of analyzing a complex protein sample, a huge amount of different tryptic peptides raised from trypsin digestion disallows a direct MS analysis similar to the PMF method. Therefore, the resulting peptide mixture has to be separated, mostly by high performance liquid chromatography (HPLC), prior to analysis on a tandem mass spectrometer. These two systems may be connected either on-line, where the HPLC capillary flows directly into the ESI ionization source, or off-line, using a fraction collecting device. This device collects the peptides eluting from a HPLC system in time-dependent fractions directly onto a MALDI target plate (Bodnar *et al.* 2003). Alternatively, a continuous elution trace may be deposited onto the MALDI plate, which results in increased chromatographic resolution, comparable to that of ESI-based MS instruments (Chen *et al.* 2005).

The mass spectrometer first acquires a MS spectrum of intact peptides, from which candidates are selected for fragmentation. In case a peptide meeting specific requirements on its intensity and charge is detected, this peptide precursor is isolated from the others, fragmented, and the resulting fragments then provide a MSMS spectrum. Information acquired from both MS and MSMS spectra is used to identify of the proteins in the original mixture (Nesvizhskii 2006). At present, several searching tools and algorithms are available. Most of these tools are based on the precursor approach (Mascot, Sequest), which uses the precursor mass value as the main search criterion (Clauser *et al.* 1999) and takes both MS and MSMS spectra equally into consideration. On the other hand, the sequence tag approach is based on partial *de novo* peptide sequencing and uses mainly the acquired MSMS spectra (Mann and Wilm 1994).

The most common peptide separation scheme nowadays is based on HPLC, using a stationary C18 reversed phase (RP) column providing excellent resolution. Along with good separation efficiency, an additional advantage of this method is the use of solvents, which do not inhibit either ESI or MALDI type of ionization (Mitulovic and Mechtler 2006). However, a single dimension RP HPLC is not powerful enough to resolve a mixture of hundreds or thousands of various peptides resulting from an enzymatic digestion of a complex protein mixture like plasma (Gilar *et al.* 2009). Therefore, various fractionation and separation methods are combined to simplify the analyzed mixture as much as possible. One of these combined schemes incorporates a strong cation-exchange chromatography (SCX) HPLC prior to the RP HPLC. This approach was denoted as Multidimensional Protein Identification Technology (MudPIT) (Wolters *et al.* 2001). Alternatively, SCX HPLC in the first dimension may be replaced by RP HPLC in basic conditions (Gilar *et al.* 2009) or even by IEF of peptides providing at least comparable fractionation efficiency (Essader *et al.* 2005). The GeLC-MSMS method combines a SDS protein electrophoresis followed by gel cutting, protein digestion and RP HPLC separation of the resulting peptides (Schirle *et al.* 2003). The HUPO Plasma Proteome Project data unambiguously showed that the shotgun approach using these multidimensional separation methods leads to a much higher number of identified proteins than does the 2D-PAGE approach (Omenn *et al.* 2005). Also, a combination of various fractionation and separation methods leads to partially redundant sets of identified proteins. In general, the more orthogonal methods are combined, the higher the number of identified proteins. On the other hand, along with the fraction count, the analysis lengthens proportionally and the procedure becomes more error prone (Hoffman *et al.* 2007).

Quantitative shotgun proteomics

The main goal of former shotgun proteomic studies was mainly protein identification. However, advances in mass spectrometry and bioinformatics enabled a focus shift towards quantitative and comparative analyses where a comparison of mutual protein concentrations in particular samples becomes possible, e.g., affected cell line versus negative control, patients with a specific disease versus healthy donors, etc. Two main quantification strategies are available at

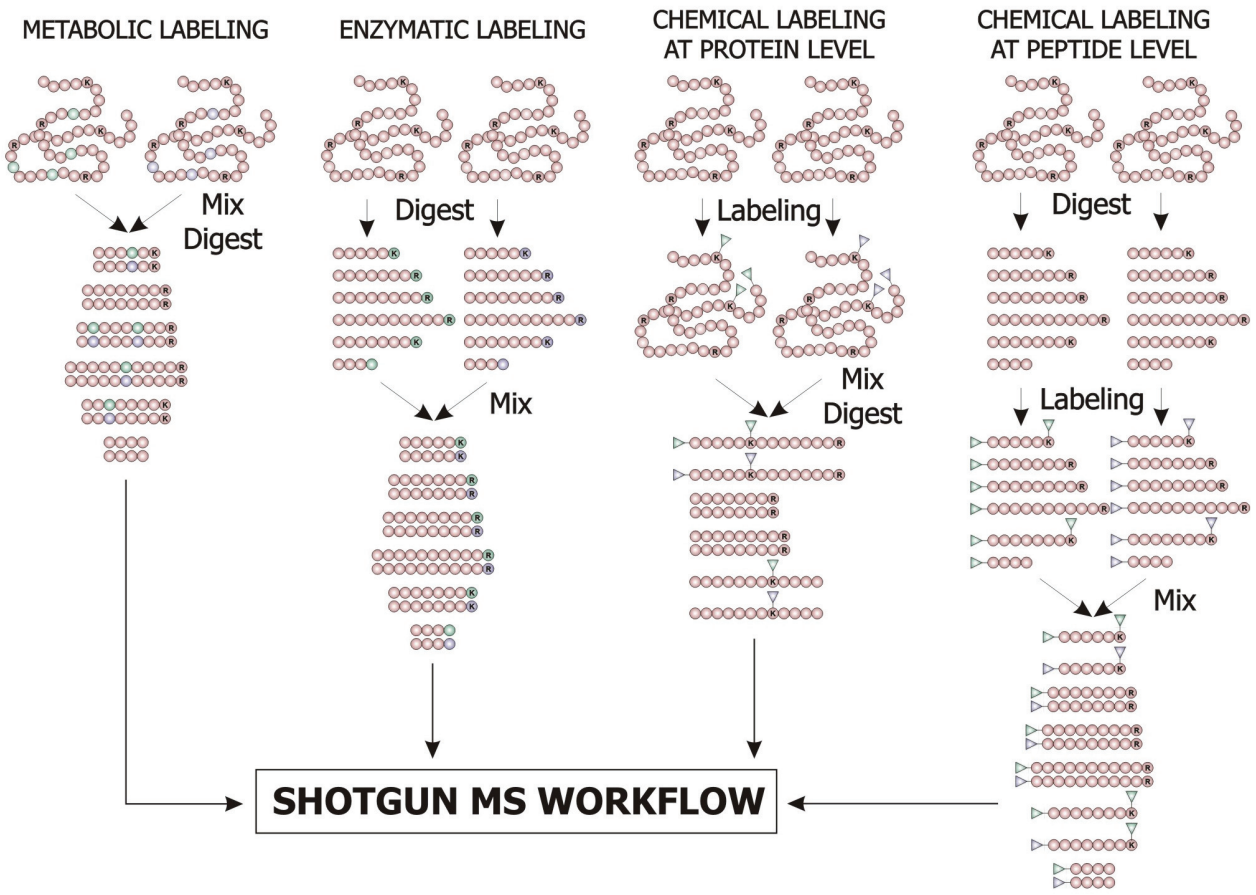


Fig. 4. Stable isotope labeling approaches. **Metabolic labeling**: Two cell cultures are grown in standard medium and a in medium containing heavy isotope labeled amino acids. After cultivation, cells are combined and are analyzed as a single sample. **Enzymatic labeling**: Two protein samples are digested by a sequence specific protease in either light (H_2O^{16}) or heavy (H_2O^{18}) water. Samples may be combined afterwards and processed as one. **Chemical labeling at protein level**: Proteins in two samples to be compared are labeled by ICAT reagents. After labeling, proteins are digested into peptides and combined. **Chemical labeling at peptide level**: Protein samples are digested separately into peptides. After digestion, each peptide sample is labeled by chemical reagents, which have identical chemical structure, but differ in stable isotope composition. After labeling, samples are combined and analyzed.

present: label-free quantification and quantification based on stable isotope labeling.

Label-free quantification

The label-free approach is based on comparison of MS signal intensities between individual experiments (Bondarenko *et al.* 2002). Semi-quantification is also possible to some extent by counting the number of peptides unambiguously identified (Ishihama *et al.* 2005). This method has several evident advantages and possible applications. The labeling step can be omitted, which both shortens and cheapens the experiment. The number of samples to be compared is virtually unlimited, which cannot be rivaled by any of the stable isotope based methods. Also, the spectral complexity is not increased, which could in turn lead to a higher number of identified proteins. Last but not least, label-free approaches are able to quantify throughout a much broader dynamic concentration range than stable isotope-based methods can. However, as different peptides ionize differently during individual experiments, their intensities may vary from run to run, making it rather difficult to correctly quantify them. Therefore, label-free methods are the least accurate, which is caused by the influence of both systematic and random errors during the experiment (Bantscheff *et al.* 2007). Nevertheless, techniques to overcome these shortcomings using bioinformatics and specialized software were suggested recently (Cox and Mann 2008). Thus, label-free quantitative proteomics particularly in combination with high resolution mass spectrometry (FT-ICR, Orbitrap) is regarded as a promising way to quantify large sets of samples even across multiple laboratories.

Labeling based on stable isotopes

Stable isotope strategies were introduced to deal

Table 3. Overview of stable isotope labeling methods in proteomics.

	SILAC	$^{16}O/^{18}O$	ICAT cICAT	Reductive alkylation	NBS	ICPL	TMT	iTRAQ
Labeling	Metabolic	Enzymatic	Chemical	Chemical	Chemical	Chemical	Chemical	Chemical
Labeling level	Proteins	Peptides	Proteins	Proteins Peptides	Peptides	Proteins Peptides	Peptides	Peptides
Target amino acid	L, R, K	C-terminus	C	N-terminus, K	W	N-terminus, K	N-terminus, K	N-terminus, K
Complexity reduction	No	No	Yes	No	Yes	No	No	No
Number of channels	2/3	2	2	2/3	2	2/3	2/6	4/8
Quantification mode	MS	MS	MS	MS	MS	MS	MSMS	MSMS

with the ionization variability of peptides and effect of errors during the workflow (Fig. 4). The samples to be compared can be mixed together and analyzed as a single one, whereas the combination of samples should be carried as soon as possible in the workflow. To distinguish the samples mixed during the analysis, they first need to be labeled with reagents containing stable isotopes, e.g. $^{12}C/^{13}C$, $^{14}N/^{15}N$, $^{16}O/^{18}O$ (Putz *et al.* 2005, Bantscheff *et al.* 2007). The proteins or peptides labeled with a substance of identical chemical structure containing stable isotopes will behave equally during all steps of the experiment since they have identical physicochemical (most importantly ionization and chromatographic) properties, but owing to a specific mass difference in their *m/z*, they can be simply recognized by a mass spectrometer. The quantification is then based on comparison of signal intensities, which differ by a specific molecular mass shift. Based on the nature of the sample, a broad range of quantification methods is at disposal. Stable isotopes may be incorporated into the samples metabolically, enzymatically, or by a chemical reaction (Table 3).

The Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC) method is based on metabolic incorporation of amino acids containing stable isotopes into the protein sequence during cell culture cultivation in a medium containing either light or heavy forms of particular amino acids, e.g. leucine or arginine (Ong *et al.* 2002, Blagoev *et al.* 2004). Owing to the metabolic nature of the labeling, the SILAC method cannot be directly employed in proteomic analyses aimed at plasma biomarker discovery. However, SILAC recently became a basis for a novel combined strategy for biomarker identification called the Stable Isotope Labeled Proteome (SILAP) method (Shah *et al.* 2009, Yu *et al.* 2009). Briefly, a cell model of the studied disease, i.e., pancreatic cancer cell line, is grown in a heavy form of the SILAC cell culture medium. Labeled proteins from these cells that are secreted into the medium are collected and subsequently combined with human plasma samples from patients suffering from pancreatic cancer and with plasma from healthy controls. Due to the heavy isotope labeling, proteins originating from the cell line are recognized in the mass spectrum as they differ by a specific mass shift from the same protein in plasma. The ratios of secretome versus control plasma and secretome versus diseased plasma are then compared, and proteins with altered ratios may than be considered as candidate markers, suitable for subsequent validation.

The next possible point in the shotgun proteomics workflow suitable for labeling is the enzymatic digestion of proteins into peptides, since certain proteases, e.g. trypsin, Glu-C, and Lys-C, catalyze exchange of two oxygen atoms at the C-termini of the peptides by two oxygen atoms coming from solvent water during the reaction (Schnolzer *et al.* 1996). When two protein samples to be compared are digested in $H_2{}^{16}O$ and $H_2{}^{18}O$ separately, the resulting peptides differ by 4 daltons (Da), which is sufficient to recognize peptide pairs properly in the mass spectrum (Heller *et al.* 2003, Havlis and Shevchenko 2004).

The incorporation of stable isotopes by a

chemical reaction represents the largest group of quantification methods. The very first chemical quantification method was the Isotope-Coded Affinity Tags (ICAT) approach, which is based on labeling cysteine-containing peptides via their thiol groups. Light and heavy ICAT labels also contain biotin; therefore the labeled peptides can be isolated using avidin. Due to the fact that approximately one quarter of all tryptic peptides contains cysteine, the enrichment results in significant reduction of the sample complexity (Liu *et al.* 2005). But as majority of proteins contain at least one cysteine in their structure, the information on the originating protein is not lost (Gygi *et al.* 1999). Cystein can be targeted also by other chemistries, such as those based on acrylamide reaction (Faca *et al.* 2008).

The majority of chemical quantification methods incorporate stable isotopes into the peptides using a reaction of $-NH_2$ groups with succinimide derivates. To illustrate, the Isotope-Coded Protein Label Triple (ICPL) method uses N-nicotinoyloxysuccinimide and offers up to three quantification channels (Schmidt *et al.* 2005). NH_2 groups may be also tagged by more stable and even less expensive chemistry based on reductive alkylation using formaldehyde (Boersema *et al.* 2008). In theory, $-NH_2$ groups-targeted labeling covers all the peptides resulting from a protein digest. A significant bottleneck of these techniques emerges during labeling at protein level because the altered side chain of lysine is not recognized by trypsin and thus incomplete cleavage occurs, resulting in fewer and larger peptides. If one wishes to preserve trypsin cleavage rather than select another protease, other functional groups must be tagged at protein level. In this case, however, peptides lacking the target group do not carry quantitative information. On the other hand, by introducing the isolation/enrichment step only or more frequently peptides carrying the tag may be analyzed – lowering the sample complexity as described in ICAT (Gygi *et al.* 1999) or NBS method (Matsuo *et al.* 2009).

Most of the labeling techniques are based on quantification at MS level, where the MS spectra are searched for signals differing by a specific *m/z* shift. The relative concentration of a given peptide is then obtained by comparing the intensities of these corresponding signals.

MS-based quantification techniques enable analysis of a limited number of samples simultaneously, whereas MSMS-based isobaric techniques offer a much higher number of possible quantification channels. The isobaric labels used in these techniques are composed of a reactive group, a reported group, and a balancer group. The sum of molecular weight of these three parts is constant, therefore a labeled peptide is observed as a single peak in MS mode. But as the individual reporter groups differ in molecular weight, the MSMS fragments originating from these reporter groups are observable as distinct peaks. Relative peptide concentration may be acquired by comparing the MSMS signal intensities of these reporter groups.

The Tandem Mass Tags were the first published isobaric technique (Thompson *et al.* 2003). The Isobaric Tag for Relative and Absolute Quantitation (iTRAQ) method, however, has gained greater popularity, as it enables an analysis of up to four samples simultaneously (Ross *et al.* 2004), the newest version even up to eight different samples in one experiment (Pierce *et al.* 2007). During the fragmentation in MSMS mode, the reporter group is released from the modified peptides and can be observed in the mass spectrum as peaks 114.1, 115.1, 116.1 or 117.1 (Fig. 5). Therefore, a tandem mass spectrometer capable of detecting MSMS fragments in low *m/z* range is mandatory. The acquired MSMS spectra are used both for peptide identification and for quantification, where the reporter group signals are used to calculate relative peptide concentrations in particular samples and the remaining fragments originating from the peptide backbone are used for peptide identification. To obtain the reporter ion signal in the spectrum, the reporter group has to be cleaved from the peptide properly. Our data show that the cleavage efficacy varies based on the peptide structure. Nevertheless, because the character of the bond is identical in all four or eight tags, respectively, the cleavage efficacy from a particular peptide is also supposed to remain constant.

Targeted shotgun proteomics

Until recently, MS has been used almost exclusively for the identification of new potential biomarkers, whereas the verification and validation steps were carried out by antibody-based techniques. Recently, a paradigm shift has been apparent, as targeted tandem mass spectrometry also known as Multiple Reaction Monitoring (LC-MRM-MSMS) is increasingly being applied into both verification and validation phases (Lange *et al.* 2008). Interestingly, this targeted approach has been also implemented into multistage strategies for biomarker identification, due to excellent sensitivity and potential to precisely quantify target molecules in complex samples. This is carried out by detecting signature peptides, which

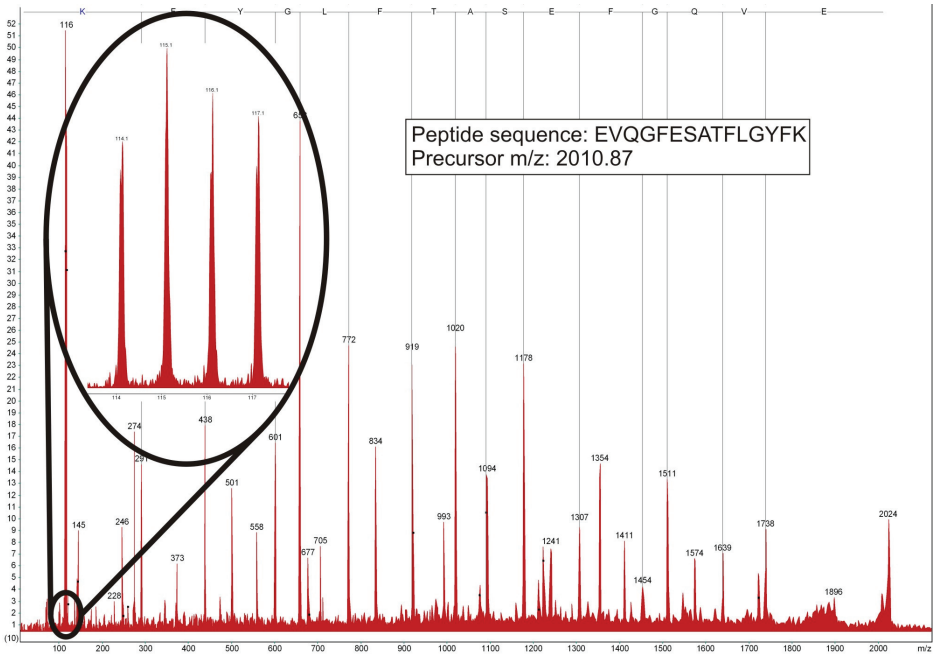


Fig. 5. Representative iTRAQ MSMS spectrum: A peptide of m/z 2010.87 was selected from a MS spectrum for fragmentation analysis. The resulting MSMS spectrum is shown. iTRAQ quantitation information can be read in the low m/z region, as shown in the magnified view. The intensity of each of the four peaks (114.1, 115.1, 116.1 and 117.1) reflects relative concentration of the given peptide in individual four samples which are to be compared. The sequence EVQGFESATFLGYFK was successfully assigned to the MSMS spectrum, resulting in unambiguous identification of Isoform 2 of Gelsolin precursor.

are unique for a given protein, by LC-MRM-MSMS (Anderson and Hunter 2006, Kitteringham *et al.* 2009). Triple quadrupole (QqQ) or hybrid quadrupole-linear ion traps (QTRAP) mass spectrometers used for this purpose are set to select only a specific precursor peptide in the first quadrupole (Q1), which is then fragmented in the collision cell (Q2) and a specific fragment is selected in the third quadrupole (Q3) (Fig. 6). As this cycle takes only a few milliseconds, tens to hundreds of different peptides may be detected and quantified in a targeted manner during a single LC run. More importantly, the detection limit for peptides in this configuration in enhanced by up to 100-fold as opposed to unbiased MS analysis (Keshishian *et al.* 2007). To reliably confirm the identity of monitored peptide, a full MSMS scan upon detecting a defined MRM transition should be acquired (Unwin *et al.* 2009). By this means, the peptide is precisely quantified by the respective chromatographic peak and confirmed by sequence acquisition from the MSMS spectrum (Fig. 7). The actual quantification is carried out by plotting the intensity of Q3 fragment ions on time axis, which results in a chromatographic peak. The most accurate way of quantifying among more samples is realized by introducing a synthetic internal standard peptide, containing a heavy amino acid, into the analyzed sample. As already described in previous chapters, these labeled peptides follow their natural counterparts during all steps of analysis, but owing to a specific mass difference, they can be easily distinguished by the mass spectrometer. The peak area of internal standard peptide, where the precise concentration is know, is compared to the peak area of peptide originating from analyzed sample and finally, absolute concentration may be calculated.

Table 4. Individual phases of a biomarker discovery pipeline (Rifai *et al.* 2006).

Phase I	Exploratory studies to identify candidate marker molecules
Phase II	Qualification – confirmation of differential abundance in samples
Phase III	Verification – assess specificity of candidate molecules
Phase IV	Validation and clinical assay development – large scale studies

The role of proteomics in biomarker candidates verification

Regardless of the method used as the first step of the biomarker discovery process, the resulting candidate markers need to be further intensively proved and tested if they are to become clinically used biomarkers. This is a multistage process and can be regarded as an analogy to the drug discovery pipeline. Starting with a large group of marker candidates, the funnel-like process eliminates low-sensitive and low-specific markers, resulting in a few final candidate molecules. Proportionally to the candidate marker count reduction, the number of tested samples grows steeply along with the project costs (Rifai *et al.* 2006) (Table 4).

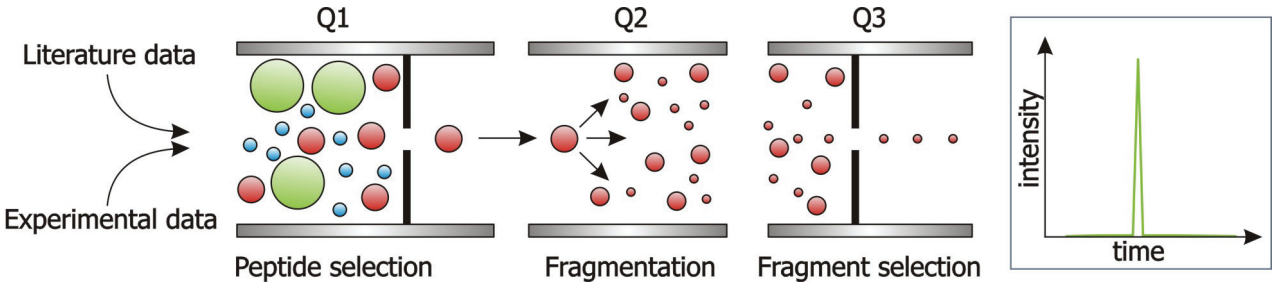


Fig. 6. Multiple Reaction Monitoring scheme. Mass spectrometers used for MRM are set to select only a specific precursor peptide in the first quadrupole (Q1), which is then fragmented in the collision cell (Q2), a specific fragment is selected in the third quadrupole (Q3) and detected. The intensity of the Q3 fragment is then plotted in time, which results in a chromatographic peak correlating with peptide amount in the sample.

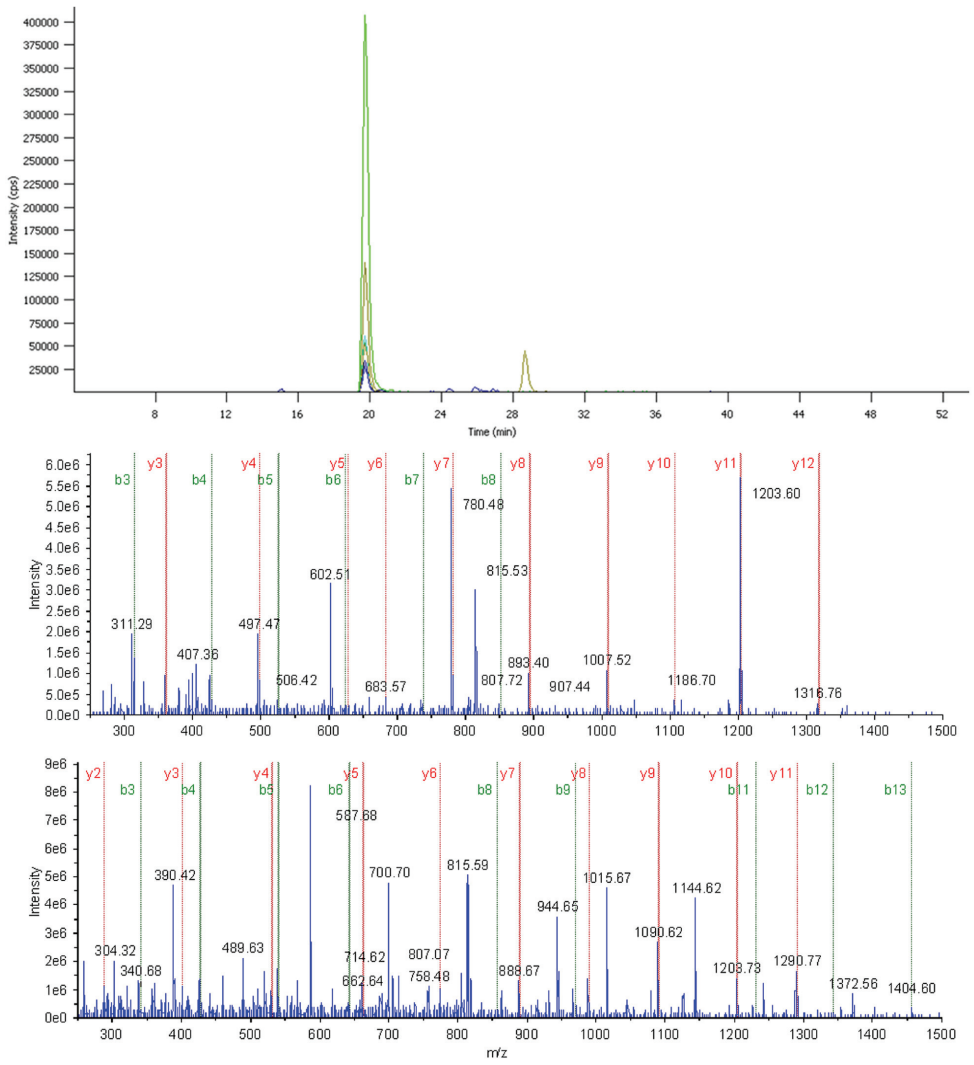


Fig. 7. MRM triggered MSMS. A confident MRM assay should be validated by confirming the identity of the chromatographic MRM peaks by additional acquisition of a MSMS spectrum. The MRM trace shows two chromatographic peaks, each eluting at a different time point. By acquiring a MRM-triggered MSMS spectrum, the targeted earlier eluting signature peptide is unambiguously identified (upper MSMS spectrum), whereas the second peak (lower MSMS spectrum) was proved to originate from a different protein.

After the first phase of the pipeline, the discovery phase, the resulting candidates need to be further proved in the qualification phase, in order to confirm their differential abundance in the tested samples. At this point, the unbiased nature of the analysis changes into a targeted one. New and unproven candidates are analyzed in a targeted manner and precisely quantified in a statistically viable number of serum or plasma samples. Unfortunately, antibodies against these newly discovered candidates are frequently unavailable, and substitutes for antibody-based detection assays (i.e., Western blotting or ELISA) have been sought in proteomics methodologies. Therefore, the method of choice in this phase is LC-MRM-MSMS (Anderson and

Hunter 2006). As the sample preparation and processing is much less extensive than in the discovery phase, the MRM sensitivity is limited by the sample complexity. The limit of quantification in undepleted plasma may reach low μg/ml level (Addona *et al.* 2009). To quantify in the ng/ml range, depletion of ~10 most abundant plasma proteins is required (Keshishian *et al.* 2007). A possible way to further increase the performance of MRM is its coupling to immunoaffinity peptide enrichment (Anderson *et al.* 2004a, Hoofnagle *et al.* 2008), which enhances both sensitivity and specificity, thus allowing analysis in complex matrices with little or no fractionation. However, it requires a specific anti-peptide antibody to be developed against each analyzed peptide. Other antibody-based approaches are unsuitable at this point, due to their low throughput, e.g., Western blotting or high development costs typical for immunoassays.

In the verification phase the specificity of candidates is addressed. The primary objective of verification is to screen potential biomarkers to ensure that only the highest-quality candidates from the discovery phase are taken forward into pre-clinical validation. This requires a larger number of tested samples, which increases approximately by an order. So as to maintain a moderate throughput, the initial candidate list has to be reduced to a few dozens. Immunoassays should be introduced at this point. The lack of high-quality antibodies, however, hinders the fast development of antibody-based assays, as highly specific antibodies are not available for most novel biomarker candidates. Therefore, LC-MRM-MSMS presents a compelling alternative to immunoassays, as it allows a moderate number of candidates to be targeted at a relatively high throughput, without a need of an immunoassay development.

The final phase of the biomarker discovery process, the validation phase, requires a clinical assay to be developed and extensively tested on thousands of clinical samples. A platform change is also required, as MS-based approaches are currently neither able to fulfill the required combination of high throughput and precision, nor are they widely available and accepted by the FDA. Therefore, the development of a suitable antibody-based assay is mandatory (Kingsmore 2006). To meet the required sensitivity, RIA or ELISA are the methods of choice.

Conflict of Interest

There is no conflict of interest.

Acknowledgements

The authors thank Martin Hubalek for critical comments and useful remarks. This article was supported by grant number NR9253-3, provided by Ministry of Health of the Czech Republic.

Abreviations

2D-PAGE – two-dimensional polyacrylamide gel electrophoresis, amu – atomic mass unit, DIGE – differential gel electrophoresis, ELISA – enzyme-linked immunosorbent assay, ESI – electrospray ionization, FDA – Food and Drug Administration, FT-ICR – Fourier-transform ion cyclotron resonance, HPLC – high performance liquid chromatography, HUPO – Human Proteome Organisation, ICAT – isotope-coded affinity tags, ICPL – isotope-coded protein label, IEF – isoelectric focusing, IPG – immobilized pH gradient, IT – ion trap, iTRAQ – isobaric tags for relative and absolute quantitation, LC – liquid chromatography, *m/z* – mass to charge ratio, MALDI – matrix-assisted laser desorption/ionization, MARS – Multiple Affinity Removal System, MRM – multiple reaction monitoring, MS – mass spectrometry, MSMS – tandem mass spectrometry, MW – molecular weight, NBS – 2-nitrobenzenesulfenyl, pI – isoelectric point, PMF – peptide mass fingerprinting, Q – quadrupole, QqQ – triple quadrupole, Q-TRAP – quadrupole-ion trap, RP – reversed-phase, SCX – strong cation exchange, SDS – sodium dodecyl sulfate, SELDI – surface-enhanced laser desorption/ionization, SILAC – stable isotope labeling with amino acids in cell culture, SILAP – stable isotope labeled proteome, TOF – time-of-flight

References

ADDONA TA, ABBATIELLO SE, SCHILLING B, SKATES SJ, MANI DR, BUNK DM, SPIEGELMAN CH, ZIMMERMAN LJ, HAM AJ, KESHISHIAN H, HALL SC, ALLEN S, BLACKMAN RK, BORCHERS CH, BUCK C, CARDASIS HL, CUSACK MP, DODDER NG, GIBSON BW, HELD JM, HILTKE T, JACKSON A, JOHANSEN EB, KINSINGER CR, LI J, MESRI M, NEUBERT TA, NILES RK, PULSIPHER TC, RANSOHOFF D, RODRIGUEZ H, RUDNICK PA, SMITH D, TABB DL, TEGELER TJ, VARIYATH AM, VEGA-MONTOTO LJ, WAHLANDER A, WALDEMARSON S, WANG M, WHITEAKER JR, ZHAO L, ANDERSON NL, FISHER SJ, LIEBLER DC, PAULOVICH AG, REGNIER FE, TEMPST P, CARR SA: Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat Biotechnol* **27**: 633-641, 2009.

ANDERSON L: Candidate-based proteomics in the search for biomarkers of cardiovascular disease. *J Physiol Lond* **563**: 23-60, 2005.

ANDERSON L, HUNTER CL: Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol Cell Proteomics* **5**: 573-588, 2006.

ANDERSON NL, ANDERSON NG: The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* **1**: 845-867, 2002.

ANDERSON NL, ANDERSON NG, HAINES LR, HARDIE DB, OLAFSON RW, PEARSON TW: Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res* **3**: 235-244, 2004a.

ANDERSON NL, POLANSKI M, PIEPER R, GATLIN T, TIRUMALAI RS, CONRADS TP, VEENSTRA TD, ADKINS JN, POUNDS JG, FAGAN R, LOBLEY A: The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol Cell Proteomics* **3**: 311-326, 2004b.

ANDRIOLE GL, CRAWFORD ED, GRUBB RL 3RD, BUYS SS, CHIA D, CHURCH TR, FOUAD MN, GELMANN EP, KVALE PA, REDING DJ, WEISSFELD JL, YOKOCHI LA, O'BRIEN B, CLAPP JD, RATHMELL JM, RILEY TL, HAYES RB, KRAMER BS, IZMIRLIAN G, MILLER AB, PINSKY PF, PROROK PC, GOHAGAN JK, BERG CD: Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* **360**: 1310-1319, 2009.

ATKINSON AJ, COLBURN WA, DEGRUTTOLA VG: Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**: 89-95, 2001.

BANKS RE: Preanalytical influences in clinical proteomic studies: raising awareness of fundamental issues in sample banking. *Clin Chem* **54**: 6-7, 2008.

BANTSCHEFF M, SCHIRLE M, SWEETMAN G, RICK J, KUSTER B: Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* **389**: 1017-1031, 2007.

BAZHIN AV, DALKE C, WILLNER N, ABSCHUTZ O, WILDBERGER HG, PHILIPPOV PP, DUMMER R, GRAW J, DE ANGELIS MH, SCHADENDORF D, UMANSKY V, EICHMULLER SB: Cancer-retina antigens as potential paraneoplastic antigens in melanoma-associated retinopathy. *Int J Cancer* **124**: 140-149, 2009.

BECK LH JR, BONEGIO RG, LAMBEAU G, BECK DM, POWELL DW, CUMMINS TD, KLEIN JB, SALANT DJ: M-type phospholipase A2 receptor as target antigen in idiopathic membranous nephropathy. *N Engl J Med* **361**: 11-21, 2009.

BERGGREN K, CHERNOKALSKAYA E, STEINBERG TH, KEMPER C, LOPEZ MF, DIWU Z, HAUGLAND RP, PATTON WF: Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. *Electrophoresis* **21**: 2509-2521, 2000.

BERTH M, MOSER FM, KOLBE M, BERNHARDT J: The state of the art in the analysis of two-dimensional gel electrophoresis images. *Appl Microbiol Biotechnol* **76**: 1223-1243, 2007.

BIJIAN K, MLYNAREK AM, BALYS RL, JIE S, XU Y, HIER MP, BLACK MJ, DI FALCO MR, LABOISSIERE S, ALAOUI-JAMALI MA: Serum proteomic approach for the identification of serum biomarkers contributed by oral squamous cell carcinoma and host tissue microenvironment. *J Proteome Res* **8**: 2173-2185, 2009.

BJELLQVIST B, EK K, RIGHETTI PG, GIANAZZA E, GORG A, WESTERMEIER R, POSTEL W: Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* **6**: 317-339, 1982.

BLAGOEV B, ONG SE, KRATCHMAROVA I, MANN M: Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol* **22**: 1139-1145, 2004.

BODNAR WM, BLACKBURN RK, KRISE JM, MOSELEY MA: Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J Am Soc Mass Spectrom* **14**: 971-979, 2003.

BOERSEMA PJ, AYE TT, VAN VEEN TA, HECK AJ, MOHAMMED S: Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* **8**: 4624-4632, 2008.

BONDARENKO PV, CHELIUS D, SHALER TA: Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal Chem* **74**: 4741-4749, 2002.

BROOKS SA, CARTER TM, ROYLE L, HARVEY DJ, FRY SA, KINCH C, DWEK RA, RUDD PM: Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies. *Anticancer Agents Med Chem* **8**: 2-21, 2008.

BRUEGEL M, PLANERT M, BAUMANN S, FOCKE A, BERGH FT, LEICHTLE A, CEGLAREK U, THIERY J, FIEDLER GM: Standardized peptidome profiling of human cerebrospinal fluid by magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Proteomics* **72**: 608-615, 2009.

CALLESEN AK, MADSEN JS, VACH W, KRUSE TA, MOGENSEN O, JENSEN ON: Serum protein profiling by solid phase extraction and mass spectrometry: a future diagnostics tool? *Proteomics* **9**: 1428-1441, 2009.

CANDIANO G, BRUSCHI M, MUSANTE L, SANTUCCI L, GHIGGERI GM, CARNEMOLLA B, ORECCHIA P, ZARDI L, RIGHETTI PG: Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* **25**: 1327-1333, 2004.

CARRETTE O, BURKHARD PR, SANCHEZ JC, HOCHSTRASSER DF: State-of-the-art two-dimensional gel electrophoresis: a key tool of proteomics research. *Nat Protoc* **1**: 812-823, 2006.

CHEN HS, REJTAR T, ANDREEV V, MOSKOVETS E, KARGER BL: High-speed, high-resolution monolithic capillary LC-MALDI MS using an off-line continuous deposition interface for proteomic analysis. *Anal Chem* **77**: 2323-2331, 2005.

CHEVALLET M, LUCHE S, DIEMER H, STRUB JM, VAN DORSSELAER A, RABILLOUD T: Sweet silver: A formaldehyde-free silver staining using aldoses as developing agents, with enhanced compatibility with mass spectrometry. *Proteomics* **8**: 4853-4861, 2008.

CHROMY BA, GONZALES AD, PERKINS J, CHOI MW, CORZETT MH, CHANG BC, CORZETT CH, MCCUTCHEN-MALONEY SL: Proteomic analysis of human serum by two-dimensional differential gel electrophoresis after depletion of high-abundant proteins. *J Proteome Res* **3**: 1120-1127, 2004.

CLAUSER KR, BAKER P, BURLINGAME AL: Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* **71**: 2871-2882, 1999.

COLLINS FS, LANDER ES, ROGERS J, WATERSTON RH: Finishing the euchromatic sequence of the human genome. *Nature* **431**: 931-945, 2004.

COMEGYS MM, LIN SH, RAND D, BRITT D, FLANAGAN D, CALLANAN H, BRILLIANT K, HIXSON DC: Two variable regions in carcinoembryonic antigen-related cell adhesion molecule1 N-terminal domains located in or next to monoclonal antibody and adhesion epitopes show evidence of recombination in rat but not in human. *J Biol Chem* **279**: 35063-35078, 2004.

COX J, MANN M: High peptide identification rates and proteome-wide quantitation via novel computational strategies. *Nat Biotechnol* **26**: 1367-1372, 2008.

DATTELBAUM AM, IYER S: Surface-assisted laser desorption/ionization mass spectrometry. *Expert Rev Proteomics* **3**: 153-161, 2006.

DESOUZA LV, ROMASCHIN AD, COLGAN TJ, SIU KW: Absolute quantification of potential cancer markers in clinical tissue homogenates using multiple reaction monitoring on a hybrid triple quadrupole/linear ion trap tandem mass spectrometer. *Anal Chem* **81**: 3462-3470, 2009.

DESOUZA LV, TAYLOR AM, LI W, MINKOFF MS, ROMASCHIN AD, COLGAN TJ, SIU KW: Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J Proteome Res* **7**: 3525-3534, 2008.

DOMON B, AEBERSOLD R: Mass spectrometry and protein analysis. *Science* **312**: 212-217, 2006.

ECHAN LA, TANG HY, ALI-KHAN N, LEE K, SPEICHER DW: Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **5**: 3292-3303, 2005.

ESSADER AS, CARGILE BJ, BUNDY JL, STEPHENSON JL JR: A comparison of immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics. *Proteomics* **5**: 24-34, 2005.

ETZIONI R, URBAN N, RAMSEY S, MCINTOSH M, SCHWARTZ S, REID B, RADICH J, ANDERSON G, HARTWELL L: The case for early detection. *Nat Rev Cancer* **3**: 243-252, 2003.

FACA VM, SONG KS, WANG H, ZHANG Q, KRASNOSELSKY AL, NEWCOMB LF, PLENTZ RR, GURUMURTHY S, REDSTON MS, PITTERI SJ, PEREIRA-FACA SR, IRETON RC, KATAYAMA H, GLUKHOVA V, PHANSTIEL D, BRENNER DE, ANDERSON MA, MISEK D, SCHOLLER N, URBAN ND, BARNETT MJ, EDELSTEIN C, GOODMAN GE, THORNQUIST MD, MCINTOSH MW, DEPINHO RA, BARDEESY N, HANASH SM: A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med* **5**: e123, 2008.

FIEDLER GM, LEICHTLE AB, KASE J, BAUMANN S, CEGLAREK U, FELIX K, CONRAD T, WITZIGMANN H, WEIMANN A, SCHUTTE C, HAUSS J, BUCHLER M, THIERY J: Serum peptidome profiling revealed platelet factor 4 as a potential discriminating Peptide associated with pancreatic cancer. *Clin Cancer Res* **15**: 3812-3819, 2009.

FINDEISEN P, ZAPATKA M, PECCERELLA T, MATZK H, NEUMAIER M, SCHADENDORF D, UGUREL S: Serum amyloid A as a prognostic marker in melanoma identified by proteomic profiling. *J Clin Oncol* **27**: 2199-2208, 2009.

GILAR M, OLIVOVA P, CHAKRABORTY AB, JAWORSKI A, GEROMANOS SJ, GEBLER JC: Comparison of 1-D and 2-D LC MS/MS methods for proteomic analysis of human serum. *Electrophoresis* **30**: 1157-1167, 2009.

GOLDMAN R, RESSOM HW, ABDEL-HAMID M, GOLDMAN L, WANG A, VARGHESE RS, AN Y, LOFFREDO CA, DRAKE SK, EISSA SA, GOUDA I, EZZAT S, MOISEIWITSCH FS: Candidate markers for the detection of hepatocellular carcinoma in low-molecular weight fraction of serum. *Carcinogenesis* **28**: 2149-2153, 2007.

GOLDMAN R, RESSOM HW, VARGHESE RS, GOLDMAN L, BASCUG G, LOFFREDO CA, ABDEL-HAMID M, GOUDA I, EZZAT S, KYSELOVA Z, MECHREF Y, NOVOTNY MV: Detection of hepatocellular carcinoma using glycomic analysis. *Clin Cancer Res* **15**: 1808-1813, 2009.

GORG A, WEISS W, DUNN MJ: Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **4**: 3665-3685, 2004.

GOVORUKHINA NI, DE VRIES M, REIJMERS TH, HORVATOVICH P, VAN DER ZEE AG, BISCHOFF R: Influence of clotting time on the protein composition of serum samples based on LC-MS data. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**: 1281-1291, 2009.

GRAMOLINI AO, KISLINGER T, ALIKHANI-KOOPAEI R, FONG V, THOMPSON NJ, ISSERLIN R, SHARMA P, OUDIT GY, TRIVIERI MG, FAGAN A, KANNAN A, HIGGINS DG, HUEDIG H, HESS G, ARAB S, SEIDMAN JG, SEIDMAN CE, FREY B, PERRY M, BACKX PH, LIU PP, MACLENNAN DH, EMILI A: Comparative proteomics profiling of a phospholamban mutant mouse model of dilated cardiomyopathy reveals progressive intracellular stress responses. *Mol Cell Proteomics* **7**: 519-533, 2008.

GUO Y, FU Z, VAN EYK JE: A proteomic primer for the clinician. *Proc Am Thorac Soc* **4**: 9-17, 2007.

GYGI SP, RIST B, GERBER SA, TURECEK F, GELB MH, AEBERSOLD R: Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**: 994-999, 1999.

HANASH SM, PITTERI SJ, FACA VM: Mining the plasma proteome for cancer biomarkers. *Nature* **452**: 571-579, 2008.

HASHIGUCHI T, TANAKA K, LEE LJ, SASAKI K, NATSUGOE S, KAWAHARA KI, ARIMURA K, MARUYAMA I: Diagnostic value of serum peptidome analyses for protease activated pathological conditions beyond cancer diagnosis. *Med Hypotheses* **73**: 760-763, 2009.

HAVLIS J, SHEVCHENKO A: Absolute quantification of proteins in solutions and in polyacrylamide gels by mass spectrometry. *Anal Chem* **76**: 3029-3036, 2004.

HELLER M, MATTOU H, MENZEL C, YAO X: Trypsin catalyzed 16O-to-18O exchange for comparative proteomics: tandem mass spectrometry comparison using MALDI-TOF, ESI-QTOF, and ESI-ion trap mass spectrometers. *J Am Soc Mass Spectrom* **14**: 704-718, 2003.

HENZEL WJ, WATANABE C, STULTS JT: Protein identification: the origins of peptide mass fingerprinting. *J Am Soc Mass Spectrom* **14**: 931-942, 2003.

HOFFMAN SA, JOO WA, ECHAN LA, SPEICHER DW: Higher dimensional (Hi-D) separation strategies dramatically improve the potential for cancer biomarker detection in serum and plasma. *J Chromatogr B Analyt Technol Biomed Life Sci* **849**: 43-52, 2007.

HONGSACHART P, HUANG-LIU R, SINCHAIKUL S, PAN FM, PHUTRAKUL S, CHUANG YM, YU CJ, CHEN ST: Glycoproteomic analysis of WGA-bound glycoprotein biomarkers in sera from patients with lung adenocarcinoma. *Electrophoresis* **30**: 1206-1220, 2009.

HOOFNAGLE AN, BECKER JO, WENER MH, HEINECKE JW: Quantification of thyroglobulin, a low-abundance serum protein, by immunoaffinity peptide enrichment and tandem mass spectrometry. *Clin Chem* **54**: 1796-1804, 2008.

HU S, LOO JA, WONG DT: Human body fluid proteome analysis. *Proteomics* **6**: 6326-6353, 2006.

HUANG HL, STASYK T, MORANDELL S, MOGG M, SCHREIBER M, FEUERSTEIN I, HUCK CW, STECHER G, BONN GK, HUBER LA: Enrichment of low-abundant serum proteins by albumin/immunoglobulin G immunoaffinity depletion under partly denaturing conditions. *Electrophoresis* **26**: 2843-2849, 2005a.

HUANG L, FANG X: Immunoaffinity fractionation of plasma proteins by chicken IgY antibodies. *Methods Mol Biol* **425**: 41-51, 2008.

HUANG L, HARVIE G, FEITELSON JS, GRAMATIKOFF K, HEROLD DA, ALLEN DL, AMUNNGAMA R, HAGLER RA, PISANO MR, ZHANG WW, FANG X: Immunoaffinity separation of plasma proteins by IgY microbeads: meeting the needs of proteomic sample preparation and analysis. *Proteomics* **5**: 3314-3328, 2005b.

ISHIHAMA Y, ODA Y, TABATA T, SATO T, NAGASU T, RAPPSILBER J, MANN M: Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* **4**: 1265-1272, 2005.

KESHISHIAN H, ADDONA T, BURGESS M, KUHN E, CARR SA: Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* **6**: 2212-2229, 2007.

KESHISHIAN H, ADDONA T, BURGESS M, MANI DR, SHI X, KUHN E, SABATINE MS, GERSZTEN RE, CARR SA: Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* **8**: 2339-2349, 2009.

KIM HJ, KANG HJ, LEE H, LEE ST, YU MH, KIM H, LEE C: Identification of S100A8 and S100A9 as serological markers for colorectal cancer. *J Proteome Res* **8**: 1368-1379, 2009.

KINGSMORE SF: Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nat Rev Drug Discov* **5**: 310-320, 2006.

KITTERINGHAM NR, JENKINS RE, LANE CS, ELLIOTT VL, PARK BK: Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**: 1229-1239, 2009.

KUI WONG N, EASTON RL, PANICO M, SUTTON-SMITH M, MORRISON JC, LATTANZIO FA, MORRIS HR, CLARK GF, DELL A, PATANKAR MS: Characterization of the oligosaccharides associated with the human ovarian tumor marker CA125. *J Biol Chem* **278**: 28619-28634, 2003.

KULASINGAM V, DIAMANDIS EP: Tissue culture-based breast cancer biomarker discovery platform. *Int J Cancer* **123**: 2007-2012, 2008.

KYSELOVA Z, MECHREF Y, KANG P, GOETZ JA, DOBROLECKI LE, SLEDGE GW, SCHNAPER L, HICKEY RJ, MALKAS LH, NOVOTNY MV: Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem* **54**: 1166-1175, 2008.

LAEMMLI UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685, 1970.

LANGE V, PICOTTI P, DOMON B, AEBERSOLD R: Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* **4**: 222, 2008.

LAU TK, LEUNG TN: Genetic screening and diagnosis. *Curr Opin Obstet Gynecol* **17**: 163-169, 2005.

LIU T, QIAN WJ, CHEN WN, JACOBS JM, MOORE RJ, ANDERSON DJ, GRITSENKO MA, MONROE ME, THRALL BD, CAMP DG 2ND, SMITH RD: Improved proteome coverage by using high efficiency cysteinyl peptide enrichment: the human mammary epithelial cell proteome. *Proteomics* **5**: 1263-1273, 2005.

LIU T, QIAN WJ, MOTTAZ HM, GRITSENKO MA, NORBECK AD, MOORE RJ, PURVINE SO, CAMP DG 2ND, SMITH RD: Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol Cell Proteomics* **5**: 2167-2174, 2006.

LOOI KS, NAKAYASU ES, DIAZ RA, TAN EM, ALMEIDA IC, ZHANG JY: Using proteomic approach to identify tumor-associated antigens as markers in hepatocellular carcinoma. *J Proteome Res* **7**: 4004-4012, 2008.

LUDWIG JA, WEINSTEIN JN: Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* **5**: 845-856, 2005.

MACKINTOSH JA, CHOI HY, BAE SH, VEAL DA, BELL PJ, FERRARI BC, VAN DYK DD, VERRILLS NM, PAIK YK, KARUSO P: A fluorescent natural product for ultra sensitive detection of proteins in one-dimensional and two-dimensional gel electrophoresis. *Proteomics* **3**: 2273-2288, 2003.

MANN M, WILM M: Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* **66**: 4390-4399, 1994.

MATSUO EI, WATANABE M, KUYAMA H, NISHIMURA O: A new strategy for protein biomarker discovery utilizing 2-nitrobenzenesulfenyl (NBS) reagent and its applications to clinical samples. *J Chromatogr B Analyt Technol Biomed Life Sci* **877**: 2607-2614, 2009.

MECHREF Y, MADERA M, NOVOTNY MV: Glycoprotein enrichment through lectin affinity techniques. *Methods Mol Biol* **424**: 373-396, 2008.

MILLER I, CRAWFORD J, GIANAZZA E: Protein stains for proteomic applications: which, when, why? *Proteomics* **6**: 5385-5408, 2006.

MITULOVIC G, MECHTLER K: HPLC techniques for proteomics analysis--a short overview of latest developments. *Brief Funct Genomic Proteomic* **5**: 249-260, 2006.

NESVIZHSKII AI: Protein identification by tandem mass spectrometry and sequence database searching. *Methods Mol Biol* **367**: 87-120, 2006.

NEUHOFF V, AROLD N, TAUBE D, EHRHARDT W: Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* **9**: 255-262, 1988.

O'FARRELL PH: High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **250**: 4007-4021, 1975.

OMENN GS, STATES DJ, ADAMSKI M, BLACKWELL TW, MENON R, HERMJAKOB H, APWEILER R, HAAB BB, SIMPSON RJ, EDDES JS, KAPP EA, MORITZ RL, CHAN DW, RAI AJ, ADMON A, AEBERSOLD R, ENG J, HANCOCK WS, HEFTA SA, MEYER H, PAIK YK, YOO JS, PING P, POUNDS J, ADKINS J, QIAN X, WANG R, WASINGER V, WU CY, ZHAO X, ZENG R, ARCHAKOV A, TSUGITA A, BEER I, PANDEY A, PISANO M, ANDREWS P, TAMMEN H, SPEICHER DW, HANASH SM: Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* **5**: 3226-3245, 2005.

ONG SE, BLAGOEV B, KRATCHMAROVA I, KRISTENSEN DB, STEEN H, PANDEY A, MANN M: Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**: 376-386, 2002.

PAPPIN DJ, HOJRUP P, BLEASBY AJ: Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* **3**: 327-332, 1993.

PIERCE A, UNWIN RD, EVANS CA, GRIFFITHS S, CARNEY L, ZHANG L, JAWORSKA E, LEE CF, BLINCO D, OKONIEWSKI MJ, MILLER CJ, BITTON DA, SPOONCER E, WHETTON AD: Eight-channel iTRAQ enables comparison of the activity of 6 leukaemogenic tyrosine kinases. *Mol Cell Proteomics* **7**: 853-863, 2008.

POON TC: Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. *Expert Rev Proteomics* **4**: 51-65, 2007.

PUTZ S, REINDERS J, REINDERS Y, SICKMANN A: Mass spectrometry-based peptide quantification: applications and limitations. *Expert Rev Proteomics* **2**: 381-392, 2005.

QUINTANA LF, CAMPISTOL JM, ALCOLEA MP, BANON-MANEUS E, SOL-GONZALEZ A, CUTILLAS PR: Application of label-free quantitative peptidomics for the identification of urinary biomarkers of kidney chronic allograft dysfunction. *Mol Cell Proteomics* **8**: 1658-1673, 2009.

RABILLOUD T, VUILLARD L, GILLY C, LAWRENCE JJ: Silver-staining of proteins in polyacrylamide gels: a general overview. *Cell Mol Biol (Noisy-le-grand)* **40**: 57-75, 1994.

RAI AJ, GELFAND CA, HAYWOOD BC, WARUNEK DJ, YI J, SCHUCHARD MD, MEHIGH RJ, COCKRILL SL, SCOTT GB, TAMMEN H, SCHULZ-KNAPPE P, SPEICHER DW, VITZTHUM F, HAAB BB, SIEST G, CHAN DW: HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. *Proteomics* **5**: 3262-3277, 2005.

RIFAI N, GILLETTE MA, CARR SA: Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* **24**: 971-983, 2006.

RIGHETTI PG, BOSCHETTI E: Sherlock Holmes and the proteome--a detective story. *FEBS J* **274**: 897-905, 2007.

RIGHETTI PG, BOSCHETTI E, LOMAS L, CITTERIO A: Protein Equalizer Technology: the quest for a "democratic proteome". *Proteomics* **6**: 3980-3992, 2006.

ROSA DD, ISMAEL G, LAGO LD, AWADA A: Molecular-targeted therapies: lessons from years of clinical development. *Cancer Treat Rev* **34**: 61-80, 2008.

ROSS PL, HUANG YN, MARCHESE JN, WILLIAMSON B, PARKER K, HATTAN S, KHAINOVSKI N, PILLAI S, DEY S, DANIELS S, PURKAYASTHA S, JUHASZ P, MARTIN S, BARTLET-JONES M, HE F, JACOBSON A, PAPPIN DJ: Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**: 1154-1169, 2004.

ROUX-DALVAI F, GONZALEZ DE PEREDO A, SIMO C, GUERRIER L, BOUYSSIE D, ZANELLA A, CITTERIO A, BURLET-SCHILTZ O, BOSCHETTI E, RIGHETTI PG, MONSARRAT B: Extensive analysis of the cytoplasmic proteome of human erythrocytes using the peptide ligand library technology and advanced mass spectrometry. *Mol Cell Proteomics* **7**: 2254-2269, 2008.

SCHIESS R, WOLLSCHEID B, AEBERSOLD R: Targeted proteomic strategy for clinical biomarker discovery. *Mol Oncol* **3**: 33-44, 2009.

SCHIRLE M, HEURTIER MA, KUSTER B: Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics* **2**: 1297-1305, 2003.

SCHMIDT A, KELLERMANN J, LOTTSPEICH F: A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5**: 4-15, 2005.

SCHNOLZER M, JEDRZEJEWSKI P, LEHMANN WD: Protease-catalyzed incorporation of 18O into peptide fragments and its application for protein sequencing by electrospray and matrix-assisted laser desorption/ionization mass spectrometry. *Electrophoresis* **17**: 945-953, 1996.

SCHRODER FH, HUGOSSON J, ROOBOL MJ, TAMMELA TL, CIATTO S, NELEN V, KWIATKOWSKI M, LUJAN M, LILJA H, ZAPPA M, DENIS LJ, RECKER F, BERENGUER A, MAATTANEN L, BANGMA CH, AUS G, VILLERS A, REBILLARD X, VAN DER KWAST T, BLIJENBERG BG, MOSS SM, DE KONING HJ, AUVINEN A: Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* **360**: 1320-1328, 2009.

SENNELS L, SALEK M, LOMAS L, BOSCHETTI E, RIGHETTI PG, RAPPSILBER J: Proteomic analysis of human blood serum using peptide library beads. *J Proteome Res* **6**: 4055-4062, 2007.

SHAH SJ, YU KH, SANGAR V, PARRY SI, BLAIR IA: Identification and quantification of preterm birth biomarkers in human cervicovaginal fluid by liquid chromatography/tandem mass spectrometry. *J Proteome Res* **8**: 2407-2417, 2009.

SHEVCHENKO A, JENSEN ON, PODTELEJNIKOV AV, SAGLIOCCO F, WILM M, VORM O, MORTENSEN P, SHEVCHENKO A, BOUCHERIE H, MANN M: Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc Natl Acad Sci USA* **93**: 14440-14445, 1996.

SPIRO RG: Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**: 43R-56R, 2002.

STATES DJ, OMENN GS, BLACKWELL TW, FERMIN D, ENG J, SPEICHER DW, HANASH SM: Challenges in deriving high-confidence protein identifications from data gathered by a HUPO plasma proteome collaborative study. *Nat Biotechnol* **24**: 333-338, 2006.

STESSL M, NOE CR, LACHMANN B: Influence of image-analysis software on quantitation of two-dimensional gel electrophoresis data. *Electrophoresis* **30**: 325-328, 2009.

TAJIRI M, OHYAMA C, WADA Y: Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: a glycopeptide approach. *Glycobiology* **18**: 2-8, 2008.

TAM SW, PIRRO J, HINERFELD D: Depletion and fractionation technologies in plasma proteomic analysis. *Expert Rev Proteomics* **1**: 411-420, 2004.

THIEDE B, HOHENWARTER W, KRAH A, MATTOW J, SCHMID M, SCHMIDT F, JUNGBLUT PR: Peptide mass fingerprinting. *Methods* **35**: 237-247, 2005.

THOMPSON A, SCHAFER J, KUHN K, KIENLE S, SCHWARZ J, SCHMIDT G, NEUMANN T, JOHNSTONE R, MOHAMMED AK, HAMON C: Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**: 1895-1904, 2003.

THULASIRAMAN V, LIN S, GHEORGHIU L, LATHROP J, LOMAS L, HAMMOND D, BOSCHETTI E: Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. *Electrophoresis* **26**: 3561-3571, 2005.

TIAN Y, ZHOU Y, ELLIOTT S, AEBERSOLD R, ZHANG H: Solid-phase extraction of N-linked glycopeptides. *Nat Protoc* **2**: 334-339, 2007.

TUCHOLSKA M, BOWDEN P, JACKS K, ZHU P, FURESZ S, DUMBROVSKY M, MARSHALL J: Human serum proteins fractionated by preparative partition chromatography prior to LC-ESI-MS/MS. *J Proteome Res* **8**: 1143-1155, 2009.

TUMANI H, PFEIFLE M, LEHMENSIEK V, RAU D, MOGEL H, LUDOLPH AC, BRETTSCHNEIDER J: Candidate biomarkers of chronic inflammatory demyelinating polyneuropathy (CIDP): Proteome analysis of cerebrospinal fluid. *J Neuroimmunol* **214**: 109-112, 2009.

UNLU M, MORGAN ME, MINDEN JS: Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**: 2071-2077, 1997.

UNWIN RD, GRIFFITHS JR, WHETTON AD: A sensitive mass spectrometric method for hypothesis-driven detection of peptide post-translational modifications: multiple reaction monitoring-initiated detection and sequencing (MIDAS). *Nat Protoc* **4**: 870-877, 2009.

VEENSTRA TD, CONRADS TP, HOOD BL, AVELLINO AM, ELLENBOGEN RG, MORRISON RS: Biomarkers: mining the biofluid proteome. *Mol Cell Proteomics* **4**: 409-418, 2005.

VILLANUEVA J, PHILIP J, ENTENBERG D, CHAPARRO CA, TANWAR MK, HOLLAND EC, TEMPST P: Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. *Anal Chem* **76**: 1560-1570, 2004.

VILLANUEVA J, SHAFFER DR, PHILIP J, CHAPARRO CA, ERDJUMENT-BROMAGE H, OLSHEN AB, FLEISHER M, LILJA H, BROGI E, BOYD J, SANCHEZ-CARBAYO M, HOLLAND EC, CORDON-CARDO C, SCHER HI, TEMPST P: Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Invest* **116**: 271-284, 2006.

WADA Y, AZADI P, COSTELLO CE, DELL A, DWEK RA, GEYER H, GEYER R, KAKEHI K, KARLSSON NG, KATO K, KAWASAKI N, KHOO KH, KIM S, KONDO A, LATTOVA E, MECHREF Y, MIYOSHI E, NAKAMURA K, NARIMATSU H, NOVOTNY MV, PACKER NH, PERREAULT H, PETER-KATALINIC J, POHLENTZ G, REINHOLD VN, RUDD PM, SUZUKI A, TANIGUCHI N: Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* **17**: 411-422, 2007.

WASINGER VC, CORDWELL SJ, CERPA-POLJAK A, YAN JX, GOOLEY AA, WILKINS MR, DUNCAN MW, HARRIS R, WILLIAMS KL, HUMPHERY-SMITH I: Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **16**: 1090-1094, 1995.

WEI YS, ZHENG YH, LIANG WB, ZHANG JZ, YANG ZH, LV ML, JIA J, ZHANG L: Identification of serum biomarkers for nasopharyngeal carcinoma by proteomic analysis. *Cancer* **112**: 544-551, 2008.

WHITEAKER JR, ZHANG H, ZHAO L, WANG P, KELLY-SPRATT KS, IVEY RG, PIENING BD, FENG LC, KASARDA E, GURLEY KE, ENG JK, CHODOSH LA, KEMP CJ, MCINTOSH MW, PAULOVICH AG: Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer. *J Proteome Res* **6**: 3962-3975, 2007.

WOLLSCHEID B, BAUSCH-FLUCK D, HENDERSON C, O'BRIEN R, BIBEL M, SCHIESS R, AEBERSOLD R, WATTS JD: Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat Biotechnol* **27**: 378-386, 2009.

WOLTERS DA, WASHBURN MP, YATES JR 3RD: An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* **73**: 5683-5690, 2001.

WORKMAN P: The opportunities and challenges of personalized genome-based molecular therapies for cancer: targets, technologies, and molecular chaperones. *Cancer Chemother Pharmacol* **52** (Suppl 1): S45-S56, 2003.

XU G, XIANG CQ, LU Y, KANG XN, LIAO P, DING Q, ZHANG YF: Application of SELDI-TOF-MS to identify serum biomarkers for renal cell carcinoma. *Cancer Lett* **282**: 205-213, 2009.

YOUNG LS, SEARLE PF, ONION D, MAUTNER V: Viral gene therapy strategies: from basic science to clinical application. *J Pathol* **208**: 299-318, 2006.

YU KH, BARRY CG, AUSTIN D, BUSCH CM, SANGAR V, RUSTGI AK, BLAIR IA: Stable isotope dilution multidimensional liquid chromatography-tandem mass spectrometry for pancreatic cancer serum biomarker discovery. *J Proteome Res* **8**: 1565-1576, 2009.

ZHANG H, LIU AY, LORIAUX P, WOLLSCHEID B, ZHOU Y, WATTS JD, AEBERSOLD R: Mass spectrometric detection of tissue proteins in plasma. *Mol Cell Proteomics* **6**: 64-71, 2007a.

ZHANG X, LI L, WEI D, YAP Y, CHEN F: Moving cancer diagnostics from bench to bedside. *Trends Biotechnol* **25**: 166-173, 2007b.

ZINKIN NT, GRALL F, BHASKAR K, OTU HH, SPENTZOS D, KALMOWITZ B, WELLS M, GUERRERO M, ASARA JM, LIBERMANN TA, AFDHAL NH: Serum proteomics and biomarkers in hepatocellular carcinoma and chronic liver disease. *Clin Cancer Res* **14**: 470-477, 2008.

ATTACHED PAPERS
PAPER II.

JOURNAL OF PROTEOMICS XX (2011) XXX–XXX

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot

CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantifying hidden proteomes

Vojtech Tambor[a], Christie L. Hunter[b], Sean L. Seymour[b], Marian Kacerovsky[c], Jiri Stulik[a], Juraj Lenco[a,*]

[a]Institute of Molecular Pathology, Faculty of Military Health Sciences, University of Defence, Hradec Kralove, Czech Republic
[b]AB SCIEX, Foster City, CA, USA
[c]Department of Obstetrics and Gynecology, Charles University in Prague, Faculty of Medicine Hradec Kralove, University Hospital Hradec Kralove, Czech Republic

ARTICLE INFO

Article history:
Received 19 August 2011
Accepted 30 September 2011

Keywords:
Cysteinyl peptide enrichment
iTRAQ reagents
Proteome coverage
Amniotic fluid proteome

ABSTRACT

Shotgun proteomics is capable of characterizing differences in both protein quality and quantity, and has been applied in various biomedical applications. Unfortunately, the high complexity and dynamic range of proteins in studied samples, clinical in particular, often hinders the identification of relevant proteins. Indeed, information-rich, low abundance proteins often remain undetected, whereas repeatedly reported altered concentrations in high abundance proteins are often ambiguous and insignificant. Several techniques have therefore been developed to overcome this obstacle and provide a deeper insight into the proteome. Here we report a novel approach, which enables iTRAQ reagent quantitation of peptides fractionated based on presence of a cysteine residue (thus CysTRAQ). For the first time, we prove that iTRAQ quantitation is fully compatible with cysteinyl peptide enrichment and is not influenced by the fractionation process. Moreover, the employment of the method combined with high-resolution TripleTOF 5600 mass spectrometer for very fast MS/MS acquisition in human amniotic fluid analysis significantly increased the number of identified proteins, which were simultaneously quantified owing to the introduction of iTRAQ labeling. We herein show that CysTRAQ is a robust and straightforward method with potential application in quantitative proteomics experiments, *i.e.* as an alternative to the ICAT reagent approach.

1. Introduction

In recent years, various proteomic technologies have been productively applied to clinical proteomics [1]. This however, is still a very challenging area of research as many disease-related and relevant proteins are hidden among the high abundance ones [2]. Substantial demands are placed on fractionation and separation techniques since even the latest proteomic technologies are not able to comprehensively interrogate a complex protein digest, and moreover, cannot cover more than 4 orders of concentration range. This is far below the expected dynamic concentration range of most of clinical samples, such as tissues or body fluids [3]. As a consequence, technologies reducing the sample complexity, and

* *Corresponding author at:* Institute of Molecular Pathology, Faculty of Military Health Sciences, University of Defence, Trebesska 1575, 50001 Hradec Kralove, Czech Republic. Tel.: +420 973 253 223; fax: +420 495 513 018.
E-mail address: lenco@pmfhk.cz (J. Lenco).

1874-3919/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.jprot.2011.09.027

thus increasing the likelihood of uncovering low-abundant proteins are required to address these challenges [4,5]. Such technologies may involve multiple rounds of fractionation both at the protein level and peptide level after digestion.

Along with separation technologies based on physico-chemical properties of proteins/peptides, techniques aimed exclusively at chemical properties have been introduced to isolate surrogate peptide representatives of proteins. To illustrate, glycopeptides may be pulled down from a protein digest using hydrazide [6] or boronic acid chemistry [7]. Researchers have also dealt with enrichment based on targeting protein *N*-terminal peptides [8] and *C*-terminal peptides [9]. Even the presence of a specific amino acid at the peptide *N*-terminus has been utilized [10]. Most tagging methods allow for specific enrichment of peptides containing a particular amino acid such as tryptophan [11], methionine [12], arginine or histidine [13]. However, due to the reactivity, simplicity of reaction and ease of reaction reversibility, cysteine is the favorite amino acid for this strategy. An *in-silico* digest of all human proteins revealed that merely 15% of all human tryptic peptides with molecular mass spanning from 800 to 3000 Da contain at least one cysteine in their sequence. As a consequence, specific enrichment of these peptides results in considerable reduction of sample complexity and therefore similar number of MS/MS events during a proteomic analysis may lead to increased number of identified proteins. On the other hand, due to the fact that 97% of all human proteins contain at least one tryptic peptide with cysteine, the major part of the proteome is still represented in such simplified mixture [14]. Moreover, a recent work combining fractionation based on cysteines and multidimensional separation in a concentration-annotated yeast standard proteome study has clearly shown increased detection of low abundance proteins [15]. All these facts and features make cysteine an ideal candidate for enrichment strategies to improve the proteome coverage. Indeed, many fractionation schemes target cysteine in order to reduce sample complexity, including the ICAT approach [16], Reversible Biotinylation of Cysteinyl Peptides technique [17], Cysteine-reactive Covalent Capture Tags [18], the Cysteinyl Peptide Enrichment (CPE) method [5] and the recently introduced Cysteine-Reactive Tandem Mass Tags [19].

Introduction of quantitative proteomic techniques based on stable isotope labeling has made simplification of a sample *via* separation techniques even more important. For MS mode based quantification, lower sample complexity may avoid errors caused by overlapping peaks (SILAC, enzymatic $^{16}O/^{18}O$ labeling). Within MS/MS mode based quantification methods (iTRAQ reagents, Tandem Mass Tags (TMT)), the precursor ion selection is not absolute. Thus, upon fragmentation, peptides close to the *m/z* of a particular precursor may contribute to the reporter ions, leading to smoothing of the quantitative ratios. By decreasing the sample complexity, more accurate quantitation may be achieved from the reporter ion intensities.

Our goal was to develop a technique, which would enable both sample simplification and quantitation, while being applicable to proteomic analysis of clinical samples, *i.e.* amniotic fluid (AF), which is one of the main focuses of our workgroup. This nourishing and protecting liquid is in direct contact with the fetus during pregnancy and is therefore one of clinical materials of interest in the research of pregnancy related disorders, including subclinical intraamniotic infection and inflammation (IAI), which may be a cause of several consecutive complications including preterm birth and preterm premature rupture of membranes (PPROM) [20].

In this study, we have designed and optimized a method, which combines the efficiency of cysteinyl peptide enrichment with the robustness of iTRAQ quantitation. In the first phase, individual parameters of CPE have been fine-tuned using radioactive ^{35}S-cysteine-labeled bacterial peptides in order to maximize capturing efficiency. The optimized protocol was consequently applied onto an iTRAQ labeled digest of bovine serum albumin (BSA) to assess the effect of CPE fractionation on iTRAQ quantitation. The final stage employed the latest QqTOF technology for very fast MS/MS acquisition and demonstrated the benefits of CysTRAQ in amniotic fluid proteome analysis.

2. Materials and methods

2.1. ^{35}S-Cysteine metabolic labeling

The *Francisella tularensis* spp. *holarctica* live vaccine strain was grown in 25 ml of chemically defined medium prepared according to Chamberlain [21] (complete medium) at 37 °C overnight. The medium was exchanged to 60 ml of fresh complete medium and cells were cultivated until OD_{600nm} reached 0.85. The cells were washed with cysteine-free medium (incomplete medium), pelleted and resuspended in 100 ml of incomplete medium. Radioactive ^{35}S-cysteine (500 μCi, 10 mCi/ml) (MP Biochemicals, Irvine, CA) was added and the cells were incubated at 37 °C for 24 h. The bacteria were washed twice in PBS, resulting pellets were re-suspended in 3 ml of 0.2% RapiGest (Waters, Milford, MA), lysed in a French pressure cell (Thermo Scientific, Milford, MA) and the lysate was centrifuged to pellet cellular debris. The supernatant was removed, total protein content was determined using BCA Protein Assay (Pierce, Rockford, IL) and the samples were stored at −80 °C until analysis. All manipulation with radioactive material was performed under strictly regulated conditions with particular detail to appropriate protection. The radioactive material was disposed according to guidelines for radioactive waste manipulation and disposal.

2.2. CPE capture optimization

Labeled bacterial proteins were digested and CPE was performed as described below. Radioactivity was measured in the following fractions: unbound non-cysteinyl peptide fraction, eluted cysteinyl peptide fraction and beads after elution. The radioactivity was determined using a Tri-Carb Liquid Scintillation Analyzer (PerkinElmer, Waltham, MA).

2.3. Amniotic fluid sample collection, classification and preparation

The study was approved by the Ethical Committee of the University Hospital in Hradec Kralove on March, 19th 2008 (No. 200804 SO1P). Amniotic fluid samples were supplied by

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

Department of Obstetrics and Gynecology, University Hospital Hradec Kralove. The samples (3–5 ml) were obtained by transabdominal amniocentesis after signing written informed consent. Samples were supplemented with Complete EDTA free protease inhibitors (Roche, Basel, Switzerland) and centrifuged at 300×*g* to remove cells and debris. The supernatants were filtered through a 0.22 μm syringe-driven filter (TPP, Trasadingen, Switzerland) and stored at −80 °C until use.

PPROM was diagnosed by sterile speculum examination confirming pooling of amniotic fluid in the vagina in association with the presence of the insulin-like growth factor binding protein (ACTIM PROM test; Medix Biochemica, Kauniainen, Finland) in the vaginal fluid. Intraamniotic inflammation was determined as the presence of histological chorioamnionitis. Histological examination of the placenta, the fetal membranes and umbilical cord was performed in all cases. The degree of polymorphonuclear leukocyte infiltration was assessed according to criteria given by Salafia et al. [22]. Intraamniotic infection was allocated as a presence of microbial invasion of the amniotic cavity, which was defined as a positive result of PCR analysis for genital mycoplasmas (*Ureaplasma urealyticum*, *Ureaplasma parvum*, *Mycoplasma hominis*) and/or positive cultivation results of any bacteria in the amniotic fluid. Only patients with confirmed intraamniotic infection and inflammation were considered as IAI positive. Patients with ruled out IAI had both parameters negative. Four representative samples were enrolled to the project, with confirmed IAI (n=2) and ruled out IAI (n=2). At this point it should be stressed out that this analysis was conducted only for purposes of showing the feasibility of the optimized protocol when applied on real-life clinical material. The initial protein concentration was determined by BCA and respective samples were pooled 1:1 according to total protein concentration (*w/w*) to create a positive and a negative sample. Both of these were then split into two equal parts to generate duplicates and further processed (Fig. 1).

2.4. *Immunoaffinity depletion of high abundance proteins*

Three milligrams of total protein from each sample was used *per* injection. First, the samples were concentrated using Amicon Ultra centrifugal filters (Millipore, Bedford, MA) with 3 kDa MW cutoff to reach protein concentration roughly equal to human plasma. The top 14 high abundance proteins

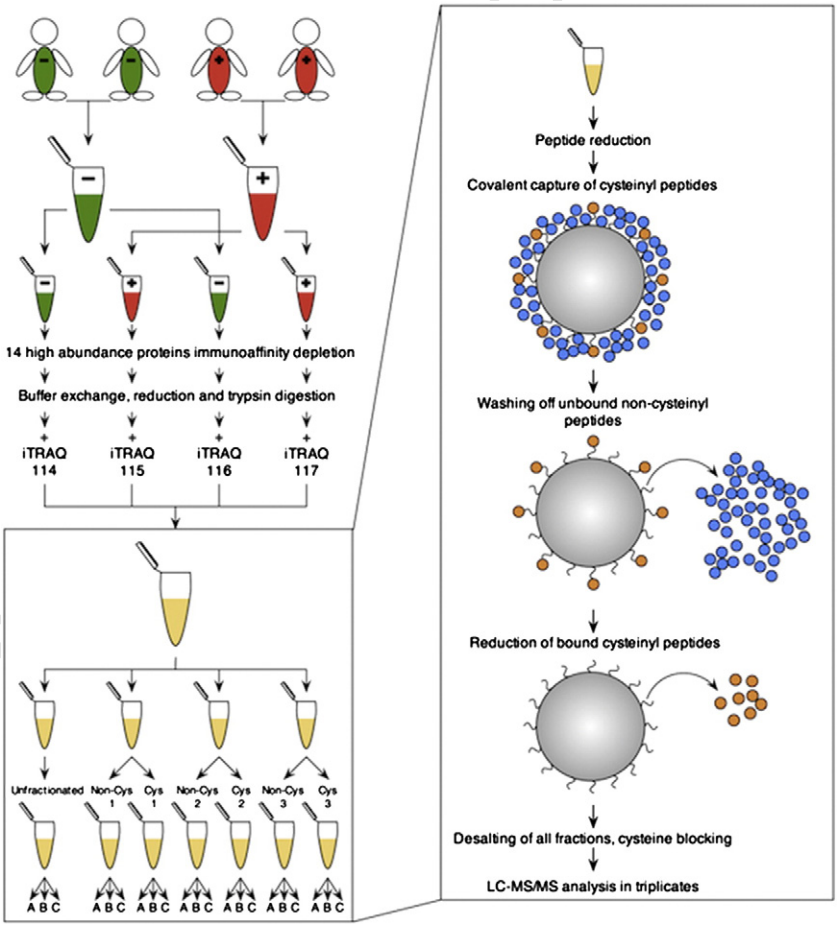


Fig. 1 – Amniotic fluid analysis workflow. Representative IAI positive and IAI negative samples were pooled and further processed as two equal duplicates. Samples were depleted from 14 high abundance proteins and digested. Resulting peptides were iTRAQ labeled, combined, desalted, split into four 50 μg aliquots and dried. Three of these aliquots were fractionated using CPE. All seven final samples were subjected to LC–MS/MS analysis (three replicates *per* sample were injected). The right part of the picture shows the enrichment of cysteinyl peptides in closer detail.

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

which constitute up to 90% of the total protein mass of human amniotic fluid were removed using the MARS Hu-14 immunoaffinity column (Agilent, Palo Alto, CA) on an Alliance 2695 HPLC system (Waters) according to manufacturer's instruction. The collected flow-through fractions containing low abundance proteins were concentrated on a 3 kDa cutoff Amicon Ultra filter (Millipore) and desalted by a prepacked PD MiniTrap G-25 desalting column (GE Healthcare, Uppsala, Sweden) equilibrated with 0.05% RapiGest. The desalted samples were further concentrated in a SpeedVac, and the final protein amount was determined by BCA.

2.5. *Trypsin digestion and iTRAQ labeling*

The proteins dissolved in 0.1% RapiGest were supplemented with 1 M triethylammonium bicarbonate buffer, pH 8.5 (Sigma, St. Louis, MO) to a final concentration of 250 mM, reduced using 5 mM tris(2-carboxyethyl) phosphine hydrochloride (TCEP) from the iTRAQ Multiplex Buffer Kit (AB SCIEX, Foster City, CA) for 1 h at 60 °C and digested overnight at 37 °C by trypsin (Promega, Madison, WI) at a 1:50 (*w*:*w*) trypsin-to-protein ratio. iTRAQ reagents dissolved in ethanol were added to the digests and incubated for 2 h at RT. Upon labeling, 300 μl of 5% TFA was added to hydrolyze both RapiGest as well as unreacted iTRAQ reagents. Samples were then combined and centrifuged to pellet the hydrophobic part of RapiGest. The supernatant was desalted using an Oasis HLB 1 cc (30 mg) Extraction Cartridge SPE column (Waters), and vacuum dried.

2.6. *CPE of iTRAQ labeled peptides*

The CPE workflow (Fig. 1) is based on the protocol described by Liu et al. [5]. All solutions were degassed prior use in order to prevent unwanted oxidation of the thiol content. Thiopropyl-Sepharose 6B (GE Healthcare) thiol affinity resin was prepared according to manufacturer's instructions. Twenty-five milligrams of the powder was rehydrated in 1 ml of water for 1 h and washed with 5 ml of water followed by 5 ml of coupling buffer (50 mM Tris, 1 mM EDTA, pH 7.5) in an empty Macro SpinColumn (Harvard Apparatus, Holliston, MA). Afterwards, the diluted slurry was transferred to a 1.5 ml microtube, allowed to settle down and the supernatant was carefully removed. Dried peptides (50 μg) were dissolved in 20 μl of freshly prepared 5 mM DTT in coupling buffer and reduced for 1 h at 60 °C. The sample was diluted to 100 μl with coupling buffer, added to the slurry and incubated for 2 h at 37 °C in an axial rotation arrangement. This ensured effective mixing, while minimizing air contact and thus unwanted oxidation. The slurry was then transferred to the Macro SpinColumn and unbound peptide fraction (non-cysteinyl peptides) was captured, followed by washing the beads twice with 100 μl of coupling buffer. These were also collected and combined with the unbound fraction. The beads were then extensively washed with 2.5 ml of each of the following solutions: washing buffer (50 mM Tris, 1 mM EDTA, pH 8.0); 2 M NaCl; 80% ACN, 0.1% TFA; and washing buffer. After transferring the diluted slurry into a new microtube and removing the supernatant as described above, 100 μl of freshly prepared 50 mM DTT in washing buffer was added and the captured peptides were released for 1 h at 60 °C with mixing. Finally, the slurry was transferred to the Macro SpinColumn and the bound peptide fraction (cysteinyl peptides) was captured, followed by washing the beads twice with 100 μl of 80% ACN, 0.1% TFA. These were also collected and combined with the bound fraction. Both fractions were desalted on Oasis HLB 1 cc (10 mg) Extraction Cartridge SPE columns and dried. The peptides were reduced with 5 mM TCEP for 1 h at 60 °C, cysteines were blocked with 10 mM methyl methanethiosulfonate (MMTS) from the iTRAQ kit for 10 min at RT and the samples were frozen at −20 °C until analysis.

2.7. *LC–MS/MS analysis*

BSA-based pilot experiments were performed using LC–MALDI-MS/MS approach (4800 MALDI-TOF/TOF Analyzer (AB SCIEX)) as described previously [23]. Amniotic fluid nanoLC peptide separation was performed on an Eksigent nanoLC Ultra and cHiPLC-nanoflex (Eksigent, Dublin, CA) in Trap Elute configuration. From each fraction, 1 μg was desalted on a 200 μm x 0.5 mm precolumn and eluted on an analytical 75 μm × 150 mm column. Both precolumn and analytical column were filled with ChromXP C_{18}-CL 3 μm 120 Å phase (Eksigent). Peptides were separated by a linear gradient formed by 2% ACN, 0.1% FA (mobile phase A) and 98% ACN, 0.1% FA (mobile phase B), from 12 to 32% of mobile phase B in 90 min at a flow rate of 300 nL/min. The MS analysis was performed on a TripleTOF 5600 system (AB SCIEX) in Information Dependent Mode. MS spectra were acquired across the mass range of 400–1500 *m*/*z* in high resolution mode (>30 000) using 250 ms accumulation time *per* spectrum. A maximum of 15 precursors *per* cycle were chosen for fragmentation from each MS spectrum with 100 ms minimum accumulation time for each precursor and dynamic exclusion for 20 s. Tandem mass spectra were recorded in high sensitivity mode (resolution >15 000) with rolling collision energy on and iTRAQ reagent collision energy adjustment on.

2.8. *Data analysis*

Peptide identification and quantification was conducted using the ProteinPilot 4.0 software (AB SCIEX) using the Paragon database search algorithm [24] and the integrated false discovery rate (FDR) analysis function [25]. The software used only unique peptide sequences as evidence for protein identification. The data were searched against UniProtKB/Swiss-Prot database (canonical and isoforms included, downloaded June 22, 2010, AB SCIEX 2007 contaminant DB was appended) containing 35 115 proteins entries. The samples were described using the following parameters in the Paragon method: Sample Type — iTRAQ 4plex (Peptide Labeled); Cys Alkylation — MMTS; Digestion — trypsin; Special Factors — no selection; Species — *Homo sapiens*. The processing was specified as follows: Quantitate — On; Bias Correction — On; Background Correction — On; ID Focus — Biological Modifications; Search Effort — Thorough; Detected Protein Threshold — 0.05 (10.0%). For FDR determination, data were searched against concatenated databases by *in silico* on-the-fly reversal for decoy sequences automatically by the

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

software. Only proteins at 1% global FDR and distinct peptides at 5% local FDR were used for further analysis of the amniotic fluid data. No FDR analysis was performed on the BSA dataset and instead, peptides identified with confidence≥95% were taken into account. For quantification, the ProteinPilot software excluded peptides with confidence<15%, peptides without iTRAQ modification, and spectra shared between different proteins — *i.e.* where a spectrum is also claimed by a different protein but with the same, similar, or unrelated peptide sequence having reasonable confidence as well. All the remaining peptides contributed to protein quantification in the ProteinPilot software. Intensities of iTRAQ reporter ions were corrected using isotope correction factors supplied with the iTRAQ kit. Automatic normalization of quantitative data (bias correction) was further performed for each iTRAQ pair to correct any experimental or systematic bias. The background correction function in ProteinPilot software was also used. This function uses an algorithm to estimate and subtract out the background contribution of many low level coeluting peptides in complex mixtures, which tends to attenuate the extremity of ratios, making them less accurate. The ProteinPilot Descriptive Statistics Template (PDST) (beta v3.001p, AB SCIEX) was used for a number of analyses, including estimation of FDR associated with quantitation results. The tool uses a target–decoy approach, where ratios of analytical replicates function as decoy ratios, to estimate global FDRs associated with single test *p*-value cutoff to provide multiple testing correction, giving the yield of differential detections at a global FDR level. The Protein Alignment Template (beta v1.995p, AB SCIEX) was used for alignment of multiple results at the protein level.

3. Results

3.1. *CPE capture optimization*

In order to optimize individual parameters of the CPE workflow, we modified the peptide capturing protocol using ^{35}S liquid scintillation counting. Unfortunately, there was not any well-established ^{35}S cysteine protein mixture standard commercially available and therefore we prepared metabolically labeled proteins in a live vaccine strain of *Francisella tularensis* bacterium which were subsequently used to generate radioactive peptides. Employment of ^{35}S labeled cysteinyl peptides enabled us to track cysteine-containing peptides in any kind of solution as well as on solid matrix and thus fine tune individual steps in the CPE workflow. The parameters which were optimized using these peptides included temperature and incubation times for individual steps of the process as well as sample handling, incubation arrangement (which ensured good mixing while minimizing unwanted cysteine oxidation) and finally the reduction conditions (various reducing agents and concentrations) for peptide elution from the beads. By tuning initial reduction and binding conditions we reached 72.8% (±0.8) of the radioactivity in the eluted cysteinyl peptide fraction in three independent replicates. After elution, 4.9% (±1.0) of the radioactivity remained on the beads. The remaining radioactivity was detected in unbound non-cysteinyl peptide fraction. Additional optimizing experiments did not lead to increased efficiency of the capturing and therefore we considered this protocol as definitive. The final optimized parameters, which provided maximum capturing efficiency and recovery, are described in the experimental section. To our knowledge, this is the first characterization of cysteine capturing efficiency among the cysteine enrichment techniques. We managed to capture ~78% of the total radioactivity, from which roughly 5% remained bound to the beads, showing excellent peptide recovery (93.6%). This is a key improvement compared to ICAT, as efficient elution of the labeled peptides from the streptavidin column was one of the drawbacks of this approach [26]. The reason why we did not enrich the entire radioactivity remains unclear. We experimentally excluded the ability of the bacterium to convert cysteine to methionine and thus we hypothesize that other sulfur substances without thiol group are responsible for the residual radioactivity in the unbound fraction, *e.g.* peptides derived from cysteine acylated lipoproteins [27]. Our assumption is supported by the LC–MS/MS results, which show almost total selectivity of the Thiopropyl-Sepharose 6B beads towards cysteinyl peptides.

3.2. *Effect of CPE on iTRAQ quantitation — BSA digest*

In the next step the optimized CPE protocol was applied onto an iTRAQ labeled BSA digest to test our assumption that the method would not interfere with iTRAQ quantitation. We labeled two identical BSA digest samples (100 μg *per* sample) using iTRAQ 114 and 117 tags and subsequently created samples containing iTRAQ 117:114 label in ratios close to 3:1, 1:1 and 1:2 to remain unbiased for further evaluation. These samples were subjected to CPE fractionation. Unfractionated sample, non-cysteinyl peptide fractions and cysteinyl peptide fractions were analyzed using LC–MALDI MS/MS approach. The quantitation is not influenced by the fractionation step, as the iTRAQ ratios and CVs are consistent across all three fractions in all samples (Fig. 2).

3.3. *Application of the CysTRAQ method on amniotic fluid sample*

The final step involved application of the CysTRAQ workflow on a real-life complex sample to evaluate the overall method performance. We thus employed this technique to analyze human amniotic fluid proteome with regard to intraamniotic infection/inflammation (IAI). To use the full potential of our method, we analyzed both amniotic fluid samples with confirmed IAI and samples from patients in which IAI diagnosis was ruled out. We depleted the 14 most abundant proteins using immunoaffinity chromatography and the resulting protein fraction was digested and iTRAQ reagent labeled. The samples were then combined, desalted and split into four aliquots. One was used as an unfractionated reference sample, whereas the remaining three were subjected to the fractionation protocol. By labeling two sets of duplicates with four iTRAQ tags, we were able to assess the effect of fractionation on iTRAQ quantitation as well as to determine the effectiveness of the method for peptide and protein identification in a complex sample. In addition, the introduction of both positive and negative clinical samples enabled us to get a basic

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

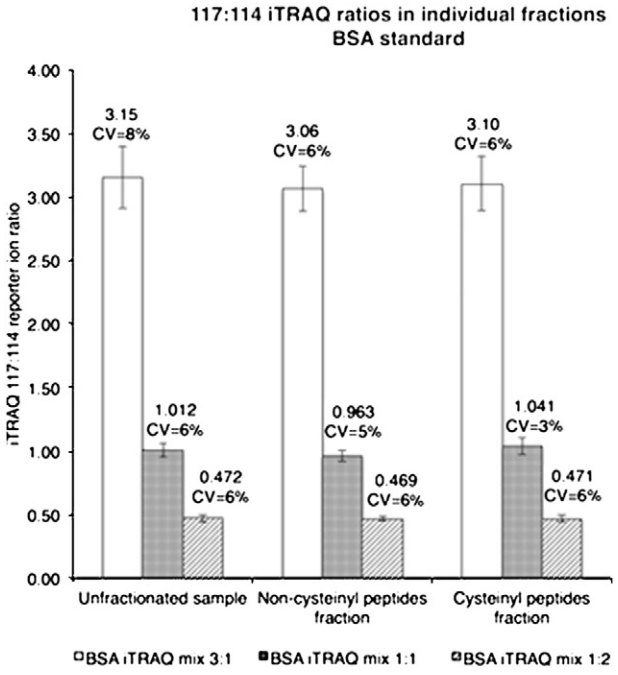


Fig. 2 – iTRAQ ratios in individual fractions of BSA samples mixed in three ratios: 3:1, 1:1 and 1:2. To remain unbiased, we mixed the ratios just roughly close to mentioned. The iTRAQ reporter ratios are constant across all three fractions, suggesting that the iTRAQ quantification is not influenced by the fractionation step (data are shown as mean ± SD).

overview of quantitative protein changes in respective groups of samples.

3.4. *Qualitative results*

All seven resulting samples were analyzed using LC–MS/MS with three injections per sample (Fig. 1). Using the high speed MS/MS capabilities of the TripleTOF 5600 system, good depth of coverage was obtained in all of the three sample types analyzed. The evaluation of three injections of the unfractionated sample yielded 544 proteins (1% global FDR) and 11 071 distinct peptides (5% local FDR). By evaluating three injections of both non-cysteinyl peptides fraction and cysteinyl peptides fraction, we identified on average 679 proteins (9 619 peptides) and 486 proteins (7 898 peptides), respectively (Fig. 3A). In our settings we clearly showed that the highest number of identified proteins was achieved in fraction depleted of cysteinyl peptides. Interestingly, the number of identified peptides was lower by 13.1% in the non-cysteinyl fraction than in unfractionated sample, but the number of proteins was higher by 24.8%. Even though both less peptides (by 28.6%) and proteins (by 10.6%) were identified in cysteinyl peptide fraction, the advantage brought by the analysis of this fractions results in 20% increase of unique proteins compared to the unfractionated sample as shown in Fig. 3E. More importantly, these could be of low abundance, as it was clearly shown in a concentration annotated yeast proteome study [15]. One major conclusion may be made from these observations. Providing that the instrument time is limited and only one injection may be performed it is worthy to inject peptides depleted of cysteinyl peptides. On the other hand, if the goal is to identify as many proteins as possible, both fractions can be readily analyzed.

The potential benefit of CPE was further assessed in regard to unique proteins identification. To remain unbiased we compared the same number of injections. We therefore compared results obtained by two injections of the same unfractionated sample against one injection of each non-cysteinyl and cysteinyl fractions. The Venn diagrams illustrate the benefits of the fractionation on uniquely identified proteins (Fig. 3B–F). One injection of each non-cysteinyl and cysteinyl fractions detected 38% more unique proteins than two injections of the unfractionated sample (Fig. 3B). This comparison revealed that CPE led to a substantial gain of unique proteins and thus leading to increased proteome coverage. By comparing three injections of the unfractionated sample *vs.* three injections of either the cysteinyl or non-cysteinyl peptides fraction, the number of identified proteins increased by 20.9% and 40.2%, respectively (Fig. 3E and F). Ultimately, the combined evaluation of three injections of both cysteinyl and non-cysteinyl peptides fractions resulted in identification of 60.8% more proteins compared to three injections of the unfractionated sample, leaving only 29 unique proteins in the unfractionated sample analysis (Fig. 3C). The contribution of increased unique protein identification rate was also accompanied by increased sequence coverage (Fig. 4). The combination of non-cysteinyl and cysteinyl fractions brings along higher sequence coverage than the analysis of unfractionated sample alone.

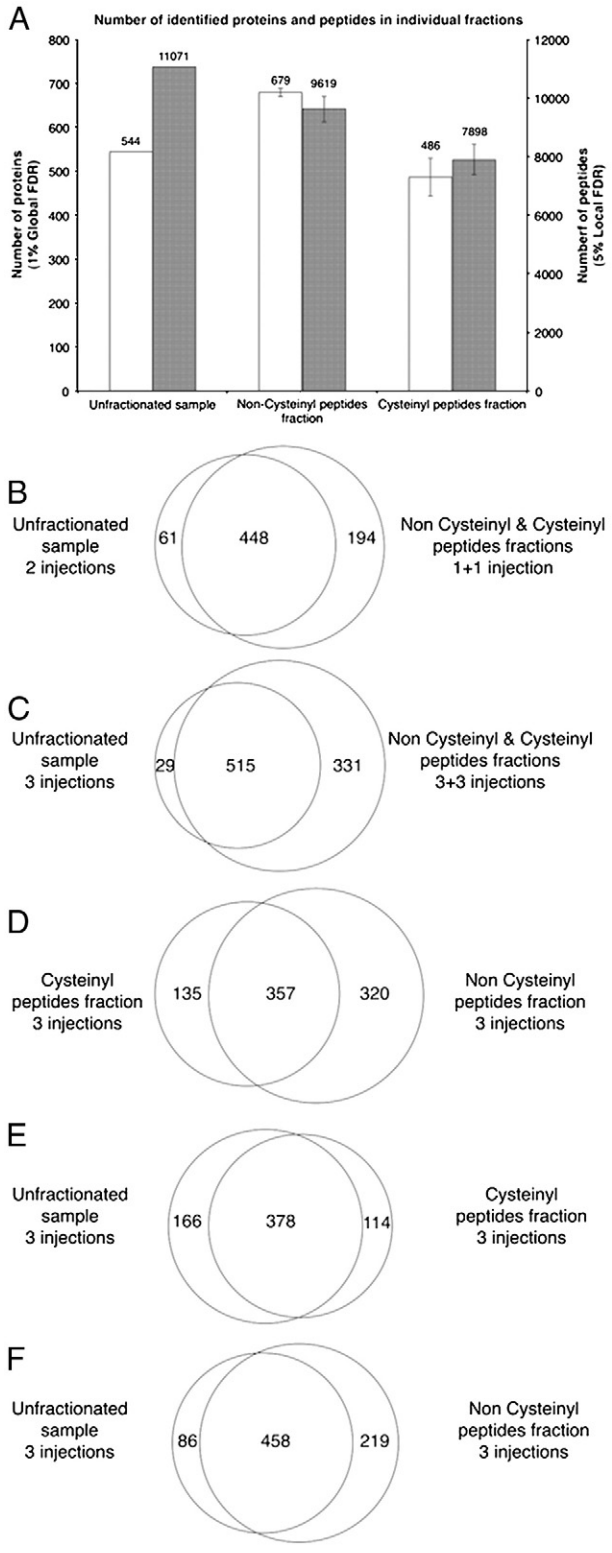
The inspection of peptides in individual fractions revealed outstanding performance of the method with respect to specificity of the cysteinyl peptide capture step. Although we were not able to bind entire loaded radioactivity on the solid support, the assessment of identified peptide sequences revealed only very low contamination with cysteinyl peptides in the non-cysteinyl fraction. On average, only 3.19% of peptides from the non-cysteinyl fraction did contain cysteine. This supports our assumption that sulfur containing compounds without thiol groups are responsible for the remaining radioactivity, which does not bind to the Thiopropyl-Sepharose 6B beads. On the other hand, 98.6% of peptides in the cysteinyl fraction did indeed contain cysteine in the sequence proving the enrichment specificity towards cysteinyl peptides. The unfractionated sample contained 37.1% of cysteinyl peptides.

3.5. *Effect of fractionation on iTRAQ quantitation*

The results from the BSA pilot experiment indicated that the fractionation step had no effect on iTRAQ quantitation. This fact was further verified on a large data set obtained by analysis of real-life protein mixture. Ratios calculated for each protein were transferred to logarithmic scale to perform parametric Pearson test. The correlation coefficient (r) was 0.837 with $p<0.001$ for the comparison of quantification in the unfractionated sample with a combination of cysteinyl and non-cysteinyl peptides (Fig. 5A). Similarly, the comparison of cysteinyl peptides quantitative results with non-cysteinyl peptides showed $r=0.795$ ($p<0.001$) (Fig. 5B). We thus conclude that the quantification is not influenced by

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

the fractionation step and is consistent across all fractions, as supported by both BSA pilot data as well as by correlation tests from large amniotic fluid datasets. We also tested the inter-experiment reproducibility by comparing either two non-cysteinyl as well as two cysteinyl fractions. Both correlation coefficients ($r=0.907$ and 0.884, respectively) as well as p-values ($p<0.001$ for both) demonstrated excellent workflow reproducibility (data not shown).



3.6. *Differentially expressed proteins: PPROM cases with confirmed IAI vs. PPROM patients with ruled out IAI*

To demonstrate the capabilities of the CysTRAQ method in terms of detecting dysregulated proteins, we compared IAI negative iTRAQ quantitation data with the IAI positive group. MS/MS spectra acquired had high resolution and sensitivity in the low mass region, providing high quality quantitation on the iTRAQ reagent reporter ions. Solid reporter ion quantitation combined with a deep interrogation of the sample allowed for many proteins to be quantified with good confidence. The analyses used for differential protein analysis were based on analysis of three technical replicates, where one evaluation consisted of the following (see Fig. 1 for details): Non-Cysteinyl fraction, 3 injections (A, B, C) and Cysteinyl fraction, 3 injections (A, B, C). All three prepared replicates were analyzed in this manner and resulted in three evaluations. Only proteins with $p<0.05$ for both 115:114 and 117:116 ratios were considered to be significantly dysregulated. However, in order to be reported in the final summary, a protein had to be significantly altered in all three evaluations. For individual evaluations of the data searches, we detected 128, 128 and 134 significantly altered proteins. An example from one of these evaluations is shown in a volcano plot in Fig. 6A, which clearly shows the relationship between extremity of change and certainty of change. Fig. 6B then shows a graphical representation of the decoy hits, used for quantitative FDR determination. For this, the 116 channel was set as decoy. Based on this analysis, all the dysregulated proteins reported in this study are with quantitative FDR lower than 5%. In total, 83 common proteins were dysregulated in all three evaluations and are reported in Table 1, Supplementary information.

4. Discussions

Despite the wide range of quantitative proteomic methods available, common drawbacks of methods based on MS mode quantitation remain i) the increment in spectral complexity and ii) the inherently lower number of sample channels available. The ICAT approach specifically addresses the increased spectral complexity issue by targeted tagging of cys-

Fig. 3 – Comparison of identified proteins/peptides count in individual fractions (panel A). Three injections of unfractionated sample were searched and evaluated as one data set and were then compared with the average yield of identified proteins/peptides obtained from three injections of the three non-cysteinyl peptide fractions, and three injections of the three cysteinyl-peptide fractions (data are shown as mean ± SD). Venn diagrams illustrating yields of identified proteins in distinct combinations (panels B–F). We present median number of identification results. The protein distribution and overlap across multiple fractions was assessed using the Protein Alignment Template (beta v1.995p, AB SCIEX).

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

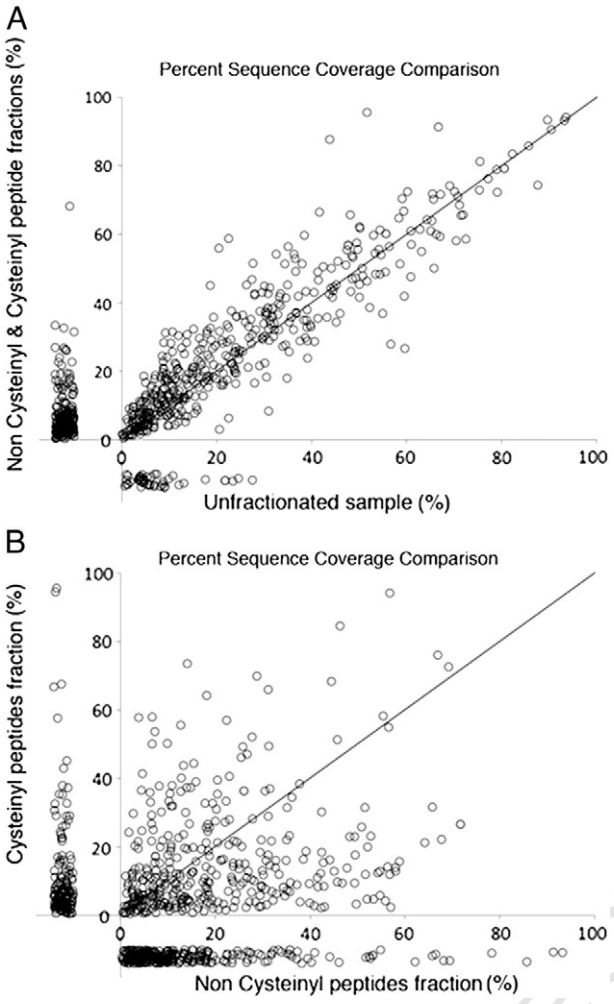


Fig. 4 – Comparison of sequence coverage for individual proteins between different fractions. Sequence coverage comparison for unfractionated sample and a combination of non-cysteinyl and cysteinyl fractions (panel A) and for non-cysteinyl and cysteinyl fractions (panel B). Each dot is a protein, with proteins unique to a particular fraction shown along the respective axes. The analysis was performed using the Protein Alignment Template (beta v1.995p, AB SCIEX).

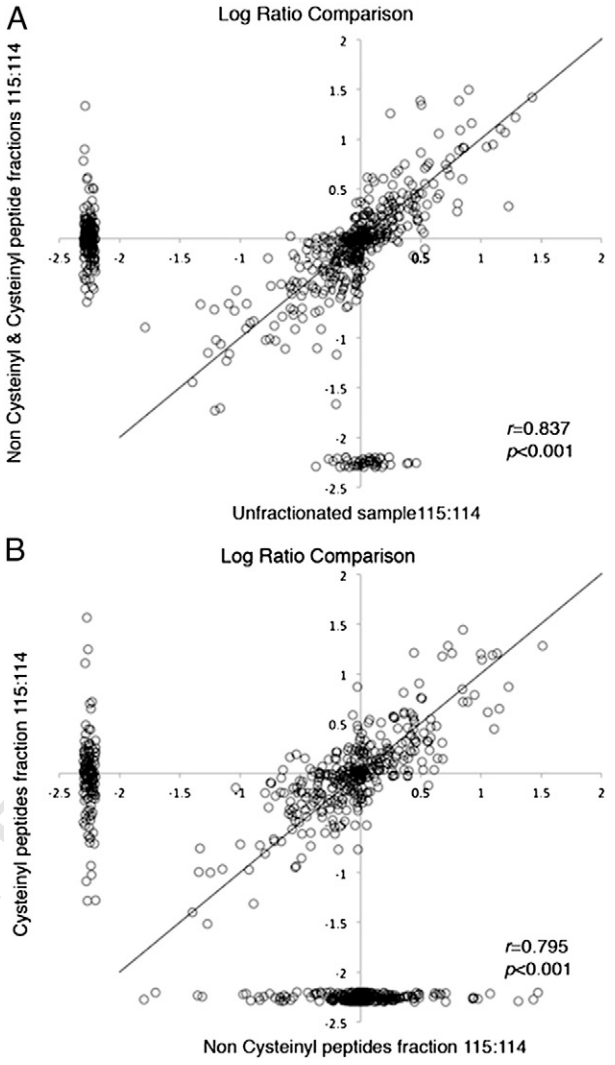


Fig. 5 – Two sample quantification comparison. Log ratio comparison of unfractionated sample (2 injections) and a combination of non-cysteinyl and cysteinyl fractions (1+1 injections) (panel A) as well as of non-cysteinyl (3 injections) and cysteinyl peptides fractions (3 injections) (panel B) show quantification reproducibility across sample preparation. Each protein ratio is shown as a dot presented in logarithmic scale, with proteins unique to the respective fraction shown along axes. Both correlation factor (r) and *p*-values are shown. The comparison was performed using the Protein Alignment Template (beta v1.995p, AB SCIEX).

teines within proteins, allowing substantial complexity reduction in downstream analysis. MS/MS mode based quantitation techniques, on the other hand, do not increase MS spectral complexity and provide a higher level of possible sample multiplexing: the TMT method provides up to six channels [28], and the latest version of the iTRAQ reagents has eight quantification channels available [29]. Being inspired by the benefits of both ICAT and iTRAQ, we attempted to combine the advantages of these methods into one method — CysTRAQ.

Another cysteine targeting approach, CPE has been previously reported to enable increased proteome coverage by fractionating the peptides in the sample according to presence or absence of cysteine [5]. In our work, we extended the technique towards quantitative proteomics by combining iTRAQ quantitation with the highly specific CPE. This specificity is enabled by covalent binding of the cysteinyl peptides to the solid support. Being truly unique among cysteine targeting approaches, the covalently bound peptides can withhold much more stringent washing out of non-specifically bound peptides. The exceptional specificity of 98.6% towards cysteinyl peptides is supported by evaluation of a dataset containing nearly 8000 unique peptides and is unmatched by any of the previous works [5,15]. In fact, the presence of cysteine in the sequence could be an additional criterion for confident peptide identification and thus for FDR rate reduction. Besides the advantage with regards to washing, the formation of disulfide bond is a simple oxidation reaction. The elution of

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

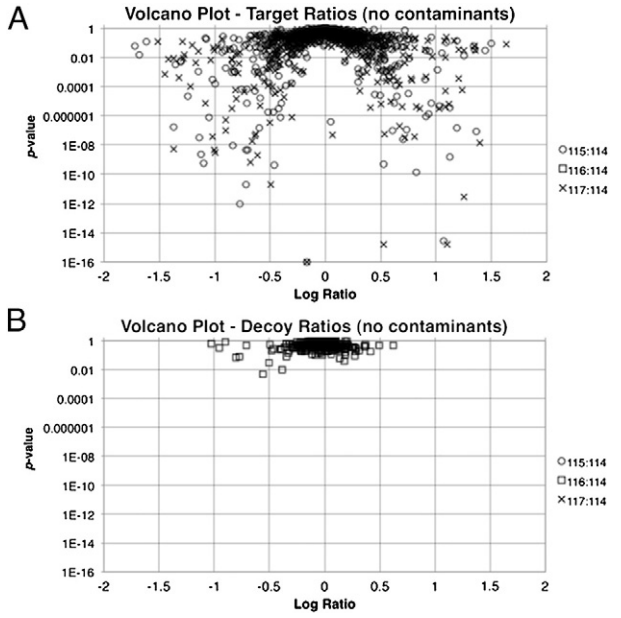


Fig. 6 – Quantitative FDR assessment. A representative example of graphical demonstration of the relationship between the extremity of change and certainty of change (panel A). Panel B shows decoy hits used for quantitative FDR calculation. Based on these data, the $p<0.05$ threshold in turn corresponds to quantitative FDR<5%. The quantitative FDR analysis was performed using the PDST tool (beta v3.001p, AB SCIEX).

the peptides of interest from the beads is therefore based on plain reduction. We support this by our radioactive peptides-based experiments, where 93.6% of the total radioactivity was recovered from the resin and provides a key advantage compared to ICAT, as efficient elution of the labeled peptides from the streptavidin column was one of the drawbacks of this approach [26]. By using the natural ability of cysteine to form disulfide bonds, we do not use any kind of tagging to mediate the capturing of cysteinyl peptides. This in turn enables us to omit any potential mass increment associated issues during MS analysis, ionization efficiency modification or shift in chromatography and simplify the whole workflow. Moreover, due to the fact that the quantitative information is not incorporated in a cysteine targeting tag, the quantitation is not limited to the cysteinyl peptides fraction, providing additional versatility. As mentioned above, the quantitative information is brought by the iTRAQ reagents, not the cysteine tag, therefore bringing the high multiplex advantages of iTRAQ MS/MS quantitation to the workflow. Also, the quantitation is not limited to the cysteinyl peptide fraction because of the iTRAQ amine labeling strategy. We herein show, that the non-cysteinyl peptide group in fact contributed with the highest number of uniquely identified proteins, making it a perfect complementary for the cysteinyl peptides analysis. In addition, iTRAQ 4plex tagged peptides were shown to provide more abundant and complete *b*- and *y*-fragment ion series and the *N*-methylpiperazine group of the iTRAQ tag seems to promote ionization of lysine-containing peptides [30], which is another advantage brought by the iTRAQ labeling. The CysTRAQ method benefits from this, providing overall better spectral quality in all three fractions.

Although the application of CysTRAQ to the real-life samples analysis should be considered as proof-a-concept of our method, our analysis pointed out on several interesting changes in AF proteins. Based on these findings, we would like to point out on three interesting phenomena that are very likely to occur in the IAI positive pregnancies, as demonstrated by our results.

We found histone H4 levels to be considerably higher in the positive patient group. Histone proteins are core components of nucleosomes and are involved in chromosomal stability, DNA replication, DNA repair, and transcription regulation. We suggest that the reason why this protein was detected in higher abundance in IAI positive samples is due to the ability of neutrophils to form extracellular fibril matrices, known as neutrophil extracellular traps (NETs). Upon activation, neutrophils actively release fibers with chromatin and granule proteins as two major components forming these trapping structures [31]. NETs bind microorganisms and thus prevent them from spreading. In addition, high local concentration of antimicrobial agents, which are incorporated into these webs, directly kills pathogens. Our hypothesis is supported by finding increased levels of these antimicrobial proteins as well as of additional proteins, which were earlier identified as integral components of NET, in the IAI positive samples (neutrophil defensin 1, myosin-9, alpha-enolase, plastin-2, protein S100-A8, protein S100-A9, transketolase) [32]. Some of these proteins were repeatedly reported in earlier studies aimed at discovering biomarkers of IAI, but without the association to NETs [33–36]. Interestingly, histone H4 has been recently reported to have direct antimicrobial effects [37].

Soluble hemoglobin scavenger receptor (sCD163) was found among proteins with considerably higher level in IAI positive samples. CD163 expression occurs mainly in late inflammatory response phase and seems to be restricted to the monocyte/macrophage lineage [38]. Besides its primary function, which is endocytosis of haptoglobin–hemoglobin complexes, it is also involved in recognition of intact bacteria [39,40]. The extramembranous part of the membrane-anchored CD163 molecule is released upon activation of the toll-like receptor signaling or in response to stimulation by other proinflammatory factors [41,42]. In our previous hypothesis-driven work we showed that PPROM women with IAI had higher median amniotic fluid sCD163 concentration than those without IAI (PPROM with IAI: 885 ng/ml, IQR: 295–1779, n=44 *vs.* PPROM without IAI 288 ng/ml, IQR: 170–499, n=45; $p<0.0001$) [43]. We believe that this work retrospectively validates results obtained in the present study and thus proved that CysTRAQ method is able to provide relevant information on proteins involved in IAI process.

One of the most down-regulated proteins found in our dataset was the chorionic somatomammotropin hormone (human placental lactogen, hPL). It is produced and secreted by the syncytiotrophoblast during pregnancy and is involved in stimulating lactation, fetal growth and metabolism. Its maternal serum levels rise in relation to the growth of the fetus and placenta [44]. Therefore, low maternal levels of human placental lactogen have been proposed as a sign of placental insufficiency, associated with intrauterine placental and

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027

fetal growth restriction [45]. Interestingly, intrauterine inflammation may be one of the causative factors responsible for these growth restrictions [46]. Furthermore, Okada et al. suggested a link between hPL and intraamniotic inflammation in a study, which found that preterm placentas affected by chorioamnionitis produced smaller amounts of both hPL and human chorionic gonadotropin [47].

5. Conclusions

Hereby, we for the first time demonstrated a successful combination of the advantages of both ICAT and iTRAQ into one method, which allows specific enrichment of cysteine-containing, iTRAQ labeled peptides. For this reason, we call this approach CysTRAQ. Based on our results, the CysTRAQ method promises several key advantages. The ability to quantify four samples in one analysis (application of the latest version of iTRAQ would enable analysis of eight samples) is unmatched compared to other cysteine targeting quantification approaches and the quantification is performed in MS/MS mode and thus the complexity of MS spectra is not increased. Furthermore, due to the fact that the quantification feature of the method is not linked to any cysteine-targeting tag, the CysTRAQ approach enables flawless analysis of both cysteinyl and non-cysteinyl peptide fractions if desired, providing even greater proteome coverage. Finally, we successfully applied the technique into amniotic fluid analysis to detect differences in protein levels caused by the presence of intraamniotic infection and inflammation.

Supplementary materials related to this article can be found online at doi:10.1016/j.jprot.2011.09.027.

Conflict of interest

There are no conflicts of interest to report.

Acknowledgment

The authors greatly acknowledge financial support from the Ministry of Health, Czech Republic (Grant No. NS/10382-3/2009) and from Czech Science Foundation GACR (Grant No. 301/09/P241). Also, the authors thank Marek Link for excellent technical assistance.

REFERENCES

[1] Tambor V, Fucikova A, Lenco J, Kacerovsky M, Rehacek V, Stulik J, et al. Application of proteomics in biomarker discovery: a primer for the clinician. Physiol Res 2009.
[2] Picotti P, Aebersold R, Domon B. The implications of proteolytic background for shotgun proteomics. Mol Cell Proteomics 2007;6:1589–98.
[3] Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 2002;1:845–67.
[4] Liu T, Qian WJ, Camp 2nd DG, Smith RD. The use of a quantitative cysteinyl-peptide enrichment technology for high-throughput quantitative proteomics. Methods Mol Biol 2007;359:107–24.
[5] Liu T, Qian WJ, Chen WN, Jacobs JM, Moore RJ, Anderson DJ, et al. Improved proteome coverage by using high efficiency cysteinyl peptide enrichment: the human mammary epithelial cell proteome. Proteomics 2005;5:1263–73.
[6] Zhou Y, Aebersold R, Zhang H. Isolation of N-linked glycopeptides from plasma. Anal Chem 2007;79:5826–37.
[7] Zhang L, Xu Y, Yao H, Xie L, Yao J, Lu H, et al. Boronic acid functionalized core-satellite composite nanoparticles for advanced enrichment of glycopeptides and glycoproteins. Chemistry 2009;15:10158–66.
[8] McDonald L, Beynon RJ. Positional proteomics: preparation of amino-terminal peptides as a strategy for proteome simplification and characterization. Nat Protoc 2006;1:1790–8.
[9] Sonomura K, Kuyama H, Matsuo E, Tsunasawa S, Nishimura O. The specific isolation of C-terminal peptides of proteins through a transamination reaction and its advantage for introducing functional groups into the peptide. Rapid Commun Mass Spectrom 2009;23:611–8.
[10] Giron P, Dayon L, David F, Sanchez JC, Rose K. Enrichment of N-terminal cysteinyl-peptides by covalent capture. J Proteomics 2009;71:647–61.
[11] Foettinger A, Leitner A, Lindner W. Selective enrichment of tryptophan-containing peptides from protein digests employing a reversible derivatization with malondialdehyde and solid-phase capture on hydrazide beads. J Proteome Res 2007;6:3827–34.
[12] Gevaert K, Van Damme J, Goethals M, Thomas GR, Hoorelbeke B, Demol H, et al. Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis: identification of more than 800 *Escherichia coli* proteins. Mol Cell Proteomics 2002;1:896–903.
[13] Betancourt L, Gil J, Besada V, Gonzalez LJ, Fernandez-de-Cossio J, Garcia L, et al. SCAPE: a new tool for the Selective CApture of PEptides in protein identification. J Proteome Res 2005;4:491–6.
[14] Giron P, Dayon L, Sanchez JC. Cysteine tagging for MS-based proteomics. Mass Spectrom Rev 2010.
[15] Lin D, Li J, Slebos RJ, Liebler DC. Cysteinyl peptide capture for shotgun proteomics: global assessment of chemoselective fractionation. J Proteome Res 2010;9:5461–72.
[16] Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 1999;17:994–9.
[17] Spahr CS, Susin SA, Bures EJ, Robinson JH, Davis MT, McGinley MD, et al. Simplification of complex peptide mixtures for proteomic analysis: reversible biotinylation of cysteinyl peptides. Electrophoresis 2000;21:1635–50.
[18] Giron P, Dayon L, Mihala N, Sanchez JC, Rose K. Cysteine-reactive covalent capture tags for enrichment of cysteine-containing peptides. Rapid Commun Mass Spectrom 2009;23:3377–86.
[19] ThermoScientific. Cysteine-reactive (CysTMT) mass tag reagents; 2011.
[20] Romero R, Espinoza J, Goncalves LF, Kusanovic JP, Friel LA, Nien JK. Inflammation in preterm and term labour and delivery. Semin Fetal Neonatal Med 2006;11:317–26.
[21] Chamberlain RE. Evaluation of live tularemia vaccine prepared in a chemically defined medium. Appl Microbiol 1965;13:232–5.
[22] Salafia CM, Weigl C, Silberman L. The prevalence and distribution of acute placental inflammation in uncomplicated term pregnancies. Obstet Gynecol 1989;73:383–9.

[23] Lenco J, Link M, Tambor V, Zakova J, Cerveny L, Stulik AJ. iTRAQ quantitative analysis of *Francisella tularensis* ssp. *holarctica* live vaccine strain and *Francisella tularensis* ssp. tularensis SCHU S4 response to different temperatures and stationary phases of growth. Proteomics 2009;9: 2875–82.

[24] Shilov IV, Seymour SL, Patel AA, Loboda A, Tang WH, Keating SP, et al. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Mol Cell Proteomics 2007;6:1638–55.

[25] Tang WH, Shilov IV, Seymour SL. Nonlinear fitting method for determining local false discovery rates from decoy database searches. J Proteome Res 2008;7:3661–7.

[26] Gygi SP, Rist B, Griffin TJ, Eng J, Aebersold R. Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. J Proteome Res 2002;1:47–54.

[27] Kovacs-Simon A, Titball RW, Michell SL. Lipoproteins of bacterial pathogens. Infect Immun 2011;79:548–61.

[28] Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 2003;75:1895–904.

[29] Pierce A, Unwin RD, Evans CA, Griffiths S, Carney L, Zhang L, et al. Eight-channel iTRAQ enables comparison of the activity of 6 leukaemogenic tyrosine kinases. Mol Cell Proteomics 2007.

[30] Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 2004;3:1154–69.

[31] Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, et al. Neutrophil extracellular traps kill bacteria. Science 2004;303:1532–5.

[32] Urban CF, Ermert D, Schmid M, Abu-Abed U, Goosmann C, Nacken W, et al. Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against *Candida albicans*. PLoS Pathog 2009;5: e1000639.

[33] Romero R, Kusanovic JP, Gotsch F, Erez O, Vaisbuch E, Mazaki-Tovi S, et al. Isobaric labeling and tandem mass spectrometry: a novel approach for profiling and quantifying proteins differentially expressed in amniotic fluid in preterm labor with and without intra-amniotic infection/inflammation. J Matern Fetal Neonatal Med 2010;23: 261–80.

[34] Buhimschi IA, Christner R, Buhimschi CS. Proteomic biomarker analysis of amniotic fluid for identification of intra-amniotic inflammation. BJOG 2005;112:173–81.

[35] Ruetschi U, Rosen A, Karlsson G, Zetterberg H, Rymo L, Hagberg H, et al. Proteomic analysis using protein chips to detect biomarkers in cervical and amniotic fluid in women with intra-amniotic inflammation. J Proteome Res 2005;4: 2236–42.

[36] Gravett MG, Thomas A, Schneider KA, Reddy AP, Dasari S, Jacob T, et al. Proteomic analysis of cervical–vaginal fluid: identification of novel biomarkers for detection of intra-amniotic infection. J Proteome Res 2007;6:89–96.

[37] Lee DY, Huang CM, Nakatsuji T, Thiboutot D, Kang SA, Monestier M, et al. Histone H4 is a major component of the antimicrobial action of human sebocytes. J Invest Dermatol 2009;129:2489–96.

[38] Buechler C, Ritter M, Orso E, Langmann T, Klucken J, Schmitz G. Regulation of scavenger receptor CD163 expression in human monocytes and macrophages by pro- and antiinflammatory stimuli. J Leukoc Biol 2000;67:97–103.

[39] Fabriek BO, van Bruggen R, Deng DM, Ligtenberg AJ, Nazmi K, Schornagel K, et al. The macrophage scavenger receptor CD163 functions as an innate immune sensor for bacteria. Blood 2009;113:887–92.

[40] Kristiansen M, Graversen JH, Jacobsen C, Sonne O, Hoffman HJ, Law SK, et al. Identification of the haemoglobin scavenger receptor. Nature 2001;409:198–201.

[41] Kodelja V, Muller C, Tenorio S, Schebesch C, Orfanos CE, Goerdt S. Differences in angiogenic potential of classically vs alternatively activated macrophages. Immunobiology 1997;197:478–93.

[42] Weaver LK, Pioli PA, Wardwell K, Vogel SN, Guyre PM. Up-regulation of human monocyte CD163 upon activation of cell-surface Toll-like receptors. J Leukoc Biol 2007;81:663–71.

[43] Kacerovsky M, Drahosova M, Andrys C, Hornychova H, Tambor V, Lenco J, et al. Amniotic fluid concentrations of soluble scavenger receptor for hemoglobin (sCD163) in pregnancy complicated by preterm premature rupture of the membranes and histologic chorioamnionitis. J Matern Fetal Neonatal Med 2010.

[44] Mannik J, Vaas P, Rull K, Teesalu P, Rebane T, Laan M. Differential expression profile of growth hormone/chorionic somatomammotropin genes in placenta of small- and large-for-gestational-age newborns. J Clin Endocrinol Metab 2010;95:2433–42.

[45] McIntyre HD, Serek R, Crane DI, Veveris-Lowe T, Parry A, Johnson S, et al. Placental growth hormone (GH), GH-binding protein, and insulin-like growth factor axis in normal, growth-retarded, and diabetic pregnancies: correlations with fetal growth. J Clin Endocrinol Metab 2000;85:1143–50.

[46] Mestan K, Yu Y, Matoba N, Cerda S, Demmin B, Pearson C, et al. Placental inflammatory response is associated with poor neonatal growth: preterm birth cohort study. Pediatrics 2010;125:e891–8.

[47] Okada T, Matsuzaki N, Sawai K, Nobunaga T, Shimoya K, Suzuki K, et al. Chorioamnionitis reduces placental endocrine functions: the role of bacterial lipopolysaccharide and superoxide anion. J Endocrinol 1997;155:401–10.

Please cite this article as: Tambor V, et al, CysTRAQ — A combination of iTRAQ and enrichment of cysteinyl peptides for uncovering and quantify..., J Prot (2011), doi:10.1016/j.jprot.2011.09.027