INFLUENCES OF EARLY-ONSET INFECTIONS ON THE NEONATAL PROTEOME

WOLFGANG BODENSOHN

MUNICH 2023

Aus der Kinderklinik und Kinderpoliklinik
im Dr. von Haunerschen Kinderspital
Klinik der Universität München
Direktor: Prof. Dr. Dr. Christoph Klein

INFLUENCES OF
EARLY-ONSET INFECTIONS
ON THE NEONATAL PROTEOME

Dissertation
zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Wolfgang Emanuel Johan Bodensohn

aus
München

2023

**Mit Genehmigung der Medizinischen Fakultät
der Universität München**

Berichterstatter:	Prof. Dr. Orsolya Genzel-Boroviczény
Mitberichterstatter:	Prof. Dr. Axel Imhof Prof. Dr. Frederick Klauschen Prof. Dr. Matthias Mann
Mitbetreuung durch den promovierten Mitarbeiter:	Dr. Susanne Pangratz-Fuehrer
Dekan:	Prof. Dr. Thomas Gudermann
Tag der mündlichen Prüfung:	27.07.2023

DEDICATED TO
ALL BABIES AND
THEIR MOTHERS

CONTENTS

APPENDIX 91

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AC	alternating current
ACTR3	actin-related protein 3
Apo	apolipoprotein
cGA	corrected gestational age
CoNS	coagulase-negative staphylococci
CRP	C-reactive protein
C-trap	curved linear trap for ion injection
DC	direct current
DDA	data-dependent acquisition
DIA	data-independent acquisition
EOS	early-onset sepsis
ESI	electrospray ionizer
FDR	false discovery rate
GBS	group B streptococci
HCD	higher-energy collisional dissociation
hnRNPK	heterogeneous nuclear ribonucleoprotein K
HP	haptoglobin
HPLC	high performance liquid chromatography
IL	interleukin
ITIH3	inter-alpha-trypsin inhibitor heavy chain H3
LFQ	label-free quantification
LC	liquid chromatography
LC-MS/MS	liquid chromatography combined with tandem mass spectrometry
LOS	late-onset sepsis
MS	mass spectrometry
MS/MS	tandem mass spectrometry

m/z   mass-to-charge ratio

NEC   necrotizing enterocolitis

PCT   procalcitonin

pp.   postpartum

RF   radio frequency (current)

SAA   serum amyloid A

VLBW   very low birth weight infants

vWF   von Willebrand factor

1D-AEA   1D annotation enrichment analysis

1. INTRODUCTION

Infectious diseases are a leading cause of morbidity and mortality in newborns worldwide. The challenges of the perinatal period and particular immunologic preconditions make neonates especially susceptible to infection and sepsis. Without medical intervention at an early stage, neonatal infections may progress rapidly with fatal or adverse long-term outcomes. Obtaining early diagnostic certainty is, however, fraught with difficulties: due to the unspecific initial symptoms in newborns, due to our incomplete understanding of their dynamic physiology, and – closely related to that – due to the unsatisfying *status quo* of diagnostic means. In the face of suspect but ambiguous symptoms, none of today's established blood tests offers a predictive accuracy that would allow neonatologists to forgo an immediate initiation of antibiotic therapy. For the sake of patient safety, antibiotics are usually administered for at least 24–48 h, normally until blood culture results are obtained. This low-threshold approach, however, leads to a significant proportion of unnecessary treatments, promotes antibiotic resistances, and takes the risk of mid- or long-term consequences due to the disruption of a developing microbiome. Therefore, neonatology would greatly benefit from novel diagnostic means, reducing the number needed to treat while enabling earlier and more accurate identification of those patients who are actually in need of therapy.

Proteomics as a systems biology approach is dedicated to the research of proteins in their totality and systematics. Its central method is mass spectrometry, enabling the identification, quantification and characterization of up to several thousand proteins in tiny amounts of most diverse specimens. Due to the principally unbiased nature and the inherent multiplexicity of such analyses, proteomics is thought to be well-suited for clinical biomarker research. The methodological progress of recent years gives hope for a breakthrough of proteomic applications and biomarkers in clinical practice. Neonatology might particularly benefit from a proteomics-driven progress in biomarker research, not least as to the described challenges with neonatal infections.

In this context, in 2017–2019 the department of neonatology at the Dr. von Hauner Children's Hospital Munich (LMU) cooperated with the Max Planck Institute of Biochemistry Martinsried (MPIB) and performed a large-scale proteomic observational study in neonates. This thesis is centred on the proteome analyses of a subgroup of late preterm and term in-

fants affected by early-onset bacterial infection. The main questions concern the differences between the protein profiles of healthy and diseased newborns as well as changes after the beginning of antibiotic therapy.

Overall, this thesis is to contribute to a deeper understanding of the neonatal blood proteome and of the benefits that proteomic methods may have for the diagnosis and treatment of neonatal infections.

1.1 Neonatal infections

Since the proteomic analyses of this thesis were conducted on neonates who were diagnosed with a **bacterial** infection, the following general remarks will focus only on these and completely omit the broad field of viral, fungal and parasitic infections.

1.1.1 Epidemiology

A child's risk of dying is never higher than during the first month of life (UNICEF, 2021). The neonatal period covers, by definition, 28 days postnatally. For 2020, the United Nations Inter-agency Group for Child Mortality Estimation (UN IGME) indicated a global **neonatal mortality rate** of 17 deaths per 1000 live births, which is an estimated total of 2.3 million fatalities worldwide. On the one hand, the global mortality rate of children has constantly been decreasing over the past decades and has never been lower than in 2020 (UN IGME, 2021). On the other hand, it could be observed that the mortality rate of neonates decreased slower than that of all other children under five years of age (Global Burden of Disease Study, 2021).

Within the neonatal period, infectious diseases are – besides prematurity and asphyxia – a leading cause of death (Lawn *et al.*, 2005; Liu *et al.*, 2015). According to a meta-analysis of global data from 1979–2016, 2202 neonates per 100 000 live births were affected by neonatal sepsis with a mortality of 11–19 % (Fleischmann-Struzek *et al.*, 2018). This results in a global incidence of roughly 3.0 million neonatal sepsis cases per year. For 2013, Liu *et al.* estimated the worldwide number of neonatal deaths due to sepsis at 0.421 million. For 2019, the Global Burden of Disease Study finds a fraction of sepsis and other infections in all neonatal fatalities of approx. 12 % (GBD, 2021). For 2000, Lawn *et al.* estimated that even 36 % of all neonatal deaths worldwide were due to any sort of infection.

Generally, neonatal infections are a major health care issue in any part of the world. In 2019, neonatal mortality rate in high-income economies was ca. ten times lower than in low-income economies (referring to World Bank ranking; UN IGME, 2021). At the same time, the proportion of neonatal deaths due to infectious diseases was only ca. two times lower (referring to sociodemographic index ranking; GBD, 2021).

For Germany in 2020, the UN IGME reports a neonatal mortality rate of 2.25 deaths per 1000 live births, which is an estimated total of 1776 fatalities in the neonatal period. These values are slightly under the European average of 2.53 deaths per 1000 live births. For 2010–2016, a Germany-wide retrospective analysis found an incidence of 1006 cases of neonatal sepsis per 100 000 live births (Born *et al.*, 2021). Case fatality rate was reported as 3.9 %.

1.1.2 Definitions

Infection is generally understood to be a proliferation of micro-organisms in normally sterile body compartments (such as blood or cerebrospinal fluid) leading to an inflammatory reaction (Zemlin *et al.*, 2021).

The term **sepsis** is to be distinguished from that, since it rather accentuates the organism's uncontrolled and harmful response to infection. In adult medicine, a revised international consensus definition for sepsis (the so-called "Sepsis-3") has been established, which terms sepsis a *"life-threatening organ dysfunction caused by a dysregulated host response to infection"* (Singer *et al.*, 2016). In neonatology, however, there is to date no such internationally consented and established definition of neonatal sepsis (Wynn, 2016; McGovern *et al.*, 2020). Among various reasons, this may be due to the very unspecific sepsis symptoms in neonates and the principal challenge of dealing with a broad spectrum of gestational ages.

In praxi, neonatal sepsis is divided into two entities according to their timely occurrence: **early-onset sepsis (EOS)** and **late-onset sepsis (LOS)**.

Table 1.1: Classification of neonatal sepsis

	Early-onset sepsis (EOS)	Late-onset sepsis (LOS)
Onset	within 72 h *pp.*	later than 72 h *pp.*
Pathogens	mainly from maternal vaginal flora (*e.g.* GBS, *E. coli*)	mainly from nosocomial flora (*e.g.* CoNS, *S. aureus*)

For further frequent causative agents of neonatal sepsis *cf.* Tab. 1.3.

The exact definition of EOS and LOS, however, strongly varies (Dong & Speer, 2015). Some authors set the cut-off generally at 48 h, 96 h or even seven days, some differentiate between preterm (72 h) and term infants (7 d), some rely on special groupings when group B streptococci (GBS) are the cause.

1.1.3 Risk factors

The risk of a bacterial infection is inversely correlated with gestational age, birth weight and socioeconomic status (Klinger *et al.*, 2009). It is in direct relation to the interval between rupture of membranes and delivery as well as to the mother's highest intrapartum temperature

(Puopolo *et al.*, 2011). A further major risk factor is vaginal colonization of the mother by GBS, leading to intrapartum transmission or even intrauterine infection (*cf.* Ch. 1.1.4; Patras & Nizet, 2018; Edmond *et al.*, 2012).

The pronounced inverse correlation of risk with birth weight had already been shown in several countries (Vergnano *et al.*, 2011; Boghossian *et al.*, 2013; Tsai *et al.*, 2014) and has recently been confirmed for Germany as well. Born *et al.* (2021) showed that among all deliveries in German hospitals in 2010–2016 a total of 17.7 % of very low birth weight infants (VLBW) were sooner or later affected by neonatal sepsis. However, this was the case for only 2.1 % of neonates with a low birth weight and 0.6 % of neonates with a normal birth weight.

Meconium-stained amniotic fluid (MSAF) during delivery in the absence of other risk factors is not a sufficient reason for starting sepsis diagnostics and antibiotic treatment (Siriwachirachai *et al.*, 2014; Zemlin *et al.*, 2021). Nevertheless, as the newborn might have swallowed or inhaled it *in utero* MSAF cannot be definitely ruled out as a risk factor for EOS (Taskin *et al.*, 2020). In combination with *e.g.* intrapartum fever MSAF is in any case a well-known warning signal for postnatal complications (Levin *et al.*, 2021).

Concerning LOS, the main risk factors (besides degree of immaturity) are strongly connected with hospitalization and invasive medical procedures, like mechanical ventilation, catheterization or surgery (Dong & Speer, 2015). Beyond that, asphyxia and postnatal respiratory complications play a major role as LOS risk factors (Leal *et al.*, 2012).

Table 1.2: Risk factors of neonatal sepsis

Maternal risk factors	Infant risk factors
invasive procedures during pregnancy	prematurity
rupture of membranes	low birth weight
prior to labour (PROM)	male sex
$> 18\,h$ prior to delivery	instrument-assisted birth
chorioamnionitis	prolonged hospitalization
fever *($> 38\,°C$)*	congenital anomalies
leukocytosis *($> 15\,000\,wbc/mm^3$)*	
tachycardia *($> 100\,bpm$)*	
uterine tenderness	
foul odor of amniotic fluid	
foetal tachycardia at delivery *($> 160\,bpm$)*	
group B streptococci (GBS)	
vaginal colonization	
bacteriuria during pregnancy	
infection in previous child	
low socioeconomic status	
poor prenatal care	
poor nutrition	
substance abuse	

Risk factors of neonatal sepsis according to Simonsen *et al.* (2014) and Genzel-Boroviczény & Roos (2019).

1.1.4 Aetiology

A neonatal bacterial infection may have its cause before, during or after birth. Invasive obstetric procedures in the amniotic cavity (*e.g.* amniocentesis or chorionic villus sampling) all bear a risk of introducing pathogens into the child's protected environment and, by that, a risk of amniotic and subsequent foetal infection (Wilkins *et al.*, 1989; Fejgin *et al.*, 1993). Moreover, transplacental hematogenous infections do not occur exclusively with viruses, but also

with certain types of bacteria, such as *Treponema pallidum* (causative agent of syphilis), *Listeria monocytogenes* (causative agent of listeriosis) or others (Lamont *et al.*, 2011; Arora *et al.*, 2017).

The most common aetiology of an EOS, however, is the infant's colonization and subsequent infection through bacteria from the maternal genitourinary and gastrointestinal tract. These may be acquired either antepartum (through an ascending intra-amniotic infection) or intrapartum (after rupture of membranes and/or through passage of birth canal). Either way, it can lead to an invasive infection with the clinical picture of an EOS (Simonsen *et al.*, 2014). Bacteraemia with GBS – the most relevant pathogen in this context – may already be prevalent at the time of delivery (Pyati *et al.*, 1983).

Table 1.3: Bacterial pathogens of neonatal sepsis

Early-onset sepsis (EOS)	Late-onset sepsis (LOS)
Group B streptococci (GBS)*	Coagulase-neg. staph. (CoNS)*
*Escherichia coli**	*Staphylococcus aureus**
Streptococcus viridans	*Escherichia coli*
Streptococcus pneumoniae	*Klebsiella spp.*
Staphylococcus aureus	*Enterobacter spp.*
Enterococcus spp.	*Pseudomonas spp.*
Enterobacter spp.	
Haemophilus spp.	
Listeria monocytogenes	

Predominant bacterial agents of EOS and LOS according to Simonsen *et al.* (2014) and Dong & Speer (2015). Most frequent ones are listed on top (*). GBS and *E. coli* together account for more than two third of all EOS cases, the same is true for CoNS among LOS cases. GBS are foremost *S. agalactiae*, CoNS are e.g. *S. epidermidis* or *S. saprophyticus*. The actual spectrum of pathogens varies significantly across countries and is constantly evolving over time (Shim *et al.*, 2011).

Possible postnatal causes for neonatal sepsis may be the immaturity of dermal/mucosal barriers in preterm infants (Darmstadt *et al.*, 2003), infections of the umbilical cord stump (Moraa *et al.*, 2019), mucosal injuries through endotracheal suctioning (Storm, 1980) or any other adverse effect of invasive medical procedures (*cf.* Ch. 1.1.3).

Within the last decades, long-term observations showed significant shifts as to the predominant pathogens (*cf.* Tab. 1.3) and the incidences of EOS and LOS. An international cohort study covering more than 140 000 term births in 2011–2016 identified GBS still to be the most frequent causal pathogen for EOS (Polcwiartek *et al.*, 2021). In the Western world, however, the fraction of cases due to GBS and *E. coli* declined since the 1980s whereas the importance of coagulase-negative staphylococci (CoNS) and *S. aureus* as EOS pathogens increased (Ronnestad *et al.*, 1998; Bizzarro *et al.*, 2005). At the same time, overall EOS cases decreased and LOS cases increased (Bizzarro *et al.*, 2005). This reduction of EOS cases and the change in the spectrum of causative agents may be attributed to the introduction of GBS prophylaxis in the early 1980s (*i.e.* antepartum screening & intrapartum antibiotics). Improved survival and prolonged hospitalization of preterm neonates may also have led to the increase of LOS cases (Bizzarro *et al.*, 2005; Shim *et al.*, 2011; Dong & Speer, 2015).

1.1.5   Immunology

Neonates are significantly more susceptible to infection and at greater risk of severe progressions than adults. Even pathogens that are very unlikely to cause illness in adults (such as GBS or RSV) can have a serious impact on newborns. However, the "immaturity" of the neonatal

immune system can be seen from at least two different angles (*cf.* Zhang *et al.*, 2017; Tsafaras *et al.*, 2020): on the one hand, it can be understood in the sense of deficient immunologic **fitness** and diverse cellular/humoral impairments ("quantity, coordination & experience"); on the other hand, the aspect of immunologic **regulation** in neonates has increasingly been taken into focus ("quality & unfavourable but necessary ways of coordination"). Two aspects may elucidate why the neonatal susceptibility for infections does not only arise from certain cellular impairments or the predominance of innate over adaptive immunity, but also from very specific ways in which the neonatal immune system is necessarily regulated.

Firstly: The immunologic challenge of pregnancy is to ensure **foeto-maternal tolerance** (PrabhuDas *et al.*, 2015). However, this effort for an immunologically "reconciled" pregnancy and a safe intrauterine development originates by far not unilaterally from the maternal side. For this mutual acceptance *e.g.* regulatory T cells seem to play a major role in both mother and child (Aluvihare *et al.*, 2004; Mold *et al.*, 2008). Mold *et al.* showed that maternal cells passing the placental barrier lead to a presentation of maternal antigens in foetal lymph nodes and subsequently to an induction of T_{reg} cells that suppress immunity against these very antigens. These T_{reg} cells persist long after birth, partly until adolescence. Immediately after birth, however, these or any residues of foeto-maternal tolerance may be disadvantageous for an effective immune response of the newborn.

Secondly: The immunologic challenge of delivery is the newborn's **transition** from a semi-sterile to a microbe-laden environment (Wassenaar & Panigrahi, 2014; Tsafaras *et al.*, 2020). Microbial colonization of all external and internal surfaces is an inevitable and crucial part of adaptation, but theoretically it could also be an overwhelming inflammatory stimulus. In view of that, maintaining a tight immune regulation postnatally also protects developing organs (as the lungs) and mucosal tissues (as in the intestine) from hyperinflammatory damage. Again, regulatory T cells (which are abundant in neonatal mucosa) may play an important role in this process (Thome *et al.*, 2016). Chronically high levels of inflammation might also impede the necessary and ongoing development of immunologic "self-tolerance" (Adkins *et al.*, 2004). In this sense, a cautious encountering of allo-antigens is of similar importance as for auto-antigens. An aggressive pathogen clearance, anyway, does not seem to be an absolute priority of the neonatal immune system (Zhang *et al.*, 2017) and – from this point of view – a certain immunologic "inertia" seems to be favourable. As soon as eventual pathogen growth exceeds a critical limit the neonatal immune system is, however, easily forced into the other extreme and tends to react with harmful hyperinflammation (Zhang *et al.*, 2017; Tsafaras *et al.*, 2020).

Summarizing the two aspects above, an overall challenge for the immune system in early life is **"balanced maturation"**. Any dysbalance in the continuous (intra- & extrauterine) adaptations of immune-regulatory processes favours adverse reactions and increases infection susceptibility (Ygberg & Nilsson, 2012). The neonatal immune system should be capable of

switching between a "moderate under-responsiveness" in normal times and full mobilization of its defensive capacities in threatening situations (Adkins *et al.*, 2004).

Besides this perspective on neonatal immune regulation, numerous aspects reflecting the cellular/humoral impairments of neonatal immunity could still be elaborated. Tab. 1.4 is supposed to highlight some prominent ones. All of these aspects underline how much neonates are in need of special diagnostic and therapeutic caution.

Table 1.4: Impairments of the neonatal immune system

Cellular immunity	
neutrophils	rolling, adhesion & diapedesis ↓ [1] − chemotaxis ↓ [2] − phagocytosis, NETosis & effects of degranulation ↓ [3]
monocytes	phagocytosis & antigen presentation ↓ [4] − adhesion, diapedesis & degranulation ↓ [5] − cytokine production ↓ [6]
dendritic cells	phagocytosis & antigen presentation ↓ [4]
natural killer cells	degranulation & cytotoxicity ↓ [7]
T cells	predominance of naive T cells [8] − T_H1-type response ↓ ("vs. intracellular pathogens") & T_H2-type response ↑ ("vs. extracellular pathogens") [9]
B cells	predominance of naive B cells [8], occurrence of class switch ↓ [11]
Humoral immunity	
cytokines	release upon stimulus ↓ [6] − pro-inflammatory potential ↓ [9, 10] − cellular polyfunctionality ↓ (*i.e.* barely more than one type of cytokine from one immune cell) [9]
antibodies	velocity, peak level, duration & affinity ↓ [11] − response to polysaccharide antigens (encapsulated bacteria) ↓ [12]
complements	serum levels ↓ [13]
Physical barriers	
cutis	permeability ↑ [14]
mucosa	permeability ↑ [15]

[1] Nussbaum & Sperandio (2011) − **[2]** Weinberger *et al.* (2001) − **[3]** Zhu *et al.* (2014) − **[4]** Nguyen *et al.* (2010) − **[5]** Kaufman *et al.* (1999) − **[6]** Yerkovich *et al.* (2007) − **[7]** Guilmot *et al.* (2011) − **[8]** Walker *et al.* (2011) − **[9]** Kollmann *et al.* (2009) − **[10]** Strunk *et al.* (2004) − **[11]** Glaesener *et al.* (2018) − **[12]** Klein Klouwenberg & Bont (2008) − **[13]** Firth *et al.* (2005) − **[14]** Darmstadt *et al.* (2003) − **[15]** Koren *et al.* (2021).

1.1.6 Diagnostics

Given the possibly foudroyant progression of neonatal infections to septic shock, reliable diagnostics enabling early therapy is absolutely crucial.

In reality, however, diagnosis can be challenging and fraught with uncertainty, beginning with the unspecific and initially subtle symptoms of affected newborns. Although they do not allow a reliable differentiation from other conditions such as maladaptation, respiratory complications or metabolic diseases, awareness for these clinical signs is of major importance in neonatology. Their occurrence necessarily initiates further diagnostics and demands for close observation of the respective infant. Tab. 1.5 provides an overview of possible clinical signs which suggest a beginning/ongoing infection. Until proven otherwise, any newborn with inexplicable respiratory distress should be considered as afflicted with infection (Zemlin *et al.*, 2021).

Table 1.5: Clinical signs of neonatal infections

General	poor feeding, hypothermia/fever, central-peripheral temperature difference >2 °C
Respiratory	apnea/dyspnea/tachypnea, grunting, nasal flaring, intercostal retractions
Circulatory	bradycardia/tachycardia, centralization (capillary refill >2 sec), arterial hypotension
Neurological	apathy/lethargy, muscular hypo-/hypertension, irritability/hyperexcitability/touch sensitivity, seizures, taut fontanel
Gastrointestinal	distended abdomen, absence of intestinal sounds, vomiting/obstipation/diarrhea, delayed gastric emptying
Dermal	paleness, cyanosis, erythema (puncture sites, catheters), pustula/abscess/omphalitis, petechial bleedings, edema, icterus

Clinical signs of neonatal infections according to Simonsen *et al.* (2014) and Genzel-Boroviczény & Roos (2019). The intensity of the symptoms above varies depending on gestational age and severity of disease. In preterm infants and Gram-positive infections, symptoms may be less obvious than in term infants or infections due to Gram-negative and fungal pathogens (Simonsen *et al.*, 2014).

Besides clinical evaluation of the neonate, basal initial diagnostics should cover a complete blood count, the C-reactive protein, the interleukines 6 & 8 and a blood culture (Zemlin *et al.*, 2021). Suprapubic or lumbar puncture as well as imaging should be conducted if the severity of symptoms gives reason to do so.

Due to the dynamic physiology of neonates the **complete blood count (CBC)** is strongly dependent on both gestational and postnatal age (Christensen *et al.*, 2009). Therefore, even when applying adapted reference values (*e.g.* for the physiologic leukocyte left shift within 48 h *pp.*; Manroe *et al.*, 1977) a CBC is generally to be interpreted with caution as far as an infection is to be evaluated. The I/T ratio (immature to total neutrophils ratio) as a measure for this leukocyte left shift may have a high negative predictive value (Murphy & Weiner, 2012; Saboohi *et al.*, 2019) and thus may help rule out neonatal sepsis, but it is lacking sensitivity. Similarly, even the total count of white blood cells (WBC) offers a merely low positive predictive value in sepsis detection due to a late increase of WBC after an onset of sepsis ("latent period" of ca. 4 h; Christensen *et al.*, 1985). Ultimately, the CBC in all its aspects has foremost a supporting diagnostic function rather than being relied upon exclusively.

Until today, the most established inflammatory biomarker in neonatology is the **C-reactive protein (CRP)**. Originally, it was discovered through and named after its ability to form precipitates with a component of the pneumococcal cell wall, the C-polysaccharide (Tillett & Francis, 1930). Today, CRP is an extensively researched serum protein produced by hepatocytes (Hurlimann *et al.*, 1966) in the so-called "acute phase", the organism's response to tissue injuries and inflammatory processes of any kind (Sproston & Ashworth, 2018). CRP's diverse physiological functions include (according to Marnell *et al.*, 2005): i) opsonization of intruded bacteria aiming at an activation of the complement system (*i.e.* fostering innate immune response); ii) mediation of immune-modulatory effects on macrophage- and T-cell-level through cytokine release; iii) marking apoptotic cells for phagocytosis and increasing their clearance; iv) binding damaged membranes and, by that, preventing autoimmunity; and

several more. Du Clos (2000) describes CRP as a "surveillance molecule for altered self and certain pathogens", comparable to a polyvalent immunoglobulin.

In daily clinical use, CRP helps assess a general "inflammatory status" and thus in identifying acute (bacterial) infections as well as evaluating responsiveness to treatment. The latter results from its property to increase reasonably soon after a suitable stimulus and to decrease just as reliably as the stimulus ceases (Du Clos, 2000). As it does not pass the placental barrier (at least not in relevant amounts) there is no risk of confusing maternal and neonatal inflammation (Rozansky & Bercovici, 1956; Nielsen *et al.*, 1990). However, there is an unspecific increase of CRP within the first days of life, presumably due to stress of delivery (Kaapa & Koistinen, 1993; Chiesa *et al.*, 2001). This physiologic rise of CRP can reach ≥ 1.0 mg/dl in up to 20 % of healthy term neonates (Mjelle *et al.*, 2019), which might complicate eventual diagnostics. While in several diseases sole CRP levels correlate well with severity and outcome, a stringent correlation for sepsis is at least controversial, both in adults (Gradel *et al.*, 2013; Kurisu *et al.*, 2020) and neonates (Philip, 1985; Vasiljević *et al.*, 2008). As for the prediction of neonatal sepsis outcomes, however, height of CRP in relation to albumin or platelet levels might be a promising complement to current diagnostic capabilities (Li *et al.*, 2021a,b). The most significant limitation of CRP is its delayed increase and its lack of sensitivity in the early phase of an acute infection. CRP levels increase 6 h after their respective stimulus and reach a peak after 48 h (Pepys & Hirschfield, 2003). *De facto*, this makes CRP a "late" marker which is more suitable for surveillance than for diagnosis itself. It necessarily needs an "early" marker as its partner which is normally interleukin-6 (IL-6), interleukin-8 (IL-8) and/or procalcitonin (PCT).

The cytokine **IL-6** originates from endothelial and immune cells at the onsite of infection and is the main inductor of hepatic CRP production and release (Weinhold & Ruther, 1997). IL-6 levels reach their peak already 2 h after their respective stimulus and return to their baseline within 8 h (Shannon *et al.*, 2007). Unlike CRP, this makes IL-6 inappropriate for monitoring. However, as CRP's precedent in the inflammatory cascade the sensitivity of IL-6 in the early phase of neonatal sepsis is high (80–100 %; Hofer *et al.*, 2012). Oftentimes, it is determined together with IL-8, another pro-inflammatory cytokine with similar properties and a high negative predictive value for neonatal sepsis in its early phase (90–94 %; Franz *et al.*, 2007).

In septicaemia – mediated through bacterial endotoxins and inflammatory cytokines – the prohormone **PCT** is released, presumably from pulmonary and intestinal neuroendocrine cells (Becker *et al.*, 1980; Maruna *et al.*, 2000). PCT levels rise at 4 h and reach their peak 6 h after their respective stimulus, which will be maintained for 24 h (Dandona *et al.*, 1994). This makes it another early biomarker for acute bacterial infections with kinetics between that of IL-6 and CRP. PCT's sensitivity and specificity in the early phase of neonatal sepsis are high (both 80 %; Vouloumanou *et al.*, 2011). Similarly to CRP, one can observe a physiological in-

crease within the first two days of life (independently from gestational age, Sachse *et al.*, 1998). Furthermore, there are rising PCT levels also in conditions associated with haemodynamic or hypoxic complications (Lapillonne *et al.*, 1998; Monneret *et al.*, 1998). Despite these limitations, PCT has proven partly superior to CRP in the prediction of neonatal sepsis (Vazzalwar *et al.*, 2005).

Although an abundance of potential further biomarkers has been evaluated in recent years, the diagnostic triumvirate out of CRP, IL-6, and PCT is *in praxi* largely unchallenged. For instance, **cell surface antigens** of bacterially activated leukocytes (as CD11b or CD64) performed in some studies with high sensitivity and specificity for neonatal sepsis (Nupponen *et al.*, 2001; Ng *et al.*, 2002). However, they are quite volatile and their determination is technically demanding (Haque, 2010). Not least, this is why CRP – despite all limitations (as its delayed sensitivity) – remains a central component to any neonatal infection diagnostics: due to its widespread availability and rapid, convenient and low-cost determination (Hofer *et al.*, 2012).

Besides blood biomarkers as indicators of a certain (patho-)physiological status, the formal gold standard in diagnosing infectious diseases are actually **microbial cultures**. However, these are fraught with decisive drawbacks: (1) the late obtainment of results due to a time to positivity commonly between 24–48 h (Kuzniewicz *et al.*, 2020); (2) a certain risk of false-negative results due to too little volumes of inoculated blood (Woodford *et al.*, 2021); (3) the high risk of false-positive results due to a contamination with cutaneous commensals during the blood sampling (Hall & Lyman, 2006). Therefore, in daily newborn healthcare the time-critical initial decision-making is based on clinical signs and blood biomarkers, whereas blood cultures serve as verification of the diagnosis, as re-evaluation of an empiric antibiotic therapy and for its targeted deescalation.

1.1.7 Therapy

In view of all inevitable shortcomings of today's diagnostic possibilities, there is an iron consensus in treating newborn patients: to react to any reasonable suspicion of a bacterial infection by immediately initiating **calculated** antibiotic therapy (Zemlin *et al.*, 2021). This urgency arises from a direct relation between delaying a necessary treatment and increased morbidity and mortality (Weiss *et al.*, 2014). An intraamniotic infection or some constellations of mere risk factors (*cf.* Ch. 1.1.3) may even justify a start of antibiotic therapy despite an absence of clinical symptoms or inflammatory markers (NICE, 2021).

In turn, this safety comes at the price of a more or less pronounced **overuse** of antibiotics in neonates, entirely depending on local standards. While a Norwegian group (Fjalstad *et al.*, 2016) reported 2.3 % of all live-born term infants of their cohort having been treated with antibiotics, a US-American group (Mukhopadhyay *et al.*, 2013) reported a fraction of 7.4 % (including late preterm infants $\geq$ 35+0). In both studies, culture-positive EOS cases were less

than 0.1 %. Kiser *et al.* (2014) drew attention to the fact that 24.2 % of a cohort of neonates born after an intraamniotic infection were (in accordance with former guidelines) treated with antibiotics, while the fraction of patients with positive blood cultures was 0.7 % in this study.

Among all treated patients of the aforementioned studies, indeed, a relevant number of "clinical" EOS cases is to be assumed, *i.e.* cases which do not allow microbiological detection of a causative pathogen despite an obvious and compelling constellation of symptoms and other findings. Nevertheless, the numbers above illustrate a great potential for savings in antibiotics, resources and not least discomforting interventions within the first days of life of many patients. More sensitive diagnostics, more solid criteria and more tailored therapies are obviously required. For VLBW, it could be shown that a prolonged application of empirical antibiotics after sterile blood culture results may even increase the risk of developing necrotizing enterocolitis (Abdel Ghany & Ali, 2012). Caution is required, not least as possible long-term health implications of an early life exposure to antibiotics come more and more into focus. These are increased odds for developing childhood asthma (Murk *et al.*, 2011), diabetes (Boursi *et al.*, 2015), inflammatory bowel disease (Ng *et al.*, 2013) or others (Schulfer & Blaser, 2015).

Thus, a further key principle – complementary to an early start of antibiotic treatment – is its early **re-evaluation**, usually after 36–72 h on the basis of clinical appearance, inflammatory markers and blood culture results (van der Hoeven *et al.*, 2022). German, British and US-American guidelines generally advise to stop empiric antibiotic therapy after 36–48 h if the suspicion of infection cannot be maintained based upon normal clinical evaluation, negative inflammatory markers and sterile blood cultures (Zemlin *et al.*, 2021; NICE, 2021; Puopolo *et al.*, 2018a,b). Referring to their (prospective, multicentre, randomized and controlled) intervention study, Stocker *et al.* (2021) argue that after 36 h of empiric antibiotic therapy and serial measurements of CRP and PCT the negative predictive values of both markers do not rise any further. According to the authors, after 36 h and provided normal CRP and PCT levels a discontinuation of antibiotic therapy is reasonable.

In clinically unambiguous or culture-proven cases, empirical antibiotic therapy ought to be changed to a de-escalated **pathogen-directed** choice of antibiotics, which is then to be maintained over the appropriate period of time. For a culture-negative but clinically unequivocal infection, an empirical therapy over up to 10 d is to be considered. In culture-proven sepsis, 10 d of treatment is generally indicated. In case of meningitis, the duration is usually 14–21 d (Nizet & Klein, 2015).

As far as the choice of antibiotics is concerned, it can be summarized that in literally every case of sepsis two aspects have to be taken into account: firstly, the onset of symptoms, as it points to an EOS- or LOS-specific spectrum of pathogens (*cf.* Ch. 1.1.2); secondly, the local conditions, as each country and ultimately each hospital has its specific microbial flora with certain patterns of susceptibility to antibiotics.

Today's probably most established combination of antibiotics for the initial empiric therapy of EOS is a broad-spectrum penicillin (mostly ampicillin) plus an aminoglycoside (mostly gentamicin) since they act synergistically against GBS and listeria (Baker *et al.*, 1981; Mac-Gowan *et al.*, 1998). A widely used alternative are third-generation cephalosporines (mostly cefotaxime) instead of the aminoglycoside. This combination, however, is increasingly criticized for the fast development of resistances (Bryan *et al.*, 1985), unfavourable selection of other pathogens (e.g. ESBL bacteria and candida; Le *et al.*, 2008; Manzoni *et al.*, 2006) and even a questionably higher mortality compared to a therapy with ampicillin/gentamicin (Clark *et al.*, 2006). However, in suspected meningitis – especially when attributable to Gram-negative bacilli – cefotaxime should always be considered due to its high capacity of penetrating tissues and cerebrospinal fluid (Begue *et al.*, 1984). Similarly as for EOS, the first line antibiotic therapy of LOS usually is a penicillin (mostly flucloxacillin or vancomycin) plus an aminoglycoside (mostly gentamicin). In the therapy of LOS, however, national standards differ more significantly than in the therapy of EOS.

As mortality rates remain high even under the most up-to-date antimicrobial and supportive therapy, there has been an intense search for adjuvant therapies. While G-CSF (granulocyte colony-stimulating factor) or IVIG (intravenous immunoglobulines) did not prove beneficial (Schibler *et al.*, 1998; INIS, 2008, 2011), pentoxifylline is still under evaluation (Lauterbach & Zembala, 1996; Peng & Xia, 2019). It is an immunomodulating agent with anti-inflammatory properties (TNF-α-inhibitor) and positive effects on microcirculation. In their meta-analysis from 2019, Peng & Xia found pentoxifylline to be associated with significantly reduced hospitalization and potentially reducing mortality among neonates suffering from sepsis.

1.1.8 Long-term morbidities

Despite all diagnostic and therapeutic endeavours, neonatal infections and sepsis are not only associated with relevant mortality (*cf.* Ch. 1.1.1) but also with unfavourable outcomes and long-term morbidities, both in preterm and term neonates.

Especially when sepsis is followed by meningitis, **neurodevelopmental outcomes** are a major issue. A prospective active surveillance study from Germany covered 347 preterm and term infants born in 2001–2003 who had survived GBS-induced sepsis (without or with meningitis). 14 % of them had neurologic sequelae at the time of discharge (Fluegge *et al.*, 2006). Among these, hydrocephalus and cerebral seizures were, by far, the most frequent manifestations. In a British cohort of preterm and term infants born in 1996–1997, among 84 survivors of a culture-positive bacterial meningitis in the neonatal period only 40 % did not show any sign of disability at five years of age (de Louvois *et al.*, 2005). 52 % showed mild or moderate, 7 % severe signs of disability.

In VLBW, the evidence for poor outcomes is even more compelling. An Israelian population-based observational study included nearly all VLBW ($n = 15\,839$) born over a period

of ten years (1995–2005). For those 2.4 % of study patients with EOS, the odds for mortality, severe intraventricular haemorrhage (IVH), bronchopulmonary dysplasia (BPD) and retinopathy of prematurity (ROP) were all significantly increased (Klinger *et al.*, 2010). Attempts to explain such observations mostly refer to the general dysregulation of inflammatory processes (*cf.* Ch. 1.1.5) and an immature blood-brain barrier in preterm infants (Adams-Chapman, 2012). Ultimately, the exact pathophysiology of these consecutive phenomena and of neonatal sepsis itself is yet to be fully understood.

1.2 Proteomics

1.2.1 Basics

The so-called **central dogma of molecular biology** going back to Francis Crick (1958) essentially deals with the flow of genetic information in any organism. Oftentimes, it is broken down to the formula *"DNA makes RNA, and RNA makes protein"*. This points to the dogma's key message that there can neither be any transfer of information between proteins and nucleic acids nor between proteins and proteins. Genetic information (*i.e.* any determination of nucleic acid sequences and thereby amino acid sequences) that once has reached protein level is virtually "caught" on it. While this postulate of an absolute unidirectionality within the flow of genetic information has long been abandoned (remember reverse transcription or the effects of prions), the conception of proteins as the central effectors of nearly all cellular functions is still valid. Performing most diverse tasks – from structuring to signalling – proteins are sometimes referred to as "working horses" of the cell or even as "nano-machines" (Zaccai, 2003). While the genome is often referred to as some sort of "construction plan", the proteome may consequently be seen as this plan's actual implementation, realized by the cellular machinery out of transcription (DNA to RNA), translation (RNA to proteins) and post-translational modifications of proteins (PTM). Caterpillar and butterfly may be genetically identical, but they totally differ in their proteome and the resulting morphology (Mishra, 2010). Therefore, the proteome is the effective link between genotype and phenotype. Or in other words: "genotype makes proteome and proteome makes phenotype".

The term **proteome** was shaped in the 1990s when, first, Marc R. Wilkins used this coinage to describe the *"entire protein complement expressed by a genome, or by a cell or tissue type"* (Wilkins *et al.*, 1996). The proteome can thus be understood as the totality of proteins of a defined system, from single cell to entire organism. Therefore, **proteomics** is a systems biology approach that deals with the expression, structure, function, modification and interaction of proteins. The most widely used method for studying the proteome is mass spectrometry (MS). Whithin the last decade, MS-based proteomics has developed into a high-throughput method. A key characteristics is the high complexity of raw data, since commonly data from multiple physi-

cal dimensions are recorded. These need to be disentangled in an elaborate raw data analysis before protein abundance information on a sample can be derived. While more conventional approaches in molecular biology often investigate only few selected gene products in a targeted manner, proteomics usually aims at a maximum of analytical depth and an unbiased detection of as many gene products as possible.

The number of protein-coding genes of the human being is said to be around 20 000 (Salzberg, 2018). Since the completion of the Human Genome Project in the early 2000s there is, however, still no definitive consensus between the leading genome databases about the exact number of protein-coding genes. Determining the total number of resulting proteins is even more delicate. Indeed, there may be just as many immediate gene products as coding genes, *i.e.* one *"representative, non-redundant"* protein for any gene locus (terminology by Uhlen & Ponten, 2005). On the other hand, the actual number of functional protein variants originating from this immediate gene product (also called **proteoforms**) can exceed the number of protein-coding genes by far, totally depending on the mode of counting. Even when only considering the number of possible protein variations through RNA splicing and proteolytic maturation, some estimations already end up with 500 000 functional proteins (Uhlen & Ponten, 2005). And yet, this does neither include the hundreds of thousands of possible proteoforms due to posttranslational modification (such as glycosylation, phosphorylation, amidation, lipidation, ubiquitinylation *etc.*) nor does it include the millions of possible protein variants due to somatic DNA rearrangements as they can be found in the synthesis of immunoglobulins (Uhlen & Ponten, 2005).

1.2.2   Challenges

Having laid out the previous numbers, it becomes obvious that the proteome is an enormously complex system. The first complete proteomic characterization of an unicellular organism was that of the yeast species *Saccharomyces cerevisiae* (de Godoy *et al.*, 2008). This was 12 years after the complete sequencing of its genome (Goffeau *et al.*, 1996). Nearly 20 years after the first publication of a whole human genome sequence (Consortium, 2004), an analogous proteomic achievement is not even in sight. The human proteome is far from being exhaustively understood or even "decrypted" (Aebersold & Mann, 2016), despite decades of research and significant progress in many aspects. State-of-the-art proteomic workflows detect around 10 000 proteins in cell culture experiments (Wang *et al.*, 2020b). Comprehensive studies aim at exploring cellular protein localizations (Itzhak *et al.*, 2016) or expression patterns in the course of cell cycle (Ly *et al.*, 2014). Others try to shed light on the complex field of signal transduction pathways (Humphrey *et al.*, 2015) and protein-protein interactions, the so-called interactome (Huttlin *et al.*, 2017). Entire cell type or organ-specific proteome maps were determined (Sharma *et al.*, 2015; Wang *et al.*, 2019) and even first attempts to draft maps of the complete human proteome were released (Kim *et al.*, 2014; Wilhelm *et al.*, 2014). And yet, all of these develop-

ments should not hide the fact that proteomic methods – with the exception of MALDI-TOF in microbiology (Croxatto *et al.*, 2012) – are still far away from a broad application in clinical routine. The promise of a widely successful biomarker discovery through proteomics is yet to be fulfilled (Geyer *et al.*, 2017).

One explanation for this gap between aspiration and reality is technological. Virtually all progress in proteomics is tied to the challenge of overcoming **technical** barriers in mass spectrometry (MS), its pivotal analysis method. Likely more than any other field of life sciences, proteomics appears to be profoundly interwoven with MS. Some authors have already risen the question *"whether technology drives biology or whether biology drives the development of new technologies"* (Aebersold, 2003). For a long time, a major barrier for a broader application of MS-based proteomics had been the fact that depth of detail had to be bought by extensive analysis time. Studies with a high protein coverage typically had a reduced sample throughput and vice versa. Newer proteomic workflows, however, address this dilemma and concentrate on a successful balance between coverage and throughput (Geyer *et al.*, 2016a). All in all, a series of challenges in terms of sample preparation, instrumentation, peptide identification, protein quantification and even study design may have contributed to the fact that the clinical benefit of proteomics has not been resounding, yet (Geyer, 2017).

Another constant challenge in proteomics is rather of **statistical** nature and is founded in its property as high-throughput method. As it usually deals with most extensive data sets, evaluating the true effect sizes of proteomic findings can be difficult. Hundreds to thousands of simultaneous statistical tests (one for each detected biomolecule) are quite in the nature of all omics technologies. But as the risk of false positive findings naturally increases with the very number of performed tests, proteomics is particularly error-prone in this regard. The correct contextual interpretation of a high number of statistical tests and their (seemingly) significant results is known as the problem of **multiple hypothesis testing** (*cf.* Ch. 1.2.6). It is, for sure, an omnipresent challenge in proteomics (Diz *et al.*, 2011).

The latter becomes all the more evident when considering some biological peculiarities of the proteome. In contrast to the relatively static genome, which is identical in virtually any cell of an organism, the proteome as the variable product of the genome is extremely dynamic. It reacts sensitively to external influences (such as diseases or drugs) and the expression level, the localization or the activity of proteins can change quickly and significantly. Of course, these dynamics can be very informative, but they have to be carefully differentiated from a **basal noise** of continuous processes on cellular level, and over-interpretation is to be strictly avoided (Friedman, 2012). Apart from that, protein levels also show a certain range of **inter- and intra-individual variability**. Some may be stable over time within one individual while the general protein levels strongly vary between several individuals (Geyer *et al.*, 2016b). Others may be variable over time for one person but the relative changes are similar between several individuals or conditions (Geyer *et al.*, 2016a).

A further proteomics-specific and even more significant issue is known as the **problem of dynamic range**. Proteins appear in extremely different abundances. In blood plasma, ten to twelve orders of magnitude can lie between the most abundant protein (albumin, ca. 50 mg/ml) and the least abundant ones (as cytokines, ca. 5 pg/ml, Hortin *et al.*, 2006). For cellular proteomes, the relation is rather six orders of magnitude (Zubarev, 2013). In both compartments, however, the abundances approximately represent a logarithmic normal distribution. Ultimately, this leads to an overlaying of the least abundant proteins by the highly abundant ones, which makes the former undetectable for MS (Wu & Han, 2006; Zubarev, 2013). As a result, even state-of-the-art MS-based proteomics has to accept this inherent **lower detection limit**.

All of the aforementioned issues can be notable confounders in proteomic analyses, which have to be faced by solid study designs, precise instruments and thorough statistics.

1.2.3 Potential

All of these complexities do, nonetheless, imply the true (or at least expected) potential of proteomics, too: which is to act as a precise "mirror" of cellular and systemic processes (Geyer *et al.*, 2016a). As laid out before, the proteome can be considered as the link between genotype and phenotype, which is not only true for the state of health but also that of disease. Proteins are thought to be the ideal indicators of any condition, be it a physiological or pathological one, and on any level – in cells, tissues, organs or the whole organism. A profound understanding of the human proteome is supposed to naturally come along with the discovery of many novel protein **biomarkers** (*i.e.* biological surrogate parameters, BDWG, 2001) or whole biomarker networks. This would certainly be of great value for many efforts in modern medicine: to diagnose diseases early and precisely, to prognosticate and monitor their progression reliably, and to identify pathways and molecular targets for new causal therapies.

According to statistics of the Department of Laboratory Medicine at Klinikum Großhadern in Munich (LMU university hospital), in 2016, for nearly 80 % of all inpatients at least one blood-based laboratory test was performed. In ca. 40 % of all blood tests, the subject were proteins with their concentrations and enzyme activities (Geyer *et al.*, 2017). These exemplary numbers indicate the eminent role of proteins in clinical decision-making, which they have been holding for a long time at least in the form of conventional single-protein tests (as immunoassays or enzymatic assays). The principal problem with such tests, however, is their often obvious lack of **specificity**. The new edition of the International Statistical Classification of Diseases and Related Health Problems (ICD-11) contains 55 000 entries, in contrast to its predecessor version with 14 000 diagnoses (ICD-10). Even when theoretically assuming only 5000 well-distinct pathological states of the human body, it would still remain more than questionable if – in view of ca. 20 000 protein-coding genes – a highly specific stand-alone biomarker existed for each of these conditions (Geyer, 2017). Despite a notable number of monogenetic

diseases, most pathologies are yet perceived as results from complex interactions of genetic, epigenetic and environmental influencing factors. Taking the specific dynamics of diseases into account as well, it becomes fully obvious that single readouts and even serial readouts of one biomarker must always be interpreted with caution.

A typically proteomic approach to face this problem of biomarker-specificity is the search for **multi-biomarker panels**. The idea is to put a blood-test founded decision on a broader basis by a contextual evaluation of a whole set of biomarkers. Combining several diagnostic findings is clearly one of the most intuitive requirements on each level of clinical decision-making, be it anamnesis, physical examination, imaging or other techniques. But especially in laboratory diagnostics – where the potential is particularly large – the idea of standardized multi-parameter panels is not as established yet as it could be. Mixed scoring instruments, that include clinical observations and single lab values, have partly been in use for a long time. The Child-Pugh score *e.g.* includes albumin and bilirubin for a classification of liver cirrhosis (Child & Turcotte, 1964; Pugh *et al.*, 1973). The SOFA-score integrates creatinine and bilirubin concentrations to assess acute organ dysfunction in sepsis (Vincent *et al.*, 1996). Several others of such mixed scoring instruments could still be mentioned, but pure biomarker panels are yet a rarity in clinical use. In gynaecology, the sFlt-1/PlGF ratio has meanwhile been widely established for the assessment of pre-eclampsia. A tandem-measurement of the anti-angiogenesis factor soluble Fms-like tyrosine kinase-1 (sFlt-1) and of the placental growth factor (PlGF) has proven to be very reliable in differentiating between pregnant women with and without pre-eclampsia (Levine *et al.*, 2004; Lapaire *et al.*, 2010).

Beyond these conventional examples the, to date, only genuine multi-biomarker panel in clinical use, that was found by means of proteomics, is the OVA1 test (Petricoin *et al.*, 2002; Rai *et al.*, 2002; Zhang *et al.*, 2004; Bristow *et al.*, 2013). In 2009, the OVA1 test was approved by the United States Food and Drug Administration (FDA) for a targeted assessment of pelvic masses suspected for ovarian cancer. It includes an evaluation of CA 125, Apolipoprotein A1, Beta-2 microglobulin, Transferrin and Pre-albumin. The OVA1 test does also illustrate that an inconspicuous bundle of unspecific parameters can potentially be combined to a significant set of markers.

The actual challenge remains to uncover the complex interactions and molecular regulation mechanisms behind. Once being able to sufficiently handle these complexities of the human protein mosaic, the theoretical potential for biomarker discovery would be immense. With tens of thousands of functional proteins the number of arithmetically possible combinations for such panels seems nearly inconceivable. Even when presuming a high number of redundant and unspecific proteins as not eligible for the formation of such panels, it is only reasonable to expect this general idea of "multiplex" laboratory diagnostics to be more powerful than the *status quo* of single-protein tests. Past decades have been characterized by the (mostly futile) search for "magic bullets" in the diagnostics of this or that type of cancer

or some other diseases (Hurst, 2009). However, the upcoming (or at least expected) era of individualized medicine and of computer-assisted clinical decision-making may be the time for a paradigm shift: from stand-alone biomarkers towards powerful multi-biomarker panels (Hurst, 2009).

The current conventional and lengthy approach of proteomics-based biomarker discovery could be summarized as follows: Few patients and controls are, first, examined in great detail leading to a large number of biomarker candidates. In subsequent validation stages, the number of biomarker candidates decreases while patient groups grow. This is what Geyer *et al.* (2017) denote as the classical triangular strategy and in contrast to which they present their concept of a **rectangular strategy**. The authors argue that – due to the achieved methodological advances in MS – the laborious multilevel scheme might be obsolete, namely when two independent, large and similarly sized cohorts are *a priori* analysed at the greatest possible depth (*cf.* Fig. 1.1). The overlapping biomarker candidates are supposed to be more significant than such filtered out by the conventional approach. This rectangular strategy relies on the assumption that the most interesting correlations are to be found when naturally emerging from huge proteomic data sets in a hypothesis-free and unbiased manner. Although it still leaves some questions of practicability, cost-effectiveness and not least of bioethics, this may be an approach which unleashes additional potential in proteomic biomarker discovery.
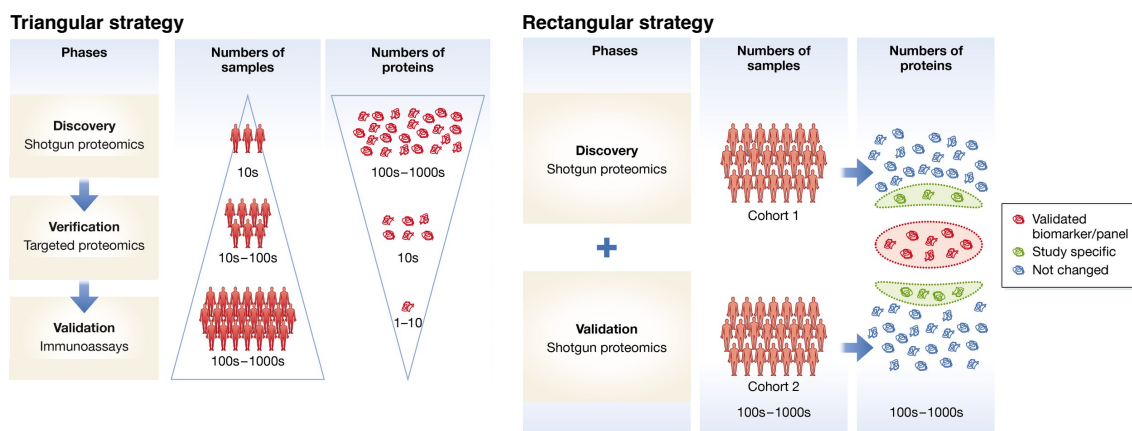
Figure 1.1: Triangular vs. rectangular strategy. Two concepts of proteomics-based biomarker discovery according to Geyer *et al.* (2017). In the first one, cohort size and protein coverage are inversely proportional over the different stages. In the second one, two large cohorts are *a priori* analysed at great depth of detail. "Shotgun proteomics" is a synonym for LC-MS/MS-based "bottom-up" proteomics. In this method, proteins are subjected to proteolytic digestion prior to their mass spectrometric identification (*cf.* "Basics of mass spectrometry" in the appendix). Figure adapted from Geyer *et al.* (2017).

1.2.4   Neonates

A perpetual challenge in paediatrics is that the majority of all medical and pharmaceutical research is conducted with a focus on adults. For instance, the "off-label use" of drugs, *i.e.* drugs that have never been specifically tested and approved for a certain purpose or patient collective, is – due to the lack of alternatives – a daily practice of paediatrics and neonatology (Schrier *et al.*, 2020). It is therefore barely surprising that all the more in the young field of
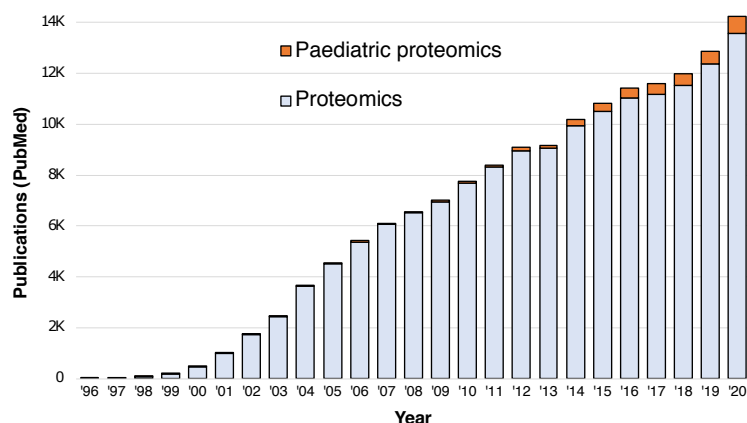
Figure 1.2: Publication numbers. A literature review in the PubMed database for 1996–2020 showed the proportions depicted above. Proteome research only deals to a very small extent with paediatric issues. The percentage of such studies, however, steadily increased (from approx. 1 % in the early 2000s to 4 % in 2020). Search terms in "all fields" were "proteomics NOT paediatrics" vs. "proteomics AND paediatrics".

clinical proteomics only a minimal percentage of studies deals with children, although this proportion has slightly increased in recent years (*cf.* Fig. 1.2). Today, proteomics may be applied to a considerable range of paediatric issues, such as genetic or metabolic diseases, haemato-oncological, cardiovascular or even surgical questions (Everett & Ignjatovic, 2014). Nevertheless, in a recent review about biomarker research in premature infants, Letunica *et al.* (2021) stated a considerable lack of proteomic evidence for this population. And yet there are elucidating contributions of MS-based proteomics in the field of inflammation and infection in neonates, of which some are to be outlined in the following.

Suski *et al.* (2018a,b) examined both cord blood and blood plasma of preterm infants ($\leq$ 30+0) and could show severe disturbances in protein networks associated with **inflammation**, immunomodulation, coagulation and complement. Compared to cord blood of term infants and to follow-up plasma samples at 36 weeks of corrected gestational age (cGA), most diverse proteins of the mentioned systems were significantly decreased or partly increased. The authors interpret this picture of overall normalizing protein patterns with increasing post-natal age as a *"progression toward recovery from perinatal perturbations"*. Referring to healthy term infants, Bennike *et al.* (2020) observe a nearly gradual increase of proteins from the classical and the terminal complement pathway in the first seven days of life. Moreover, the mere postnatal adaptation on the first day of life seems to be characterized by a significant temporary acute phase reaction, mainly through an increase of SAA1 (Bennike *et al.*, 2020; Marchini *et al.*, 2000; Levy, 2007; Pettengill *et al.*, 2014). Any search for novel infection biomarkers in newborns has to take account of this dynamic physiology.

These and other peculiarities of the neonatal proteome may have contributed to the so far limited success in biomarker discovery for newborns. However, some promising approaches and results from recent time that deal with three of the most significant complications in neonatology, EOS, LOS and NEC (necrotizing enterocolitis), shall be introduced in more detail in the following.

In several studies published in 2005–2010, Buhimschi *et al.* examined the proteome of amniotic fluid and found neutrophil defensin-1 & -2 and S100-A8 & -A12 (calgranulin A & C) to be effective markers for intra-amniotic infection and/or inflammation, given that at least two

of four proteins are measurable (Buhimschi *et al.*, 2005). S100-A8 & -A12 in amniotic fluid is associated with a subsequent **EOS** and poor neurological outcomes (Buhimschi *et al.*, 2007a,b, 2009; Buhimschi & Buhimschi, 2010). Furthermore, Buhimschi *et al.* (2011) assessed cord blood of a cohort of 180 preterm infants, several of whom had blood-culture-positive EOS. Multistage analyses showed that antenatal exposure to intra-amniotic infection and/or inflammation leads to a genetic "switch-on" pattern in haptoglobin expression. Highly increased levels of haptoglobin (HP) and HP-related protein in cord blood reflect a foetal adaptation to intrauterine immunologic stress. They may be used as early biomarkers of an emerging EOS with positive effects on early and targeted treatment. The observed "switch-on" pattern was independent of gestational age and birthweight, suggesting it to be a powerful and clinically useful indicator.

Since 2010, at least five major proteomic studies attempted to find reliable indicators for NEC and LOS in preterm infants. Similarly to EOS, both conditions are clinically severe and associated with a high risk of death or poor long-term outcomes, such as bronchopulmonary dysplasia, neurodevelopmental impairment or gastrointestinal complications (Sharma & Hudak, 2013; Shah *et al.*, 2015).

Ng *et al.* (2010) identified the pro-apolipoprotein C2 (Pro-Apo-C2) and a variant of the serum amyloid A (SAA) as promising indicators in blood plasma for **NEC** and **LOS**. The suggested "ApoSAA score" combines both proteins. It was reported to identify cases and to stratify them into risk categories, ultimately allowing for a more targeted initiation and termination of antibiotic treatment. Murgas Torrazza *et al.* (2013), however, criticized the ApoSAA score for enabling only late diagnoses and for not distinguishing between NEC and LOS. A safe distinction between these life-threatening conditions with overlapping symptoms is critical since it affects the decision on enteral feeding, eventual surgery and details in antibiotic treatment.

Murgas Torrazza *et al.* (2013) themselves approached the diagnosis of NEC by buccal swabs, obtaining buccal epithelium as a representative of the intestinal epithelium (both have similar embryologic roots). Proteomic analyses of the buccal proteins suggested the interleukin-1 receptor antagonist (IL-1RA) as a promising predictor of NEC. Another non-plasma-based approach of proteomic NEC diagnosis was presented by Sylvester *et al.* (2014) who identified a panel of fibrinogen peptides in urine (FGA1826, FGA1883 and FGA2659) as promising indicators of an emerging NEC (AUC = 0.856). By combining it with a specifically developed scoring instrument out of 27 clinical parameters, the outcome of all 65 included NEC cases could correctly be predicted as either conservatively or surgically treatable. Additionally, non-invasive sampling of urine or buccal epithelium would be a great practical benefit, enabling easier monitoring of patients at risk than by repeated blood sampling.

Stewart *et al.* (2016) examined the serum proteomes of ten NEC and LOS patients longitudinally over three standardized time points. At the time of diagnosis, they state associations of

both diseases with several proteins, such as CRP, SAA2, HP and others. One of the main conclusions, however, is the low likelihood for the existence of a single, highly specific biomarker for NEC and/or LOS.

Chatziioannou *et al.* (2018) compared the serum proteomes of 25 NEC and 18 LOS patients to matched controls and to each other. They presented two biomarker panels of three proteins each, which they claim to have excellent accuracy in diagnosing LOS and in distinguishing NEC from LOS. The first panel out of CRP, Apo-A4 and CETP (cholesterol ester transfer protein) allowed for a highly precise distinction between LOS and control at the time of diagnosis (AUC = 0.98; 95 % CI = 0.917–1). The second panel out of Apo-A4, Apo-C1 and LCAT (lecithin-cholesterol acyltransferase) was even more accurate in discriminating between NEC and LOS patients at the time of diagnosis (AUC = 0.999; 95 % CI = 0.987–1). With an AUC = 0.784 CRP as a solitary biomarker was reported to be not sufficient for distinguishing NEC from LOS patients. Notably, despite numerous overlapping findings as for the lipid metabolism in infection/inflammation, Chatziioannou *et al.* (2018) did not observe the Apo-C2 (described by Ng *et al.*, 2010) to be a significant marker.

The aforementioned studies from the NEC & LOS complex also underline the potential of multi-biomarker panels (*cf.* Ch. 1.2.3). In their review of infection and inflammation biomarkers in febrile children Zandstra *et al.* (2021) emphasize the progress that has indeed been made with respect to transcriptomic, proteomic or flow-cytometric multi-biomarker panels. Different procedures from all of these dynamic fields bear great promise to facilitate clinical decision making through novel laboratory or even **point-of-care tests (POCT)**. There are already POCTs on the market which support the distinction between bacterial and viral infections, a daily issue *e.g.* in ambulatory health care. An example is the ImmunoXpert™ triple assay, which evaluates the plasma proteins TRAIL (TNF-related apoptosis-inducing ligand), IP-10 (IFN-γ-induced protein-10) and CRP (Oved *et al.*, 2015; van Houten *et al.*, 2017; Srugo *et al.*, 2017), or the FebriDx® double assay, which measures MxA (myxovirus resistance protein A) and CRP (Self *et al.*, 2017; Shapiro *et al.*, 2018; Shirley, 2019). A third example is a novel POCT measuring a two-transcript host RNA signature out of EMR1-ADGRE1 and IFI44L, enabling an on-site transcriptomics-based distinction between a bacterial and viral genesis of infection (Pennisi *et al.*, 2021).

Besides this multitude of proteomics-based clinical observation studies and first established diagnostic concepts, two exemplary animal studies of recent time ought to be mentioned at this point since they may have high relevance for neonatology and the understanding of sepsis.

Muk *et al.* (2019) investigated plasma and cerebrospinal fluid alterations in preterm piglets, resulting from a systemic infection with *S. epidermidis* immediately after caesarean section (n = 15 vs. control n = 11). The plasma proteomes of 6, 12 & 24 h after inoculation showed elevated protein levels and significant dynamics for HP, lipopolysaccharide binding protein (LBP), von

Willebrand factor (vWF), Apo-A4, SERPIN A3-6, soluble CD14 (sCD14) and mannose-binding lectin 1 (MBL1). The cerebrospinal fluid proteomes of 24 h after inoculation showed elevations of neuroinflammation-related neuropeptide Y (NPY), interleukin 18 (IL-18) and matrix metalloproteinase 14 (MMP14). In a third group of piglets (n = 14), all infected with *S. epidermidis* but also enterally fed with bovine colostrum, the described alterations in plasma were reduced. This dampening effect, however, was much weaker in the cerebrospinal fluid. All in all, the observed proteome alterations elucidate the pathophysiology of premature organisms in sepsis.

Malmstrom *et al.* (2016) presented a detailed MS-acquired tissue protein atlas for a mouse model. With the help of such an extensive **reference proteome** the massive changes in blood plasma composition during sepsis may be more comprehensible. Generally, organ damage leads to an increase of so-called **tissue leakage proteins** in plasma. Although not having an original functional role in plasma they are yet well suited as indicators for the underlying organ damage (*e.g.* AST and ALT in hepatic injury *etc.*). For a proteomic investigation of sepsis, a whole tissue protein atlas may be particularly useful since sepsis –"*a life-threatening organ dysfunction caused by a dysregulated host response to infection*" (Singer *et al.*, 2016) – is characterized by pronounced tissue damage. Provided that the generation of such reference proteomes for humans is feasible, their clinical application in monitoring sepsis and organ injury is conceivable.

Altogether, the presented proteomic knowledge about infection, inflammation, and sepsis in neonates is promising, even though the way to clinical application may still be long. MS-based proteomics can provide deep insights into the perinatal period's complex physiology. In order to do so, it only requires minimal sample amounts (some dozens of µg), which is especially advantageous with neonatal patients. Therefore, the further investigation of the neonatal proteome is most worthwhile and the present work is intended to contribute to this.

1.2.5   Mass spectrometry

Mass spectrometry (MS) has long been an indispensable tool in many fields, particularly in the entire domain of life sciences. The connection between proteomics and MS, however, appears to be so close today that any engagement with proteomics is almost "incomplete" without some basic knowledge of MS. Therefore, some major principles of MS shall be briefly outlined here, while the appendix contains a more systematic introduction to the topic, which resulted from my personal interest in the technology.

So, what is mass spectrometry? Basically, MS is not one isolated method, but rather a whole family of approaches, techniques and devices, grown over the 20th century. Their common goal is identifying, quantifying and characterizing unknown components of any chemical mixture, be it an elemental or a molecular one, be it a pure one or a complex mix of compounds as in proteomics (Gross, 2017).

MS ultimately achieves these goals by determining the mass of molecules. While masses, which are reasonably large (from a human perspective), can be classified by observing their response to gravity, this is not possible with the minimal masses of peptides and proteins. Therefore, MS takes a different physical approach. Tiny amounts of the desired analyte are transferred to gas phase as ions and then exposed to arbitrarily applied electromagnetic forces in a vacuum environment. Observing their response to these electromagnetic forces replaces the evaluation of their response to gravity (Savaryn *et al.*, 2016). In the controlled surroundings of a mass spectrometer, ions in gas phase behave in strict dependence of their mass and charge. Hence, the **mass-to-charge ratio** (m/z) can be used to identify and finally quantify the ion in question (Gross, 2017).

What, most of the time, is simply called "a mass spectrometer" is in fact a multi-modular setup unifying several analysis methods. At present, one of the leading approaches in proteomics (and the approach of this thesis) is **LC-MS/MS**. It is a combination of liquid chromatography (LC) and tandem mass spectrometry (MS/MS). In such a setup, a prepared mixture of cleaved proteins is first separated in an LC column into its single components. In **chromatography**, all constituents of a mixture show an individual affinity to the material of the separation module, and therefore leave it in a certain order (according to their retention time). By applying a voltage at the end of the LC column (**electrospray ionization**) the separated analyte is transformed into an ionized aerosol. These ions in gas phase finally enter the electric field of the actual mass spectrometer, in which they will first be focused to a concentrated beam, purified from adducts and neutrons and then filtered according to the user-defined mode of analysis. In devices that belong to the realm of **tandem mass spectrometry**, the remaining peptides can (optionally) still be fragmented before meeting the actual analysis module, which determines the m/z. This ability to analyze each peptide species on multiple levels of detail is a pivotal strength of the approach. A peptide can once be detected in its original shape and once as its fragments, providing additional, almost fingerprint-like information for its identification. By combining all given information (retention time, mode of analysis, intensity, m/z of each level etc.), specialized bioinformatics software can finally infer from fragments to peptides and from peptides to proteins. Ultimately, **label-free quantification** is a widespread computational approach to quantify the measured intensities without any external standards and only in relation to the present measurements, providing a calculable quantity for biostatistical downstream analyses.

1.2.6 Multiple hypothesis testing

Multiple hypothesis testing is a crucial challenge in proteomics (as laid out in Ch. 1.2.2) and statistical precautions have to be taken to prevent misleading conclusions. Such are often based on the avoidance of false-positive test results by adjusting the **false discovery rate (FDR)**. Three of these correction mechanisms (all applied in the analyses of this thesis)

will be outlined in the following: the Benjamini-Hochberg corrected FDR, the Permutation-based FDR and the SAM algorithm. At last, the theory behind the 1D annotation enrichment analysis (1D-AEA) will be introduced.

Benjamini-Hochberg corrected FDR   Here, an adjusted significance level is individually applied on each p-value of a test series. In a first step, all calculated p-values of the data set are sorted and ranked in ascending order. The new significance threshold for any single p-value derives from its position in this ranking (*cf.* Fig. 1.3). Generally, the Benjamini-Hochberg correction is a rather conservative approach tending to generate false-negatives. A normal distribution of data is required.

p-values of all tests ($n = 10$)
ranked in ascending order (i)

original significance level $\alpha = 0.01$

adjusted significance level: $p\text{-}value_i \overset{!}{\leq} \dfrac{i}{n}\alpha$

Rank i	p-value	adj. signif. level	Reject H$_0$?
1	0.0001	0.001	yes
2	0.001	0.002	yes
3	0.002	0.003	yes
4	0.003	0.004	yes
5	0.005	0.005	yes
6	0.008	0.006	no
7	0.010	0.007	no
8	0.017	0.008	no
9	0.038	0.009	no
10	0.050	0.010	no

Figure 1.3: Benjamini-Hochberg correction. Formula for adjusted significance level embedded in an exemplary data set of ten p-values. As it can be seen, two H$_0$ hypotheses less are rejected than without a Benjamini-Hochberg correction. Figure adapted from Cox (2018).

Permutation-based FDR   In this concept, the FDR is estimated by performing randomized tests with the two groups in question (*cf.* Fig. 1.4). The procedure is repeated up to thousands of times. For each of these tests group 1 & 2 are evenly mixed up, creating new groups with equal numbers of group 1 & 2 samples. Thus, the natural "stimulus-control experiment" is intentionally altered to an artificial "control-control experiment", which can be expected to satisfy the H$_0$ hypothesis. Yet occurring significances are to be regarded as false-positives and their distribution is used for an estimation of the FDR. In contrast to the rather conservative Benjamini-Hochberg correction, the permutation-based approach is an approximation to the true FDR with less false-negatives. A further advantage is that it relies on the actual distribution of the available data and does not require a normal distribution.
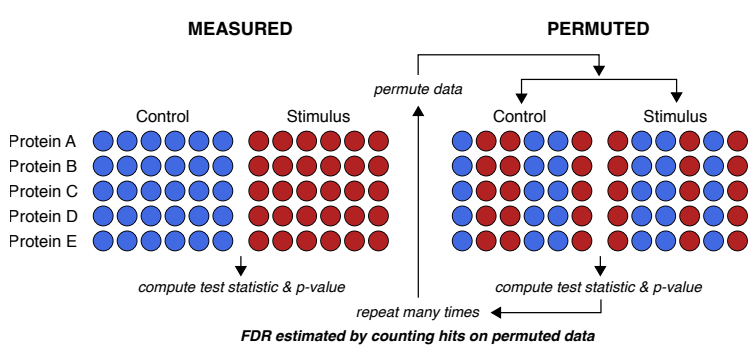

Figure 1.4: Permutation-based FDR. The FDR is estimated by repeated testing of intentionally permuted groups. Figure adapted from Cox (2018).

SAM algorithm   The permutation-based FDR can be combined with the so-called significance analysis of microarrays (SAM, Tusher *et al.*, 2001). It takes account of the plausible

consideration that reproducibility and effect size of any finding should always be perceived in context. The one might outweigh the other in a relevant manner. For example: A result with high reproducibility (high p-value) but small effect size (low fold change) should of course be seen as biologically less relevant than another result with equally high reproducibility and big effect size. On the other hand, findings with very high reproducibility should not be too easily rejected even when their effect size is relatively small. The SAM algorithm enables such a dynamic discrimination of findings by considering their p-values not only in relation to the FDR but also to their fold change.

$$ d(i) = \frac{\bar{x}_I(i) - \bar{x}_U(i)}{s(i) + s_0} $$

$$ s(i) = \sqrt{a\left\{ \sum_m [x_m(i) - \bar{x}_I(i)]^2 + \sum_n [x_n(i) - \bar{x}_U(i)]^2 \right\}} $$

Figure 1.5: SAM algorithm. The SAM algorithm is an adjustment of the two-sample t-test by the s_0 parameter, an artificial background variability (Tusher *et al.*, 2001). Its effect is illustrated in Fig. 1.6. $d(i)$ is the difference of gene expression/protein abundance, $\bar{x}_I(i)$ and $\bar{x}_U(i)$ are the average levels of gene expression/protein abundance for gene/protein (i) in two different states I and U. $s(i)$ represents the standard deviation of these measurements. $a = (1/n_1 + 1/n_2)/(n_1 + n_2 - 2)$ with n_1 and n_2 as the count of measurements of I and U. s_0 can be any positive rational value.

1D annotation enrichment analysis Besides the omnipresent problem of multiple hypothesis testing, it is challenging not to lose track of the bigger picture and to identify systematic trends in proteomic data. Naturally, findings of interest may not only occur on single protein level but on a functional level as well. Single proteins which are significantly altered between two conditions may indeed come into question as meaningful markers. However, no less interesting are (even small) consistent changes across a whole functional group of proteins, *i.e.* proteins with similar regulation, properties or function.
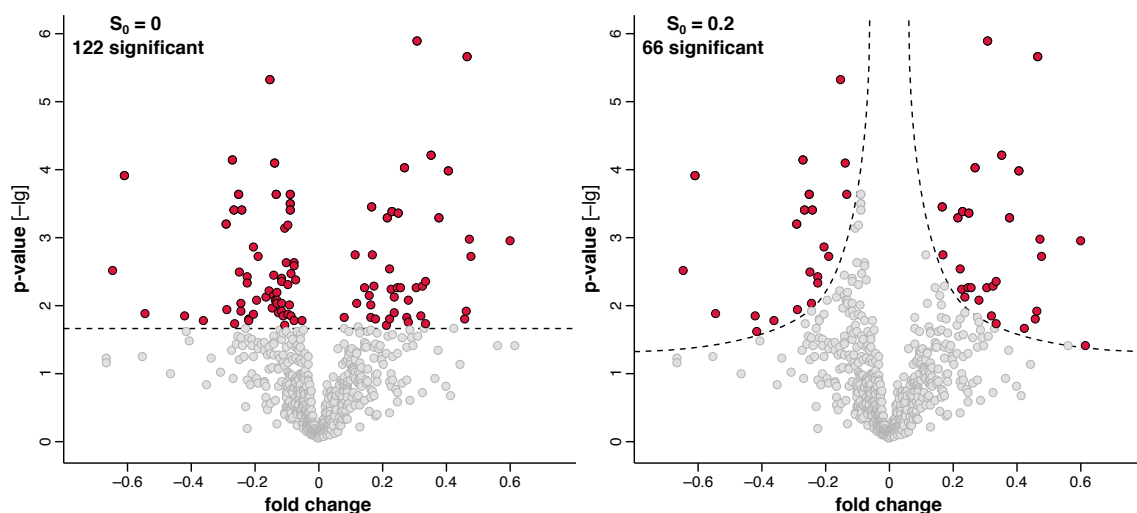
Figure 1.6: Effect of SAM. In an exemplary volcano plot with no SAM-correction ($s_0 = 0$, adjustment only by permutation-based FDR) a high number of proteins (all above the horizontal threshold) is declared significant. The same volcano plot with SAM-correction ($s_0 = 0.2$) dynamically assesses p-values in relation to the fold change and declares clearly less proteins significant (all above the curved threshold). Theoretically, $s_0 = \infty$ would completely exclude the weight of p-values in this figure and induce a vertical cut-off with respect to the fold change, only. Figure adapted from Cox (2018).

An important tool for uncovering such trends in proteomic data sets are annotation terms. Freely accessible online resources (such as geneontology.org or uniprot.org) provide whole libraries of such keywords (or "tags") reflecting the up-to-date knowledge about most proteins in a very condensed form. Hence, not only the mere quantity of a certain protein may be subjected to statistical tests, but also the associated annotation terms.

Such a statistical concept is the 1D annotation enrichment analysis (1D-AEA, Cox & Mann, 2012), which plays a role for the results of this thesis, too. In its core, this test is based on the comparison of two distributions: the overall distribution of all available annotation terms versus the distribution of one annotation term of interest. Technically, any numerical value assigned to each protein (or gene *etc.*) of the data set can serve as input for the x-axis (in proteomics most commonly a form of protein abundance). The y-axis displays the mere counts of
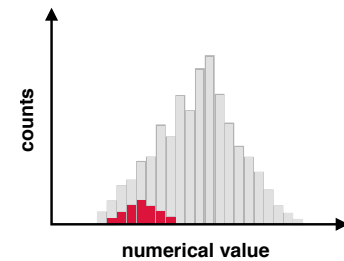
Figure 1.7: 1D-AEA Basic principle of the 1D-AEA is the comparison of two distributions.

annotation terms being present at the respective numerical value. Thereby, the overall distribution of terms can be opposed to the distribution of the examined term (*cf.* Fig. 1.7). The more the centres of both distributions differ, the more one can assume a differential regulation of those proteins annotated with the term in question. The 1D-AEA algorithm reflects this deviation by a "score" between –1 and +1. Negative values indicate a deviation towards the lower end of the overall distribution, positive ones indicate the opposite. Usually, a 1D-AEA is not performed for selected annotation terms but rather in the form of multiple hypothesis testing for all available annotation terms.

2. OBJECTIVES

Neonatal infections are a major challenge in paediatrics. The prevention of life-threatening disease progressions necessitates very early and accurate diagnostics as well as targeted therapies. The *status quo* of diagnostic means, however, clearly leaves room for improvement. The exploration of the blood proteome, an insightful mirror of cellular processes and human physiology, may be helpful in the search for novel inflammation markers or other diagnostic approaches with the required predictive accuracy.

The overall objective of this thesis is to deliver new insights into infection and inflammation in newborns at protein level and to contribute to improvements of the respective diagnostic resources in the long term. For that reason, the whole blood proteomes of newborns with early-onset bacterial infection were compared to those of healthy newborns.

The following questions will be addressed in detail (*cf.* graphical abstract, p. 85)**:**

1. Are there significant **differences** between the whole blood proteomes of neonates ($\geq$ 35+0) *with* and *without* early-onset infection?

2. Are there significant **dynamics** in the whole blood proteomes of neonates ($\geq$ 35+0) with prolonged early-onset infection in the early phase of their antibiotic treatment?

3. Are there significant features for a **distinction** between the whole blood proteomes of neonates ($\geq$ 35+0) with early-onset infection whose antibiotic treatment lasted *at least* 96 h or *at most* 72 h, respectively?

---

3. METHODS

This thesis deals with a subset of proteomic data originating from a cooperation between the Dr. von Hauner Children's Hospital Munich (LMU) and the Max Planck Institute of Biochemistry Martinsried (MPIB). From February 2017 to June 2019, in three study phases 501 term and 161 preterm infants were included in a prospective cohort study at the LMU perinatal centre of Campus Innenstadt in downtown Munich. Descriptive data on this cohort were published as Munich Preterm and Term Clinical study (MUNICH-PreTCl) by Pangratz-Fuehrer *et al.* (2021). When the entire proteomic data set will be analysed and published, it will presumably be the largest proteome study in paediatrics to date. Its overall objective is to enhance the general understanding of the neonatal blood proteome and its diverse perinatal influencing factors.

The present proteomic data on neonatal infections, however, were obtained from a **pilot study** within the MUNICH-PreTCl cohort comprising a subset of patients from study phase 2. This thesis is centred on the whole blood proteome analyses of a group ($n = 20$) of **late preterm** ($\geq$ 35+0) and **term infants** affected by (non-culture-proven) **early-onset bacterial infection** (onset $\leq$ 36 h *pp.*).

3.1 Ethics committee vote

The conduct of the cohort study and its proteome analyses is covered by an approval of the Ethics Committee of the LMU (approved addendum to request 735-15). The principles of the Helsinki Declaration in its revision of Somerset West (1996) were followed. The study is registered in the German Clinical Trials Register (DRKS 00024189). Strictly confidential handling of any individual-related data has been ensured.

3.2 Participation criteria

The following conditions were the applied **inclusion criteria**:
- Birth at LMU perinatal centre of Campus Innenstadt in study period
- Explanatory meeting with at least one parent and some hours of consideration time
- Written informed consent of at least one parent (*cf.* additional materials in appendix)

The following conditions had been previously defined as **exclusion criteria**:
- Severe congenital anomalies of the newborn
- Hereditary immunodeficiency of the newborn
- Infectious viral diseases of the mother (HIV, Hep-B, Hep-C)

The parents could withdraw their consent verbally at any time. It was explicitly pointed out that any blood specimens would only be taken in the course of clinical routine blood sampling (with minimal extra amounts), that there would not be any form of experimental intervention, and that neither a medical risk nor an individual advantage would result from study participation.

3.3 Study phases

MUNICH-PreTCl was conducted in three stages (*cf.* Fig. 3.1):
- **Stage 1**: initial phase – eligible: preterm only – establishing of workflow, data collection & adapted proteomic analysis methods – no questionnaires
- **Stage 2**: main phase – eligible: preterm & term – study questionnaires – **pilot study**
- **Stage 3**: closing phase – eligible: gestational age $< 30+0$, birth weight $< 10^{\text{th}}$ percentile, newborn infection, maternal (gestational) diabetes – study questionnaires
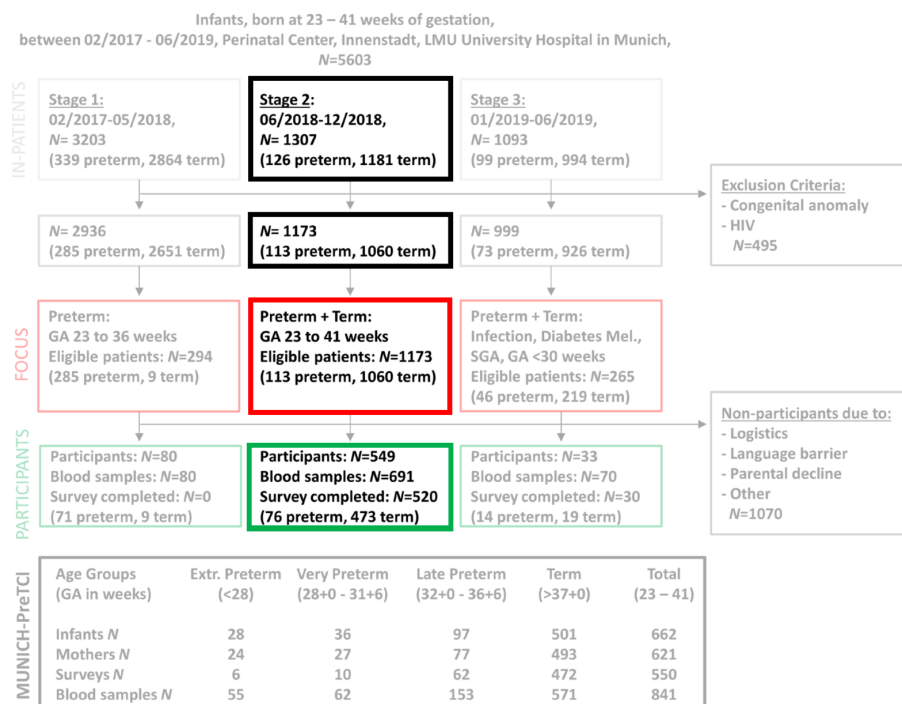

Figure 3.1: Flow chart MUNICH-PreTCl study. This thesis is based on a reduced patient collective from stage 2 (pilot study). Figure adapted from Pangratz-Fuehrer *et al.* (2021).

3.4   Pilot study

After completion of stage 2 (June to December 2018), a pilot study with first MS runs was conducted on a **reduced patient collective** (115 out of 549 stage 2 patients, *cf.* Fig. 3.1). It comprises three groups with particular characteristics (*cf.* Tab. 3.1) and a control group of term infants. All patients with neonatal infections were primarily included in the pilot study ($n = 26$). However, six of them had to be excluded from this thesis ($n = 20$) for their too low gestational age ($< 35+0$). Four late preterm infants (35+0 to 36+6) without neonatal infection were added to the control group of this thesis ($n = 30 + 4$) as counterparts for the late preterm infants in the infection group.

Table 3.1: Pilot study groups & thesis patients

Group	n_{pilot} / n_{thesis}	Samples	Description
Preterm	29 / 4	2–5	25 preterm infants $< 35+0$, 4 late preterm infants $\geq 35+0$
Infection	26 / 20	2	6 preterm infants $< 35+0$, 5 late preterm infants $\geq 35+0$, 15 term infants
Allergy	30 / 0	1	term infants, no intrapartum complications, parental allergic conditions
Control	30 / 30	1	term infants, no intrapartum complications, no major parental conditions (allergies, diabetes, hereditary diseases or others)

n_{pilot}/n_{thesis} = number of pilot study patients included in the analyses of this thesis. **Samples** = number of acquired specimens for each patient. **Description** = Parental conditions were assessed according to study questionnaire.

3.5   Thesis concept

Within the superior study protocol of MUNICH-PreTCl the analyses of this thesis had been designed as follows (*cf.* graphical abstract, p. 85):

- **Eligible patients:**
    - gestational age $\geq 35+0$ (late preterm & term infants)
    - onset of symptoms $\leq 72\,$h *pp.* (early-onset infection/sepsis)
- **Blood samples infection group:**
    - prior to 1st administration of antibiotics ("I1")
    - ca. 48 h after 1st administration of antibiotics ("I2")
- **Blood samples control group:**
    - 36–72 h *pp.* as part of standardized newborn screening ("C")

Table 3.2: Patient collective

Group	Total	Term	Late pr.	Samples
Infection	20	15	5	2
Control	34	30	4	1

Three out of four preterm control patients were sampled more than once, resulting in 38 control samples (*cf.* Ch. 4.3.3, p. 58)

3.6   Patient collective

Under the conditions above two study groups (as depicted in Tab. 3.1 and Tab. 3.2) could be formed from the pilot study.

Descriptively, for all cases in the infection group applies:

- Onset of symptoms and start of antibiotic therapy was $\leq$ **36 h** *pp.* ("early-onset")
- Blood cultures remained **sterile** ("non-culture-proven" or "clinical" infection)

3.7 Clinical data

An extensive collection of clinical and anamnestic data had been performed for MUNICH-PreTCl, also by using an optional study questionnaire distributed to the majority of participating parents (stage 2 & 3, *cf.* Ch. 3.3). From all available data only the following (mainly numeric) data points were used for this thesis:

- **Neonatal data:** sex, gestational age, APGAR score, birth weight, body length, head circumference (*cf.* Tab. 4.1)
- **Clinical data:** time of blood sampling, duration and type of antibiotic treatment; CRP, IL-6, WBC & PC at diagnosis of infection (*cf.* Tab. 4.1)
- **Maternal data:** age, BMI at 1st gyn. examination of pregnancy; durations of rupture of membranes, delivery and nicotine consumption prior to pregnancy (*cf.* Fig. 4.7)

3.8 Blood samples

Blood samples were only taken in the course of clinical routine by the responsible medical personnel. No specimen was solely taken for study purposes. Whole blood was conserved on Whatman® 903 Protein Saver Snap-apart Cards (Cytiva). One of four circles was filled with one or two drops of blood (according to manufacturer ca. 80 µl).

Two varieties of **whole blood** were collected:

- **Capillary** blood
- **Venous** blood

Immediately after blood taking all filter cards were:

1. **Pseudonymized**
2. **Dried** at room temperature over **24 h**
3. **Conserved** at **–80 °C**

Figure 3.2: Filter card. Since the 1960s, cotton fibre cards are established in screening for congenital metabolic diseases. Due to easy handling of minimal blood volumes and sample stability without refrigeration they are advantageous in scientific settings as well. Photo from Sheler (2007).

3.9 Mass spectrometry

The underlying mass spectrometric workflow belongs to the domain of **bottom-up proteomics** (proteolytic digestion prior to mass spectrometry, *cf.* "Basics of mass spectrometry" in the appendix). It is largely based on the methods described by Geyer *et al.* (2016a). Our work-

ing group's detailed laboratory protocol is described in the "Additional materials" in the appendix.

In brief: 14 Gauge punches were taken from each dried blood spot. After dissolving the contained whole blood, buffering, thermic denaturation, tryptic proteolysis and several steps of peptide purification, samples were measured with a mass spectrometric setup (LC-MS/MS) including the following key instruments:

1. Liquid chromatography – Evosep One (Evosep)
2. Electrospray ionizer – Nanospray ion source (Thermo Fisher Scientific)
3. Hybrid mass spectrometer – Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap™ mass spectrometry system (Thermo Fisher Scientific)

All mass spectra were acquired by **data-independent acquisition (DIA)** and quantified by **label-free quantification (LFQ)**.

3.10 Software

Spectronaut™   Version 12.0.20491.12.30484. For primary analyses of detected mass spectra (peptide identification, protein inference, label-free quantification, *cf.* Ch. B.6).

Perseus   Version 1.6.2.3. For secondary biostatistical analyses, especially in omics technologies. An in-house development of the cooperating MPIB (Tyanova *et al.*, 2016).

Python™   For the only paired t-test of this thesis (Fig. 4.11, p. 54). Data afterwards transferred into *Perseus* for further analyses and visualizations.

Microsoft Excel®   For collection of clinical raw data and descriptive statistics of patient groups.

3.11 Biostatistics

Normal distribution   Logarithmic normal distribution of protein abundances is assumed (Zubarev, 2013).

Decadic logarithm   The notation of decadic logarithms follows $\log_{10} x = \lg x$. All indicated LFQs are meant as $\lg(LFQ)$.

Valid values   Only proteins detectable in $\geq 70\,\%$ of patients in at least one of two groups were included in comparative statistical tests.

Unpaired t-test   Mean LFQs of any included protein in two unrelated groups were compared by using unpaired (two-sided) two-sample t-tests. The chosen form of visualization is the volcano plot. It displays the difference of mean LFQs in both groups ("fold change") on the x-axis and the p-value of the respective t-test on the y-axis.

Paired t**-test**   Mean LFQs of any included protein in two related groups were compared by using paired (two-sided) two-sample t-tests. The kind of visualization is principally the same as described for the unpaired t-test, however, displaying the mean of paired LFQ differences between both groups on the x-axis and the p-value of the respective t-test on the y-axis.

Pearson correlation   Between certain proteins of interest (identified in previous t-tests) and selected clinical or anamnestic parameters (*cf.* Ch. 3.7) a Pearson correlation (r) was conducted, visualized by scatter plots or a heatmap.

1D annotation enrichment analyses   Proteomic data was standardly annotated with the following sets of terms:

- *GOBP name* (Gene Ontology Biological Processes, full term)
- *GOMF name* (Gene Ontology Molecular Functions, full term)
- *GOCC name* (Gene Ontology Cellular Compartments, full term)
- *KEGG name* (Kyoto Encyclopedia of Genes and Genomes, full term)
- *UP keywords* (UniProt keywords to structures, functions and pathways)

Enrichment analyses only refer to these five classes of annotations. The chosen form of visualization for the 1D-AEA is the histogram.

Significance levels   With respect to the problem of multiple hypothesis testing (*cf.* Ch. 1.2.6), the following applies:

1. Volcano plots were generated with **250 permutations**, **FDR = 0.05** and s_0 **= 0.1** (default settings *Perseus*). Data points above the resulting significance curves are called "significant". Data points below with a **p-value** $\leq$ **0.01** ($-\lg p \geq 2$) are reported.

2. In the only volcano plot based on paired t-tests (Fig. 4.11, p. 54) data points with a **q-value** $\leq$ **0.01** are called "significant" ($-\lg q \geq 2$). Data points with a higher q-value but with a **p-value** $\leq$ **0.01** ($-\lg p \geq 2$) are reported.

3. In 1D-AEA only annotation terms with a Benjamini-Hochberg **FDR** $<$ **0.02** are called "significant" and are reported (default settings *Perseus*).

4. Pearson correlations with a **p-value** $\leq$ **0.05** are reported.

4. RESULTS

4.1 Patient collective and groupings

The patient collective of this thesis (*cf.* Ch. 3.4 & 3.6) comprises 20 infants with a (non-culture-proven) early-onset bacterial infection (onset of symptoms and of antibiotics in all cases $\leq$ 36 h *pp.*, no detection of causative pathogens by blood culture). 15 patients were term infants ($\geq$37+0), 5 were late preterm infants (35+0 to 36+9). 18 patients were male, 2 were female. The mean birthweight was 3335 g (SD 558 g, *cf.* Tab. 4.2). The median APGAR scores of 8, 10 & 10 (for 1, 5 & 10 min *pp.*) suggest mainly uncomplicated births and primarily uncomplicated adaptations.

16 out of 20 patients were treated with an identical scheme of antibiotic therapy, which was ampicillin and cefotaxime in a dosage of 150 mg/kg/day each (*i.v.*, intervals of 8 h). 2 patients were temporarily treated with the double dosage of ampicillin. 2 patients were treated with other antibiotics: in one case with meropenem in a dosage of 100 mg/kg/day (*i.v.*, intervals of 8 h), in the other case with piperacillin/tazobactam in a dosage of 300 mg/kg/day (*i.v.*, intervals of 8 h, *cf.* Fig. 4.2).

Furthermore, I divided the homogeneous main group of 16 patients according to the duration of antibiotic therapy. 7 of them had been treated for max. 72 h, 9 of them for min. 96 h. This approach enabled separate analyses of the severely diseased subgroup (with prolonged treatment) and a comparison of both subgroups.

The control group of this thesis (*cf.* Ch. 3.4 & 3.6) comprises 34 infants without infectious disease, complications during birth or major parental conditions (such as allergies, diabetes, hereditary diseases or others). 30 patients were term infants ($\geq$ 37+0), 4 were late preterm infants (35+0 to 36+9). 16 patients were male, 18 were female. The mean birthweight was 3215 g (SD 577 g, *cf.* Tab. 4.2). The median APGAR scores were 9, 10 & 10 (for 1, 5 & 10 min *pp.*).

Out of this major control group, I formed 3 minor control groups ("C-I1(all), C-I1(72) & C-I1(96)") for the infection group and both of its subgroups ("I1(all), I1(72) & I1(96)", *cf.* Tab. 4.1 & Tab. 4.2). The control groups were composed in such a way that equal numbers of patients with a maximum of similarity in 3 selected key characteristics were matched to each other

(matched pairs). These key characteristics were (in descending order of importance): cGA at the time of blood sampling, birth weight and sex. With respect to cGA and birth weight, appropriately comparable distributions between the matched groups could be ensured (*cf.* Tab. 4.2) & the additional figures in the appendix). Due to the higher priority of the cGA, however, the numbers of females and males in the paired groups are unequal. Generally, there are significantly more male than female newborns in the infection groups while the control groups are slightly more balanced, in this regard. As a result of clinical practice, the blood samples of the infection groups are mainly of venous origin while those of the control groups are mainly of capillary origin. Furthermore, the infection groups are characterized by higher maternal ages, higher maternal BMIs, longer intervals between rupture of membranes and birth, and longer durations of delivery. Comprehensibly, the initial levels of CRP and IL-6 were higher in I1(96) than in I1(72), while for the abundance of platelets the opposite was true (*cf.* Tab. 4.2).

Only for one specific analysis (*cf.* Ch. 4.3.3), I formed 2 equally sized non-matched control groups with 19 blood samples each ("Cnm-I1(96) & Cnm-I2(96)"). This was possible since 3 out of 4 late preterm control patients were represented with more than one blood sample, resulting in 38 control samples (*cf.* Tab. 3.2). In this specific case, the available 38 samples were separated in two groups by their median time of blood taking to simulate a longitudinal control proteome.

In general, for each patient in the infection group two blood samples ("I1 & I2") had been collected, one before and one ca. 48 h after the first administration of antibiotics (median interval 45 h, range 37–85 h). From each patient in the control group one blood sample had been collected, normally at the respective newborn screening 36–72 h *pp.* (median 44 h *pp.*).

Tab. 4.1 displays the used abbreviations for all groups, Tab. 4.2 offers a more detailed descriptive statistics.
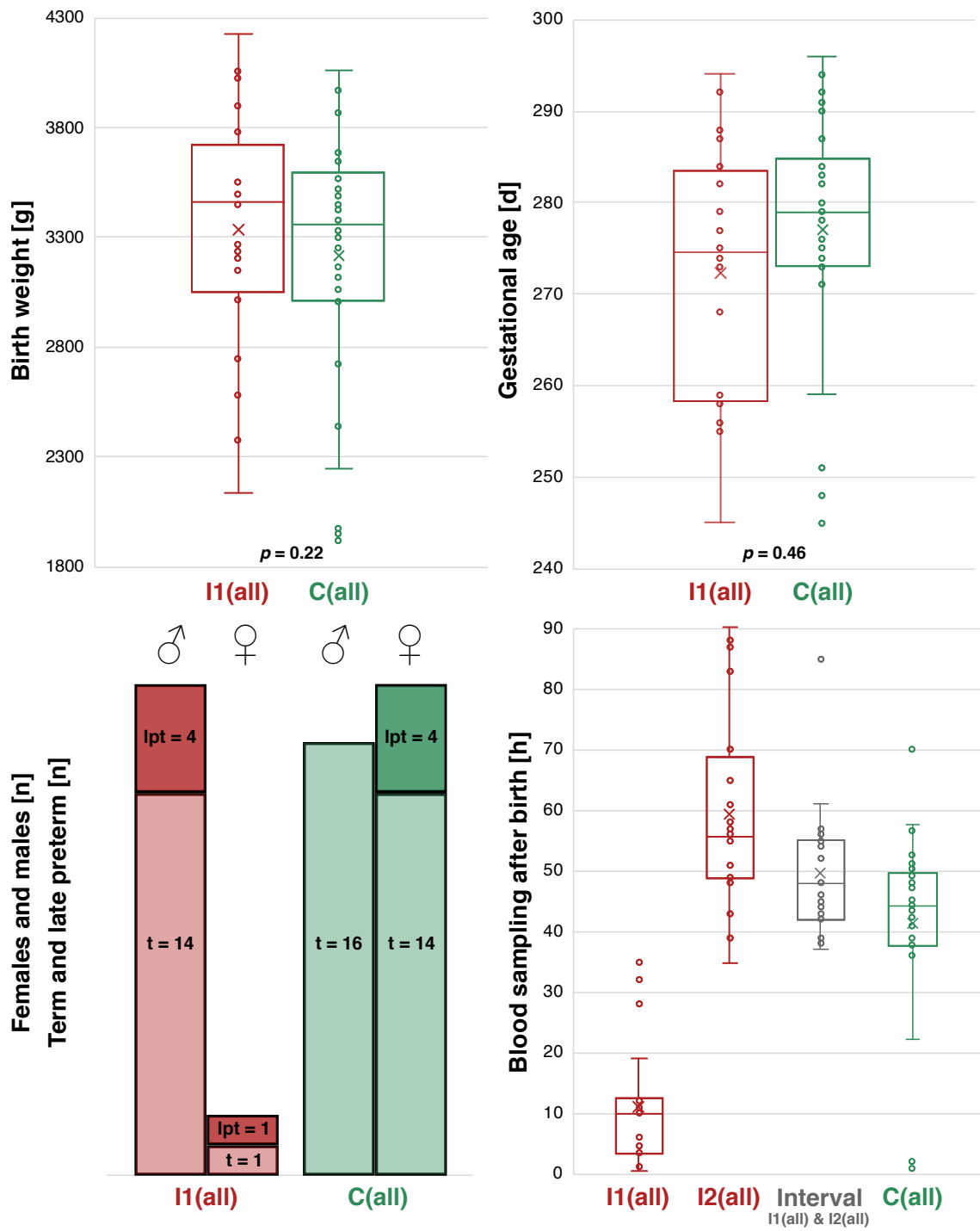
Figure 4.1: Patient collective (1). Comparison of all included patients with early-onset bacterial infection I1(all) with all control patients C(all) with respect to key characteristics. The p-values in the upper box plots refer to two-sided two-sample t-tests between the given groups. For detailed insight into the subgroups please see the additional figures in the appendix.
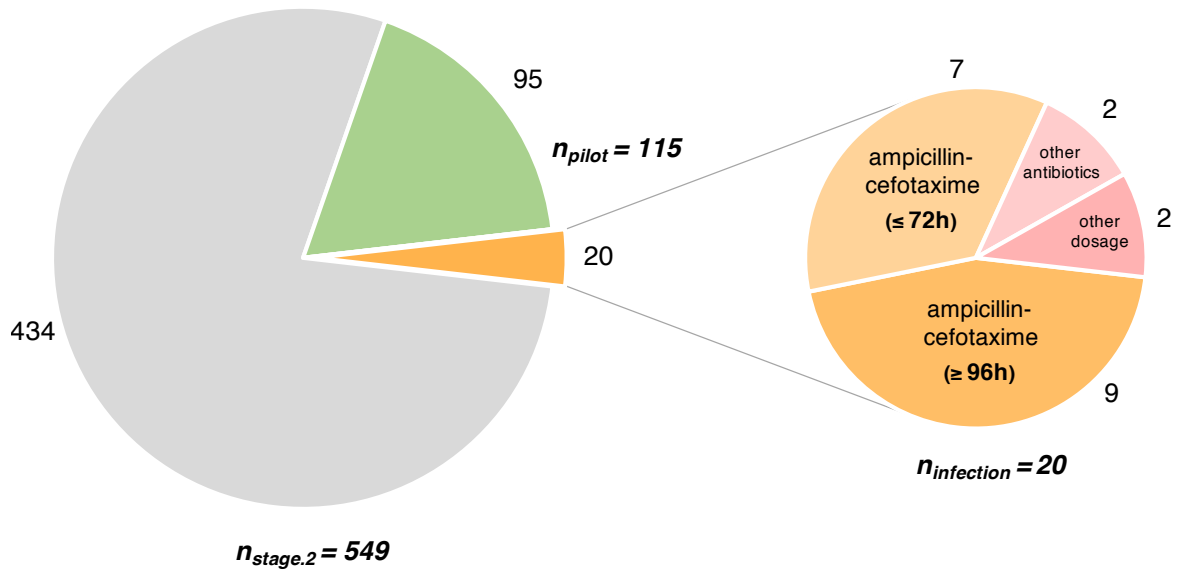
Figure 4.2: Patient collective (2). 115 out of 549 patients from stage 2 of the MUNICH-PreTCl study (*cf.* Fig. 3.1) were included in a pilot study. 20 of them received an antibiotic treatment for their (non-culture-proven) early-onset bacterial infection and were included with a pair of blood samples (before & 36–72 h after the first administration of antibiotics).

Table 4.1: Groupings

Abbreviation	Patients	Explanation
I1(all)	20	1st samples of all patients under *any* antibiotic treatment
I1(72)	7	1st samples of patients under ampicillin-cefotaxime $\leq$ 72 h
I1(96)	9	1st samples of patients under ampicillin-cefotaxime $\geq$ 96 h
I2(96)	9	2nd samples of patients under ampicillin-cefotaxime $\geq$ 96 h
C-I1(all)	20	matched control for I1(all)
C-I1(72)	7	matched control for I1(72)
C-I1(96)	9	matched control for I1(96)
Cnm-I1(96)	19	non-matched control for I1(96)
Cnm-I2(96)	19	non-matched control for I2(96)
C(all)	34	entire control patient collective (only for descriptive statistics)

Table 4.2: Descriptive statistics

	I1(all)	C-I1(all)	I1(72)	C-I1(72)	I1(96)	C-I1(96)	Cnm-I1(96)	Cnm-I2(96)	C(all)
Basics									
Patients, female vs. male | n, n	2, 18	8, 12	0, 7	3, 4	2, 7	3, 6	11, 8	11, 8	18, 16
Gestational age, weeks | mean (SD)	38.9 (2.0)	38.8 (2.0)	38.6 (2.6)	38.4 (2.6)	39.0 (1.6)	39.2 (1.0)	39.0 (2.3)	39.3 (2.3)	39.2 (1.9)
APGAR score 1 min | median (range)	8 (6, 10)	9 (6, 10)	9 (7, 10)	9 (8, 10)	8 (6, 10)	9 (8, 10)	9 (6, 10)	9 (8, 10)	9 (6, 10)
5 min | median (range)	10 (7, 10)	10 (10, 10)	10 (9, 10)	10 (10, 10)	9 (7, 10)	10 (9, 10)	10 (10, 10)	10 (9, 10)	10 (9, 10)
10 min | median (range)	10 (8, 10)	10 (9, 10)	10 (9, 10)	10 (10, 10)	10 (8, 10)	10 (10, 10)	10 (9, 10)	10 (9, 10)	10 (9, 10)
Blood sampling, cap. vs. ven. | n, n	4, 16	19, 1	0, 7	6, 1	1, 8	9, 0	14, 5	14, 5	29, 5
Blood sampling, h *pp.* | median (range)	10 (0, 35)	43 (1, 58)	6 (0, 35)	39 (2, 51)	11 (1, 13)	48 (36, 50)	39 (1, 44)	50 (45, 79)	44 (1, 70)
Duration antibiotics, h | median (range)	106 (38, 310)		64 (38, 72)		144 (104, 192)				
Maternal age, years | mean (SD)	35 (4.4)	33 (4.5)	33 (2.4)	32 (3.6)	36 (6.0)	32 (6.2)	33 (4.2)	34 (5.1)	33 (4.6)
Maternal BMI, n | mean (SD)	24.2 (4.6)	21.9 (4.5)	23.0 (4.9)	21.1 (2.9)	25.2 (5.2)	22.0 (1.9)	22.6 (3.1)	21.2 (1.6)	22.1 (2.5)
Rupture of membranes, h | mean (SD)	28.8 (73.5)	7.8 (12.0)	23.5 (4.9)	3.4 (2.8)	44.0 (105.2)	8.3 (15.6)	20.9 (55.4)	35.8 (70.0)	16.4 (40.4)
Duration of birth, h | mean (SD)	10.1 (8.1)	5.9 (5.6)	12.6 (9.1)	5.3 (4.3)	10.4 (8.4)	3.4 (4.8)	7.3 (6.0)	8.7 (6.7)	7.8 (6.4)
Percentiles (Fenton)									
Birth weight, g | mean (SD)	3335 (558)	3103 (640)	3484 (721)	3136 (849)	3310 (465)	3298 (474)	3048 (742)	3116 (613)	3215 (577)
percentile | mean (SD)	51 (26)	39 (26)	63 (27)	46 (36)	49 (24)	48 (28)	34 (26)	35 (25)	37 (25)
z-score | mean (SD)	0.00 (0.82)	-0.40 (0.91)	0.42 (0.82)	-0.16 (1.14)	-0.08 (0.72)	-0.17 (1.06)	-0.52 (0.92)	-0.48 (0.80)	-0.43 (0.87)
Weight length ratio, g/cm | mean (SD)	65.3 (7.7)	60.0 (12.5)	67.1 (9.0)	62.9 (13.5)	65.1 (7.0)	60.4 (14.2)	58.7 (13.5)	62.1 (9.4)	62 (11)
percentile | mean (SD)	47 (26)	38 (31)	63 (20)	49 (37)	40 (27)	44 (31)	35 (33)	37 (26)	39 (29)
z-score | mean (SD)	-0.05 (0.79)	-0.50 (1.18)	0.36 (0.55)	-0.07 (1.20)	-0.14 (0.79)	-0.43 (1.35)	-0.61 (1.22)	-0.45 (0.83)	-0.43 (1.04)
Body length, cm | mean (SD)	51 (3)	50 (4)	51 (5)	49 (4)	51 (2)	52 (1)	50 (4)	50 (3)	51 (3)
percentile | mean (SD)	57 (29)	55 (27)	62 (34)	48 (28)	59 (23)	72 (21)	54 (27)	49 (30)	53 (29)
z-score | mean (SD)	0.36 (1.05)	0.21 (0.93)	0.74 (1.35)	-0.02 (0.82)	0.26 (0.68)	0.68 (0.66)	0.20 (0.97)	-0.08 (1.09)	0.11 (1.07)
Head circumference, cm | mean (SD)	35 (2)	34 (2)	35 (2)	34 (2)	34 (1)	35 (1)	34 (2)	34 (2)	34 (2)
percentile | mean (SD)	56 (27)	55 (23)	71 (27)	59 (35)	52 (22)	66 (21)	44 (24)	46 (28)	47 (26)
z-score | mean (SD)	0.23 (-0.98)	0.17 (0.77)	0.85 (1.04)	0.36 (1.23)	0.03 (0.69)	0.54 (0.78)	-0.32 (1.02)	-0.25 (1.12)	-0.24 (1.10)
Diagnostics									
Hematocrit (cord blood), % | mean (SD)	49.7 (4.6)	50.1 (5.8)	48.4 (2.7)	47.5 (4.0)	49.5 (4.8)	51.1 (5.1)	50.3 (6.3)	50.8 (4.8)	50.9 (5.4)
Hemoglobin (cord blood), mg/dl | mean (SD)	16.2 (1.5)	16.4 (1.9)	15.8 (0.9)	15.5 (1.3)	16.2 (1.6)	16.7 (1.7)	16.4 (2.1)	16.6 (1.6)	16.6 (1.8)
pH (cord blood), n | mean (SD)	7.28 (0.08)	7.34 (0.07)	7.30 (0.09)	7.33 (0.07)	7.29 (0.08)	7.28 (0.08)	7.35 (0.08)	7.27 (0.08)	7.31 (0.09)
CRP, mg/dl | median (range)	2.7 (0.1, 6.6)		2.0 (0.1, 3.0)		3.0 (0.1, 4.7)				
IL-6, pg/dl | median (range)	413 (18, 5144)		199 (18, 530)		735 (192, 2387)				
WBC, G/L | median (range)	19.5 (7.5, 31.7)		19.5 (16.7, 26.1)		20.7 (13.2, 31.7)				
PC, G/L | median (range)	236 (113, 403)		270 (262, 403)		180 (113, 353)				

For detailed insight into the distributions of cGA, birth weight, and of females and males in the paired groups please see the additional figures in the appendix (p. 93). Please note that I1(96) and I2(96) are principally identical, however, the indicated number of capillary and venous samples, CRP, IL-6, white blood cell count (WBC) and platelet count (PC) in the I1(96) column refer to I1(96) only.

4.2 Question 1 – "Differences"

Are there significant **differences** between the whole blood proteomes of neonates ($\geq$ 35+0) *with* and *without* early-onset infection? (*cf.* Ch. 2. Objectives)

I approached this question in four steps, as depicted in Fig. 4.3. Firstly, I compared all available patients with differently severe early-onset infections and different types of treatment to a matched control group of healthy neonates. Secondly, I compared a smaller but more homogenous group of patients with severe early-onset infection and prolonged treatment (min. 96 h, ampicillin-cefotaxime only) to a matched control group. On the same group I performed an enrichment analysis in a third step, aiming to identify significantly changed protein patterns between both groups. Lastly, I correlated proteins of interest from the previous three steps with a number of selected clinical parameters.
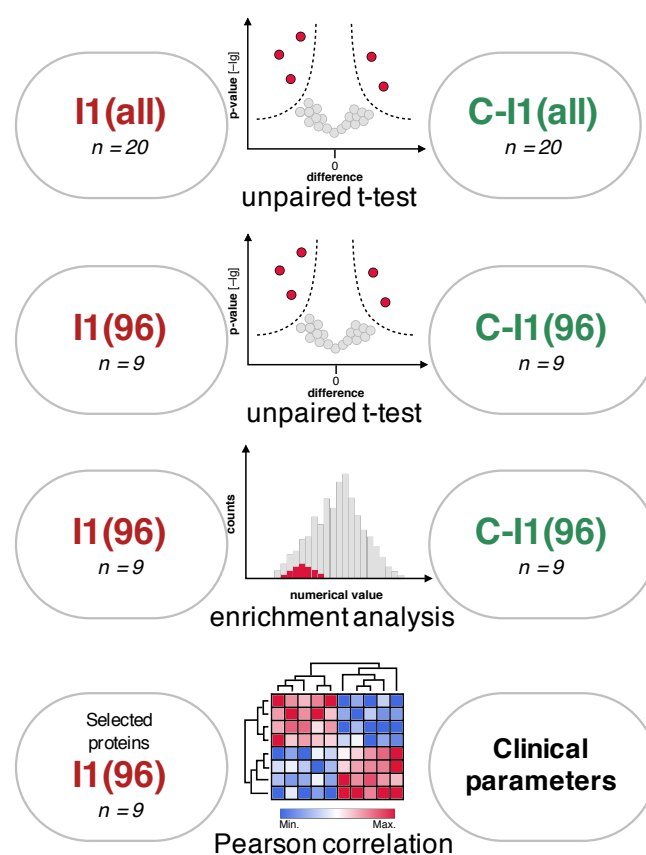
Figure 4.3: Graphical abstract (question 1)

4.2.1   I1(all) vs. C-I1(all)

The first (*i.e.* "pre-antibiotic") blood samples of all patients under antibiotic treatment of any kind and of any duration (I1(all), n = 20) were compared with a matched control (C-I1(all), n = 20). Mean time points of blood taking were 11 h *pp.* for I1(all) and 40 h for C-I1(all). An unpaired t-test for all included proteins (*cf.* Ch. 3.11) produces the results shown in Fig. 4.4. Of these, 24 proteins are differentially abundant by an absolute p-value of ≤ 0.01, but none of them is significant after s_0 correction ("all below significance curves"). Tab. 4.3 reports these 24 proteins in detail.
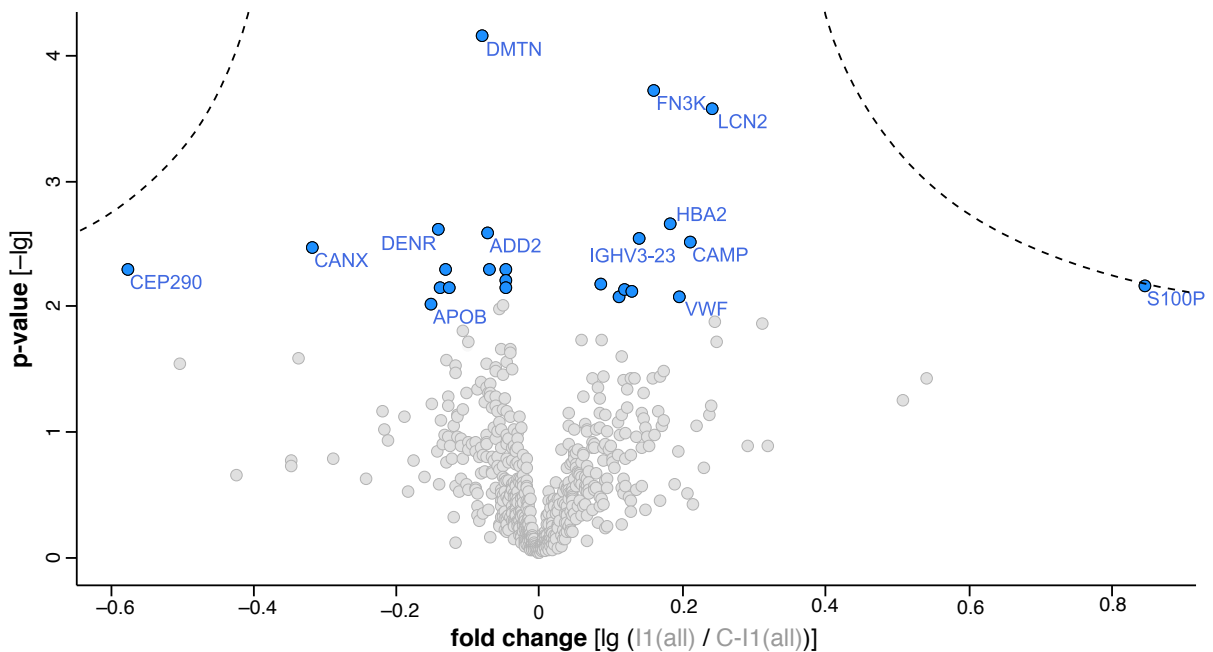
Figure 4.4: Volcano plot I1(all) vs. C-I1(all). x-axis = fold change ("difference of means": mean ($\log_{10}$) LFQ of protein a in 1st group minus mean ($\log_{10}$) LFQ of protein a in 2nd group; or as indicated: logarithmic ratio of mean absolute LFQs of protein a in 1st & 2nd group; *i.e.* pos. x-values reflect a higher abundance of protein a in 1st group and vice versa; absolute fold change of protein a equals $10^{x(a)}$). y-axis = p-value (for respective unpaired t-test on the fold change of protein a). Blue = abs. p-value ≤ 0.01. For further details see Tab. 4.3.

Table 4.3: Protein outliers volcano plot I1(all) vs. C-I1(all) (Fig. 4.4)

○	−lg *p*	Fold change	Gene name	UniProt ID	Protein description
●	4.16	−0.08 (0.83)	DMTN	Q08495	Dematin
●	3.73	0.16 (1.45)	FN3K	Q9H479	Fructosamine-3-kinase
●	3.58	0.24 (1.74)	LCN2	P80188	Isoform 2 of Neutrophil gelatinase-assoc. lipocalin
●	2.66	0.18 (1.51)	HBA2	G3V1N2	Isoform of Hemoglobin subunit alpha
●	2.62	−0.14 (0.72)	DENR	F8VVL1	Density-regulated protein
●	2.58	−0.07 (0.85)	ADD2	P35612	Beta-adducin
●	2.54	0.14 (1.38)	IGHV3-23	A0A0B4J2B7	Ig heavy chain V-III region 23
●	2.51	0.21 (1.62)	CAMP	J3KNB4	Cathelicidin antimicrobial peptide
●	2.47	−0.32 (0.48)	CANX	P27824	Calnexin
●	2.30	−0.05 (0.89)	CFL1	E9PK25	Cofilin-1
●	2.30	−0.13 (0.74)	GLIPR2	Q9H4G4	Golgi-assoc. plant pathogenesis-related protein 1
●	2.30	−0.57 (0.27)	CEP290	J3KNF5	Centrosomal protein of 290 kDa
●	2.29	−0.07 (0.85)	UBAC1	Q9BSL1	Ubiquitin-assoc. domain-containing protein 1
●	2.21	−0.05 (0.89)	ANK1	P16157	Ankyrin-1
●	2.18	0.09 (1.23)	VTN	P04004	Vitronectin
●	2.16	0.85 (7.08)	S100P	P25815	Protein S100-P
●	2.16	−0.05 (0.89)	SPTA1	P02549	Spectrin alpha chain
●	2.15	−0.14 (0.72)	H2AFJ	P04908	Histone H2A (different types)
●	2.15	−0.13 (0.74)	SRI	C9J0K6	Sorcin
●	2.14	0.12 (1.31)	CAPZB	B1AK87	Capping protein muscle Z-line
●	2.13	0.13 (1.35)	CLU	P10909	Clusterin
●	2.08	0.20 (1.58)	VWF	P04275	von Willebrand factor
●	2.08	0.11 (1.29)	IGLC2	A0A075B6K9	Ig lambda-2 chain C regions
●	2.02	−0.15 (0.71)	APOB	P04114	Apolipoprotein B-100

Proteins of Fig. 4.4 with an absolute p-value ≤ 0.01. Fold change = $\log_{10}$ value, in parentheses absolute value.

4.2.2   I1(96) vs. C-I1(96)

The first (*i.e.* "pre-antibiotic") blood samples of patients who had been treated for min. 96 h with ampicillin-cefotaxime (I1(96), $n = 9$) were compared with a matched control (C-I1(96), $n = 9$). Mean time points of blood taking were 8 h *pp.* for I1(96) and 46 h for C-I1(96). An unpaired t-test for all included proteins (*cf.* Ch. 3.11) produces the results shown in Fig. 4.5. Of these, 27 proteins are differentially abundant by an absolute p-value of ≤ 0.01, three of which are significant after s_0 correction ("above significance curves"). These are protein S100-P, haptoglobin (HP) and actin-related protein 3 (ACTR3). Tab. 4.4 reports these 27 proteins in detail.
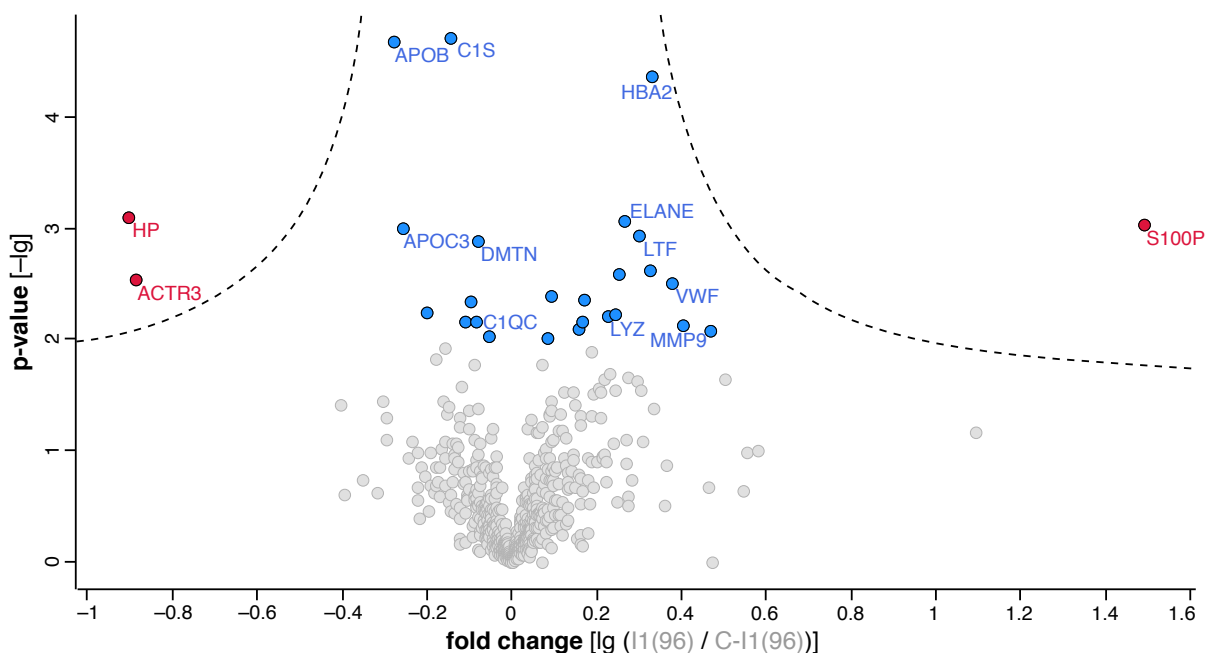
Figure 4.5: Volcano plot I1(96) vs. C-I1(96). x-axis = fold change ("difference of means": mean $(\log_{10})$ LFQ of protein a in 1st group minus mean $(\log_{10})$ LFQ of protein a in 2nd group; or as indicated: logarithmic ratio of mean absolute LFQs of protein a in 1st & 2nd group; *i.e.* pos. x-values reflect a higher abundance of protein a in 1st group and vice versa; absolute fold change of protein a equals $10^{x(a)}$). y-axis = p-value (for respective unpaired t-test on the fold change of protein a). Blue = abs. p-value ≤ 0.01. Red = abs. p-value ≤ 0.01 and fold changes over s_0-corrected significance curves. For further details see Tab. 4.4.

Table 4.4: Protein outliers volcano plot I1(96) vs. C-I1(96) (Fig. 4.5)

○	–lg *p*	Fold change	Gene name	UniProt ID	Protein description
🔵	4.71	−0.15 (0.71)	C1S	P09871	Complement C1s subcomponent
🔵	4.68	−0.28 (0.52)	APOB	P04114	Apolipoprotein B-100
🔵	4.36	0.33 (2.14)	HBA2	G3V1N2	Isoform of Hemoglobin subunit alpha
🔴	3.10	−0.90 (0.13)	HP	P00738	Haptoglobin
🔵	3.06	0.26 (1.82)	ELANE	P08246	Neutrophil elastase
🔴	3.02	1.49 (30.9)	S100P	P25815	Protein S100-P
🔵	2.99	−0.26 (0.55)	APOC3	B0YIW2	Apolipoprotein C-III
🔵	2.93	0.30 (2.00)	LTF	E7EQB2	Lactotransferrin
🔵	2.89	−0.08 (0.83)	DMTN	Q08495	Dematin
🔵	2.62	0.33 (2.14)	MPO	P05164	Myeloperoxidase
🔵	2.58	0.25 (1.78)	H2BFS	O60814	Histone H2B (different types)
🔴	2.53	−0.89 (0.13)	ACTR3	P61158	Actin-related protein 3
🔵	2.50	0.38 (2.40)	VWF	P04275	von Willebrand factor
🔵	2.39	0.09 (1.23)	HNRNPK	P61978	Heterogeneous nuclear ribonucleoprotein K
🔵	2.35	0.17 (1.48)	HK1	P19367	Hexokinase-1 (different isoforms)
🔵	2.34	−0.10 (0.79)	ADD2	P35612	Beta-adducin
🔵	2.23	−0.20 (0.63)	H2AFJ	P04908	Histone H2A (different types)
🔵	2.22	0.24 (1.74)	LCN2	P80188	Neutrophil gelatinase-assoc. lipocalin
🔵	2.21	0.23 (1.70)	LYZ	A0A0B4J259	Lysozyme C
🔵	2.16	−0.09 (0.81)	C1QC	P02747	Complement C1q subcomponent subunit C
🔵	2.16	0.17 (1.48)	TF	P02787	Serotransferrin
🔵	2.15	−0.11 (0.78)	EIF5A	I3L397	Eukaryotic translation initiation factor 5A-1
🔵	2.12	0.41 (2.57)	MMP9	P14780	Matrix metalloproteinase-9
🔵	2.09	0.16 (1.45)	HIST1H1B	P16401	Histone H1.5
🔵	2.08	0.47 (2.95)	TBCEL	Q5QJ74	Tubulin-specific chaperone cofactor E-like protein
🔵	2.03	−0.05 (0.89)	PA2G4	Q9UQ80	Proliferation-assoc. protein 2G4
🔵	2.00	0.09 (1.23)	PFN1	P07737	Profilin-1

Proteins of Fig. 4.5 with an absolute p-value ≤ 0.01. Fold change = $\log_{10}$ value, in parentheses absolute value.

4.2.3   1D annotation enrichment analysis I1(96) vs. C-I1(96)

As introduced in Ch. 1.2.6, a 1D-AEA can be performed on any numerical value of an an-
notated omics data set. Here, it was first conducted on the fold change between I1(96) and
C-I1(96). This 1D-AEA delivered *no* significant results.

A second 1D-AEA on the mere LFQ ("protein intensity") within I1(96) patients showed
the 27 enriched terms in Tab. 4.5. Obviously, two of them are directly linked to the immune
system, the terms "immunoglobulindomain" and "antigen binding".

Table 4.5: 1D annotation enrichment analysis I1(96)

–lg *p*	FDR	Size	Score	Mean	Median	Type	Name
8.42	5.67	78	0.41	5.01	4.77	GOCC name	Blood microparticle
6.35	3.32	10	0.93	6.99	7.41	GOMF name	Oxygen binding
6.35	3.63	10	0.93	6.99	7.41	GOMF name	Oxygen transporter activity
6.35	3.72	10	0.93	6.99	7.41	UP Keywords	Oxygentransport
5.86	3.41	12	0.81	6.63	6.49	GOCC name	Hemoglobin complex
5.41	3.14	145	0.25	4.69	4.63	GOCC name	Extracellular space
5.39	3.06	16	0.67	5.27	5.18	UP Keywords	Immunoglobulindomain
5.29	3.14	18	0.63	5.37	5.03	UP Keywords	Hereditaryhemolyticanemia
5.23	3.20	454	0.21	4.57	4.44	UP Keywords	Directproteinsequencing
4.79	2.63	129	0.24	4.73	4.59	GOCC name	Extracellular region
4.74	2.80	16	0.63	6.07	5.47	UP Keywords	Heme
4.46	1.91	19	0.56	5.85	5.36	GOMF name	Heme binding
4.46	2.03	19	0.56	5.85	5.36	GOMF name	Tetrapyrrole binding
4.39	2.06	61	0.32	5.07	4.71	GOMF name	Transition metal ion binding
4.20	1.95	14	0.62	5.06	4.85	GOMF name	Antigen binding
4.15	2.09	36	0.39	5.08	4.85	GOCC name	Cytopl. mb.-bound vesicle lumen
4.15	2.17	36	0.39	5.08	4.85	GOCC name	Vesicle lumen
4.06	2.20	6	0.93	6.66	6.54	UP Keywords	Glycation
4.02	1.83	26	0.45	5.55	5.11	GOMF name	Iron ion binding
3.86	1.95	29	−0.42	4.06	4.10	GOCC name	Golgi apparatus part
3.85	2.06	7	0.84	5.41	5.51	UP Keywords	Oxidation
3.65	1.92	15	−0.56	3.91	3.97	UP Keywords	Endosome
3.46	1.78	12	−0.60	3.92	3.94	UP Keywords	Prenylation
3.33	1.70	14	−0.55	3.86	3.91	UP Keywords	Golgiapparatus

Enriched annotation terms of proteins in I1(96). Red = terms with direct link to the immune system. **–lg *p*** =
$-\lg p$ for performed 1D-AEA. **FDR** = $-\lg$ of Benjamini-Hochberg FDR for performed 1D-AEA. **Size** =
total number of proteins in I1(96) annotated with this term. **Score** = reflects distance between centres of
distributions. **Mean** = mean LFQ of proteins annotated with this term. **Median** = median LFQ of proteins
annotated with this term. **Type** = category of annotation term. **Name** = term identified as enriched. For
further methodical details of the 1D-AEA *cf.* Ch. 1.2.6 or Cox & Mann (2012).

Fig. 4.6 illustrates to what extent the distributions of the annotation terms "immunoglobulin" & "antigen binding" deviate from the centre of the overall distribution of annotation terms. In the subfigures (A1) & (B1), the bars in red colour represent the distribution of the indicated annotation term, whereas the bars in grey colour represent the distribution of all other annotation terms (of the respective annotation category "Type" in Tab. 4.5). The terms "immunoglobulin" (score = 0.67) and "antigen binding" (score = 0.62) are clearly located in the upper half of the overall distribution. In the corresponding extracts (A2) & (B2) from the volcano plot I1(96) vs. C-I1(96) (Fig. 4.5) those proteins annotated with the respective term are highlighted in red colour as well.

The two histograms (A1) & (B1) show the comparatively high abundance of proteins with the respective annotation term *within* the patient group I1(96). The two corresponding volcano plots (A2) & (B2), however, show the relatively higher abundance of proteins with these annotation terms in I1(96) *compared* to its matched control C-I1(96).
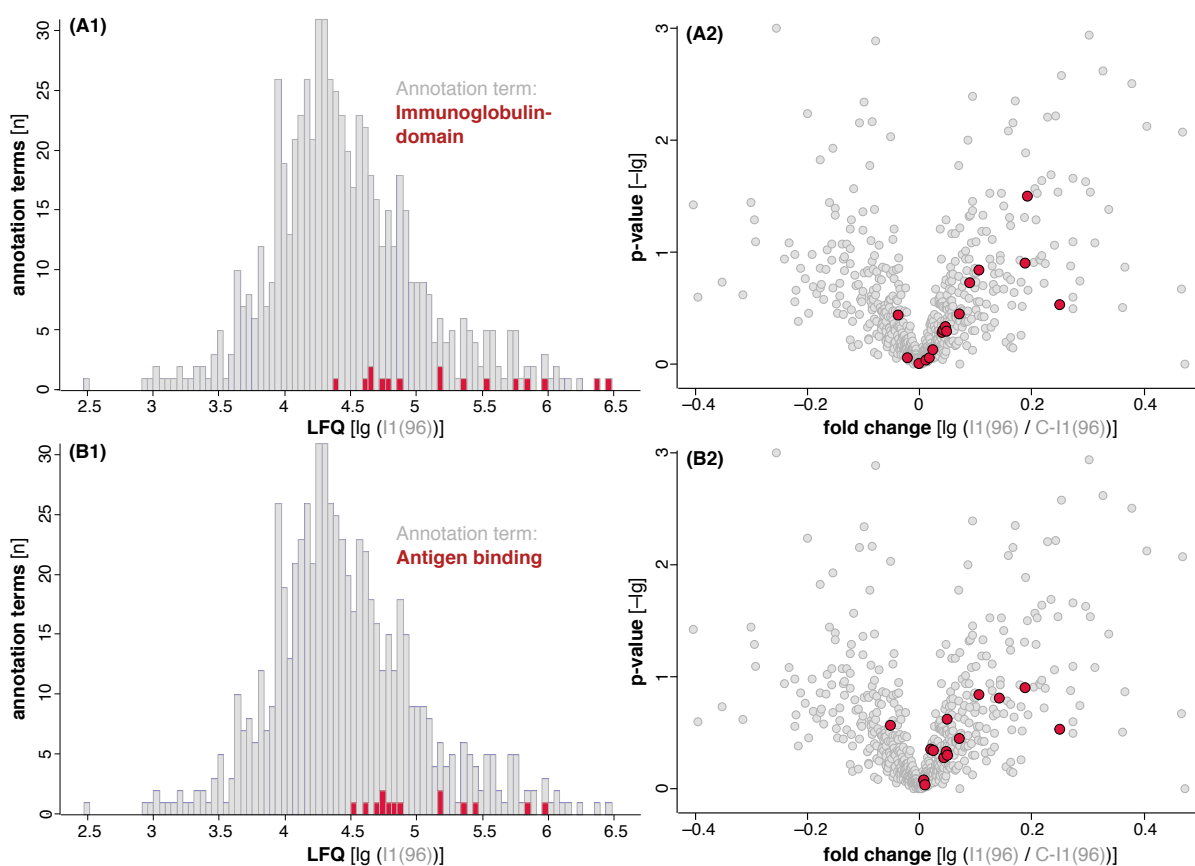
Figure 4.6: 1D annotation enrichment analysis I1(96). Histograms (A1) & (B1): x-axis = LFQ of proteins measured in I1(96) patients. y-axis = counted annotation terms for all proteins with an LFQ within the range of one bin. Total of bins = 100. Volcano plots (A2) & (B2) are extracts from Fig. 4.5 showing only the centre of the original plot and highlighting those proteins annotated with the respective term. For methodical details of the 1D-AEA *cf.* Ch. 1.2.6 or Cox & Mann (2012).

4.2.4 Correlations with clinical parameters I1(96)

All 27 proteins from the volcano plot I1(96) vs. C-I1(96) (Fig. 4.5) with an absolute p-value ≤ 0.01 were subjected to a correlation analysis. Fig. 4.7, a Euclidean distance heatmap in which all inputs are hierarchically clustered by their Pearson correlation coefficient, illustrates in which way these 27 proteins correlate with few selected clinical data points or anamnestic values taken from the study questionnaire.
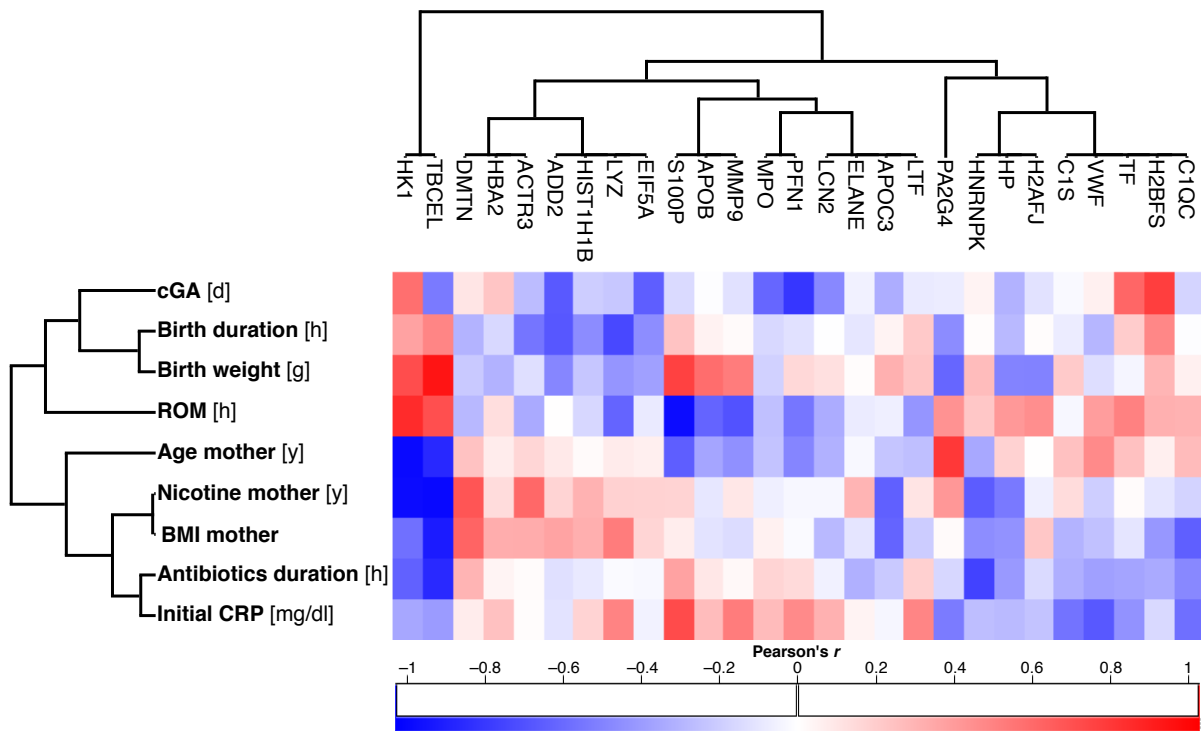

Figure 4.7: Correlation heatmap I1(96). Euclidean distance heatmap, hierarchical clustering by Pearson's correlation coefficient (r). Correlated were selected clinical and anamnestic data points with the LFQ of 27 proteins in I1(96) patients, namely those of Fig. 4.5 with an abs. p-value ≤ 0.01. **cGA** = corrected gestational age. **ROM** = "rupture of membranes", interval between amniorrhexis and delivery. **Nicotine mother** = duration of nicotine consumption before pregnancy. **BMI mother** = body mass index at first gynaecological check-up of pregnancy. **Antibiotics duration** = total duration of therapy. **Initial CRP** = value of CRP decisive for the start of antibiotic treatment.

For two of the selected clinical data points, the total duration of antibiotic therapy and the decisive CRP value at the onset of antibiotic treatment, Pearson's r was calculated and plotted against its respective p-value (Fig. 4.8).

For heterogeneous nuclear ribonucleoprotein K (hnRNPK), this kind of illustration shows that the observed negative correlation with the total duration of antibiotic therapy ($r = -0.76$ at $p = 0.02$) is not only pronounced but also relatively reliable. This relation is clearly more marked than it is between hnRNPK and the initial CRP (*cf.* Fig. 4.7 or Fig. 4.8 B).

S100-P and vWF, in contrast, strongly correlate with the initial CRP (S100-P: $r = 0.73$ at $p = 0.03$; vWF: $r = -0.67$ at $p = 0.05$), but they do not with the total duration of antibiotic therapy. For the respective regression analysis of all three proteins see Fig. 4.9.
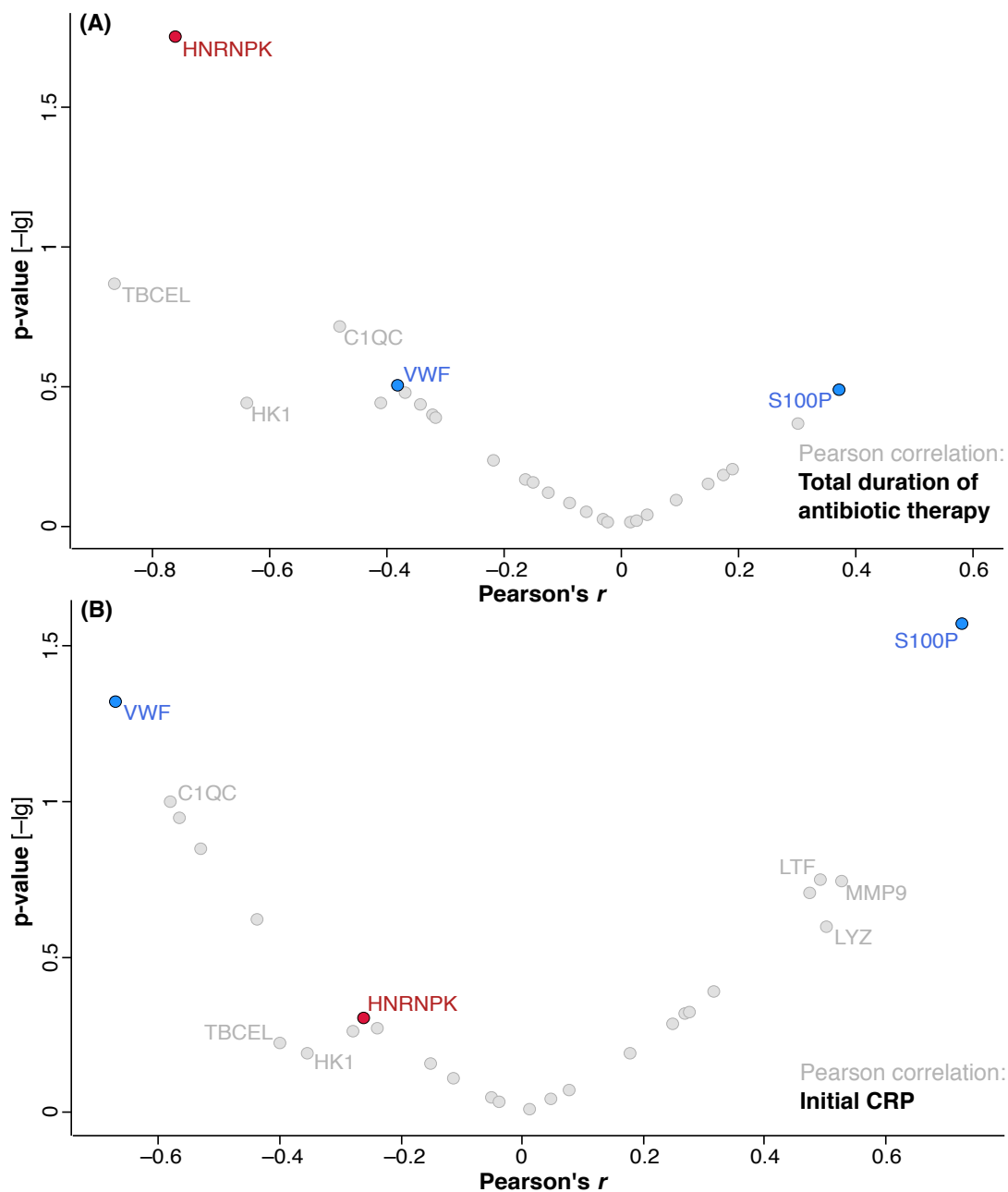
Figure 4.8: Correlation scatter plots I1(96). x-axis = Pearson's r (correlation between indicated category and protein abundance, *i.e.* LFQ). y-axis = $-\lg p$ (for Pearson's r). Subject to this correlation are again all 27 proteins from Fig. 4.5 with an abs. p-value ≤ 0.01.
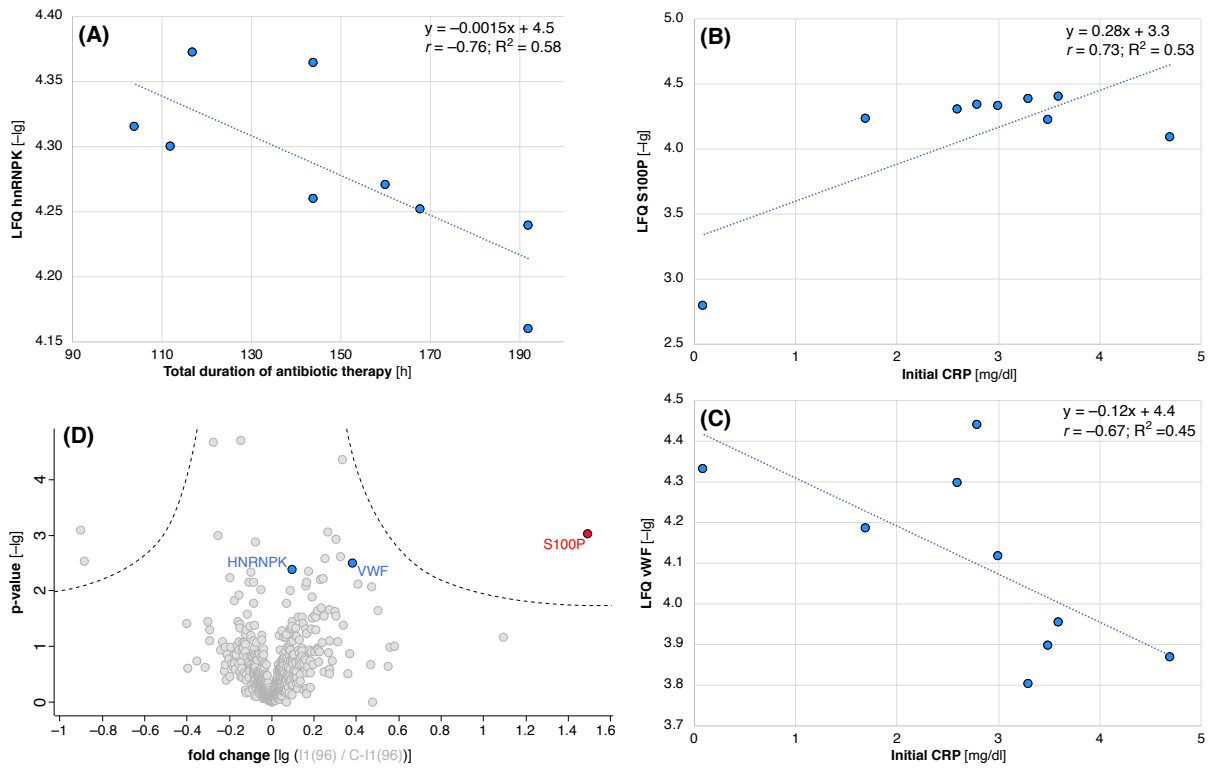
Figure 4.9: Regression scatter plots I1(96). The scatter plots (A)-(C) show the regression analysis for those three proteins that significantly correlate with the total duration of antibiotic therapy or the initial CRP, respectively (*cf.* Fig. 4.8). Subfigure (D) additionally shows their location in the volcano plot I1(96) vs. C-I1(96) (*cf.* Fig. 4.5).

4.3 Question 2 – "Dynamics"

Are there significant **dynamics** in the whole blood proteomes of neonates ($\geq$ 35+0) with prolonged early-onset infection in the early phase of their antibiotic treatment? (*cf.* Ch. 2. Objectives)

I approached this question in three steps, as depicted in Fig. 4.10. The subject of all considerations was the small but homogeneous group of neonates which had been severely affected by an early-onset infection and treated for min. 96 h (exclusively with ampicillin-cefotaxime). In a first step, their available blood samples (one before the first administration of antibiotics, one ca. 48 h after) underwent a paired t-test. Next, I performed a 1D-AEA on the mean of differences between t_1 & t_2. Third, I compared the differences between both time points in the infection group to the differences between two unrelated groups of control patients which had been matched to the infection group as accurately as possible in their time points of blood taking.
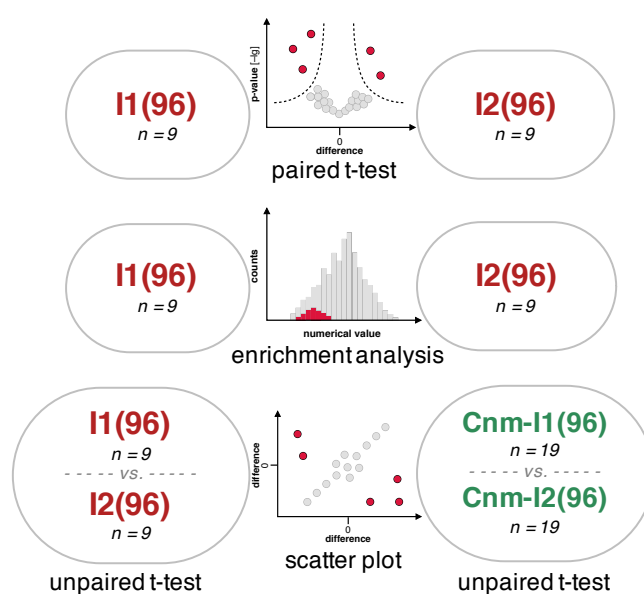

Figure 4.10: Graphical abstract (question 2)

4.3.1   I1(96) vs. I2(96)

The paired t-test for all included proteins (*cf.* Ch. 3.11) of the pre-antibiotic samples I1(96) versus their follow-up measurement I2(96) in the course of antibiotic treatment is displayed in Fig. 4.11. Mean time points of blood taking were 8 h *pp.* for I1(96) and 55 h for I2(96). For practical reasons, in the present volcano plot significant outliers are primarily defined by their Benjamini-Hochberg correction derived q-values. The three proteins that are increased in I2(96) with the highest significance are the apolipoproteins E & B-100 and Inter-alpha-trypsin inhibitor heavy chain H3 (ITIH3). Among the proteins with a medium or low significant increase in I2(96), there are two further apolipoproteins and several well-known acute-phase reactants. Among the decreased proteins in I2(96), protein S100-A8 & S100-A9 are to be pointed out.
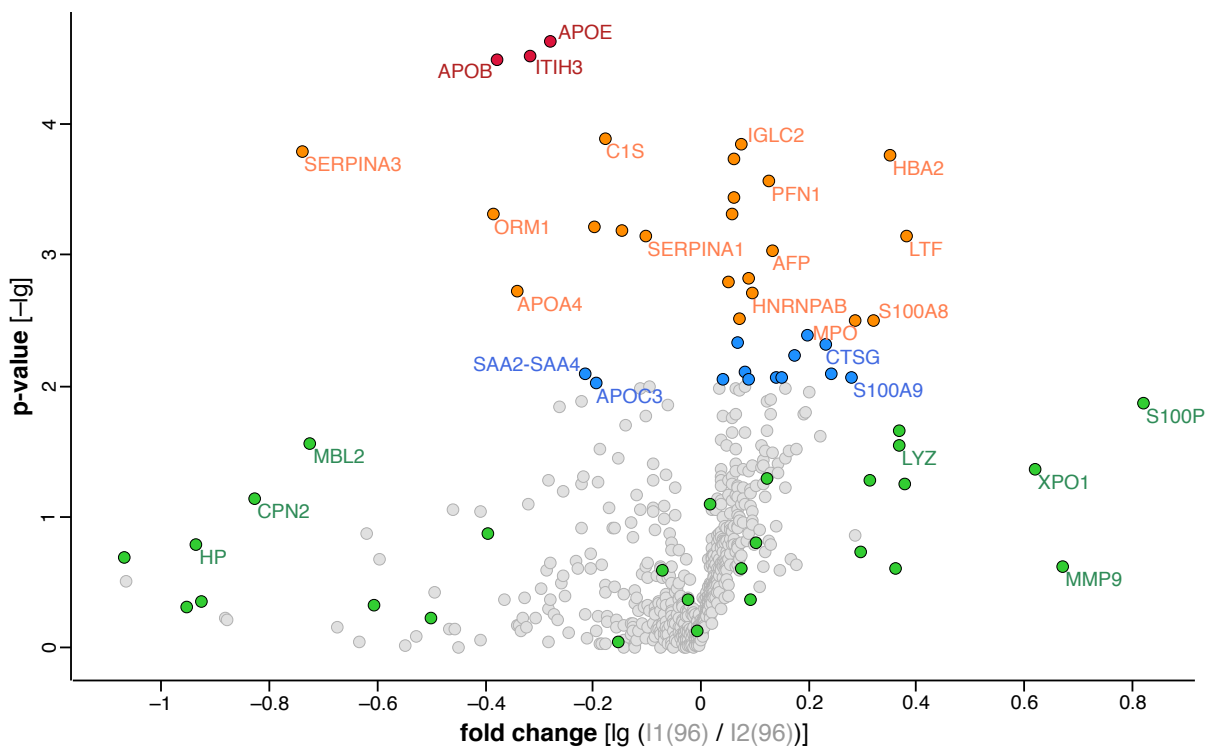
Figure 4.11: Volcano plot I1(96) vs. I2(96). x-axis = fold change ("mean of differences", mean of all paired differences between ($\log_{10}$) LFQ of protein a at t_1 minus ($\log_{10}$) LFQ of protein a at t_2 in same patient; or as indicated: mean of all logarithmic ratios of absolute LFQs of protein a at t_1 and t_2 in same patient; *i.e.* pos. values reflect a higher abundance of protein a at t_1 and vice versa; absolute fold change equals $10^{x(a)}$). y-axis = $-\lg p$ (for respective paired t-test on the fold change of protein a). Blue = p-value ≤ 0.01. Orange = p-value ≤ 0.01 & q-value ≤ 0.1. Red = p-value ≤ 0.0001 & q-value ≤ 0.01. Green = all proteins whose x- and y-values in Fig. 4.13 differ by at least 0.301 (and do not have a p-value ≤ 0.01). For further details see Tab. 4.6.

Table 4.6: Protein outliers volcano plot I1(96) vs. I2(96)

○	–lg *p*	–lg *q*	Fold change	Gene name	UniProt ID	Protein description
●	4.64	2.13	−0.28 (0.52)	APOE	P02649	Apolipoprotein E
●	4.52	2.13	−0.32 (0.48)	ITIH3	Q06033	Inter-alpha-trypsin inhibitor heavy chain H3
●	4.49	2.13	−0.38 (0.42)	APOB	P04114	Apolipoprotein B-100
○	3.89	1.81	−0.18 (0.66)	C1S	P09871	Complement C1s subcomponent
○	3.85	1.81	0.08 (1.20)	IGLC2	A0A075B6K9	Ig lambda-2 chain C regions
○	3.79	1.81	−0.74 (0.18)	SERPINA3	P01011	Alpha-1-antichymotrypsin
○	3.76	1.81	0.35 (2.24)	HBA2	G3V1N2	Isoform of Hemoglobin subunit alpha
○	3.74	1.81	0.06 (1.15)	CCT4	P50991	T-complex protein 1 subunit delta
○	3.56	1.69	0.13 (1.35)	PFN1	P07737	Profilin-1
○	3.43	1.60	0.06 (1.15)	IGLV8-61	A0A075B6I0	Protein IGLV8-61 (Fragment)
○	3.32	1.56	−0.38 (0.42)	ORM1	P02763	Alpha-1-acid glycoprotein 1
○	3.31	1.56	0.06 (1.15)	MYH9	P35579	Myosin-9
○	3.21	1.52	−0.20 (0.63)	FGA	P02671	Fibrinogen alpha chain
○	3.19	1.52	−0.15 (0.70)	FGG	P02679	Fibrinogen gamma chain
○	3.15	1.52	0.38 (2.40)	LTF	E7EQB2	Lactotransferrin
○	3.15	1.52	−0.10 (0.79)	SERPINA1	P01009	Alpha-1-antitrypsin
○	3.04	1.43	0.13 (1.35)	AFP	P02771	Alpha-fetoprotein
○	2.88	1.31	−0.22 (0.60)	–	B4E1Z4	Uncharacterized protein
○	2.82	1.27	0.09 (1.23)	PSMD14	O00487	26S proteasome non-ATPase regulatory subunit 14
○	2.80	1.26	0.05 (1.12)	HSP90AA1	P07900	Heat shock protein HSP 90-alpha
○	2.72	1.22	−0.34 (0.46)	APOA4	P06727	Apolipoprotein A-IV
○	2.71	1.22	0.10 (1.26)	HNRNPAB	A0A087WZV1	Heterogeneous nuclear ribonucleoprotein A/B
○	2.52	1.06	0.07 (1.17)	IGKV3-15	A0A0B4J1T9	Protein IGKV3-15 (Fragment)
○	2.50	1.06	0.32 (2.09)	S100A8	P05109	Protein S100-A8
○	2.50	1.06	0.29 (1.95)	MPO	P05164	Myeloperoxidase
○	2.38	0.97	0.20 (1.58)	VIM	P08670	Vimentin
○	2.33	0.93	0.07 (1.17)	LTA4H	P09960	Leukotriene A-4 hydrolase
○	2.31	0.93	0.23 (1.70)	CTSG	P08311	Cathepsin G
○	2.23	0.86	0.17 (1.48)	CORO1A	P31146	Coronin-1A
○	2.11	0.79	0.08 (1.20)	LCP1	P13796	Plastin-2
○	2.09	0.79	0.24 (1.74)	ATP5A1	P25705	ATP synthase subunit alpha
○	2.09	0.79	−0.21 (0.62)	SAA2-SAA4	A0A096LPE2	Protein SAA2-SAA4
○	2.06	0.79	0.14 (1.38)	IGKV3-11	A0A087WZW8	Protein IGKV3-11
○	2.06	0.79	0.28 (1.91)	S100A9	P06702	Protein S100-A9
○	2.06	0.79	0.15 (1.41)	GANAB	E9PKU7	Neutral alpha-glucosidase AB
○	2.05	0.79	0.09 (1.23)	IGHV3-23	A0A0B4J2B7	Ig heavy chain V-III region 23
○	2.05	0.79	0.04 (1.10)	IGHG2	A0A0G2JN06	Ig gamma-2 chain C region
○	2.03	0.79	−0.19 (0.65)	APOC3	B0YIW2	Apolipoprotein C-III

Proteins of Fig. 4.11 with an absolute p-value ≤ 0.01. Fold change = $\log_{10}$ value, in parentheses absolute value.

4.3.2 1D annotation enrichment analysis I1(96) vs. I2(96)

The present 1D-AEA on the fold change between I1(96) and I2(96) delivered 12 significantly enriched annotation terms listed in Tab. 4.7. With one exception ("acetylation") all of them are enriched towards t_2 (*i.e.* the abundances of proteins with these annotations increased between t_1 & t_2). Among them, the terms "complement and coagulation cascades", "acute-phase response" and "staphylococcus aureus infection" can be found (*cf.* Fig. 4.12).

Table 4.7: 1D annotation enrichment analysis I1(96) vs. I2(96)

–lg *p*	FDR	Size	Score	Mean	Median	Type	Name
7.70	5.41	34	-0.57	-0.14	-0.10	KEGG name	Complement and coag. cascades
6.56	3.93	117	-0.30	-0.09	-0.07	UP Keywords	Secreted
6.14	2.49	15	-0.75	-0.27	-0.16	GOBP name	Acute-phase response
5.02	1.85	14	-0.69	-0.15	-0.16	GOBP name	Negative regulation of coagulation
4.96	2.63	133	-0.25	-0.08	-0.06	UP Keywords	Signal
4.80	1.76	20	-0.57	-0.13	-0.13	GOBP name	Regulation of wound healing
4.78	2.79	14	-0.67	-0.18	-0.11	KEGG name	Staphylococcus aureus infection
4.55	2.40	10	-0.77	-0.29	-0.16	UP Keywords	Acutephase
4.16	2.13	386	0.18	-0.01	0.01	UP Keywords	Acetylation
3.97	2.04	141	-0.21	-0.07	-0.04	UP Keywords	Glycoprotein
3.91	2.05	18	-0.53	-0.11	-0.12	UP Keywords	Bloodcoagulation
3.91	2.12	18	-0.53	-0.11	-0.12	UP Keywords	Hemostasis

Enriched annotation terms of proteins for the paired t-test I1(96) vs. I2(96). Red = terms with direct link to the innate immune response. **–lg *p*** = $-\lg p$ for performed 1D-AEA. **FDR** = $-\lg$ of Benjamini-Hochberg FDR for performed 1D-AEA. **Size** = total number of proteins in I1(96) annotated with this term. **Score** = reflects distance between centres of distributions. **Mean** = mean LFQ of proteins annotated with this term. **Median** = median LFQ of proteins annotated with this term. **Type** = category of annotation term. **Name** = term identified as enriched. For further methodical details of the 1D-AEA *cf.* Ch. 1.2.6 or Cox & Mann (2012).

Fig. 4.12 illustrates to what extent the distributions of the indicated annotation terms deviate from the centre of the overall distribution. All three terms (scores = –0.57 to –0.75) are clearly located in the lower half of the overall distribution showing the increase of proteins associated with these very terms between t_1 & t_2. In the corresponding copies (A2), (B2) & (C2) of the volcano plot I1(96) vs. I2(96) (Fig. 4.11) those proteins annotated with the respective term are highlighted in red colour.
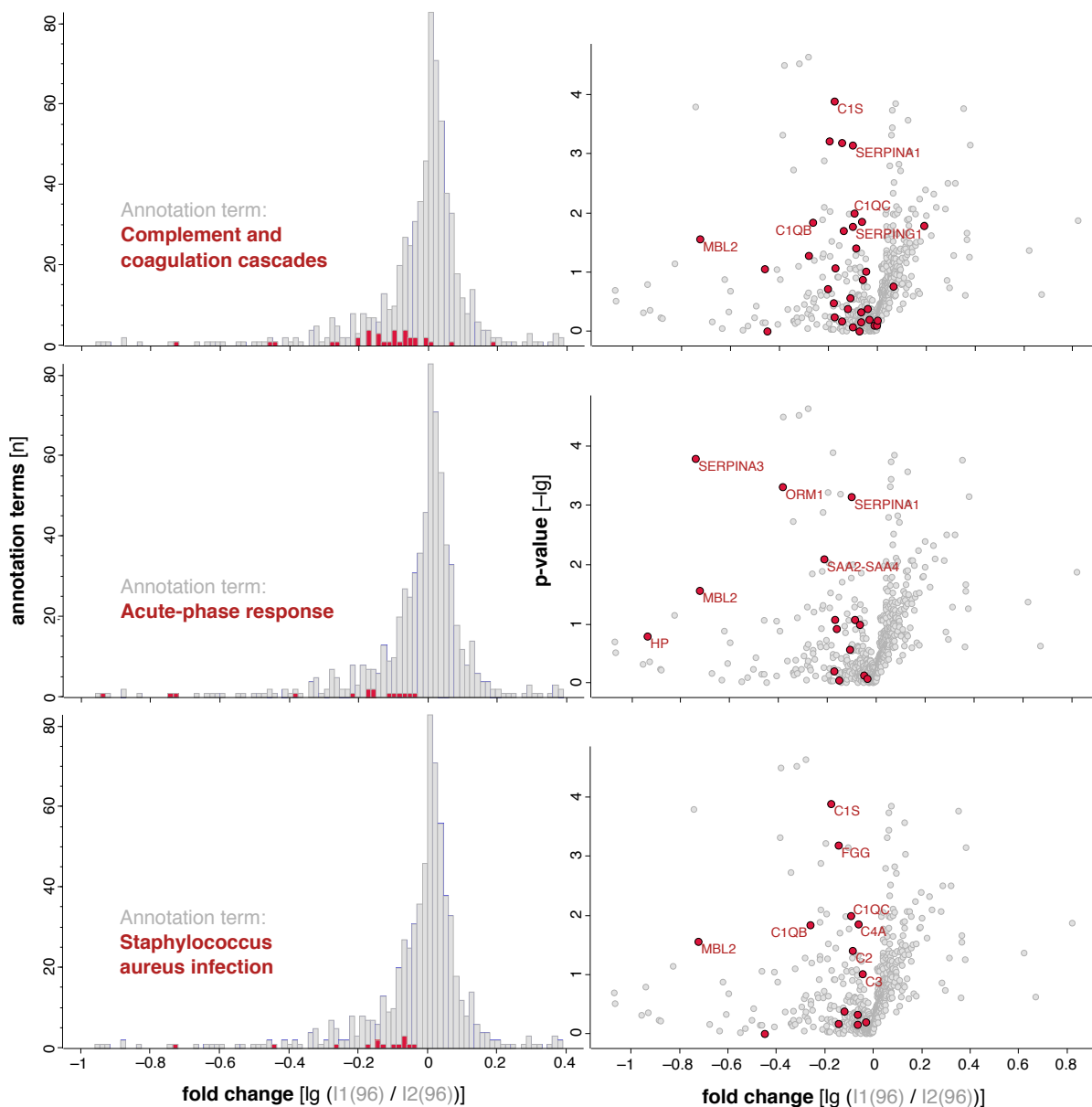
Figure 4.12: 1D annotation enrichment analysis I1(96) vs. I2(96). Histograms (A1), (B1) & (C1): x-axis = mean of differences calculated for the paired t-test I1(96) vs. I2(96), *i.e.* mean change of LFQ in each patient. y-axis = counted annotation terms for all proteins with an LFQ within the range of one bin. Total of bins = 100. Volcano plots (A2), (B2) & (C2) show the position of those proteins annotated with the respective term in the volcano plot of the original paired t-test (Fig. 4.11). For methodical details of the 1D-AEA *cf.* Ch. 1.2.6 or Cox & Mann (2012).

4.3.3 I1(96) to I2(96) vs. Cnm-I1(96) to Cnm-I2(96)

Validation of the protein dynamics seen in the volcano plot I1(96) vs. I2(96) (*cf.* Fig. 4.11) is only possible by comparing them to the physiological postnatal dynamics in healthy newborns. Here, the practical difficulty is that no neonate at best health requires two consecutive blood takings. The formation of a suitable longitudinal control group is, thus, almost impossible. Due to the limited number of blood samples available in the pilot study (*cf.* Ch. 3.4) it was neither possible to form two unrelated control groups which matched the infection group in terms of both cGA and the mean time interval between the samples. A reasonable grouping could only be conducted with respect to the mere time interval. To achieve at least sufficient quantities, all 38 available control blood samples of healthy newborns were separated in two groups (n = 19) by their median time of blood taking (44 h *pp.*). The fold changes from unpaired t-tests between both control and both infection groups, respectively, were plotted against each other. Mean time points of blood taking were: 8 h *pp.* for I1(96), 55 h for I2(96), 39 h for Cnm-I1(96) and 53 h for Cnm-I2(96).
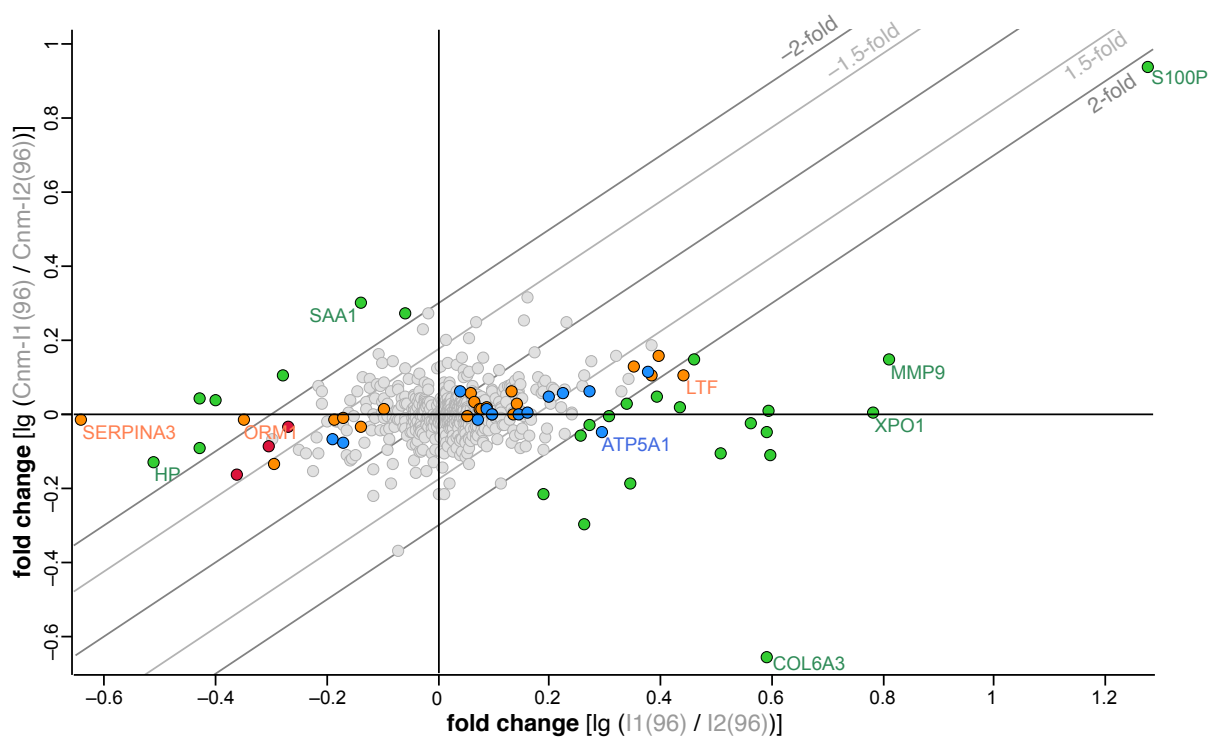

Figure 4.13: Scatter plot I1(96) to I2(96) vs. Cnm-I1(96) to Cnm-I2(96). x-axis & y-axis = fold change ("difference of means": mean ($\log_{10}$) LFQ of protein a in 1st group minus mean ($\log_{10}$) LFQ of protein a in 2nd group; or as indicated: logarithmic ratio of mean absolute LFQs of protein a in 1st & 2nd group; *i.e.* pos. x- or y-values reflect a higher abundance of protein a in 1st group and vice versa; absolute fold change of protein a equals $10^{x(a)}$ or $10^{y(a)}$, respectively). x- and y-axis both display the fold changes from unpaired t-tests. The colour highlighting though refers to the paired t-test displayed in Fig. 4.11. Green = all other proteins whose x- and y-values differ by at least 0.301, *i.e.* proteins that differ in their absolute intensity at least by factor 2. For a better visual interpretation, the following straight lines are drawn: $y = x \pm 0.301$, $y = x \pm 0.176$ and $y = x$ ($lg2 = 0.301$ & $lg1.5 = 0.176$).

The formation of data points in Fig. 4.11 resembles a diffuse cloud around the centre of the plot. Since there is no obvious alignment of points at $y = x$, this method does not allow

a direct identification of proteins whose dynamics between t_1 & t_2 could now be traced to presumably physiological changes within the first days of life. At least based on this type of control plot, none of the significant proteins from Tab. 4.6 need to be excluded from further considerations.

On the other hand, four proteins with a significant fold change in the volcano plot I1(96) vs. I2(96) (*cf.* Tab. 4.6) are among those proteins of Fig. 4.13 whose x- and y-values differ by at least 0.301. As x- and y-axis are scaled in the decadic logarithm these four proteins – LTF, ATP5A1, ORM1 & SERPINA3 – differ in their absolute protein intensity at least by factor 2 ($lg2 = 0.301$). Their significant (Fig. 4.11) and at the same time specific change (Fig. 4.13) underlines these proteins' potentially important role in innate immune response.

This consideration can be extended to altogether eight significant proteins whose x- and y-values differ by 0.176–0.301, reflecting an absolute difference by factor 1.5–2 ($lg1.5 = 0.176$). Among these eight proteins are those three proteins with the highest significance in Fig. 4.11 – APOE, ITIH3 & APOB – and furthermore S100A8, S100A9, HBA2, MPO & CTSG.

4.4 Question 3 – "Distinction"

Are there significant features for a **distinction** between the whole blood proteomes of neonates ($\geq 35+0$) with early-onset infection whose antibiotic treatment lasted *at least* 96 h or *at most* 72 h, respectively? (*cf.* Ch. 2. Objectives)

I approached this question in two steps, as depicted in Fig. 4.14. First, I compared t_1 of the group of patients treated for min. 96 h (exclusively with ampicillin-cefotaxime) to t_1 of the group of patients which had been treated for max. 72 h (with the same antibiotics). In a second step, I contrasted the differences between both infection groups and their matched control group with each other.
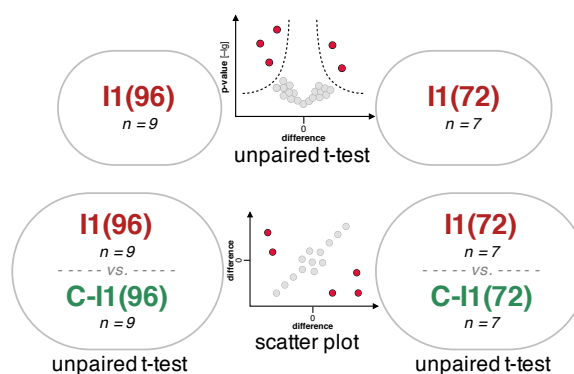

Figure 4.14: Graphical abstract (question 3)

4.4.1   I1(96) vs. I1(72)

Comparing the first (*i.e.* "pre-antibiotic") blood samples of those two groups of patients who were both treated with ampicillin-cefotaxime only, but for different periods of time, the unpaired t-test for all included proteins (*cf.* Ch. 3.11) produces the results shown in Fig. 4.15. Mean time points of blood taking were 8 h *pp.* for I1(96) and 15 h for I1(72). Five proteins are differentially abundant by an absolute p-value of ≤ 0.01, among them ACTR3 (*cf.* Ch. 4.2.2, p. 46).
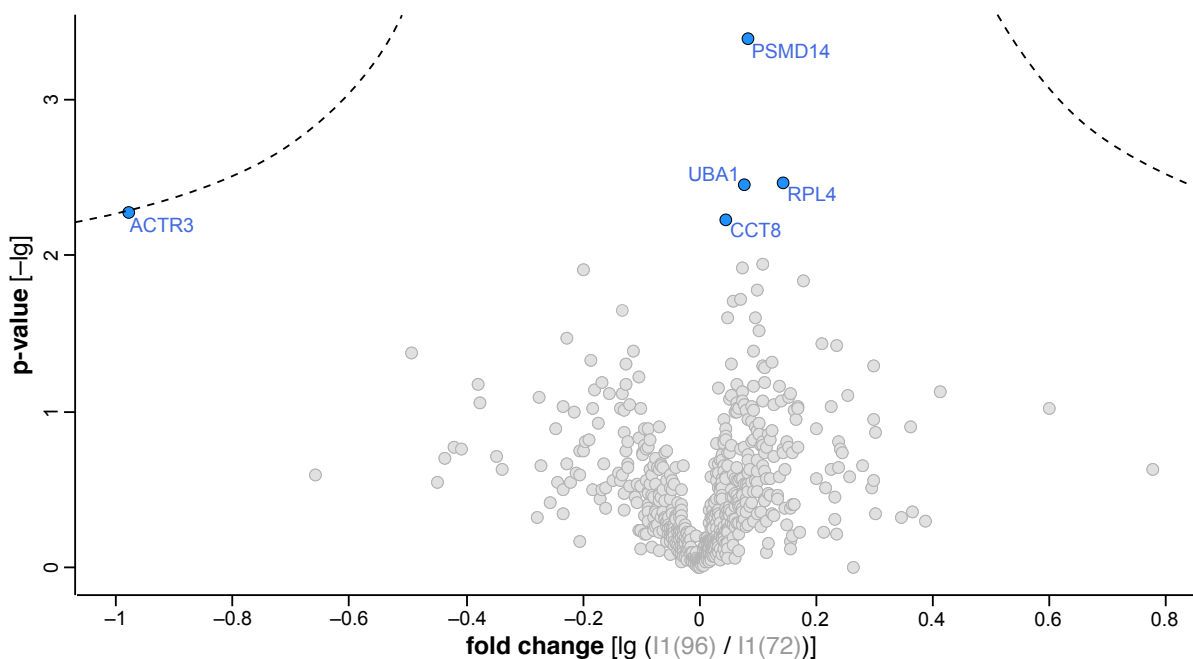

Figure 4.15: Volcano plot I1(96) vs. I1(72). x-axis = fold change ("difference of means": mean $(\log_{10})$ LFQ of protein a in 1st group minus mean $(\log_{10})$ LFQ of protein a in 2nd group; or as indicated: logarithmic ratio of mean absolute LFQs of protein a in 1st & 2nd group; *i.e.* pos. x-values reflect a higher abundance of protein a in 1st group and vice versa; absolute fold change of protein a equals $10^{x(a)}$). y-axis = p-value (for respective unpaired t-test on the fold change of protein a). Blue = abs. p-value ≤ 0.01. For further details see Tab. 4.8.

Table 4.8: Protein outliers volcano plot I1(96) vs. I1(72)

○	–lg *p*	Fold change	Gene name	UniProt ID	Protein description
●	3.39	0.08 (1.20)	PSMD14	O00487	26S proteasome non-ATPase regulatory subunit 14
●	2.46	0.14 (1.38)	RPL4	P36578	60S ribosomal protein L4
●	2.45	0.08 (1.20)	UBA1	P22314	Ubiquitin-like modifier-activating enzyme 1
●	2.28	−0.98 (0.10)	ACTR3	P61158	Actin-related protein 3
●	2.23	0.04 (1.10)	CCT8	P50990	T-complex protein 1 subunit theta

Proteins of Fig. 4.15 with an absolute p-value ≤ 0.01. Fold change = $\log_{10}$ value, in parentheses absolute value.

Fig. 4.16 with its subfigures (A)-(C) puts the volcano plot I1(96) vs. I1(72) in a context with findings from question 1 & 2 (Ch. 4.3 & 4.4)

In Subfigure (A), the five proteins of interest from question 1 are highlighted (significant fold change in Fig. 4.5, p. 46, or remarkable correlation in Fig. 4.8, p. 51). Obviously, ACTR3 is the only one of them which is also nearly significantly differential between I1(96) and I1(72). Its decrease is in accordance with Fig. 4.5. Subfigure (B) illustrates the position of seven proteins of interest from question 2, namely such with a significant fold change in Fig. 4.11, p. 54, or with a remarkable position in Fig. 4.13, p. 58. Subfigure (C) highlights those proteins annotated with the significantly enriched terms from question 2 "Complement and coagulation cascades", "Acute-phase response" and "Staphylococcus aureus infection" (*cf.* Tab. 4.7, p. 56). In subfigure (D), such proteins annotated with the Uniprot Keyword "antimicrobial" are highlighted red. Although the obvious shift to the right, *i.e.* higher abundances in I1(96), could not be detected statistically in a 1D-AEA it shall be reported here as a visual finding.
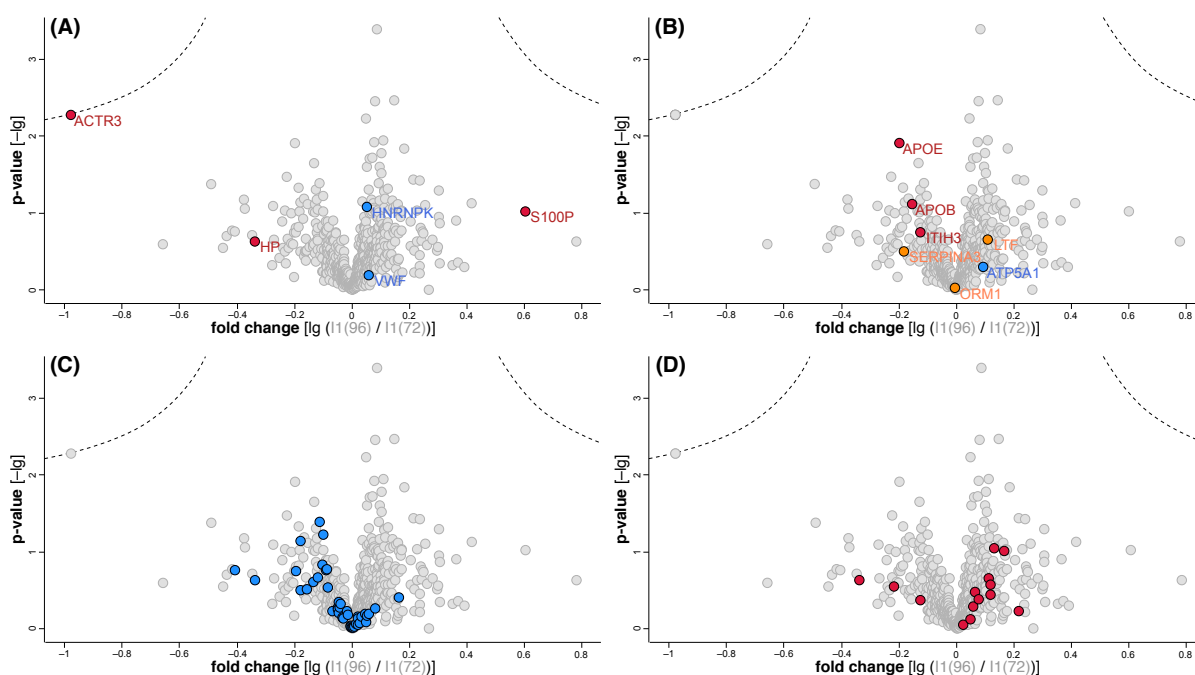

Figure 4.16: Volcano plot I1(96) vs. I1(72) with cross-references. All four subfigures are copies of Fig. 4.15 and highlight a different interpretative aspect taken from question 1 & 2 (Ch. 4.2 & 4.3). Subfig. **(A)** refers to question 1: Red are those three significant proteins of Fig. 4.5, p. 46, blue are two of those proteins with a significant correlation with the duration of antibiotic therapy or the initial CRP in Fig. 4.8, p. 51. Subfig. **(B)** refers to question 2: Red are those three proteins of Fig. 4.11, p. 54, with the highest significance, orange & blue are those significant proteins with remarkable positions in Fig. 4.13, p. 58. Subfig. **(C)** refers to question 2: Blue are all proteins associated with three relevant annotation terms from Fig. 4.12, p. 57 ("complement and coagulation cascades", "acute-phase response" and "staphylococcus aureus infection"). Subfig. **(D)**: Red are all proteins associated with the Uniprot keyword "antimicrobial".

4.4.2 I1(96) to C-I1(96) vs. I1(72) to C-I1(72)

The previous consideration, a direct comparison of the infection groups, is extended to a comparison of both groups' differences to their respective matched controls. The fold changes are plotted against each other, and selected proteins of interest from question 1 & 2 are highlighted (Fig. 4.17 & 4.18). These are the seven proteins with significant dynamics from I1(96) to I2(96) (*cf.* Fig. 4.11 & 4.13, couloured red, orange & blue) and those three proteins with remarkable correlations to the initial CRP or the total duration of antibiotic therapy (*cf.* Fig. 4.8 & 4.9, coloured pink). Mean time points of blood taking were: 8 h *pp.* for I1(96), 46 h for C-I1(96), 15 h for I1(72) and 38 h for C-I1(72).

In general, Fig. 4.17 does not reflect a significantly similar behaviour of fold changes in I1(96) and I1(72). Rather than aligning with $y = x$, the data points form a symmetric and centred point cloud. There is no accumulation of points in the I. or III. quadrant ($x/y = +/+$ or $-/-$). However, a such is at least roughly implied for the coloured proteins in the scatter plot. Among them, the most prominent ones are S100-P, ACTR3 and HP, whose (significant) fold changes in I1(96) vs. C-I1(96) (x-axis, *cf.* Fig. 4.5) are at least 2-fold higher than those in I1(72) vs. C-I1(72) (y-axis).
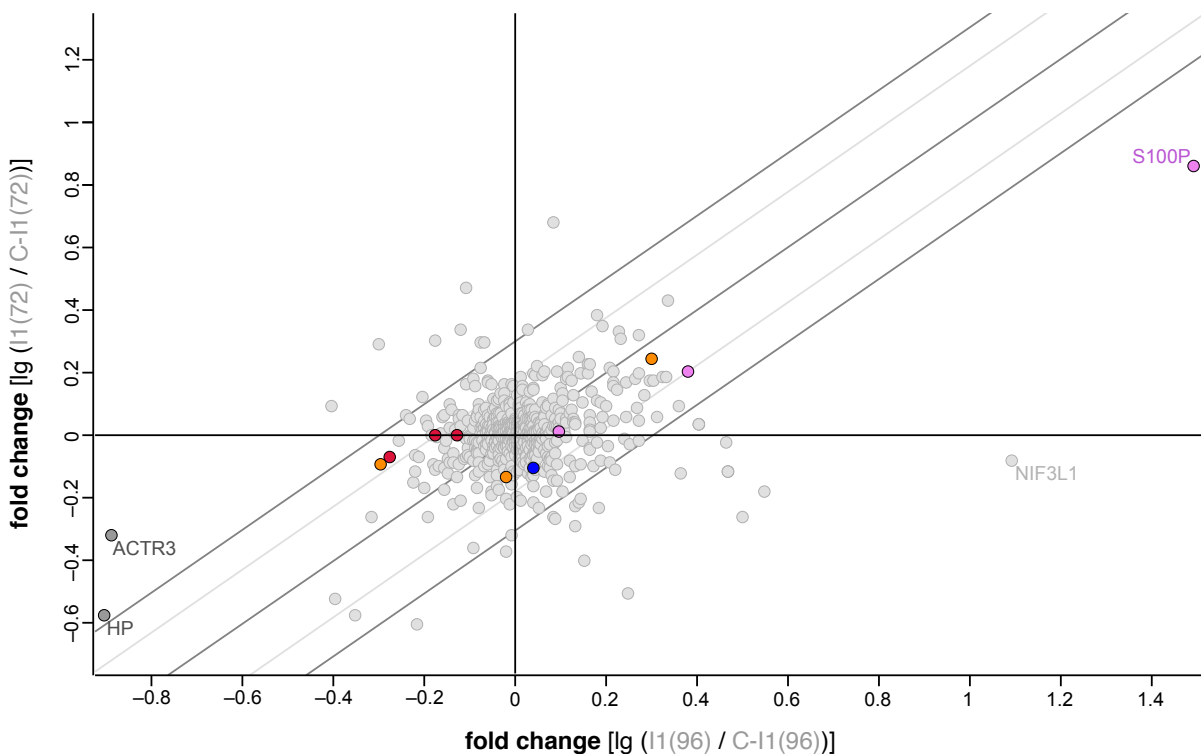
Figure 4.17: Scatter plot I1(96) to C-I1(96) vs. I1(72) to C-I1(72) (overview). This figure shows all available proteins whereas Fig. 4.18 focuses on the centre of the plot, cutting out some outliers as S100-P, ACTR3 or HP (*cf.* Fig. 4.5). For details of the colour coding or the drawn straight lines see also Fig. 4.18.

Apo-E and ITIH3 are two of those proteins with the highest significance ($q \leq 0.01$) for an increase from I1(96) to I2(96), as displayed in the paired t-test of Fig. 4.11, p. 54. While they are decreased in I1(96), a respective difference is absent for I1(72) (almost aligned at $y = 0$).

hnRNPK shows barely any difference between I1(72) and C-I1(72) either (almost aligned at $y = 0$). It is that protein with a negative correlation to the total duration of antibiotic therapy in I1(96) patients (*cf.* Fig. 4.8 & 4.9) and a decent fold change in I1(96) vs. C-I1(96).
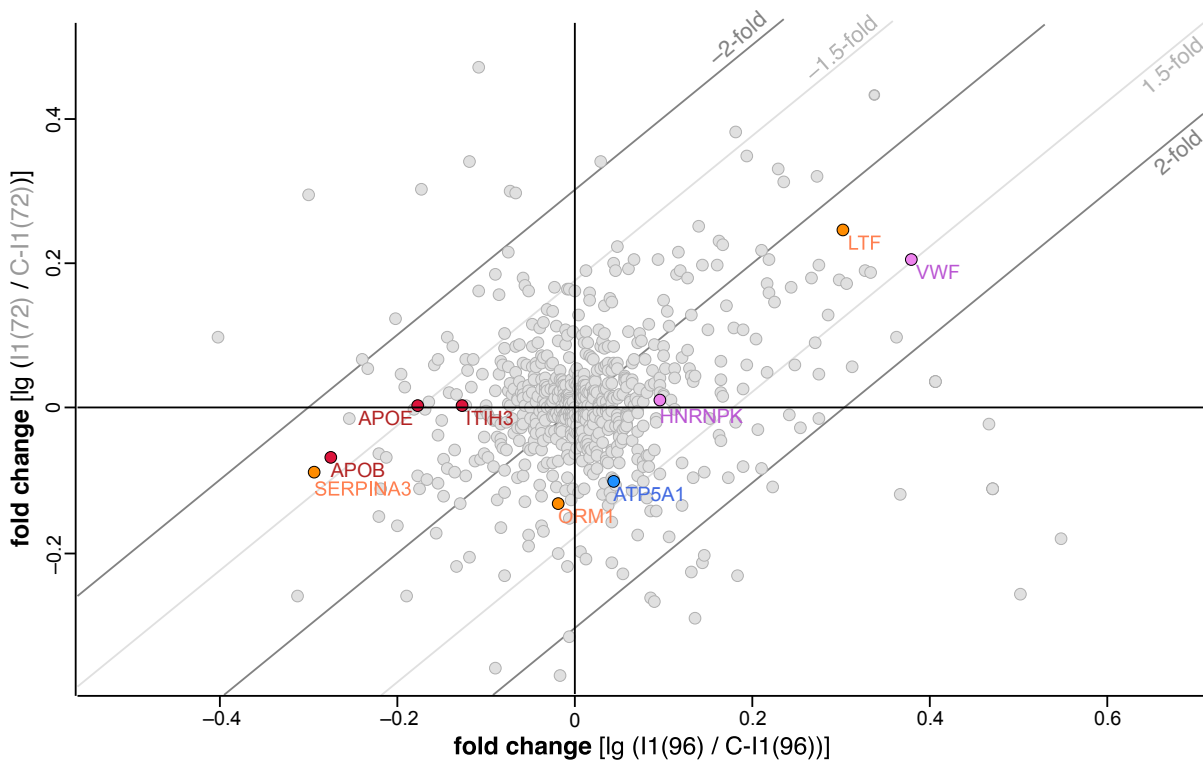

Figure 4.18: Scatter plot I1(96) to C-I1(96) vs. I1(72) to C-I1(72) (centre). x-axis & y-axis = fold change ("difference of means": mean $(\log_{10})$ LFQ of protein a in 1st group minus mean $(\log_{10})$ LFQ of protein a in 2nd group; or as indicated: logarithmic ratio of mean absolute LFQs of protein a in 1st & 2nd group; *i.e.* pos. x- or y-values reflect a higher abundance of protein a in 1st group and vice versa; absolute fold change of protein a equals $10^{x(a)}$ or $10^{y(a)}$, respectively). x- and y-axis both display the fold changes from unpaired t-tests. For a better visual interpretation, the following straight lines are drawn: $y = x \pm 0.301$; $y = x \pm 0.176$; $y = x$ ($lg2 = 0.301$ & $lg1.5 = 0.176$). Red = significant fold change at $q \leq 0.01$ in Fig. 4.11, p. 54. Orange = significant fold change at $p \leq 0.01$ & $q \leq 0.1$ in Fig. 4.11 and a difference of min. factor 2 in Fig. 4.13, p. 58. Blue = significant fold change at $p \leq 0.01$ in Fig. 4.11 and difference of min. factor 2 in Fig. 4.13. Pink = negative correlation with total duration of antibiotic therapy or initial CRP in I1(96), *cf.* Fig. 4.9, p. 52.

In the case of vWF, which shows a negative correlation to the initially measured CRP in I1(96) patients (*cf.* Fig. 4.8 & 4.9), the position in the plot suggests a 1.5-fold higher fold change in I1(96) than in I1(72). The same can be stated for Apo-B100 and SERPINA3, which are proteins with significant dynamics from I1(96) to I2(96) (*cf.* Fig. 4.11, p. 54).

However, ORM1 and ATP5A1 (*cf.* Fig. 4.11 & 4.13) show a higher fold change for I1(72). The fold change of LTF is relatively similar in both groups.

5. DISCUSSION

5.1 Question 1 – "Differences"

In question 1 (Ch. 4.2, p. 43), I compared the blood samples of infected newborns with those of healthy control patients.

Main findings in the severely affected neonates were a significant alteration of S100-P (31-fold increased, $p = 0.001$), ACTR3 (8-fold decreased, $p = 0.003$) and HP (8-fold decreased, $p = 0.0008$). S100-P also correlated positively with the CRP before initiation of antibiotic therapy ($r = 0.73$, $p = 0.03$). Furthermore, hnRNPK showed a negative correlation with the total duration of antibiotic treatment ($r = -0.76$, $p = 0.02$). The haemostatic glycoprotein vWF had, unlike S100-P, a negative correlation with the initial CRP ($r = -0.67$, $p = 0.05$).

Proteins with abundance differences in question 1 may theoretically serve as diagnostic marker in an early phase of neonatal inflammation/infection. Thus, of all three questions the results of question 1 have the most immediate relevance. A major limitation, however, arises from the distribution of blood takings on the postnatal timeline (*cf.* Ch. 5.5). The mean time points were 11 h *pp.* for I1(all) and 40 h for C-I1(all) (*cf.* Fig. 4.1, p. 39). Despite a matched pairing, this discrepancy of nearly 30 h may have lowered significance in a relevant manner and reflects well the practical difficulties of clinical research in neonates (*cf.* Ch. 5.5).

Generally, it should be noted that in all of the presented analyses (questions 1–3) CRP is not contained. Technically, this is due to the chosen minimum of valid values (70 % in at least one of two compared groups, *cf.* Ch. 3.11). Among all samples of the entire infection group (at both time points), CRP was measurable in only less than 70 % of samples. Partially, this may be explained by CRP's late increase in infection. However, even in I2(96) it was measurable in less than the required number of samples. In view of the sufficiently late sampling and the severe course of these patients CRP normally would have been expected to be detectable in I2(96). That it was *not* suggests that, in the majority of cases, it may still have been under a certain detection limit for the present whole blood analyses. Similarly, IL-6 and other interleukines do generally escape a mass spectrometric detection because of their minimal plasma abundance and the problem of dynamic range (*cf.* Ch. 1.2.2). Probably, the absence of PCT is also due to reasons of timing and/or dynamic range.

Volcano plots (p. 44 *et seq.*)   In Fig. 4.4 all available patients with infection in an early stage of disease are compared with their matched control group. For each protein, the fold change between both groups is plotted against the p-value of its respective t-test. Presumably, for the reason that I1(all) comprises patients of diverse conditions (short & long treatment, different antibiotics) and thereby represents an inhomogeneous group, it delivers not a single significant protein.

Fig. 4.5, in contrast, is based only on the group of nine patients who received the same antibiotics in the same dosage for at least 96 h and who were apparently seriously ill. Comparing them with their matched control means comparing two conditions that differ more clearly. Here, the protein S100-P traverses the s_0 significance curve as well as haptoglobin (HP) and actin-related protein 3 (ACTR3). Mean time points of blood taking, however, were in this case similarly unfavourable, namely 8 h *pp.* for I1(96) and 46 h for C-I1(96).

Protein S100-P is a calcium binding protein from the S100 family which is involved in most diverse cellular functions and, to date, foremost known as unspecific tumor marker expressed in several tumor types (Tóthová & Gibadulinová, 2013). More recently, its role in implantation and embryonic development has been reported (Prica *et al.*, 2016). A clear involvement in inflammatory processes has not been described yet. Its highly significant increase in I1(96) compared to its matched control, however, suggests to further evaluate an eventual role as pro-inflammatory agent. A related protein, S100-A12 has recently been proposed as novel biomarker for the evaluation of neonatal sepsis (sensitivity 96.8 %, specificity 93.3 %; Tosson *et al.*, 2019).

ACTR3 is a regulator protein of the cytoskeleton and of cell motility. It is involved in ciliogenesis (Kim *et al.*, 2010) as well as in carcinogenesis and metastasis of several gastrointestinal tumors (Yin *et al.*, 2015; Hu *et al.*, 2021). In brains of aborted foetuses with Down syndrome, a lack of ACTR3 could be observed (Weitzdoerfer *et al.*, 2002). A specific involvement in inflammatory processes, however, has not been reported yet. The measurement of this and many other cellular proteins in our study setting is due to the fact that whole blood was subjected to analysis, not plasma (*cf.* Ch. 3.8). After all, the significant reduction of ACTR3 in I1(96) compared to its matched control is hard to explain. Possibly, it is a consequence of the unequal distribution of venous and capillary blood samples between infection and control groups. This circumstance can be seen in Tab. 4.2, p. 41, and it will be further discussed in Ch. 5.5, p. 80. Since venous blood samples have lower hematocrits than capillary ones and since the former are overrepresented in the infection patients, it can not be ruled out that the reduction of the cellular protein ACTR3 in I1(all) is merely a statistically significant example for a bias towards higher abundances of cellular proteins in the control groups.

HP is an antioxidant plasma protein which is broadly known for binding free haemoglobin from lytic erythrocytes and thus preventing it from unfolding its oxidative activity in a harmful way. Furthermore, it is an acute-phase protein with increased levels in infection, inflam-

mation and malignant diseases (Dobryszycka, 1997). Its anti-inflammatory potential even includes a bacteriostatic function as part of innate immune response (Eaton *et al.*, 1982). More than 40 years ago, it was already hypothesized that elevated levels of HP could be used for an early prediction of neonatal sepsis (Philip & Hewitt, 1980). Although it has never asserted itself against the diagnostic performance of CRP, the interest in HP has not faded. Kalenka *et al.* (2006) reported higher HP-up-regulation in adult sepsis survivors than in non-survivors. And Buhimschi *et al.* (2011) found that intraamniotic infections lead to elevated levels of HP in cord blood, which might be used as an EOS screening biomarker and could be helpful in the selection of patients who are actually in need of early treatment.

The fact that HP is decreased in I1(96) compared to its matched control is opposed to HP's actual role as "positive" acute-phase protein (*i.e.* increase in the acute phase). Most probably, this seemingly contradictory picture is to be explained by a gradual increase of HP over the first seven days of life, presumably in preparation for the required clearance of foetal haemoglobin (Bennike *et al.*, 2020). All control samples from healthy newborns were standardly collected 36–72 h *pp.* (mean for C-I1(96) was 46 h). I1(96) patients, however, were sampled for the first time within 13 h *pp.* (mean 8 h). Therefore, at least in the evaluation of HP a "physiological bias" and the weakness of the (necessary) groupings by "corrected gestational age" becomes manifest. Without close-meshed longitudinal HP reference values from healthy neonates starting at the time of delivery, a use of HP as EOS biomarker will presumably remain doubtful. A second (less probable) explanation for the relative absence of HP in I1(96) could be that it reflects a consumption of HP as a result of its antioxidant and bacteriostatic activity. This assumption is at least suggestively supported by the fact that HP levels in I1(96) patients correlated slightly negatively with their initial CRP ($r = -0.28$, $p = 0.55$) and their total duration of antibiotic therapy ($r = -0.41$, $p = 0.36$; *cf.* Fig. 4.7). Thirdly, the overall decrease of HP in I1(96) might hypothetically reflect a low-expression genotype of HP that contributes to an increased infection susceptibility of the respective neonates. The latter hypothesis has already been pursued in preterm neonates with inconclusive result (Kessel *et al.*, 2016).

Enrichment analysis (p. 48 *et seq.*)   Since the primary 1D-AEA over the fold change of proteins between infection and control group remained without significant term, a major difference in expression patterns between these two groups is apparently not detectable by this analysis method.

A secondary 1D-AEA, however, which was performed for the protein intensities in I1(96) alone, showed at least two significant and relevant terms, "immunoglobulindomain" and "antigen binding". Although such an effect was not detectable by the primary 1D-AEA, Fig. 4.6 demonstrates: firstly, to what extent the respective proteins (annotated with these two terms) are over-represented among all proteins within I1(96) (*cf.* Subfig. A1 & B1); and

secondly, that these very proteins are also factually over-represented in I1(96) compared to C-I1(96), although only in a low-significance area (*cf.* Subfig. A2 & B2, only the central area of the volcano plot with small fold changes and low *p*-values is depicted).

Nonetheless, the meaning of the terms "immunoglobulindomain" and "antigen binding" in this context remains questionable, since they are found in an otherwise very unspecific list of enriched terms (Tab. 4.5). The vast majority of them is simply associated with intra-erythrocytic functions (due to the analysis of whole blood, *cf.* Ch. 5.6, p. 82).

After all, the above findings are barely significant. Though, they may demonstrate a potential diagnostic concept. When proteome profiles are analysed – beyond focusing on solitary biomarkers or multivariate biomarker panels – it is also possible to analyse the behaviour of whole functional protein networks, *i.e.* groups of proteins sharing a function and (if known) the respective annotation term. For instance: In several years, there might be hundreds of proteins whose involvement in "innate immune response" will be proven. Then, diagnosing an EOS hypothetically may not only be possible when established markers as CRP are elevated, but also in situations when these "first-line" markers are (still) ambiguous and, instead, several "background" proteins associated with "innate immune response" are shifted towards inflammation.

Correlation analysis (p. 50 *et seq.*)   Among the selected clinical and anamnestic parameters for Fig. 4.7, the "total duration of antibiotic therapy" and the "value of CRP at the beginning of antibiotic therapy" are certainly the most relevant in the context of question 1. This is why for these two categories in Fig. 4.8 Pearson's r of all 27 proteins was plotted against its respective *p*-value.

S100-P's strong positive correlation with the initial CRP ($r = 0.73$; $p = 0.02$) might underline its potential function as an indicator of EOS. However, its positive correlation with birth weight ($r = 0.76$; $p = 0.02$) and its negative correlation with the interval between amniorrhexis and delivery ($r = -0.98$; $p = \leq 0.001$) may weaken this hypothesis. The former might indicate a confounding effect, the latter is in some way contradictory to the correlation with CRP, since a longer time span after the rupture of membranes is obviously associated with a higher risk of bacterial colonization/infection of the neonate. Nonetheless, it should still be noted that in the hierarchical clustering of Fig. 4.7 S100-P is adjacent to Apo-B100 and matrix metallopeptidase 9. MMP9 is presumably an anti-inflammatory agent (Zhang *et al.*, 2019) and the role of Apo-B100 in innate immune response will be discussed en detail with the findings of question 2.

hnRNPK (heterogeneous nuclear ribonucleoprotein K) is a cellular protein with numerous regulatory functions. It is foremost a DNA/RNA-binding protein that influences transcription, but also proliferation, DNA repair and apoptosis (Xu *et al.*, 2019). hnRNPK has already been associated with both oncogene and tumor-suppressive processes (Wang *et al.*, 2020a). In a mouse model, there was no embryonic survival under bi-allelic deletion of hnRNPK, whereas

mono-allelic deletion led to developmental defects and short postnatal survival (Gallardo *et al.*, 2015). Since it is also involved in the formation of osteoclasts, mutations of hnRNPK in humans are associated with skeletal dysorders and can be found in the Kabuki-like and the Au-Kline syndrome (Au *et al.*, 2018; Dentici *et al.*, 2018). Feng *et al.* (2020) showed that hnRNPK is up-regulated in the NLRP3 inflammasomes of LPS-stimulated macrophages. At the same time, the formation of these inflammasomes (cytoplasmic receptor-like protein complexes) with all its resulting effector molecules (such as Caspase-1, IL-1β & IL-18) is mitigated when hnRNPK is knocked-out. These findings suggest that hnRNPK has, in addition to its many fundamental cellular functions, a pro-inflammatory component.

The volcano plot I1(96) vs. C-I1(96) (*cf.* Fig. 4.5) shows an hnRNPK fold change of 0.09 at $-\lg p = 2.39$. In absolute values, this is a 1.2-fold increase in the diseased group at $p = 0.004$. Simultaneously, there is a negative correlation of the hnRNPK level in the diseased group with the total duration of antibiotic therapy ($r = -0.76$, $p = 0.02$). Interestingly, the hierarchical clustering of Fig. 4.7 shows similar correlations for hnRNPK, HP, C1q and C1s. The latter are two proteins at the beginning of the classical complement pathway. Taking into account the available knowledge about hnRNPK, all of these findings could cautiously be interpreted as follows: i) an increase of hnRNPK as a presumable representative of proinflammatory processes can also be observed in neonates at an early stage of infection; ii) there is, however, an indirect relationship between hnRNPK level and duration of antibiotic therapy; iii) a too small increase of hnRNPK could therefore be interpreted as an expression of inadequate immune response leading to prolonged disease. *In praxi* however, a "positive" biomarker (increasing under the examined condition) with a "negative" correlation to presumed disease severity (here: duration of therapy) would not be optimal.

Von Willebrand factor (vWF) is a protein expressed in endothelial cells, subendothelial connective tissue and megakaryocytes which is primarily known for its pivotal function in haemostasis (Bryckaert *et al.*, 2014). It is an adapter protein that can link platelets to each other (cohesion) and to thrombogenic surfaces such as injured endothelium (adhesion), which constitutes the initiation of primary coagulation. Additionally, vWF is a carrier protein for factor VIII, a component of the intrinsic coagulation pathway and thus secondary coagulation. In the past few decades, the understanding of the linkage between coagulation and inflammation has continuously increased. Among other adhesion proteins, vWF's role as an acute-phase reactant and as potential inflammatory marker has been evaluated in numerous studies with inconsistent conclusions (Reinhart *et al.*, 2002; Paulus *et al.*, 2011). Principally, under pro-inflammatory circumstances vWF levels increase due to reduced cleavage of its multimeric precursors through the plasmatic metalloprotease ADAMTS13. In theory, this makes vWF a valid indicator of systemic inflammation. A correlation with the severity of inflammation and the extent of organ failure could indeed be shown by Claus *et al.* (2010). However, as vWF

generally would not allow a discrimination between septicaemia and specific tissue damage its use as inflammatory marker is seen critically (Paulus *et al.*, 2011; Kawecki *et al.*, 2017).

The volcano plot I1(96) vs. C-I1(96) (*cf.* Fig. 4.5) showed a vWF fold change of 0.38 at $-\lg p$ = 2.50. In absolute values, this is a 2.4-fold increase in the diseased group at $p = 0.003$. This clearly reflects vWF's inflammation-induced accumulation (as described above). Additionally, it clusters together with C1s & C1q in Fig. 4.7. Its negative correlation, however, with the initially measured CRP (r = –0.67; p = 0.05) is hard to explain. Hypothetically, it could be associated with the late increase of CRP and an earlier beginning of re-normalization of vWF levels. Overall, as in the case of hnRNPK the correlation behaves counter-intuitive, in some sense.

5.2   Question 2 – "Dynamics"

In question 2 (Ch. 4.3, p. 53), I examined the blood samples of newborns with severe infection longitudinally over the time points I1(96) & I2(96) (immediately before and ca. 48 h after the first administration of antibiotics, also "t_1" & "t_2").

Main findings were significant dynamics of apolipoproteins, complement and acute-phase proteins towards t_2. Apo-B100, Apo-E and also ITIH3 were all ca. 2-fold lower at t_1 than at t_2 (q = 0.008).

Likewise, the protein abundances of complement proteins and acute-phase proteins, such as SERPIN-A3 & -A1, ORM1 and SAA2–4, were lower at t_1 than at t_2. Among all of these significantly altered proteins SERPIN-A3 and ORM1 reflected the dynamics towards t_2 most accurately, since there was no detectable change of these two proteins between the examined control groups either. S100-A8 and S100-A9 were 2-fold higher at t_1 (p = 0.003 & 0.01).

A strength of the analyses of question 2 is the insight into the inner dynamics of the infection group, (partially) without the need for more or less well-matched controls and the inevitably associated inaccuracies (*cf.* Ch. 5.5, p. 80). These analyses have the potential to validate findings from question 1. At the same time, they may also direct the focus on relevant inflammatory proteins that remain non-significant in a single comparison with a healthy control.

A remaining limitation, on the other hand, is that due to the present study design it is almost impossible to assess with certainty to what extent the observed protein dynamics are specific consequences of the infection or of physiological postnatal adaptation. This would only be possible by a control group with two longitudinally related blood samples each, which would have to be matchable to the infection group in their cGA and their time interval. Since such a control group of healthy newborns with an early and a late blood sample each is, indeed, impossible to acquire (*cf.* Ch. 4.3.3, p. 58), the distinction between infection-

and adaptation-associated proteome alterations could only be approached in a suboptimal manner in Ch. 4.3.3 (which will still be discussed *en detail*).

Mean time points of blood taking were 8 h *pp.* for I1(96) (t_1) and 55 h for I2(96) (t_2).

Volcano plot (p. 54 *et seq.*)   Notably, two of three proteins which are relatively decreased at t_1 (or increased at t_2, respectively) with the highest significance are apolipoproteins (Apo-E & Apo-B100, $q < 0.01$, marked red in Fig. 4.11). Two further apolipoproteins (Apo-A4 & Apo-C3) are less abundant at t_1 than at t_2 with medium and low significance. Although the subsequent 1D-AEA does not reveal a systematic shift of terms associated with lipid metabolism (*cf.* Tab. 4.7, p. 56), this accumulation is notable.

In the last decades, the role of lipid metabolism in sepsis attracted increasing attention. Lipid components of bacterial cell walls, such as lipopolysaccharide (LPS, from Gram-negative bacteria) and lipoteichoic acid (LTA, from Gram-positive bacteria), are essential in the pathogenesis of excessive inflammation and sepsis (van der Poll *et al.*, 2017). These pathogen-associated lipids (PAL) can be bound and cleared by lipoproteins in the course of their normal function as lipid transporters (usually for triglycerides and cholesterols). Meanwhile, this mechanism of detoxification is understood as part of innate immune defence (Trinder *et al.*, 2019). During inflammation and sepsis, however, for unclear reasons a decrease of many lipoprotein species can be observed, whereas triglycerides are increased due to a diminished lipoprotein lipase activity (Trinder *et al.*, 2019). Buhimschi *et al.* (2011) reported significant down-regulations of Apo-A4, Apo-E and Apo-H in the cord blood of neonates with culture-confirmed EOS, too. Ng *et al.* (2010) even proposed a diagnostic score for NEC and LOS in preterm infants, that is based on (decreased) pro-apolipoprotein C2 and (increased) serum amyloid A. On the other hand, an LPS- and TNF-α-dependent upregulation of Apo-B100 was shown for rodent hepatocytes (Bartolomé *et al.*, 2007, 2008). Sigel *et al.* (2012) reported Apo-B100 to be a potent inhibitor of cytokine release from murine and human immune cells stimulated by staphylococcal LTA (*S. aureus*). In summary, these findings may suggest that apolipoproteins and Apo-B100, in particular, have a protective function against PAL-induced excessive inflammation and that this protection can, however, be compromised in the early phase of sepsis (for unknown reasons).

In the present volcano plot, the increase of apolipoproteins from I1(96) to I2(96) presumably reflects an overcoming of this very initial phase of infection and a beginning decline of the inflammatory response under antibiotic therapy. In view of the evidence above and other findings in this thesis, a physiological or nutrition-associated increase can rather be ruled out. Namely, Apo-B100 and Apo-C3 are also clearly ($p < 0.01$) decreased in I1(96) compared to its matched control group (*cf.* Fig. 4.5, p. 46). Hypothetically, Apo-B100 (2-fold decreased, $p = 0.00002$) would have crossed the significance threshold of Fig. 4.5 in bigger sample sizes. Even in the generally insignificant comparison of I1(all) with its matched control (*cf.* Tab. 4.3, p. 45)

Apo-B100 is 1.4-fold decreased at $p = 0.01$. In summary, these previous findings may support the assumption that the volcano plot I1(96) vs. I2(96) depicts an increase of apolipoproteins from initially (pathologically) reduced levels. In addition, Fig. 4.13 on p. 58 (a figure which will still be discussed en detail) shows that the two most significantly increasing apolipoproteins (Apo-B100 and Apo-E) have a positive dynamics towards t_2 that is more than 1.5-fold stronger than that of the respective controls, affirming the specifity of this apolipoprotein dynamics for the infection groups.

After all, the present data from neonatal patients strongly support the significant role of (apo-)lipoproteins in inflammation and sepsis. A further examination of Apo-B100 (and its upstream regulators) seems to be most worthwhile with respect to its diagnostic (or even therapeutic) potential.

A second key finding of the volcano plot I1(96) vs. I2(96) is the increase of several acute-phase proteins towards t_2, such as alpha-1-antichymotrypsin (SERPIN-A3), alpha-1-antitrypsin (SERPIN-A1), alpha-1-acid-glycoprotein or "orosomucoid" (ORM1), fibrinogen (FGA & FGB) and haptoglobin (HP). Major acute-phase reactants which are also increased are the serum amyloid A proteins (SAA2-4). Elevated levels of the latter are associated with acute and chronic inflammation as well as with malignant conditions. Furthermore, they are agents of the innate immune response through opsonization, toxin inactivation, bactericidal/fungicidal/virustatic effects and leukocyte recruitment (Salama *et al.*, 2021). Referring to alpha-1-anti-(chymo)trypsin, Muk *et al.* (2019) showed a linear increase of serpins over 24 h in the blood plasma of preterm piglets infected with *S. epidermidis*. This linearity may suggest serpins as supportive markers in the early phase of an infection. As already discussed in Ch. 5.1, CRP is not contained in Fig. 4.11 although it would have been expected to be sufficiently detectable at least in I2(96) (*cf.* Ch. 3.11). The fact that it was *not* suggests that even in most cases of I2(96) CRP may still have been under the detection limit of whole blood proteomics.

Among those three proteins which are increased towards t_2 with the highest significance ($q < 0.01$, marked red in Fig. 4.11) is also ITIH3, the inter-alpha-trypsin inhibitor heavy chain H3. It belongs to a family of five variants (ITIH1-5) which can be found in the extracellular matrix and in blood plasma. ITIH3 is supposed to have a stabilising effect on the extracellular matrix (Zhuo *et al.*, 2004) and furthermore tumorsuppressive properties (Paris *et al.*, 2002). It was suggested to be an early biomarker for gastric cancer (Chong *et al.*, 2010). Notably, Muk *et al.* (2019) reported an increase of ITIH4 and a decrease of ITIH1 in the blood plasma of preterm piglets infected with *S. epidermidis* after 24 h. ITIH4 has already been characterized as an acute-phase reactant (Piñeiro *et al.*, 1999). An increase of ITIH1 alongside SERPIN-A1 and SERPIN-A3 and other acute-phase proteins is documented for mice with focal segmental glomerulosclerosis (Bukosza *et al.*, 2020). A connection between ITIH3 and the acute-phase – as visible in Fig. 4.11 – has been outstanding, so far.

Finally, the 2-fold decrease of Protein S100-A8 and S100-A9 towards t_2 should be high-lighted. These mostly phagocytic proteins can be counted to the damage-associated molecular patterns (DAMPs or "alarmins") and are strong promoters of inflammation (Vogl *et al.*, 2007). The S100-A8/A9 dimer is also known as "calprotectin" and has long been used in fae-cal samples as a diagnostic and monitoring biomarker for inflammatory bowel disease, both in adults and in children (Ayling & Kok, 2018). In a proteomic examination of amniotic fluid from women with intra-amniotic infection, the level of S100-A8 correlated with a subsequent development of EOS (Buhimschi & Buhimschi, 2010). Moreover, Pirr *et al.* (2021) claimed to have validated serum S100-A8/A9 at delivery as the first predictive biomarker for LOS in preterm infants. Their findings suggested that high perinatal S100-A8/A9 (foremost due to labour-associated stress) is protective against a later occurrence of neonatal sepsis. Thus, measurements of calprotectin were said to support early decisions on a beginning of empir-ical antibiotic therapy. The negative predictive value of postnatal S100-A8/A9 levels above 1000 ng/ml was 0.95 and, in this concern, was supposed to be superior to IL-6 (individual baselines) and CRP (late increase). In the present volcano plot, the decrease of S100-A8/A9 (medium and low significance) may be interpreted as an early indicator for an effective an-tibiotic treatment and for a decline of the overall inflammation level (prior to a decrease of the classical acute-phase reactants). In Fig. 4.13, p. 58, S100-A8/A9 have a more than 1.5-fold stronger dynamics in the infection group than in the respective controls, which makes it at least unlikely that the decrease over time is a merely physiological postnatal decline. How-ever, it should be remembered that – in contrast to the study setting of Pirr *et al.* – the patients of I1/I2(96) are both late preterm and term neonates with a clinical EOS (not LOS). Finally, it must also be critically noted that there is no significant difference in calprotectin levels be-tween I1(96) and its matched control (*cf.* Fig. 4.5, p. 46).

Enrichment analysis (p. 56 *et seq.*)   Unlike in question 1, the 1D-AEA conducted over the fold change between I1(96) and I2(96) (*cf.* Fig. 4.11, p. 54) delivered a remarkably specific pic-ture. With one exception, all significant terms of Tab. 4.7 were enriched towards t_2, indicating generally higher abundances of the associated proteins at t_2 than at t_1. Obviously, I1(96) and I2(96) display an interval which is – despite the beginning of antibiotic therapy – character-ized by an increasing infection dynamics and the organism's response to it. The majority of the enriched terms are namely such that are associated with complement, acute phase and haemostasis. Partly, related terms appear simultaneously in different annotation categories (*cf.* "Type" column), which may underline the validity of the enrichment.

The appearance of the KEGG name "Staphylococcus aureus infection" shall be empha-sized by name as it is the most direct connection to innate immune defence against a bacte-rial pathogen. In the present patients, "Staphylococcus aureus infection" comprises foremost elements of the different complement pathways (such as C1, MBL2 and C3; *cf.* Fig. 4.12, Sub-

fig. C2). Therefore, there is a large intersection with the proteins of the KEGG name "complement and coagulation cascades" (*cf.* Fig. 4.12, Subfig. A2). Both enriched terms reveal the systemic increase of complement proteins towards t_2, which is in some way "hidden" under the significance thresholds of the respective volcano plot (Fig. 4.11, p. 54). Tab. 4.6, p. 55, barely contains complement proteins although the general trend of their increase is obvious in the 1D-AEA. This may, again, underline the general methodical and possibly even diagnostic value of 1D-AEA. Of course, the fact alone that complement levels (relatively) rise from t_1 to t_2 does not reveal this dynamic's (absolute) relation to normal levels of complement protein in neonates. However, the comparison of I1(96) and C-I1(96) in Fig. 4.5, p. 46, has already shown diminished levels of C1s and C1q in the infection group. Hence, the increase of complement between I1(96) and I2(96) might also be interpreted as a normalization (or escalation) of initially reduced levels (in the sense of an early complement consumption during innate immune response and a consecutive recovery of complement levels).

The increase of several well-known acute-phase reactants towards t_2 has already been discussed above (*cf.* Fig. 4.11 and Fig. 4.12, Subfig. B2). Nonetheless, the enrichment of the GOBP name "Acute-phase response" is additional proof of their systemic increase.

Scatter plot (p. 58 *et seq.*)   Fig. 4.13 primarily aims at a critical questioning of the significant proteins from volcano plot I1(96) vs. I2(96) (Fig. 4.11, p. 54). Theoretically, such proteins that align in the present scatter plot at $y = x$ are most likely to do so due to physiological dynamics. Thereby, they could be excluded from further considerations, in spite of seemingly "significant" dynamics in the corresponding volcano plot. Inversely, an alignment of proteins at $y = 0$ may underline their possible significance, especially when $x \geq \pm 0.301$. The latter would, namely, indicate a presumable absence of any physiological dynamics in the control groups ($y = 0$) with an at least 2-fold ($x \geq \pm 0.301$) presumably pathological dynamics in the infection group for this specific protein.

However, there are several important limitations to this scatter plot. First, as described in Ch. 4.3.3, p. 58, the control are two unrelated groups of samples, taken as surrogate for a control group of healthy newborns with two blood samples each (which is absolutely not available, in reality). However, the inter-individual variability of protein abundances (as introduced in Ch. 1.2.2) gives reason to question the validity of such a "pseudo-longitudinal" control group (composed of two unrelated samples). Moreover, these control groups (Cnm = "control non-matched") could hardly be matched to the infection group: not at all in their cGA at the time of blood taking, and merely sub-optimally with respect to the mean postnatal time of blood taking. This was 8 h *pp.* for I1(96) and 55 h for I2(96). In contrast, it was 39 h *pp.* for Cnm-I1(96) and 53 h for Cnm-I2(96). Thus, infection and control group also cover different intervals, which additionally restricts their comparability. Weak dynamics in the control group might actually be more pronounced than depicted. Completely absent dynamics, however,

can hypothetically be taken as such. In exchange for a suitable matching, at least relatively big sample sizes for the two unrelated control groups could be ensured ($n = 19$). Finally, it has also to be noted that the t-test for I1(96) vs. I2(96) is repeated as an *unpaired* (two-sided) two-sample t-test to ensure comparability to the controls. It is not identical with the fold change from Fig. 4.11, p. 54, which is based on a *paired* (two-sided) two-sample t-test. Despite all these limitations, the present type of analysis is informative and its results are worth discussing.

Generally, numerous proteins of interest from Fig. 4.11 are part of the diffuse cloud in the centre of the scatter plot. A clear alignment of these proteins at $y = x$ is not observable. The majority of proteins fulfilling $y = x \pm 0.301$ are not significant in Fig. 4.11 (green in both figures).

Alpha-1-antichymotrypsin (SERPIN-A3) and alpha-1-acid-glycoprotein (ORM1), however, have a notable increase from t_1 to t_2 (4-fold & 2-fold, respectively) with almost no measurable dynamics in the control groups. This sharp contrast may confirm their specific role as acute-phase reactants with a possibly high relevance in neonatal infections. Since both of them are, however, not included in the significant proteins of the volcano plot I1(96) vs. C-I1(96) (*cf.* Fig. 4.5, p. 46), their potential for an early distinction between diseased and non-diseased neonates seems to be restricted.

The increase of Apo-E, Apo-B100 and ITIH3 (the three most significantly altered proteins from Fig. 4.11, red in both figures) is more than 1.5-fold stronger in the diseased group than in the control groups. Although the relations are not as evident as for SERPIN-A3 and ORM1 and a physiological postnatal dynamics for these proteins cannot be totally ruled out, the present picture additionally underlines the relevance of Apo-E, Apo-B100 and ITIH3 and potentially validates them as markers. Especially Apo-B100 is of interest, here, since it also shows a nearly significant negative fold change in the volcano plot I1(96) vs. C-I1(96) (2-fold decrease, $p = 0.00002$; *cf.* Fig. 4.5, p. 46). A certain decrease of Apo-B100 is even visible in the more unspecific volcano plot I1(all) vs. C-I1(all) (1.4-fold decreased, $p = 0.01$; *cf.* Fig. 4.4, p. 44). The reasons for a suppression of apolipoproteins in inflammation may still be elusive. However, even if (hypothetically) all neonates showed increasing levels of Apo-B100 immediately after birth (due to recovery from birth stress) all present data still suggest that particularly low levels of Apo-B100 could be distinctive for an early-onset infection/sepsis, and that Apo-B100 most significantly increases in the examined patients in the course of their antibiotic therapy.

Lactotransferrin (LTF) is an iron-binding protein foremost in exocrine secretions such as saliva or milk (especially in colostrum; Sanchez *et al.*, 1992), but it is also to be found in granules from neutrophils. For its both bacteriostatic and bactericidal functions right at the scene of inflammation it is often counted as an element of innate immune response (Legrand *et al.*, 2005) and as a non-hepatocytic acute-phase reactant (Schrödl *et al.*, 2016). The depicted more than 2-fold dynamics of LTF in the infection group can be harmonized well with the nearly significant ($p = 0.001$) 2-fold positive change in the volcano plot I1(96) vs. C-I1(96) (*cf.* Fig. 4.5).

Overall decreasing levels of LTF between t_1 & t_2 could – besides recovery from birth stress – also be attributed to the first feeding with LTF-rich colostrum right after delivery, and the continuous transformation of colostrum to more mature forms of breast milk with decreasing LTF-concentrations. In the infection group, higher postnatal levels of inflammation and their effective treatment may cause the 2-fold stronger decrease displayed in the present scatter plot (Fig. 4.13).

ATP synthase subunit alpha (ATP5A1) is a mitochondrial membrane bound protein that is involved in ATP production and thus provision of cellular energy. Recently, it was also reported to bind enterobactin (Ent), a bacterial siderophore (Qi & Han, 2018). Siderophores are iron-binding compounds, secreted by either bacteria or fungi, that scavenge Fe^{3+} from hosts to appropriate them for their own supply. ATP5A1, however, is supposed to counteract this disadvantageous mechanism and to enhance cellular iron haemostasis by simply taking up Ent-Fe^{3+} complexes and utilizing them in the mitochondrial metabolism. Qi & Han (2018) presented the theory of a constant "thug of war" for iron between host and commensals/pathogens, in which ATP5A1 is an effective measure on the host side. Notably, Fig. 4.13 shows relatively specific dynamics for ATP5A1 in the infection group (1.7-fold increase in I1(96), *cf.* Fig. 4.11, p. 54). Whereas Qi & Han see the role of ATP5A1 foremost as a counterweight to gastrointestinal commensal bacteria, the present data at least raise the question if there may be an immune response associated up-regulation as well. When considering the importance of iron metabolism to the immune system, it should also be noted that lipocalin (LCN2) is elevated in both comparisons of infection to control groups, too (*cf.* Tab. 4.3 & 4.4, p. 45 *et seq.*). LCN2 is able to bind siderophores and thereby block the pathogenic uptake (Xiao *et al.*, 2017), which makes it a bacteriostatic agent of innate immune response. An evaluation of a potential co-regulation of ATP5A1 & LCN2 and their eventually joint involvement in innate immune response may be worthwhile.

At last, it should be noted that calprotectin (S100-A8/A9), which was *not* significant in the second volcano plot of question 1 (Fig. 4.5, p. 46) but significant in the volcano plot of question 2 (Fig. 4.11, p. 54), has more than 1.5-fold dynamics in Fig. 4.13 in the infection group compared to the controls. Furthermore, S100-P, which was on the other hand significant in question 1 but *not* in question 2, does show more than 2-fold dynamics in Fig. 4.13 as well. In addition to all what has been said so far about these three members of the S100 family, the present scatter plot again hints at their notable dynamics in adaptation and immune response, and strengthens the hypothesis of their diagnostic potential in neonates.

5.3   Question 3 – "Distinction"

In question 3 (Ch. 4.4, p. 59), I compared the pre-antibiotic blood samples of newborns who were treated for more than four days, *i.e.* I1(96), to those of newborns who were treated for

less than three days, *i.e.* I1(72). As laid out in the introduction, a major challenge in neonatology is the early and correct identification of infected newborns. Ensuring the greatest possible patient safety almost inevitably leads to a worldwide overuse of antibiotics in newborns. This is due to the fact that today's diagnostic resources often cannot provide sufficient safety for an early decision *against* an antibiotic treatment. With respect, however, to the mid- and long-term health implications which early exposure to antibiotics may have (*cf.* Ch. 1.1.7), searching for reliable means of distinction between "seemingly" ill and "certainly" ill neonates is highly relevant. Principally, one could object that such a diagnostic criterion would best be found by comparing very distinct groups, such as "diseased" versus "non-diseased" patients (question 1, Ch. 4.2). Nonetheless, question 3 originated from the interest if high-resolution proteomes gained by mass spectrometry might yet uncover a distinguishing feature for these rather similar groups of "seemingly" and "certainly" ill neonates. Presumably, among the patients of I1(72) there are such whose severe disease progression could be prevented effectively in its very beginnings and whose short antibiotic therapy was totally adequate. Equally, there may be such patients who were (retrospectively) treated as a precautionary measure and whose treatment would not have been initiated in an utopian setting with flawless diagnostics. This precondition, however, of a *per se* inhomogeneous group and the small sample sizes may be the reasons for the relatively fruitless results of question 3. At least through their mean time points of blood taking at 8 h *pp.* for I1(96) and 15 h for I1(72) both groups are relatively well comparable. The later blood taking in I1(72) may also reflect the lower clinical urgency and the, ultimately, less severe course of disease.

Volcano plot (p. 60 *et seq.*)   Reduced levels of ACTR3 in I1(96) compared to I1(72) are in line with the picture of volcano plot I1(96) vs. C-I1(96) (Fig. 4.5, p. 46), which shows a significant reduction of ACTR3 in I1(96), too. Hypothetically, with bigger sample sizes ACTR3 would also cross the s_0-corrected significance threshold of Fig. 4.15. Moreover, Fig. 4.17 (the next figure to discuss) does also support this observation of a relevant reduction of ACTR3 in I1(96), which is visible for I1(72), too, but clearly more pronounced in I1(96). Although one can only speculate about ACTR3's actual role in inflammatory processes (*cf.* Ch. 5.1) and if it would stand up to a more targeted validation as infection marker, the statistical tests of question 3 suggest, in any case, a gradual reduction of ACTR3 according to the severity of disease (derived from the duration of treatment, 72 h or 96 h).

Fig. 4.16, p. 61, highlights several aspects that attract attention when the major findings from question 1 & 2 are put into context.

Fig. 4.16 **(A)** shows that all five proteins of interest from question 1 (with a significant fold change or a remarkable correlation) behave similarly as they did in Fig. 4.5, p. 46, and in Fig. 4.9, p. 52, respectively. While these five proteins are principally in- or decreased in the same way, the *p*-values for their fold changes are much lower. Comparing I1(96) with

I1(72) thus delivers a similar but "dampened" picture of what I1(96) vs. C-I1(96) shows. The distinctive features are less pronounced (especially in their significance) but they can still be traced.

Fig. 4.16 **(B)** illustrates the position of seven proteins with significant fold changes from question 2 (*cf.* Fig. 4.11, p. 54, and Fig. 4.13, p. 58). Therefore, Apo-E, Apo-B100 and ITIH3 do not only increase over time from I1(96) to I2(96) but they are also decreased in I1(96) compared to I1(72). This might additionally support the hypothesis of their role as negative biomarkers in the early phase of infection. Remember that Apo-B100 is nearly significantly reduced in I1(96) vs. C-I1(96), too (*cf.* Fig. 4.5, p. 46). Overall, similarly to subfigure (A), trends from previous findings can be reconstructed in subfigure (B), but again in an area of statistical non-significance.

Fig. 4.16 **(C)** contains an alike statement of what has been described for subfigure (B), merely referring to those proteins that are associated with the significant annotation terms from question 2 ("complement and coagulation cascades", "acute-phase response" and "staphylococcus aureus infection"). Proteins with these terms (in majority complement proteins) are decreased in I1(96) compared to I2(96) (*cf.* Fig. 4.12, p. 57), and so they are in I1(96) compared to I1(72). Most probably, this is an expression for a more pronounced consumption of complement proteins in the more severely affected group.

Fig. 4.16 **(D)**, in contrast, shows a relative increase of such proteins in I1(96) that are annotated with the UniProt keyword "antimicrobial". Although visually notable, this finding can not be proven by a 1D-AEA. Nevertheless it may represent a beginning expression of innate immune proteins that is slightly more pronounced in I1(96) than it is in I1(72).

Scatter plot (p. 62 *et seq.*) In Fig. 4.17 the fold changes of two unpaired t-tests are plotted against each other, that of I1(96) vs. C-I1(96) on the x-axis against that of I1(72) vs. C-I1(72) on the y-axis. While Fig. 4.15, p. 60, examines the direct difference between the infection groups, this scatter plot examines the difference between their differences to control. Mean time points of blood taking were: 8 h *pp.* for I1(96), 46 h for C-I1(96), 15 h for I1(72) and 38 h for C-I1(72). The principal temporal discrepancy between infection and control groups remains (*cf.* Ch. 5.5). It is, however, similarly pronounced in both cases, which may be favourable for comparability.

At first sight, Fig. 4.17 presents an unspecific picture due to its central point cloud. However, the roughly implied distribution of those 12 highlighted proteins, that have already proven a prominent role in question 1 & 2, is notable in a certain way. They are foremost to be found in the I. or III. quadrant ($x/y = +/+$ or $-/-$) and the majority of them is additionally defined by $|x| > |y|$. This means nothing else than generally more pronounced fold changes in the more severely affected group of patients, *i.e.* I1(96). In this sense, Fig. 4.17 supports the above impression from Fig. 4.16 that major trends, which have been described so far for I1(96) can be discerned for I1(72) as well, but in a weakened manner. Therefore, this scatter plot may

be regarded as a further validation of the previous results, since the respective proteins seem to be specific for both infection groups while they are rather more pronounced in the more severely affected one.

Proteins such as S100-P, HP or ACTR3, whose fold changes are significant in I1(96) vs. C-I1(96) (*cf.* Fig. 4.5, p. 46), are developed in the same direction in both infection groups, but to a much lesser extent in I1(72). This comprehensible grading may additionally underline the potential usefulness of S100-P, HP and ACTR3 in the distinction between diseased and non-diseased patients. The same can be said about Apo-B100 (nearly significant in Fig. 4.5, p. 46) and SERPIN-A3. Apo-E and ITIH3 are obviously more specific for I1(96) compared to I1(72) since there is hardly any deviation on the y-axis of Fig. 4.18. Overall, this plot may emphasize once more the importance of a further exploration of lipid metabolism in infection and sepsis.

5.4   Study design

As a result of the ongoing technological advances in mass spectrometry, proteomic approaches are more and more conceptualized „hypothesis-free" (*cf.* Ch. 1.2.3). Unbiased questions (and their answers) are supposed to emerge from the proteomic data rather than the other way around. The MUNICH-PreTCl cohort study was also designed according to this paradigm. Its broad and principally hypothesis-free approach may be an attractive aspect for a proteomics project of its dimension. However, for a clearly hypothesis-driven subgroup analysis and a "nested case-control study" as the present work, MUNICH-PreTCl's extensive design had some disadvantages. Time and resources were tied up by the collection of hundreds of informed consents, filter cards, plasma samples, questionnaires and a plethora of clinical data. In its practical realization, the specific study design of this thesis would yet have required a more determined focus on its own challenges, such as including diseased patients in unexpected and time-critical situations ("I1") and ensuring their follow-up ("I2"). A separate and tailor-made workflow – without the parallel challenge of sampling hundreds of mainly healthy newborns at their metabolic screening (36–72 h *pp.*) – would have led to more homogeneous intervals between I1 & I2 and to fewer missed patients with neonatal infection/sepsis. Within the relevant study period (stage 2, *cf.* Ch. 3.4, p. 33), 16 out of 48 term infants with a bacterial sepsis (coded as P36.9 in ICD-10) could be included in the infection group. The proportions for the respective late preterm infants (35+0 to 36+9) are assumed to be similar.

Furthermore, one might disagree with the formation of a hand-selected patient collective for the pilot study (*cf.* Ch. 3.4). Strictly speaking, the objection of a selection bias may be justified for all groups apart from the infection group itself since all eligible patients from stage 2 (neonatal infection, gestational age $\geq$ 35+0, two samples each) were indeed included. The actual objection certainly is the above: that the "infection workflow" of the overall study pro-

tocol was not sufficiently implemented into clinical routine to ensure *a priori* higher numbers of eligible patients in the pilot study.

Hypothetically, a key element of the present study design, its longitudinal character, would gain further analytical depth if cord blood and a maternal blood sample prior to delivery could be included in the proteomic examination as well. The time axis would thereby be prolonged by two well-defined time points, and insights into maternal influences on the neonatal blood proteome would be possible. At the same time, the still higher complexity of such a study should probably be balanced by further restricting the eligible gestational ages or other characteristics. In any case, many promising and more developed approaches are imaginable.

5.5 Patient collective

Generally, the recruited patient collective can be considered suitable for the analyses performed. Tab. 4.2 on p. 41 and the additional figures on p. 93 reveal a primarily homogeneous collective of affected patients and controls. All groupings could be performed with appropriately comparable distributions of (corrected) gestational ages and birth weight. The principal exclusion of preterm infants $< 35{+}0$ was certainly crucial for assuring homogeneity and comparability within the thesis collective, not least with respect to the developmental and immunologic differences between early and late gestational ages. For the whole infection group, the intervals between the blood samplings I1 and I2 could effectively be narrowed to a consistent corridor (median interval of 48 hours, *cf.* Fig. 4.1, p. 39). The APGAR scores of all patients suggest consistently uncomplicated (immediate) postnatal adaptations. The infection group is also consistent insofar as 16 out of 20 patients were treated with exactly the same scheme of antibiotic therapy. The further separation of these patients in two groups according to the duration of their antibiotic treatment (as surrogate parameter for severity of disease) may be considered as reasonable since the initial levels of CRP, IL-6 and platelets differ clearly between these two groups. Comprehensibly, the intervals between rupture of membranes and delivery and the duration of delivery itself are significantly longer in the infection groups than in the control groups. Moreover, the maternal ages and BMIs are relatively higher, too.

Certainly, a central limitation of this thesis is sample size ($n = 20$, *cf.* Ch. 3.6, p. 33). Some findings at the borderline to significance might have been confirmed (or relativized) in a more targeted setting with higher statistical power. Nonetheless, the present group sizes from 7 to 20 individuals (*cf.* Tab. 4.2, p. 41) are not unusual and can also be found in other studies at the intersection of immunology and clinical proteomics, especially when neonates are involved (Stewart *et al.*, 2016; Chatziioannou *et al.*, 2018; Suski *et al.*, 2018a,b). Furthermore, it has to be noted that female & male newborns and capillary & venous blood samples are unequally distributed between infection & control groups. The majority of males in the infection groups results from the higher priority of cGA for the process of group pairing and the well-known

higher susceptibility of male newborns for infections (Simonsen *et al.*, 2014). The majority of venous blood samples in the infection groups was probably a result of intensified care for ill-appearing newborns, and of the fact that in the metabolic screening healthy newborns are rather sampled by heel pricks than by venipunctures. However, since venous blood samples have lower hematocrits than capillary ones (Kayiran *et al.*, 2003) a systemic bias with respect to the abundance of cellular and plasma proteins in the infection and control groups cannot be ruled out, and if so the size of this confounding effect remains elusive.

Although it reflects well the clinical routine with late preterm and term infants, it has to be outlined, too, that none of the included patients' blood cultures turned positive. Probably, in a group with mainly culture-proven infections the associated higher bacterial loads would have induced stronger inflammatory responses. Under such conditions, the observed proteomic effects would possibly have been more significant. Furthermore, to simply speak of an EOS without proof of bacteraemia would not be quite accurate, even though definitions are vague (*cf.* Ch. 1.1.2) and in all included cases onset of symptoms and start of antibiotic therapy were within the first 36 h *pp.* Therefore, mainly the description as "non-culture-proven" or simply as "clinical" infection was used.

At last, a fundamental problem in terms of grouping, which could not be solved in a satisfying way, is the problem of two time axes: the postconceptional and the postnatal one. From a strictly physiological point of view, the blood samples of two newborns are best comparable if both patients match in their gestational age *and* both samples match in their postnatal time point. In reality, however, hardly any clinical experiment can perfectly respect both of these aspects. In the present work, this issue was faced by pairing all specimens according to the patients' "corrected" gestational ages (cGA) at the time of blood sampling. This achieves at least a theoretical alignment of the postconceptional time axes. An alignment of the postnatal time axes, however, was not possible at all when the early blood samples of diseased neonates (ca. 12 h *pp.*) were compared to healthy controls, since control samples were taken as extra drops in the course of the standardized metabolic newborn screening (36–72 h *pp.*). Since there is no need (and no justification) for an earlier blood sampling on healthy neonates after an uncomplicated delivery, there cannot be a perfect control group for infected neonates in the presented study design, either. Technically, all group pairings could be performed with appropriate accuracy, as far as corrected gestational age and birth weight are concerned (*cf.* additional figures in appendix). The temporal gap, however, between I1 and C1 blood samples on the postnatal timeline, and the yet poorly understood extent of early postnatal proteome changes may have lowered the significance of several results.

5.6 Methods

The usage of filter cards as conserving medium for all blood samples was an effective means for a best possible implementation of the study workflow into clinical routine and for achieving the largest possible group sizes. Due to the daily metabolic screenings on neonatal wards, nurses are experienced in the application of filter cards, which surely increased the consistency of sample quality. On the other hand, whole blood (in contrast to plasma or serum) diminishes the analytical depth of MS measurements even under the most sophisticated sample preparation. The proteomic insights are partly blurred by the abundance of cellular proteins. Functional plasma proteins which should actually be measurable may be overlayed by the sheer mass of erythrocytic proteins (as described for the general problem of "dynamic range", *cf.* Ch. 1.2.2). Nonetheless, using filter cards for proteomic workflows generally represents a suitable compromise between pragmatism and quality.

The sample preparation techniques and the MS workflow were largely based on the methods described in Geyer *et al.* (2016a). They are characterized by a high degree of robustness, automation and speed without loss of quality. Up to 1000 proteins are said to be measurable in minimal amounts of plasma or prepared whole blood (5 µl) in no time ("3 h from blood droplet to results"). For non-cell-culture-derived samples from clinical routine these numbers are indeed remarkable. All in all, at the time of analysing the pilot study samples in early 2019, the applied methods – including *Spectronaut*™ and *Perseus*, whose capacities are described and cited in dozens of biostatistical papers – were state-of-the-art.

On the level of biostatistics itself it has to be pointed out again that a large portion of proteins, which are highlighted in the volcano plots of Ch. 4, are not declared "significant" by *Perseus* itself. As defined in Ch. 3.11, p. 35, only those proteins in volcano plots which hold up to all of *Perseus'* correction mechanisms for multiple hypothesis testing would be called genuinely "significant", but t-test results with a p-value ≤ 0.01 ($-\lg p \geq 2$, respectively) would still be reported. In this sense, some of the discussed findings have to be seen as indicators rather than significant results.

5.7 Proteomics in neonatology

Ultimately, the question arises if an implementation of proteomic methods into neonatological routine may be imaginable in a near future, and if so, under which conditions.

Principally, the minimal sample amounts required for MS-based analyses fit well with the important condition of sparing the vulnerable organism of a newborn. Concerning a concrete application in the field of neonatal infection diagnostics, it is obvious, however, that any new biomarker or technique has to compete with the firmly established and so far unchallenged diagnostic triumvirate of CRP, IL-6, and PCT. The benefit of a novel technique would have

to be overwhelming if these easily determinable, rapidly available and relatively inexpensive markers were to be replaced, irrespective of their existing limitations. It should also be remembered that interleukines generally escape a mass spectrometric detection due to their minimal plasma abundance and the problem of dynamic range (*cf.* Ch. 1.2.2). Thus, IL-6 as an established and very early infection marker is not available in MS-based diagnostics and research. A rapidly progressing and life-threatening condition as sepsis also illustrates that even the most sophisticated and most accurate biomarker panel will hardly be applicable in clinical practice if its determination takes longer than 30 minutes. Considering this timely dimension, MS-based proteomic methods still seem to have a long way until an application in acute diagnostics will be feasible. Nonetheless, a translation of MS-acquired proteomic evidence into more easily available POCT techniques may be realistic at an earlier point. A reasonable example would be an MS-based validation of a set of 10–20 predictive EOS biomarkers in cord blood which could be measured in a classical immunoassay- or aptamer-based POCT device right in the delivery room.

In the diagnostics of less acute diseases or for screening approaches, the situation could be (and in parts already *is*) different. 12 out of 19 diseases covered by the Germany-wide standardized metabolic newborn screening (36–72 h *pp.*) are obligatorily performed by means of MS/MS (GBA, 2022). In contrast to their targeted detection, however, there are many hopes that unbiased and complete "plasma proteome profiling" might soon be established in a far more prominent role within clinical diagnostics, be it in the sense of an all-round screening method or as complementary diagnostics for patients of all ages and situations (Geyer *et al.*, 2016a). Though, for any diagnostic approach which inherently generates a more comprehensive output than necessary for a specific diagnosis (as in genomics & proteomics), numerous bioethical, legal and practical questions should be answered before a widespread clinical use is within reach. In proteomics, as a relatively young field of life sciences, the discussion about ethical guidelines and *e.g.* the handling of sensitive incidental findings are still in its infancy (Geyer *et al.*, 2021; Porsdam Mann *et al.*, 2021). The principles of proportionality, of proper purpose and of economic efficiency raise at least the question whether complete and unblinded proteome analyses – as presented in this thesis, too – will become standard clinical practice in a foreseeable future (or ever). Their contributions to basic and clinical research, however, are already today of great value and the further development of proteomic biomarker research deserves to be watched closely in any case.

It also remains to be seen to what extent the diagnostic potential of proteomics may one day ally itself with visionary therapeutic approaches. For instance, Schneider *et al.* (2018) reported a significantly alleviated phenotype of ectodermal dysplasia in three affected infants after intraamniotic injections of the crucial missing gene product, Ectodysplasin A (EDA). A recombinant fusion protein out of the receptor-binding part of EDA and the constant region of IgG$_1$ could enter the foetal organisms (presumably intestinally mediated) and unfold its EDA-

like effect on dependent structures in a critical phase of their development. This case report impressively illustrates both the central effector function of proteins and that a specific loss can be replaced to a certain degree in therapeutic intention. Proteins can not only be *measured*, they can also be *applied*, a perspective which might permeate proteomics one day in a more consistent manner, and which may add a further fascinating facet to this rich and dynamic field of science.

6. Summary

Infection and sepsis are major causes of global morbidity and mortality in the neonatal period. Their management is challenging since initial symptoms are unspecific and the diagnostic accuracy of established blood tests is suboptimal. Proteomics, the systemic investigation of proteins by mass spectrometry (MS), may provide new blood biomarkers – single proteins or entire biological networks – facilitating early diagnosis and improving the outcome of neonatal infection and sepsis.

In 2017–2019, the Dr. von Hauner Children's Hospital Munich (LMU) and the Max Planck Institute of Biochemistry Martinsried (MPIB) conducted a large-scale proteomic observational study in neonates ($n = 662$). By means of a cutting-edge MS-workflow, up to 700 proteins each were measured in minimal amounts of whole blood collected on conventional filter cards. Within the broader framework of this study, I dedicated the present thesis to a subgroup of late preterm ($\geq$ 35+0) and term infants affected by (non-culture-proven) early-onset bacterial infection ($n = 20$). From each of them, a pair of blood samples had been collected: a first sample as part of initial diagnostics and before the beginning of antibiotic therapy (around 10 h after birth), a second sample ca. 48 h later. I evaluated the respective protein profiles from three points of view: **(1)**, in comparison to control samples from healthy neonates; **(2)**, longitudinally over the two time points; **(3)**, by comparing the severely diseased patients with the less severely affected ones.
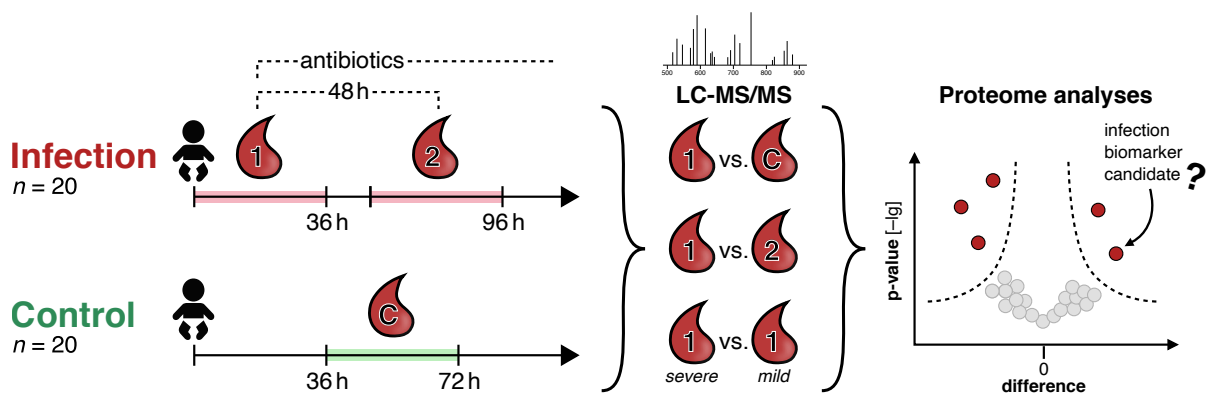
Figure 6.1: Graphical abstract

(1) S100-P was the most significantly altered protein (31-fold increased, $p = 0.001$). It also correlated positively with CRP (C-reactive protein) at initiation of antibiotic therapy ($r = 0.73$, $p = 0.03$). Moreover, the cytoskeletal regulator protein ACTR3 (actin-related protein 3) and the acute-phase protein HP (haptoglobin) were both 8-fold decreased ($p = 0.003$ & 0.0008). The DNA-binding protein hnRNPK (heterogeneous nuclear ribonucleoprotein K) showed a negative correlation with the total duration of antibiotic treatment ($r = -0.76$, $p = 0.02$). The haemostatic glycoprotein vWF (von Willebrand factor) had, unlike S100-P, a negative correlation with the initial CRP ($r = -0.67$, $p = 0.05$).

(2) The protein abundances of apolipoprotein B100 & E and of ITIH3 (inter-alpha-trypsin inhibitor heavy chain H3) were most significantly lower at t_1 than at t_2 (all ca. 2-fold lower, $q = 0.008$). Moreover, complement proteins and acute-phase proteins were significantly less abundant at t_1 as well, such as SERPIN-A3 & -A1 (alpha-1-anti[chymo]trypsin), ORM1 (alpha-1-acid-glycoprotein) and SAA2–4 (serum amyloid A 2–4). Among all of these, SERPIN-A3 and ORM1 reflected the dynamics towards t_2 most accurately, since they did not show any longitudinal change between the two examined control groups. S100-A8 and S100-A9 (also known as calprotectin dimer) were 2-fold more abundant at t_1 than at t_2 ($p = 0.003$ & 0.01).

(3) I could not identify any further potential markers, but I could reproduce some of the main findings from the two previous considerations in a low significance area.

Table 6.1: Main findings

	Proteins	Findings
Question 1 *diseased vs. healthy*	S100-P ACTR3 & HP hnRNPK vWF	↑ in infection group, positive correlation with initial CRP *(p. 46, 66, 52, 68)* ↓ in infection group *(p. 46, 66)* negative correlation with duration of antibiotic treatment *(p. 52, 68)* negative correlation with initial CRP *(p. 52, 69)*
Question 2 *diseased: early vs. late*	APOB, APOE & ITIH3 S100-A8/A9 Complement proteins Acute-phase proteins SERPIN-A3 & ORM1	↓ at early stage of infection *(p. 54, 71)* ↑ at early stage of infection *(p. 54, 72)* systemically ↓ at early stage *(p. 56, 73)* systemically ↓ at early stage *(p. 56, 73)* specific increase ↑ over time *(p. 58, 75)*
Question 3 *diseased: severe vs. mild*	ACTR3	(↓) in severely affected infection group *(p. 60, 77)*

Overall, despite small to medium sample sizes the present data grants deep insight into the proteome profiles of neonatal patients affected by early-onset bacterial infection. The significant role of (apo-)lipoproteins in inflammation and sepsis was confirmed (particularly for Apo-B100). A further examination of their diagnostic (or even therapeutic) potential in neonates seems to be very promising. Moreover, protein S100-P was shown to be significantly increased in the early phase of infection. For the proteins S100-A8 & A9 (calprotectin), there are at least indications that their usage in monitoring inflammation levels could be applicable to neonatal sepsis as well.

This thesis is to contribute to a more comprehensive understanding of the neonatal blood proteome. It is part of the ongoing search for the ideal set of biomarkers for neonatal infection and sepsis. Thereby, it is dedicated to an issue with room for significant improvement, for the greater good of many patients at an early point in their life.

7. ZUSAMMENFASSUNG

Infektion und Sepsis gehören weltweit zu den bedeutendsten Ursachen von Morbidität und Mortalität in der Neonatalperiode. Ihre Behandlung ist eine Herausforderung, da die Initialsymptome unspezifisch und die diagnostisch Genauigkeit etablierter Bluttests suboptimal sind. Proteomics, die systemische Erforschung von Proteinen mittels Massenspektrometrie (MS), könnte neue Blut-Biomarker liefern – einzelne Proteine oder ganze biologische Netzwerke – welche eine frühe Diagnose erleichtern und die Behandlungsergebnisse bei neonatalen Infektionen und Sepsen verbessern würden.

Von 2017–2019 führten das Dr. von Haunersche Kinderspital (LMU) und das Max-Planck-Institut für Biochemie Martinsried (MPIB) eine großangelegte proteomische Beobachtungsstudie an Neugeborenen durch (n = 662). Mit hochmodernen massenspektrometrischen Methoden wurden bis zu 700 Proteine in minimalen Mengen Vollbluts detektiert, welches auf konventionellem Filterpapier konserviert worden war. Im größeren Rahmen dieser Studie beschäftigte ich mich in der vorliegenden Promotionsarbeit mit einer Subgruppe von Frühgeborenen ($\geq$ 35+0) und Reifgeborenen, die von bakteriellen early-onset Infektionen (ohne mikrobiologischen Nachweis) betroffen waren (n = 20). Von jedem von Ihnen war ein Blutproben-Paar gewonnen worden: eine erste Probe als Teil der initialen Diagnostik und vor dem Beginn einer antibiotischen Therapie (ca. 10 h nach Geburt), eine zweite Probe ca. 48 h später. Ich untersuchte die jeweiligen Proteinprofile unter drei Gesichtspunkten: **(1)**, im Vergleich zu Kontrollproben gesunder Neugeborener; **(2)**, longitudinal über die zwei erhobenen Zeitpunkte; **(3)**, indem ich die schwer erkrankten Patienten mit den weniger schwer betroffenen verglich.

(1) S100-P war dasjenige Protein mit der signifikantesten Differenz (31-fach erhöht, p = 0.001). Es korrelierte außerdem positiv mit dem CRP (C-reaktives Protein) zu Beginn der antibiotischen Therapie (r = 0.73, p = 0.03). Darüberhinaus waren sowohl das cytoskelettale Regulatorprotein ACTR3 (Aktin-ähnliches Protein 3) und das Akute-Phase-Protein HP (Haptoglobin) beide 8-fach vermindert (p = 0.003 bzw. 0.0008). Das DNA-bindende Protein hnRNPK (Heterogenes nukleäres Ribonukleoprotein K) zeigte eine negative Korrelation mit der Gesamtdauer der antibiotischen Therapie (r = –0.76, p = 0.02). Das hämostatische Glyko-
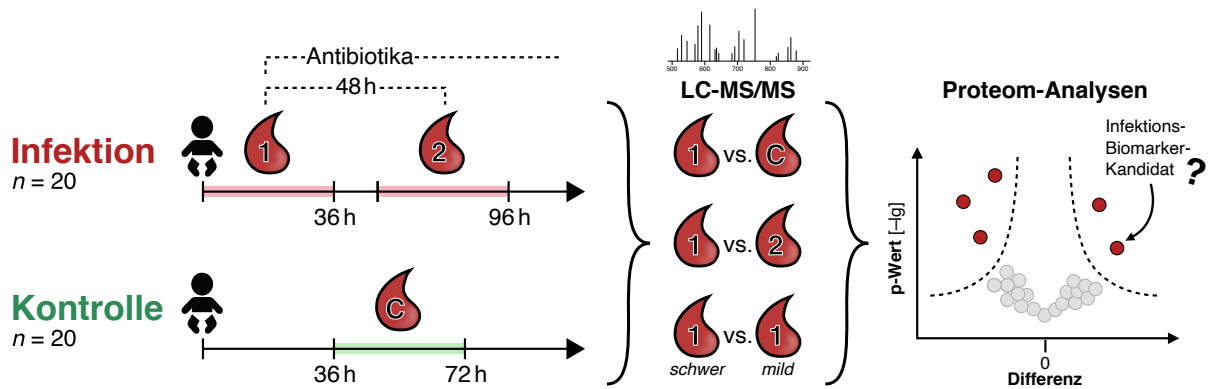
Figure 7.1: Graphical abstract

protein vWF (Von-Willebrand-Faktor) korrelierte anders als S100-P negativ mit dem initial gemessenen CRP ($r = -0.67$, $p = 0.05$).

(2) Die Apolipoproteine B100 & E sowie ITIH3 (inter-alpha-trypsin inhibitor heavy chain H3) waren bei t_1 in hoch-signifikanter Weise weniger stark vorhanden als bei t_2 (alle ca. 2-fach niedriger, $q = 0.008$). Darüberhinaus waren Komplement- und Akute-Phase-Proteine bei t_1 ebenfalls signifikant niedriger messbar, so z. B. SERPIN-A3 & -A1 (Alpha-1-anti[chymo]-trypsin), ORM1 (Saures Alpha-1-Glykoprotein) und SAA2–4 (Serum-Amyloid A 2–4). Unter all diesen gaben SERPIN-A3 und ORM1 die Dynamik zu t_2 am genausten wieder, da diese keine longitudinale Veränderung zwischen den beiden untersuchten Kontrollgruppen zeigten. S100-A8 und S100-A9 (auch bekannt als Calprotectin-Dimer) waren bei t_1 2-fach höher messbar als bei t_2 ($p = 0.003$ & 0.01).

(3) Ich konnte keine weiteren potentiellen Marker identifizieren, aber ich konnte manche der Hauptergebnisse aus den zwei vorigen Fragestellungen mit niedriger Signifikanz reproduzieren.

Alles in allem erlauben die vorliegenden Daten trotz kleiner bis mittlerer Stichprobengrößen tiefe Einblicke in die Proteinprofile von Neugeborenen mit bakteriellen early-onset Infektionen. Die signifikante Rolle von (Apo-)Lipoproteinen in Inflammation und Sepsis wurde bestätigt (insbesondere die von Apo-B100). Eine weitere Untersuchung ihres diagnostischen (oder gar therapeutischen) Potentials bei Neugeborenen erscheint äußerst vielversprechend.

Table 7.1: Hauptergebnisse.

|	Proteine	Ergebnisse
Frage 1 *erkrankt vs. gesund*	S100-P ACTR3 & HP hnRNPK vWF	↑ in Infektionsgruppe, positive Korrelation mit initialem CRP (p. 46, 66, 52, 68) ↓ in Infektionsgruppe (p. 46, 66) negative Korrelation mit Dauer antibiotischer Behandlung (p. 52, 68) negative Korrelation mit initialem CRP (p. 52, 69)
Frage 2 *erkrankt: früh vs. spät*	APOB, APOE & ITIH3 S100-A8/A9 Komplement-Proteine Akute-Phase-Proteine SERPIN-A3 & ORM1	↓ in Frühphase der Infektion (p. 54, 71) ↑ in Frühphase der Infektion (p. 54, 72) systematisch ↓ in Frühphase der Infektion (p. 56, 73) systematisch ↓ in Frühphase der Infektion (p. 56, 73) spezifischer Anstieg ↑ über die Zeit (p. 58, 75)
Frage 3 *erkrankt: schwer vs. mild*	ACTR3	(↓) in schwer betroffener Infektionsgruppe (p. 60, 77)

Darüberhinaus konnte gezeigt werden, dass S100-P in der Frühphase der Infektion stark signifikant erhöht ist. Für die Proteine S100-A8 & A9 (Calprotectin) ergaben sich zumindest Hinweise darauf, dass deren Einsatz bei der Verlaufskontrolle von Inflammationslevels auch auf die neonatale Sepsis übertragbar sein könnte.

Diese Arbeit will einen Beitrag zu einem umfassenderen Verständnis des neonatalen Blut-Proteoms leisten. Sie ist Teil der anhaltenden Suche nach den idealen Biomarkern für die Infektion und Sepsis des Neugeborenen. So gesehen ist sie einem Thema gewidmet, welches noch Raum für deutliche Verbesserungen birgt, zum Wohle vieler Patienten an einem frühen Punkt ihres Lebens.

APPENDIX

ADDITIONAL FIGURES

Among the depicted characteristics the corrected gestational age at the time of blood sampling had the highest priority for the formation of matched control groups. Moreover, the figures show the appropriately comparable distributions of birth weights and the resulting distribution of females and males for each grouping (cf. Ch. 4.1). The p-values in the box plots refer to two-sided two-sample t-tests between the given groups.

Grouping I1(72) vs. C-I1(72)

Grouping I1(96) vs. Cnm-I1(96)

Grouping I2(96) vs. Cnm-I2(96)
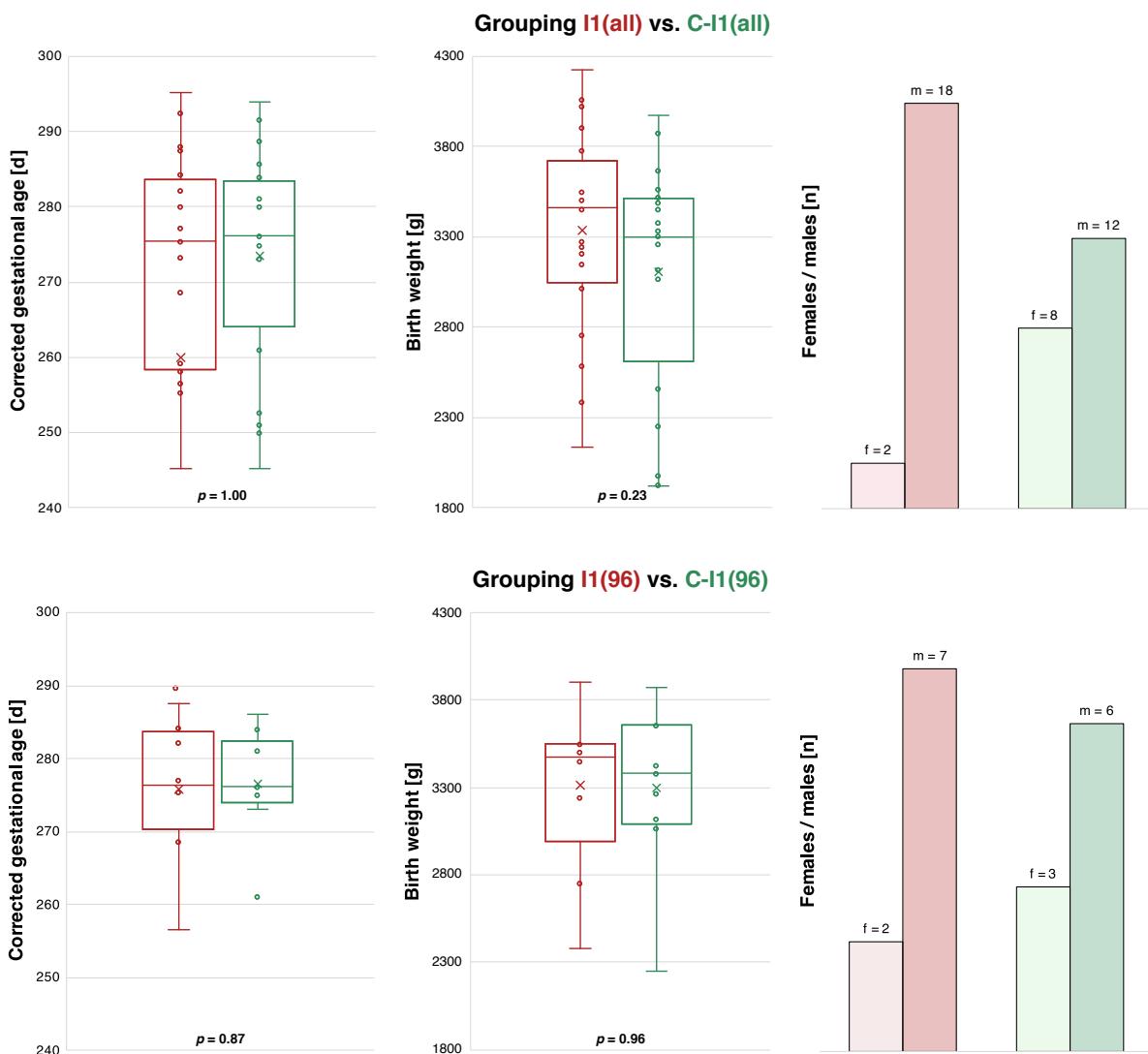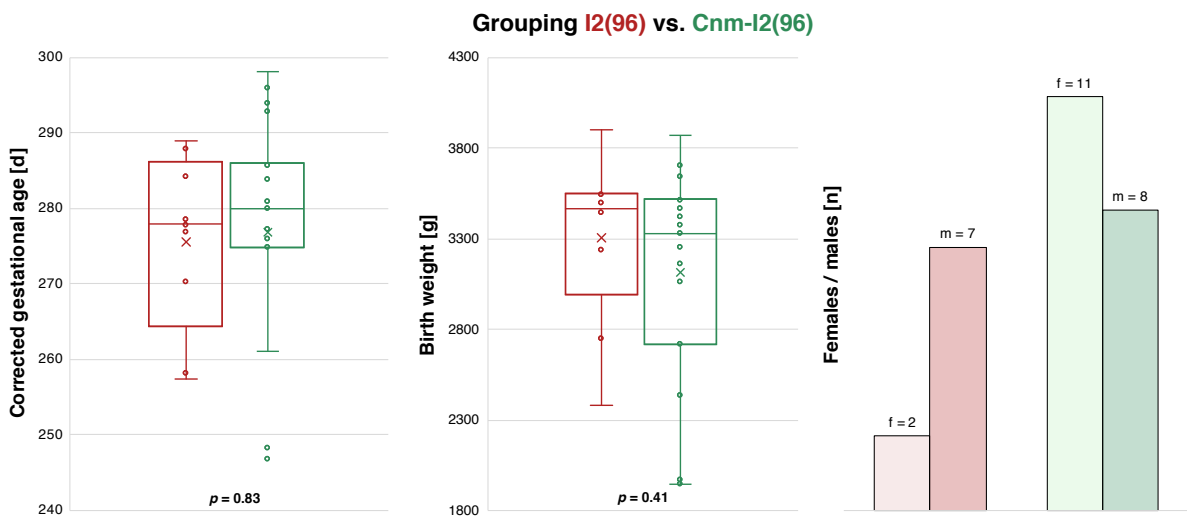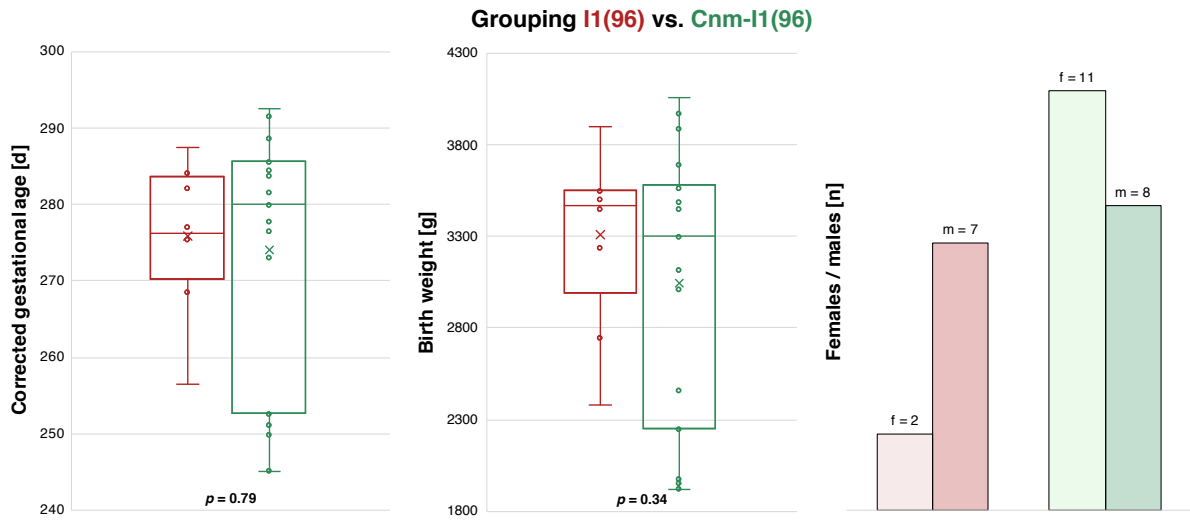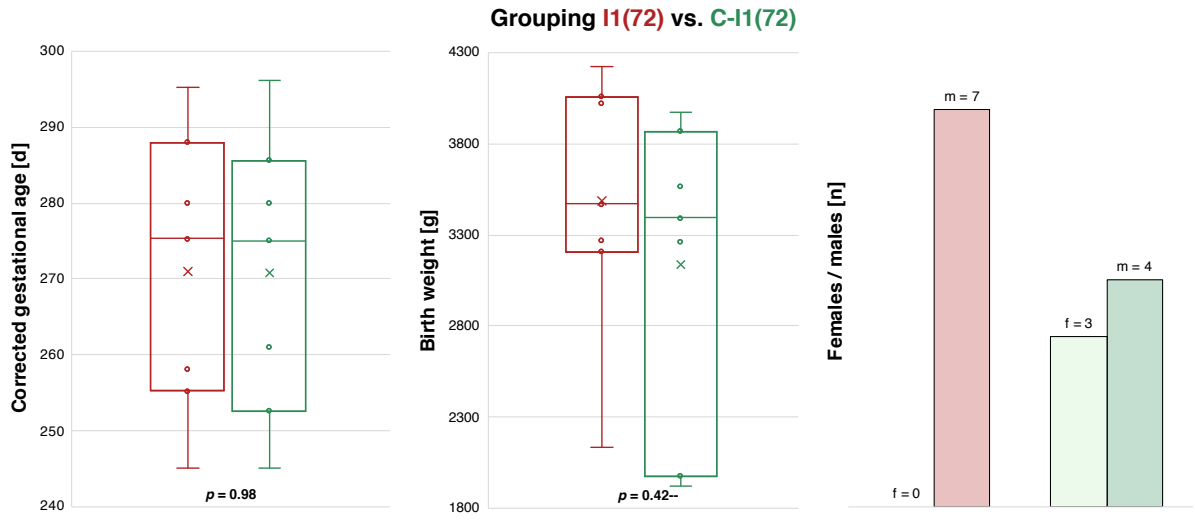
ADDITIONAL MATERIALS

Basics of mass spectrometry

As briefly laid out in Ch. 1.2.5, mass spectrometry can be characterized as "weighing" (determining the mass of) molecules by "charging" (ionizing) them and observing their response to an artificial electric field. The following subchapters address these basics of MS in some more detail. They originated from my personal engagement with this method and might be helpful for other beginners in this complex field as well.

From here, the major steps of a prototypical LC-MS/MS workflow (similar to that of the presented study, *cf.* Ch. 3.9) and its basic operating principles shall be explained (*cf.* Fig. B.1). The following pages theoretically trace the journey of an unknown analyte from sample preparation throughout the mass spectrometric key modules up to its computational identification and quantification.
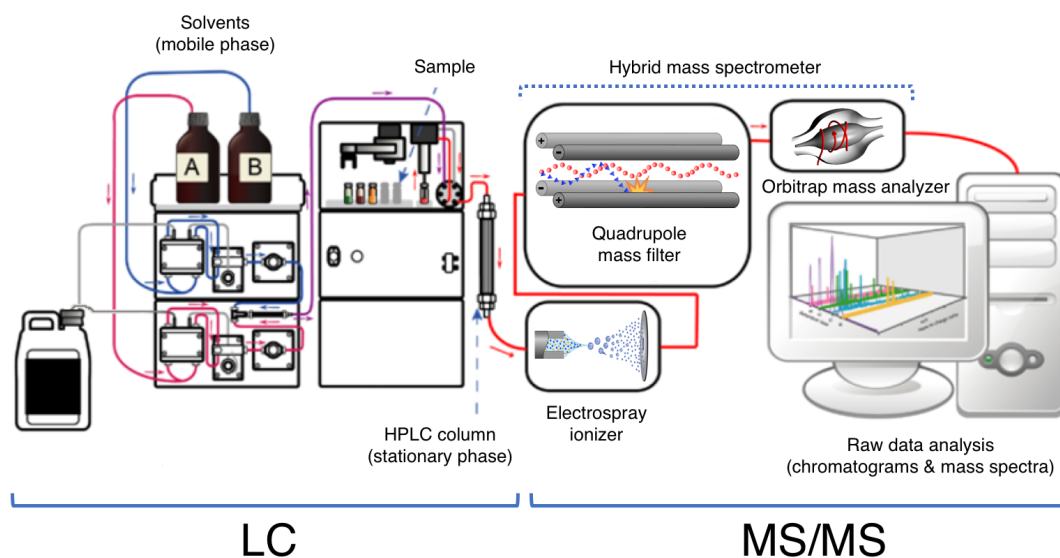

Figure B.1: Schematic LC-MS/MS setup. LC-MS/MS is a combination of liquid chromatography (LC) and tandem mass spectrometry (MS/MS). The structure of this chapter follows the illustrated order of major steps in a typical LC-MS/MS system (as used in this thesis). Figure adapted from Norena-Caro (2017a).

B.1 Sample preparation

In MS-based proteomics, the stage of sample preparation is already a conceptually decisive step. Similarly to other scientific fields, there are two complementary major approaches in proteomics called **top-down** and **bottom-up** (Aebersold & Mann, 2016).

In top-down proteomics, samples with intact proteins enter analysis, whereas in bottom-up proteomics samples with cleaved proteins, *i.e.* peptides, are fed into the mass spectrometric setup. To this end, after having dissolved, buffered and denatured the samples, they undergo enzymatic proteolysis with a sequence-specific protease (*e.g.* trypsin). Since it cleaves the amino acid chain only at stereotypic places a later inference to the original protein will still be possible by means of bioinformatics. While complex posttranslational modifications and protein isoforms can be determined better in top-down proteomics, a significant strength of bottom-up proteomics is its higher resolution. Therefore, it is the more widespread approach at present.

All following descriptions of MS working principles mainly relate to bottom-up proteomics, since it is the approach of this thesis (*cf.* Ch. 3.9)

B.2 Liquid chromatography

Chromatography is a procedure which separates mixtures into their single constituents. It is used in numerous analytical applications, not only as preparation for MS. A common characteristic of all types of chromatography is the interaction of a so-called **mobile phase** with a **stationary phase** (*cf.* Fig. B.2 and Ritgen, 2019). Usually, the analyte is dissolved in the mobile phase (a form of solvent or gas) and follows its current through a structure (like a column, plane or sheet) which is firmly loaded by the stationary phase (a gel or another material with a specific chemical property). Each dissolved component has an individually pronounced affinity to both phases. A molecule either adheres to molecules of the mobile phase (and is carried farther) or it adheres to the stationary phase (and

Figure B.2: Chromatography. Basic principle of chromatography is the interaction of a mobile phase with a stationary phase.

stands still), but it never does both at the same time. According to its **partition coefficient**, it is always a certain portion of one kind of molecules that is retained by the stationary phase. This equilibrium determines the individual velocity of all types of constituents during the passage,
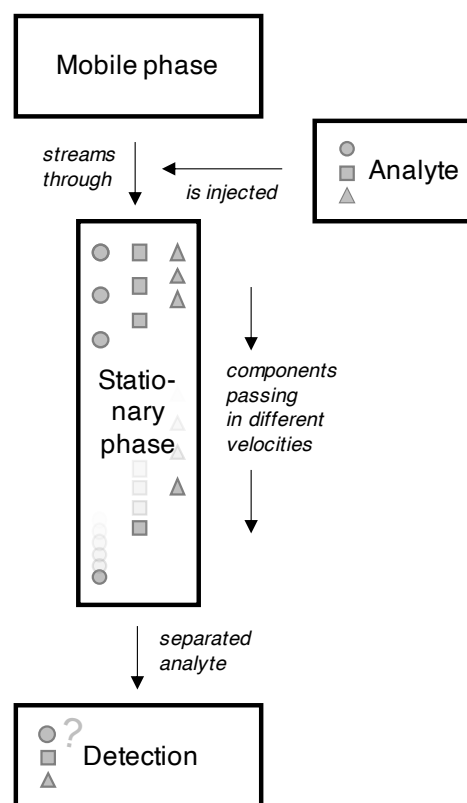
their **retention time** (Newton, 1988). Thus, chemical interactions and resulting retention times are the reason for which the different groups of analytes are separated.

Typically, in liquid chromatography (LC) the mobile phase is a fluid that streams under pressure through a column containing the stationary phase. Current methods as in this thesis are often called high performance liquid chromatography (HPLC). They use high pressures (several hundred bar) to pump the mobile phase (an aqueous solvent) at constant flow rates (a few hundred nl/min) through a thin column (inside diameter a few µm to mm, length some dm). The stationary phases are mainly silica- or polymer-based micro-structures inside the column. Due to their individual retention time all of the mixture's components leave the column in a certain order and can then be identified with another technique.

B.3   Electrospray ionizer

Mass spectrometry has already been characterized as determining the mass of molecules by ionizing them and evaluating their response to a controlled electric field. Since this is only possible for ions in gas phase, the eluted analytes still have to be converted from liquid to gas phase. Therefore, an **ion source** as an electrospray ionizer (Fenn *et al.*, 1989) may connect the LC module with the actual mass spectrometer. An electrospray ionizer (ESI) nebulizes and ionizes the analyte, leading to a fine, charged aerosol. MALDI (matrix-assisted laser desorption/ionization), the other widespread form of sample ionization (Karas *et al.*, 1985; Tanaka *et al.*, 1988), will be omitted here (*cf.* Ch. 3.9).
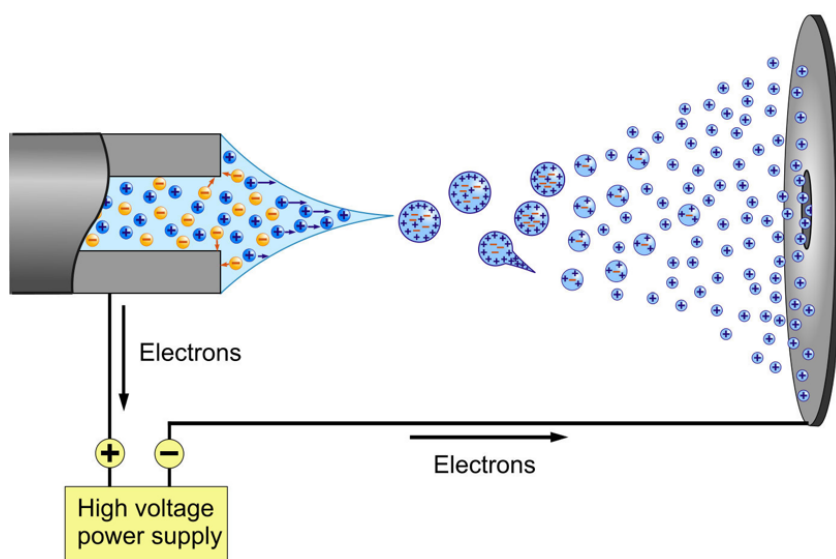
Figure B.3: Electrospray ionizer (ESI). A schematic ESI chamber with voltage (positive mode) between metal capillary and counter electrode. The *Taylor cone* at the tip of the capillary develops due to an equilibrium of forces between electric field forces, surface tension, hydrostatic pressure, external gas pressure and gravitation. The shrinking ionized droplets traverse the chamber driven by the potential difference. Figure adapted from Dahlin (2015).

In a typical MS setup, the LC module directly elutes into the ESI chamber. A voltage (2–4 kV) is applied between the very end of the HPLC column and a plate-shaped counter electrode. Depending on the voltage permeating the solvent, ionized droplets of only one kind (either positively or negatively charged ones) form a so-called **Taylor cone** at the tip of the capillary (*cf.* Fig. B.3 and Taylor & McEwan, 1965). Due to the repulsion of identical

charges within the cone the flow of liquid steadily disintegrates and single droplets come off. They follow the electric field forces and travel towards the counter electrode. A neutral gas in the ESI chamber (*e.g.* nitrogen) supports their evaporation, leading to gradually decreasing volumes of the ionized droplets. While their volume decreases, their charge stays the same. Any time the density of identical charges within a shrinking droplet exceeds the **Rayleigh limit** (Rayleigh, 1882), the droplet separates itself into even smaller droplets – a phenomenon known as **Coulomb explosions** (Kebarle & Verkerk, 2009). This repeats itself until only single ions arrive at the counter electrode, pass its central orifice and enter the next MS module.

B.4 Hybrid mass spectrometer

Given the immense variety of available devices, describing universal operating principles of mass spectrometers is difficult. Instead, this subchapter deals, by means of example, with the basics of the Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap™ MS system (by Thermo Fisher Scientific, *cf.* Kelstrup *et al.*, 2018). It is the system which was used for the measurements of this thesis. From here, the route of analytes through a device such as the Q Exactive™ HF-X is followed step by step.
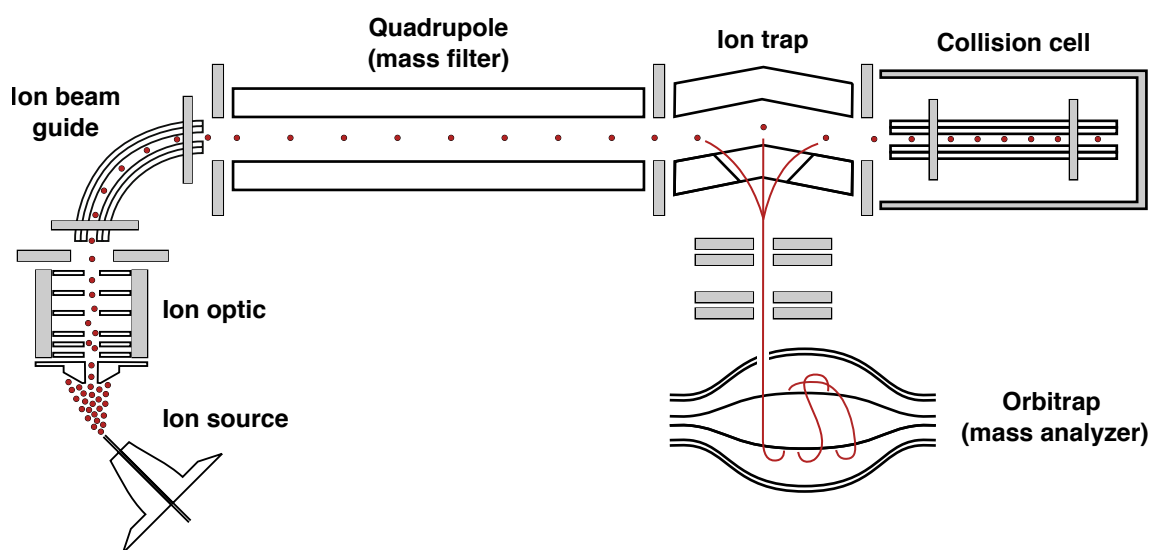
Figure B.4: Q Exactive™ series MS system. Scheme of a hybrid MS system from the Q Exactive™ series by Thermo Fisher Scientific (Kelstrup *et al.*, 2018) with the major modules introduced in this subchapter.

The ultimate goal of a mass spectrometer is to separate the introduced ions according to their mass-to-charge ratio (m/z), to quantify this ratio and to infer from it to the identity of the underlying particle. "Hybrid" mass spectrometers are setups which contain at least two different m/z separation components (as quadrupole and orbitrap in the Q Exactive™ HF-X, *cf.* Fig. B.4). The modules' largely linear arrangement defines the flight path for the entering ions. Most of the apparatus is a vacuum environment. This is to ensure that as few ions as possible get lost to analysis due to collisions with other gas particles. The latter would either

lead to an ion's neutralization (its complete loss) or at least to the distortion of its actual mass spectrum. At the same time, the vacuum is one of two factors responsible for the necessary propulsion of ions. By establishing a low-vacuum at the system's entry and a high-vacuum at its end, a flow throughout all compartments is generated. The second driving force for the ion's propulsion are electric potential gradients which will still be mentioned several times.

Typically, the first modules after the ESI have two goals: one, focusing the entering "ion cloud" to a concentrated "ion beam", and two, cleaning it from adducts and neutral molecules, which resulted from gas dynamics or charge loss and which could enter the instrument.

The focusing is achieved by parts summarized as **ion optics**. Not quite different from lenses that focus light, the disc-shaped electrodes of an ion optic (called "lenses" as well) are able to focus ions. This is possible due to their funnel-like arrangement and their electric configuration. The applied alternating current (AC), offset between neighboured electrodes, confines the ions radially. The simultaneous application of direct current (DC) decreasing towards the rear electrodes generates an electric potential gradient which propels the ions axially. This axial motion of ions in the electric field can be compared to balls that can not help but roll "downhill" (Savaryn *et al.*, 2016).

The cleaning can be taken on by modules sometimes referred to as **ion beam guides**. A specific application of currents in several metal rods brings out the ion's further axial propulsion through this channel-like arrangement. Although similar to the quadrupole (described below), an obvious difference is the ion beam guide's curved shape. Neutral particles, that have come so far, will not be able to hold the required bent trajectory due to their inertia. They collide with the rods and are thus prevented from entering the quadrupole.
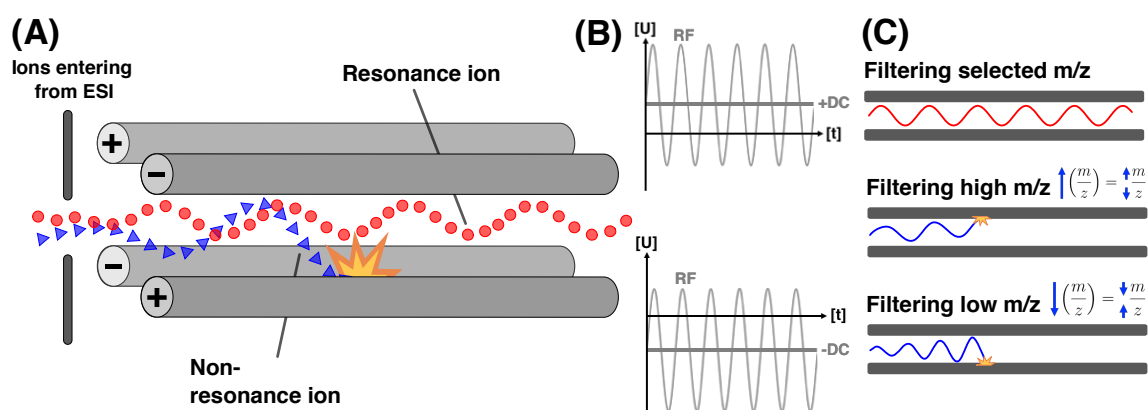
Figure B.5: Quadrupole mass filter. (A) Schematic illustration of the charged rods' crosswise pairing and the propulsion of ions through the electric field. Depending on the settings of DC and RF *resonance ions* take on a stable helical trajectory while *non-resonance ions* do not and collide with the rods. **(B)** The offset DC baselines (between the pairs of rods) are periodically exceeded by their concomitant RF. **(C)** An ion's trajectory is a direct consequence of the applied currents, its mass-to-charge ratio (m/z) and inertia. Heavy molecules with light charge (high m/z) are hardly deflected (low amplitude), whereas light molecules with high charge (low m/z) are easily deflected (large amplitude). The quadrupole ultimately achieves its mass filter function by ensuring a stable trajectory for ions of the desired m/z (or range of m/z). Figure adapted from Gadolinq (2020).

Although being very versatile, in the majority of all applications the **quadrupole** serves as a mass filter. It can be set in a way that allows either all or only selected sorts of ions to

pass. A quadrupole is a symmetrical arrangement of four charged metal rods. In terms of the applied currents, they are "crosswise" paired (*cf.* Fig. B.5). A positive DC flows through two opposite rods, a negative one through the other two. At the same time, both DC baselines are combined with an AC reversing polarity at radio frequencies (RF current). An oscillating electric field is created (similar to ion optic and ion beam guide) forcing the entering ions onto a helical trajectory. However, depending on how exactly AC and RF are combined, not necessarily all ions will have a stable trajectory and pass the whole length of the quadrupole as so-called **resonance ions**. The user-defined scanning window may allow passage to all ions, or only such of a certain m/z range or one defined m/z value. The **non-resonance ions** will get onto unstable trajectories with increasing radius and will sooner or later collide with the rods (which illustrates that MS is indeed "sample-consuming").

In the Q Exactive™ HF-X system, the quadrupole is followed by an **ion trap**, which can hold and then transmit incoming ions either to a **collision cell** or directly to the **orbitrap**, the actual place of m/z detection (*cf.* Fig. B.4). The possibility to send ions in either direction is a pivotal characteristic and a precondition for the concept of tandem mass spectrometry (MS/MS), which will be introduced in the following subchapter (*cf.* Ch. B.5). In their operating principles, ion traps, collision cells and orbitraps have a lot in common with the previous modules and their interplay of voltages, leading to arbitrarily controllable electric fields.

Ion traps as those of the C-trap type (curved linear trap for ion injection) can effectively accumulate, stabilize and store ions without loss before sending them in packets in one of the two possible directions (collision cell or orbitrap). Ions that take the detour and enter the collision cell first, will undergo higher-energy collisional dissociation (HCD). This means they are accelerated and collide with the neutral molecules of a gas (such as nitrogen or helium) filling out the HCD cell. Due to the associated conversion of kinetic to internal energy, the incoming ions' amino acid chains break at specific low-energetic bonds, depending on the energy applied. These stereotypical breakages are a condition for the upcoming identification of the peptide's amino acid sequence. To this end, the resulting

Figure B.6: Orbitrap mass analyzer. The C-trap transmits ion packets to the Orbitrap mass analyzer for final detection. The ions take on a helical trajectory around the inner electrode. A *Fourier transformed* record of the axial oscillation signal reveals the underlying m/z. Figure adapted from Thermo Fisher Scientific (2012).

fragment ions are sent back to the C-trap, and from there to the orbitrap.

The orbitrap (Zubarev & Makarov, 2013) is a mass spectrometric analyzer consisting of a spindle-shaped inner electrode and two cup-shaped counter-electrodes, which form some kind of case for the inner one. By applying voltage between outer and inner electrodes, an electric field with an axial and a radial component is created. By that, entering ions are
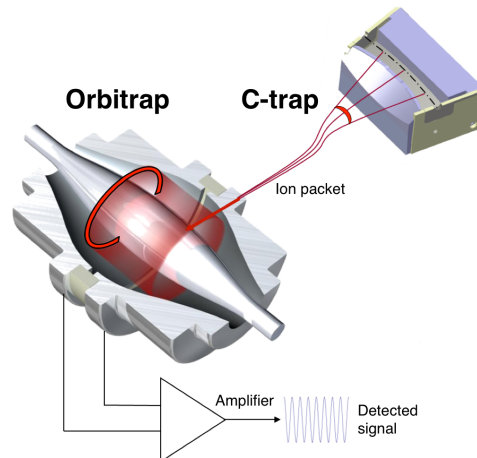
trapped in an orbital motion around the inner electrode. The ions oscillate radially to the spindle (around it) while also moving axially back and forth (along it). The axial oscillation is purely harmonic and independent from the frequency of rotation, but related to the m/z. By recording an image of these stabilized, axial oscillation signals for a certain period (millisecond range) and subjecting them to a **Fourier transformation**, it is ultimately possible to infer to the underlying m/z of any injected ion.

B.5   Tandem mass spectrometry

In the previous subchapter an analyte's journey through a mass spectrometer such as the Q Exactive™ HF-X was followed. However, without some complementary remarks on the *temporal* dimensions of this analytic process called tandem mass spectrometry (MS/MS), these explanations would remain incomplete.

In proteomics, "tandem" MS refers to the examination of each peptide species in (at least) two steps: once as intact peptide (**MS^1 level**) and once (or more often) as its fragments (**MS^2** or **MS^n level**). This can be achieved either by the spatial arrangement of two or more mass analyzers in a row with an interposed collision cell (**tandem in space**) or by the targeted re-examination of peptide species in the same and only analyzer (**tandem in time**), *i.e.* first without and then with previous HCD (Gross, 2017). After all, the merged information from these multiple analyses in MS/MS delivers more comprehensive information enabling a far more accurate peptide and protein identification than by MS^1 spectra alone.
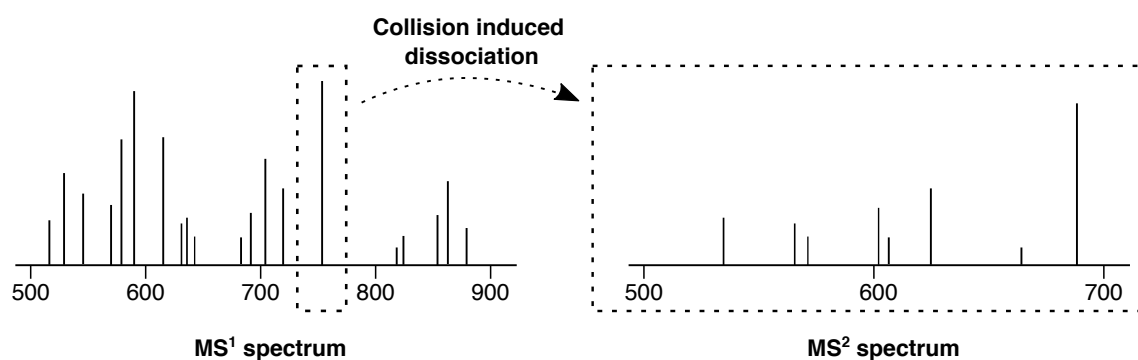
Figure B.7: MS^1 & MS^2 spectra. MS^1 spectra refer to the measurement of *precursor ions* and MS^2 spectra display the measurement of their *fragment ions*. Interpreting both in synopsis is a central feature of tandem mass spectrometry, leading to greater analytical depth.

No matter if in tandem in space or tandem in time, in the first stage an unknown peptide is analyzed as a whole (together with other peptides), which delivers the so-called MS^1 spectrum. However, the contained m/z peaks of these **precursor ions** are more or less like "silhouettes", since they may give an idea of what is visible but might not be sufficient for an undoubted peptide identification, yet. In the second stage of analysis, each detected m/z is further explored in a targeted manner. Through filtering the ions of a certain m/z and by fragmenting them prior to detection, a complementary MS^2 spectrum is yielded, which displays

the respective **fragment ions**. Due to the stereotypical breakages introduced in the HCD cell all of these differently long peptide fragments can be combined to uncover vast parts of the precursor ion's amino acid sequence. With some bioinformatic effort (*e.g.* comparing the questionable MS^2 spectrum with all known and all theoretically possible MS^2 spectra), the identity of the underlying peptide can finally be inferred. Thus, together an MS^1 spectrum and its sequence-specific MS^2 spectrum result in a "sharp" and "well-illuminated" picture instead of a mere „silhouette".

Based on this principle, there are broadly two strategies in bottom-up proteomics how MS devices can actually be set to acquire these tandem mass spectra: the **data-dependent acquisition (DDA)** and the **data-independent acquisition (DIA)**. Both scan modes will shortly be illustrated here, again referring to an MS system as the Q Exactive™ HF-X.

DDA – data dependent acquisition
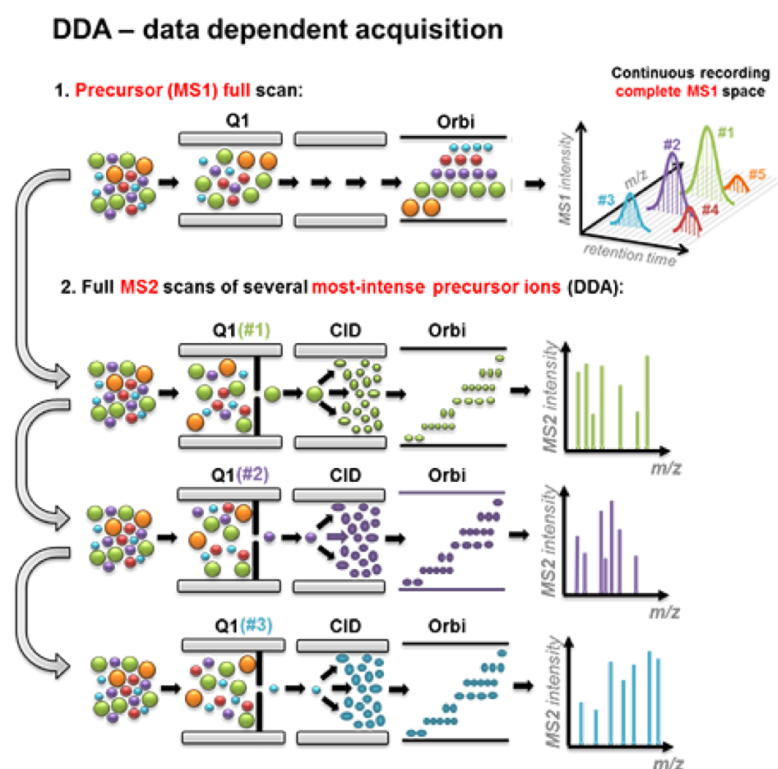
Figure B.8: Data-dependent acquisition (DDA). Scheme of the basic concept of DDA. After one primary full scan (MS^1 level) the *n* most intense precursor ion populations are one after another isolated, fragmented and measured (MS^2 level). Figure adapted from Ludwig *et al.* (2018) and reprinted by courtesy of Dr. Stefanie Hauck (Helmholtz Munich).

In a typical DDA, first, a full-scan MS^1 spectrum is acquired for all available peptides of one time point in the LC run. Concretely, this means that the quadrupole mass filter is first set to let pass all incoming ions, and that these will be sent to the orbitrap mass analyzer without previous fragmentation in the HCD cell. In the obtained MS^1 spectrum, a user-defined number n of the most intense precursor ions will be selected for further analysis on MS^2 level (top n method). The mass spectrometer will then work off this selection and step by step acquire the corresponding MS^2 spectra. The scheme is always the same: the quadrupole allows passage only to ions of the desired m/z; the C-trap will transmit these ions to the HCD cell; the resulting fragment ions will be returned and detected by the orbitrap. By and by, these cycles

generate a set of complementary MS^2 spectra to the original MS^1 spectrum. A considerable drawback of this approach is that there can hardly ever be enough analysis time to run an experiment without the arbitrary limitation to a certain number of top precursors. Ultimately, not all present peptides will reach the MS^2 level, and – although comprehensive – the peptide coverage in the DDA strategy is prone to be incomplete (Sinha & Mann, 2020).

In DIA, by contrast, a full-scan MS^1 spectrum as a starting point of analysis is not mandatory. DIA rather focuses on MS^2 spectra for whole groups of peptide species (not isolated peptide species as in DDA). *A priori*, a relatively broad search window (20–40 m/z) is defined. This will be moved by the quadrupole step by step through the entire mass range. Concretely, this means that in all MS cycles the quadrupole is open for all ions of the current search window; a heterogeneous group of peptides will reach C-trap and HCD cell; the returned mix of fragment ions will enter the orbitrap for a simultaneous detection. This scheme results in complex MS^2 spectra containing the data for several ions which still have to be unravelled. However, an immense benefit of this approach is its unbiased nature since it represents an attempt to gradually cover the whole mass range by MS^2 spectra. Despite the bioinformatic effort required to decipher them, DIA is the method of choice in many large experimental setups (as in the study underlying this thesis). That is because it offers high consistency, reproducibility and a most comprehensive peptide detection at the same time (Ludwig *et al.*, 2018).
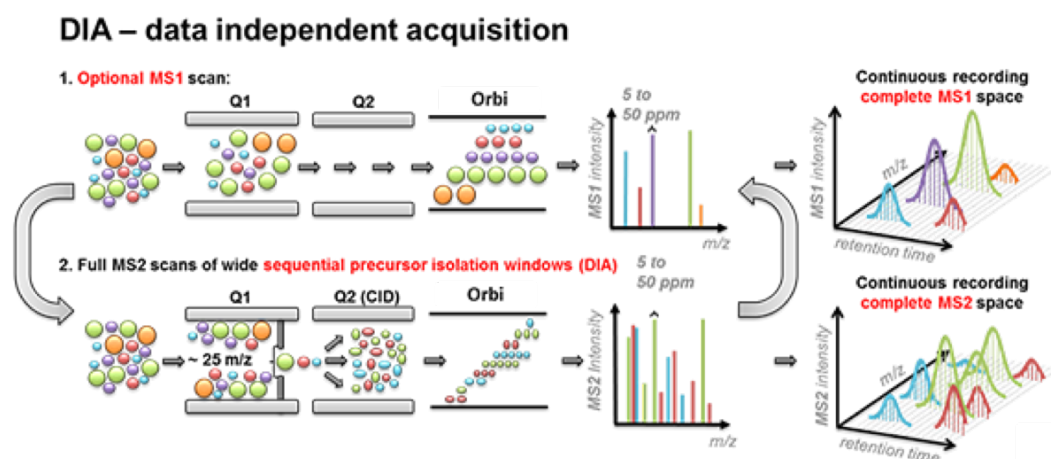
Figure B.9: Data-independent acquisition (DIA). Scheme of the basic concept of DIA. A defined mass window is gradually moved over the entire mass range. The mixed groups of peptides, which are comprised by this mass window at a certain moment and pass the mass filter, are always led to fragmentation, measurement and identification together. Figure adapted from Ludwig *et al.* (2018) and reprinted by courtesy of Dr. Stefanie Hauck (Helmholtz Munich).

B.6 Raw data analysis

Regardless of the strategy chosen for data acquisition, in the end any obtained mass spectrum still has to be traced back to its original analyte. In the case of bottom-up proteomics, however, the challenge is twofold since the analyte changes its shape twice in the course of analysis: As part of sample preparation all proteins are subjected to sequence-specific enzy-

matic cleavage (*cf.* Ch. B.1), and to get insight into the MS^2 level these peptides are further fragmented in the HCD cell (*cf.* Ch. B.4). Therefore, not only an inference from mass spectra to peptides is required, but also from peptides to original proteins, which is known as **inference problem**. The computational processing of MS raw data in bottom-up proteomics can thus be broken down to four major steps: mass spectra acquisition, peptide identification, protein identification and protein quantification (Sinha & Mann, 2020).

The identification of peptides (after HCD) is largely based on the formation of **b- and y-ion series** for each precursor ion's fragments. Essentially, this means to sort all available (randomly created) peptide fragments by length, to read out all mass differences within the fragment series and, thereby, to virtually "sequence" the amino acid chain of the precursor ion. The b-ion series comprises all fragments with the original N-terminus, the y-ion series those with the C-terminus. By comparing the merged data on both series with pre-existing proteome databases, the identity of the underlying precursor ion can be inferred.

The major principle behind the inference from identified peptides to their original proteins has already been outlined in Ch. B.1. In bottom-up proteomics, the initial proteolytic digestion of samples is typically conducted with a single enzyme, a highly **sequence-specific protease** (*e.g.* trypsin or LysC) that cleaves the amino acid chains only at stereotypic places. By this constant initial condition and some pre-existing bioinformatic resources, a correct re-assembling of the original proteins becomes only a matter of computing capacity.

A further key feature of LC-MS/MS which is helpful for the interpretation of mass spectra is their relation to the initially obtained chromatogram. The **retention time** from LC enriches the mass spectra (ultimately an m/z vs. the intensity of measurement) by a third dimension. This "fanning-out" is especially helpful for comparisons between different runs of an experiment, when analytes are supposed to be reproducibly identified, and also when it comes to the problem of peptide quantification.
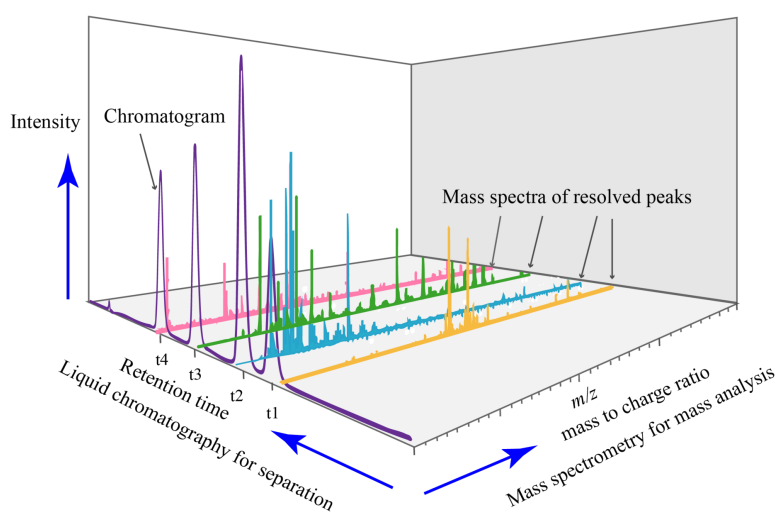

Figure B.10: Mass spectra and retention time. A major strength of LC-MS/MS is the possibility to combine knowledge about the obtained mass spectra (m/z) with knowledge about the chromatogram (*retention time*). LC is introducing a third dimension into the mass spectrometric analysis, which decisively increases the overall depth of analysis. From Norena-Caro (2017b).

Concerning this challenge of quantifying MS measurements, broadly two strategies can be distinguished: label-based and label-free quantification.

The common characteristic of **label-based** approaches is the introduction of isotopic labels into portions of all samples prior to LC-MS/MS. Thereby, a reference sample condition is created. By the signal intensity ratios between the resulting pairs of peptides, the tagged ones and their non-tagged (or differently tagged) counterparts, a relative quantification is possible. Only to mention a few examples: In the ICAT approach (isotope-coded affinity tags, Gygi *et al.*, 1999), one portion of the sample is chemically labelled with a heavy isotope and another portion with a light one. The iTRAQ approach (isobaric tags for relative and absolute quantitation, Ross *et al.*, 2004) works similarly but allows higher sample throughput. In contrast to these *in vitro* strategies, in the *in vivo* SILAC approach (stable isotope labelling by amino acids in cell culture, Ong *et al.*, 2002) the desired cell line is grown in a special medium in which an essential amino acid is totally replaced by an isotopically labelled form. The modified amino acid will be incorporated and will by and by displace the normal form in the entire cell metabolism. After cell harvest and proper sample preparation the experimental specimens are mixed with the normally cultured control specimens. Both will be measured together by LC-MS/MS and a relative quantification of each peptide species is possible as described above.

On the contrary, **label-free** quantification (LFQ) does not require extended sample preparation. It is a purely computational inference from the detected mass spectra to the original peptide proportions. This can be achieved either by evaluating the precursor ions' signal intensities (height of peaks, area under the curve) or by just counting the spectra of any identified peptide. As most label-based approaches, this is a form of relative quantification. The measurements are neither assigned any units nor are they compared to external standards. The intensity or count of the measured will only be seen in relation to the own measurement series. At the price of losing a direct comparability between different measurement series, LFQ is advantageous in large-scale experiments since it saves time and resources (shortened sample preparation, no labelling reagents needed, applicable to any kind of specimen, flexibility in study design). Thanks to the bioinformatic progress of recent years the accuracy of LFQ is meanwhile more than satisfying (Bantscheff *et al.*, 2012; Patel *et al.*, 2009). All these aspects make it one of the most commonly applied quantitative methods, especially in bottom-up proteomics.

With the quantification completed, be it label-based or label-free, the proteomic raw data set is finally made accessible to biostatistical downstream analyses.

Laboratory protocol

Proteomics sample preparation: Protein extraction, digestion and in-StageTip purification
For sample preparation, samples of the DBS filter cards were randomized, punched out with a 14-Gauge syringe and aliquoted into 96 well plates (Eppendorf TwinTec, low bind), prepared with 20 µl H_2O per well.

We performed batch sample preparation according to the in-StageTip protocol (Kulak *et al.*, 2014). In brief, 20 µl 2x lysis buffer (80 mM Chloroacetmide, 20 mM TCEP, 200 mM Tris, 2 % Glycocholic acid (w/v), pH 8.0) was added to the wells, and samples were incubated for reduction of disulfide bridges, cysteine alkylation and protein denaturation at 95 °C for 10 min. Samples digested with Trypsin and LysC (1:100 ratio, µg of enzyme to µg of sample protein) at 37 °C for 4 h. Protein amount per punch-out was pre-determined on test samples at ∼200 µg.

Digestion was quenched with 100 µl Buffer A (1 % FA in H_2O) and 500 ng peptide digests were loaded on Evotips (Evosep, Odensee, Denmark) following the manufacturer's protocol.

High-pressure liquid chromatography and mass spectrometry We analyzed the samples applying LC-MS instrumentation consisting of an Evosep One chromatographic system (Evosep, Odensee, Denmark), which was coupled to an Exploris HF-X mass spectrometer by a nano-electrospray ion source (Thermo Fisher Scientific).

We used an in-house-packed 15 cm HPLC-column (150 µm ID, packed into a pulled tip with ReproSil-Pur C18-AQ 1.9 µm resin, Dr. Maisch GmbH) in combination with the 21 min 60 SPD gradient provided by the manufacturer. Column temperature was kept at 60 °C by an in-house-developed oven containing a Peltier element, and parametres were monitored in real time by the SprayQC software (Scheltema & Mann, 2012).

To acquire MS data, the data-independent acquisition (DIA) scan mode was used for single-shot patient samples, whereas a pool of the samples was fractionated and acquired with a top-12 data-dependent acquisition (DDA) scan mode. The mass spectrometer was operated by the Xcalibur software (Thermo Fisher). DDA scan settings on full MS level included an automatic gain control target value set at 3E6 in the 300–1 650 m/z range with a maximum injection time of 50 ms and a resolution of 60 000 at m/z 200. At the MS/MS level, the automatic gain control target value was 5E4 with a maximum injection time of 25 ms and a resolution of 15 000 at m/z 200. For MS/MS events only, precursor ions with 2–6 charges which were not on the 15 s dynamic exclusion list were isolated in a 1.4 m/z window. Fragmentation was performed by higher-energy collisional dissociation (HCD) set at 27 eV. DIA was performed with one full MS event followed by 10 MS/MS windows in one cycle. The full MS settings included an automatic gain control target value set at 3E6 in the 300–1 650 m/z range with a maximum injection time of 50 ms and a resolution of 120 000 at m/z 200. DIA precursor windows ranged from 350 m/z (lower boundary of the first window) to 1650 m/z

(upper boundary of the 60th window). MS/MS settings included an automatic gain control target set at 3E6 for the precursor windows with a maximum injection time of 54 ms and a resolution of 30 000 at m/z 200.

Raw data analysis   The MS data of the fractionated pools were used to generate a DDA-library with the MaxQuant software (version 1.6.1.13, Cox & Mann, 2008). The spectral library was used to search the MS data of the single-shot patient samples in the Spectronaut software, version 12.0.20491.12.30484 (Biognosys AG), for final protein identification and quantitation. All searches were performed against the human UniProt reference proteome of canonical and isoform sequences with 93 786 entries downloaded in March 2018. Searches used carbamidomethylation as fixed modification and acetylation of the protein N-terminus, oxidation of methionines and deamidation of asparagine or glutamine as variable modifications. Default settings were used for other parameters. In brief, a trypsin/P proteolytic cleavage rule was used, permitting a maximum of two miscleavages and a peptide length of 7–52 amino acids.
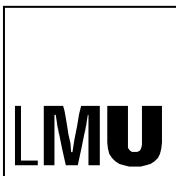
KLINIKUM

DER UNIVERSITÄT MÜNCHEN

CAMPUS INNENSTADT | PERINATALZENTRUM
NEONATOLOGIE DER KINDERKLINIK
AN DER FRAUENKLINIK MAISTRASSE

LMU

PROTEOM-Studie
bei Neu- und Frühgeborenen

Informationsschrift
und Einwilligungserklärung

Eltern

STUDIENLEITER

Prof. Dr. med. Orsolya Genzel-Boroviczény

Perinatalzentrum – Neonatologie der Kinderklinik

Dr. von Haunersches Kinderspital

Kinderklinik und Kinderpoliklinik der

Ludwig-Maximilians-Universität München

Maistraße 11

80337 München

Tel.: 089-4400-54535

E-Mail: genzel@med.uni-muenchen.de

Sehr geehrte Eltern,

wir bitten Sie um die Teilnahme Ihres Kindes an der Proteom-Studie. Die folgenden Seiten sollen Sie über unsere Studie informieren und darüber, was eine Teilnahme für Sie und Ihr Kind bedeuten würde. Nachdem Sie dieses Schreiben gelesen und mit Ihrem Partner besprochen haben, können Sie sich frei entscheiden, ob Sie teilnehmen möchten.

Warum interessieren wir uns für Proteine?

Proteine erfüllen im menschlichen Körper die vielfältigsten Funktionen. Es wird immer deutlicher, dass unser Körper auf bestimmte Zustände oder Krankheiten sogar mit „Schutz-Proteinen" reagieren kann. Zum Proteom von Neugeborenen liegen aber noch kaum Daten vor.

Warum fragen wir Sie?

Ziel unserer Studie ist es, Zusammenhänge zwischen Erkrankungen – seien sie nun metabolischen, neurologischen, allergischen oder sonstigen Ursprungs – und der Zusammensetzung der Proteine im Blut von Neugeborenen zu finden. Jedes Ergebnis soll mit der Blutzusammensetzung gesunder Neugeborener bzw. Neugeborener ohne familiäre Vorbelastung verglichen werden. Es werden daher sowohl gesunde wie weniger gesunde Kinder als Teilnehmer benötigt.

Die Teilnahme ist freiwillig!

Ihre Teilnahme und die Ihres Kindes ist freiwillig. Sie können jederzeit, auch ohne Angabe von Gründen, die Teilnahme widerrufen, ohne dass Ihrem Kind oder Ihnen dadurch irgendwelche Nachteile für die weitere ärztliche Versorgung entstehen würden. Einen Widerruf der Einwilligung bzw. den Rücktritt von der Studie müssten Sie nicht begründen. Im Falle eines Widerrufes werden die bis zu diesem Zeitpunkt erhobenen Daten irreversibel anonymisiert, sodass keine Rückschlüsse auf Ihr Kind mehr möglich sind.

Wie läuft die Studie ab?

Im Rahmen der Studie erfolgen keine zusätzlichen Blutentnahmen oder Interventionen. Es werden lediglich im Rahmen von routinemäßigen Blutabnahmen 1-2 zusätzliche Blutstropfen entnommen. Schon anhand dieser kleinen Menge Blut kann eine enorme Zahl an Proteinen im Blut Ihres Kindes bestimmt werden. Dabei untersuchen wir:

1. Bei gesunden Neugeborenen: 1-2 Blutstropfen beim Neugeborenen-Screening

2. Bei Frühgeborenen: 1-2 Blutstropfen im Rahmen klinisch notwendiger Blutabnahmen am 1. Lebenstag, beim Neugeborenenscreening und vor Entlassung

3. Bei Kindern mit Asphyxie: 1-2 Blutstropfen im Rahmen klinisch notwendiger Blutabnahmen am 1. und 2. Lebenstag

4. Bei Kindern mit Infektionen: 1-2 Blutstropfen, abgenommen vor sowie 48h nach Behandlungsbeginn im Rahmen klinisch notwendiger Blutabnahmen

Welche Daten werden erhoben?

Im Rahmen der Studie sollen auch Daten zum Gesundheitszustand Ihres Kindes erhoben werden. Dies wären u.a. die Diagnose, Laborwerte, sowie eben die Zusammensetzung der Blut-Proteine. Zusätzlich bitten wir Sie als Eltern, einige Fragen zu Ihrem eigenen Gesundheitszustand und dem von engen Familienangehörigen zu beantworten. Diese Daten werden mit einer fortlaufenden Nummer pseudonymisiert und in einer Datenbank gesammelt.

Gibt es irgendwelche Risiken?

Da es sich bei der Studie um eine reine Datenerfassung anhand der gesammelten Blutstropfen handelt, bestehen keine studienspezifischen Risiken. Zusätzliche Blutentnahmen nur zum Zwecke dieser Studie wird es nicht geben, lediglich die Entnahme der je 1-2 zusätzlichen Blutstropfen bei ohnehin anstehenden Blutentnahmen.

Gibt es irgendwelche Vorteile?

Sie oder Ihr Kind werden keinen unmittelbaren persönlichen Nutzen aus dieser Studie ziehen. Die Studie möchte dazu beitragen, im Allgemeinen mehr über mögliche Schutzfunktionen von Proteinen und deren Zusammenhang mit Erkrankungen unterschiedlicher Art herauszufinden. Dies könnte aber in Zukunft möglicherweise die Behandlung von Patienten mit vergleichbaren Erkrankungen verbessern.

Wer kann die persönlichen Daten Ihres Kindes einsehen?

Die zusammengetragenen Studienergebnisse werden streng vertraulich behandelt. Unsere Datenbank wird nur auf den geschützten Rechnern des Klinikums der Universität München gespeichert. Ausschließlich Studienbefugte, also Ihre behandelnden Ärzte, die Studienleitung und die Doktoranden, haben Zugriff auf diese verschlüsselten Daten.

Ihre persönlichen Daten oder die Ihres Kindes werden nie in irgendeiner Form veröffentlicht oder weitergegeben werden. Diejenigen Daten, die keinen Rückschluss auf Ihr Kind zulassen, werden gemeinsam von der Studiengruppe ausgewertet. Im Falle eines Widerrufes Ihrer Einwilligungserklärung würde eine irreversible Anonymisierung aller Daten erfolgen, sodass definitiv keine Rückschlüsse mehr auf Ihr Kind möglich sind – nicht einmal mehr für die Mitglieder der Studiengruppe.

Wie lange werden die Daten Ihres Kindes aufgehoben?

Die Behandlungsakten Ihres Kindes werden entsprechend der gesetzlichen Bestimmungen verwahrt, unabhängig von einer Teilnahme an der Studie. Die Unterlagen speziell zu dieser Studie werden nach Abschluss der Kohortenstudie 10 Jahre verwahrt und dann irreversibel anonymisiert.

Gibt es ein Follow-Up?

Bei unserer Studie gibt es kein klassisches Follow-Up. D.h., dass es nach Ihrem Krankenhausaufenthalt keine weiteren Termine, Untersuchungen, Blutentnahmen o.Ä. als Verlaufsbeobachtung geben wird. Wir bitten Sie dennoch um Ihre Erlaubnis, Sie unter Umständen kontaktieren zu dürfen, um Sie in knapper Form zur Entwicklung Ihres Kindes zu befragen.

Gibt es einen speziellen Versicherungsschutz?

Für die Studie besteht keine Patientenversicherung, da im Rahmen der Studie keine Intervention erfolgt.

Bei offen gebliebenen Fragen können Sie sich gerne jederzeit an uns wenden!

Datenschutz

Bei dieser Studie werden die Vorschriften über die ärztliche Schweigepflicht und den Daten-schutz eingehalten. Es werden persönliche Daten erhoben, gespeichert und verschlüsselt (pseu-donymisiert) ausgewertet, d.h. weder der Name noch die Initialen oder das Geburtsdatum er-scheinen im Verschlüsselungscode.

Im Falle des Widerrufs Ihrer Einwilligung bzw. 10 Jahre nach Abschluss der Kohortenstudie werden die gespeicherten Daten irreversibel anonymisiert und die Originalunterlagen vernich-tet.

Eine Entschlüsselung erfolgt lediglich in Fällen, in denen es die Sicherheit Ihres Kindes erfordert („medizinische Gründe") oder falls es zu Änderungen in der wissenschaftlichen Fragestellung kommt („wissenschaftliche Gründe").

Sie haben jederzeit die Möglichkeit, Auskunft über die Ihr Kind betreffenden personenbezoge-nen Daten zu erhalten sowie ggf. deren Berichtigung oder Löschung zu verlangen.

Im Falle von Veröffentlichungen der Studienergebnisse bleibt die Vertraulichkeit der persönli-chen Daten gewährleistet.

Bei Beschwerden bzgl. des Datenschutzes haben Sie die Möglichkeit, sich an die Datenschutz-aufsichtsbehörde zu wenden.

Datenschutzverantwortliche:

Datenschutzbeauftragter zentral:
Gerhard Mayer, Behördlicher
Datenschutzbeauftragter,
Pettenkoferstrasse 8, 80336 München
Tel. 089/4400-58454, Fax –58452
datenschutz@med.uni-muenchen.de

Datenschutzbeauftragter lokal:
Prof. Dr. Johannes Hübner
Dr. von Haunersches Kinderspital
Johannes.hübner@med.uni-muenchen.de

Datenschutzaufsichtsbehörde:
Prof. Dr. T. Petri, Bayerische Landesbeauftragte
für den Datenschutz
Postfach 22 12 19, 80502 München
Wagmüllerstraße 18, 80538 München
Tel. 089 212672-0, Fax 089 212672-50
poststelle@datenschutz-bayern.de

KLINIKUM

DER UNIVERSITÄT MÜNCHEN

CAMPUS INNENSTADT | PERINATALZENTRUM
NEONATOLOGIE DER KINDERKLINIK
AN DER FRAUENKLINIK MAISTRASSE

Einwilligungserklärung

PROTEOM-Studie bei Neu- und Frühgeborenen

Wir wurden gebeten, dass unser Kind an der Proteom-Studie zur Untersuchung einer möglichen Korrelation zwischen der Protein-Zusammensetzung im Blut unseres Kindes mit verschiedenen Erkrankungen teilnimmt. Der Zweck der Teilnahme wurde uns erklärt. Wir wurden über mögliche Vorteile und Risiken einer Teilnahme informiert. Wir hatten Gelegenheit, Fragen zu dem Projekt zu stellen und haben auf diese ausreichend Antworten bekommen. Wir sind uns bewusst, dass unsere Teilnahme freiwillig ist und wir jederzeit ohne besondere Begründung unsere Zustimmung widerrufen können. Daraus entsteht uns und insbesondere unserem Kind kein Nachteil. Die Ethikkommission des Fachbereiches Medizin der Ludwig-Maximilians-Universität hat der Durchführung zugestimmt.

Bei dieser Studie werden die Vorschriften über die ärztliche Schweigepflicht und den Datenschutz eingehalten. Es werden persönliche Daten und Befunde erhoben, gespeichert und in pseudonymisierter Form nach Maßgabe der Patientenaufklärung ausgewertet.

Wir/Ich, _____ **(Name/n)**

☐ **sind/bin mit der Erhebung und Verwendung persönlicher Daten und Befunddaten nach Maßgabe der Patienteninformation einverstanden und nehme/n an der Studie teil.**

☐ **lehne/n die Teilnahme ab.**

Name des **Kindes**	
Geburtsdatum: (T) |_____| (M) |_____| (J) |_____|	
Name der **Mutter**	Unterschrift der **Mutter**
	Datum (heute): (T) |_____| (M) |_____| (J) |_____|
Name des **Vaters**	Unterschrift des **Vaters**
	Datum (heute): (T) |_____| (M) |_____| (J) |_____|
Name der Person, die das Aufklärungsgespräch geführt hat	Unterschrift der aufklärenden Person
	Datum (heute): (T) |_____| (M) |_____| (J) |_____|

Vorwort

Vielen Dank, dass Sie sich die Zeit nehmen, folgenden Fragebogen zu unserer Proteom-Studie bei Neu- und Frühgeborenen auszufüllen. Sie werden dafür lediglich etwa **5-15 Minuten** benötigen.

Zum Zwecke der Vollständigkeit unserer Datenerhebung möchten wir Sie darum bitten, auch bei nicht-zutreffenden Sachverhalten die freien Felder **durchzustreichen** bzw. „**Nein**" anzukreuzen. Dadurch können wir diejenigen Dinge, die nicht zutreffen, besser von jenen unterscheiden, zu denen Sie tatsächlich keine Angaben machen konnten oder wollten.

In unsere finale Datenauswertung werden alle Ihre Angaben ausschließlich in streng **pseudonymisierter** Form eingehen. Alles, was Sie uns im Rahmen der Studie mitteilen, unterliegt der **ärztlichen Schweigepflicht.**

Bei Fragen stehen wir Ihnen gerne zur Verfügung.

Angaben zur Person (Mutter)

Name, Vorname ....…………………………………………….   Geburtsdatum …………………….............

Gesundheit von Mutter und Vater

1. Ist bei Ihnen, der Mutter des Babys bzw. dem (leiblichen) Vater, jemals eine der folgenden Erkrankungen von einem Arzt diagnostiziert worden? Wenn ja welche?

	Mutter	**Vater**
- Herzkreislauf-Erkrankungen	………………………….	………………………….
- Diabetes mellitus	Typ 1 ☐ 2 ☐ Nein ☐	1 ☐ 2 ☐ Nein ☐
- Schilddrüsenerkrankungen	………………………….	………………………….
- Gerinnungsstörungen	………………………….	………………………….
- Herpes simplex (z.B. Lippenherpes)	Ja ☐ Nein ☐	Ja ☐ Nein ☐
- Autoimmunerkrankungen (z.B. Rheumatoide Arthritis)	………………………….	………………………….
- Magen-Darmerkrankungen (M. Crohn, ulzerative Colitis)	………………………….	………………………….
- Erbkrankheiten	………………………….	………………………….

   Falls eine Erbkrankheit vorliegt: Sind Sie mit dem Vater des Kindes verwandt?

   Nein   ☐   Ja   ☐   Wie? (z.B. Cousin 3. Grades) ……………………………………………

   - Sonstige Erkrankungen (z.B. Stoffwechsel-/neurolog. Erkr.) …………………………   …………………………

2. Wurde jemals eine der folgenden Erkrankungen bei Ihnen oder dem Vater des Kindes von einem Arzt diagnostiziert?

	Mutter	**Vater**
Allergisches Asthma	Ja ☐ Nein ☐	Ja ☐ Nein ☐
Ekzem/Neurodermitis	Ja ☐ Nein ☐	Ja ☐ Nein ☐
Heuschnupfen	Ja ☐ Nein ☐	Ja ☐ Nein ☐
Welches Allergen? ………………………		…………………………
Sonstige bekannte Allergien / Unverträglichkeiten	Ja ☐ Nein ☐	Ja ☐ Nein ☐
Welche? …………………………		…………………………

3. Falls Sie, die Mutter, an einer allergischen Erkrankung leiden (Heuschnupfen, allergisches Asthma bronchiale, Neurodermitis/atopische Dermatitis/endogenes Ekzem), waren die Symptome **während der Schwangerschaft** verändert?

Nein, gleich ☐   Weniger/schwächere Symptome ☐   Mehr/stärkere Symptome ☐
Keine Allergie (o.Ä.) ☐

Krankheiten in der Familie

1. Wurde bei **einem Ihrer Verwandten** (Verwandte der Mutter) eine der folgenden Erkrankungen von einem Arzt diagnostiziert? (Eltern, Geschwister)

mütterlicherseits

Allergie / Unverträglichkeit	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	
Allergisches Asthma	Ja ☐	Bei wem? ……………………………………………
	Nein ☐	
Neurodermitis	Ja ☐	Bei wem? ……………………………………………
	Nein ☐	
Diabetes	Ja ☐	Typ 1 ☐ Bei wem? …………….………………
	Nein ☐	Typ 2 ☐ Bei wem? …………….………………
		(= „Altersdiabetes")
Erbkrankheiten	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	
Andere	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	

2. Wurde bei einem **Verwandten des Vaters** des Kindes eine der folgenden Erkrankungen von einem Arzt diagnostiziert? (Eltern, Geschwister)

väterlicherseits

Allergie / Unverträglichkeit	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	
Allergisches Asthma	Ja ☐	Bei wem? ……………………………………………
	Nein ☐	
Neurodermitis	Ja ☐	Bei wem? ……………………………………………
	Nein ☐	
Diabetes	Ja ☐	Typ 1 ☐ Bei wem? …………….………………
	Nein ☐	Typ 2 ☐ Bei wem? …………….………………
		(= „Altersdiabetes")
Erbkrankheiten	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	
Andere	Ja ☐	Welche? ……………………………………………
		Bei wem? ……………………………………………
	Nein ☐	

Gynäkologische Fragen

1. Wie viel haben Sie **vor** Beginn der Schwangerschaft gewogen?   ……………………… kg

2. In welchem Alter hatten Sie das erste Mal Ihre Menstruationsblutung?   ……………………… Jahre

3. Wie lange dauert Ihr durchschnittlicher **Zyklus** (vom 1. Tag Ihrer Regelblutung bis zum 1. Tag der nächsten Regelblutung)?   ……………………… Tage

4. Haben Sie die **Pille** eingenommen?

   Ja ☐   Für wie lange?   ……………………… Jahre

   Einnahme wurde beendet vor: ……………………… Jahr(en)

   Nein ☐   Name des zuletzt verwendeten Präparats: ……………………………………

5. Haben Sie **andere** hormonelle oder pharmakologische Verhütungsmaßnahmen angewendet?

   Ja ☐   Für wie lange?   ……………………… Jahre

   Art der Verhütungsmethode   …………………………………

   Nein ☐

6. Hatten Sie bereits relevante **gynäkologische Voroperationen**?

   Ja ☐   Welche?   ……………………………
   Wie viele?   ……………………………

   Nein ☐

Das Jahr   V O R   der Schwangerschaft

1. Wurden Sie im Jahr vor Ihrer Schwangerschaft **geimpft**?

   Ja ☐   Welche Impfung(en)?   ……………………………………………………
   Nein ☐

2. Haben Sie im Jahr vor Ihrer Schwangerschaft **Medikamente** eingenommen, darunter evtl. Schmerz- oder Schlafmittel, Vitamine oder Mineralstoffe? Wenn ja, welche? (Name des Präparats, falls möglich)

   Ja ☐   Medikamente   ……………………………………………
   Schmerzmittel   ……………………………………………
   Schlafmittel   ……………………………………………
   Vitaminpräparate   ……………………………………………
   Folsäure   ……………………………………………
   Mineralstoffe   ……………………………………………

   Nein ☐

3. Hatten Sie im Jahr vor Ihrer Schwangerschaft **Operationen** jeglicher Art?

   Ja ☐   Welche?   ……………………………………………………………

   Nein ☐

4. Haben Sie im Jahr vor Ihrer Schwangerschaft eine größere **Reise** ins Ausland unternommen?

   Ja ☐   Wohin?   ……………………………………………………………

   Nein ☐   Wann? Wie viele Monate etwa vor Beginn der Schw.?   ………

Schwangerschaftsverlauf

1. Wurden Sie während Ihrer Schwangerschaft **geimpft**?

   Ja ☐   Welche Impfung(en)?   …………………………………………………

   Nein ☐

2. Haben Sie während Ihrer Schwangerschaft **Medikamente** eingenommen, darunter evtl. Schmerz- oder Schlafmittel, Vitamine oder Mineralstoffe? Wenn ja, welche? (Name des Präparats, falls möglich)

   Ja ☐   Medikamente   …………………………………………………

                 Schmerzmittel   …………………………………………………

                 Schlafmittel   …………………………………………………

                 Vitaminpräparate   …………………………………………………

                 Folsäure   …………………………………………………

                 Mineralstoffe   …………………………………………………

   Nein ☐

3. Hatten Sie während Ihrer Schwangerschaft (über die üblichen Vorsorgeuntersuchungen hinaus) irgendwelche **medizinischen Eingriffe**?

   Ja ☐   Welche?   ……………………………………………………………

   Nein ☐

4. Mussten Sie während Ihrer Schwangerschaft **stationär** in einem Krankenhaus aufgenommen werden?

   Ja ☐   Ursache?   ……………………………………………………………

   Nein ☐

5. Haben Sie während der Schwangerschaft in **städtischem** oder **ländlichem** Umfeld gelebt?

   städtisch ☐   ländlich ☐   gemischt ☐

   Seit wie vielen Jahren leben Sie bereits in dieser Umgebung? ………… Jahre

6. Haben Sie während Ihrer Schwangerschaft eine größere **Reise** ins Ausland unternommen?

   Ja ☐   Wohin?   ……………………………………………………………

   Nein ☐   Wann?   1. Drittel ☐   2. Drittel ☐   3. Drittel ☐   der Schw.

Überblick Schwangerschaften

(War dies Ihre allererste Schwangerschaft überhaupt? Ja? ☐ Dann können Sie dieses Kapitel überspringen.)

1. Waren unter allen Ihren bisherigen Schwangerschaften auch Zwillings- bzw. Drillings oder Vierlings-Geburten?

   Ja ☐   ………… Zwillingsgeburt   ………… Drillingsgeburt   ………… Vierlingsgeburt

   Nein ☐

2. Wie viele Ihrer Kinder kamen **zu früh** auf die Welt (vor der 37. Schwangerschaftswoche, einschließlich Ihres gerade geborenen Kindes)?

   ……………………… Frühgeburten   Keine ☐

3. Wie haben Sie Ihre Kinder geboren (einschließlich Ihres gerade geborenen Kindes)?

   <u>Anzahl</u> Spontangeburten ………… Zangengeburten/Vakuumextraktionen ………. Kaiserschnitte ………….

4. Hatten Sie bereits einen Abort, eine stille Geburt oder eine Bauchhöhlenschwangerschaft?

       Ja   ☐   <u>Anzahl</u> …….. Abort   …….. Stille Geburt   …….. Bauchhöhlenschw.
       Nein ☐

5. Falls es leibliche Geschwister gibt, wie viele?   ………… Mädchen   ………… Jungen

6. Leidet eines bzw. mehrere dieser Geschwister an einer von einem Arzt diagnostizierten **Grunderkrankung**?

     (Wenn keine Geschwister, überspringen Sie Frage 6)

Allergie / Unverträglichkeit	Ja	☐	Welche?	………………………………………………………
			Bei wem?	………………………………………………………
	Nein	☐		
Asthma	Ja	☐	Bei wem?	………………………………………………………
	Nein	☐		
Neurodermitis	Ja	☐	Bei wem?	………………………………………………………
	Nein	☐		
Diabetes	Ja	☐	Typ 1 ☐	Bei wem? ………………………………………….
	Nein	☐	Typ 2 ☐	
Erbkrankheiten	Ja	☐	Welche?	………………………………………………………
			Bei wem?	………………………………………………………
	Nein	☐		
Andere	Ja	☐	Welche?	………………………………………………………
			Bei wem?	………………………………………………………
	Nein	☐		

Reproduktionsmedizin

1. Haben Sie bei diesem Kind medizinische Hilfe in Anspruch genommen, um schwanger zu werden?

    Nein ☐   *(Wenn nein, können Sie dieses Kapitel überspringen)*

                                          Gesamtzahl Versuche/eingesetzte Embryonen

Ja	☐	☐	Ausschließlich Hormontherapie	
		☐	In-vitro- Fertilisation (IVF)	…………………../…………………..
		☐	Intra-zytoplasmatische Spermieninjektion (ICSI)	…………………../…………………..
		☐	Testikuläre Spermienextraktion mit ICSI (TESE)	…………………../…………………..
		☐	Mikrochir. epididymale Spermatozoenaspiration mit ICSI (MESA)	………./……..

2. Falls eine Form der künstlichen Befruchtung stattgefunden hat, stammte der Samen von Ihrem Mann?

       Ja   ☐   Nein, von einem Spender   ☐   Keine künstliche Befr. ☐

3. Haben Sie im Rahmen der Maßnahmen eine Hormontherapie erhalten?

       Ja   ☐   Welche? / Name des Präparats ……………………………………………………………

       Nein ☐

4. Hat der Vater des Kindes im Rahmen der Maßnahmen eine Hormontherapie erhalten?

       Ja   ☐   Welche? / Name des Präparats ……………………………………………………
       Nein ☐

Konsumverhalten

1. Sind Sie oder Ihr Mann **Raucher**?

Nicht-Raucher	Mutter ☐	Vater ☐
Ehemaliger Raucher	Mutter ☐	Vater ☐
Raucher	Mutter ☐	Vater ☐

   (Sollten Sie beide schon immer Nicht-Raucher gewesen sein, springen Sie zu Frage 6)

2. Haben Sie, die Mutter, (jemals) **vor** der Schwangerschaft geraucht?

   Ja ☐   Wie viele Zigaretten? ………………… pro Tag / ……………… pro Woche
   Wie viele Jahre haben Sie geraucht? ……………………………………… Jahre
   Ggf.: Wann haben Sie mit dem Rauchen aufgehört? Vor ……. Jahren

   Nein ☐

3. Haben Sie, die Mutter, **während** der Schwangerschaft geraucht?

   Ja ☐   Wie viele Zigaretten pro Tag?   ……… pro Tag / ……… pro Woche
   Nein ☐

4. Wurde in Ihrer **Wohnung** / Ihrem **Haus** in den letzten 12 Monaten geraucht?
   Zigaretten die auf dem Balkon / der Terasse geraucht wurden, sollen nicht mitgezählt werden.

   Ja ☐

Anzahl der Zigaretten pro Tag	………………… pro Tag
Davon Mutter	………………… pro Tag
Davon Partner	………………… pro Tag
Davon andere	………………… pro Tag

   Nein ☐

5. Wurde **zuvor** in Ihrer Wohnung / Ihrem Haus geraucht, in den letzten 12 Monaten jedoch nicht mehr?

   Ja, bis vor 12 Monaten wurde im Haushalt geraucht ☐

   Nein, es wurde auch in den letzten 12 Monaten geraucht ☐

   Ich/wir habe/n nie in der Wohnung / im Haus geraucht ☐

6. Haben Sie **vor** Beginn der Schwangerschaft regelmäßig Alkohol konsumiert?

   Ja ☐   Anzahl der alkoholischen Getränke   ……… pro Tag / ……… pro Woche
   Nein ☐

7. Haben Sie **während** der Schwangerschaft regelmäßig Alkohol konsumiert?

   Ja ☐   Anzahl der alkoholischen Getränke   ……… pro Tag / ……… pro Woche
   Nein ☐

Tierkontakte

1. Sind Sie zeitweise mit Haustieren / engem Tierkontakt aufgewachsen? Mutter ☐ Vater ☐   Nein ☐

2. Gab es im Jahr **vor** der Schw. ein Haustier in Ihrem Haushalt? ………………………………………………………….

3. Gab es **während** der Schw. ein Haustier in Ihrem Haushalt? ………………………………………………………

4. Haben Sie, die **Mutter**, beruflich oder in Ihrer Freizeit mit Tieren zu tun? …………………………………….

5. Hat der **Vater** beruflich oder in seiner Freizeit mit Tieren zu tun? ……………………………………………….

   Ggf.: Ihr Kommentar zum Thema Tierkontakte: ………………………………………………………………………….

Allgemeine Fragen

Die Fragen 1-4 beziehen sich v.a. auf Ihre ethnische **Abstammung**, weniger auf Ihre Staatsangehörigkeit.

(wenn immer gleich, gerne Abkürzungen verwenden)

mütter-licherseits

1. In welchem Land wurden **Sie** geboren?   Mutter …………………………………

2. In welchem Land wurden Ihre **Eltern** geboren?   Großmutter …………………………………

   Großvater …………………………………

väter-licherseits

3. In welchem Land wurde der **Vater** des Kindes geboren?   Vater …………………………………

4. In welchem Land wurden die **Eltern** des Vaters geboren?   Großmutter …………………………………

   Großvater …………………………………

5. Was sind Sie von **Beruf**?

   Mutter ………………………………………….   Vater ………………………………………….

6. Welchen **Bildungsweg** haben Sie genommen?

	Mutter	Vater
Kein Schulabschluss	☐	☐
Mittelschule	☐	☐
Mittlere Reife	☐	☐
(Fach-)Abitur	☐	☐
(Fach-)Hochschule, Universität	☐	☐
Anderer Bildungsweg	☐	☐

Fragen zum Vater

1. Größe des Vaters   ……. , ………….. m

2. Gewicht des Vaters   ………………….. kg

3. Alter des Vaters ………………….… Jahre

Abschlussfrage

Wenn Sie die hinter Ihnen liegende Schwangerschaft auf einer Skala von 1 („keine Belastung") bis 10 („unerträgliche Belastung") bewerten müssten, welche Punktzahl würden Sie vergeben?

…………………………

Nachwort

Wir möchten uns herzlich für Ihre Geduld und Gewissenhaftigkeit bei der Beantwortung unseres Fragebogens bedanken. Ihnen und Ihrem neuen Familienmitglied wünschen wir von Herzen alles Gute für Ihre gemeinsame Zukunft!

Optional: Ihre E-mail-Adresse für evtl. Rückfragen: ………………………………………………………………

Möchten Sie in ca. 2 Jahren per E-Mail über Ergebnisse unserer Studie informiert werden?   Ja ☐   Nein ☐

BIBLIOGRAPHY

Abdel Ghany, E. A., & Ali, A. A. 2012. Empirical antibiotic treatment and the risk of necrotizing enterocolitis and death in very low birth weight neonates. *Ann Saudi Med*, **32**(5), 521–6. (*cf.* p. 13)

Adams-Chapman, I. 2012. Long-term impact of infection on the preterm neonate. *Semin Perinatol*, **36**(6), 462–70. (*cf.* p. 15)

Adkins, B., Leclerc, C., & Marshall-Clarke, S. 2004. Neonatal adaptive immunity comes of age. *Nature Reviews Immunology*, **4**(7), 553–564. (*cf.* pp. 8, 9)

Aebersold, R. 2003. A mass spectrometric journey into protein and proteome research. *J Am Soc Mass Spectrom*, **14**(7), 685–95. (*cf.* p. 17)

Aebersold, R., & Mann, M. 2016. Mass-spectrometric exploration of proteome structure and function. *Nature*, **537**(7620), 347–55. (*cf.* pp. 16, 96)

Aluvihare, V. R., Kallikourdis, M., & Betz, A. G. 2004. Regulatory T cells mediate maternal tolerance to the fetus. *Nat Immunol*, **5**(3), 266–71. (*cf.* p. 8)

Arora, N., Sadovsky, Y., Dermody, T. S., & Coyne, C. B. 2017. Microbial Vertical Transmission during Human Pregnancy. *Cell Host Microbe*, **21**(5), 561–567. (*cf.* p. 7)

Au, P. Y. B., Goedhart, C., Ferguson, M., Breckpot, J., Devriendt, K., Wierenga, K., Fanning, E., Grange, D. K., Graham, G. E., Galarreta, C., Jones, M. C., Kini, U., Stewart, H., Parboosingh, J. S., Kline, A. D., Innes, A. M., & Care for Rare Canada Consortium. 2018. Phenotypic spectrum of Au-Kline syndrome: a report of six new cases and review of the literature. *Eur J Hum Genet*, **26**(9), 1272–1281. (*cf.* p. 69)

Ayling, R. M., & Kok, K. 2018. Fecal Calprotectin. *Pages 161–190 of: Advances in Clinical Chemistry*. Elsevier. (*cf.* p. 73)

Baker, C. N., Thornsberry, C., & Facklam, R. R. 1981. Synergism, killing kinetics, and antimicrobial susceptibility of group A and B streptococci. *Antimicrob Agents Chemother*, **19**(5), 716–25. (*cf.* p. 14)

Bantscheff, M., Lemeer, S., Savitski, M. M., & Kuster, B. 2012. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem*, **404**(4), 939–65. (*cf.* p. 105)

Bartolomé, N., Rodríguez, L., Martínez, M. J., Ochoa, B., & Chico, Y. 2007. Upregulation of apolipoprotein B secretion, but not lipid, by tumor necrosis factor-alpha in rat hepatocyte cultures in the absence of extracellular fatty acids. *Ann N Y Acad Sci*, **1096**(Jan), 55–69. (*cf.* p. 71)

Bartolomé, N., Arteta, b., Martínez, M. J., Chico, Y., & Ochoa, B. 2008. Kupffer cell products and interleukin 1beta directly promote VLDL secretion and apoB mRNA up-regulation in rodent hepatocytes. *Innate Immun*, **14**(4), 255–66. (*cf.* p. 71)

BDWG, Biomarkers Definitions Working Group. 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics*, **69**(3), 89–95. (*cf.* p. 18)

Becker, K. L., Monaghan, K. G., & Silva, O. L. 1980. Immunocytochemical localization of calcitonin in Kulchitsky cells of human lung. *Arch Pathol Lab Med*, **104**(4), 196–8. (*cf.* p. 11)

Begue, P., Floret, D., Mallet, E., Raynaud, E. J., Safran, C., Sarlangues, J., & Teyssier, G. 1984. Pharmacokinetics and clinical evaluation of cefotaxime in children suffering with purulent meningitis. *J Antimicrob Chemother*, **14 Suppl B**, 161–5. (*cf.* p. 14)

Bennike, T. B., Fatou, B., Angelidou, A., Diray-Arce, J., Falsafi, R., Ford, R., Gill, E. E., van Haren, S. D., Idoko, O. T., Lee, A. H., Ben-Othman, R., Pomat, W. S., Shannon, C. P., Smolen, K. K., Tebbutt, S. J., Ozonoff, A., Richmond, P. C., van den Biggelaar, A. H. J., Hancock, R. E. W., Kampmann, B., Kollmann, T. R., Levy, O., & Steen, H. 2020. Preparing for Life: Plasma Proteome Changes and Immune System Development During the First Week of Human Life. *Front Immunol*, **11**, 578505. (*cf.* pp. 21, 67)

Bizzarro, M. J., Raskind, C., Baltimore, R. S., & Gallagher, P. G. 2005. Seventy-five years of neonatal sepsis at Yale: 1928-2003. *Pediatrics*, **116**(3), 595–602. (*cf.* p. 7)

Boghossian, N. S., Page, G. P., Bell, E. F., Stoll, B. J., Murray, J. C., Cotten, C. M., Shankaran, S., Walsh, M. C., Laptook, A. R., Newman, N. S., Hale, E. C., McDonald, S. A., Das, A., Higgins, R. D., Eunice Kennedy Shriver National Institute of Child, Health, & Human Development Neonatal Research, Network. 2013. Late-onset sepsis in very low birth weight infants from singleton and multiple-gestation births. *J Pediatr*, **162**(6), 1120–4, 1124 e1. (*cf.* p. 6)

Born, S., Dame, C., Matthäus-Krämer, C., Schlapbach, L. J., Reichert, F., Schettler, A., Schwarzkopf, D., Thomas-Rüddel, D., Proquitté, H., Reinhart, K., & Fleischmann-Struzek, C. 2021. Epidemiology of Sepsis Among Children and Neonates in Germany: Results From an Observational Study Based on Nationwide Diagnosis-Related Groups Data Between 2010 and 2016. *Crit Care Med*, **49**(7), 1049–1057. (*cf.* pp. 5, 6)

Boursi, B., Mamtani, R., Haynes, K., & Yang, Y. X. 2015. The effect of past antibiotic exposure on diabetes risk. *Eur J Endocrinol*, **172**(6), 639–48. (*cf.* p. 13)

Bristow, R. E., Smith, A., Zhang, Z., Chan, D. W., Crutcher, G., Fung, E. T., & Munroe, D. G. 2013. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecol Oncol*, **128**(2), 252–9. (*cf.* p. 19)

Bryan, C. S., John, J. F., Jr., Pai, M. S., & Austin, T. L. 1985. Gentamicin vs cefotaxime for therapy of neonatal sepsis. Relationship to drug resistance. *Am J Dis Child*, **139**(11), 1086–9. (*cf.* p. 14)

Bryckaert, M., Rosa, J.-P., Denis, C. V., & Lenting, P. J. 2014. Of von Willebrand factor and platelets. *Cellular and Molecular Life Sciences*, **72**(2), 307–326. (*cf.* p. 69)

Buhimschi, C. S., Buhimschi, I. A., Abdel-Razeq, S., Rosenberg, V. A., Thung, S. F., Zhao, G., Wang, E., & Bhandari, V. 2007a. Proteomic biomarkers of intra-amniotic inflammation: relationship with funisitis and early-onset sepsis in the premature neonate. *Pediatr Res*, **61**(3), 318–24. (*cf.* p. 22)

Buhimschi, C. S., Bhandari, V., Hamar, B. D., Bahtiyar, M. O., Zhao, G., Sfakianaki, A. K., Pettker, C. M., Magloire, L., Funai, E., Norwitz, E. R., Paidas, M., Copel, J. A., Weiner, C. P., Lockwood, C. J., & Buhimschi, I. A. 2007b. Proteomic profiling of the amniotic fluid to detect inflammation, infection, and neonatal sepsis. *PLoS Med*, **4**(1), e18. (*cf.* p. 22)

Buhimschi, C. S., Bhandari, V., Han, Y. W., Dulay, A. T., Baumbusch, M. A., Madri, J. A., & Buhimschi, I. A. 2009. Using proteomics in perinatal and neonatal sepsis: hopes and challenges for the future. *Curr Opin Infect Dis*, **22**(3), 235–43. (*cf.* p. 22)

Buhimschi, C. S., Bhandari, V., Dulay, A. T., Nayeri, U. A., Abdel-Razeq, S. S., Pettker, C. M., Thung, S., Zhao, G., Han, Y. W., Bizzarro, M., & Buhimschi, I. A. 2011. Proteomics mapping of cord blood identifies haptoglobin "switch-on" pattern as biomarker of early-onset neonatal sepsis in preterm newborns. *PLoS One*, **6**(10), e26111. (*cf.* pp. 22, 67, 71)

Buhimschi, I. A., & Buhimschi, C. S. 2010. The role of proteomics in the diagnosis of chorioamnionitis and early-onset neonatal sepsis. *Clin Perinatol*, **37**(2), 355–74. (*cf.* pp. 22, 73)

Buhimschi, I. A., Christner, R., & Buhimschi, C. S. 2005. Proteomic biomarker analysis of amniotic fluid for identification of intra-amniotic inflammation. *BJOG*, **112**(2), 173–81. (*cf.* pp. 21, 22)

Bukosza, E. N., Kornauth, C., Hummel, K., Schachner, H., Huttary, N., Krieger, S., Nöbauer, K., Oszwald, A., Razzazi Fazeli, E., Kratochwill, K., Aufricht, C., Szénási, G., Hamar, P., & Gebeshuber, C. A. 2020. ECM Characterization Reveals a Massive Activation of Acute Phase Response during FSGS. *Int J Mol Sci*, **21**(6). (*cf.* p. 72)

Chatziioannou, A. C., Wolters, J. C., Sarafidis, K., Thomaidou, A., Agakidis, C., Govorukhina, N., Kuivenhoven, J. A., Bischoff, R., & Theodoridis, G. 2018. Targeted LC-MS/MS for the evaluation of proteomics biomarkers in the blood of neonates with necrotizing enterocolitis and late-onset sepsis. *Anal Bioanal Chem*, **410**(27), 7163–7175. (*cf.* pp. 23, 80)

Chiesa, C., Signore, F., Assumma, M., Buffone, E., Tramontozzi, P., Osborn, J. F., & Pacifico, L. 2001. Serial measurements of C-reactive protein and interleukin-6 in the immediate postnatal period: reference intervals and analysis of maternal and perinatal confounders. *Clin Chem*, **47**(6), 1016–22. (*cf.* p. 11)

Child, C. G., & Turcotte, J. G. 1964. Surgery and portal hypertension. *Major Probl Clin Surg*, **1**, 1–85. (*cf.* p. 19)

Chong, P. K., Lee, H., Zhou, J., Liu, S.-C., Loh, M. C. S., Wang, T. T., Chan, S. P., Smoot, D. T., Ashktorab, H., So, J. B. Y., Lim, K. H., Yeoh, K. G., & Lim, Y. P. 2010. ITIH3 is a potential biomarker for early detection of gastric cancer. *J Proteome Res*, **9**(7), 3671–9. (*cf.* p. 72)

Christensen, R. D., Rothstein, G., Hill, H. R., & Hall, R. T. 1985. Fatal early onset group B streptococcal sepsis with normal leukocyte counts. *Pediatr Infect Dis*, **4**(3), 242–5. (*cf.* p. 10)

Christensen, R. D., Henry, E., Jopling, J., & Wiedmeier, S. E. 2009. The CBC: reference ranges for neonates. *Semin Perinatol*, **33**(1), 3–11. (*cf.* p. 10)

Clark, R. H., Bloom, B. T., Spitzer, A. R., & Gerstmann, D. R. 2006. Empiric use of ampicillin and cefotaxime, compared with ampicillin and gentamicin, for neonates at risk for sepsis is associated with an increased risk of neonatal death. *Pediatrics*, **117**(1), 67–74. (*cf.* p. 14)

Claus, R. A., Bockmeyer, C. L., Sossdorf, M., & Lösche, W. 2010. The balance between von-Willebrand factor and its cleaving protease ADAMTS13: biomarker in systemic inflammation and development of organ failure? *Curr Mol Med*, **10**(2), 236–48. (*cf.* p. 69)

Consortium, International Human Genome Sequencing. 2004. Finishing the euchromatic sequence of the human genome. *Nature*, **431**(7011), 931–945. (*cf.* p. 16)

Cox, J. 2018. Perseus and statistics. Lecture (L21) at 10th MaxQuant Summer School on Computational Mass Spectrometry-Based Proteomics (July 8th to 13th 2018, Barcelona, Spain). https://www.youtube.com/watch?v=V0evsE0CQgA. (Last accessed: 2022-06-24). (*cf.* pp. 26, 27)

Cox, J., & Mann, M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*, **26**(12), 1367–72. (*cf.* p. 107)

Cox, J., & Mann, M. 2012. 1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data. *BMC Bioinformatics*, **13 Suppl 16**, S12. (*cf.* pp. 28, 48, 49, 56, 57)

Crick, F. H. 1958. On protein synthesis. *Symp Soc Exp Biol*, **12**, 138–63. (*cf.* p. 15)

Croxatto, A., Prod'hom, G., & Greub, G. 2012. Applications of MALDI-TOF mass spectrometry in clinical diagnostic microbiology. *FEMS Microbiology Reviews*, **36**(2), 380–407. (*cf.* p. 17)

Cytiva, Cytiva Global Life Sciences Solutions USA LLC (formerly part of GE Healthcare). LE Whatman®903 Protein Saver Snap-apart Card. https://www.cytivalifesciences.com/en/de/shop/whatman-laboratory-

`filtration/whatman-dx-components/blood-collection-cards-and-accessories/903-proteinsaver-snap-apart-card-p-00677`. (Last accessed: 2022-06-24). (*cf.* p. 34)

Dahlin, A. 2015. ESI positive mode (21589986840).jpg. `https://commons.wikimedia.org/wiki/File:ESI_positive_mode_(21589986840).jpg`. Graphic licensed under Creative Commons (CC BY 2.0 ↗) (Last accessed: 2022-06-24). (*cf.* p. 97)

Dandona, P., Nix, D., Wilson, M. F., Aljada, A., Love, J., Assicot, M., & Bohuon, C. 1994. Procalcitonin increase after endotoxin injection in normal subjects. *J Clin Endocrinol Metab*, **79**(6), 1605–8. (*cf.* p. 11)

Darmstadt, G. L., Saha, Samir, Ahmed, Nawshad, Khatun, M., & Chowdhury, M. A. K. 2003. The skin as a potential portal of entry for invasive infections in neonates. *Perinatology*, **5**, 205–212. (*cf.* pp. 7, 9)

de Godoy, L. M. F., Olsen, J. V., Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., & Mann, M. 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, **455**(7217), 1251–1254. (*cf.* p. 16)

de Louvois, J., Halket, S., & Harvey, D. 2005. Neonatal meningitis in England and Wales: sequelae at 5 years of age. *Eur J Pediatr*, **164**(12), 730–4. (*cf.* p. 14)

Dentici, M. L., Barresi, S., Niceta, M., Pantaleoni, F., Pizzi, S., Dallapiccola, B., Tartaglia, M., & Digilio, M. C. 2018. Clinical spectrum of Kabuki-like syndrome caused by HNRNPK haploinsufficiency. *Clin Genet*, **93**(2), 401–407. (*cf.* p. 69)

Diz, A. P., Carvajal-Rodriguez, A., & Skibinski, D. O. 2011. Multiple hypothesis testing in proteomics: a strategy for experimental work. *Mol Cell Proteomics*, **10**(3), M110 004374. (*cf.* p. 17)

Dobryszycka, W. 1997. Biological functions of haptoglobin–new pieces to an old puzzle. *Eur J Clin Chem Clin Biochem*, **35**(9), 647–54. (*cf.* p. 67)

Dong, Y., & Speer, C. P. 2015. Late-onset neonatal sepsis: recent developments. *Arch Dis Child Fetal Neonatal Ed*, **100**(3), F257–63. (*cf.* pp. 5, 6, 7)

Du Clos, T. W. 2000. Function of C-reactive protein. *Ann Med*, **32**(4), 274–8. (*cf.* p. 11)

Eaton, J. W., Brandt, P., Mahoney, J. R., & Lee, J. T. 1982. Haptoglobin: A Natural Bacteriostat. *Science*, **215**(4533), 691–693. (*cf.* p. 67)

Edmond, K. M., Kortsalioudaki, C., Scott, S., Schrag, S. J., Zaidi, A. K., Cousens, S., & Heath, P. T. 2012. Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *Lancet*, **379**(9815), 547–56. (*cf.* p. 6)

Everett, A., & Ignjatovic, V. 2014. Paediatrics and proteomics: back to the beginning. *Proteomics Clin Appl*, **8**(11-12), 805–6. (*cf.* p. 21)

Fejgin, M., Amiel, A., Kaneti, H., Ben-Nun, I., & Beyth, Y. 1993. Fulminant sepsis due to group B beta-hemolytic streptococci following transcervical chorionic villi sampling. *Clin Infect Dis*, **17**(1), 142–3. (*cf.* p. 6)

Feng, J., Li, H., Li, J., Meng, P., Wang, L., Liu, C., Zhao, S., Sun, W., & Zhang, Y. 2020. hnRNPK knockdown alleviates NLRP3 inflammasome priming by repressing FLIP expression in Raw264.7 macrophages. *Redox Rep*, **25**(1), 104–111. (*cf.* p. 69)

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, **246**(4926), 64–71. (*cf.* p. 97)

Firth, M. A., Shewen, P. E., & Hodgins, D. C. 2005. Passive and active components of neonatal innate immune defenses. *Anim Health Res Rev*, **6**(2), 143–58. (*cf.* p. 9)

Fjalstad, J. W., Stensvold, H. J., Bergseng, H., Simonsen, G. S., Salvesen, B., Ronnestad, A. E., & Klingenberg, C. 2016. Early-onset Sepsis and Antibiotic Exposure in Term Infants: A Nationwide Population-based Study in Norway. *Pediatr Infect Dis J*, **35**(1), 1–6. (*cf.* p. 12)

Fleischmann-Struzek, C., Goldfarb, D. M., Schlattmann, P., Schlapbach, L. J., Reinhart, K., & Kissoon, N. 2018. The global burden of paediatric and neonatal sepsis: a systematic review. *Lancet Respir Med*, **6**(3), 223–230. (*cf.* p. 4)

Fluegge, K., Siedler, A., Heinrich, B., Schulte-Moenting, J., Moennig, M. J., Bartels, D. B., Dammann, O., von Kries, R., Berner, R., & German Pediatric Surveillance Unit Study, Group. 2006. Incidence and clinical presentation of invasive neonatal group B streptococcal infections in Germany. *Pediatrics*, **117**(6), e1139–45. (*cf.* p. 14)

Franz, A.R., Steinbach, G., Kron, M., & Pohlandt, F. 2007. Interleukin-8: a valuable tool to restrict antibiotic therapy in newborn infants. *Acta Paediatrica*, **90**(9), 1025–1032. (*cf.* p. 11)

Friedman, D. B. 2012. Assessing signal-to-noise in quantitative proteomics: multivariate statistical analysis in DIGE experiments. *Methods Mol Biol*, **854**, 31–45. (*cf.* p. 17)

Gadolinq. 2020. Quadrupole ion trajectory. `https://commons.wikimedia.org/wiki/File:Quadrupole_ion_trajectory.svg#/media/File:Quadrupole_ion_trajectory.svg`. Graphic licensed under Creative Commons (CC BY-SA 4.0 ↗) (Last accessed: 2022-06-24). (*cf.* p. 99)

Gallardo, M., Lee, H. J., Zhang, X., Bueso-Ramos, C., Pageon, L. R., McArthur, M., Multani, A., Nazha, A., Manshouri, T., Parker-Thornburg, J., Rapado, I., Quintas-Cardama, A., Kornblau, S. M., Martinez-Lopez, J., & Post, S. M. 2015. hnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer Cell*, **28**(4), 486–499. (*cf.* p. 69)

GBA, Gemeinsamer Bundesausschuss. 2022. Richtlinie des Gemeinsamen Bundesausschusses über die Früherkennung von Krankheiten bei Kindern (Kinder-Richtlinie) `https://www.g-ba.de/downloads/62-492-2848/Kinder-RL_2022-04-21_iK-2022-06-23.pdf`. (Last accessed: 2022-06-24). (*cf.* p. 83)

GBD, Global Burden of Disease Study 2019 under-5 Mortality Collaborators, Paulson, K. R., Kamath, A. M., Alam, T., Bienhoff, K., Abady, G. G., Abbas, J., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abd-Elsalam, S. M., Abdoli, A., Abedi, A., Abolhassani, H., Abreu, L. G., Abu-Gharbieh, E., Abu-Rmeileh, N. M. E., Abushouk, A. I., Adamu, A. L., Adebayo, O. M., Adegbosin, A. E., Adekanmbi, V., Adetokunboh, O. O., Adeyinka, D. A., Adsuar, J. C., Afshari, K., Aghaali, M., Agudelo-Botero, M., Ahinkorah, B. O., Ahmad, T., Ahmadi, K., Ahmed, M. B., Aji, B., Akalu, Y., Akinyemi, O. O., Aklilu, A., Al-Aly, Z., Alam, K., Alanezi, F. M., Alanzi, T. M., Alcalde-Rabanal, J. E., Al-Eyadhy, A., Ali, T., Alicandro, G., Alif, S. M., Alipour, V., Alizade, H., Aljunid, S. M., Almasi-Hashiani, A., Almasri, N. A., Al-Mekhlafi, H. M., Alonso, J., Al-Raddadi, R. M., Altirkawi, K. A., Alumran, A. Khalid, Alvis-Guzman, N., Alvis-Zakzuk, N. J., Ameyaw, E. K., Amini, S., Amini-Rarani, M., Amit, A. M. L., Amugsi, D. A., Ancuceanu, R., Anderlini, D., Andrei, C. L., Ansari, F., Ansari-Moghaddam, A., Antonio, C. A. T., Antriyandarti, E., Anvari, D., Anwer, R., Aqeel, M., Arabloo, J., Arab-Zozani, M., Aripov, T., Ärnlöv, J., Artanti, K. D., Arzani, A., Asaad, M., Asadi-Aliabadi, M., Asadi-Pooya, A. A., Asghari Jafarabadi, M., Athari, S. S., Athari, S. S., Atnafu, D. D., Atreya, A., Atteraya, M. S., Ausloos, M., Awan, S. T., Ayala Quintanilla, B. P., Ayano, G., Ayanore, M. A., Aynalem, Y. A., Azari, S., Azarian, G., Azene, Z. N., Darshan, B., Babaee, E., Badiye, A. D., Baig, A. A., Banach, M., *et al.* . 2021. Global, regional, and national progress towards Sustainable Development Goal 3.2 for neonatal and child health: all-cause and cause-specific mortality findings from the Global Burden of Disease Study 2019. *The Lancet*, **398**(10303), 870–905. (*cf.* p. 4)

Genzel-Boroviczény, O., & Roos, R. 2019. *Checkliste Neonatologie*. 6th revised edn. Georg Thieme Verlag (Stuttgart, New York). Chap. 11. Infektionen, pages 245–292. (*cf.* pp. 6, 10)

Geyer, P. E. 2017. *Plasma proteome profiling to assess human health and disease*. Thesis, Ludwig-Maximilians-Universität München, Fakultät für Chemie und Pharmazie. (*cf.* pp. 17, 18)

Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., & Mann, M. 2016a. Plasma Proteome Profiling to Assess Human Health and Disease. *Cell Syst*, **2**(3), 185–95. (*cf.* pp. 17, 18, 34, 82, 83)

Geyer, P. E., Wewer Albrechtsen, N. J., Tyanova, S., Grassl, N., Iepsen, E. W., Lundgren, J., Madsbad, S., Holst, J. J., Torekov, S. S., & Mann, M. 2016b. Proteomics reveals the effects of sustained weight loss on the human plasma proteome. *Mol Syst Biol*, **12**(12), 901. (*cf.* p. 17)

Geyer, P. E., Holdt, L. M., Teupser, D., & Mann, M. 2017. Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol*, **13**(9), 942. Graphic licensed under Creative Commons (CC BY-NC 4.0 ⬀). (*cf.* pp. 17, 18, 20)

Geyer, P. E., Porsdam Mann, S., Treit, P. V., & Mann, M. 2021. Plasma Proteomes Can Be Reidentifiable and Potentially Contain Personally Sensitive and Incidental Findings. *Molecular & Cellular Proteomics*, **20**. doi: 10.1074/mcp.RA120.002359. (*cf.* p. 83)

Glaesener, S., Jaenke, C., Habener, A., Geffers, R., Hagendorff, P., Witzlau, K., Imelmann, E., Krueger, A., & Meyer-Bahlburg, A. 2018. Decreased production of class-switched antibodies in neonatal B cells is associated with increased expression of miR-181b. *PLOS ONE*, **13**(2), e0192230. (*cf.* p. 9)

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., & Oliver, S. G. 1996. Life with 6000 genes. *Science*, **274**(5287), 546, 563–7. (*cf.* p. 16)

Gradel, K. O., Jensen, T. G., Kolmos, H. J., Pedersen, C., Vinholt, P. J., & Lassen, A. T. 2013. Does C-reactive protein independently predict mortality in adult community-acquired bacteremia patients with known sepsis severity? *Apmis*, **121**(9), 835–42. (*cf.* p. 11)

Gross, J. H. 2017. *Mass spectrometry - A Textbook*. 3 edn. Springer International Publishing. (*cf.* pp. 24, 25, 101)

Guilmot, A., Hermann, E., Braud, V. M., Carlier, Y., & Truyens, C. 2011. Natural Killer Cell Responses to Infections in Early Life. *Journal of Innate Immunity*, **3**(3), 280–288. (*cf.* p. 9)

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., & Aebersold, R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol*, **17**(10), 994–9. (*cf.* p. 105)

Hall, K. K., & Lyman, J. A. 2006. Updated review of blood culture contamination. *Clin Microbiol Rev*, **19**(4), 788–802. (*cf.* p. 12)

Haque, K. 2010. Neonatal Sepsis in the Very Low Birth Weight Preterm Infants: Part 2: Review of Definition, Diagnosis and Management. *Journal of Medical Sciences*, **3(1)**, 11–27. (*cf.* p. 12)

Hofer, N., Zacharias, E., Müller, W., & Resch, B. 2012. An Update on the Use of C-Reactive Protein in Early-Onset Neonatal Sepsis: Current Insights and New Tasks. *Neonatology*, **102**(1), 25–36. (*cf.* pp. 11, 12)

Hortin, G. L., Jortani, S. A., Ritchie, J. C. Jr., Valdes, R. Jr., & Chan, D. Wolfgang. 2006. Proteomics: A New Diagnostic Frontier. *Clinical Chemistry*, **52**(7), 1218–1222. (*cf.* p. 18)

Hu, H., Zhang, S., Xiong, S., Hu, B., He, Y., & Gu, Y. 2021. ACTR3 promotes cell migration and invasion by inducing epithelial mesenchymal transition in pancreatic ductal adenocarcinoma. *J Gastrointest Oncol*, **12**(5), 2325–2333. (*cf.* p. 66)

Humphrey, S. J., Azimifar, S. B., & Mann, M. 2015. High-throughput phosphoproteomics reveals in vivo insulin signaling dynamics. *Nat Biotechnol*, **33**(9), 990–5. (*cf.* p. 16)

Hurlimann, J., Thorbecke, G. J., & Hochwald, G. M. 1966. The liver as the site of C-reactive protein formation. *J Exp Med*, **123**(2), 365–78. (*cf.* p. 10)

Hurst, R. E. 2009. Does the biomarker search paradigm need re-booting? *BMC Urol*, **9**, 1. (*cf.* p. 20)

Huttlin, E. L., Bruckner, R. J., Paulo, J. A., Cannon, J. R., Ting, L., Baltier, K., Colby, G., Gebreab, F., Gygi, M. P., Parzen, H., Szpyt, J., Tam, S., Zarraga, G., Pontano-Vaites, L., Swarup, S., White, A. E., Schweppe, D. K., Rad, R., Erickson, B. K., Obar, R. A., Guruharsha, K. G., Li, K., Artavanis-Tsakonas, S., Gygi, S. P., & Harper, J. W. 2017. Architecture of the human interactome defines protein communities and disease networks. *Nature*, **545**(7655), 505–509. (*cf.* p. 16)

INIS, The INIS Study Collaborative Group. 2008. The INIS Study. International Neonatal Immunotherapy Study: non-specific intravenous immunoglobulin therapy for suspected or proven neonatal sepsis: an international, placebo controlled, multicentre randomised trial. *BMC Pregnancy and Childbirth*, **8**(1), 52. (*cf.* p. 14)

INIS, The INIS Study Collaborative Group. 2011. Treatment of Neonatal Sepsis with Intravenous Immune Globulin. *New England Journal of Medicine*, **365**(13), 1201–1211. (*cf.* p. 14)

Itzhak, D. N., Tyanova, S., Cox, J., & Borner, G. H. 2016. Global, quantitative and dynamic mapping of protein subcellular localization. *Elife*, **5**. (*cf.* p. 16)

Kaapa, P., & Koistinen, E. 1993. Maternal and neonatal C-reactive protein after interventions during delivery. *Acta Obstet Gynecol Scand*, **72**(7), 543–6. (*cf.* p. 11)

Kalenka, A., Feldmann, R. E., Otero, K., Maurer, M. H., Waschke, K. F., & Fiedler, F. 2006. Changes in the Serum Proteome of Patients with Sepsis and Septic Shock. *Anesthesia & Analgesia*, **103**(6), 1522–1526. (*cf.* p. 67)

Karas, M., Bachmann, D., & Hillenkamp, F. 1985. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry*, **57**(14), 2935–2939. (*cf.* p. 97)

Kaufman, D., Kilpatrick, L., Hudson, R. G., Campbell, D. E., Kaufman, A., Douglas, S. D., & Harris, M. C. 1999. Decreased Superoxide Production, Degranulation, Tumor Necrosis Factor Alpha Secretion, and CD11b/CD18 Receptor Expression by Adherent Monocytes from Preterm Infants. *Clinical Diagnostic Laboratory Immunology*, **6**(4), 525–529. (*cf.* p. 9)

Kawecki, C., Lenting, P. J., & Denis, C. V. 2017. von Willebrand factor and inflammation. *Journal of Thrombosis and Haemostasis*, **15**(7), 1285–1294. (*cf.* p. 70)

Kayiran, S.M., Özbek, N., Turan, M., & Gürakan, B. 2003. Significant differences between capillary and venous complete blood counts in the neonatal period. *Clinical & Laboratory Haematology*, **25**(1), 9–16. (*cf.* p. 81)

Kebarle, P., & Verkerk, U. H. 2009. Electrospray: from ions in solution to ions in the gas phase, what we know now. *Mass Spectrom Rev*, **28**(6), 898–917. (*cf.* p. 98)

Kelstrup, C. D., Bekker-Jensen, D. B., Arrey, T. N., Hogrebe, A., Harder, A., & Olsen, J. V. 2018. Performance Evaluation of the Q Exactive HF-X for Shotgun Proteomics. *J Proteome Res*, **17**(1), 727–738. (*cf.* p. 98)

Kessel, I., Leib, M., Levy, A., Miller-Lotan, R., Waisman, D., Jacobson, E., Rotschild, A., & Blum, S. 2016. Does Haptoglobin Phenotype Influence Postnatal Morbidity in Preterm Neonates? *Am J Perinatol*, **33**(02), 130–135. (*cf.* p. 67)

Kim, J., Lee, J. E., Heynen-Genel, S., Suyama, E., Ono, K., Lee, K. Y., Ideker, T., Aza-Blanc, P., & Gleeson, J. G. 2010. Functional genomic screen for modulators of ciliogenesis and cilium length. *Nature*, **464**(7291), 1048–1051. (*cf.* p. 66)

Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L. D., Patil, A. H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T., Raju, R., Kumar, M., Sreenivasamurthy, S. K., Marimuthu, A., Sathe, G. J., Chavan, S., Datta, K. K., Subbannayya, Y., Sahu, A., Yelamanchi, S. D., Jayaram, S., Rajagopalan, P., Sharma, J., Murthy, K. R., Syed, N., Goel, R., Khan, A. A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T. C., Zhong, J., Wu, X., Shaw, P. G., Freed, D., Zahari, M. S., Mukherjee, K. K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C. J., Shankar, S. K., Satishchandra, P., Schroeder, J. T., Sirdeshmukh, R., Maitra, A., Leach, S. D., Drake, C. G., Halushka, M. K., Prasad, T. S., Hruban, R. H., Kerr, C. L., Bader, G. D., Iacobuzio-Donahue, C. A., Gowda, H., & Pandey, A. 2014. A draft map of the human proteome. *Nature*, **509**(7502), 575–81. (*cf.* p. 16)

Kiser, C., Nawab, U., McKenna, K., & Aghai, Z. H. 2014. Role of guidelines on length of therapy in chorioamnionitis and neonatal sepsis. *Pediatrics*, **133**(6), 992–8. (*cf.* p. 13)

Klein Klouwenberg, P., & Bont, L. 2008. Neonatal and Infantile Immune Responses to Encapsulated Bacteria and Conjugate Vaccines. *Clinical and Developmental Immunology*, **2008**, 628963. (*cf.* p. 9)

Klinger, G., Levy, I., Sirota, L., Boyko, V., Reichman, B., Lerner-Geva, L., & Israel Neonatal, Network. 2009. Epidemiology and risk factors for early onset sepsis among very-low-birthweight infants. *Am J Obstet Gynecol*, **201**(1), 38 e1–6. (*cf.* p. 5)

Klinger, G., Levy, I., Sirota, L., Boyko, V., Lerner-Geva, L., Reichman, B., & Israel Neonatal, Network. 2010. Outcome of early-onset sepsis in a national cohort of very low birth weight infants. *Pediatrics*, **125**(4), e736–40. (*cf.* p. 15)

Kollmann, T. R., Crabtree, J., Rein-Weston, A., Blimkie, D., Thommai, F., Wang, X. Y., Lavoie, P. M., Furlong, J., Fortuno, E. S., 3rd, Hajjar, A. M., Hawkins, N. R., Self, S. G., & Wilson, C. B. 2009. Neonatal innate TLR-mediated responses are distinct from those of adults. *J Immunol*, **183**(11), 7150–60. (*cf.* p. 9)

Koren, N., Zubeidat, K., Saba, Y., Horev, Y., Barel, O., Wilharm, A., Heyman, O., Wald, S., Eli-Berchoer, L., Shapiro, H., Nadler, C., Elinav, E., Wilensky, A., Prinz, I., Bercovier, H., & Hovav, A. H. 2021. Maturation of the neonatal oral mucosa involves unique epithelium-microbiota interactions. *Cell Host Microbe*, **29**(2), 197–209 e5. (*cf.* p. 9)

Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., & Mann, M. 2014. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods*, **11**(3), 319–24. (*cf.* p. 106)

Kurisu, K., Yoshiuchi, K., Ogino, K., Okada, Y., & Oda, T. 2020. Peak C-reactive protein levels do not predict 30-day mortality for bacteremia: A retrospective cohort study. *Journal of Infection and Chemotherapy*, **26**(1), 23–27. (*cf.* p. 11)

Kuzniewicz, M. W., Mukhopadhyay, S., Li, S., Walsh, E. M., & Puopolo, K. M. 2020. Time to Positivity of Neonatal Blood Cultures for Early-onset Sepsis. *Pediatr Infect Dis J*, **39**(7), 634–640. (*cf.* p. 12)

Lamont, R. F., Sobel, J., Mazaki-Tovi, S., Kusanovic, J. P., Vaisbuch, E., Kim, S. K., Uldbjerg, N., & Romero, R. 2011. Listeriosis in human pregnancy: a systematic review. *J Perinat Med*, **39**(3), 227–36. (*cf.* p. 7)

Lapaire, O., Shennan, A., & Stepan, H. 2010. The preeclampsia biomarkers soluble fms-like tyrosine kinase-1 and placental growth factor: current knowledge, clinical implications and future application. *Eur J Obstet Gynecol Reprod Biol*, **151**(2), 122–9. (*cf.* p. 19)

Lapillonne, A., Basson, E., Monneret, G., Bienvenu, J., & Salle, B. L. 1998. Lack of specificity of procalcitonin for sepsis diagnosis in premature infants. *Lancet*, **351**(9110), 1211–2. (*cf.* p. 12)

Lauterbach, R., & Zembala, M. 1996. Pentoxifylline reduces plasma tumour necrosis factor-alpha concentration in premature infants with sepsis. *Eur J Pediatr*, **155**(5), 404–9. (*cf.* p. 14)

Lawn, J. E., Cousens, S., Zupan, J., & Lancet Neonatal Survival Steering, Team. 2005. 4 million neonatal deaths: when? Where? Why? *Lancet*, **365**(9462), 891–900. (*cf.* p. 4)

Le, J., Nguyen, T., Okamoto, M., McKamy, S., & Lieberman, J. M. 2008. Impact of empiric antibiotic use on development of infections caused by extended-spectrum beta-lactamase bacteria in a neonatal intensive care unit. *Pediatr Infect Dis J*, **27**(4), 314–8. (*cf.* p. 14)

Leal, Y. A., Alvarez-Nemegyei, J., Velazquez, J. R., Rosado-Quiab, U., Diego-Rodriguez, N., Paz-Baeza, E., & Davila-Velazquez, J. 2012. Risk factors and prognosis for neonatal sepsis in southeastern Mexico: analysis of a four-year historic cohort follow-up. *BMC Pregnancy Childbirth*, **12**, 48. (*cf.* p. 6)

Legrand, D., Elass, E., Carpentier, M., & Mazurier, J. 2005. Lactoferrin: a modulator of immune and inflammatory responses. *Cell Mol Life Sci*, **62**(22), 2549–59. (*cf.* p. 75)

Letunica, N., Cai, T., Cheong, J. L. Y., Doyle, L. W., Monagle, P., & Ignjatovic, V. 2021. The use of proteomics for blood biomarker research in premature infants: a scoping review. *Clin Proteomics*, **18**(1), 13. (*cf.* p. 21)

Levin, G., Tsur, A., Shai, D., Cahan, T., Shapira, M., & Meyer, R. 2021. Prediction of adverse neonatal outcome among newborns born through meconium-stained amniotic fluid. *Int J Gynaecol Obstet*, **154**(3), 515–520. (*cf.* p. 6)

Levine, R. J., Maynard, S. E., Qian, C., Lim, K. H., England, L. J., Yu, K. F., Schisterman, E. F., Thadhani, R., Sachs, B. P., Epstein, F. H., Sibai, B. M., Sukhatme, V. P., & Karumanchi, S. A. 2004. Circulating angiogenic factors and the risk of preeclampsia. *N Engl J Med*, **350**(7), 672–83. (*cf.* p. 19)

Levy, O. 2007. Innate immunity of the newborn: basic mechanisms and clinical correlates. *Nat Rev Immunol*, **7**(5), 379–90. (*cf.* p. 21)

Li, T., Li, X., Wei, Y., Dong, G., Yang, J., Yang, J., Fang, P., & Qi, M. 2021a. Predictive Value of C-Reactive Protein-to-Albumin Ratio for Neonatal Sepsis. *J Inflamm Res*, **14**, 3207–3215. (*cf.* p. 11)

Li, X., Li, T., Wang, J., Feng, Y., Ren, C., Xu, Z., Yang, J., Zhang, Q., & An, C. 2021b. Clinical Value of C-Reactive Protein/Platelet Ratio in Neonatal Sepsis: A Cross-Sectional Study. *J Inflamm Res*, **14**, 5123–5129. (*cf.* p. 11)

Liu, L, Oza, S., Hogan, D., Perin, J., Rudan, I., Lawn, J. E., Cousens, S., Mathers, C., & Black, R. E. 2015. Global, regional, and national causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *The Lancet*, **385**(9966), 430–440. (*cf.* p. 4)

Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B. C., & Aebersold, R. 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol*, **14**(8), e8126. (*cf.* pp. 102, 103)

Ly, T., Ahmad, Y., Shlien, A., Soroka, D., Mills, A., Emanuele, M. J., Stratton, M. R., & Lamond, A. I. 2014. A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *Elife*, **3**, e01630. (*cf.* p. 16)

MacGowan, A., Wootton, M., Bowker, K., Holt, H. A., & Reeves, D. 1998. Ampicillin-aminoglycoside interaction studies using Listeria monocytogenes. *J Antimicrob Chemother*, **41**(3), 417–8. (*cf.* p. 14)

Malmstrom, E., Kilsgard, O., Hauri, S., Smeds, E., Herwald, H., Malmstrom, L., & Malmstrom, J. 2016. Large-scale inference of protein tissue origin in gram-positive sepsis plasma using quantitative targeted proteomics. *Nat Commun*, **7**, 10261. (*cf.* p. 24)

Manroe, B. L., Rosenfeld, C. R., Weinberg, A. G., & Browne, R. 1977. The differential leukocyte count in the assessment and outcome of early-onset neonatal group B streptococcal disease. *The Journal of Pediatrics*, **91**(4), 632–637. (*cf.* p. 10)

Manzoni, P., Farina, D., Leonessa, M., d'Oulx, E. A., Galletto, P., Mostert, M., Miniero, R., & Gomirato, G. 2006. Risk factors for progression to invasive fungal infection in preterm neonates with fungal colonization. *Pediatrics*, **118**(6), 2359–64. (*cf.* p. 14)

Marchini, G., Berggren, V., Djilali-Merzoug, R., & Hansson, L. O. 2000. The birth process initiates an acute phase reaction in the fetus-newborn infant. *Acta Paediatr*, **89**(9), 1082–6. (*cf.* p. 21)

Marnell, L., Mold, C., & Du Clos, T. W. 2005. C-reactive protein: ligands, receptors and role in inflammation. *Clin Immunol*, **117**(2), 104–11. (*cf.* p. 10)

Maruna, P., Nedelnikova, K., & Gurlich, R. 2000. Physiology and genetics of procalcitonin. *Physiol Res*, **49 Suppl 1**, S57–61. (*cf.* p. 11)

McGovern, M., Giannoni, E., Kuester, H., Turner, M. A., van den Hoogen, A., Bliss, J. M., Koenig, J. M., Keij, F. M., Mazela, J., Finnegan, R., Degtyareva, M., Simons, S. H. P., de Boode, W. P., Strunk, T., Reiss, I. K. M., Wynn, J. L., & Molloy, E. J. 2020. Challenges in developing a consensus definition of neonatal sepsis. *Pediatr Res*, **88**(1), 14–26. (*cf.* p. 5)

Mishra, N. 2010. *Introduction to Proteomics: Principles and Applications*. Wiley. (*cf.* p. 15)

Mjelle, A. B., Guthe, H. J. T., Reigstad, H., Bjorke-Monsen, A. L., & Markestad, T. 2019. Serum concentrations of C-reactive protein in healthy term-born Norwegian infants 48-72 hours after birth. *Acta Paediatr*, **108**(5), 849–854. (*cf.* p. 11)

Mold, J. E., Michaelsson, J., Burt, T. D., Muench, M. O., Beckerman, K. P., Busch, M. P., Lee, T. H., Nixon, D. F., & McCune, J. M. 2008. Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero. *Science*, **322**(5907), 1562–5. (*cf.* p. 8)

Monneret, G., Labaune, J. M., Isaac, C., Bienvenu, F., Putet, G., & Bienvenu, J. 1998. Increased serum procalcitonin levels are not specific to sepsis in neonates. *Clin Infect Dis*, **27**(6), 1559–61. (*cf.* p. 12)

Moraa, P. K., Mweu, M. M., & Njoroge, P. K. 2019. Association between umbilical cord hygiene and neonatal sepsis among neonates presenting to a primary care facility in Nairobi County, Kenya: a case-control study. *F1000Res*, **8**, 920. (*cf.* p. 7)

Muk, T., Stensballe, A., Pankratova, S., Nguyen, D. N., Brunse, A., Sangild, P. T., & Jiang, P. P. 2019. Rapid Proteome Changes in Plasma and Cerebrospinal Fluid Following Bacterial Infection in Preterm Newborn Pigs. *Front Immunol*, **10**, 2651. (*cf.* pp. 23, 72)

Mukhopadhyay, S., Eichenwald, E. C., & Puopolo, K. M. 2013. Neonatal early-onset sepsis evaluations among well-appearing infants: projected impact of changes in CDC GBS guidelines. *J Perinatol*, **33**(3), 198–205. (*cf.* p. 12)

Murgas Torrazza, R., Li, N., Young, C., Kobeissy, F., Chow, M., Chen, S., Mai, V., Sharma, R., Hudak, M., Shuster, J., & Neu, J. 2013. Pilot study using proteomics to identify predictive biomarkers of necrotizing enterocolitis from buccal swabs in very low birth weight infants. *Neonatology*, **104**(3), 234–42. (*cf.* p. 22)

Murk, W., Risnes, K. R., & Bracken, M. B. 2011. Prenatal or early-life exposure to antibiotics and risk of childhood asthma: a systematic review. *Pediatrics*, **127**(6), 1125–38. (*cf.* p. 13)

Murphy, K., & Weiner, J. 2012. Use of Leukocyte Counts in Evaluation of Early-onset Neonatal Sepsis. *The Pediatric Infectious Disease Journal*, **31**(1), 16–19. (*cf.* p. 10)

Newton, R. 1988. Introduction to the theory and practice of high-performance liquid chromatography. *Monogr Endocrinol*, **30**, 1–31. (*cf.* p. 97)

Ng, P. C., Li, K., Wong, R. P., Chui, K. M., Wong, E., & Fok, T. F. 2002. Neutrophil CD64 expression: a sensitive diagnostic marker for late-onset nosocomial infection in very low birthweight infants. *Pediatr Res*, **51**(3), 296–303. (*cf.* p. 12)

Ng, P. C., Ang, I. L., Chiu, R. W., Li, K., Lam, H. S., Wong, R. P., Chui, K. M., Cheung, H. M., Ng, E. W., Fok, T. F., Sung, J. J., Lo, Y. M., & Poon, T. C. 2010. Host-response biomarkers for diagnosis of late-onset septicemia and necrotizing enterocolitis in preterm infants. *J Clin Invest*, **120**(8), 2989–3000. (*cf.* pp. 22, 23, 71)

Ng, S. C., Bernstein, C. N., Vatn, M. H., Lakatos, P. L., Loftus, E. V., Jr., Tysk, C., O'Morain, C., Moum, B., Colombel, J. F., Epidemiology, & Natural History Task Force of the International Organization of Inflammatory Bowel, Disease. 2013. Geographical variability and environmental risk factors in inflammatory bowel disease. *Gut*, **62**(4), 630–49. (*cf.* p. 13)

Nguyen, M., Leuridan, E., Zhang, T., De Wit, D., Willems, F., Van Damme, P., Goldman, M., & Goriely, S. 2010. Acquisition of Adult-Like TLR4 and TLR9 Responses during the First Year of Life. *PLOS ONE*, **5**(4), e10407. (*cf.* p. 9)

NICE. 2021. National Institute for Health and Care Excellence (Clinical Guidelines) – Neonatal infection: antibiotics for prevention and treatment. https://www.nice.org.uk/guidance/ng195/resources/neonatal-infection-antibiotics-for-prevention-and-treatment-pdf-66142083827653. (Last accessed: 2022-06-24). (*cf.* pp. 12, 13)

Nielsen, F. R., Bek, K. M., Rasmussen, P. E., Qvist, I., & Tobiassen, M. 1990. C-reactive protein during normal pregnancy. *Eur J Obstet Gynecol Reprod Biol*, **35**(1), 23–7. (*cf.* p. 11)

Nizet, V., & Klein, J.O. 2015. *Remington and Klein's Infectious Diseases of the Fetus and Newborn Infant*. 7 edn. Elsevier Health Sciences. Book Bacterial Sepsis and Meningitis, pages 222–275. (*cf.* p. 13)

Norena-Caro, D. 2017a. Liquid Chromatography Mass Spectrometer. https://en.wikipedia.org/wiki/File:Liquid_chromatography_tandem_Mass_spectrometry_diagram.png. Graphic in the public domain (CC0 1.0 ⤤) (Last accessed: 2022-06-24). (*cf.* p. 95)

Norena-Caro, D. 2017b. Liquid chromatography MS spectrum 3D analysis. https://en.wikipedia.org/wiki/File:Liquid_chromatography_MS_spectrum_3D_analysis.png. Graphic in the public domain (CC0 1.0 ⤤) (Last accessed: 2022-06-24). (*cf.* p. 104)

Nupponen, I., Andersson, S., Jarvenpaa, A. L., Kautiainen, H., & Repo, H. 2001. Neutrophil CD11b expression and circulating interleukin-8 as diagnostic markers for early-onset neonatal sepsis. *Pediatrics*, **108**(1), E12. (*cf.* p. 12)

Nussbaum, C., & Sperandio, M. 2011. Innate immune cell recruitment in the fetus and neonate. *J Reprod Immunol*, **90**(1), 74–81. (*cf.* p. 9)

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., & Mann, M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, **1**(5), 376–86. (*cf.* p. 105)

Oved, K., Cohen, A., Boico, O., Navon, R., Friedman, T., Etshtein, L., Kriger, O., Bamberger, E., Fonar, Y., Yacobov, R., Wolchinsky, R., Denkberg, G., Dotan, Y., Hochberg, A., Reiter, Y., Grupper, M., Srugo, I., Feigin, P., Gorfine, M., Chistyakov, I., Dagan, R., Klein, A., Potasman, I., & Eden, E. 2015. A novel host-proteome signature for distinguishing between acute bacterial and viral infections. *PLoS One*, **10**(3), e0120012. (*cf.* p. 23)

Pangratz-Fuehrer, S., Genzel-Boroviczény, O., Bodensohn, W., Eisenburger, R., Scharpenack, J., Geyer, P. E., Müller-Reif, J. B., van Hagen, N., Müller, A. M., Jensen, M. K., Klein, C., Mann, M., & Nussbaum, C. 2021. Cohort profile: the MUNICH Preterm and Term Clinical study (MUNICH-PreTCl), a neonatal birth cohort with focus on prenatal and postnatal determinants of infant and childhood morbidity. *BMJ Open*, **11**(6), e050652. Graphic licensed under Creative Commons (CC BY-NC 4.0 ⤤). (*cf.* pp. 31, 32)

Paris, S., Sesboüé, R., Delpech, B., Chauzy, C., Thiberville, L., Martin, J.-P., Frébourg, T., & Diarra-Mehrpour, M. 2002. Inhibition of tumor growth and metastatic spreading by overexpression of inter-alpha-trypsin inhibitor family chains. *Int J Cancer*, **97**(5), 615–20. (*cf.* p. 72)

Patel, V. J., Thalassinos, K., Slade, S. E., Connolly, J. B., Crombie, A., Murrell, J. C., & Scrivens, J. H. 2009. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res*, **8**(7), 3752–9. (*cf.* p. 105)

Patras, K. A., & Nizet, V. 2018. Group B Streptococcal Maternal Colonization and Neonatal Disease: Molecular Mechanisms and Preventative Approaches. *Front Pediatr*, **6**, 27. (*cf.* p. 6)

Paulus, P., Jennewein, C., & Zacharowski, K. 2011. Biomarkers of endothelial dysfunction: can they help us deciphering systemic inflammation and sepsis? *Biomarkers*, **16**(sup1), S11–S21. (*cf.* pp. 69, 70)

Peng, P., & Xia, Y. 2019. Influency of pentoxifylline treatment for neonatal sepsis: a meta-analysis of randomized controlled studies. *Hong Kong Journal of Emergency Medicine*, 1024907919864248. (*cf.* p. 14)

Pennisi, I., Rodriguez-Manzano, J., Moniri, A., Kaforou, M., Herberg, J. A., Levin, M., & Georgiou, P. 2021. Translation of a Host Blood RNA Signature Distinguishing Bacterial From Viral Infection Into a Platform Suitable for Development as a Point-of-Care Test. *JAMA Pediatr*, **175**(4), 417–419. (*cf.* p. 23)

Pepys, M. B., & Hirschfield, G. M. 2003. C-reactive protein: a critical update. *J Clin Invest*, **111**(12), 1805–12. (*cf.* p. 11)

Petricoin, E. F., Ardekani, A. M., Hitt, B. A., Levine, P. J., Fusaro, V. A., Steinberg, S. M., Mills, G. B., Simone, C., Fishman, D. A., Kohn, E. C., & Liotta, L. A. 2002. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, **359**(9306), 572–7. (*cf.* p. 19)

Pettengill, M. A., van Haren, S. D., & Levy, O. 2014. Soluble mediators regulating immunity in early life. *Front Immunol*, **5**, 457. (*cf.* p. 21)

Philip, A. G., & Hewitt, J. R. 1980. Early diagnosis of neonatal sepsis. *Pediatrics*, **65**(5), 1036–41. Philip, A G Hewitt, J R eng 1980/05/01 Pediatrics. 1980 May;65(5):1036-41. (*cf.* p. 67)

Philip, A. G. S. 1985. Response of C-reactive protein in neonatal Group B streptococcal infection. *The Pediatric Infectious Disease Journal*, **4**(2), 145–148. (*cf.* p. 11)

Piñeiro, M., Alava, M. A., González-Ramón, N., Osada, J., Lasierra, P., Larrad, L., Piñeiro, A., & Lampreave, F. 1999. ITIH4 serum concentration increases during acute-phase processes in human patients and is up-regulated by interleukin-6 in hepatocarcinoma HepG2 cells. *Biochem Biophys Res Commun*, **263**(1), 224–9. (*cf.* p. 72)

Pirr, S., Dauter, L., Vogl, T., Ulas, T., Bohnhorst, B., Roth, J., & Viemann, D. 2021. S100A8/A9 is the first predictive marker for neonatal sepsis. *Clinical and Translational Medicine*, **11**(4). (*cf.* p. 73)

Polcwiartek, L. B., Smith, P. B., Benjamin, D. K., Zimmerman, K., Love, A., Tiu, L., Murray, S., Kang, P., Ebbesen, F., Hagstrøm, S., Clark, R. H., & Greenberg, R. G. 2021. Early-onset sepsis in term infants admitted to neonatal intensive care units (2011–2016). *Journal of Perinatology*, **41**(1), 157–163. (*cf.* p. 7)

Porsdam Mann, S., Treit, P. V., Geyer, P. E., Omenn, G. S., & Mann, M. 2021. Ethical Principles, Constraints, and Opportunities in Clinical Proteomics. *Molecular & Cellular Proteomics*, **20**. doi: 10.1016/j.mcpro.2021.100046. (*cf.* p. 83)

PrabhuDas, M., Bonney, E., Caron, K., Dey, S., Erlebacher, A., Fazleabas, A., Fisher, S., Golos, T., Matzuk, M., McCune, J. M., Mor, G., Schulz, L., Soares, M., Spencer, T., Strominger, J., Way, S. S., & Yoshinaga, K. 2015. Immune mechanisms at the maternal-fetal interface: perspectives and challenges. *Nat Immunol*, **16**(4), 328–34. (*cf.* p. 8)

Prica, F., Radon, T., Cheng, Y., & Crnogorac-Jurcevic, T. 2016. The life and works of S100P - from conception to cancer. *American journal of cancer research*, **6**(2), 562–576. (*cf.* p. 66)

Pugh, R. N., Murray-Lyon, I. M., Dawson, J. L., Pietroni, M. C., & Williams, R. 1973. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg*, **60**(8), 646–9. (*cf.* p. 19)

Puopolo, K. M., Draper, D., Wi, S., Newman, T. B., Zupancic, J., Lieberman, E., Smith, M., & Escobar, G. J. 2011. Estimating the probability of neonatal early-onset infection on the basis of maternal risk factors. *Pediatrics*, **128**(5), e1155–63. (*cf.* p. 6)

Puopolo, K. M., Benitz, W. E., Zaoutis, T. E., Committee On, Fetus, Newborn, & Committee On Infectious, Diseases. 2018a. Management of Neonates Born at </=34 6/7 Weeks' Gestation With Suspected or Proven Early-Onset Bacterial Sepsis. *Pediatrics*, **142**(6). (*cf.* p. 13)

Puopolo, K. M., Benitz, W. E., Zaoutis, T. E., Committee On, Fetus, Newborn, & Committee On Infectious, Diseases. 2018b. Management of Neonates Born at >/=35 0/7 Weeks' Gestation With Suspected or Proven Early-Onset Bacterial Sepsis. *Pediatrics*, **142**(6). (*cf.* p. 13)

Pyati, S. P., Pildes, R. S., Jacobs, N. M., Ramamurthy, R. S., Yeh, T. F., Raval, D. S., Lilien, L. D., Amma, P., & Metzger, W. I. 1983. Penicillin in infants weighing two kilograms or less with early-onset Group B streptococcal disease. *N Engl J Med*, **308**(23), 1383–9. (*cf.* p. 7)

Qi, B., & Han, M. 2018. Microbial Siderophore Enterobactin Promotes Mitochondrial Iron Uptake and Development of the Host via Interaction with ATP Synthase. *Cell*, **175**(2), 571–582.e11. (*cf.* p. 76)

Rai, A. J., Zhang, Z., Rosenzweig, J., Shih Ie, M., Pham, T., Fung, E. T., Sokoll, L. J., & Chan, D. W. 2002. Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med*, **126**(12), 1518–26. (*cf.* p. 19)

Rayleigh, L. 1882. XX. On the equilibrium of liquid conducting masses charged with electricity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **14**(87), 184–186. (*cf.* p. 98)

Reinhart, K., Bayer, O., Brunkhorst, F., & Meisner, M. 2002. Markers of endothelial damage in organ dysfunction and sepsis. *Critical Care Medicine*, **30**(Supplement), S302–S312. (*cf.* p. 69)

Ritgen, U. 2019. *Allgemeines zur Chromatographie*. Berlin, Heidelberg: Springer Berlin Heidelberg. Pages 127–129. (*cf.* p. 96)

Ronnestad, A., Abrahamsen, T. G., Gaustad, P., & Finne, P. H. 1998. Blood culture isolates during 6 years in a tertiary neonatal intensive care unit. *Scand J Infect Dis*, **30**(3), 245–51. (*cf.* p. 7)

Ross, P. L., Huang, Y. N., Marchese, J. N., 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, D. J. 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*, **3**(12), 1154–69. (*cf.* p. 105)

Rozansky, R., & Bercovici, B. 1956. C-reactive protein during pregnancy and in cord blood. *Proc Soc Exp Biol Med*, **92**(1), 4–6. (*cf.* p. 11)

Saboohi, E., Saeed, F., Khan, R. N., & Khan, M. A. 2019. Immature to total neutrophil ratio as an early indicator of early neonatal sepsis. *Pak J Med Sci*, **35**(1), 241–246. (*cf.* p. 10)

Sachse, C., Dressler, F., & Henkel, E. 1998. Increased serum procalcitonin in newborn infants without infection. *Clin Chem*, **44**(6 Pt 1), 1343–4. (*cf.* p. 12)

Salama, S. A., Gouwy, M., van Damme, J., & Struyf, S. 2021. The turning away of serum amyloid A biological activities and receptor usage. *Immunology*, **163**(2), 115–127. (*cf.* p. 72)

Salzberg, S. L. 2018. Open questions: How many genes do we have? *BMC Biol*, **16**(1), 94. (*cf.* p. 16)

Sanchez, L, Calvo, M, & Brock, J H. 1992. Biological role of lactoferrin. *Archives of Disease in Childhood*, **67**(5), 657–661. (*cf.* p. 75)

Savaryn, J. P., Toby, T. K., & Kelleher, N. L. 2016. A researcher's guide to mass spectrometry-based proteomics. *Proteomics*, **16**(18), 2435–43. (*cf.* pp. 25, 99)

Scheltema, R. A., & Mann, M. 2012. SprayQc: A Real-Time LC–MS/MS Quality Monitoring System To Maximize Uptime Using Off the Shelf Components. *Journal of Proteome Research*, **11**(6), 3458–3466. PMID: 22515319. (*cf.* p. 106)

Schibler, K. R., Osborne, K. A., Leung, L. Y., Le, T. V., Baker, S. I., & Thompson, D. D. 1998. A Randomized, Placebo-Controlled Trial of Granulocyte Colony-stimulating Factor Administration to Newborn Infants With Neutropenia and Clinical Signs of Early-onset Sepsis. *Pediatrics*, **102**(1), 6–13. (*cf.* p. 14)

Schneider, H., Faschingbauer, F., Schuepbach-Mallepell, S., Körber, I., Wohlfart, S., Dick, A., Wahlbuhl, M., Kowalczyk-Quintas, C., Vigolo, M., Kirby, N., Tannert, C., Rompel, O., Rascher, W., Beckmann, M. W., & Schneider, P. 2018. Prenatal Correction of X-Linked Hypohidrotic Ectodermal Dysplasia. *New England Journal of Medicine*, **378**(17), 1604–1610. PMID: 29694819. (*cf.* p. 83)

Schrier, L., Hadjipanayis, A., Stiris, T., Ross-Russell, R. I., Valiulis, A., Turner, M. A., Zhao, W., De Cock, P., de Wildt, S. N., Allegaert, K., & van den Anker, J. 2020. Off-label use of medicines in neonates, infants, children, and adolescents: a joint policy statement by the European Academy of Paediatrics and the European society for Developmental Perinatal and Pediatric Pharmacology. *Eur J Pediatr*, **179**(5), 839–847. (*cf.* p. 20)

Schrödl, W., Büchler, R., Wendler, S., Reinhold, P., Muckova, P., Reindl, J., & Rhode, H. 2016. Acute phase proteins as promising biomarkers: Perspectives and limitations for human and veterinary medicine. *PROTEOMICS - Clinical Applications*, **10**(11), 1077–1092. (*cf.* p. 75)

Schulfer, A., & Blaser, M. J. 2015. Risks of Antibiotic Exposures Early in Life on the Developing Microbiome. *PLoS Pathog*, **11**(7), e1004903. (*cf.* p. 13)

Self, W. H., Rosen, J., Sharp, S. C., Filbin, M. R., Hou, P. C., Parekh, A. D., Kurz, M. C., & Shapiro, N. I. 2017. Diagnostic Accuracy of FebriDx: A Rapid Test to Detect Immune Responses to Viral and Bacterial Upper Respiratory Infections. *J Clin Med*, **6**(10). (*cf.* p. 23)

Shah, J., Jefferies, A. L., Yoon, E. W., Lee, S. K., Shah, P. S., & Canadian Neonatal, Network. 2015. Risk Factors and Outcomes of Late-Onset Bacterial Sepsis in Preterm Neonates Born at ¡ 32 Weeks' Gestation. *Am J Perinatol*, **32**(7), 675–82. (*cf.* p. 22)

Shannon, E., Noveck, R., Sandoval, F., Kamath, B., & Kearney, M. 2007. Thalidomide suppressed interleukin-6 but not tumor necrosis factor-alpha in volunteers with experimental endotoxemia. *Transl Res*, **150**(5), 275–80. (*cf.* p. 11)

Shapiro, N. I., Self, W. H., Rosen, J., Sharp, S. C., Filbin, M. R., Hou, P. C., Parekh, A. D., Kurz, M. C., & Sambursky, R. 2018. A prospective, multi-centre US clinical trial to determine accuracy of FebriDx point-of-care testing for acute upper respiratory infections with and without a confirmed fever. *Ann Med*, **50**(5), 420–429. (*cf.* p. 23)

Sharma, K., Schmitt, S., Bergner, C. G., Tyanova, S., Kannaiyan, N., Manrique-Hoyos, N., Kongi, K., Cantuti, L., Hanisch, U. K., Philips, M. A., Rossner, M. J., Mann, M., & Simons, M. 2015. Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci*, **18**(12), 1819–31. (*cf.* p. 16)

Sharma, R., & Hudak, M. L. 2013. A clinical perspective of necrotizing enterocolitis: past, present, and future. *Clin Perinatol*, **40**(1), 27–51. (*cf.* p. 22)

Sheler, E. T. 2007. Phenylketonuria testing. https://commons.wikimedia.org/wiki/File:Phenylketonuria_testing.jpg. Graphic in the public domain (CC0 1.0 ↗) (Last accessed: 2022-06-24). (*cf.* p. 34)

Shim, G. H., Kim, S. D., Kim, H. S., Kim, E. S., Lee, H. J., Lee, J. A., Choi, C. W., Kim, E. K., Choi, E. H., Kim, B. I., Lee, H. J., & Choi, J. H. 2011. Trends in epidemiology of neonatal sepsis in a tertiary center in Korea: a 26-year longitudinal analysis, 1980-2005. *J Korean Med Sci*, **26**(2), 284–9. (*cf.* p. 7)

Shirley, M. 2019. FebriDx(®): A Rapid Diagnostic Test for Differentiating Bacterial and Viral Aetiologies in Acute Respiratory Infections. *Mol Diagn Ther*, **23**(6), 803–809. (*cf.* p. 23)

Sigel, S., Bunk, S., Meergans, T., Doninger, B., Stich, K., Stulnig, T., Derfler, K., Hoffmann, J., Deininger, S., von Aulock, S., & Knapp, S. 2012. Apolipoprotein B100 is a suppressor of Staphylococcus aureus-induced innate immune responses in humans and mice. *European Journal of Immunology*, **42**(11), 2983–2989. (*cf.* p. 71)

Simonsen, K. A., Anderson-Berry, A. L., Delair, S. F., & Davies, H. D. 2014. Early-onset neonatal sepsis. *Clin Microbiol Rev*, **27**(1), 21–47. (*cf.* pp. 6, 7, 10, 81)

Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G. R., Chiche, J. D., Coopersmith, C. M., Hotchkiss, R. S., Levy, M. M., Marshall, J. C., Martin, G. S., Opal, S. M., Rubenfeld, G. D., van der Poll, T., Vincent, J. L., & Angus, D. C. 2016. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*, **315**(8), 801–10. (*cf.* pp. 5, 24)

Sinha, A., & Mann, M. 2020. A beginner's guide to mass spectrometry–based proteomics. *The Biochemist*, **42**(5), 64–69. (*cf.* pp. 103, 104)

Siriwachirachai, T., Sangkomkamhang, U. S., Lumbiganon, P., & Laopaiboon, M. 2014. Antibiotics for meconium-stained amniotic fluid in labour for preventing maternal and neonatal infections. *Cochrane Database Syst Rev*, CD007772. (*cf.* p. 6)

Sproston, N. R., & Ashworth, J. J. 2018. Role of C-Reactive Protein at Sites of Inflammation and Infection. *Front Immunol*, **9**, 754. (*cf.* p. 10)

Srugo, I., Klein, A., Stein, M., Golan-Shany, O., Kerem, N., Chistyakov, I., Genizi, J., Glazer, O., Yaniv, L., German, A., Miron, D., Shachor-Meyouhas, Y., Bamberger, E., Oved, K., Gottlieb, T. M., Navon, R., Paz, M., Etshtein, L., Boico, O., Kronenfeld, G., Eden, E., Cohen, R., Chappuy, H., Angoulvant, F., Lacroix, L., & Gervaix, A. 2017. Validation of a Novel Assay to Distinguish Bacterial and Viral Infections. *Pediatrics*, **140**(4). (*cf.* p. 23)

Stewart, C. J., Nelson, A., Treumann, A., Skeath, T., Cummings, S. P., Embleton, N. D., & Berrington, J. E. 2016. Metabolomic and proteomic analysis of serum from preterm infants with necrotising enterocolitis and late-onset sepsis. *Pediatr Res*, **79**(3), 425–31. (*cf.* pp. 22, 80)

Stocker, M., van Herk, W., El Helou, S., Dutta, S., Schuerman, Faba, van den Tooren-de Groot, R. K., Wieringa, J. W., Janota, J., van der Meer-Kappelle, L. H., Moonen, R., Sie, S. D., de Vries, E., Donker, A. E., Zimmerman, U., Schlapbach, L. J., de Mol, A. C., Hoffman-Haringsma, A., Roy, M., Tomaske, M., R, F. Kornelisse, van Gijsel, J., Visser, E. G., Plotz, F. B., Heath, P., Achten, N. B., Lehnick, D., & van Rossum, A. M. C. 2021. C-Reactive Protein, Procalcitonin, and White Blood Count to Rule Out Neonatal Early-onset Sepsis Within 36 Hours: A Secondary Analysis of the Neonatal Procalcitonin Intervention Study. *Clin Infect Dis*, **73**(2), e383–e390. (*cf.* p. 13)

Storm, W. 1980. Transient bacteremia following endotracheal suctioning in ventilated newborns. *Pediatrics*, **65**(3), 487–90. (*cf.* p. 7)

Strunk, T., Temming, P., Gembruch, U., Reiss, I., Bucsky, P., & Schultz, C. 2004. Differential Maturation of the Innate Immune Response in Human Fetuses. *Pediatric Research*, **56**(2), 219–226. (*cf.* p. 9)

Suski, M., Bokiniec, R., Szwarc-Duma, M., Madej, J., Bujak-Gizycka, B., Borszewska-Kornacka, M. K., Ksiazek, T., Grabowska, A., Revhaug, C., Baumbusch, L. O., Saugstad, O. D., Pietrzyk, J. J., & Kwinta, P. 2018a. Plasma proteome changes in cord blood samples from preterm infants. *J Perinatol*. (*cf.* pp. 21, 80)

Suski, M., Bokiniec, R., Szwarc-Duma, M., Madej, J., Bujak-Gizycka, B., Kwinta, P., Borszewska-Kornacka, M. K., Revhaug, C., Baumbusch, L. O., Saugstad, O. D., & Pietrzyk, J. J. 2018b. Prospective plasma proteome changes in preterm infants with different gestational ages. *Pediatr Res*. (*cf.* pp. 21, 80)

Sylvester, K. G., Ling, X. B., Liu, G. Y., Kastenberg, Z. J., Ji, J., Hu, Z., Peng, S., Lau, K., Abdullah, F., Brandt, M. L., Ehrenkranz, R. A., Harris, M. C., Lee, T. C., Simpson, J., Bowers, C., & Moss, R. L. 2014. A novel urine peptide biomarker-based algorithm for the prognosis of necrotising enterocolitis in human infants. *Gut*, **63**(8), 1284–92. (*cf.* p. 22)

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., & Matsuo, T. 1988. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, **2**(8), 151–153. (*cf.* p. 97)

Taskin, A., Can, E., & Hamilcikan, S. 2020. Suspected or Proven Early-Onset Sepsis and NLR, PLR, and MPV Parameters in Neonates with Born through MSAF. *Am J Perinatol*. (*cf.* p. 6)

Taylor, G. I., & McEwan, A. D. 1965. The stability of a horizontal fluid interface in a vertical electric field. *Journal of Fluid Mechanics*, **22**(1), 1–15. (*cf.* p. 97)

Thermo Fisher Scientific. 2012. Orbitrap mass analyzer and injector. https://commons.wikimedia.org/wiki/File:OrbitrapMA%26Injector.png. Graphic licensed under Creative Commons (CC BY-SA 3.0 ⌕) (Last accessed: 2022-06-24). (*cf.* p. 100)

Thome, J. J., Bickham, K. L., Ohmura, Y., Kubota, M., Matsuoka, N., Gordon, C., Granot, T., Griesemer, A., Lerner, H., Kato, T., & Farber, D. L. 2016. Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues. *Nat Med*, **22**(1), 72–7. (*cf.* p. 8)

Tillett, W. S., & Francis, T. 1930. Serological Reactions in Pneumonia with a Non-Protein Somatic Fraction of Pneumococcus. *J Exp Med*, **52**(4), 561–71. (*cf.* p. 10)

Tosson, A. M. S., Glaser, K., Weinhage, T., Foell, D., Aboualam, M. S., Edris, A. A., Ansary, M. El, Lotfy, S., & Speer, C. P. 2019. Evaluation of the S100 protein A12 as a biomarker of neonatal sepsis. *The Journal of Maternal-Fetal & Neonatal Medicine*, **33**(16), 2768–2774. (*cf.* p. 66)

Tóthová, V., & Gibadulinová, A. 2013. S100P, a peculiar member of S100 family of calcium-binding proteins implicated in cancer. *Acta virologica*, **57**(02), 238–246. (*cf.* p. 66)

Trinder, M., Boyd, J. H., & Brunham, L. R. 2019. Molecular regulation of plasma lipid levels during systemic inflammation and sepsis. *Current Opinion in Lipidology*, **30**(2), 108–116. (*cf.* p. 71)

Tsafaras, G. P., Ntontsi, P., & Xanthou, G. 2020. Advantages and Limitations of the Neonatal Immune System. *Front Pediatr*, **8**, 5. (*cf.* p. 8)

Tsai, M. H., Hsu, J. F., Chu, S. M., Lien, R., Huang, H. R., Chiang, M. C., Fu, R. H., Lee, C. W., & Huang, Y. C. 2014. Incidence, clinical characteristics and risk factors for adverse outcome in neonates with late-onset sepsis. *Pediatr Infect Dis J*, **33**(1), e7–e13. (*cf.* p. 6)

Tusher, V. G., Tibshirani, R., & Chu, G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*, **98**(9), 5116–21. (*cf.* pp. 26, 27)

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, **13**(9), 731–740. (*cf.* p. 35)

Uhlen, M., & Ponten, F. 2005. Antibody-based proteomics for human tissue profiling. *Mol Cell Proteomics*, **4**(4), 384–93. (*cf.* p. 16)

UN IGME, United Nations Inter-agency Group for Child Mortality Estimation. 2021. https://childmortality.org. (Last accessed: 2022-06-24). (*cf.* p. 4)

UNICEF, United Nations International Children's Emergency Fund. 2021. https://data.unicef.org/topic/child-survival/neonatal-mortality/. (Last accessed: 2022-06-24). (*cf.* p. 4)

van der Hoeven, A., van der Beek, M. T., Lopriore, E., Steggerda, S. J., & Bekker, V. 2022. Predicting Neonatal Early Onset Sepsis: A 14-Year Cohort Study. *Pediatr Infect Dis J*, **41**(1), 72–77. (*cf.* p. 13)

van der Poll, T., van de Veerdonk, F. L., Scicluna, B. P., & Netea, M. G. 2017. The immunopathology of sepsis and potential therapeutic targets. *Nature Reviews Immunology*, **17**(7), 407–420. (*cf.* p. 71)

van Houten, C. B., de Groot, J. A. H., Klein, A., Srugo, I., Chistyakov, I., de Waal, W., Meijssen, C. B., Avis, W., Wolfs, T. F. W., Shachor-Meyouhas, Y., Stein, M., Sanders, E. A. M., & Bont, L. J. 2017. A host-protein based assay to differentiate between bacterial and viral infections in preschool children (OPPORTUNITY): a double-blind, multicentre, validation study. *The Lancet Infectious Diseases*, **17**(4), 431–440. (*cf.* p. 23)

Vasiljević, B., Antonović, O., Maglajlić-Djukić, S., & Gojnić, M. 2008. [The serum level of C-reactive protein in neonatal sepsis]. *Srp Arh Celok Lek*, **136**(5-6), 253–7. (*cf.* p. 11)

Vazzalwar, R., Pina-Rodrigues, E., Puppala, B. L., Angst, D. B., & Schweig, L. 2005. Procalcitonin as a screening test for late-onset sepsis in preterm very low birth weight infants. *J Perinatol*, **25**(6), 397–402. (*cf.* p. 12)

Vergnano, S., Menson, E., Kennea, N., Embleton, N., Russell, A. B., Watts, T., Robinson, M. J., Collinson, A., & Heath, P. T. 2011. Neonatal infections in England: the NeonIN surveillance network. *Arch Dis Child Fetal Neonatal Ed*, **96**(1), F9–F14. (*cf.* p. 6)

Vincent, J. L., Moreno, R., Takala, J., Willatts, S., De Mendonca, A., Bruining, H., Reinhart, C. K., Suter, P. M., & Thijs, L. G. 1996. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med*, **22**(7), 707–10. (*cf.* p. 19)

Vogl, T., Tenbrock, K., Ludwig, S., Leukert, N., Ehrhardt, C., van Zoelen, M. A. D., Nacken, W., Foell, D., van der Poll, T., Sorg, C., & Roth, J. 2007. Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nature Medicine*, **13**(9), 1042–1049. (*cf.* p. 73)

Vouloumanou, E. K., Plessa, E., Karageorgopoulos, D. E., Mantadakis, E., & Falagas, M. E. 2011. Serum procalcitonin as a diagnostic marker for neonatal sepsis: a systematic review and meta-analysis. *Intensive Care Med*, **37**(5), 747–62. (*cf.* p. 11)

Walker, J. C., Smolders, M. A., Gemen, E. F., Antonius, T. A., Leuvenink, J., & de Vries, E. 2011. Development of lymphocyte subpopulations in preterm infants. *Scand J Immunol*, **73**(1), 53–8. (*cf.* p. 9)

Wang, D., Eraslan, B., Wieland, T., Hallstrom, B., Hopf, T., Zolg, D. P., Zecha, J., Asplund, A., Li, L. H., Meng, C., Frejno, M., Schmidt, T., Schnatbaum, K., Wilhelm, M., Ponten, F., Uhlen, M., Gagneur, J., Hahne, H., & Kuster, B. 2019. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol Syst Biol*, **15**(2), e8503. (*cf.* p. 16)

Wang, Z., Qiu, H., He, J., Liu, L., Xue, W., Fox, A., Tickner, J., & Xu, J. 2020a. The emerging roles of hnRNPK. *J Cell Physiol*, **235**(3), 1995–2008. (*cf.* p. 68)

Wang, Z., Kavdia, K., Dey, K. K., Pagala, V. R., Kodali, K., Liu, D., Lee, D. G., Sun, H., Chepyala, S. R., Cho, J. H., Niu, M., High, A. A., & Peng, J. 2020b. High-throughput and Deep-proteome Profiling by 16-plex Tandem Mass Tag Labeling Coupled with Two-dimensional Chromatography and Mass Spectrometry. *J Vis Exp*. (*cf.* p. 16)

Wassenaar, T. M., & Panigrahi, P. 2014. Is a foetus developing in a sterile environment? *Lett Appl Microbiol*, **59**(6), 572–9. (*cf.* p. 8)

Weinberger, B., Laskin, D. L., Mariano, T. M., Sunil, V. R., DeCoste, C. J., Heck, D. E., Gardner, C. R., & Laskin, J. D. 2001. Mechanisms underlying reduced responsiveness of neonatal neutrophils to distinct chemoattractants. *J Leukoc Biol*, **70**(6), 969–76. (*cf.* p. 9)

Weinhold, B., & Ruther, U. 1997. Interleukin-6-dependent and -independent regulation of the human C-reactive protein gene. *Biochem J*, **327 (Pt 2)**, 425–9. (*cf.* p. 11)

Weiss, S. L., Fitzgerald, J. C., Balamuth, F., Alpern, E. R., Lavelle, J., Chilutti, M., Grundmeier, R., Nadkarni, V. M., & Thomas, N. J. 2014. Delayed antimicrobial therapy increases mortality and organ dysfunction duration in pediatric sepsis. *Crit Care Med*, **42**(11), 2409–17. (*cf.* p. 12)

Weitzdoerfer, R., Fountoulakis, M., & Lubec, G. 2002. Reduction of actin-related protein complex 2/3 in fetal Down syndrome brain. *Biochemical and Biophysical Research Communications*, **293**(2), 836–841. (*cf.* p. 66)

Wilhelm, M., Schlegl, J., Hahne, H., Gholami, A. M., Lieberenz, M., Savitski, M. M., Ziegler, E., Butzmann, L., Gessulat, S., Marx, H., Mathieson, T., Lemeer, S., Schnatbaum, K., Reimer, U., Wenschuh, H., Mollenhauer, M., Slotta-Huspenina, J., Boese, J. H., Bantscheff, M., Gerstmair, A., Faerber, F., & Kuster, B. 2014. Mass-spectrometry-based draft of the human proteome. *Nature*, **509**(7502), 582–7. (*cf.* p. 16)

Wilkins, I., Mezrow, G., Lynch, L., Bottone, E. J., & Berkowitz, R. L. 1989. Amnionitis and life-threatening respiratory distress after percutaneous umbilical blood sampling. *Am J Obstet Gynecol*, **160**(2), 427–8. (*cf.* p. 6)

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., & Williams, K. L. 1996. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev*, **13**, 19–50. (*cf.* p. 15)

Woodford, E. C., Dhudasia, M. B., Puopolo, K. M., Skerritt, L. A., Bhavsar, M., DeLuca, J., & Mukhopadhyay, S. 2021. Neonatal blood culture inoculant volume: feasibility and challenges. *Pediatr Res*, **90**(5), 1086–1092. (*cf.* p. 12)

Wu, L., & Han, D. K. 2006. Overcoming the dynamic range problem in mass spectrometry-based shotgun proteomics. *Expert Rev Proteomics*, **3**(6), 611–9. (*cf.* p. 18)

Wynn, J. L. 2016. Defining neonatal sepsis. *Curr Opin Pediatr*, **28**(2), 135–40. (*cf.* p. 5)

Xiao, X., Yeoh, B. S., & Vijay-Kumar, M. 2017. Lipocalin 2: An Emerging Player in Iron Homeostasis and Inflammation. *Annual Review of Nutrition*, **37**(1), 103–130. (*cf.* p. 76)

Xu, Y., Wu, W., Han, Q., Wang, Y., Li, C., Zhang, P., & Xu, H. 2019. Post-translational modification control of RNA-binding protein hnRNPK function. *Open Biology*, **9**(3), 180239. (*cf.* p. 68)

Yerkovich, S. T., Wikström, M. E., Suriyaarachchi, D., Prescott, S. L., Upham, J. W., & Holt, P. G. 2007. Postnatal Development of Monocyte Cytokine Responses to Bacterial Lipopolysaccharide. *Pediatric Research*, **62**(5), 547–552. (*cf.* p. 9)

Ygberg, S., & Nilsson, A. 2012. The developing immune system - from foetus to toddler. *Acta Paediatr*, **101**(2), 120–7. (*cf.* p. 8)

Yin, X., Zhang, Y., Guo, S., Jin, H., Wang, W., & Yang, P. 2015. Large scale systematic proteomic quantification from non-metastatic to metastatic colorectal cancer. *Scientific Reports*, **5**(1). (*cf.* p. 66)

Zaccai, G. 2003. Proteins as nano-machines: dynamics function relations studied by neutron scattering. *Journal of Physics: Condensed Matter*, **15**(18), S1673–S1682. (*cf.* p. 15)

Zandstra, J., Jongerius, I., & Kuijpers, T. W. 2021. Future Biomarkers for Infection and Inflammation in Febrile Children. *Front Immunol*, **12**, 631308. (*cf.* p. 23)

Zemlin, M., Berger, A., Franz, A., Gille, C., Härtel, C., Küster, H., Müller, A., Pohlandt, F., Simon, A., & Merz, W. 2021. Bakterielle Infektionen bei Neugeborenen – Leitlinie der Gesellschaft für Neonatologie und Pädiatrische Intensivmedizin (GNPI), der Deutschen Gesellschaft für Pädiatrische Infektiologie (DGPI), der Deutschen Gesellschaft für Kinder- und Jugendmedizin (DGKJ) und der Deutschen Gesellschaft für Gynäkologie und Geburtshilfe (DGGG). **024-008**. (*cf.* pp. 5, 6, 9, 10, 12, 13)

Zhang, H., Liu, L., Jiang, C., Pan, K., Deng, J., & Wan, C. 2019. MMP9 protects against LPS-induced inflammation in osteoblasts. *Innate Immunity*, **26**(4), 259–269. (*cf.* p. 68)

Zhang, X., Zhivaki, D., & Lo-Man, R. 2017. Unique aspects of the perinatal immune system. *Nat Rev Immunol*, **17**(8), 495–507. (*cf.* p. 8)

Zhang, Z., Bast, R. C., Jr., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., Rosenzweig, J. M., Cameron, B., Wang, Y. Y., Meng, X. Y., Berchuck, A., Van Haaften-Day, C., Hacker, N. F., de Bruijn, H. W., van der Zee, A. G., Jacobs, I. J., Fung, E. T., & Chan, D. W. 2004. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res*, **64**(16), 5882–90. (*cf.* p. 19)

Zhu, J., Zhang, H., Guo, T., Li, W., Li, H., Zhu, Y., & Huang, S. 2014. Quantitative proteomics reveals differential biological processes in healthy neonatal cord neutrophils and adult neutrophils. *Proteomics*, **14**(13-14), 1688–1697. (*cf.* p. 9)

Zhuo, L., Hascall, V. C., & Kimata, K. 2004. Inter-alpha-trypsin inhibitor, a covalent protein-glycosaminoglycan-protein complex. *J Biol Chem*, **279**(37), 38079–82. (*cf.* p. 72)

Zubarev, R. A. 2013. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics*, **13**(5), 723–6. (*cf.* pp. 18, 35)

Zubarev, R.A., & Makarov, A. 2013. Orbitrap Mass Spectrometry. *Analytical Chemistry*, **85**(11), 5288–5296. (*cf.* p. 100)

Danksagung

Vor allen anderen möchte ich hier denjenigen danken, ohne die kein Buchstabe dieser Arbeit überhaupt denkbar gewesen wäre: den Eltern aller Neugeborenen, die uns ihr Vertrauen geschenkt und einer Studienteilnahme zugestimmt haben. In sehr besonderen, für manche Eltern aber auch in schweren Stunden ihres Lebens haben wir viel Geduld und Unterstützung erfahren, oft mehr als man hätte erwarten können. Ich hoffe sehr, dass es allen Kindern seitdem gut ergangen ist, und wünsche jedem einzelnen von ihnen – wie groß die Herausforderungen auch immer sein mögen – ein erfülltes Leben.

Darüberhinaus möchte ich von ganzem Herzen danken ...

... Prof. Dr. Genzel-Boroviczény – für ihr stetes Vertrauen in unser Projekt, die rechten Weichenstellungen zur rechten Zeit und alle wegweisenden, persönlich bereichernden Gespräche.

... Dr. Susanne Pangratz-Fuehrer – für ihre selbstlose Unterstützung und ihre ansteckende wissenschaftliche Neugier, ohne die so vieles unerreichbar geblieben wäre.

... Robin Eisenburger & Janne Scharpenack – für das gemeinsam Erreichte.

... Dr. Johannes Müller-Reif – für seine brillianten Erklärungen.

... Dr. Philipp Geyer – für die wertschätzende Zusammenarbeit.

... PD Dr. Claudia Nußbaum – für alle messerscharfen Denkanstöße.

... allen Kinderärzten, Kinderkrankenschwestern und Hebammen in der damaligen Frauenklinik Maistraße – für jede Hilfe bei der Durchführung unserer Studie.

... meinen Brüdern – für ihre Unterstützung auf den letzten Metern dieser Arbeit.

... meinen Eltern – für alles, insbesondere aber dafür, dass sie mir den Weg in den Arztberuf gezeigt und mich auf diesem immer unterstützt haben.

Mein größter Dank aber gehört meiner Frau und meinem Sohn – für unsere Einheit und unser gemeinsames Glück, das sich weder in Zahlen noch Worte jemals fassen ließe.

AFFIDAVIT

Eidesstattliche Versicherung

Ich, WOLFGANG BODENSOHN, erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel

INFLUENCES OF
EARLY-ONSET INFECTIONS
ON THE NEONATAL PROTEOME

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 28. August 2023

Wolfgang Bodensohn

—————————————————————
Ort, Datum

—————————————————————
Doktorand