Johan Löfgren

GENETIC POLYMORPHISMS IN COLLECTINS AND TOLL-LIKE RECEPTOR 4 AS FACTORS INFLUENCING SUSCEPTIBILITY TO SEVERE RSV INFECTIONS AND OTITIS MEDIA

FACULTY OF MEDICINE, INSTITUTE OF CLINICAL MEDICINE, DEPARTMENT OF PEDIATRICS, UNIVERSITY OF OULU; BIOCENTER OULU; UNIVERSITY OF OULU
JOHAN LÖFGREN

GENETIC POLYMORPHISMS IN COLLECTINS AND TOLL-LIKE RECEPTOR 4 AS FACTORS INFLUENCING SUSCEPTIBILITY TO SEVERE RSV INFECTIONS AND OTITIS MEDIA

Academic dissertation to be presented with the assent of the Faculty of Medicine of the University of Oulu for public defence in Auditorium 12 of Oulu University Hospital, on 3 April 2009, at 12 noon
Löfgren, Johan, Genetic polymorphisms in collectins and Toll-like receptor 4 as factors influencing susceptibility to severe RSV infections and otitis media
Faculty of Medicine, Institute of Clinical Medicine, Department of Paediatrics, University of Oulu, P.O.Box 5000, FI-90014 University of Oulu, Finland; Biocenter Oulu, University of Oulu, P.O. Box 5000, FI-90014 University of Oulu, Finland
Acta Univ. Oul. D 1010, 2009
Oulu, Finland

Abstract
Respiratory syncytial virus (RSV) is one of the most common pathogens for early childhood respiratory tract infections. Most children are infected by RSV, approximately 1% of infants require hospital care. RSV infections are often accompanied by otitis media (OM). Surfactant proteins A and D (SP-A and SP-D) are involved in lung function and innate immunity. The proteins are capable of recognizing surface patterns in pathogens and function as a defense before the acquired immunity is developed. The Toll-like receptor (TLR) family is associated to several pathogens. The role of the TLR-family is to recognize pathogens and activate the immune defense.

The aim of the thesis was to investigate the genetic association of RSV and OM infections in infants. A candidate gene approach was used study SP-A, SP-D and TLR4 in severe RSV bronchiolitis. Association between SP-A and OM was also studied. A case-control study setup was used for all studies. 1700 samples were collected for the studies.

The results revealed genetic association between SP-A gene variation and severe RSV infections. SP-A allele 1A3 was overrepresented in RSV infants, the allele 1A was present more often in the control population. SP-D allele Met11 and genotype Met/Met were predisposing to severe RSV infections. The TLR4 gene did not show direct association with severe RSV. However, we showed for the first time difference in association in two separate epidemics. In the OM study, an association was shown between SP-A gene variations and otitis media in children.

The present results have brought new information about innate defense and the genetic variations and associations involved. The results will help understand the mechanisms of innate defense and predisposition to infections. The results also present possibilities to investigate and develop new treatment strategies.

Keywords: bronchiolitis, genetic polymorphism, innate immunity, otitis media, respiratory infection, RS-virus, surfactant protein
To my family
Acknowledgements

First, I wish to thank Professor Mikko Hallman, MD for the opportunity to work in his laboratory and for sharing his extensive knowledge in this and other fields of science. I also wish to thank him for his endless patience, support, professional guidance, discussions and the personal support received when times were thorny. I can only imagine how difficult it must be to transform a medical student with no knowledge to something that resembles a research scientist.

My deepest gratitude goes to Ville Peltola and Harri Saxén for accepting to be the reviewers for my thesis work. Thank you for your expert comments and suggestions regarding my work. I also want to thank docent Mika Rämet MD, he guided me during my first years in the secrets of genetics and the fine arts of genotyping. He provided me with the basic tools upon which I could build my own knowledge and my own experiments. His devotion to medical science has been a tremendous inspiration to do my own research. A most kind thanks for Meri Rova, PhD, we have had very successful collaboration with surfactant protein research and wish her a successful career in science. Riitta Martila, MD and Marjo Renko, MD thank you for helping out with sample collections at the Seinäjoki and Oulu pediatric wards during these years. With your help we were able to collect an impressive amount of samples for the studies. You have also been good company to discuss matters regarding research, medicine and family. Ritva Haataja, thank you for the numerous inputs in genetics and especially statistics. Your extensive knowledge in this field has been an invaluable help.

I also wish to thank our entire laboratory for their support and for creating the “laboratory spirit”. So a big thanks in no particular order to Sanna Eilola, Outi Seppänen, Elsi Jokelanen, Virpi Glumoff, Reetta Vuolteenaho, Marja Ojaniemi, Reija Paananen, Mari Liljeroos, Sirpa Kangas, Johanna Ulvila, Mataleena Parikka, Saija Taponen, Heidi Enwald, Minna Karjalainen, Juhani Metsola, Annamari Salminen. A special thanks to Dr. Samuli Rounioja, with who it was always been a pleasure to discuss science, aviation and medical matters. Special thanks to our two excellent laboratory technicians Maarit Haarala and Mirkka Ovaska for their assistance. Their help with the endless number of genotyping and sample processing’s has been priceless.

This entire thesis would not have been possible without all the volunteer patients. All the small infants and their parents that participated in the studies receive my great gratitude and I thank them for their supportive outlook to medical research.
I wish to thank my parents Tuija and Peter for providing me with a genetic heritage and upbringing that seems to spring an unlimited strive for knowledge, passion for learning and caring of patients and family in me. Also thanks for the support during these years, even though we have been thousands of miles apart on the other sides of this planet for most of the time. To my sister Kristina a kind thanks for support and cheering email from time to time, keeping the spirits high on those dark Monday mornings and everlasting Friday afternoons.

Finally, to my dear wife Eeva and our two endless sources of inspiration Anna and Anton, you have been the most understanding and loving family a researcher can get. It is not easy to run a family with two researchers, doing clinical work day and night and on top of that raise two children. Time together has been sparse at times, but now we can look back at those times with pride, feel of achievement and satisfaction. Without your endless and unlimited love and support this thesis would never have seen daylight. I cannot ever thank you enough for the support received.

This thesis has been supported by following instances, Maud Kuistila foundation, Alma and K.A. Snellman foundation, Pediatric Research Foundation in Finland, Finnish medical research foundation and Duodecim ry.

Oulu, 24 February 2009

Johan Löfgren

"When once you have tasted flight, you will forever walk the Earth with your eyes turned skyward, for there you have been, and there you will always long to return"

– Leonardo Da Vinci
Abbreviations

aa amino acid
AOM acute otitis media
BAL bronchoalveolar lavage
bp base pair
BPD bronchopulmonary dysplasia
CD cluster of differentiation
CF cystic fibrosis
CL-P1 collectin placenta 1
CL-L1 collectin liver 1
CL-43 collectin of 43 kDa
CL-46 collectin of 46 kDa
CMI cell-mediated immunity
COPD chronic obstructive pulmonary disease
CRD carbohydrate recognition domain
CTL cytotoxic T lymphocytes
DNA deoxyribonucleic acid
DPPC dipalmitoylphosphatidylcholine
ER endoplasmic reticulum
ET Eustachian tube
fMLP N-formyl peptides
G-CSF granulocyte-colony stimulating factor
GI gastrointestinal
HWE Hardy–Weinberg equilibrium
IFN interferon
IL interleukin
IRAK interleukin-1 receptor-associated kinase
LPS lipopolysaccharide
MBL mannose-binding lectin
MEFs middle ear fluids
mRNA messenger RNA
NK natural killer (cells)
NSAID non-steroidal anti-inflammatory drug
OM otitis media
OME otitis media with effusion
PAMP pathogen-associated molecular pattern
<table>
<thead>
<tr>
<th>Abbr.</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC</td>
<td>phosphatidylcholine</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PMN</td>
<td>polymorphonuclear leucocytes</td>
</tr>
<tr>
<td>RDS</td>
<td>respiratory distress syndrome</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RIP</td>
<td>receptor-interacting protein</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>ROM</td>
<td>recurrent otitis media</td>
</tr>
<tr>
<td>RSV</td>
<td>respiratory syncytial virus</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>SP</td>
<td>surfactant protein</td>
</tr>
<tr>
<td>TDT</td>
<td>transmission disequilibrium test</td>
</tr>
<tr>
<td>TGF</td>
<td>transforming growth factor</td>
</tr>
<tr>
<td>Th</td>
<td>T helper cell</td>
</tr>
<tr>
<td>TIRAP</td>
<td>Toll-interleukin 1 receptor (TIR) domain containing adaptor protein</td>
</tr>
<tr>
<td>TLRs</td>
<td>Toll-like receptors</td>
</tr>
<tr>
<td>TNF</td>
<td>tumour necrosis factor</td>
</tr>
<tr>
<td>TRAF</td>
<td>TNF receptor-associated factor</td>
</tr>
</tbody>
</table>
List of original papers

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:


## Contents

Abstract
Abbreviations ................................................................. 9
List of original papers ...................................................... 11
Contents ............................................................................. 13

1 Introduction ..................................................................... 17

2 Review of the literature .................................................. 19

2.1 Immunity ....................................................................... 19

2.1.1 The human immune system ........................................ 19

2.1.2 Acquired immunity ................................................... 20

2.1.3 Innate immunity ....................................................... 20

The complement system ................................................... 21

2.2 Lungs/airways ............................................................... 22

2.3 Respiratory syncytial virus infections in infants ............... 24

2.3.1 Respiratory syncytial virus ......................................... 24

2.3.2 RSV infections in infants .......................................... 25

2.3.3 Epidemiology of RSV and RSV bronchiolitis ............... 26

Prophylaxis, treatment and outcome of severe RSV infection 27

2.4 Otitis media ................................................................. 28

2.4.1 Acute and recurrent otitis media ................................. 28

2.4.2 Otitis media and viral infections ................................ 30

2.5 Genetic risk factors of respiratory disease ...................... 31

2.6 Soluble proteins involved in innate immunity .................. 31

2.6.1 Mannose-binding lectin ............................................. 31

2.6.2 CD14 ....................................................................... 32

2.6.3 Cytokines .................................................................. 32

2.6.4 Chemokines ............................................................. 33

2.6.5 Cytokines associated with RSV ................................. 33

2.7 Collectins ...................................................................... 34

2.7.1 Collectins in innate immunity ..................................... 36

2.8 Human pulmonary surfactant ........................................ 36

2.9 SP-B and SP-C ............................................................ 38

2.10 Surfactant protein A ..................................................... 38

2.10.1 SP-A structure ........................................................ 38

2.10.2 SP-A production and secretion ................................. 39

2.10.3 Functions of SP-A .................................................. 40
6 Discussion
6.1 Role of SP-A and SP-D in severe RSV infections ........................................ 67
6.2 TLR4 receptor 299 polymorphism in RSV infection................................. 70
6.3 Otitis media and SP-A gene variation .................................................... 72

7 Summary and conclusions ................................................................. 75

References ......................................................................................... 77

Original papers ............................................................................... 103
1 Introduction

The respiratory syncytial virus (RSV) is one of the most common pathogens in early childhood respiratory tract infections. It infects most children by their third year of life and approximately 1% of infants require hospital care because of severe RSV bronchiolitis. The clinical manifestations of RSV infections vary from rhinitis to lethal respiratory insufficiency. The mortality rate in developed countries associated with RSV-related infections in hospitalized infants is 0.5–1%. Predisposing factors of severe RSV infections are male gender, premature birth, congenital and chronic diseases, parental smoking and immunodeficiency. Treatment of RSV bronchiolitis is today still symptomatic, but some prophylactic treatments are available. Infection with RSV is often accompanied by otitis media (OM), a very common upper respiratory tract infection in children. The RS virus can be found in OM infections as a co-pathogen.

Surfactant proteins A and D (SP-A and SP-D) are produced primarily in lung alveoli and on respiratory tract epithelium. They are involved in regulation of lung surfactant secretion and surface activity. Both proteins also have an important role in inborn host defence functions. They are capable of binding to pathogens and function as a first line of defence before the development of acquired immunity in infants. Toll-like receptors (TLRs) are a family of receptors capable of recognizing pathogens in the human body and activating a host immune response. Toll-like receptor 4 (TLR4) has been shown to be associated with RSV and it is involved in innate immune responses.

The aim of the present work was to investigate the genetic association and heritability of innate defence proteins in connection with RSV and OM infections in infants. Earlier laboratory and animal studies had suggested evidence of an association between innate defence proteins and the two diseases.
2 Review of the literature

2.1 Immunity

2.1.1 The human immune system

The human immune system has been estimated to have developed over hundreds of millions of years (Marchalonis et al. 2002). The function of this system is to respond to attacks by pathogens and defend the body against other external threats. There are similarities between the human immune system and those of species that are more primitive. This system will keep on gradually developing into the future, depending on the environmental challenges that have arisen even in the recent past.

The human organs that are involved in the immune system are not centralized in one place, but scattered throughout the body. They are called lymphoid organs, as they are involved in lymphocyte cell development and deployment. Lymphocytes are white blood cells; they are the major cells in the immune system. The major organs that are involved in lymphoid immunity are the bone marrow, the thymus, spleen, adenoids, tonsils, appendix and lymph nodes throughout the body. Small patches in the small intestine and lymphatic and blood vessels are also counted as belonging to the organs involved in immunity (Janeway CA et al. 2001).

The bone marrow produces the necessary T and B lymphocytes from hematopoietic stem cells. The thymus is involved in teaching the T-cells to distinguish alien cells from the body’s own ones. In case of failure to learn, T-cells would attack the body’s own cells. The readied cells migrate to the lymphoid organs and nodes. The bone marrow also produces phagocytes, which include monocytes and macrophages. Granulocytes, erythrocytes and thrombocytes are also products of the pluripotent stem cells (Janeway CA et al. 2001). In addition to the specialized bone marrow-derived cells, other cells carry out functions supporting the immune system.

The human immune system can be divided into two components, covering acquired and innate immunity. Today we know that these are not separate systems, but they work together in various ways to provide protection from outside threats to the body.
2.1.2 Acquired immunity

Acquired immunity involves B lymphocytes and T lymphocytes. These two cell types divide acquired immunity into cell-mediated (T cells) and humoral (B cells) immunity.

There are three types of cells involved in cell-mediated immunity (CMI). The T cells can be suppressor T cells or regulatory T cells. As the name indicates, the suppressor cells suppress and regulate the immune response and they are involved in the control of autoimmunity (Bluestone & Tang 2005). Helper T cells direct B cells and macrophages to their target areas where they kill pathogenic microbes. The third T cell type includes the cytotoxic T lymphocytes (CTLs). These cells are capable of destroying pathogen-infected cells. The CTLs do this by secretion of agents that destroy the target cells. They recognize their targets by the foreign antigen fragments on the cell surface. Natural killer (NK) cells are also T cells; these cells kill pathogens and cancer cells without recognizing the target cells.

Acquired immunity requires previous exposure to a certain pathogen so that large quantities of antibodies are produced to deter invaders. The acquired immune system has a memory of how it has dealt with the pathogen earlier and can this way react quickly to threats.

Human B cells (lymphocytes) are responsible for humoral immunity. These cells secrete antibodies against a specific antigen. A B cell can produce only one type of unique antibody. The immunological memory is provided by T and B cells that turn into memory cells that are then activated when needed (Janeway CA et al. 2001).

2.1.3 Innate immunity

Innate immunity is an ancient form of host defence that relies on a small number of germ line-encoded receptors. These receptors are encoded in our genome and therefore are principally functional after they have been synthesized. They have evolved to recognize conserved microbial components, also called pathogen-associated molecular patterns (PAMPs), thereby enabling the immune system to distinguish between microbes and the body’s own cells. The PAMPs include lipopolysaccharides (LPSs), aldehyde-derivatized proteins, mannans, teichoic acids, denatured DNA and bacterial DNA (Medzhitov & Janeway, Jr. 2002). The innate immune system uses several different receptors to recognize pathogens. Toll-like receptors (TLRs) and the collectin family of proteins provide a means
for the body to sense foreign cells. Their function is to recognize certain molecular patterns belonging to pathogens and to initiate immune responses.

The innate immunity system involves the use of phagocytes as soldiers that ingest and kill microbes. There are two types of phagocytes. The first type is represented by microphages, also called polymorphonuclear leucocytes (PMNs), which form in the bone marrow and circulate in the bloodstream. They have a short half-life, as they participate in the inflammatory process during tissue damage and trauma. They can also prevent local bacterial growth (Kobayashi & DeLeo 2004, Alves-Filho et al. 2006). The second type is represented by macrophages and these develop from monocytes. Tissue macrophages are located throughout the body at strategic locations, such as lung alveoli, the skin, intestine and the abdominal and chest cavities (Janeway CA et al. 2001). Macrophages function as the first line of defence in these areas. After phagocytosis, the microbe must be killed. This can be done either aerobically by producing oxygen radicals that disrupt the microbe, or anaerobically, when the microbe is deprived of essential iron. Another anaerobic method is to change the acidity of the phagocyte’s internal environment. After killing the microbe, the macrophages accumulate in lymph nodes to present the microbe to the acquired immunity system (Janeway CA et al. 2001).

**The complement system**

The complement system comprises a number of proteins in the bloodstream that react in sequence with one another to induce an inflammatory reaction and to opsonize microbes. The complement proteins are activated in a cascade where the active complement enzyme cleaves its substrate, which in turn becomes the active enzyme for the next substrate downstream. The cascade provides a rapid and strong response at the location of inflammation. The complement on pathogen surfaces can be activated in three pathways (Figure 1), the classic pathway where the pathogen is bound directly to the first complement protein (C1q) in sequence or through adaptive immunity activation with antibody-antigen complexes. The second pathway is through the mannan-binding lectin, when it binds to microbes. As a third option, we have an alternative pathway, where a complement protein is activated spontaneously and binds to a microbe. All three pathways lead to activation of the complement system. The active complement produces chemotaxins and recruits phagocytes and activates them. It can also opsonize or kill the microbes by creating pores in the membrane of the bacteria. To aid the
defence reaction, proinflammatory chemicals are produced (Janeway CA et al. 2001, Carroll 2004, Gasque 2004).

![Diagram of complement activation pathways]

**Fig. 1.** The complement activation pathways. C1q: first protein in the complement cascade. MBL: mannose-binding lectin. SAC: spontaneously activated complement.

### 2.2 Lungs/airways

Human airways are divided into upper and lower parts. The upper airway consists of the oral cavity, nasal cavity, pharynx, larynx and epiglottis. The lower airways include the trachea, bronchi, the smaller airways and the alveoli (Figure 2).
Fig. 2. Human lung airways.

The airways are covered with airway epithelial cells. The histology is different at the proximal and distal ends of the airway. The oral cavity has a stratified squamous type of epithelium and is covered with a thick mucus membrane to protect it. The histology of the nasal cavity, larynx and pharynx is characterized by a pseudostratified columnar ciliated epithelium with goblet cells and there are a large number of serous and mucous glands. The trachea in the lower respiratory tract has histological features similar to those in the larynx. In the primary bronchii, the epithelium is somewhat thinner than that in the upper airways and the number of goblet cells is decreased. The bronchial epithelium is pseudostratified, the underlying lamina containing large amounts of elastin and mast cells. In bronchioles less than 1 mm in diameter, the epithelium type is simple ciliated columnar, with very few goblet cells. The smooth muscle layer is
very prominent and involved, for example, in asthma. Finally, the terminal alveoli are mostly lined with flat type I pneumocytes (96%) (Castranova et al. 1988). There are also round type II pneumocytes that are involved in surfactant metabolism (Burkitt H.G. et al. 1993).

The upper airways have several functions related to exchange of air. They provide passage to the lower airways and they also clean, warm and moisturize the air before it enters the lower airways and lungs. All along the way to the alveoli, the airway epithelial cilia and mucous membranes also clean the air. At the terminal alveoli, oxygen and carbon dioxide are exchanged through the type I pneumocytes that create a thin permeable layer between the alveolar wall and the bloodstream (Ramirez et al. 1999).

2.3 Respiratory syncytial virus infections in infants

2.3.1 Respiratory syncytial virus

The respiratory syncytial virus (RSV) is an RNA virus belonging to the Paramyxoviridae virus family. It has two fusion proteins, F and G that can interact with outside molecules. The RS virus has an envelope around it, a helical nucleocapsid. It is pleomorphic, presenting in various morphological forms in the population. The spherical form of RSV ranges between 80 and 350 nm in diameter, whereas the length of the filamentous form varies between 60–100 nm. The virus has a single stranded RNA chain, consisting of 15 200 nucleotides. An RNA polymerase and modifying enzymes are contained in the virion. This RNA chain is capable of producing 11 functional proteins (Hacking & Hull 2002). The sequence of the genes encoding the proteins is shown in Figure 3.

![Fig. 3. The encoded genes of the RS-virus.](image)

The RS virus has four nucleocapsid proteins, N, P, L and M2. The N nucleoprotein is needed for transcriptional activity and binds tightly along genomic RNA. The P nucleoprotein is also needed for transcriptional activity; it associates with free N and L proteins for maintenance. Protein L is the RNA polymerase enzyme. The protein M2 has two forms, M2-1 and M2-2. The first is a transcription elongation factor and second regulates viral transcription in the
host cell. The three transmembrane virion proteins of RSV are the F, G and SH proteins. The F protein is the fusion protein that enables viral entry into the host cell and aids in the formation of the RSV syncytia. The G transmembrane glycoprotein is used for viral attachment to the host cell surface. The function of the small hydrophobic SH protein is unknown. The NS1 and NS2 proteins are non-structural ones; they have shown anti-interferon α and β activity. The M protein is a matrix protein involved in viral assembly (Johnson et al. 1987, Johnson & Collins 1988, Johnson & Collins 1989, Collins et al. 1990, Garcia et al. 1993, Grosfeld et al. 1995, Hacking & Hull 2002).

Several genes of the RS virus can be altered somewhat without destroying viability, including those of the NS1, NS2, SH, M2-2 and G proteins (Schmidt et al. 2001). However, mutation of the genes for the N1 and N2 proteins have an effect on viral replication efficiency (Whitehead et al. 1999, Teng et al. 2000).

The human RS virus has been shown to have genetic variation and there are two main groups of RSV, A and B. These two groups differ in both nucleotide sequence and antigen variation. Both groups can further be divided into several genotypes or strains. It has been proposed that there might be a difference in virulence between the two groups. Group A and B infections can also to some extent explain reinfection with RSV.

### 2.3.2 RSV infections in infants

The respiratory epithelial cells are the first line of innate immunity defence against RSV infections. They are capable of producing opsonins, cytokines and collectins (surfactant proteins A and D) (Persson et al. 1988, Voorhout et al. 1992, Patel et al. 1995, Noah et al. 1995b, Becker et al. 1997). The airway epithelial cells are the primary targets of RSV infection. Alveolar macrophages in the lungs react to the infection and regulate the immune response; they also produce and secrete pro-inflammatory cytokines (Becker et al. 1991, Midulla et al. 1993). Monocyte CD14 receptor and Toll-like receptor 4 have been shown to bind to the RS virus. They both have been shown to stimulate cytokine responses (IL-1β, IL-6, IL-8 and TNF-α) (Kurt-Jones et al. 2000). However, Welliver et al. demonstrated that RSV infection does not seem to activate the classical T lymphocyte production of the cytokines IL-2, IL-4, IL-17 and IFN-γ (Welliver et al. 2007). The pathogenesis of RSV in infants has not been studied very much. It has been suggested that the lung function deficiency could be a result of cell
sloughing as the RS virus replicates and causes direct lung injury (DeVincenzo 2007).

Exposure to RSV does not appear to occur before birth and adaptive immune insufficiency or greater innate responses may account for early life RSV-induced illnesses (Krishnan et al. 2003). An infant develops a cellular immune response against RSV with CTLs within 10 days of infection if it is the infant’s first RSV infection. Cytotoxic T lymphocytes are not capable of recognizing all of the RSV proteins (Chiba et al. 1989, Cherrie et al. 1992).

2.3.3 Epidemiology of RSV and RSV bronchiolitis

Respiratory syncytial viruses cause annual epidemics during the winter months that vary geographically. Such epidemics occur in Finland in cyclic periods. Every other year there is a larger RSV epidemic during the winter months, with the start in the autumn of the previous year. In the spring of intervening years there is an epidemic peak, usually ending before the summer (Ruuskanen & Ogra 1993).

Respiratory syncytial virus is the major causative agent of bronchiolitis, a disease that affects infants under 1 year old. In older children, RSV infection more usually causes obstructive bronchitis. RSV is the most significant respiratory pathogen in early childhood (Parrott et al. 1973). It is estimated that every child has an RSV infection during the three first years of life (Ruuskanen & Ogra 1993). In adults RSV causes only mild symptoms. The incubation period for RSV is 5–8 days. The virus is very contagious and spread mostly by hand-to-eye and hand-to-nose contacts. It can survive on surfaces for several hours. Other viruses have also been detected in acute expiratory wheezing. Jartti et al. showed that enteroviruses and rhinovirus were present in up to 42% of cases in their study (Jartti et al. 2004). Human metapneumovirus has been detected in children with clinical bronchiolitis symptoms. The metapneumovirus belongs to the same Paramyxoviridae family as RSV (van den Hoogen et al. 2001, Williams et al. 2004). Co-infections with RSV have been shown in severe bronchiolitis, with metapneumovirus and Bordetella pertussis. Patients infected with RSV alone have been reported to have a longer hospitalization time and more severe hypoxia than patients with RSV and metapneumovirus co-infection (Korppi & Hiltunen 2007, Canducci et al. 2008).

Respiratory syncytial virus bronchiolitis is diagnosed on the basis of clinical symptoms and seasonal occurrence. Definitive confirmation is obtained by
antigen testing, PCR or virus culture. In most infants, RSV manifests itself as congestion and rhinitis, possibly with mild fever. In more difficult forms, coughing, wheezing and dyspnoea may occur. The symptoms usually resolve in 5–7 days. RSV bronchiolitis affects the lower airways, producing plugs of hyper-secreting mucus, causing inflammation and oedema in the mucosal membrane and virus-induced necrosis in the epithelium. This restricts normal airflow to and from the lungs, causing laboured breathing and decreased oxygenation in infants (McIntosh K 1966). In at least 1% of infants with RSV infection the patients require hospital treatment as a result of respiratory distress suggesting bronchiolitis. About 10% of these infants are in need of intensive care. The reported mortality rate of infants with RSV bronchiolitis is 0.2–0.71% (Shay et al. 2001, Simon et al. 2006).

The severity of RSV bronchiolitis depends on age. The younger the patient is, the smaller the airways are and the disease is more difficult (Anderson et al. 1990, American Academy of Pediatrics Committee on Infectious Diseases: Use of ribavirin in the treatment of respiratory syncytial virus infection 1993). Elderly people and high-risk adults (with chronic lung or heart disease) are also prone to RSV infection. A small number of patients require intensive care and the mortality rate is 8% in hospitalized cases (Falsey et al. 2005). Contributing factors to a more difficult disease form are premature birth, bronchopulmonary dysplasia (BPD), cystic fibrosis (CF), congenital heart diseases, male sex, parental smoking, communal or group day-care, immunodeficiency, low levels of maternally received RSV antibodies and low socioeconomic status (Glezen et al. 1981, MacDonald et al. 1982, McConnochie & Roghmann 1986, Groothuis et al. 1988, Meert et al. 1990, Griese et al. 1997). Breastfeeding has been shown to reduce the concentrations of soluble CD25 (soluble IL-2 receptor), which again is associated with the severity of RSV infections (Roine et al. 2005).

Prophylaxis, treatment and outcome of severe RSV infection

Attempts to develop a vaccine against RSV have not been successful. According to current evidence, prophylactic treatment of high-risk groups with monoclonal antibodies against RSV decreases the risk of severe RSV infections (Groothuis et al. 1993, Feltes et al. 2003, Romero 2003). Such treatments are very expensive and limited to very select cases, mostly to premature infants during their first months of life. The humanized monoclonal antibody Palivizumab binds to the RSV F protein, which mediates virus entry into cells and the formation of
syncytia. Recently a more potent antibody (Motavizumab) has been under development and in clinical trials. This antibody could also be effective in preventing RSV infection in the upper respiratory tract (Wu et al. 2007).

Several treatments have been tried for bronchiolitis patients at pediatric wards. These include mono- and polytherapies of $\beta_2$ agonists, i.e. bronchodilators, inhaled and systemic corticosteroids, ribavirin, non-steroidal anti-inflammatory drugs (NSAIDs) and racemic epinephrine (adrenalin). However, the latest treatment guidelines do not recommend using these as routinely (American Academy of Pediatrics 2006).

Nasopharyngeal suction is used to help breathing, as approximately 60% of the breathing resistance caused by RSV is induced in the upper airways (Kneyber et al. 2000, Black 2003). Supplemental oxygen is given as required, on the basis of transcutaneous measurement of oxygen saturation. The prognosis for infants is good if there are no confounding factors and oxygenation is good. Most infants hospitalized spend only 2–4 days on the ward. The most difficult RSV cases usually need intensive care and in some cases respiratory assistance (Green et al. 1989, Stretton et al. 1992).

2.4 Otitis media

2.4.1 Acute and recurrent otitis media

The middle ear is the space behind the tympanic membrane, where the hearing chain bones are also located (Figure 4). The function of the middle ear is to amplify sounds and pass them forward to the inner ear where the sensory cells are located, which transform pressure changes into neural input.

Acute and recurrent middle ear infections are very common in early childhood. The peak age is one year in the Finnish population (Alho et al. 1991). With approximately 500 000 cases of AOM yearly in Finland, these infections are a very large burden on the primary healthcare system and a major economic concern (Niemela et al. 1999). Otitis media (OM) is inflammation of the middle ear. There are different types of otitis media; acute otitis media (AOM) is described as the presence of fluid, typically pus, in the middle ear and an abnormal looking tympanic membrane. The effusions in the middle ear can be purulent, serous, mucoid or combinations. Additionally, there can be symptoms such as pain in the ear region, fever, coughing, restlessness, irritability, hearing
problems, nausea and vomiting. There is also a more chronic form of otitis media called otitis media with effusion (OME), where there is fluid in the middle ear for a longer time, but the middle ear is not necessarily infected.

![Anatomical view of the human ear](image)

**Fig. 4. Anatomical view of the human ear.** The ear is divided into three parts, the outer, middle and inner ear. The outer ear starts from the tympanic membrane and continues outwards, including the ear leaf. The middle ear consists of the space behind the tympanic membrane and also the Eustachian tube (ET). The middle ear contains the hearing chain bones. The inner ear contains the neural sensory organs that detect sound and the ones that produce positional and accelerative information.

In AOM, bacteria are found in approximately 50%–90% of cases (Ruohola et al. 2006). The most common bacteria found in middle ear infections are *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis* (Kilpi et al. 2001). Viruses can be found in 20%–49% of middle ear fluids and co-infections with bacteria in up to 66% of cases (Heikkinen et al. 1999, Heikkinen & Chonmaitree 2003, Ruohola et al. 2006).

Risk factors of AOM are parental smoking (Stahlberg et al. 1986), use of pacifiers (Niemela et al. 1995), attending large group day-care centres, bottle feeding (Duffy et al. 1997), male sex, large families, sibling history of recurrent otitis media (ROM) (Pukander et al. 1985, Teele et al. 1989) and viral respiratory infections (Ruuskanen et al. 1989). Interference in Eustachian tube (ET) function
can predispose infants to OM. The Eustachian tubes in infants are shorter and more horizontal than those of adults and they are also narrower and less stiff (Chantzi et al. 2005). There is also evidence that the muscular opening function increases with age and that children with otological diseases have poorer ET function than healthy children (Bylander 1984). Expression of surfactant proteins (SPs) has been shown in the ET (Paananen et al. 1999, Paananen et al. 2001). Such proteins have been shown to lower opening pressures of the ET and shorten the duration of OME and the severity and duration of AOM (Chandrasekhar & Mautone 2004).

### 2.4.2 Otitis media and viral infections

Several viruses have been shown to be present in middle ear fluids (MEFs) during OM. According to Ruohola et al., picornaviruses (rhinoviruses or enteroviruses) are present in 60% of cases of AOM; other viruses found are bocavirus, coronavirus, influenza virus and RSV (Ruohola et al. 2006). The results of other studies suggest that RSV is the most common respiratory virus found in MEF. It can be detected in MEF in 74% of children with RSV infection (Heikkinen et al. 1999). Respiratory viruses have been shown to augment adhesion of common bacteria to the respiratory epithelium. Parainfluenza virus and RSV have also been found to increase the expression of bacterial receptors (Avadhanula et al. 2006). The RSV G protein has been shown to be a binding receptor for both *Haemophilus influenzae* and *Streptococcus pneumoniae* (Avadhanula et al. 2007). Studies on twins and triplets have shown that there is a formidable genetic component in susceptibility to OM and RSV infections (Casselbrant et al. 1999, Casselbrant et al. 2004). RSV has been shown to increase the risk of OM (Henderson et al. 1982, Uhari et al. 1995, Andrade et al. 1998). Interferon-γ (IFN-γ) genotypes have been shown to be associated with OM in RSV-infected infants (Gentile et al. 2003). Tumour necrosis factor-α and interleukin 6 polymorphisms are associated with an elevated risk of OM (Patel et al. 2006). It has been suggested that patients with RSV-associated AOM could be at an elevated risk of relapse, possibly as a result of a weakened host defence after primary infection (Sagai et al. 2004).
2.5 Genetic risk factors of respiratory disease

Many respiratory diseases have risk factors that include both environmental aspects and polygenic variations. There are a few single gene diseases where the genetic error is clearly the cause of the disease. Cystic fibrosis and α1-antitrypsin deficiency are two examples of such single gene diseases (Kolb et al. 2006). One of the most studied polygenic diseases in infancy is respiratory distress syndrome (RDS). Several genes have been shown to be associated with RDS, most importantly the surfactant protein genes (Hallman & Haataja 2007). Asthma has also been studied extensively; several genes have been shown to be associated with the disease and several environmental factors are also involved (Meurer et al. 2006). Infection with RSV, and OM, are both polygenic diseases that have been associated with genes of proteins that are involved in human immunity. Among these are the collectins, Toll-like receptors, cytokines and immunoglobulins (Abreu & Arditi 2004, Patel et al. 2006, Bossuyt et al. 2007).

2.6 Soluble proteins involved in innate immunity

When our bodies are attacked by pathogens, an immune response is initiated as the first line of defence. The innate response system includes both soluble and cellular products. The soluble ones include cytokines, chemokines, complement products and other proteins.

2.6.1 Mannose-binding lectin

Mannose-binding lectin (MBL) belongs to the lectin family of proteins capable of recognizing pathogens and activating the lectin pathway of the complement system. It binds to surface carbohydrates of bacteria, fungi, viruses and parasites (van Emmerik et al. 1994, Neth et al. 2000, Townsend et al. 2001, Neth et al. 2002). It is produced in the liver and circulates in the bloodstream (Thiel et al. 1992). A deficiency of MBL seems to be a common trait and it is a predisposing factor as regards infection (Turner 1996). The amount of MBL in the bloodstream is associated with three SNPs in the structural region of the MBL2 gene at codons 52, 54 and 57, and two gene promoter polymorphisms at positions -221 and -550 (Minchinton et al. 2002). No association between RSV and decreased MBL levels has been found (Nielsen et al. 2003, Kristensen et al. 2004). In addition, no association has been shown between low MBL levels and AOM/ROM, even
though acute respiratory infection is associated with the above-mentioned SNPs (Homoe et al. 1999, Koch et al. 2001).

### 2.6.2 CD14


The CD14-TLR4 complex has been shown to interact with the RSV virus (Wright et al. 1990, Pugin et al. 1994, Kurt-Jones et al. 2000). The CD14 gene polymorphism CD14/–159 is not associated with severe RSV infections (Tal et al. 2004).

### 2.6.3 Cytokines

Cytokines are used in the body to provide cells and tissues with intra- and intercellular tools to communicate. Such communication is used in several functions, for example innate and adaptive immunity, inflammation and infection, development, haematopoiesis and tissue repair (Haddad 2002). The major cytokines are the interleukins (ILs) 1–15, interferon (IFN) α, β and γ and tumour necrosis factor (TNF) α and β. Helper T cells (Ths) have two important functions, they promote inflammation and immune responses in cells (Th type 1 cells) and help B cells to produce antibodies (Th type 2 cells). Th1 cells produce IL-2, IFN-γ, IL-12 and TNF-β. Th2 cells produce IL-4, IL-5, IL-6, IL-10 and IL-13 (Haddad 2002).
2.6.4 Chemokines

Chemokines (chemotactic cytokines) are a family of pro-inflammatory inducible cytokines. They are glycoproteins with molecular weights of 6–14 kDa and contain 70–125 amino acids (Rollins 1997, Baggiolini et al. 1997). There are 4 subfamilies of chemokines (C, CC, CXC and CX3C), differentiated on the basis of cysteine residue configuration. The C chemokines are lymphocyte-specific and there are two species. The CC chemokines (28) attract basophils, eosinophils, lymphocytes (NK and T cells) and monocytes. The CXC chemokines (16) are neutrophil chemoattractants, but basophils, eosinophils and T cells also respond to some extent. The last group, the CX3C chemokines, induce adhesion and migration of leucocytes (Rollins 1997, Baggiolini et al. 1997, Moser & Willimann 2004, Smit & Lukacs 2006). The function of chemokines is to induce chemotaxis in autocrine, paracrine or endocrine systems and direct cells involved in innate immunity (Moser & Willimann 2004, Tripp et al. 2005). Chemokine production is usually initiated in microbial infections or damage of tissue. Several cell types can release chemokines. The CX3C and CC subfamilies have been shown to be associated with RSV infections and with Th1/Th2 cytokine balance (Tripp et al. 2005).

2.6.5 Cytokines associated with RSV

The interleukin-10 receptor is a member of the interferon receptor family. Interleukin-10 plays a central role in immunological balancing; it is an anti-inflammatory cytokine. It is capable of suppressing Th1-type harmful immune responses and can promote protective Th2 responses. It is capable of increasing the proliferation of T cells and B cells (Moore et al. 2001). The severity of RSV bronchiolitis has been shown to be associated with the IL-10 gene locus. Two single nucleotide polymorphisms (-1117 and -3585 downstream) have been found to be associated with a need for respirator treatment. The IL-10 allele -529C has been shown to be overrepresented in hospitalized RSV patients. Parental transmission of the allele has also been found to be common in such patients (Hoebbe et al. 2004). In addition, this genotype has also been proposed to be associated with decreased expression of IL-10 (Turner et al. 1997, Westendorp et al. 1997, Wilson et al. 2005). Levels of IL-10 in nasal aspirate have been shown to be increased during RSV infection (Bont et al. 2002).
Interleukin-4 (IL-4) is involved in Th2 cell stimulation, promoting B cell proliferation. It also counteracts interferon gamma (IFN-γ) production in Th1 cells. Interferon-γ has been shown to be protective in cases of severe RSV infection (Aberle et al. 1999, Bont et al. 2001). Interleukin-4 has been shown to be associated with severe RSV bronchiolitis in two studies. In both studies the IL-4 C-590T (or C-589T) polymorphism showed significant overrepresentation of the -590T promoter variant (Choi et al. 2002, Hoebee et al. 2003).

Interleukin-8 (IL-8) has been shown to be associated with severe RSV bronchiolitis. High concentrations of IL-8 have been found in infant nasal lavage samples and in blood plasma during the infection (Biswas et al. 1995, Abu-Harb et al. 1999). Epithelial cells have also been shown to produce large quantities of IL-8 when infected with RSV (Fiedler et al. 1995, Noah et al. 1995a, Fiedler et al. 1996). The IL-8 -251A allele has been found to be significantly more transmitted to affected children than expected (p=0.014). The analysis was carried out by means of the transmission disequilibrium test (TDT) (Hull et al. 2000).

Tumour necrosis factor alpha (TNF-α) is a major proinflammatory cytokine in infections, produced by macrophages, NK and B cells. Interleukin-9 (IL-9) again promotes Th2-type cell proliferation and activates bronchial epithelial cells. However, IL-9 A-345G and TNF-α G-308A polymorphisms have not shown significant genetic association with RSV bronchiolitis infections (Hoebee et al. 2004).

Interleukin-13 (IL-13) C-1112T polymorphism has been shown to have an association with severe RSV infection. This same polymorphism has also been shown to be associated with gene expression and nuclear factor binding to the IL-13 promoter (Puthothu et al. 2006). Disproportionate IL-13 expression has been shown to cause mucus over-expression in the airways, causing breathing difficulties (Lukacs et al. 2001, Tekkanat et al. 2002, Park et al. 2003, Rudd et al. 2006).

2.7 Collectins

Collectins are a family of molecules that are involved in host innate defence. They belong to the mammalian C-type lectins and all have a similar structure. The collectins have four regions, the carbohydrate-recognition domain (CRD) which is the C-terminal lectin, an α-helical neck domain, a collagen-like region and a cysteine-rich N-terminal (Holmskov et al. 1994) (Figure 5). The basic functional unit is a trimer and the different collectins have different numbers of these trimers
Collectins are produced and released at several locations around the body. Eight collectins have been recognized and all have typical properties of the family. Mannose-binding lectin (MBL), surfactant protein A (SP-A), surfactant protein D (SP-D), collectin placenta 1 (CL-P1) and collectin liver 1 (CL-L1) are found in humans and animals (Kawasaki et al. 1978, White et al. 1985, Persson et al. 1988, Ohtani et al. 1999, Ohtani et al. 2001). A collectin of 43 kDa (CL-43), one of 46 kDa (CL-46) and conglutinin are found only in bovine species (Hansen & Holmskov 2002).

Fig. 5. SP-D and SP-A protein domains, trimeric subunits and tertiary structures.
2.7.1 Collectins in innate immunity

The collectins are capable of recognizing conserved surface carbohydrate structures on pathogens via their CRDs and can in this way distinguish between foreign molecules and the body’s own ones (van de Wetering et al. 2004). Collectins participate in innate host defense in several ways. They have been shown to activate the complement cascade, act as opsonins and activate phagocytosis through phagocytic receptors. They also bring about agglutination of microbes and can inhibit their growth (McNeely & Coonrod 1994, O'Riordan et al. 1995, van Rozendaal et al. 2000b, Petersen et al. 2001). The collectins can modulate different responses of the immune system. Inflammatory, adaptive immune and allergic responses are known to be modulated by MBL, SP-A and SP-D (Malhotra et al. 1993, Wang et al. 1996b, Brinker et al. 2001, Brinker et al. 2003).

2.8 Human pulmonary surfactant

Human pulmonary surfactant is a surface-active substance that is found in the human lung, lining the interior. Its primary function is to lower the surface tension in the alveolar air-liquid interface and thus prevent alveolar/lung collapse when expiration occurs in the lung (Clements 1977). Disease states or congenital lack of surfactant lead to severe respiratory distress as a result of atelectasis and lung oedema. The most common neonatal problem associated with lack of surfactant is respiratory distress syndrome (RDS) (AVERY & MEAD 1959). Surfactant is also involved in innate host defense through surfactant proteins. These proteins are capable of pathogen recognition and can initiate immune responses (Holmskov 2000, Sano & Kuroki 2005).

Alveolar surfactant complex is produced by type II alveolar cells. The surfactant complex secreted into the alveolar space forms a net-like structure called tubular myelin. This is the film-like form that surfactant takes on at the alveolar air-liquid interface. The alveolus is lined by a lipid monolayer (Figure 6). Human pulmonary surfactant consists of approximately 90% lipid and 10% protein (King 1982). The lipids comprise approximately 90% phospholipids, cholesterol-carrying neutral lipoproteins being the most abundant.

Proteins associated with the surfactant complex are called surfactant proteins A, B, C and D (SP-A, SP-B, SP-C and SP-D). These proteins are produced by alveolar type II cells and in bronchiolar epithelial cells (Clara cells), except for

Fig. 6. Schematic diagram of an alveolus. Modified from Hawgood & Clements 1990.

The surfactants at the different locations vary in composition and function. Conductive airway surfactant has a similar composition as alveolar surfactant, but concentrations of surfactant proteins are lower and the surface tension-lowering capacity is less profound (Bernhard et al. 1997). The surfactant in the ET has a different composition from the pulmonary ones and there are lower concentrations of phosphatidylcholine (PC) in the ET. Eustachian tube surfactant does not have a very low surface tension, allowing closure of the tube (Paananen et al. 2002). The composition of gastric mucosa surfactant is similar to that of the conductive airways (Bernhard et al. 1995).
2.9 SP-B and SP-C

Surfactant protein B is an 8 kDa hydrophobic protein. The gene transcripts produce a 381-amino acid preprotein that is processed further into a mature 79-amino acid protein (Weaver & Whitsett 1989, Guttentag et al. 1998, Beck et al. 2000). SP-B is always associated with the phospholipids of surfactant (Weaver & Conkright 2001). SP-C is synthesized to either a 191- or a 197-amino acid proprotein, depending on alternative splicing at exon 5 (Warr et al. 1987, Glasser et al. 1988a, Glasser et al. 1988b). The proprotein is then processed to a mature 35-amino acid protein with a molecular mass of 3.5 kDa (Nogee 1998, Weaver & Conkright 2001).

The main function of SP-B and SP-C, together with the lipids, is to lower the surface tension at the alveolar air-liquid interface and prevent alveolar collapse at the end of expiration. Alveolar surfactant is critically dependent on the presence of SP-B, but the role of SP-C is not fully understood (Beck et al. 2000). It is produced exclusively in the lungs, whereas SP-B is additionally expressed in Clara cells and the Eustachian tube (Phelps & Floros 1988, Paananen et al. 2001).

2.10 Surfactant protein A

2.10.1 SP-A structure

Surfactant protein A is the most abundant of the surfactant proteins (5.3%) (Weaver & Whitsett 1991, Kishore et al. 2006). It is a large hydrophilic protein with a molecular mass of 28–36 kDa, depending on glycosylation (Whitsett et al. 1985). The primary structure of SP-A is conserved and analogous throughout species in nature. The SP-A mature translation product has 248 amino acids (Hoppe & Reid 1994, Kishore et al. 2006). The primary product has an N-terminal signalling peptide. This peptide is removed during the processing of SP-A in the cell (Phelps et al. 1986, Drickamer et al. 1986). The SP-A protein monomer has four primary structures, the amino terminal domain, the collagenous domain in the middle, the neck domain and the carbohydrate recognition domain (CRD). SP-A forms trimers and SP-A normally exists as an octadecameric form with six trimers in a flower bouquet formation (Voss et al. 1988) (Figure 5). The trimers have been suggested to consist of two SP-A1 units and one SP-A2 unit (Voss et al. 1991).
The amino terminal (N-terminal) segment is composed of 7–11 amino acids, depending on species, and contains two cysteine (Cys) residues. The Cys residues have been shown in rat SP-A to be required for multimerization of SP-A and they are involved in interchain disulphide bonding (Elhalwagi et al. 1997). The N-terminal segment is involved in arranging the trimeric subunits of SP-A in a parallel conformation (Palaniyar et al. 2001).

The human SP-A collagen-like domain contains 73 amino acids and consists of 23 Gly (glycine)-X-Y repeats, where X and Y in SP-A are proline or hydroxyproline residues (McCormack et al. 1997, McCormack 1998, Kishore et al. 2006). The Gly-X-Y repeat can be easily interrupted by inserting a single amino acid. This is evident in human SP-A at Gly 44, which causes a typical bend in the SP-A collagen domain (Voss et al. 1988). SP-A1 contains an additional Cys 85 residue in the collagen domain, whereas SP-A2 does not (Berg et al. 2000). This residue could promote formation of larger multimers via intertrimer bonds and enhance stability in products composed of both SP-A1 and SP-A2 (Palaniyar et al. 2001, Kishore et al. 2006).

The neck domain of SP-A contains 34 amino acids and connects the collagen region to the CRD. The neck domain is critical in protein folding, mediating intertrimeric protein-protein interactions (Palaniyar et al. 2000). It may also be involved in orienting the CRDs for selective ligand binding (McCormack 1998).

The carbohydrate recognition domain has 123 amino acids that fold into a ligand binding unit. The structure of the CRD is stabilized by two pairs of disulphide bonds (Haagsman et al. 1989, Berg et al. 2000, Palaniyar et al. 2000, Kishore et al. 2006). There are two Ca\(^{2+}\)-binding sites in the CRD and its functions are dependent on Ca\(^{2+}\) (Haagsman et al. 1990, Sohma et al. 1992). Binding of calcium to the CRD results in translocation and conformational change in the domain (Haagsman et al. 1990, Casals et al. 1993).

### 2.10.2 SP-A production and secretion

Surfactant protein A has been shown to be produced in the lungs by alveolar type II cells and bronchiolar epithelial cells (Clara cells) (Persson et al. 1988, Khoor et al. 1993). Additionally, SP-A is expressed in the Eustachian tube and other locations (Paananen et al. 1999). It can be detected as early as in amniotic fluid and after birth its production increases greatly (Hallman et al. 1989). Production of SP-A involves several post-translational changes to the preprotein first synthesized, these modifications taking place in the rough endoplasmic reticulum.
(ER) and the Golgi apparatus. Folding and assembly of SP-A is mostly done before exiting the ER (McCormack 1998). About 10% of SP-A is secreted from the lamellar bodies, an organelle in type II cells (Froh et al. 1993). Transport of SP-A from the Golgi complex to the lamellar bodies is carried out via multivesicular bodies (Voorhout et al. 1993). Most of the SP-A is secreted through other routes, bypassing the lamellar body (Ikegami et al. 1992, Froh et al. 1993, Osanai et al. 2006).

2.10.3 Functions of SP-A

Surfactant protein A has several functions ranging from surfactant homeostasis to innate host defence. It is involved in regulating surfactant and it inhibits type II cell secretion of surfactant phospholipids (Dobbs et al. 1987, Rice et al. 1987). It also enhances dipalmitoylphosphatidylcholine (DPPC) uptake by type II cells and improves surface tension through adsorption of phospholipids at the air-liquid interface (Hawgood et al. 1987, Weaver & Whitsett 1991, Haagsman & van Golde 1991, Kuroki & Akino 1991). SP-A has been shown to be important in promoting formation of tubular myelin (Poulain et al. 1992).

Surfactant protein A binds to carbohydrate moieties of pathogens, providing innate recognition and defence for the body (Holmskov 2000, van de Wetering et al. 2004). It has been shown to have anti-viral, anti-bacterial and anti-fungal activities (Lawson & Reid 2000, Wright 2005). It can also directly inhibit bacterial and fungal growth (Wu et al. 2003). Macrophage phagocytosis is directly activated by SP-A and the amount of macrophage receptors increases under the influence of SP-A (Beharka et al. 2002, Kuronuma et al. 2004). SP-A binds to apoptotic cells (Schagat et al. 2001) and it is involved in cytokine regulation; it can suppress Th2-type cytokines and enhance the production of Th1 cytokines (Madan et al. 2001).

2.10.4 SP-A genes and genetic variation

Human SP-A has two functional proteins SP-A1 and SP-A2, also known as SFPT1/SFTP1A and SFTP1B/SFTP1A respectively. The two genes for these proteins have very strong nucleotide (94%) and amino acid (96%) homology (White et al. 1985, Floros et al. 1986, Katyal et al. 1992, Floros & Hoover 1998). There is also a non-functional SP-A pseudogene (SFTP1P1) (Korfhagen et al. 1991). The genes are located on the long arm of chromosome 10 (10q22.2-23.1).
(Figure 7) (Hoover & Floros 1998) and the two functional SP-A genes are in linkage disequilibrium (Floros et al. 1996). Both of the genes have 4 exons that are translated and 5 introns that are not translated. The most common genetic variations studied in the SP-A genes are located in exons 1, 2 and 4 (Table 1, figure 8). The SP-A genes are highly polymorphic and contain several single nucleotide polymorphisms (SNPs) (McCormick et al. 1994, Karinch & Floros 1995). The combinations of polymorphisms in an individual’s gene determine the alleles, one maternally and one paternally inherited. The two alleles form the genotype. SP-A1 and SP-A2 are in strong linkage disequilibrium, i.e. a specific set of SP-A1 and SP-A2 alleles reside in the same chromosome. SP-A alleles are denoted 6A^n and the SP-A2 alleles 1A^m. In the same manner, the genotypes are denoted 6A^n/6A^n and 1A^m/1A^m. There are differences in the distribution of individual haplotypes between races and ethnic groups (Liu et al. 2003). In the Finnish population the number of haplotypes is lower than in other populations, suggesting population homogeneity (Ramet et al. 2000).

Table 1. SP-A polymorphisms.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene location</th>
<th>AA affected</th>
<th>AA change</th>
<th>Nucleotide change</th>
<th>AA characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP-A1</td>
<td>Exon 1</td>
<td>19</td>
<td>Ala/Val</td>
<td>C/T</td>
<td>Phob/Phob</td>
</tr>
<tr>
<td></td>
<td>Exon 1</td>
<td>50</td>
<td>Leu/Val</td>
<td>C/G</td>
<td>Phob/Phob</td>
</tr>
<tr>
<td></td>
<td>Exon 2</td>
<td>62</td>
<td>Pro</td>
<td>A/G</td>
<td>Phob</td>
</tr>
<tr>
<td></td>
<td>Exon 4</td>
<td>133</td>
<td>Pro</td>
<td>A/G</td>
<td>Phob</td>
</tr>
<tr>
<td></td>
<td>Exon 4</td>
<td>219</td>
<td>Arg/Trp</td>
<td>C/T</td>
<td>Phob/Phob</td>
</tr>
<tr>
<td>SP-A2</td>
<td>Exon 1</td>
<td>9</td>
<td>Asn/Thr</td>
<td>A/C</td>
<td>Phil/Phil</td>
</tr>
<tr>
<td></td>
<td>Exon 2</td>
<td>91</td>
<td>Pro/Ala</td>
<td>C/G</td>
<td>Phob/Phob</td>
</tr>
<tr>
<td></td>
<td>Exon 4</td>
<td>140</td>
<td>Ser</td>
<td>C/T</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exon 4</td>
<td>223</td>
<td>Lys/Gln</td>
<td>A/C</td>
<td>Phil/Phil</td>
</tr>
</tbody>
</table>

AA, amino acid; Ala, alanine; Val, valine; Leu, leucine; Arg, arginine; Trp, tryptophan; Asn, asparagine; Thr, threonine; Pro, proline; Lys, lysine; Ser, serine; Gin, glutamine; A, adenine; C, cytosine; G, guanine; T, thymine; Phob, hydrophobic amino acid; Phil, hydrophilic amino acid.
Fig. 7. Human chromosome 10. The collectins SP-A, SP-D and MBL are located close together between 10q21.1–23.1. P is the SP-A pseudogene.
Surfactant protein A has been shown to be associated with RSV. The distribution of SP-A protein in the lungs during RSV infection changes and the total amount of surfactant protein in bronchoalveolar lavage (BAL) fluid increases. Levels of SP-A mRNA increase 3-fold early in the infection. No significant association has been shown between SP-A levels and the severity of RSV infection (Kerr & Paton 1999, Alcorn et al. 2005). The susceptibility of preterm infants to RSV could be a
result of low expression of SP-A in the immature lung (Meyerholz et al. 2006). In
SP-A knock-out mice pulmonary clearance of RSV is decreased, viral titres are
higher and lung inflammation is more severe. Levels of proinflammatory
cytokines (IL-6 & TNF-α) are elevated. The production of oxygen radicals is also
decreased. Exogenous SP-A reduces inflammation and viral titres (LeVine et al.
1999). SP-A increases the uptake of RSV by monocytes and macrophages.

Surfactant protein A binds to RSV through the F (fusion) glycoprotein,
interacting with the highly glycosylated F2 subunit. The binding is Ca²⁺-
dependent, suggesting involvement of the CRD. SP-A might neutralize RSV by
blocking the F protein, which is essential for viral entry. SP-A can act as an
opsonin with other viruses; thus it has been suggested that SP-A could do this
with RSV (Ghildyal et al. 1999, Sano et al. 2003). It has been shown that SP-A
enhances attachment and entry of RSV into cells. As SP-A also activates viral
haemolysis, it considerably decreases infectivity (Hickling et al. 2000, Sano et al.
2003).

2.11 Surfactant protein D

2.11.1 SP-D structure

Surfactant protein D (SP-D) is a 43 kDa hydrophilic glycoprotein (Persson et al.
1989). A 50 kDa monomer form has been reported in BAL fluid, possibly
resulting from further post-translational glycosylation (Mason et al. 1998). SP-D
has a 25-amino acid N-terminal domain, a 177-amino acid collagen-like domain,
a neck domain and a carbohydrate recognition domain. The long collagen-like
domain consists of 59 Gly-X-Y repeats and unlike SP-A it has no interruptions,
and is therefore linear (Rust et al. 1991, Crouch et al. 1994). SP-D also forms
basic trimers, as does SP-A, but the trimers in SP-D assemble themselves into a
cross-shaped tetramer of 520 kDa and also more complex oligomers (Kishore et
al. 2006). The dodecamer form is the most common for SP-D (Crouch et al.
1994). The dodecamer is stabilized by trimer interchain disulphide bonds. The
collagen-like domain and neck domain are necessary in dodecamer formation; the
neck domain is also involved in trimer organization (Ogasawara & Voelker 1995,
Zhang et al. 2001).
2.11.2 SP-D production and secretion

Surfactant protein D is produced in several locations throughout the body, but main production is in the lungs by Clara cells and type II pneumocytes. SP-D has also been shown to exist in the human Eustachian tube and in mouse conductive airways (Persson et al. 1988, Voorhout et al. 1992, Crouch et al. 1992, Wong et al. 1996, Paananen et al. 1999). Production of SP-D can be detected at 16 weeks of gestation, before SP-A can be detected. The levels of SP-D increase towards the end of gestation (Dulkerian et al. 1996). SP-D undergoes post-translational modifications; a signal peptide is removed and partial hydroxylation of collagen region prolines and lysines occurs (Holmskov et al. 1995). In rat lung, SP-D has been detected in the endoplasmic reticulum, Golgi complex and multivesicular bodies, but not in the lamellar bodies (Voorhout et al. 1992).

2.11.3 Function of SP-D

Surfactant protein D is involved in both innate host defence and in metabolic and homeostatic functions. Surfactant homeostasis is disrupted in SP-D knockout mice, causing accumulation of SP-A and SP-B in the alveolar space and a 10-fold increase in alveolar macrophages. In addition, type II cells and lamellar bodies show changes in morphology and tubular myelin is decreased (Korfhagen et al. 1998, Botas et al. 1998). SP-D does not seem to be involved in surfactant surface activity and lack of SP-D is not associated with decreased survival (Botas et al. 1998, Ikegami et al. 2001).

Surfactant protein D can bind to alveolar macrophages and enhance the production of oxygen radicals (van Iwaarden et al. 1992, Miyamura et al. 1994). It is capable of recognizing a wide range of pathogens including gram-positive and gram-negative bacteria, viruses, yeasts and fungi. It is capable of agglutinating pathogens and the binding is calcium-dependent (Kuan et al. 1992, Hartshorn et al. 1994, Schelenz et al. 1995, Madan et al. 1997, Hartshorn et al. 1998, van de Wetering et al. 2001). In vitro, SP-D can inhibit viral infectivity and it can enhance phagocytosis and killing of fungi (Madan et al. 1997, Reading et al. 1997). SP-D can inhibit bacterial and yeast growth by increasing pathogen membrane permeability and it has also been shown that SP-D inhibits fungal growth (van Rozendaal et al. 2000a, Wu et al. 2003, McCormack et al. 2003). Like SP-A, SP-D is capable of binding to apoptotic cells and it enhances the uptake of these by alveolar macrophages (Schagat et al. 2001). SP-D is also

### 2.11.4 SP-D and RSV bronchiolitis

Surfactant protein D has been shown to be associated with respiratory syncytial virus. The level of SP-D in BAL fluid decreases during RSV infection in infants and already low levels of SP-D could lead to increased susceptibility to RSV bronchiolitis (Kerr & Paton 1999). Premature lamb lungs have been shown to have decreased amounts of SP-D (Meyerholz et al. 2006). Early studies with recombinant human SP-D indicated that it would not bind to RSV glycoproteins F or G, whereas SP-A bound to the fusion protein (Ghildyal et al. 1999). Later studies showed that recombinant human SP-D binds to the attachment glycoprotein of human RSV in Hep-2C cells, presumably at the CRD (Hickling et al. 1999). Rat SP-D has been shown to bind to both the attachment and the fusion proteins of RSV (LeVine et al. 2004). Intranasal administration of SP-D to RSV-infected mice has been shown to reduce the amount of virus and to inhibit RSV replication (Hickling et al. 1999).

Surfactant protein D knock-out mice have shown decreased clearance of RSV virus together with increased inflammation and augmented recruitment of inflammatory cells. Both knock-out and protein-producing mice showed decreased oxygen radical production after RSV infection. SP-D enhances phagocytosis of RSV in vitro and in vivo lack of SP-D causes a decrease of alveolar macrophage function (LeVine et al. 2004). Levels of SP-D mRNA in type II pneumocyte cells have been found not to show an increase when challenged with RSV, whereas SP-A mRNA showed a three-fold increase (Alcorn et al. 2005).

### 2.11.5 The SP-D gene and genetic variation

The SP-D protein is encoded from a single 11 kb gene (*SFTPD*, earlier denoted *SFTP4*) located at 10q23.3 (Figure 7). The gene has seven exons, exon 1 encoding the amino terminal domain and a short part of the collagen-like domain. Exons 2–5 encode the rest of the collagen-like domain, the neck region is encoded by exon 6 and the CRD by exon 7 (Crouch et al. 1993) (Figure 9).
Like the SP-A gene, the SP-D gene is polymorphic and contains several SNPs. These gene variations have been studied in pulmonary diseases (Floros et al. 2000, Guo et al. 2001, Rova et al. 2004, Krueger et al. 2006). The two most common SNPs studied are Met11Thr and Ala160Thr (DiAngelo et al. 1999) (Table 2). The SP-D gene has also been shown to have differences in allele distribution between ethnic groups and different races (Liu et al. 2003). The gene variation at Met11Thr affects serum concentrations of SP-D and the oligomerization of subunits in serum. The allelic variants of SP-D have also been shown bind microbes with varying degrees of potency (Leth-Larsen et al. 2005, Heidinger et al. 2005).

Table 2. SP-D polymorphisms.

<table>
<thead>
<tr>
<th>Gene Location</th>
<th>AA Affected</th>
<th>AA Change</th>
<th>Nucleotide Change</th>
<th>Hydrophobic/hydrophilic AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1</td>
<td>Met/Thr</td>
<td>T/C</td>
<td>Phob/Phil</td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>Ala/Thr</td>
<td>G/A</td>
<td>Phob/Phil</td>
<td></td>
</tr>
<tr>
<td>Exon 7</td>
<td>Ser/Thr</td>
<td>T/A</td>
<td>Phil/Phil</td>
<td></td>
</tr>
</tbody>
</table>

AA, amino acid; Met, methionine; Thr, threonine; Ala, alanine; Ser, serine; A, adenine; C, cytosine; G, guanine; T, thymine; Phob, hydrophobic amino acid; Phil, hydrophilic amino acid.
2.12 Toll-like receptors

The Toll receptor protein was first found in *Drosophila melanogaster*, the fruit fly, when studying the genetics of the dorso-ventral polarity of embryonic Drosophila. The protein appeared to have extracellular, transmembrane and cytoplasmic domains (Anderson *et al.* 1985, Hashimoto *et al.* 1988, Morisato & Anderson 1995). The findings led to a search for a mammalian equivalent; the interleukin 1 receptor (IL-1R) was found to have a similar cytoplasmic domain, but it did not have the pattern-recognition functions of Toll (Gay & Keith 1991). The first human Toll-like receptor (hTLR) was reported in 1997. This TLR is today named TLR4 (Medzhitov *et al.* 1997). So far 10 mammalian hTLRs have been identified (Chaudhary *et al.* 1998, Rock *et al.* 1998, Takeuchi *et al.* 1999, Chuang & Ulevitch 2000, Du *et al.* 2000, Sebastiani *et al.* 2000, Chuang & Ulevitch 2001).

The TLRs all have a similar structure, with a cytoplasmic domain, a transmembrane domain and an extracellular domain with several leucine-rich repeats. The signalling pathways are similar in Toll, TLRs and IL-1R, due to homology in the cytoplasmic domain. The DNA sequences of the TLRs are highly conserved between different species (Hallman *et al.* 2001). TLRs are expressed in several locations – in endothelial, epithelial and myelomocytic cells. They can also be found in several organs. How the TLRs behave when activated depends on the location and receptor (Muzio *et al.* 2000c, Zarember & Godowski 2002). Several ligands have been identified for the TLRs. These pathogen-associated molecular patterns (PAMPs) include bacteria, viruses and fungi and TLRs can also recognize double- and single-stranded RNA (Kopp & Medzhitov 2003, Takeda *et al.* 2003, Heil *et al.* 2004).

2.13 TLR4

2.13.1 Structure of TLR4

Toll-like receptor 4 is an 839–841 amino acid member of the TLR family of receptors which are type I transmembrane proteins (Medzhitov *et al.* 1997, da Silva & Ulevitch 2002). The exact structure of the human TLR4 ectodomain is still unknown, although some predictions have been made (Rallabhandi *et al.* 2006). TLR4 has the typical domains of the TLRs – the extracellular domain has leucine-rich repeats (LRRs) and the intracellular signalling domain is analogous
to that of IL-1R. The human TLR4 extracellular domain comprises 21 LRRs, at the N-terminal end there is a 31-aa-long flanking region and the C-terminal end is flanked by a cysteine-rich domain (Medzhitov et al. 1997). The extracellular domain contains 9 N-linked glycosylation sites that are essential in maintaining the function of the receptor. It has been found that glycosylation site-mutated TLR4s cannot be transported to the cell surface (da Silva & Ulevitch 2002). The intracellular domain comprises the 160-aa Toll-interleukin 1 receptor (TIR) domain, which is essential for signalling. The TIR domain has three important regions (boxes 1, 2, 3). Their function is to serve as a signature sequence, to engage with downstream elements and to transmit signals (Xu et al. 2000, Slack et al. 2000, O'Neill et al. 2003). TLR4 has been shown to form homodimers both by ligand-induction and without it and these dimers can activate downstream pathways (Zhang et al. 2002, Saitoh et al. 2004).

2.13.2 Expression of TLR4

Toll-like receptor 4 has been found in many organs, including the lungs, the heart, placenta, small intestine, stomach, bone, the eye, smooth muscle, liver, skin and other organs. The expression of TLR4 is regulated by several ligands – microbial components, LPSs and inflammatory mediators. The up- or down-regulation of expression depends on tissue and cell type (Muzio et al. 2000a, Muzio et al. 2000c, Armstrong et al. 2004a, Armstrong et al. 2004b). Many tissues have been shown to express TLR4, such as retinal pigment cells and corneal epithelial cells (Song et al. 2001, Elner et al. 2005), gingival fibroblasts of the gums (Wang et al. 2000), cardiomyocytes and coronary endothelia in the heart (Frantz et al. 1999), lung bronchial epithelium and alveolar septal cells (Janardhan et al. 2006). Expression of TLR4 in the gastrointestinal region has been shown for gastric pit cells of the stomach (Kawahara et al. 2001), hepatic Kupffer and stellate cells (Su et al. 2000, Paik et al. 2003), biliary and pancreatic epithelial cells (Li et al. 2005, Yokoyama et al. 2006, Li et al. 2006), intestinal epithelial cells and foetal enterocytes (Caro et al. 2000, Fusunyan et al. 2001). Prostate epithelial cells, placental trophoblasts and cervical smooth-muscle cells also show expression of TLR4 (Watari et al. 2000, Holmlund et al. 2002, Gatti et al. 2006). In the skin, keratinocytes, dermal endothelial cells and fat adipocytes show TLR4 protein activity (Faure et al. 2000, Lin et al. 2000, Pivarcsi et al. 2003). Mouse osteoblasts in the bone show activation of TLR4 by LPS (Kikuchi et al. 2001).
TLR4 is found on immunocompetent cells like monocytes, macrophages and dendritic cells (Muzio et al. 2000b).

### 2.13.3 Function and signalling of TLR4

![Diagram of TLR4 signalling pathways]

**Fig. 10. TLR4 MyD88-dependent and -independent signalling pathways**

The main function of TLR4 is to provide pathogen recognition at an early time point after microbial entry into the host. TLR4 recognizes microbes and quickly (within minutes) initiates an immune response; it also assists in the killing of the pathogens. TLR4 can stimulate cytokine and chemokine secretion. Adaptive immunity takes days to respond – this shows the importance of this innate system. Newborn infants do not have any form of adaptive immunity and rely solely on the innate system and maternal antibodies. TLR4 and the innate system are involved in the development of adaptive immunity (Beutler 2005).

The mammalian TLR4 signalling cascade is similar to the *Drosophila* pathway. The signalling can be divided into the myeloid differentiation factor (MyD)88-dependent and MyD88-independent pathways (Figure 10). The TIR
domain-containing adapter protein (TIRAP) functions as the first chain between TLR4 and MyD88 downstream. MyD88 recruits IL-1R-associated kinases (IRAKs) to form a receptor complex. There are four IRAKs, of which IRAK-4 is the most critically involved in the pathway. IRAK-4 associates with TNF receptor-associated factor 6 (TRAF6), this in turn activates TGF-β-activated kinase (TAK) 1, which activates the IκB kinase (IKK) complex. IκB inhibition is reversed through phosphorylation. Finally, NF-κB is activated to start gene expression, leading to cytokine and IFN-β production (Takeda & Akira 2004, Kaisho & Akira 2006).

The MyD88-independent pathway starts with TLR4 signalling the TRIF-related adaptor molecule (TRAM). This in turn binds to the TIR-domain-containing adaptor protein, inducing IFN-β (TRIF, also known as TICAM-1). TRIF interacts with receptor-interacting protein (RIP) 1, which activates the IKK complex as in the MyD88-dependent pathway. TRIF can also interact with TRAF6 and start the pathway from there (Takeda & Akira 2004, Kaisho & Akira 2006).

IRAK-4 deficiency has been shown to cause selective predisposition to bacterial infections and it impairs TLR-dependent innate antiviral immune responses. IRAK-4 deficiency seems particularly predispose individuals to invasive pneumococcal disease (McDonald et al. 2006, Ku et al. 2007).

2.13.4 TLR4 and RSV

Toll-like receptor 4 has been shown to be involved in RSV infections. It mediates the innate immunity response to RSV fusion protein. TLR4 knock-out mice have shown inability or difficulty in clearing the lungs of RSV and the viral load in the lungs is higher than in wild-type mice (Kurt-Jones et al. 2000). Further work on knock-out mice has shown a deficiency in natural killer cell function and pulmonary trafficking, a decrease of IL-12 expression and confirmed the difficulty in clearing the lungs of RSV (Haynes et al. 2001). The primary response to RSV infection is dependent on TLR4 and alveolar macrophages, whereas the secondary response is mediated by inflammatory cells and epithelial cells (Haeberle et al. 2002). RSV up-regulates TLR4 expression in monocytes during the acute phase of infection in vivo. It can be seen as increased mRNA and protein and cell membrane localization. These changes also sensitize airway epithelial cells to other environmental factors, such as LPSs (Monick et al. 2003, Gagro et al. 2004). Studies have shown that TLR4 mRNA levels increase during
gestation; preterm infants could be at a disadvantage as a result of to reduced expression of TLR4 (Meyerholz et al. 2006). Contrasting results as regards the role of TLR4 have also been published. In a knock-out mouse study TLR4 had no impact on virus elimination and there was no effect on NK cell recruitment or activity. In addition, no pulmonary inflammatory cells were recruited (Ehl et al. 2004). Genetic association studies have shown that TLR4 polymorphisms (299 & 399) are correlated to severe RSV bronchiolitis (Tal et al. 2004, Mandelberg et al. 2006).

2.13.5 The TLR4 gene and genetic variation

The human TLR4 gene is 19 kb long, it is located at 9q32-33 and it has four exons (Rock et al. 1998, Smirnova et al. 2000). The function of the human TLR4 fourth exon is unknown and it is believed that translation of this exon leads to production of a non-functional protein. Different mammalian species show homology in the amino acid sequence of TLR4 – chimpanzee and baboon TLR4 show over 93% similarity to human TLR4 (Smirnova et al. 2000). The human TLR4 gene is polymorphic (White et al. 2003); the most studied SNPs are located at aa 299 and aa 399. These two polymorphisms are located in the extracellular domain of TLR4 (Arbour et al. 2000, Smirnova et al. 2001). It has been predicted that these polymorphisms could affect ligand and co-receptor binding. Both amino acids seem to be on the same face of the domain (White et al. 2003, Rallabhandi et al. 2006). The Asp299Gly and Thr399Ile SNPs have been studied in connection with several diseases.
3 Outline of the present study

In early life, infants are susceptible to infections, as their host defence functions are insufficiently developed, the amount of antibodies received transplacentally decreases and because of lack of exposure to infections the cells involved in acquired immunity are naive. Thus, the role of the innate immune system is of critical importance in the absence of acquired immunity. During the first few months, infants solely depend on maternal antibodies and the innate immunity system. The airways have a critical function in defence against infections and indeed the most serious and most common infections in infancy and early childhood involve the upper and the lower airways. RSV causes annual epidemics during the winter months. Recurrent middle ear infections are the most common infections during childhood. Constitutional and environmental factors play an important role in the susceptibility to common infections. According to preliminary evidence, the susceptibility to common serious infections may in part depend on genetic factors. The individual genes encoding the components of the host defence system reveal polymorphism.

Here we proposed that the proteins serving as receptors for microbes influence the predisposition to specific airway infections. Experimental studies have shown that mice deficient in molecules from this innate system have more infections and they are severe. Among the receptors for microbes, SP-A, SP-D and TLR4 molecules have been shown to play specific roles in the innate defence system of the airway epithelium and SP-A has been shown to be present in the epithelium of the Eustachian tube. Furthermore, polymorphism is likely to influence the function of proteins encoded by these genes.

These findings prompted us to set the following aims for this study:

1. To study whether susceptibility to severe RSV infections is influenced by polymorphism of SP-A1 and SP-A2.
2. To investigate the association between the alleles and genotypes of SP-D, and severe RSV bronchiolitis in infants.
3. To investigate the genetic variation of TLR4 and its association with severe RSV bronchiolitis in Finnish infants.
4. Establishing SP-A as a protein component expressed in the epithelium of the Eustachian tube led us to study the association between SP-A haplotypes and recurrent otitis media.
4 Subjects and methods

4.1 Ethical considerations

The studies in this thesis conform to the 1995 Helsinki accord. The study protocols were approved by all hospital ethics committees participating in the studies. Patients were given information about the study and its aim in a written form. Clinical data was obtained by means of a questionnaire filled in by the patient’s parents or guardians. Some data was obtained from hospital records, including hospitalization time, ward of treatment and laboratory results. In addition, the patient’s records were screened for concomitant genetic disease and neonatal/infection history. Informed consent documents were signed by all participants or guardians in all studies.

4.2 Study design

The present work is based on candidate gene approaches to determine genetic associations concerning the chosen molecules and cases of severe RSV bronchiolitis and recurrent otitis media.

The population for the RSV study was collected over a time period of seven years (1999–2005). Altogether, 1322 RSV study samples were collected. Samples were collected on site at Oulu University Hospital Children’s Hospital and Southern Ostrobothnia Central Hospital. A large portion of samples was collected retrospectively from the Northern Ostrobothnia Hospital District and a smaller batch from Southern Ostrobothnia.

The inclusion criteria for the RSV group of infants were: positive antigen test in nasopharyngeal secretion (carried out by the hospital laboratory as a normal procedure for confirming RSV; in the same antigen-immunofluorescence test were included adenovirus, influenza A and B, and parainfluenza 1–3), admission to the ward with severe clinical symptoms of RSV bronchiolitis, including difficulty in breathing, decreased oxygenation measured by pulse oximeter or a need for supplemental oxygen. In addition the children were required to be under one year of age and not to have any other concomitant infectious diseases. The selected subjects had no genetic diseases and patients who had received prophylactic RSV treatment (Palivizumab) were excluded. All cases, controls and their parents were required to be of Finnish origin.
The controls were recruited from hospital birth records, searching for matched infants of the same sex and with the same place of residence, birthday (± 2 weeks) and length of gestation at birth (± 2 weeks). In addition, the selected controls had no upper or lower airway infections or diseases, with the exception of 1 case of otitis media with or without effusion. Clinical data was collected in the same way as for the RSV-positive infants. The questionnaire was used to collect data concerning neonatal diseases, acute and chronic lung diseases, middle ear infections, allergies, type of day-care, parental smoking, family structure and information about the birth and length of breastfeeding.

The infants in the otitis media study were recruited at the Department of Otolaryngology, Oulu University Hospital. They were patients undergoing elective surgery for tympanostomy and/or adenoidectomy because of recurrent middle ear infections. The controls for this study were collected at the same hospital from consecutively born infants. No concomitant diseases or genetic disorders were present in either population. The infants qualifying for the ROM population had had more than five acute otitis media infections.

4.3 Study populations

In the SP-A association study (I) the population consisted of 86 infected infants and their 95 closely matched controls (Table 3). The original group was larger, but because of our stringent matching criteria, the number of cases decreased to the above-mentioned number. These samples were collected both prospectively (12/1999–5/2000) and retrospectively (1/1998–11/1999).

For the SP-D paper (II) the same base population was used as a starting point as in the SP-A (I) study. The final number of matched and successfully genotyped cases was 84 and there were 93 controls. The matching criteria were sex, place of residence, birthday and length of gestation at birth within 2 weeks (Table 3).

The TLR4 study (III) involved a non-matched base population of 345 RSV-positive infants and 383 controls (Table 3). These samples were collected between 1999–2004. Samples from the parents were also collected at the same time for future projects.

For study IV, 147 children with recurrent otitis media were recruited. The 278 controls for this study were collected at Oulu University Hospital (Table 3).
4.4 DNA extraction and genotyping

The samples for the studies came in as whole blood and buccal smear samples. All samples from Oulu University Hospital and Southern Ostrobothnia came as buccal samples, except for the Southern Ostrobothnia samples during the 12/1999–5/2000 period, when they came as whole blood samples.

The buccal samples were collected using sterile cotton swabs (Copan) from the infant’s cheek mucus membrane. Two buccal swabs were collected from all subjects. The blood samples were collected in connection with normal hospital laboratory diagnostics, using normal sample collection techniques.

The DNA from the buccal smear swabs was extracted using 10% Chelex-100 medium ((Bio-Rad Laboratories) and proteinase K. Negative controls of sterile water, Chelex medium and proteinase K were used to ensure contamination-free samples. Visibly contaminated swabs were excluded. From whole blood samples genomic DNA was extracted using Puregene DNA Isolation Kits (Gentra Systems) and 0.5–3 ml of whole blood was used for this. The whole blood samples were shipped frozen and stored at -70 °C until processing at the Paediatric Research Laboratory, Oulu University. The buccal samples were processed as soon as they arrived at the laboratory.

Genetic analysis of the three candidate genes was performed using the polymerase chain reaction (PCR) technique. Water controls were used to ensure that no contamination was present in the reactions.

4.4.1 SP-A

For SP-A and TLR4, allelic analysis/genotyping was carried out using the restriction fragment length polymorphism (RFLP) method. For SP-A, 9 polymorphic sites (SNPs) were analysed to determine the genotypes for SP-A1 (5 sites) and SP-A2 (4 sites) (DiAngelo et al. 1999). Polymerase chain reactions were carried out with specific primers; the products were digested with specific restriction enzymes for all sites. Polyacrylamide gel electrophoresis (PAGE) and ethidium bromide (EtBr) staining were used to separate and analyse digested fragments.
4.4.2 SP-D

SP-D genotyping was carried out in connection with three polymorphic sites. These were located at amino acids 11, 160 and 270. Met11Thr was genotyped as described earlier (DiAngelo et al. 1999). For Ala160Thr a PCR-RFLP method with restriction enzyme digestion and gel electrophoresis was used to determine the genotype. The Ser270Thr genotype was investigated by using a single PCR reaction and enzymatic digestion of the product. The product was visualized by means of agar gel electrophoresis (II).

4.4.3 TLR4

TLR4 genotyping was performed by way of two nested PCRs that were used to amplify the SNP at amino acid 299. The site-specific primers used for the first PCR were forward CCTGTGCAATTGTACCATTG and reverse TCATGGTAATAACACCATTGAAGC. The nested primers for the second PCR were forward GATTAGCATACTTTAGACTACTACCTCCCATG and reverse GATCAACTTTCTGAAAAAGCATTCCAC. The conditions for the first reaction were 95 °C for 1 min (denaturation), 60 °C for 1 min (annealing) and 72 °C for 1 min (elongation) and for the second PCR, 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec. The obtained PCR product was digested with NcoI restriction enzyme for 5 hours at 37 °C. The digested product was separated on 2% agar gel, stained with EtBr and photographed under UV light.

4.5 Statistical analysis

Allele and genotype frequency comparisons were carried out using 2 × 2 and 2 × k tables. $\chi^2$ analysis was performed using these tables. When the expected number of cases was less than 5, Fisher’s exact test was used for calculations. Conditional logistic regression analysis was used to see whether confounding factors affected the results and associations.

The software used for these statistical calculations comprised SPSS for Windows (basic statistical calculations, logistical regression analysis), Arcus Quickstat ($\chi^2$ analysis for 2 × 2 and 2 × k tables, Fisher’s exact test), Egret for Windows (conditional regression analysis for matched data) and Microsoft Excel for Windows (basic statistical calculations and data manipulation).
5 Results

5.1 SP-A genetic association with severe RSV

5.1.1 Patient characteristics

We genotyped and analysed 86 specimens from patients with severe RSV and 95 control samples. The samples were matched for sex, gestational age at birth, date of birth and hospital district. With the exception of a greater number of otitis media episodes among the patients, characteristics were similar in both groups. The percentage of premature infants was 6% in both groups, reflecting the low prematurity rate. The clinical data are presented in Table 3. The mean age at the time of hospitalization at Oulu Children’s Hospital was $87 \pm 68$ (SD) days and the length of hospitalization $4.3 \pm 2.4$ days. In all 11% of the RSV patients in Oulu required treatment in the paediatric intensive care unit at some point.

Table 3. Patient characteristics in the studies.

<table>
<thead>
<tr>
<th>Study/Characteristics</th>
<th>Severe RSV/OM</th>
<th>Controls</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP-A/RSV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>86</td>
<td>95</td>
<td>Matched</td>
</tr>
<tr>
<td>Mean GA ± SD (days)</td>
<td>278 ± 12</td>
<td>278 ± 11</td>
<td></td>
</tr>
<tr>
<td>Male/Female %</td>
<td>64 / 36</td>
<td>64 / 36</td>
<td></td>
</tr>
<tr>
<td>SP-D/RSV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>84</td>
<td>93</td>
<td>Matched</td>
</tr>
<tr>
<td>Mean GA ± SD (days)</td>
<td>277 ± 12</td>
<td>277 ± 10</td>
<td></td>
</tr>
<tr>
<td>Male/Female %</td>
<td>64 / 36</td>
<td>65 / 35</td>
<td></td>
</tr>
<tr>
<td>TLR4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>312</td>
<td>356</td>
<td>Unmatched</td>
</tr>
<tr>
<td>Mean ± SD GA (days)</td>
<td>274 ± 18</td>
<td>278 ± 13</td>
<td></td>
</tr>
<tr>
<td>Male/Female %</td>
<td>60 / 40</td>
<td>57 / 43</td>
<td></td>
</tr>
<tr>
<td>SP-A/OM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>147</td>
<td>278</td>
<td>Unmatched</td>
</tr>
<tr>
<td>Male/Female %</td>
<td>66 / 34</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RSV, respiratory syncytial virus; OM, otitis media; SP, surfactant protein; GA, gestational age; TLR4, Toll-like receptor 4.
5.1.2 Allele analysis

There was a significant difference between the RSV and control groups in allele distribution of the SP-A2 gene ($p=0.006$, $\chi^2$-analysis) and a trend in the SP-A1 gene ($p=0.071$) (Table 4). The SP-A1 allele 6A was underrepresented in the RSV children compared with the healthy controls (frequency 0.01 vs. 0.06, $p=0.011$, OR 0.17, 95% CI 0.04–0.8). There were significant differences in individual SP-A2 alleles. The RSV group showed underrepresentation of the 1A allele (0.01 vs. 0.06, $p=0.011$, OR 0.17, 95% CI 0.04–0.8) and significant overrepresentation of the 1A3 allele (0.05 vs. 0.005, $p=0.006$, OR 10.4, 95% CI 1.3–83.2) when compared with the healthy controls. In terms of frequency, the 1A1 allele tended to be overrepresented (0.23 vs. 0.17, $p=0.207$) and the 1A5 allele underrepresented (0.04 vs. 0.08, $p=0.128$) when the infants with severe RSV infection were compared with the controls. Based on similar changes in a critical amino acid (Glu223Lys), the 1A1 and 1A3 alleles were pooled. The population with the 1A1 and/or 1A3 allele was overrepresented in the RSV group (0.28 vs. 0.18, $p=0.023$, OR 1.78, 95% CI 1.1–2.9) (Figure 11).

5.1.3 Genotype analysis

The homozygous genotypes 6A3/6A3 and 6A2/6A2 were overrepresented in the RSV group vs. controls but not significantly so (0.14 vs. 0.06, $p=0.086$ and 0.35 vs. 0.24, $p=0.115$ respectively). The SP-A1 6A/6A2, 6A/6A4 and 6A3/6A4 genotypes were not significantly underrepresented in RSV either (0.01 vs. 0.05, $p=0.130$; 0.0 vs. 0.04, $p=0.074$ and 0.02 vs. 0.06, $p=0.174$, respectively). In contrast, the SP-A2 genotypes revealed a significant frequency difference. The 1A1/1A1 and 1A0/1A3 genotypes showed overrepresentation in the RSV group (0.07 vs. 0.0, $p=0.011$; 0.05 vs. 0.0, $p=0.049$, respectively). The SP-A2 genotype 1A1/1A5 showed a trend towards underrepresentation in the RSV group (0.0 vs. 0.04, $p=0.074$). Neither the RSV nor the control group differed from Hardy–Weinberg equilibrium.
5.1.4 Haplotype analysis

Haplotype analysis revealed the RSV group to have significant underrepresentation of the 6A/1A haplotype (0.01 vs. 0.06, p=0.011, OR 0.17, 95% CI 0.04–0.8) and overrepresentation of the 6A2/1A3 haplotype (0.05 vs. 0.005, OR 10.4, 95% CI 1.3–83.2). The 6A4/1A5 haplotype showed a trend towards underrepresentation (0.04 vs. 0.07, p=0.18).
### Table 4. Distribution of SP-A alleles in cases of severe RSV and control infants

<table>
<thead>
<tr>
<th>Gene/Allele</th>
<th>Severe RSV (n=86)</th>
<th>Controls (n=95)</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SP-A1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6A</td>
<td>1.2</td>
<td>6.3</td>
<td>0.17 (0.04–0.79)</td>
<td>0.01</td>
</tr>
<tr>
<td>6A2</td>
<td>58.1</td>
<td>51.6</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>6A3</td>
<td>34.3</td>
<td>32.6</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>6A4</td>
<td>6.4</td>
<td>9.0</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>6A5</td>
<td>0.0</td>
<td>0.5</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td><strong>SP-A2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A0</td>
<td>52.9</td>
<td>51.1</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>1A</td>
<td>1.2</td>
<td>6.3</td>
<td>0.17 (0.04–0.79)</td>
<td>0.011</td>
</tr>
<tr>
<td>1A1</td>
<td>22.7</td>
<td>17.4</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>1A2</td>
<td>13.4</td>
<td>14.7</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>1A3</td>
<td>5.2</td>
<td>0.5</td>
<td>10.44 (1.3–83.2)</td>
<td>0.006</td>
</tr>
<tr>
<td>1A5</td>
<td>4.1</td>
<td>7.9</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td>0.6</td>
<td>2.1</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval; 6An denotes Sp-A1 alleles; 1Am denotes SP-A2 alleles.

### 5.1.5 Amino acid analysis

Amino acid (aa) analysis showed statistical difference in the aa changes at codons 19 (SP-A1), 91 and 223 (SP-A2). Alanine at aa19 was underrepresented in the RSV group (0.01 vs. 0.07, p=0.007, OR 0.16, 95% CI 0.04–0.7). Likewise, proline was underrepresented in the RSV group at aa91 (0.05 vs. 0.16, p=0.001, OR 0.29, 95% CI 0.14–0.64). Lysine was significantly overrepresented in severe RSV (0.28 vs. 0.18, p=0.023, OR 1.78, 95% CI 1.1–2.9) at codon 223.

### 5.2 SP-D gene variation in RSV

Significant differences in SP-D allele frequencies for amino acid 11 were observed between the 84 infants with severe RSV and the 93 matched controls, the frequency of the allele coding for 11 Met being 0.72 in the RSV group and 0.61 in the control group (p=0.033) (Table 5). The frequency of the homozygous genotype Met/Met for amino acid 11 was increased in the RSV group relative to the controls (0.55 versus 0.37), whereas the heterozygous genotype tended to be less frequent among the RSV cases than in the matched controls (0.35 versus 0.49). No significant differences were observed in allele or genotype frequencies as regards amino acids 160 or 270.
Table 5. Distribution of SP-D alleles in cases of severe RSV and control infants.

<table>
<thead>
<tr>
<th>Gene/Allele</th>
<th>Severe RSV (n=84)</th>
<th>Controls (n=93)</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Met</td>
<td>0.72</td>
<td>0.61</td>
<td>1.63 (1.04–2.54)</td>
<td>0.033</td>
</tr>
<tr>
<td>11 Thr</td>
<td>0.28</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>160 Ala</td>
<td>0.54</td>
<td>0.60</td>
<td>0.78 (0.51–1.19)</td>
<td>0.251</td>
</tr>
<tr>
<td>160 Thr</td>
<td>0.46</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>270 Ser</td>
<td>0.96</td>
<td>0.93</td>
<td>2.20 (0.82–5.86)</td>
<td>0.108</td>
</tr>
<tr>
<td>270 Thr</td>
<td>0.04</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval

Both groups studied were in Hardy–Weinberg equilibrium. Conditional logistic regression analysis was used to study whether the confounding factors, i.e. smoking and number of children in the family, influenced the association between the SP-D genotype and the risk of RSV bronchiolitis. The results of these analyses confirmed the association between the homozygous SP-D genotype 11 Met/Met and the risk of severe RSV infection (p=0.028, OR 2.29, 95% CI 1.09–4.81).

5.3 TLR4 in severe RSV infections

For the study, 312 RSV and 356 control samples were successfully genotyped, out of 345 RSV and 383 control samples. The patients and controls were not matched, but similar dates of birth, gestational ages at birth and places of residence were used when possible. The demographic data is presented in Table 3. The proportion of preterm patients was 10.1%, in the RSV group and 6.0% in the control group.

No differences in allele frequency were observed when comparing the RSV and control groups (89.3 vs. 89.7 for allele A; 10.7 vs. 10.3 for allele G) (see Table 2).

Genotype analysis of the whole material also revealed no statistically significant differences as regards genotype GG when comparing the RSV group vs. the controls (1.9 vs. 2.0). The other genotypes (AA and AG) also did not show any significant differences (80.4 vs. 81.5 and 17.6 vs. 16.6 respectively) (Table 6).
Table 6. Distribution of SNP 299 TLR4 genotypes in cases of severe RSV and control infants

<table>
<thead>
<tr>
<th>Allele/Genotype</th>
<th>Severe RSV (n=312)</th>
<th>Controls (n=356)</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alleles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.89</td>
<td>0.90</td>
<td>0.95 (0.7–1.3)</td>
<td>0.773</td>
</tr>
<tr>
<td>G</td>
<td>0.11</td>
<td>0.10</td>
<td>1.05 (0.7–1.5)</td>
<td>0.773</td>
</tr>
<tr>
<td><strong>Genotypes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.80</td>
<td>0.82</td>
<td>0.94 (0.6–1.4)</td>
<td>0.740</td>
</tr>
<tr>
<td>AG</td>
<td>0.18</td>
<td>0.16</td>
<td>1.08 (0.7–1.6)</td>
<td>0.718</td>
</tr>
<tr>
<td>GG</td>
<td>0.02</td>
<td>0.02</td>
<td>0.98 (0.3–2.9)</td>
<td>0.968</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval

The material was divided according to separate epidemics; the two largest were those of 1999/2000 and 2003/2004. These were separately analysed. The results showed significant differences between the RSV and control groups and also differences between the epidemics. In the 1999/2000 epidemic allele A was overrepresented (92.1 vs. 87.0, p=0.0490) and allele G underrepresented (7.9 vs. 13.0, p=0.0490) in the RSV group. The genotypes AA and AG did not show any significant differences. The GG genotype was underrepresented in the RSV group (0.0 vs. 4.2, p=0.020) (Table 7).

The 2003/2004 epidemic showed underrepresentation of the A allele (86.1 vs. 92.0, p=0.014) and overrepresentation of allele G (13.9 vs. 8.0, p=0.014) in the RSV group compared with the controls, in contrast to the results found in connection with the 1999/2000 epidemic. The genotypes AA and AG showed no significant differences between the RSV group and the controls. Genotype GG was overrepresented in the RSV population (4.0 vs. 0.0, p=0.009) (Table 7).

Conditional logistic regression analysis was used to determine if confounding factors affected the results. This analysis revealed that there was a difference in the distribution of alleles in the different epidemics. Other factors did not affect the results. Hardy–Weinberg testing was carried out and although the total material was in equilibrium, that associated with individual epidemics was not.
Table 7. Distribution of TLR4 alleles and genotypes in cases of severe RSV and controls in two individual epidemics

<table>
<thead>
<tr>
<th>Allele/Genotype</th>
<th>Severe RSV</th>
<th>Controls</th>
<th>OR (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999–2000 epidemic</td>
<td>(n=161)</td>
<td>(n=180)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.92</td>
<td>0.87</td>
<td>1.76 (1.0–3.1)</td>
<td>0.049</td>
</tr>
<tr>
<td>G</td>
<td>0.08</td>
<td>0.13</td>
<td>1.76 (1.0–3.1)</td>
<td>0.049</td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.84</td>
<td>0.78</td>
<td>1.50 (0.8–2.8)</td>
<td>0.195</td>
</tr>
<tr>
<td>AG</td>
<td>0.16</td>
<td>0.18</td>
<td>0.90 (0.5–1.6)</td>
<td>0.675</td>
</tr>
<tr>
<td>GG</td>
<td>0.00</td>
<td>0.04</td>
<td></td>
<td>0.020</td>
</tr>
<tr>
<td>2003–2004 epidemic</td>
<td>(n=151)</td>
<td>(n=176)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.86</td>
<td>0.92</td>
<td>0.5 (0.3–0.9)</td>
<td>0.014</td>
</tr>
<tr>
<td>G</td>
<td>0.14</td>
<td>0.08</td>
<td>0.5 (0.3–0.9)</td>
<td>0.014</td>
</tr>
<tr>
<td>Genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.76</td>
<td>0.84</td>
<td>0.6 (0.3–1.04)</td>
<td>0.072</td>
</tr>
<tr>
<td>AG</td>
<td>0.20</td>
<td>0.16</td>
<td>1.3 (0.7–2.3)</td>
<td>0.350</td>
</tr>
<tr>
<td>GG</td>
<td>0.04</td>
<td>0.00</td>
<td></td>
<td>0.009</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval

5.4 SP-A allelic association in otitis media

In this study, 96 male and 51 female patients with recurrent otitis media (ROM) were recruited. As a control group, 278 consecutive infant samples were collected at the same hospital. The genotyping results showed that no allelic association was present. The homozygous genotype 6A2/6A2 was overrepresented (p=0.05) in the general population compared with the ROM group. In haplotype analysis 6A4/1A5 was overrepresented in the ROM and AOM groups (p=0.003 and p=0.02 respectively). The 6A2/1A0 haplotype was underrepresented in the AOM group compared with controls (p=0.03). When comparing haplotype distributions, both the AOM and ROM groups differed from the controls (p=0.04 in both cases).
6 Discussion

Respiratory syncytial virus bronchiolitis, acute otitis media and recurrent otitis media are multifactorial diseases with both environmental and genetic factors influencing predisposition, severity and recurrence. These factors interact to form a complex disease aetiology. Lower airway disease due to RSV infection and OM as an upper airway disease share some of the risk factors. This complexity and transient symptoms represent a challenge. The purpose of this study was to evaluate selected innate immunity candidate gene polymorphisms as interactive risk factors. We chose to investigate the genes for surfactant proteins A and D and Toll-like receptor 4 on the basis of previous studies that showed promising results with knockout mice, and other laboratory investigations. Here we show that SP-A, SP-D and TLR 4 revealed an association with severe RSV bronchiolitis and SP-A was associated with otitis media.

6.1 Role of SP-A and SP-D in severe RSV infections

Bronchiolitis caused by RSV is the most common lower airway disease in early infancy. Our results show, for the first time, a genetic association between both surfactant proteins A and D and severe RSV virus infection. The study population was racially homogeneous. The SP-A gene locus has been shown to be sufficiently polymorphic for genetic studies in American and Finnish populations (Ramet et al. 2000). SP-D gene frequencies have been shown to vary slightly between different ethnic groups, indicating that caution must be used when comparing the results of studies carried out different ethnic groups (Liu et al. 2003, Leth-Larsen et al. 2005).

Both SP-A and SP-D are part of the innate immunity system and these proteins bind and agglutinate specific lung pathogens in the respiratory tract. The proteins seem to have a very critical role during infancy, when maternal anti-RSV antibodies are at a low level and the respiratory tract is narrow and clogs easily with inflammation. The levels of surfactant proteins A and D have been shown to decrease during infection. The introduction of SP-A and SP-D into the lungs of mice infected with RSV increases clearance of the viral load (LeVine et al. 1999, LeVine et al. 2004).

Several environmental and social factors predisposing infants to severe RSV infections could affect the study results. To ensure that these effects would be minimal, the patients were selected prospectively, on the basis of known risk
factors. By matching, we tried to ensure highly comparable data as regards sex, gestational age at birth, time of birth and hospital district. Factors such as allergies and parental smoking did not reveal any significant differences between the two study groups. However, there was a difference in the number of children in the families of infants with severe RSV versus their controls. Conditional logistic regression analysis showed that this difference had no effect on the association between SP-A and D gene variances and the risk of severe RSV. As a result of good social and medical care in Finland, predisposing factors such as day care and lower social class did not affect the results. Most of the infants were at home for their first nine months to a year of life. In addition, the community and hospital medical services were easily available at a low cost in all cases.

Table 8. A) SP-A1 allele amino acid changes.

<table>
<thead>
<tr>
<th>SP-A1</th>
<th>aa19</th>
<th>aa 50</th>
<th>aa 62</th>
<th>aa 133</th>
<th>aa 219</th>
</tr>
</thead>
<tbody>
<tr>
<td>6A</td>
<td>Ala</td>
<td>Leu</td>
<td>NC</td>
<td>NC</td>
<td>Arg</td>
</tr>
<tr>
<td>6A2</td>
<td>Val</td>
<td>Val</td>
<td>NC</td>
<td>NC</td>
<td>Arg</td>
</tr>
<tr>
<td>6A3</td>
<td>Val</td>
<td>Leu</td>
<td>NC</td>
<td>NC</td>
<td>Arg</td>
</tr>
<tr>
<td>6A4</td>
<td>Val</td>
<td>Leu</td>
<td>NC</td>
<td>NC</td>
<td>Trp</td>
</tr>
<tr>
<td>6A5</td>
<td>Ala</td>
<td>Leu</td>
<td>NC</td>
<td>NC</td>
<td>Trp</td>
</tr>
</tbody>
</table>

aa, amino acid; Ala, alanine; Val, valine; Leu, leucine; Arg, arginine; Trp, tryptophan; Asn, asparagine; Thr, threonine; Pro, proline; Lys, lysine; Gln, glutamine; NC, no change in amino acid.

Table 8. B) SP-A2 allele amino acid changes.

<table>
<thead>
<tr>
<th>SP-A2</th>
<th>aa 9</th>
<th>aa 91</th>
<th>aa 140</th>
<th>aa 223</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>Thr</td>
<td>Pro</td>
<td>NC</td>
<td>Gln</td>
</tr>
<tr>
<td>1A0</td>
<td>Asn</td>
<td>Ala</td>
<td>NC</td>
<td>Gln</td>
</tr>
<tr>
<td>1A1</td>
<td>Thr</td>
<td>Ala</td>
<td>NC</td>
<td>Lys</td>
</tr>
<tr>
<td>1A2</td>
<td>Thr</td>
<td>Ala</td>
<td>NC</td>
<td>Gln</td>
</tr>
<tr>
<td>1A3</td>
<td>Asn</td>
<td>Ala</td>
<td>NC</td>
<td>Lys</td>
</tr>
<tr>
<td>1A5</td>
<td>Thr</td>
<td>Pro</td>
<td>NC</td>
<td>Gln</td>
</tr>
</tbody>
</table>

aa, amino acid; Ala, alanine; Val, valine; Leu, leucine; Arg, arginine; Trp, tryptophan; Asn, asparagine; Thr, threonine; Pro, proline; Lys, lysine; Gln, glutamine; NC, no change in amino acid.

The risk of severe RSV was associated with certain SP-A2 alleles and SP-A haplotypes. The study showed that specific SP-A alleles, genotypes and haplotypes appeared to predispose infants to or protect them from severe RSV bronchiolitis. Earlier it has been shown that the SP-A1 and SP-A2 genes are in strong linkage disequilibrium, which means that alleles of SP-A1 are linked to the same chromosome as corresponding SP-A2 alleles. Alleles 6A and 1A had similar
risk associations in connection with severe RSV due to close proximity of these two genes. Logically, an association between the haplotype 6A/1A and severe RSV was found. Variation in SP-D was also shown to be associated with severe RSV. The allele Met11 was significantly overrepresented in severe RSV as well as the genotypes Met11Met and Met11Thr. The Thr11 allele and the Thr/Thr genotype showed no association with risk. The 160 and 270 polymorphisms were not associated with severe RSV. We propose, on the basis of our results, that both specific SP-A and SP-D alleles, genotypes and haplotypes are involved in susceptibility to severe RSV infection during infancy when acquired immunity is weak. The main alleles and haplotypes of SP-A in the general population were not associated with predisposition to severe RSV infections. In contrast, the major SP-D allele 11Met is a risk allele with relatively low penetrance.

A non-synonymous single nucleotide polymorphism causes a change in amino acid. In SP-A amino acid 19 (alanine) in the N-terminal domain of SP-A1 protein and amino acid 91 (proline) in the collagen-like domain of the SP-A2 protein were both underrepresented in severe RSV. As regards SP-D, a significant change occurs at amino acid 11. At this location, the two amino acids Met and Thr vary according to the SNP. The mature SP-D molecules have different hydrophobic properties depending on amino acid composition. In the case of Met11Thr, the location is near two conserved amino acids Cys^{15} and Cys^{20} that are actively involved in the inter-chain disulphide bonding of SP-D trimeric subunits. The two disulphide bonds are involved in stabilizing the multimeric structure of SP-D. It has been proposed that the antiviral properties of the protein increase with increased multimerisation of SP-D. The functional consequences of these changes in the amino acid sequence are still unknown.

Previous laboratory and animal research has shown the essential roles of SP-A and SP-D in the pathogenesis of RSV. Surfactant protein knockout mice more often have severe RSV infections than wild-type mice. SP-A has been shown to enhance the uptake of RSV by neutrophils and alveolar macrophages and it also increases opsonization. During infant RSV infections, reduced levels of SP-A have been reported (Kerr & Paton 1999, Barr et al. 2000).

A pilot study with exogenous surfactant has been conducted in which patients with severe RSV infection were treated. SP-A and SP-D could be considered for the treatment of severe airway diseases, but currently SP-A and SP-D are not included in commercial surfactant. Allele-specific surfactant proteins could be considered as a prospective means to enhance the effectiveness and properties of both SP-A and SP-D in the treatment of severe RSV infections.
The present studies provide the first evidence of two surfactant gene loci associated with severe RSV infections. Further research with a large cohort is required to assess the complex relationships between the genetic and non-genetic factors involved in susceptibility to severe RSV bronchiolitis and to evaluate the rare alleles and haplotypes. It is also necessary to see if the genotypes of SP-A and SP-D have effects on protein concentrations during RSV infection and if the antiviral properties change. Those infants genetically and clinically predisposed to RSV could be considered for advanced prophylactic RSV treatment.

6.2 TLR4 receptor 299 polymorphism in RSV infection

The association between TLR4 and severe RSV infections was studied in a Finnish population between 1999 and 2004. The patients and controls were from Southern and Northern Ostrobothnia. During this period, several RSV epidemics occurred in recurring manners typical of the areas.

Our hypothesis was that an association would be found between severe RSV bronchiolitis and TLR4 gene variation. The allele and genotype data showed no significant associations between TLR4 and severe RSV bronchiolitis when the whole patient and control populations were examined. No matching of the material was carried out for this study, in order to maximize the number of cases. Paulus et al. studied a slightly smaller sized population and found no significant allele or genotype associations between RSV cases and controls (Paulus et al. 2007).

The patient and control populations were split up according to separate epidemics and analysed separately. The two largest epidemics, in 1999–2000 and in 2003–2004 were selected for further study and analysis. When analysed, the data showed a small but significant difference in both epidemics between RSV and TLR4 allele distribution and the GG genotype. The more unexpected finding was that the two largest epidemics showed opposite results (Table 7). In the 2003–2004 epidemic the G allele and GG genotype were risk factors. This risk was also earlier shown by Tal et al. (Tal et al. 2004). In contrast, the 1999–2000 epidemic results indicated that the G allele and the GG genotype were protective as regards severe RSV. The number of studied cases and controls were essentially equal in both epidemics.

The present results show evidence of a variable association between severe RSV infections and TLR4, implying that the antigenic diversity of RSV may influence the ability of individual TLR4 allelic variants to initiate a proper innate
response. Severe RSV bronchiolitis is associated with recurrent respiratory tract infections and wheezing, which, according to the present findings, may be a consequence of the interaction between the genotype of the host and the constitution of the pathogen.

When examining the entire study population, Hardy–Weinberg equilibrium (HWE) was observed, but in studies of the individual epidemics, this was not the case. However, natural populations in HWE are uncommon. Usually one or several of the assumptions regarding HWE are violated. In our study, population size is one criterion that did not meet the HWE prerequisites (infinite/large size). The number of cases with the GG genotype was small, but they were in opposite groups in the two epidemics, with no cases in the other group. A larger amount of study material would be needed to prove the results with confidence.

Previous data has shown that TLR4 gene variation is associated with a higher risk of premature birth; the risk of severe infection is also elevated (Lorenz et al. 2002a, Lorenz et al. 2002b). These infections have shown similarities to RSV infection in that the TLR4 G allele and genotypes AG and GG have been overrepresented in patients (Tal et al. 2004). Amino acid 299 changes along with the SNP. It has been proposed that the function of the receptor is dependent on a single amino acid in the extracellular domain of TLR4 (Arbour et al. 2000).

Prior studies have revealed that RSV exhibits genetic variation and there are two main strains (A and B) (Sullender 2000). Both strains show genetic variability in the G glycoprotein. Of the two strains, group A viruses have shown more variability than the B group (Parveen et al. 2006). It has been suggested that the variants could be the cause of RSV re-infection and that they play a role in allowing RSV to evade immune responses (Sullender 2000). The dominating RSV subgroup (A/B) has been shown to have altered with two-year cycles in Finland between 1981–1990 (Waris 1991), but data on strains for the time period of the present study is not available. In Sweden, however, this kind of data on virus genotype is available for 1990 to 2004, with small gaps. This data shows some diversity in the subgroup of A genotypes but the dominating genotype was GA5 (Rafiefard et al. 2004).

According to the present results, genetic variability of TLR4 influences immune responses and the susceptibility to severe RSV bronchiolitis. Based on the results, it is likely that susceptibility to severe RVS bronchiolitis could also depend on genetic variation of the host, the most likely site of variation being the binding site of the microbial ligand.
Further studies are required to confirm the hypothesis that the variable genotypes of the TLR4 extracellular domain could influence susceptibility to different RSV subtypes. These latest findings on Asp299Gly-induced conformational changes in the extracellular domain of TLR4 are stimulating further functional studies using different TLR4 genotypes and different RSV strains/genotypes. Resistance against serious RSV infections may offer an evolutionary selection advantage, as the antigenic diversity of RSV favours heterogeneity in the extracellular domain of TLR4, driving towards allelic heterogeneity of the gene encoding TLR4.

6.3 Otitis media and SP-A gene variation

Otitis media is one of the most common childhood infections. The infection is affected by many environmental and epidemiological factors. These factors must be considered when analyzing results. The pathogen causing OM varies and there are allelic variants as well. Effectiveness in binding or launching an immune attack against the causative microbes could differ according to the SP-A variant. *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis* are the most common bacterial pathogens in OM (Kilpi et al. 2001). RSV infections are associated with OM; several studies have shown RSV to be present in middle ear effusion together with bacteria. RSV has been shown to be one of the most common upper respiratory viral pathogens (Sagai et al. 2004, Bulut et al. 2007).

The present results showed that acute and recurrent otitis media were associated with SP-A genotypes and haplotypes. The 6A2/6A2 homozygous genotype was shown to be overrepresented in the general population compared with the ROM group. The haplotype 6A2/1A0 showed similar overrepresentation. The only predisposing haplotype was 6A4/1A5; it showed overrepresentation in both the ROM and AOM groups. Pettigrew et al. reported opposite results in their study. They found that the 6A4/1A5 haplotype was underrepresented in children with AOM and ROM. Their material consisted of a cohort of patients at an elevated risk of asthma and the material was of variable ethnic background (Pettigrew et al. 2006).

Surfactant protein A is expressed in both the upper and lower airways. It has been shown to be associated with RSV and the results show that different SP-A genotypes and haplotypes are associated with OM and RSV bronchiolitis. The haplotypes 6A2/* and 6A4/1A5 appear to be distributed in opposing ways in
these two conditions. The RSV material was investigated by using conditional logistic regression analysis. The results showed that the number of otitis media infections had no effect on the association between the SP-A alleles and severe RSV bronchiolitis (data not shown). These results indicate possible differences in microbe-selectivity of individual SP-A genotypes or differences in the expression of various SP-A genotypes within the upper and the lower airways.

Further studies with larger material are necessary to confirm the findings. In future studies, the causative microbes of OM should be included as factors in statistical analysis. The presence of RS virus in middle ear effusion should also be investigated.
7 Summary and conclusions

The results of the present studies suggest that collectins (surfactant protein A and D) and receptors (Toll-like receptor 4) of the innate immune system are involved in defence against infection in the upper and lower respiratory tracts. The diseases studied were severe Respiratory Syncytial Virus bronchiolitis and otitis media in infants and young children. Both diseases have a substantial impact on healthcare resources.

We showed, for the first time, associations of several gene variations with severe RSV bronchiolitis and one in connection with otitis media. RSV bronchiolitis and otitis media had no common genetic associations with collectins, even though otitis media is associated with RSV. Interestingly RSV showed a possible annual variation in association with TLR4. This hypothesis requires further study.

Future challenges are to show how these associations are related to conformational changes and binding properties of the innate proteins. Again, a change in surface structure of the microbe may influence its ligation to a receptor and ability to initiate proper immune response. It is to be hoped that progress in understanding innate immunity and the mechanisms involved will help improve treatment strategies and develop new drugs and vaccines.

The four papers in this study suggest the following conclusions:

1. Certain surfactant protein A alleles, genotypes and haplotypes are associated with severe RSV infection in infants.
2. Susceptibility to severe RSV infection was found in association with the surfactant protein D allele Met11 and genotypes Met11Met and Met11Thr.
3. The TLR4 SNP 299 showed no association with severe RSV bronchiolitis when investigated in the material as a whole. However, there was an association when annual epidemics were analysed separately. Opposite results were found in two different epidemics, suggesting possible annual RSV strain differences.
4. Recurrent and acute otitis media was shown for the first time to be associated with certain SP-A genotypes and haplotypes.
References


Grosfeld H, Hill MG & Collins PL (1995) RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. J Virol 69: 5677–5686.


Johnson PR & Collins PL (1989) The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (RSV) of antigenic subgroups A and B: sequence conservation and divergence within RSV genomic RNA. J Gen Virol 70 (Pt 6): 1539–1547.


Original papers

This thesis is based on the following articles, which are referred to in the text by their Roman numerals:


The original articles are reprinted with permission from (I) The Infectious Diseases Society of America/University of Chicago Press, (II) The International Pediatric Research Foundation, Inc. and (IV) Elsevier Inc.

Original publications are not included in the electronic version of the dissertation.


1005. Inmaninen, Sari (2009) Fall accidents and exercise among a very old home-dwelling population.


1010. Löfgren, Johan (2009) Genetic polymorphisms in collectins and Toll-like receptor 4 as factors influencing susceptibility to severe RSV infections and otitis media.

Johan Löfgren

GENETIC POLYMORPHISMS IN COLLECTINS AND TOLL-LIKE RECEPTOR 4 AS FACTORS INFLUENCING SUSCEPTIBILITY TO SEVERE RSV INFECTIONS AND OTITIS MEDIA