CANDIDATE GENES OTHER THAN THE CFTR GENE AS POSSIBLE MODIFIERS OF PULMONARY DISEASE SEVERITY IN CYSTIC FIBROSIS

BY

DESPINA DAISY FRANGOLIAS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

THE FACULTY OF GRADUATE STUDIES

(EXPERIMENTAL MEDICINE)

THE UNIVERSITY OF BRITISH COLUMBIA
(VANCOUVER)

FEBRUARY 2008

ABSTRACT

Cystic fibrosis (CF) is a single gene Mendelian disorder characterized by pulmonary disease and pancreatic insufficiency. Pulmonary disease is the major cause of death in CF patients. Although some cystic fibrosis transmembrane conductance regulator (CFTR) genotypes are associated with less severe disease, patients possessing the same genotype show great variation in pulmonary disease severity and progression. Genes involved in modulating the inflammatory response and genes increasing susceptibility to infection are proposed as modifiers of pulmonary disease severity. Polymorphisms selected for based on evidence that they affect the function of the gene and prevalence of the putative risk allele: 1) antiprotease gene alpha-1-antitrypsin (α_1-AT), 2) innate immunity genes: mannose binding lectin (MBL2) (promoter [G→C] at -221 and codon 52 (Arg52Cys, D allele), 54 (Gly54Asp, B allele), and 57 (Gly57Glu, C allele), and pulmonary surfactant genes SPA-1 (Arg219Trp), SPA-2 (Thr9Asn, Lys223Gln) and SPD (Thr11Met), 3) antioxidant genes GSTM1 and T1 (gene deletion polymorphisms), GSTP1 (Ile105Val) and GCLC repeats, 4) mucin genes (MUC2 and MUC5B). Pulmonary disease progression and survival in patients with chronic *Burkholderia cepacia* complex (BCC) infection were also investigated controlling for genomovar and RAPD type of the organism. BCC infection was associated with more severe pulmonary disease progression and worse survival. α_1-AT genotype was not a major contributor to variability of pulmonary disease severity, but the results suggest that α_1-AT plasma levels during pulmonary infections may be affected by poor nutritional status. We showed similar pulmonary disease progression and MBL2 genotype. Contrary to the previous literature, wild-type MBL2 genotype was associated with steeper decline in pulmonary disease over time following chronic infection with BCC, but genotype was not associated with increased susceptibility to BCC infection. We showed inconsistant results for the pulmonary surfactant gene polymorphisms, GSTM1, T1 and GSTP1 polymorphisms, and number of repeats for GCLC and MUC5B depending on the phenotype investigated. We conclude that some of the variability in pulmonary disease severity and progression in CF is explained by polymorphisms in secondary genes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIDS	Acquired immune deficiency syndrome
ANOVA	Univariate analysis of variance
ASP PCR	Allele-specific oligonucleotide PCR
α_1-AT	Alpha-1-antitrypsin
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
ASLF	Airway surface lining fluid
ATS	American Thoracic Society
BAL fluid	Bronchoalveolar lavage fluid
BCC	Burkholderia cepacia complex
BMI	Body mass index measured as Weight/Height^2 (units: kg/m^2)
CEPH	Centre d'Etude du Polymorphisme Humain
CHO	Carbohydrates
COPD	Chronic obstructive lung disease
cAMP	Cyclic adenosine monophosphate
CF	Cystic fibrosis
CFTR	CF transmembrane conductance regulator
CI	Confidence interval
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphates
EM	Expectation-maximization
ER	Endoplasmic reticulum
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
GEN	Genomovar
GSH	Reduced glutathione
GSSG	Glutathione disulfide
GCLC	Glutamate cysteine ligase catalytic subunit
GST	Glutathione S-transferase
H. influenza	Haemophilus influenza

HWE	Hardy Weinberg equilibrium
Ig	Immunoglobulin
IGEPAL	Octylphenyl-polyethylene glycol
IV	Intravenous
IFN-γ	Interferon gamma
IL	Interleukin
kb	Kilobase pair
kDa	Kilo Dalton
LD	Linkage disequilibrium
LPS	Lipopolysacharide
MASP	MBL2-associated protein
MHC	Major histocompatibility complex
MBL2	Mannose binding lectin
mRNA	Messenger ribonucleic acid
ml	Milliliters
µl	Microliter
MUC	Mucin (eg, MUC2 mucin 2 gene)
NBF	Nucleotide binding fold
NADPH	Nicotinamide adenine dinucleotide phosphate
NE	neutrophil elastase
PAMP	pathogen-associated molecular patterns
%predFEV_1	% of predicted forced expiratory volume in 1 second
OD	Optical density
PA infection status	Pseudomonas aeruginosa infection status
PHASE	Inferred haplotype probabilities
PCR	Polymerase chain reaction
PSS	Pancreatic sufficiency status
P. aeruginosa	Pseudomonas aeruginosa
RDS	Respiratory distress syndrome
ROS	Reactive oxygen species
S-K score	Schwachman-Kulczycki score
SSP PCR	sequence-specific priming PCR

SEM	Standard error of the mean
Staph. aureus	Staphylococcus aureus
SDM PCR	Site Directed Mutagenesis Polymerase Chain Reaction
SP	Pulmonary surfactant
SNP	Single nucleotide polymorphism
RAPD	Random amplification of polymorphic DNA
RFLP	Restriction fragment length polymorphism
RPM	Revolutions per minute
RNA	Ribonucleic acid
TGF-β1	Transforming growth factor-β1
TMD	Transmembrane domains
TNFα	Tumor necrosis factor alpha
UV	Ultraviolet light
VNTR	Variable number tandem repeat
yr	Year
≠	Not equal (different)

ACKNOWLEDGEMENTS

This study was funded by research grants by the Canadian Cysic Fibrosis Foundation, Bayer/Red Cross and B.C. Lung Foundation.

I acknowledge and thank the Canadian CF Foundation and Michael Smith Foundation for their financial support.

A very special thank-you to:
Dr. Jiang Ruan and Ms. Tracey Weir for their technical support with the laboratory work.
Ms. Eugenia Wu for her statistical help and extensive assistance in the development of the pulmonary infections statistical models for analysis.
Mr. Ryan Woods and Ms. Kelly Burkett for their initial statistical help on the project.
Dr. Speert's Laboratory and Ms. Debora Henry for the BCC genotyping data.

I would like to to express my sincere gratitude and appreciation to the staff and directors of the CF clinics involved in the study and the CF patients who volunteered their time and consented to participate in this study.

I extend my sincere appreciation to my committee members (Drs. Joel Singer, Keith Walley, John Hill) and supervisor Dr. Peter Pare for their help and support. A special thank-you to Drs. Andy Sandford and Pearce Wilcox (committee members) and Drs. George Davidson, David Speert and Yves Berthiume for their guidance.

I would also like to acknowledge and thank my husband's company Nortrak who lent me and serviced a laptop for the first year of my clinical data collection period. And I should not forget my husband Brian Charles (P.Eng.) for coming along for the ride and 'tolerating' my absurd work hours and me. Thank-you for sticking by me.

DEDICATION

This project is dedictated to the CF community and their commitment to research and advancement of quality of life for the patient and their families. There are certain CF patients who have participated in studies I have been involved in that have impressed me with their spirit. This thesis is dedicated to them and their families: Lloyd, Collette and Bernadette, Helena, Bob and Sheldon.

A special dedication to Andrea Herzer and her family. I met Andrea when she transitioned to the Adult clinic. I got to know her very well through her participation in numerous research projects. She was a delite to be around. I would sometimes pick her up from her home for scheduled research project testing and had the opportunity to meet her mother as well. I last saw her as she was to leave for the U.S. for her lung transplant and I was expecting my first baby. We sat in her hospital room and talked about our upcoming life changing events and remember saying our good byes with a hug telling each other to take care of ourselves and promised we see each other after it all. I had Stevie after some complications, but ultimately both of us were fine. Andrea's transplant was not successful; she was too young when she lost her battle with CF.

Lastly, I dedicate this work to my children (Steven and Thomas) who I acquired during this process and who at this time may have no clue what mommy has been doing on this little black toy (laptop) which they seem to have made their mission to take possession.

CO-AUTHORSHIP STATEMENT

I, Despina Frangolias, performed the following tasks on the 2003 publication presented in chapter 3 entitled**:**

'Frangolias D.D., Ruan J., Wilcox P.G., Berthiaume Y., Davidson G., Hennessey R., Corey M., Tullis E., Zielenski J., Wilson W.M., Freitag A., Sandford A.. Alpha-1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *American Journal of Respiratory Cell and Molecular Biology*. 29:390-396, 2003.'

.

1. Identification of gene to investigate was collaboration between Drs. Peter Pare, Andrew Sandford and myself, Despina Frangolias.
2. Designed the study and composed and submitted ethics forms for study approval
3. Recruited Vancouver site subjects and recruited additional centers across Canada for the study, except for Toronto (recruited by my supervisor Dr. Peter Pare).
4. Defined the study variables, created the clinical data collection sheets and collected the clinical data for Vancouver centers and entered hardcopy data sent from centers Hamilton and Montreal.
5. Performed approximatey 90% of the laboratory experiments (i.e., extraction of DNA and genotyping for the alpha-1-antitrypsin gene polymorphisms).
6. Performed 50% of statistical analyses (longitudinal data analyses and analysis of substudy data on alpha-1-antitrypsin) and Mr. Ryan Woods performed the survival analyses and confirmed my statistical analyses findings.
7. Wrote the manuscript, which was then passed on to the co-authors for their review prior to submission.

I have written and performed all the work for Chapters 4, 5 and 6 with the advice and consulatation of my supervisor Dr. Peter Pare and committee members Drs. Andrew Sandford and Dr. Wilcox.

CHAPTER 1: OVERVIEW AND OBJECTIVES

1.0 INTRODUCTION

Cystic fibrosis (CF) affects approximately 1 in 2500 live births in Canada and the United States (1, 2). Cystic fibrosis is a systemic disorder transmitted as an autosomal recessive trait (1). In the late 1980s the gene responsible for the CF phenotype was identified (3-6). The CF gene codes for a cellular membrane protein, the CF transmembrane conductance regulator (CFTR), which forms a chloride channel and also regulates other channel proteins. The CF gene is expressed in the epithelial cells of many organs including the sweat glands, pancreas, lungs, gastrointestinal, and reproductive tract. The mutation disrupts exocrine function by causing abnormal regulation of epithelial ion transport. Chronic and recurrent pulmonary infections and digestive disorders, infertility, and 'salty sweat' characterize the cystic fibrosis phenotype. The classic CF phenotype includes pulmonary disease, pancreatic exocrine insufficiency and abnormal sweat gland function. The onset of pulmonary disease in CF patients is associated with colonization of the airways, most commonly by the pathogen *Pseudomonas aeruginosa* (7, 8). This provokes a vigorous local inflammatory response that prevents the spread of the infection beyond the lung, however by doing so it likely contributes to lung destruction(7-9). Precisely how the CF electrolyte transport defect leads to persistent lung infection and inflammation is still not well understood.

With advances in treatment, patients who have CF are living beyond their adolescent years into adulthood. The median age of survival has increased to well into the third decade. The large variability in disease progression and degree of organ involvement has become even more evident with the increased survival. Only a portion of this variability can be explained by the individual's CFTR genotype. Although some CFTR genotypes are associated with less severe disease, patients possessing the same genotype show great variation in disease severity and progression (1). Over 900 mutations have been identified in the CFTR gene, but ΔF508 is the most common(10). Environment plays a significant role in explaining some of the variability in disease severity and progression. Environmental influences include the availability of therapeutic modalities as well as the psychosocial issues such as the patient's compliance to therapy and the patient's coping mechanisms with their illness. However, other genes may play a significant role in explaining this wide range of clinical outcomes and disease progression. Variants in other secondary genes may act in concert to positively

or negatively contribute to disease severity and progression in CF. An understanding of the complex interaction of other genes with the CFTR gene may help us gain a better understanding of the cystic fibrosis phenotype and develop specific therapies. This chapter will focus on a brief historical review of the discovery of the cystic fibrosis gene, and a description of the gene's structure, function, and the effect of the CFTR defect on the main exocrine organ it affects, the lung. A review of the literature on possible modifier genes in CF will be explored and this section will conclude with the study purpose, rationale, and main and secondary hypotheses.

1.2 THE CF GENE

1.2.1 Discovery of the Cystic Fibrosis Gene

From early on it was known from population studies that cystic fibrosis was an autosomal recessive disorder and that it affected a number of organs, including the lung airways, pancreas and sweat glands. Knowledge that this disorder was recessive came from the observation that offspring who had CF were born from apparently unaffected parents. The basic defect is associated with decreased chloride ion conductance across the apical membrane of epithelial cells and the defect persisted in cultured cells derived from numerous epithelial tissues suggesting that the CF gene was expressed in these cells. It was also known that the defect lay in epithelial cells due to the observation of increased chloride in sweat, exocrine pancreatic obstruction leading to pancreatic insufficiency and intestinal obstruction (meconium ileus in newborns). Unlike other genetic disorders where the amino acid sequence of the protein that was defective was known (e.g. haemoglobinopathies, phenylketonuria and alpha-1-antitrypsin), in CF the basic protein defect resulting in the disorder was still unknown until the gene was discovered using positional cloning. From as early as 1986 it was hypothesized that the defect was probably due to a failure of an outwardly rectifying anion channel. Research in these early years focused on candidate genes. These included a sodium inhibitory factor identified from CF saliva (11), ciliary dyskinesia factor (12, 13), CF antigen on chromosome one and HLA (14). These studies failed to show any positive associations.

In the late 1980s linkage analysis was used to search for the CF gene. The first step to identify the primary defect in CF was to establish the chromosomal localization of the

disease locus. Development and refinement of polymerase chain reaction during this era facilitated this search; restriction fragment length polymorphism (RFLP) markers were used to screen DNA that was obtained from families with 2 or more affected individuals to. The relatively straightforward CF phenotype, knowledge that it was a recessive disorder (since affected offspring were born to normal parents), and the large number of CF families facilitated this work. Early work did not yield candidates.

In the mid 1980's Eiberg and associates reported linkage between the CF locus and a polymorphic locus which controlled activity of the serum aryl esterase paraoxonase (PON), however although the location of the CF gene was narrowed down to one third of the genome, the chromosomal location of PON was not known (15). The location of the CF gene was narrowed down to one percent of the genome by Tsui and associates(3) who showed that a DNA marker called D0CRI-917 was genetically linked in a set of 35 families and was also linked to the PON locus, which by independent evidence had been linked to the CF locus. The search for the CF gene had been narrowed down to 30 million base pairs; the estimates of the genetic distances were calculated to be 5 centimorgans between the DNA marker and PON and 15 centimorgans between the DNA marker and the CF locus(3). In 1985 linkage was established between CF and polymorphic markers (i.e., *D7S15*, *MET*, *D7S8*, *COL1A2*), all known to be located between bands 22 and 31 on the long arm of chromosome 7 (3-6, 16-21). This still represented a very large region and two approaches were used to pinpoint the CF gene. One approach was to proceed to isolate the CF gene by directly looking for sequences that were preferentially expressed by epithelial cells and the second approach was to define the physical map of the region. Genetic and physical mapping studies showed the order and distance of the four markers that had been identified (i.e., *MET-D7S340-D7S122-D7S8* with distances between them of 500, 10 and 980 kb, respectively) (18).

Allelic and haplotype associations were shown between the CF locus and closely linked DNA markers. Beaudet and associates investigated recombination events between CF and these linked markers in 100 CF families with 2 or more affected children, the order between these markers and the CF gene was established and flanking markers were *MET* and *pj3.11*(*D7S8*) (22). Pancreatic insufficiency is a common pancreatic abnormality shared by many CF patients but not all. Kerem and associates used two clinical subgroups of CF

patients based on pancreatic sufficiency status (i.e., pancreatic insufficient (PI) and pancreatic sufficient (PS)) and showed that pancreatic dysfunction could be explained by different mutations on the basis of family studies and haplotype data (23, 24). PI CF patients (85%) were more homogeneous than PS patients (15%); however these studies were based on linked DNA markers whose exact relation to CF was not entirely certain (24). Chromosome walking and jumping were used to link DNA molecules at great distances from one another (400-500kb). This process was facilitated by the development of pulse-field gel electrophoresis and the discovery of additional restriction enzymes. A restriction map of the region was constructed which localized the CF locus with a number of closely linked markers. Each one of these closely linked markers was used as a starting point for a series of chromosomal walking and jumping experiments in order to clone and sequence large DNA regions (19). These cloned regions (which were overlapping) were used to look for candidate coding sequences that would be conserved in other animal species (rodent, bovine, mouse, and chicken). They found one such sequence that was conserved but it was fairly small (only 113bp long) and consequently represented only a small section of the gene. Ultimately this coding sequence was identified as exon 1 of the CF gene. By successive screening with cDNA libraries generated from a number of cell lines (i.e., a cell carcinoma line, normal and CF sweat gland cells, pancreas and adult lung), they were able to isolate an additional 18 clones. Ultimately Dr. Tsui and his collaborators were able to deduce the coding region of the CF locus from the overlapping DNA clones (3). Together these clones spanned 6.1 kb and encoded a protein which was 1480 amino acids long (3).

To visualize the transcript of the CF gene, mRNA samples were prepared from various tissues. They found high mRNA levels in tissues which were affected by CF (e.g. nasal polyps, pancreas, lung, sweat glands, colon, placenta, liver, parotid gland) and no detectable mRNA expression in tissues not affected by the disease, such as the brain and adrenal gland. They consequently surmised that the expression of the CF gene occurred in many of the tissues examined, with higher levels in those tissues severely affected by CF (20).

RFLPs associated with the CF locus were used to establish the relation to CF of the DNA segments isolated from the chromosome walking and jumping experiments and family studies. Families where cross-over events between CF and other flanking DNA markers had previously been discovered were used (24). The recombination breakpoints were localized

in 2 families that were informative for the DNA markers tested and the CF gene localized to a region between 2 markers (i.e., KM19 and D7S424) (25). Linkage disequilibrium was detected for markers that were close to the CF gene (markers within 300kb interval) and not for the markers further away. A similar conclusion was made by the investigators when they utilized haplotype analysis (24). For these analyses, investigators used PI patients because they appeared to be more homogeneous genetically. Most of these patients would later be shown to be carriers of the most common and severe mutation in CF, deltaF508 (ΔF508).

Comparisons between cDNA clone sequences derived from CF and unaffected individuals showed the most striking difference was a 3 base pair deletion, which resulted in the loss of a phenyalanine residue (ΔF508 mutation). Sixty eight percent of CF chromosomes in the general patient population had this 3 base pair deletion. In contrast none of the unaffected individuals had this deletion. Extended haplotypes based on 23 DNA markers were generated for the CF and wild-type (i.e., unaffected) chromosomes in the collection of families previously used for linkage analysis. Five major groups of wild-type and CF haplotypes were identified within the region of the CF gene, with one of them associating with the most frequent CFTR mutation, the ΔF508 mutation.

The protein structure was next characterized and from the cDNA clones the characteristic features of this protein were described to be two repeat motifs which included a transmembrane spanning domain and sequence resembling ATP-binding folds. From these characteristics and similarities with other membrane associated proteins it was predicted that the CF gene product was likely involved in the transport of ions across a membrane, although it was unclear how CFTR was involved in the regulation of ion conductance across the apical membrane of epithelial cells. Validation that the gene identified was the correct gene responsible for the disease came with the *in vitro* work showing the association of mutant CFTR and defective chloride transport(26-28).

1.2.2 Structure, function and localization of the CFTR gene

The CFTR gene is a 250kb gene and has 27 exons that are transcribed into a 6.5kb mRNA, which is translated to a 1480 amino acid product. The protein is a member of the ATP binding cassette family. The protein is composed of two repeat motifs, each with a 6 transmembrane (loops) domains (TMD) and an intracellular nucleotide binding fold (NBF)

separated by an intracellular hydrophilic regulatory domain (R). The protein is a chloride channel activated by cyclic AMP mediated protein kinase A phosphorylation of the R domain and ATP hydrolysis by the NBFs. The mature CFTR protein forms an apical epithelial channel. The two transmembrane domains form the pore of the channel.

The regulatory domain and the NBFs are situated intracellularly. The regulatory domain, which is encoded by exon 13, contains phosphorylation sites for protein kinase A and for protein kinase C. Phosphorylation of the regulatory domain results in a conformational change in this domain which moves it away from the pore of the channel and thus allows the flow of chloride through the channel (i.e., phosphorylation of the regulatory domain results in the opening of the channel). ATP hydrolysis occurs on the NBFs; hydrolysis of NBF-1 opens the CFTR channel and ATP hydrolysis of by NBF-2 closes the CFTR channel.

The activity of the channel is regulated by intracellular cAMP levels. Epithelial cells contain receptors and respond to beta-adrenergic agonists, prostaglandins, adenosine and vasoactive intestinal peptide. Cellular calcium concentrations are controlled by bradykinin, substance P, leukotrienes and nucleotides such as ATP. These extracellular signals regulate the two main second messengers: cAMP and calcium. Increase in cAMP activates cAMP dependent protein kinase, which phosphorylates and thus activates CFTR chloride channels, basolateral Na/K/Cl cotransporters and likely some of the basolateral K channels. In contrast to the cAMP mediated chloride secretion, intracellular events that mediate chloride secretion stimulated by increased cytosolic calcium are less clear. The increase in membrane anion conductance by the calcium activated chloride channel is hypothesized to be mediated by a calcium/calmodulin dependent protein kinase II (2, 29). One possible mechanism described for the channel in intestinal epithelial cells is that increases in cytosolic calcium are stimulated by acetylcholine through the M3 muscarinic receptor which increases conductance of a basolateral membrane potassium channel (distinct from the cAMP-regulated potassium channel) which results in membrane hyperpolarization and creates an increased driving force for chloride exit through these calcium activated (alternative) chloride channels (2, 30).

It is also thought that CFTR functions in ATP efflux and concomitant regulation of the calcium activated chloride channel and as a cAMP dependent negative regulator of the

epithelial Na channel. Loss of this regulatory function may account for the abnormal Na transport characteristic of tissues which are affected by CF; however it is not clear if this is a consequence of direct CFTR interaction with the epithelial sodium channel or a secondary effect mediated by other regulatory proteins. The promoter region of the CFTR gene is 3.5kb and is a GC rich region with a major transcription start site and multiple minor transcription start sites. A number of SP-1 and SP-2 binding sites have been identified as well as sequences for cAMP and glucocorticoid response elements.

1.2.3 Classification of CFTR mutations

Close to 1000 mutations in the CF gene have been identified (10, 31). Defective CFTR can result in defective protein production, defective processing and degradation in the endoplasmic reticulum (ER), or a defective channel pore or gating properties. CFTR mutations are grouped into 5 mechanistic classes based on demonstrated or predicted molecular dysfunction.

Class 1: Includes mutations which cause defective synthesis. These mutations prevent transcription into a stable full length mRNA, leading to defective protein products. There is little or no full length protein that is produced and this causes loss of CFTR function.

Class II: Includes mutations with defective maturation of the protein (i.e., protein processing). CFTR mRNA is formed, but the protein fails to mature and does not traffic to the cell membrane. ΔF508 belongs to this class since this mutation disrupts the proper folding of the CFTR protein such that the protein is rapidly degraded before leaving the endoplasmic reticulum. At sub-physiological temperatures (23-30degrees Celsius) or upon chemical modification (e.g. with glycerol) ΔF508 CFTR is folded and processed and is functional at the apical membrane.

Class III: Includes mutations that have a defect in the channel regulation (blocked activation). CFTR protein is produced and traffics to the cell membrane but fails to respond to cyclic AMP stimulation. CFTR mutants are fully processed and properly localized but are not activated by cAMP. A number of mutations in the NBF1 and NBF2 belong to this class and these mutations disrupt binding and hydrolysis. Class III mutations can also affect phosphorylation sites in the R domain, but these sites are thought to be redundant and the domain confirmation may be sufficiently flexible to accommodate a mutation in the R domain without significant functional consequences.

Class IV: Mutations in this category are defective in conductance through the channel and represent milder mutations. CFTR protein is produced and traffics to the apical cell membrane, however mutations result in decreased chloride conductance due to altered ion conductance. Some of the mutations in this class (such as R117H, R334W, R347P) occur in the TMDs that form the channel pore and reduce the amount of chloride current by altering the rate of ion flow (R347P) or by changing the amount of time the channel remains open (R117H). The mutation P574H occurs in NBF1 and also alters the duration of channel opening.

Class V: These mutations are the result of abnormal splicing (decreased abundance). Mutations influence the quantity of full length mRNA transcript and protein required for normal function. This class includes promoter mutations, mutations that contribute to alternate splicing or mutations that cause inefficient protein processing and consequently reduced levels of functional protein.

1.2.4 Cellular mechanism of ion transport by normal and CF airway epithelia

There are at least three channels operating on the apical epithelial membrane: the epithelial sodium channel, the CFTR channel and the calcium activated chloride channel. The following channels are found on the basolateral membrane of airway epithelial cells: the sodium-potassium ATPase (Na^+-K^+-ATPase) pump, the sodium-chloride-potassium ($Na^+/2Cl^-/K^+$) co-transporter, and the potassium pump.

In normal airway epithelial cells sodium is absorbed through the sodium channels on the apical membrane and extruded through the basolateral membrane by the Na^+-K^+-ATPase pump. This transport of ions results in the movement of water by osmosis, thus this cycle of Na+ absorption also results in water absorption. These processes operate to limit the volume of the periciliary fluid layer. Chloride also passes paracellularly following sodium movement in an attempt to maintain electrical neutrality. Ion transport mechanisms also exist to re-hydrate the airway surface. Low intracellular Na^+ concentration created by the Na^+-K^+-ATPase triggers movement of Na^+ into the cell from the basolateral surface through the sodium-chloride-potassium ($Na^+/2Cl^-/K^+$) co-transporters. Chloride thus enters the cell down a favorable Na^+ gradient created by the ATPase and exits the cell down this created gradient toward the luminal surface through chloride channels. Water follows the chloride movement enabling rehydration of the airway surface. Chloride transport into the airway

lumen occurs via the CFTR channel, but this channel also stimulates extrusion through another apical chloride channel named the calcium activated chloride channel.

The CFTR channel also controls activity of the epithelial sodium channel and inhibits sodium movement through the epithelial sodium channel by switching the response to cAMP from an increasing to a decreasing channel opening probability. Under basal conditions sodium absorption predominates, but during periods of activity during which the potential for airway dehydration exists, chloride secretion can be activated to maintain periciliary fluid volume.

In CF this regulatory process is lost thus producing the increased sodium absorption; the epithelial sodium channels have a greater open probability that results in increased net sodium absorption, but also results in increased water absorption via osmosis and increased paracellular movement of chloride. Sodium exits the cell through the basolateral Na^+-K^+-ATPase pump. During periods of activity, chloride enters the cell through the basolateral $Na^+/2Cl^-/K^+$ co-transporters; however the Cl^- cannot be extruded into the airway through the defective CFTR channel. Sodium absorption (via epithelial sodium channel) and chloride extrusion (via calcium activated chloride channel) are both also compromised in CF.

1.2.5 Model for the development of airway disease in cystic fibrosis

CFTR controls fluid secretion from the apical surface of epithelial cells since the movement of chloride and sodium determine the osmotic forces of water movement. There are two competing theories to explain how these changes in airway surface lining fluid cause the abnormalities which are responsible for the pulmonary phenotype in CF. Boucher's group (7, 8) postulates that the volume of the airway surface fluid is too little (or insufficient) for two reasons:

- The defective CFTR cannot properly regulate fluid production.
- Fluid absorption is enhanced by the unregulated epithelial sodium channel.

Smith and associates (32-34) postulate that the volume of the airway surface lining fluid is the same in CF and non-CF patients, but that sodium and chloride concentrations are what differ. In their studies, measurements made on cultured normal and CF epithelial cells indicated differences in the sodium and chloride concentrations between normal and CF

cells, with higher salt concentration in CF cells. It is this high salt environment that is postulated to inactivate naturally occurring antibiotics.

The hypothesis most supported is the depletion of the airway surface lining fluid layer. It is hypothesized that airway destruction and obstruction in CF is due to volume depletion, decreased mucocilliary clearance and resultant infection (2, 35). The cause of the reduced ASL fluid is hypothesized to be due to the CFTR defect which results in accelerated sodium transport, failure to regulate cAMP dependent chloride secretion which leads to reduced periciliary lining fluid and failure of mechanical mucus clearance.

On normal airway epithelia a thin layer of mucus resides on top of the periciliary fluid layer. The mucous layer is capable of donating and accepting water and when there is depletion of liquid from the airway surface the mucous layer donates fluid until its height/volume ratio is reduced to approximately 50 percent and thereafter liquid is donated by the periciliary liquid layer(2, 36, 37). In the early stages of volume depletion, water is donated without it being replenished from the mucous layer to the periciliary liquid layer. A point is reached when the mucous layer can no longer donate liquid to the periciliary liquid layer and the periciliary liquid layer is absorbed.

Depletion of the periciliary liquid layer has a number of effects. Depletion of the periciliary liquid layer prevents the cilia from extending normally and beating and therefore compromises ciliary-dependent mucus clearance. The concentrated mucins in the mucous layer become more adhesive due to the volume depletion and this increased viscosity of the mucous layer also contributes to hindering ciliary-dependent transport. As a result of the periciliary liquid depletion there is adhesion of the previously mobile mucous layer to the cell surface and this allows the contact of mucus with the cell surface glycocalyx which is believed to effectively adhere the previously mobile mucous layer to the airway surfaces(2, 8, 36). This has two effects (36):

- It abolishes mucus clearance due to failure of the cilia to extend and beat.
- It degrades the ability of cough to remove mucus due to the physical adhesion of mucus to airway epithelia surfaces.

There is continued mucus secretion by goblet cells into immobile airway surface mucous plaques. The net effect of this accumulation is to extend the height of the mucous plaques and/or create mucous plugs. These mucous plaques in CF airways are characterized by:

- Increased mucus viscosity, which decreases the ability of soluble antimicrobial factors (e.g., lysozyme, lactoferin) to diffuse within the mucous plaque.
- Increased viscosity that degrades the capacity of neutrophils to penetrate this niche.
- Increased height of the mucous plaques on airway surfaces coupled with accelerated oxygen consumption by CF airway epithelia that fuels an increased sodium transport and creates hypoxic niches at the bases of the mucous plaques.

Boucher and associates have shown that motile *Pseudomonas aeruginosa* is particularly suited to this environment and the pathogen adheres to respiratory mucus (8). The increase in CF airway epithelial oxygen consumption is due to the absence of the CFTR's normal inhibitory activity on the epithelial sodium channel. The consequential accelerated sodium absorption on the apical membrane is fueled by an increased turnover rate of the ATP consuming Na^+-K^+-ATPase pump leading to the two- or three-fold increase in epithelial oxygen consumption(8).

The persistent mucin secretion into stationary mucus generates these thick airway plaques and/or plugs. The combination of thickened mucus and raised oxygen consumption of CF airway epithelium generates steep oxygen gradients within adherent mucin. Bacteria deposited on the thickened mucus can penetrate into these hypoxic zones. The excessive volume absorption having obliterated normal rotational mucus transport allows motile *P. aeruginosa* to penetrate these static plaques and grow in hypoxic/anaerobic CF mucus. *P. aeruginosa* responds to hypoxic mucus with alginate production and biofilm-like macrocolony formation, which is the dominant phenotype of *P. aeruginosa* in CF airways (2). Clearance of mucus is likely a key feature of innate lung defense and a fundamental defect leading to chronic CF pulmonary infections is failure to clear mucus that contains bound bacteria from the airway lumen.

1.2.6 Determinants of pulmonary disease severity in cystic fibrosis

It would seem that at some point early in a CF patient's life infection sets up an inflammatory response. Precisely how the CF electrolyte transport defect leads to persistent

lung infection and inflammation is still not well understood. The most common CFTR mutation (i.e., ΔF508) is generally associated with more severe pulmonary disease (38, 39). However, even among ΔF508 homozygous individuals there is a wide variation in disease severity (38) suggesting that there are other genetic and/or environmental factors that influence pulmonary function. In contrast, a good correlation has been shown between pancreatic sufficiency and CFTR mutation (40). A series of mutations have been identified which are classified as mild for pancreatic status (i.e., pancreatic sufficiency) (41) and these mutations are predominately missense mutations involving exons 5, 7, and 17 of the CFTR gene (42, 43) which result in amino acid changes in the TMDs (44). The lack of a clear association between CFTR genotype and pulmonary phenotype suggests that additional factors are involved and these could be genetic modifiers. Santis et al. investigated the level of lung function (FEV_1 % predicted) in pairs of siblings who had CF and showed significant association in %predFEV_1 (r=0.84) between the siblings (45). Concordance of pulmonary disease severity in first-degree relatives is suggestive evidence for shared genetic factors. In addition, residual chloride secretion in intestinal tissue was more concordant in monozygous than in dizygous twin pairs (46).

Some of the heterogeneity observed in pulmonary disease can be explained by CFTR mutations; that is variability in disease severity based on the class of the CFTR mutation. There are two other factors which also influence pulmonary disease severity. One such factor is the environment, specifically availability of therapeutic modalities to the patient, patient compliance to therapy recomendations and *Burkholderia cepacia* complex (BCC) cross-infection. The other factor is genes other than CFTR that may contribute to pulmonary disease progression. The concept of modifier genes as predictors of pulmonary disease heterogeneity and recent modifier gene studies will be reviewed below.

1.3 CANDIDATE MODIFIER GENES INVESTIGATED

Currently, one modifier gene locus has been identified for meconium ileus. In a genome-wide search for modifier genes in a CFTR-deficient mouse model of meconium ileus, a major modifier locus was detected near the centromere of mouse chromosome 7 (47). The orthologous region in humans is chromosome 19q13. A multicenter study of 197 CF sibling

pairs confirmed linkage of meconium ileus to this region (10). The gene responsible for this linkage has yet to be identified. This locus was not shown to modulate pulmonary disease severity either in mouse models or in humans.

1.3.1 Alpha-1-antitrypsin gene

Alpha-1-antitrypsin (α_1-AT) was one of the first candidate modifier genes investigated in CF. Studies had shown extremely high levels of neutrophil elastase (NE) in the airways of CF patients, which indicated that there is an imbalance between α_1-AT and elastase in the airways of patients with CF. The inflammatory process in response to pulmonary infection in CF airways is characterized by a massive influx of neutrophils (48). Neutrophils represent less than 5% of the cells recovered in bronchoalveolar lavage fluid in normal individuals but in adults and children (1-5 years of age) who have CF, neutrophils may comprise up to 95% of the cell population (49). Neutrophils contain a number of proteolytic enzymes one of which, neutrophil elastase (NE), has been implicated in excessive pulmonary damage observed in cigarette smokers and in CF patients. Elevated levels of NE have been reported in the sputum of patients who have CF (50, 51). NE is capable of causing direct lung damage by hydrolyzing all the major connective tissue proteins that make up the lung and airway matrix. NE also affects adherence of *P. aeruginosa* to the airway epithelium (52-54), impairs complement-mediated phagocytosis (particularly of *P. aeruginosa)* (55-58), stimulates interleukin-8 (IL-8) secretion (59) and stimulates mucus production (58, 60, 61). In normal hosts, the actions of NE are prevented primarily by alpha-1-antitrypsin (α_1-AT), a serine protease inhibitor that binds to NE and inhibits the breakdown of elastic tissue in the lung. Normal to elevated levels of α_1-AT have been reported in the airway secretions (49) and plasma (49, 54) of CF patients. Elevated levels of α_1-AT have been reported during pulmonary infections in this patient population (53).

The α_1-AT MZ genotype had been shown to be a risk factor for COPD (62). This candidate gene has been investigated as a possible modifier gene in CF in the last decade with conflicting results concerning the role that α_1-AT may play in pulmonary disease progression in CF. It is reasonable to hypothesize that individuals who have lower than normal levels of α_1-AT may be at increased risk for lung damage. Several mutations of the

α_1-AT gene result in a deficiency of this antiprotease. There is also evidence that α_1-AT genotype influences the acute phase response (55).

Doring and colleagues (57) showed no association between α_1-AT S and Z alleles and pulmonary disease severity in CF but they did show an earlier age of onset of *P. aeruginosa* infection in CF subjects with these deficiency alleles (6 out of a total sample of 215). Mahadeva and associates questioned this association and in fact showed that CF patients who were heterozygous for the S and Z alleles (19 out of a total sample of 147) had higher levels of pulmonary function than wild type individuals (63). In another study, the same authors showed that the α_1-AT Z and S deficiency alleles were not more prevalent in those CF patients with severe pulmonary disease (dead or lung transplanted CF patients)(64). Meyer and associates more recently showed similar age of onset of *P. aeruginosa* infection in MZ (N=5) or MS (N=16) CF patients compared with matched wild-type CF patients (65).

Another polymorphism in the α_1-AT gene studied in pulmonary diseases is the A allele of the polymorphism (G1237A) in the 3' region of the gene. Kalsheker and associates showed that the 3' mutation was associated with COPD (66). Morgan and colleagues provided *in vitro* evidence that the association with COPD may be due to deficiency in the α_1-AT acute phase response (55). Sandford and associates (67), however, did not find that the 3' mutation attenuated the acute phase rise in α_1-AT in their study of patients undergoing open-heart surgery. Similarly, Madadeva and associates showed that the 3' mutation had no effect on α_1-AT levels in CF patients (63), whereas Henry and associates (68) showed less severe pulmonary disease and fewer infective pulmonary exacerbations over 2 years in CF patients who were heterozygous for the A allele. These data suggest that heterozygotes may have a slower disease progression.

In CF cohorts, the association of α_1-AT genotype and pulmonary disease severity is unclear (57, 63, 64), the main limitation of these studies being their small sample sizes (number of subjects with the deficiency alleles 6-20) and therefore the high possibility of type 2 error (false negative)(69).

1.3.2 Innate immunity genes

The majority of children diagnosed as having CF develop chronic pulmonary infection, most often with *P. aeruginosa*, in late childhood. There is a positive association between age of first colonization with this pathogen and ultimate respiratory failure in CF patients. Genes regulating the first defense to bacterial and viral infections, especially during the first years of life can have a major impact on survival. There are at least four potential genes, which may increase the risk for developing bacterial and viral infections during infancy. Homozygosity or compound heterozygosity for mannose binding lectin (MBL2) variants has been shown to predispose young children to pulmonary infections with pathogens commonly cultured from CF patients (70, 71). MBL2 is synthesized by the liver and released upon IL-6 stimulation; it provides protection against bacterial and viral infections especially in infancy before adaptive immunity becomes established and provides non-specific defense against pathogens continually. MBL2 protein binds to mannose containing proteins or carbohydrates on bacterial and viral surfaces which are then recognized by alveolar macrophages which induce phagocytosis and release inflammatory cytokines.

The MBL2 gene is located on chromosome 10 and there are 3 missense polymorphisms at codons 54, 57 and 52 of exon 1. Further variation due to a polymorphism in the promoter region at position –221 has been shown to cause down regulation of gene expression and consequently lower serum MBL2 concentrations (72). Another polymorphism in the promoter region (at position at –550) is also associated with reduced MBL2 expression (73). Garred and associates (72) showed reduced serum MBL2 levels in heterozygotes for the coding polymorphisms, no MBL2 production in homozygotes for the variant alleles and lower gene expression with the promoter region polymorphism in a CF cohort. They also showed reduced lung function, earlier infection with *P. aeruginosa*, and also showed a trend suggesting that CF patients who had the variant structural and promoter alleles were more likely to become colonized with BCC. A closer look at their data however shows that the main difference in pulmonary function over the eight year longitudinal interval studied was that the MBL2 deficient group had a lower average lung function and one could not see a progressive decline in pulmonary function, in fact the graphs in their study showed a steady increase in pulmonary function in both groups which the investigators contributed to changes in pulmonary function testing equipment. Further evidence of increased susceptibility to BCC infection and MBL2 deficiency has come from Davies and colleagues

(74) who showed that MBL2 binds to BCC and activates complement. Specifically, they obtained BCC isolates from 10 CF patients and showed increased binding (61%) of MBL2 to BCC compared with non-mucoid strains of *P. aeruginosa* (2.9% binding). Gabolde and associates (75) in their case-control study of CF patients, matched for age and sex with normal healthy controls, showed a significant decline in lung function in those homozygous or heterozygous for the variant structural alleles. In all studies, sample size limited the power of their studies.

Three other innate immunity genes which are proposed as candidate modifier genes are located in close proximity to MBL2; pulmonary surfactant protein genes SP-A1, SP-A2, and SP-D are located in succession on the long arm of chromosome 10. The SP-A and SP-D genes have been localized at 10q22-q23 and MBL2 placed at 10q21. As these genes are located in close proximity to one another, it cannot be dismissed that any association found between disease severity and MBL2 may be due to linkage disequilibrium with the pulmonary surfactant genes. The SP-A1 and SP-A2 genes are separated by an SP-A pseudogene. The SP-A1 gene is flanked by the SP-D gene.

The pulmonary surfactant proteins SP-A (SP-A1 and SP-A2) and SP-D activate complement through the alternative pathway of complement activation. Unlike MBL2, SP-A and SP-D are synthesized within the lung predominately by alveolar type II cells and Clara cells. The pulmonary surfactant genes function as opsonins by binding to specific receptors on macrophages (76) and are also suggested to have both pro- and anti-inflammatory properties (77, 78). These surfactant proteins have been shown to bind to a large number of pathogens, including common pathogens found in CF sputum, although they differ in their interaction with the pathogens.

Vandivier and associates (79) showed impaired clearance of apoptotic cells from CF airways and reduced clearance of apoptotic cells in SP-D knockout mice. The authors speculated that the severity of the inflammatory response may be enhanced by defective clearance due to SP-A or SP-D deficiency. Results from knockout mice have shown that SP-A deficiency and SP-D deficiency are associated with pulmonary infection (80-82). Polymorphisms within the SP-A locus have been associated with pulmonary disease including respiratory distress syndrome (83) and chronic obstructive pulmonary disease (84). SP-A and SP-D

polymorphisms (9 polymorphisms in the SP-A genes and 2 in the SP-D gene) have shown negative results in association studies of acute respiratory distress syndrome (85).

In summary the association of MBL2 genotype and pulmonary disease severity is unclear in CF cohorts (72, 74, 75, 86). The pulmonary surfactant genes are good candidates as modifier genes in CF.

1.3.3 Antioxidant genes

Glutathione (GSH) is an important factor in the prevention of oxidant-induced lung injury. Reduced GSH has been shown in a number of pulmonary diseases including idiopathic pulmonary fibrosis (87-89), acute respiratory distress syndrome (90, 91), COPD (92), idiopathic respiratory distress syndrome (93) and in HIV-positive patients (94-97). Decreased GSH levels have also been shown in CF patients (98) and in the CF mouse (99). There are two groups of enzymes in the GSH synthesis pathway which have been investigated as affecting GSH function. The glutathione S-transferases (GSTs) are involved in detoxification of hydroperoxides by conjugating them with glutathione (H_20_2; the reaction is: 2 GSH + H_20_2 →GSSG + 2H_20) (100). There are several families of GSTs which have been identified in humans and are referred to as α, μ, π, and θ (101). GSTM1, GSTP1 and GSTT1 genes are polymorphic. The μ class GST, GSTM1, is located on chromosome 1 and the π class, GSTP1 is located on chromosome 11, and both are synthesized by the liver (102). GSTT1 is located on chromosome 22 and is synthesized by the liver and red blood cells**.** A rate limiting step in the glutathione synthesis pathway is the enzyme glutathione cysteine ligase. The catalytic subunit of the gene contains a polymorphic GAG trinucleotide repeat. In humans three alleles containing 7, 8, or 9 repeats have been identified.

Roum and associates (103) showed reduced levels of GSH in airway surface lining fluid, but also showed reduced levels of oxidized GSH, thereby suggesting that the low GSH was likely due to low GSH production and not excessive GSH oxidation. The CFTR is permeable to oxidized GSH, and defective CFTR results in altered permeability to GSH. Aerosolized GSH has been shown to reduce airway inflammation in CF (98). Transfection of normal CFTR or synthetic chloride channels to CF epithelial cells restores GSH secretion (104, 105).

There are two functionally active GSTM1 alleles and a null allele which produces no protein. Those homozygous for the null allele have been shown to have increased risk for some cancers, pulmonary emphysema and chronic bronchitis. Hull and Thomson (106)showed that CF patients who have the null allele exhibited somewhat worse lung function, clinical scores of disease severity and nutritional status. These patients were also more likely to be positive for *P. aeruginosa* infection(106). Baranov and associates (107)earlier showed CF patients who were homozygous for the null allele were more likely to die before the age of five years. Of CF patients who were pancreatic sufficient and manifested pulmonary symptoms in their study 51.1% were homozygous for the null allele (total sample N=194)(107). Henrion-Caude and colleagues (108) showed a significant association between CF liver disease and a GSTP1 polymorphism (exon 5 Ile105Val) in 106 CF children. No association with liver disease was shown for GSTM1 for this CF cohort.

1.3.4 Pro-inflammatory and anti-inflammatory mediators

Another group of studies have focused on investigating inflammatory and anti-inflammatory genes as candidate modifier genes in CF. CF is characterized by chronic inflammation focused on the airway lumen. Arkwright and associates (109) investigated inflammatory (tumor necrosis factor alpha (TNFα), interferon gamma (IFN-γ)), and anti-inflammatory (interleukin 10 (IL-10), transforming growth factor-β1 (TGF-β1)) cytokine genes as candidate modifier genes in 261 CF patients. They showed the high producer TGF-β1 (+915) genotypes to be associated with more severe pulmonary disease progression and earlier age of colonization with *P.aeruginosa* (109). In an earlier study Arkwright and associates (110) showed more rapid pulmonary disease progression in CF patients homozygous for the ΔF508 mutation (N=171) and another TGF-β1 polymorphism (Leu+869Pro). This same polymorphism was investigated in the group's more recent study but no association was shown with pulmonary disease decline (109). Arkwright and associates did not show a significant association for TNF-α and decline in pulmonary function (109), which has been previously reported in a cohort of 53 CF children (106).

Major histocompatibility complex class II genes have been investigated as possible modifier genes in CF. Aron and colleagues (111)investigated the contribution of this group of genes to the inflammation seen in this population. MHC class II antigen has been shown to be

expressed in CF nasal polyps (112). The investigators reported a significant association of HLA-DR7 with the presence of *P. aeruginosa* infection implying that the HLA class II region may modulate susceptibility to this pathogen (112). However, they did not show an association of HLA-DR alleles with level of lung function. In summary, study sample sizes are small and further study to replicate results on independent study sample of larger sample size is necessary.

1.3.5 Additional modifier genes

Another family of genes which we propose as candidate modifier genes are the mucin genes. Twenty mucin genes that have been characterized, six (MUC2, MUC4, MUC5AC, MUC5B, MUC8 and MUC18) are known to be expressed in the respiratory tract (113-115). The main characteristic of all mucin genes is the presence of a central domain composed of a variable number of tandem repeats. Each repeat unit contains multiple sites for O-linked glycosylation. There is marked inter-individual variation in the number of glycosylated side chains on the mucus proteins MUC4 and MUC5AC due to inherited variability in the number of repeats. The MUC4 gene is one of the most polymorphic and the number of tandem repeats varies from 145-395 between individuals (116). MUC5AC is also highly polymorphic (117) and the protein forms a major constituent of respiratory mucin (118). Although correlation of the degree of glycosylation and the visco-elastic properties of mucus from individual subjects has not been studied it is likely that such a relationship does exist. Increased glycosylation would be expected to increase mucus visco-elasticity and slow mucociliary clearance of purulent broncho-pulmonary secretions. There is increased secretion of mucus in CF which is stimulated by inflammation. There is also hypertrophy of submucosal glands and goblet cell hyperplasia and metaplasia (extension of goblet cells into the bronchioles where they normally do not occur). These phenomena also account for some of the increase of mucus secretion. The increased viscosity of the mucus can be partially explained by the CFTR defect and the effects of chronic inflammation. The mucin 2 gene (MUC2) contains a variable number of tandem 69bp repeats (119, 120). These repeats contain multiple glycosylation sites and were previously believed to be monomorphic (119), but preliminary work in our laboratory has shown that this region is polymorphic. This may have functional consequences since a mucin with a higher number of repeats will have a higher number of glycosylation sites and this may alter the viscoelastic properties of the

molecule. The MUC5B gene contains highly variable number of tandem repeats in intron 36 (121).

1.4 STATEMENT OF THE PROBLEM

Variability in pulmonary disease progression in CF cannot be explained by the CFTR defect alone. The CFTR defect in association with gene defects in other genes involved in modulating the severity of the inflammatory response and response to infection may contribute to more severe and rapid phenotypic progression of pulmonary disease in CF. The main questions that this study addressed were:

- Are genes other than the CFTR gene responsible for the variability in pulmonary disease severity and pulmonary disease progression seen in CF?
- Are genes other than the CFTR gene responsible for differential survival (i.e., clinical outcome; i.e., that is early lung transplantation or death) in CF?
- Are genes other than the CFTR gene responsible for susceptibility to recurrent pulmonary infections?
- Are genes other than the CFTR gene responsible for age of first and chronic infection with the respiratory pathogen *P. aeruginosa*?
- Are genes other that the CFTR gene responsible for the variability in frequency of pulmonary infections requiring intravenous therapy seen in CF patients?
- Are genes (i.e., MBL2) other than the CFTR gene responsible for the variability in pulmonary disease severity and pulmonary disease progression seen in CF patients chronically infected with BCC?
- Are genes (i.e., MBL2) other than the CFTR gene responsible for differential survival in CF patients chronically infected with BCC?

The main secondary questions that this study addressed were:

- Is there a difference in pulmonary disease severity and progression following chronic infection with BCC based on genomovar group and random amplification of polymorphic DNA (RAPD)-type?

- Is there a difference in pulmonary disease severity and progression in CF patients infected with BCC only versus infected only with *P. aeruginosa* or infected with both pathogens?
- Is there a difference in survival following chronic infection with BCC based on genomovar group and RAPD-type?
- Is there a difference in survival among CF patients who are chronically infected with BCC versus *P. aeruginosa* or infected with both pathogens and neither pathogen?
- Are genes (i.e., MBL2) other than the CFTR gene responsible for increased susceptibility to BCC colonization in CF?

1.5 MAIN STUDY HYPOTHESIS

Polymorphisms in genes other than CFTR influence CF disease severity and progression and are associated with:

- Worse pulmonary disease.
- Accelerated pulmonary disease progression
- Earlier infection with *P.aeruginosa*.
- Susceptibility to *B. cepacia* complex colonization
- Worse clinical outcome (early lung transplantation, death).

Specific Aims:

Primary aims:

1. Determine whether alpha-1-antitrypsin deficiency is associated with more severe pulmonary disease and pulmonary disease progression.
2. Determine whether MBL2 deficiency is associated with more severe pulmonary disease and pulmonary disease progression, earlier chronic colonization with *P.aeruginosa* and increased susceptibility to *B. cepacia* colonization.
3. Perform haplotypes analysis for MBL2 and determine if MBL2 is in linkage disequilibrium with the adjacent pulmonary surfactant genes.
4. Determine whether the pulmonary surfactant genes are in linkage disequilibrium.
5. Determine whether pulmonary surfactant gene variants are associated more severe pulmonary disease and pulmonary disease progression and earlier chronic colonization with *P. aeruginosa*.

6. Determine whether there is a relationship between length of the MUC2 and MUC5AC tandem repeats and the severity of pulmonary disease in CF.
7. Determine whether the GST gene polymorphisms and the catalytic subunit of the glutathione cysteine ligase gene are associated with more severe pulmonary disease, pulmonary disease progression and earlier chronic colonization with *P.aeruginosa*.

Secondary aims:

1. Determine whether BCC colonization is associated with a more rapid decline in pulmonary function and more severe clinical outcome taking into consideration genomovar grouping of BCC.
2. Determine whether MBL2 deficiency and BCC colonization with genomovar III versus II are associated with more severe pulmonary disease progression following colonization.
3. Determine whether the pulmonary surfactant gene polymorphisms are associated with susceptibility to *P.aeruginosa* colonization and with more severe pulmonary disease progression following colonization.
4. Determine whether MBL2 deficiency and the pulmonary surfactant gene polymorphisms are associated with differences in the frequencies of pulmonary infections requiring IV therapy.

1.6 SIGNIFICANCE

The proposed study will extend current knowledge of secondary genetic factors in pulmonary disease progression in CF. Polymorphisms in genes which modulate the severity of inflammation and increase susceptibility to infection could contribute to a more rapid progression of pulmonary disease if such polymorphism can be identified and this information could be used in patients' care. Our larger relative sample will help correct for confounding genetic and environmental parameters, which have been a problem in previous studies.

1.7 BIBLIOGRAPHY

1.Zielenski, J., and L. C. Tsui. 1995. Cystic fibrosis: genotypic and phenotypic variations. *Annu Rev Genet* 29:777-807.

2.Boucher, R. C. 2004. New concepts of the pathogenesis of cystic fibrosis lung disease. *Eur Respir J* 23(1):146-58.

3.Tsui, L. C., M. Buchwald, D. Barker, J. C. Braman, R. Knowlton, J. W. Schumm, H. Eiberg, J. Mohr, D. Kennedy, N. Plavsic, and et al. 1985. Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science* 230(4729):1054-7.

4.Knowlton, R. G., O. Cohen-Haguenauer, N. Van Cong, J. Frezal, V. A. Brown, D. Barker, J. C. Braman, J. W. Schumm, L. C. Tsui, M. Buchwald, and et al. 1985. A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. *Nature* 318(6044):380-2.

5.Wainwright, B. J., P. J. Scambler, J. Schmidtke, E. A. Watson, H. Y. Law, M. Farrall, H. J. Cooke, H. Eiberg, and R. Williamson. 1985. Localization of cystic fibrosis locus to human chromosome 7cen-q22. *Nature* 318(6044):384-5.

6.White, R., S. Woodward, M. Leppert, P. O'Connell, M. Hoff, J. Herbst, J. M. Lalouel, M. Dean, and G. Vande Woude. 1985. A closely linked genetic marker for cystic fibrosis. *Nature* 318(6044):382-4.

7.Boucher, R. C. 2001. Pathogenesis of cystic fibrosis airways disease. *Trans Am Clin Climatol Assoc* 112:99-107.

8.Boucher, R. C. 2002. An overview of the pathogenesis of cystic fibrosis lung disease. *Adv Drug Deliv Rev* 54(11):1359-71.

9.Hoiby, N., E. W. Flensborg, B. Beck, B. Friis, S. V. Jacobsen, and L. Jacobsen. 1977. Pseudomonas aeruginosa infection in cystic fibrosis. Diagnostic and prognostic significance of Pseudomonas aeruginosa precipitins determined by means of crossed immunoelectrophoresis. *Scand J Respir Dis* 58(2):65-79.

10.Zielenski, J. 2000. Genotype and phenotype in cystic fibrosis. *Respiration* 67(2):117-33.

11.Mangos, J. A., N. R. McSherry, and P. J. Benke. 1967. A sodium transport inhibitory factor in the saliva of patients with cystic fibrosis of the pancreas. *Pediatr Res* 1(6):436-42.

12.Spock, A., H. M. Heick, H. Cress, and W. S. Logan. 1967. Abnormal serum factor in patients with cystic fibrosis of the pancreas. *Pediatr Res* 1(3):173-7.

13.Spock, A., H. M. Heick, H. Cress, and W. Logan. 1967. In vitro study of ciliary motility to detect individuals with active cystic fibrosis and carriers of the disease. *Bibl Paediatr* 86:200-6.

14.Goodchild, M. C., J. H. Edwards, K. P. Glenn, C. Grindey, R. Harris, P. Mackintosh, and J. Wentzel. 1976. A search for linkage in cystic fibrosis. *J Med Genet* 13(6):417-9.

15.Eiberg, H., J. Mohr, K. Schmiegelow, L. S. Nielsen, and R. Williamson. 1985. Linkage relationships of paraoxonase (PON) with other markers: indication of PON-cystic fibrosis synteny. *Clin Genet* 28(4):265-71.

16.Schmiegelow, K., H. Eiberg, L. C. Tsui, M. Buchwald, P. D. Phelan, R. Williamson, W. Warwick, E. Niebuhr, J. Mohr, M. Schwartz, and et al. 1986. Linkage between the loci for cystic fibrosis and paraoxonase. *Clin Genet* 29(5):374-7.

17.Tsui, L. C., M. Zsiga, D. Kennedy, N. Plavsic, D. Markiewicz, and M. Buchwald. 1985. Cystic fibrosis: progress in mapping the disease locus using polymorphic DNA markers. I. *Cytogenet Cell Genet* 39(4):299-301.

18.Rommens, J. M., S. Zengerling-Lentes, B. Kerem, G. Melmer, M. Buchwald, and L. C. Tsui. 1989. Physical localization of two DNA markers closely linked to the cystic fibrosis locus by pulsed-field gel electrophoresis. *Am J Hum Genet* 45(6):932-41.

19.Rommens, J. M., M. C. Iannuzzi, B. Kerem, M. L. Drumm, G. Melmer, M. Dean, R. Rozmahel, J. L. Cole, D. Kennedy, N. Hidaka, and et al. 1989. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* 245(4922):1059-65.

20.Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J. L. Chou, and et al. 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245(4922):1066-73.

21.Eiberg, H., K. Schmiegelow, C. Koch, J. Mohr, M. Schwartz, and E. Niebuhr. 1985. Cystic fibrosis; hint of linkage with F13B. *Clin Genet* 27(2):206.

22.Beaudet, A., A. Bowcock, M. Buchwald, L. Cavalli-Sforza, M. Farrall, M. C. King, K. Klinger, J. M. Lalouel, G. Lathrop, S. Naylor, and et al. 1986. Linkage of cystic fibrosis to two tightly linked DNA markers: joint report from a collaborative study. *Am J Hum Genet* 39(6):681-93.

23.Kerem, B. S., J. A. Buchanan, P. Durie, M. L. Corey, H. Levison, J. M. Rommens, M. Buchwald, and L. C. Tsui. 1989. DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. *Am J Hum Genet* 44(6):827-34.

24.Kerem, B., J. M. Rommens, J. A. Buchanan, D. Markiewicz, T. K. Cox, A. Chakravarti, M. Buchwald, and L. C. Tsui. 1989. Identification of the cystic fibrosis gene: genetic analysis. *Science* 245(4922):1073-80.

25.Farrall, M., B. J. Wainwright, G. L. Feldman, A. Beaudet, Z. Sretenovic, D. Halley, M. Simon, L. Dickerman, M. Devoto, and G. Romeo. 1988. Recombinations between IRP and cystic fibrosis. *Am J Hum Genet* 43(4):471-5.

26.Rich, D. P., M. P. Anderson, R. J. Gregory, S. H. Cheng, S. Paul, D. M. Jefferson, J. D. McCann, K. W. Klinger, A. E. Smith, and M. J. Welsh. 1990. Expression of cystic fibrosis transmembrane conductance regulator corrects defective chloride channel regulation in cystic fibrosis airway epithelial cells. *Nature* 347(6291):358-63.

27.Gregory, R. J., S. H. Cheng, D. P. Rich, J. Marshall, S. Paul, K. Hehir, L. Ostedgaard, K. W. Klinger, M. J. Welsh, and A. E. Smith. 1990. Expression and characterization of the cystic fibrosis transmembrane conductance regulator. *Nature* 347(6291):382-6.

28.Drumm, M. L., H. A. Pope, W. H. Cliff, J. M. Rommens, S. A. Marvin, L. C. Tsui, F. S. Collins, R. A. Frizzell, and J. M. Wilson. 1990. Correction of the cystic fibrosis defect in vitro by retrovirus-mediated gene transfer. *Cell* 62(6):1227-33.

29.Boucher, R. C. 2004. Relationship of airway epithelial ion transport to chronic bronchitis. *Proc Am Thorac Soc* 1(1):66-70.

30.Dharmsathaphorn, K., J. Cohn, and G. Beuerlein. 1989. Multiple calcium-mediated effector mechanisms regulate chloride secretory responses in T84-cells. *Am J Physiol* 256(6 Pt 1):C1224-30.

31.Zielenski, J., P. Patrizio, D. Markiewicz, R. H. Asch, and L. C. Tsui. 1997. Identification of two mutations (S50Y and 4173delC) in the CFTR gene from patients with congenital bilateral absence of vas deferens (CBAVD). *Hum Mutat* 9(2):183-4.

32.Smith, J. J., P. H. Karp, and M. J. Welsh. 1994. Defective fluid transport by cystic fibrosis airway epithelia. *J Clin Invest* 93(3):1307-11.

33.Smith, J. J., and M. J. Welsh. 1993. Fluid and electrolyte transport by cultured human airway epithelia. *J Clin Invest* 91(4):1590-7.

34.Smith, J. J., S. M. Travis, E. P. Greenberg, and M. J. Welsh. 1996. Cystic fibrosis airway epithelia fail to kill bacteria because of abnormal airway surface fluid. *Cell* 85(2):229-36.

35.Knowles, M. R., J. M. Robinson, R. E. Wood, C. A. Pue, W. M. Mentz, G. C. Wager, J. T. Gatzy, and R. C. Boucher. 1997. Ion composition of airway surface liquid of patients with

cystic fibrosis as compared with normal and disease-control subjects. *J Clin Invest* 100(10):2588-95.

36.Matsui, H., B. R. Grubb, R. Tarran, S. H. Randell, J. T. Gatzy, C. W. Davis, and R. C. Boucher. 1998. Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. *Cell* 95(7):1005-15.

37.Matsui, H., S. H. Randell, S. W. Peretti, C. W. Davis, and R. C. Boucher. 1998. Coordinated clearance of periciliary liquid and mucus from airway surfaces. *J Clin Invest* 102(6):1125-31.

38.Burke, W., M. L. Aitken, S. H. Chen, and C. R. Scott. 1992. Variable severity of pulmonary disease in adults with identical cystic fibrosis mutations. *Chest* 102(2):506-9.

39.Wilmott, R. W., S. L. Tyson, and D. J. Matthew. 1985. Cystic fibrosis survival rates. The influences of allergy and Pseudomonas aeruginosa. *Am J Dis Child* 139(7):669-71.

40.1993. Correlation between genotype and phenotype in patients with cystic fibrosis. The Cystic Fibrosis Genotype-Phenotype Consortium. *N Engl J Med* 329(18):1308-13.

41.Kerem, E., M. Corey, B. Kerem, P. Durie, L. C. Tsui, and H. Levison. 1989. Clinical and genetic comparisons of patients with cystic fibrosis, with or without meconium ileus. *J Pediatr* 114(5):767-73.

42.Ferrari, M., and L. Cremonesi. 1996. Genotype-phenotype correlation in cystic fibrosis patients. *Ann Biol Clin (Paris)* 54(6):235-41.

43.Welsh, M. J., and A. E. Smith. 1993. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell* 73(7):1251-4.

44.Brancolini, V., L. Cremonesi, E. Belloni, E. Pappalardo, R. Bordoni, M. Seia, S. Russo, R. Padoan, A. Giunta, and M. Ferrari. 1995. Search for mutations in pancreatic sufficient cystic fibrosis Italian patients: detection of 90% of molecular defects and identification of three novel mutations. *Hum Genet* 96(3):312-8.

45.Santis, G., L. Osborne, R. A. Knight, and M. E. Hodson. 1990. Independent genetic determinants of pancreatic and pulmonary status in cystic fibrosis. *Lancet* 336(8723):1081-4.

46.Bronsveld, I., F. Mekus, J. Bijman, M. Ballmann, J. Greipel, J. Hundrieser, D. J. Halley, U. Laabs, R. Busche, H. R. De Jonge, B. Tummler, and H. J. Veeze. 2000. Residual chloride secretion in intestinal tissue of deltaF508 homozygous twins and siblings with cystic fibrosis. The European CF Twin and Sibling Study Consortium. *Gastroenterology* 119(1):32-40.

47.Zielenski, J., M. Corey, R. Rozmahel, D. Markiewicz, I. Aznarez, T. Casals, S. Larriba, B. Mercier, G. R. Cutting, A. Krebsova, M. Macek, Jr., E. Langfelder-Schwind, B. C. Marshall, J. DeCelie-Germana, M. Claustres, A. Palacio, J. Bal, A. Nowakowska, C. Ferec, X. Estivill, P. Durie, and L. C. Tsui. 1999. Detection of a cystic fibrosis modifier locus for meconium ileus on human chromosome 19q13. *Nat Genet* 22(2):128-9.

48.Cantin, A. M., S. Lafrenaye, and R. O. Begin. 1991. Antineutrophil elastase activity in cystic fibrosis serum. *Pediatric Pulmonology* 11(3):249-53.

49.Birrer, P., N. G. McElvaney, A. Rudeberg, C. W. Sommer, S. Liechti-Gallati, R. Kraemer, R. Hubbard, and R. G. Crystal. 1994. Protease-antiprotease imbalance in the lungs of children with cystic fibrosis. *American Journal of Respiratory & Critical Care Medicine* 150(1):207-13.

50.O'Connor, C. M., K. Gaffney, J. Keane, A. Southey, N. Byrne, S. O'Mahoney, and M. X. Fitzgerald. 1993. alpha 1-Proteinase inhibitor, elastase activity, and lung disease severity in cystic fibrosis. *Am Rev Respir Dis* 148(6 Pt 1):1665-70.

51.Goldstein, W., and G. Doring. 1986. Lysosomal enzymes from polymorphonuclear leukocytes and proteinase inhibitors in patients with cystic fibrosis. *American Review of Respiratory Disease* 134(1):49-56.

52.Suter, S. 1986. Imbalance between polymorphonuclear leukocyte proteases and antiproteases in chronic pyogenic infections and its relation to the proteolysis of complement component C3. *Complement* 3(1):1-24.

53.Suter, S., U. B. Schaad, H. Tegner, K. Ohlsson, D. Desgrandchamps, and F. A. Waldvogel. 1986. Levels of free granulocyte elastase in bronchial secretions from patients with cystic fibrosis: effect of antimicrobial treatment against Pseudomonas aeruginosa. *Journal of Infectious Diseases* 153(5):902-9.

54.Brown, M. A., W. J. Morgan, P. R. Finley, and P. Scuderi. 1991. Circulating levels of tumor necrosis factor and interleukin-1 in cystic fibrosis. *Pediatric Pulmonology* 10(2):86-91.

55.Morgan, K., G. Scobie, P. Marsters, and N. A. Kalsheker. 1997. Mutation in an alpha1-antitrypsin enhancer results in an interleukin-6 deficient acute-phase response due to loss of cooperativity between transcription factors. *Biochim Biophys Acta* 1362(1):67-76.

56.Hilman, B. C. 1997. Genetic and immunologic aspects of cystic fibrosis. *Ann Allergy Asthma Immunol* 79(5):379-90; quiz 391-4.

57.Doring, G., H. Krogh-Johansen, S. Weidinger, and N. Hoiby. 1994. Allotypes of alpha 1-antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for deltaF508. *Pediatric Pulmonology* 18(1):3-7.

58.Plotkowski, M. C., G. Beck, J. M. Tournier, M. Bernardo-Filho, E. A. Marques, and E. Puchelle. 1989. Adherence of Pseudomonas aeruginosa to respiratory epithelium and the effect of leucocyte elastase. *J Med Microbiol* 30(4):285-93.

59.Santis, G., L. Osborne, R. Knight, M. E. Hodson, and M. Ramsay. 1990. Genetic influences on pulmonary severity in cystic fibrosis. *Lancet* 335(8684):294.

60.Santis, G., L. Osborne, R. Knight, M. Ramsay, R. Williamson, and M. Hodson. 1990. Cystic fibrosis haplotype association and the delta F508 mutation in adult British CF patients. *Hum Genet* 85(4):424-5.

61.Plotkowski, M. C., A. M. Saliba, S. H. Pereira, M. P. Cervante, and O. Bajolet-Laudinat. 1994. Pseudomonas aeruginosa selective adherence to and entry into human endothelial cells. *Infect Immun* 62(12):5456-63.

62.McElvaney, N. G., R. C. Hubbard, P. Birrer, M. S. Chernick, D. B. Caplan, M. M. Frank, and R. G. Crystal. 1991. Aerosol alpha 1-antitrypsin treatment for cystic fibrosis. *Lancet* 337(8738):392-4.

63.Mahadeva, R., R. C. Westerbeek, D. J. Perry, J. U. Lovegrove, D. B. Whitehouse, N. R. Carroll, R. I. Ross-Russell, A. K. Webb, D. Bilton, and D. A. Lomas. 1998. Alpha1-antitrypsin deficiency alleles and the Taq-I G-->A allele in cystic fibrosis lung disease. *Eur Respir J* 11(4):873-9.

64.Mahadeva, R., S. Stewart, D. Bilton, and D. A. Lomas. 1998. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. *Thorax* 53(12):1022-4.

65.Meyer, P., A. Braun, and A. A. Roscher. 2002. Analysis of the two common alpha-1-antitrypsin deficiency alleles PiMS and PiMZ as modifiers of Pseudomonas aeruginosa susceptibility in cystic fibrosis. *Clin Genet* 62(4):325-7.

66.Kalsheker, N. A., G. L. Watkins, S. Hill, K. Morgan, R. A. Stockley, and R. B. Fick. 1990. Independent mutations in the flanking sequence of the alpha-1- antitrypsin gene are associated with chronic obstructive airways disease. *Dis Markers* 8(3):151-7.

67.Sandford, A. J., T. Chagani, J. J. Spinelli, and P. D. Pare. 1999. alpha1-antitrypsin genotypes and the acute-phase response to open heart surgery. *Am J Respir Crit Care Med* 159(5 Pt 1):1624-8.

68.Henry, M. T., S. Cave, J. Rendall, C. M. O'Connor, K. Morgan, M. X. FitzGerald, and N. Kalsheker. 2001. An alpha(1)-antitrypsin enhancer polymorphism is a genetic modifier of pulmonary outcome in cystic fibrosis. *Eur J Hum Genet* 9(4):273-8.

69.Wacholder, S., S. Chanock, M. Garcia-Closas, L. El Ghormli, and N. Rothman. 2004. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 96(6):434-42.

70.Koch, A., M. Melbye, P. Sorensen, P. Homoe, H. O. Madsen, K. Molbak, C. H. Hansen, L. H. Andersen, G. W. Hahn, and P. Garred. 2001. Acute respiratory tract infections and mannose-binding lectin insufficiency during early childhood. *Jama* 285(10):1316-21.

71.Summerfield, J. A., M. Sumiya, M. Levin, and M. W. Turner. 1997. Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series. *Bmj* 314(7089):1229-32.

72.Garred, P., T. Pressler, H. O. Madsen, B. Frederiksen, A. Svejgaard, N. Hoiby, M. Schwartz, and C. Koch. 1999. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest* 104(4):431-7.

73.Madsen, H. O., P. Garred, S. Thiel, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, and A. Svejgaard. 1995. Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol* 155(6):3013-20.

74.Davies, J., O. Neth, E. Alton, N. Klein, and M. Turner. 2000. Differential binding of mannose-binding lectin to respiratory pathogens in cystic fibrosis. *Lancet* 355(9218):1885-6.

75.Gabolde, M., M. Guilloud-Bataille, J. Feingold, and C. Besmond. 1999. Association of variant alleles of mannose binding lectin with severity of pulmonary disease in cystic fibrosis: cohort study. *Bmj* 319(7218):1166-7.

76.Tenner, A. J. 1999. Membrane receptors for soluble defense collagens. *Curr Opin Immunol* 11(1):34-41.

77.Eggleton, P., and K. B. Reid. 1999. Lung surfactant proteins involved in innate immunity. *Curr Opin Immunol* 11(1):28-33.

78.Crouch, E. C. 1998. Structure, biologic properties, and expression of surfactant protein D (SP-D). *Biochim Biophys Acta* 1408(2-3):278-89.

79.Vandivier, R. W., V. A. Fadok, C. A. Ogden, P. R. Hoffmann, J. D. Brain, F. J. Accurso, J. H. Fisher, K. E. Greene, and P. M. Henson. 2002. Impaired clearance of apoptotic cells from cystic fibrosis airways. *Chest* 121(3 Suppl):89S.

80.Wert, S. E., M. Yoshida, A. M. LeVine, M. Ikegami, T. Jones, G. F. Ross, J. H. Fisher, T. R. Korfhagen, and J. A. Whitsett. 2000. Increased metalloprotenase activity, oxidant production, and emphysema in surfactant protein D gene-inactivated mice. *Proceeds of the National Academy of Science* 97:5972-.

81.LeVine, A. M., K. E. Kurak, M. D. Bruno, J. M. Stark, J. A. Whitsett, and T. R. Korfhagen. 1998. Surfactant protein-A-deficient mice are succeptible to *Pseudomonas aeruginosa* infection. *Am J Respir Cell Mol Biol* 19:700-708.

82.LeVine, A. M., M. D. Bruno, K. M. Huelsman, G. F. Ross, J. A. Whitsett, and T. R. Korfhagen. 1997. Surfactant protein A-deficient mice are susceptible to group B streptococcal infection. *J Immunol* 158(9):4336-40.

83.Kala, P., T. Ten Have, H. Nielsen, M. Dunn, and J. Floros. 1998. Association of pulmonary surfactant protein A (SP-A) gene and respiratory distress syndrome: interaction with SP-B. *Pediatr Res* 43(2):169-77.

84.Guo, X., H. M. Lin, Z. Lin, M. Montano, R. Sansores, G. Wang, S. DiAngelo, A. Pardo, M. Selman, and J. Floros. 2001. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 18(3):482-90.

85.Lin, Z., C. Pearson, V. Chinchilli, S. M. Pietschmann, J. Luo, U. Pison, and J. Floros. 2000. Polymorphisms of human SP-A, SP-B, and SP-D genes: association of SP-B Thr131Ile with ARDS. *Clin Genet* 58(3):181-91.

86.Davies, J. C., M. W. Turner, N. Klein, and t. L. M. S. Group. 2004. Impaired pulmonary status in cystic fibrosis adults with two mutated *MBL-2* alleles. *Eur Respir J* 24:798-804.

87.MacNee, W., and I. Rahman. 1995. Oxidants/antioxidants in idiopathic pulmonary fibrosis. *Thorax* 50 Suppl 1:S53-8.

88.Cantin, A. M., R. C. Hubbard, and R. G. Crystal. 1989. Glutathione deficiency in the epithelial lining fluid of the lower respiratory tract in idiopathic pulmonary fibrosis. *Am Rev Respir Dis* 139(2):370-2.

89.Rahman, I., E. Skwarska, M. Henry, M. Davis, C. M. O'Connor, M. X. FitzGerald, A. Greening, and W. MacNee. 1999. Systemic and pulmonary oxidative stress in idiopathic pulmonary fibrosis. *Free Radic Biol Med* 27(1-2):60-8.

90.Pacht, E. R., A. P. Timerman, M. G. Lykens, and A. J. Merola. 1991. Deficiency of alveolar fluid glutathione in patients with sepsis and the adult respiratory distress syndrome. *Chest* 100(5):1397-403.

91.Moss, M., D. M. Guidot, M. Wong-Lambertina, T. Ten Hoor, R. L. Perez, and L. A. Brown. 2000. The effects of chronic alcohol abuse on pulmonary glutathione homeostasis. *Am J Respir Crit Care Med* 161(2 Pt 1):414-9.

92.Rahman, I., S. K. Biswas, and A. Kode. 2006. Oxidant and antioxidant balance in the airways and airway diseases. *Eur J Pharmacol* 533(1-3):222-39.

93.Boda, D., I. Nemeth, and S. Pinter. 1998. Surface tension, glutathione content and redox ratio of the tracheal aspirate fluid of premature infants with IRDS. *Biol Neonate* 74(4):281-8.

94.Droge, W., and E. Holm. 1997. Role of cysteine and glutathione in HIV infection and other diseases associated with muscle wasting and immunological dysfunction. *Faseb J* 11(13):1077-89.

95.Mihm, S., J. Ennen, U. Pessara, R. Kurth, and W. Droge. 1991. Inhibition of HIV-1 replication and NF-kappa B activity by cysteine and cysteine derivatives. *Aids* 5(5):497-503.

96.Pacht, E. R., P. Diaz, T. Clanton, J. Hart, and J. E. Gadek. 1997. Alveolar fluid glutathione decreases in asymptomatic HIV-seropositive subjects over time. *Chest* 112(3):785-8.

97.Staal, F. J., S. W. Ela, M. Roederer, M. T. Anderson, and L. A. Herzenberg. 1992. Glutathione deficiency and human immunodeficiency virus infection. *Lancet* 339(8798):909-12.

98.Roum, J. H., Z. Borok, N. G. McElvaney, G. J. Grimes, A. D. Bokser, R. Buhl, and R. G. Crystal. 1999. Glutathione aerosol suppresses lung epithelial surface inflammatory cell-derived oxidants in cystic fibrosis. *J Appl Physiol* 87(1):438-43.

99.Velsor, L. W., A. van Heeckeren, and B. J. Day. 2001. Antioxidant imbalance in the lungs of cystic fibrosis transmembrane conductance regulator protein mutant mice. *Am J Physiol Lung Cell Mol Physiol* 281(1):L31-8.

100.Boxer, L. A., J. M. Oliver, S. P. Spielberg, J. M. Allen, and J. D. Schulman. 1979. Protection of granulocytes by vitamin E in glutathione synthetase deficiency. *N Engl J Med* 301(17):901-5.

101.Hayes, J. D., and L. I. McLellan. 1999. Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free Radic Res* 31(4):273-300.

102.Feher, J., G. Lengyel, and A. Blazovics. 1998. Oxidative stress in the liver and biliary tract diseases. *Scand J Gastroenterol Suppl* 228:38-46.

103. Roum, J. H., R. Buhl, N. G. McElvaney, Z. Borok, and R. G. Crystal. 1993. Systemic deficiency of glutathione in cystic fibrosis. *J Appl Physiol* 75(6):2419-24.

104. Gao, L., J. R. Broughman, T. Iwamoto, J. M. Tomich, C. J. Venglarik, and H. J. Forman. 2001. Synthetic chloride channel restores glutathione secretion in cystic fibrosis airway epithelia. *Am J Physiol Lung Cell Mol Physiol* 281(1):L24-30.

105. Gao, L., K. J. Kim, J. R. Yankaskas, and H. J. Forman. 1999. Abnormal glutathione transport in cystic fibrosis airway epithelia. *Am J Physiol* 277(1 Pt 1):L113-8.

106. Hull, J., and A. H. Thomson. 1998. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. *Thorax* 53(12):1018-21.

107. Baranov, V. S., T. Ivaschenko, B. Bakay, M. Aseev, R. Belotserkovskaya, H. Baranova, P. Malet, J. Perriot, P. Mouraire, V. N. Baskakov, G. A. Savitskyi, S. Gorbushin, S. I. Deyneka, E. Michnin, A. Barchuck, V. Vakharlovsky, G. Pavlov, V. I. Shilko, T. Guembitzkaya, and L. Kovaleva. 1996. Proportion of the GSTM1 0/0 genotype in some Slavic populations and its correlation with cystic fibrosis and some multifactorial diseases. *Hum Genet* 97(4):516-20.

108. Henrion-Caude, A., C. Flamant, M. Roussey, C. Housset, A. Flahault, A. A. Fryer, K. Chadelat, R. C. Strange, and A. Clement. 2002. Liver disease in pediatric patients with cystic fibrosis is associated with glutathione S-transferase P1 polymorphism. *Hepatology* 36(4 Pt 1):913-7.

109. Arkwright, P. D., V. Pravica, P. J. Geraghty, M. Super, A. K. Webb, M. Schwarz, and I. V. Hutchinson. 2003. End-organ dysfunction in cystic fibrosis: association with angiotensin I converting enzyme and cytokine gene polymorphisms. *Am J Respir Crit Care Med* 167(3):384-9.

110. Arkwright, P. D., S. Laurie, M. Super, V. Pravica, M. J. Schwarz, A. K. Webb, and I. V. Hutchinson. 2000. TGF-beta(1) genotype and accelerated decline in lung function of patients with cystic fibrosis. *Thorax* 55(6):459-62.

111. Aron, Y., B. S. Polla, T. Bienvenu, J. Dall'ava, D. Dusser, and D. Hubert. 1999. HLA class II polymorphism in cystic fibrosis. A possible modifier of pulmonary phenotype. *Am J Respir Crit Care Med* 159(5 Pt 1):1464-8.

112. Wang, D., G. M. Levasseur-Acker, R. Jankowski, G. Kanny, D. A. Moneret-Vautrin, D. Charron, A. Lockhart, and E. Swierczewski. 1997. HLA class II antigens and T lymphocytes in human nasal epithelial cells. Modulation of the HLA class II gene transcripts by gamma interferon. *Clin Exp Allergy* 27(3):306-14.

113. Seregni, E., C. Botti, S. Massaron, C. Lombardo, A. Capobianco, A. Bogni, and E. Bombardieri. 1997. Structure, function and gene expression of epithelial mucins. *Tumori* 83(3):625-32.

114. Shankar, V., P. Pichan, R. L. Eddy, Jr., V. Tonk, N. Nowak, S. N. Sait, T. B. Shows, R. E. Schultz, G. Gotway, R. C. Elkins, M. S. Gilmore, and G. P. Sachdev. 1997. Chromosomal localization of a human mucin gene (MUC8) and cloning of the cDNA corresponding to the carboxy terminus. *Am J Respir Cell Mol Biol* 16(3):232-41.

115. Rose, M. C., and J. A. Voynow. 2006. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* 86(1):245-78.

116. Nollet, S., N. Moniaux, J. Maury, D. Petitprez, P. Degand, A. Laine, N. Porchet, and J. P. Aubert. 1998. Human mucin gene MUC4: organization of its 5'-region and polymorphism of its central tandem repeat array. *Biochem J* 332 (Pt 3):739-48.

117. Pigny, P., V. Guyonnet-Duperat, A. S. Hill, W. S. Pratt, S. Galiegue-Zouitina, M. C. d'Hooge, A. Laine, I. Van-Seuningen, P. Degand, J. R. Gum, Y. S. Kim, D. M. Swallow, J. P. Aubert, and N. Porchet. 1996. Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 38(3):340-52.

118. Hovenberg, H. W., J. R. Davies, A. Herrmann, C. J. Linden, and I. Carlstedt. 1996. MUC5AC, but not MUC2, is a prominent mucin in respiratory secretions. *Glycoconj J* 13(5):839-47.

119. Toribara, N. W., J. R. Gum, Jr., P. J. Culhane, R. E. Lagace, J. W. Hicks, G. M. Petersen, and Y. S. Kim. 1991. MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorphism. *J Clin Invest* 88(3):1005-13.

120. Debailleul, V., A. Laine, G. Huet, P. Mathon, M. C. d'Hooghe, J. P. Aubert, and N. Porchet. 1998. Human mucin genes MUC2, MUC3, MUC4, MUC5AC, MUC5B, and MUC6 express stable and extremely large mRNAs and exhibit a variable length polymorphism. An improved method to analyze large mRNAs. *J Biol Chem* 273(2):881-90.

121. Desseyn, J. L., K. Rousseau, and A. Laine. 1999. Fifty-nine bp repeat polymorphism in the uncommon intron 36 of the human mucin gene MUC5B. *Electrophoresis* 20(3):493-6.

CHAPTER 2: STUDY METHODOLOGY

2.0 METHODS AND PROCEDURES

The chapter describes clinic and patient study population recruitment, laboratory techniques performed, clinical data (definition of variables, description of outcome and independent variables) collected, hypotheses tested and statistical data analysis techniques.

2.1 SUBJECT AND ELIGIBILITY CRITERIA

Patients attending the following Canadian CF clinics were recruited:
Adult CF clinic at St. Paul's Hospital (Vancouver, B.C.).
Children's CF clinic at B.C. Children's Hospital (Vancouver, B.C.),
Adult CF clinic at Victoria General Hospital (Victoria, B.C.),
Children's CF clinic at Victoria Children's Hospital (Victoria, B.C.).
Adult CF clinic at Hamilton Health Sciences (Hamilton, Ontario).
Children's CF clinic at Hamilton Health Sciences (Hamilton, Ontario).
Adult CF clinic at Hôtel-Dieu du Chum (Montreal, Quebec).
Children's CF Clinic at Hospital for Sick Children (Toronto, Ontario).
Adult CF clinic at St. Michael's Hospital (Toronto, Ontario).
Adult CF clinic in Seattle (Seattle, Washington, U.S.A.).
An attempt was made to recruit all patients attending the participating clinics. The potential study sample based on clinic patient numbers was 1265. Table 2.1 provides a description of clinic participation for each gene studied. Patients attending the Toronto clinics only participated in the investigation of the alpha-1-antitrypsin (α_1-AT) gene as a potential modifier gene and provided 2 years of clinical follow-up data. Patients attending the Seattle clinic only participated in the investigation of the GST and GCLC genes as potential modifier genes and only provided cross-sectional data.

The remaining genes were investigated on patients attending the Vancouver, Hamilton, and Montreal clinics. The Victoria clinics did not participate in the investigation of the alpha-1-antitrypsin gene and were recruited subsequently to increase sample size and used in our investigation of innate immunity and antioxidant genes. Table 2.1 includes the number of CF patients genotyped for each polymorphism by the center attended.

Patients with a diagnosis of CF on the basis of clinical signs, elevated sweat chloride values and/or positive genotyping for mutant CFTR allele(s) were recruited for the study. CF patients who had received a lung transplant were also recruited and pulmonary function data from prior to transplantation were collected for these CF patients.

In a separate sub-study we also recruited 31 consecutive patients from the St. Paul's Hospital adult CF clinic (mean age (±SEM) 27.5(1.1) years) who developed an acute pulmonary exacerbation, to measure serum α_1-AT levels during the acute phase and 2-3 months later during a stable phase.

Details of the experimental procedures, the risks and benefits involved were explained to subjects before obtaining written consent, which was approved by the Ethics Committees of the institutions participating in this study (Appendix A).

2.2 LABORATORY TECHNIQUES

2.2.1 Extraction of genomic DNA from human whole blood

Ten milliliters of whole blood was collected from consenting CF patients. Samples for DNA were collected in one 10ml or two 7ml EDTA lavender topped tubes and stored in a 4°C fridge for between 0-4 days and then either immediately extracted or stored in a -20°C freezer prior to extraction.

Thawed whole blood for each sample was transferred into a 50ml sterile centrifuge Falcon tube and the tube was rinsed with 10ml of 0.1% IGEPAL CA-630 (octylphenyl-polyethylene glycol) in sterile water and added to the Falcon tube. Additional 0.1% IGEPAL was added to the Falcon tube to a final volume of 40ml. The tube with the sample mixture was mixed and allowed to sit for 20 minutes for complete lysis of the red blood cells. The tube was then spun for 20 minutes (in a Beckman Model TJ-6 Centrifuge at 2300RPM (1100g)) to pellet the white cells. The supernatant was poured out and the cell pellet washed with 40ml of 0.1% IGEPAL and mixed on vortex for 5 -10 seconds. The sample tube was spun again (at 2300RPM for 20 minutes), the supernatant (which contained the remaining lysed red blood cells) was poured off. The cell pellet was suspended in 10ml of digestion buffer containing

50µl of 20mg/ml proteinase K. The sample was incubated overnight at 50°C with agitation at 250 RPM. On the second day an equal volume (10ml) of phenol/chloroform/isoamyl alcohol (25:24:1) and 1ml of 2M NaCl were added to each sample and mixed gently using a vortex. The tubes were spun in the centrifuge (at 2300RPM for 20 minutes), then the upper aqueous layer was transferred to a new 50ml Falcon tube. To precipitate the DNA, 2 volumes of 95% ethanol were added and the solution gently mixed. A 1.5ml sterile microfuge tube with 1ml 70% ethanol was prepared and using a sterile pipette tip the DNA was spooled out of the falcon tube and added to the microfuge tube. The DNA-ethanol mixture was centrifuged for 5 minutes at 14,000 RPM and the ethanol poured off and the pellet air-dried under Kimwipes. The DNA pellets were finally dissolved in 500-800µl of TE buffer (pH=8.0) at room temperature.

2.2.2 DNA Quantification

The optical density (OD) of each sample was measured at a wavelength of 260 nm and 280 nm using an ultraviolet spectrophotometer. We measured a 100-fold dilution of the original DNA solution. One OD260 unit corresponds to a double strand DNA concentration of 50 µg/ml. Proteins or RNA also absorb UV light at 260nm and 280nm. The purity of DNA samples were estimated by the ratio A260/A280. A ratio of 1.7 to 1.9 was indicative of a pure DNA sample.

2.3 GENOTYPING METHODS

2.3.1 Amplification of DNA using the Polymerase Chain Reaction

The purpose of the polymerase chain reaction (PCR) is to artificially replicate a short sequence of DNA, in order to make millions of copies of this DNA sequence. Two oligonucleotides are used as primers for a series of synthetic reactions that are catalyzed by a thermostable DNA polymerase (e.g., *Taq* DNA polymerase). The oligonucleotides are complementary to sequences that lie on opposite strands of the template DNA and flank the segment we want to amplify. Genomic DNA is first heated (approximately 94°C) to denature the double stranded DNA molecules in a solution containing an excess of the two oligonucleotide primers and the four deoxynucleoside triphosphates (dNTPs). The reaction mixture is then cooled (45-65°C) to allow the primers to anneal to their target sequences and

then heated to an intermediate temperature (i.e., 70-75°C) which allows extension of the annealed primers from their 3' termini by *Taq* DNA polymerase. The cycle of denaturation, annealing and primer extension is repeated 35-40 times using a thermal cycler. The products of each round of amplification serve as the template for the next, effectively doubling the amount of synthesized DNA with each cycle. Typically, this exponential amplification is terminated when the amount of active *Taq* DNA polymerase is exhausted at which time the selected DNA sequence has been amplified 10^6 times.

2.3.2 Genotyping using Restriction Fragment Length Polymorphisms

Restriction enzymes (restriction endonucleases) bind specifically to, and cleave double stranded DNA at specific sites that are within or adjacent to particular sequences of nucleotides. Restriction enzymes are purified from bacteria; many have been isolated and they recognize specific sequences of double stranded DNA which are 4-8 nucleotides long, although there are a few enzymes which recognize longer sequences. These enzymes cut DNA at specific sequences and produce DNA fragments of specific sizes which are called restriction fragments. Single nucleotide polymorphisms can create or destroy restriction enzyme recognition sites. Therefore the presence or the absence of a recognition site can be used to detect a polymorphism. The restriction enzyme is added to the PCR mixture and digested in a specific restriction enzyme buffer and then analyzed on a gel. When more than one PCR is performed in one reaction, this is defined as a multiplex PCR.

2.3.3 Amplification of DNA using Site Directed Mutagenesis Polymerase Chain Reaction followed by Restriction Digestion

In this technique, a primer is used which contains a one base mismatch and gives rise to a new restriction enzyme site following the PCR. Then an appropriate restriction enzyme is used to cut the DNA fragment at this new site. This method is used when the polymorphism does not lie within a naturally occurring restriction enzyme recognition site; therefore we create a site which can be recognized by a known restriction enzyme.

2.3.4 Amplification of DNA using allele-specific oligonucleotide PCR or Sequence Specific Priming Polymerase Chain Reaction

In allele-specific oligonucleotide (ASP) PCR or sequence-specific priming (SSP) PCR, the primer contains the specific polymorphic site. We used two allele specific primers that were complementary to the respective two alleles. If the allele contains the specific polymorphism then we obtain a PCR product. This is visualized on a gel where we obtain a band when the allele is present. If the allele is not present no band is seen. In order to ensure that the PCR has worked an internal control is used.

2.4 GENE POLYMORPHISMS INVESTIGATED AND GENOTYPING METHODS

2.4.1 Gene Polymorphisms investigated and genotyping techniques

Patients were genotyped for polymorphisms in the following genes:

α_1-AT gene: We studied the S and Z mutations in exons 3 and 5, respectively and the *Taq*I polymorphism in the 3' untranslated region.

Innate immunity genes:

MBL gene: promoter polymorphism (Y or X [G→C] at -221) and the 3 single base structural polymorphisms at codons 52 (Arg→Cys, D allele), 54 (Gly→Asp, B allele), and 57 (Gly→Glu, C allele).

SPA-1: we studied a polymorphism in exon 4 (C655T; Arg219Trp).

SPA-2: we studied two polymorphisms. One polymorphism in exon 2 (A26C;Thr9Asn) and one in exon 4 (A667C; Lys223 Gln).

SPD: we studied a polymorphism in exon 1 (C32T; Thr11Met).

Other genes:

MUC2 gene: we studied the imperfectly conserved repeats (mRNA sequence accession number: NM_002457 (1)).

MUC5b gene: we studied the variable number tandem repeat region in intron 36 (genomic DNA accession number: Y09788 (2, 3)).

GSTP1 gene: we studied the polymorphism in exon 5 (A313G; Ile105Val).

GSTM1 and GSTT1 genes: we studied deletions of these two genes.

GCLC gene: we studied the trinucleotide repeat (GAC) upstream from the translation start codon(4).

Table 2.2 shows the gene polymorphisms studied and genotyping methods used. The PHASE program (5, 6)was used to identify haplotypes for the MBL gene and the surfactant genes.

2.4.2 Alpha-1-antitrypsin gene

The Z and S polymorphisms were examined by multiplex site directed mutagenesis PCR / restriction fragment length polymorphisms (RFLP) assays using primers which amplified regions of exon 5 and 3, respectively(7). Two mixtures were prepared; the first mixture consisted of 0.5µl of 10µM EX55 upstream primer (5'-TAA GGC TGT GCT GAC CAT CGT C-3') and 0.5µl of 10µM BYZ downstream primer (5'-CAA AGG GTT TGT TGA ACT TGA CC-3') for the Z mutation, and 1µl of 10µM EX35 upstream primer with a mismatch to produce a *Taq*I site (5'-GAG GGG AAA CTA CAG CAC CTC G-3') and 1µl of 10µM EX33 downstream primer (5'-ACC CTC AGG TTG GGG AAT CAC C-3') for the S mutation, 1µl (0.1µg) DNA and water (6 µl ddH_20) to a final volume of 10µl was heated (7 minute denaturation step at 94°C followed by 10 minutes at 80°C). Mixture 2 of the PCR reaction contained 2µl 10xPCR buffer (by Invitrogen containing 500mM KCl, 200mM Tris-Cl, pH 8.4), 2µl 2mM deoxynucleoside triphosphates, 0.6µl 50mM $MgCl_2$, 0.1µl *Taq* DNA polymerase (0.5U), with water (5.3 µl ddH_20) added to a final volume of 10µl. Mixture 2 was added to mixture 1 followed by thermal cycling (35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 58°C, 30 second extension at 72°C and a final 5 minute extension at 72°C). Digestion of PCR products was performed with restriction enzyme *Taq*I (0.5µl (10U)), 3µl 10x*Taq*I buffer, 0.3µl 100X BSA and 6.2µl ddH_20) at 65°C for five hours. At the Z mutation site, the restriction enzyme cuts the wild-type allele into 2 fragments (123bp and 21bp) and leaves the Z allele uncut (144bp). At the S mutation site, the restriction enzyme cuts the wild-type allele into 2 fragments (78bp and 20bp) and leaves the S allele uncut (98bp). The PCR product was electrophoresed on a 2.5% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.1).

The 3 prime *Taq*I polymorphism ($G_{1237}\rightarrow A$) was examined by PCR/RFLP using primers which amplified a 373bp region. Each PCR reaction containing 1µl of 10µM P2 forward

primer (5'-CTC TCA GGT CTG GTG TCA TCC C-3') and 1µl of 10µM P4 reverse primer (5'-GAC ACA GCA GCC AGG AAG TCC-3'), 1µl (0.1µg) DNA and 7 µl ddH_20 to a final volume of 10µl was heated (7 minute denaturation step at 94°C followed by 10 minutes at 80°C) (Mixture 1). Mixture 2 of the PCR reaction contained 2µl 10xPCR buffer (by Invitrogen containing 500mM KCl, 200mM Tris-Cl, pH 8.4), 2µl 2mM deoxynucleoside triphosphates, 0.6µl 50mM $MgCl_2$, 0.1µl *Taq* DNA polymerase (0.5U), with 5.3 µl ddH_20 added to a final volume of 10µl. Mixture 2 was added to mixture 1 followed by thermal cycling (40 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C). Digestion of PCR products was performed with the restriction enzyme *Taq*I (0.5µl (10U) *Taq*I, 3µl 10x *Taq*I buffer, 0.3µl 100xBSA and 6.2µl ddH_20) at 65°C for five hours. The restriction enzyme cuts the wild-type G allele into 2 fragments (191bp and 182bp) and leaves the A allele uncut. The PCR product was electrophoresed on a 2.0% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.2).

2.4.3 Mannose-Binding Lectin gene

MBL2-B, C and D allele genotyping was performed by PCR-RFLP as described by Madsen et al. (8) with some modification. The B and C alleles were detected respectively by *Ban*I and *Mbo*II restriction enzyme digestion of the 329bp product that had been amplified by the MBP1left upstream primer (5'-GTA GGA CAG AGG GCA TGC TC-3') and the MBP1right downstream primer (5'-CAG GCA GTT TCC TCT GGA AGG-3'), followed by a 2% agarose gel electrophoresis in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light. *Ban*I cleaves the A allele into two fragments (245bp and 84bp) and leaves the B allele undigested. *Mbo*II cleaves the C allele into two fragments (266bp and 63bp) and leaves the A allele undigested (Figure 2.3 A and B).

Specifically for MBL2-B, the PCR reaction contained 2µl 10xPCR buffer (by Invitrogen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7), 0.6µl 50mM $MgCl_2$, 2µl 2mM deoxynucleoside triphosphates (dNTPs), 1µl of 10µM primer MBP1left and 1µl of 10µM primer MBP1right, 1µl (0.25µg) DNA, 0.1µl *Taq* DNA polymerase (by Invitrogen), with 12.3 µl ddH_20 added to a final volume of 20µl. Following a 2 minute denaturation step at

94°C, 35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C. Digestion of PCR products was performed with 0.25µl (10U/µl) of the restriction enzyme 0.25µl *Ban*I (10U/µl), 2.5µl 10x NE buffer 4, and 2.25µl ddH_20 to a final volume of 5µl at 37°C overnight (Figure 2.3 A).

Specifically for MBL-C, the PCR reaction contained 2µl 10xPCR buffer (by Invitrogen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7), 0.6µl 50mM $MgCl_2$, 2µl 2mM dNTPs, 1µl of 10µM primer MBP1left and 1µl of 10µM primer MBP1right, 1µl (0.25µg) DNA, 0.1µl *Taq* DNA polymerase (by Invitrogen), with 12.3 µl ddH_20 added to a final volume of 20µl. Following a 2 minute denaturation step at 94°C, 35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C. Digestion of PCR products was performed with 0.5µl (5U/µl) of the restriction enzyme 0.5µl *Mbo*II (5U/µl), 2.5µl 10x NE buffer 2, and 2µl ddH_20 to a final volume of 5µl at 37°C overnight (Figure 2.3B).

A *Mlu*I restriction enzyme site was introduced into the amplification product of the D allele by site-directed mutagenesis (SDM)-PCR. The PCR reaction contained 2µl 10xPCR buffer (by Invitrogen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7), 0.6µl 50mM $MgCl_2$, 2µl 2mM dNTPs, 1µl of 10µM primer MBP1left (5'-CAA CGG CTT CCC AGG CAA AGA CGC G-3') and 1µl of 10µM primer MBP1right (5'-ATC CCC AGG CAG TTT CCT CTG GAA GG-3'), 1µl (0.25µg) DNA, 0.1µl *Taq* DNA polymerase (by Invitrogen), with 12.7 µl ddH_20 added to a final volume of 20µl. Following a 2 minute denaturation step at 94°C, 35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C. Digestion of PCR products was performed with 0.5µl (10U/µl) of the restriction enzyme 0.5µl *Mlu*I (10U/µl), 2.5µl 10x NE buffer 3, and 2µl ddH_20 to a final volume of 5µl at 37°C overnight. The 121bp product was visualized with ethidium bromide staining under ultraviolet light on a 3% agarose gel following electrophoresis in a 0.5xTBE buffer. *Mlu*I cleaves the D allele into two bands (21 bp and 100 bp) and leaves the A allele uncut (Figure 2.4A).

MBL2-X/Y genotyping was performed by PCR using sequence-specific priming (SSP) PCR as described by Madsen et al. (9) with some modification. Two upstream primers contain alleles X and Y respectively (MBP-X: 5'-CAT TTG TTC TCA CTG CCA CC-3', MBP-Y: 5'-CAT TTG TTC TCA CTG CCA CG-3') and one downstream primer matches the sequence of MBL2 (5'-ACA TTC CTT GTG ACA CTG CG-3'). A PCR of the β-globulin gene was used as a positive control (Figure 4B). The PCR was performed in a total volume of 20 µl, containing 2µl 10xPCR buffer (by Invitrogen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7), 2µl 2mM dNTPs, 1 µl (0.25µg) genomic DNA, 0.5µl β-globulin-F (sense: 5'-CAA CTT CAT CCA CGT TCA CC-3), 0.5µl β-globulin-R (antisense: 5'-GAA GAG CCA AGG ACA GGT AC-3) and 0.75µl 10µM of the specific primers described above in the presence of 0.6µl 1.5 mM $MgCl_2$ and 1µl HotStarTaqDNA Polymerase (by Qiagen) with 12.4 µl ddH_20 added to a final volume of 20µl. PCRs were initiated by a 15 min polymerase activation step at 95^0C and completed by a final 5 min extension step at 72^0C. The temperature cycles for the PCRs were as follows: 35 cycles of 20 seconds at 94^0C, 50 seconds at 65^0C, and 20 sec at 72^0C. The 154bp product was visualized with ethidium bromide staining under ultraviolet light on a 1.5% agarose gel following electrophoresis in a 0.5xTBE buffer (Figure 2.4 B).

2.4.4 Pulmonary surfactant protein-A1 gene

The SPA-1 polymorphism in exon 4 (C655T; Arg219Trp) was examined by site directed mutagenesis PCR/RFLP using primers which amplified a 159bp region (10). Each PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15mM $MgCl_2$), 2µl 2mM deoxynucleoside triphosphates, 1µl of 10µM sense (5'-GCC ATT GCA AGC TTC GTG AA-3') with a mismatch to produce a *Taq*I site for the C allele and 10µM antisense (5'-CAC ACA CTG CTC TTT TCC TC-3') primers, 1µl (0.25µg) DNA, 0.1 hot start *Taq* DNA polymerase (Qiagen), with 12.9 µl ddH_20 added to a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 35 cycles were performed consisting of a 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C.

Digestion of PCR products with restriction enzyme *Taq*1 (0.5µl (10U), 2.5µl 10x*Taq*1 buffer, 0.25µl 100xBSA and 1.75µl ddH_20) was at 65°C for three hours. The restriction

enzyme cuts the C allele into 2 fragments (138bp and 21bp) and leaves the T allele uncut (159bp). The digested products were electrophoresed on a 3% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.5).

2.4.5 Pulmonary surfactant protein-A2 gene

Two polymorphisms were examined in the SPA-2 gene. The polymorphism in exon 4 (A667C; Lys223Gln) was examined by site directed mutagenesis PCR/RFLP using primers which amplified a 235bp region (11). Each PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15mM $MgCl_2$,), 2µl 2mM deoxynucleoside triphosphates, 1µl of 10µM sense (5'-GAG CCT GCA GGT CGG GGA AAA GC-3') with a mismatch to produce a *Hha*I site for the C allele and 10µM antisense (5'-CCT CCA GCT CTA ATA GCC ACA AGT-3') primers, 1µl (0.25µg) DNA, 0.1µl (0.5U) hot start *Taq* DNA polymerase (5U/µl , QIAGEN), with added water (12.9 µl ddH_20) to a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C. Digestion of PCR products with restriction enzyme *Hha*I (0.25µl (5U)), 2.5µl 10xNEB_4, 0.25µl 100xBSA and 2µl ddH_20 at 37°C overnight. The restriction enzyme cuts the C allele into 2 fragments (24bp and 211bp) and leaves the A allele uncut (235bp). The digested products were separated on a 3% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.6A).

The second SPA-2 polymorphism investigated in exon 2 (A26C; Thr9Asn) was examined by site directed mutagenesis PCR/RFLP using primers which amplified a 150bp region (11). Each PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15 mM $MgCl_2$), 2µl 2mM deoxynucleoside triphosphates, 1µl of 10µM sense (5'-GCT GTG CCC TCT GGC CCT TA-3') with mismatch to produce a *Mse*I site for the A allele and 10µM antisense (5'-TCC TTT GAC ACC ATC TC-3') primers, 1µl (0.1µg) genomic DNA, 0.1µl hot start *Taq* DNA polymerase (5U/µl, Qiagen), with 12.9 µl ddH_20 added to give a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 35

cycles were performed consisting of 30 second denaturation at 95°C, 30 second annealing at 56°C, 30 second extension at 72°C and a final 5 minute extension at 72°C.

Digestion of PCR products with 5 µl restriction enzyme *Mse*1 mixture (0.5µl (10U/µl) *Mse*1, 2.5µl 10xNEB2 buffer, 0.25µl 100xBSA and 1.75µl ddH_20) was performed at 37°C overnight. The restriction enzyme cuts the A allele into 2 fragments (18bp and 132bp) and leaves the C allele uncut (150bp). The digested products were electrophoresed on a 3% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.6 B).

2.4.6 Pulmonary surfactant protein D gene

The SPD polymorphism in exon 1 (32C/T, 11Thr/Met) was examined by site directed mutagenesis PCR/RFLP using primers which amplified a 101bp region. Sequence was obtained from Genbank (AH005286). Each PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15mM $MgCl_2$), 2µl 2mM deoxynucleoside triphosphates, 1µl of 10µM sense (5'-CTC CTC TCT GCA CTG GTC CT-3') with mismatch to produce a *Fsp*I site for the T allele and 10µM antisense (5'-ACC AGG GTG CAA GCA CTG CG-3') primers, 1µl (0.1µg) genomic DNA, 0.1µl (0.5U) hot start *Taq* DNA polymerase (by Qiagen), with added 12.9 µl ddH_20 added to a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 35 cycles were performed consisting of 30 second denaturation at 94°C, 30 second annealing at 60°C, 30 second extension at 72°C and a final 5 minute extension at 72°C.

Digestion of PCR products with restriction enzyme *Fsp*I (0.5µl (2.5U) *Fsp*I, 2.5µl 10xNEB4 buffer, 0.25µl 100xBSA and 2µl ddH_20) was performed at 37°C overnight. The restriction enzyme cuts the T allele into 2 fragments (82bp and 19bp) and leaves the C allele uncut (101bp). The digested products were electrophoresed on a 3% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.7).

2.4.7 Other genes

2.4.7.1 MUC2 gene

The MUC2 polymorphism was examined by PCR using primers which amplified the repetitive threonine/serine/proline-rich subdomain (Genbank NM 002457). There is a region of imperfectly conserved repeats (nucleotides 4228-5268 in GenBank accession # NM 002457) in this subdomain. Each PCR reaction contained 2µl 10xPCR buffer (500mM KCl, 200mM Tris-HCl, (pH 8.4)), 0.56µl 50mM $MgCl_2$ (final concentration 1.4mM), 3µl 2mM dNTPs, 2µl of 10µM sense (5'-GTGTCAATTGTTGCTGGCCC-3' nucleotides 4181-4200 in Genbank NM 002457) and 10µM antisense (5'-CCAGCCAGTCCAATGCAGA-3' nucleotides 5381-5400 in Genbank NM 002457) primers, 2µl (0.2µg) DNA, 0.1µl (0.5U) *Taq* DNA polymerase (*Invitrogen*), and 3µl 50% glycerol, 2µl 0.1% Triton x-100, with 3.34 µl ddH_20 added to give a final volume of 20µl. Following a 3 minute denaturation step at 94°C, 40 cycles were performed consisting of 30 second denaturation at 95°C, 30 second annealing at 64°C, 2 minute extension at 70°C and a final 6 minute extension at 72°C. The PCR product was electrophoresed on a 1% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.8).

2.4.7.2 MUC5B gene

The variable number tandem repeat (VNTR) in intron 36 of the MUC5B gene (Genbank Y09788) was examined by PCR using primers which amplified the entire VNTR region (nt 4000-4451 - again these numbers are meaningless unless we know which sequence they refer to). Each PCR reaction contained 2µl 10xPCR buffer (500mM KCl, 200mM Tris-HCl, (pH 8.4)), 2µl 2mM deoxynucleoside triphosphates, 2µl of 10µM sense (5'-AGT GTG CAG TGA CTG GCG AG-3' nt 3967-3986) and 10µM antisense (5'-CTA GAG TTG CAG GTG GCA GG-3' nt 4655-4674) primers, 0.56µl 50mM $MgCl_2$ (final concentration 1.4mM), 2µl (200ng) DNA, 0.1 (0.5U) *Taq* DNA polymerase (Invitrogen), 2µl 0.1% Triton x-100, and 3µl 50% glycerol, with 4.34 µl ddH_20 added to a final volume of 20µl. Following a 3 minute denaturation step at 94°C, 30 cycles were performed consisting of 30 second denaturation at 95°C, 30 second annealing at 64°C, 2 minute extension at 70°C and a final 6 minute extension at 72°C. The PCR product was electrophoresed on a 1.5% agarose gel in a

0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.9).

2.4.7.3 GSTP1 gene

The GSTP1 polymorphism in exon 5 (A313G; Ile105Val) was examined by PCR/RFLP using primers which amplified a 433bp region. Each PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15mM $MgCl_2$), 2µl 2mM deoxynucleoside triphosphates, 1µl of 10µM sense (5'-GTA GTT TGC CCA AGG TCA AG-3') and 10µM antisense (5'-AGC CAC CTG AGG GGT AAG-3') primers, 1µl (0.1µg) DNA, 0.1µg (0.5U) hot start *Taq* DNA polymerase (Qiagen), with added ddH_20 to a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 15 cycles were performed consisting of 30 second denaturation at 95°C, 30 second annealing at 60°C, 60 second extension at 72°C and another 20 cycles were performed consisting of 30 second denaturation at 95°C, 30 second annealing at 57°C, 30 second extension at 72°C and a final 5 minute extension at 72°C.

Digestion of PCR products with restriction enzyme *Bsm*AI (1.5µl *Bsm*AI (7.5U), 2.5µl 10xNEB buffer3, and 1.0µl ddH_20) at room temperature overnight. The restriction enzyme cuts the A allele into 2 fragments (104bp and 329bp) and the G allele at into three fragments (104bp, 222bp and 107bp). The digested products were electrophoresed on a 1.5% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light (Figure 2.10 A).

2.4.7.4 GSTM1 and GSTT1 genes

The GSTM1 and GSTT1 gene deletions were examined by multiplex PCR as described by Yim and associates (12)using the β-globulin gene as an internal control. The primer pairs for each gene were as follows:

GSTM1F (sense):5'-GAA CTC CCT GAA AAG CTA AAG C-3'
GSTM1R (antisense): 5'-GTT GGG CTC AAA TAT ACG GTG G-3'
GSTT1F (sense): 5'-TTC CTT ACT GGT CCT CAC ATC TC-3'
GSTT1R (antisense): 5'-TCA CCG GAT CAT GGC CAG CA-3'
β-Globulin F (sense): 5'-CAA CTT CAT CCA CGT TCA CC-3'

β-Globulin R (antisense): 5'-GAA GAG CCA AGG ACA GGT AC-3'

The PCR reaction contained 2µl 10xPCR buffer (by Qiagen containing KCl, $(NH_4)_2SO_4$, Tris-Cl, pH 8.7, and 15mM $MgCl_2$), 2µl 2mM deoxynucleoside triphosphates, 0.5µl β-globulin F and 0.5µl β-globulin R, 0.75µl GSTM1-F and 0.75µl GSTM1-R, 0.5µl GSTT1-F and 0.5µl GSTT1-R primers, 1µl (0.1µg) DNA, 0.1µl (0.5U) hot start *Taq* DNA polymerase (Qiagen), with 11.4 µl ddH_20 added to a final volume of 20µl. Following a 15 minute denaturation step at 95°C, 35 cycles were performed consisting of 45 second denaturation at 94°C, 45 second annealing at 60°C, 45 second extension at 72°C and a final 5 minute extension at 72°C. The PCR product was electrophoresed on a 2% agarose gel in a 0.5xTBE buffer and visualized with ethidium bromide staining under ultraviolet light. Visualized on the gel were: a 215bp fragment for GSTM1, a 480bp fragment for GSTT1, and a 268bp fragment for the β-globulin gene (Figure 2.10 B).

2.4.7.5 GCLC gene

The trinucleotide repeat guanine-adenine-guanine (GAG) was examined by microsatellite PCR as described by McKone and associates (4) for the Canadian cohort and the Seattle cohort.

2.5. PHENOTYPIC DATA COLLECTION

2.5.1 Pulmonary function

The subjects performed post-bronchodilator spirometry (forced expiratory volume in one second (FEV_1) and forced vital capacity (FVC)) in accordance with ATS criteria (13). A standard protocol was used by all centers. Values were expressed as a percent of the predicted normal values based on age, gender and height (14). The best-recorded post-bronchodilator measurements were used for this study. Predicted values were calculated from equations derived for adults (14) and children (15). For patients who had received lung transplantation, lung function data prior to transplantation were collected and used for statistical analyses.

2.5.2 Cross-sectional and longitudinal data sheets

Cross-sectional and longitudinal data were collected for the study. A list of variables and the format of the variables are presented in Table 2.3 and 2.4 for cross-sectional data and Table 2.5 for longitudinal data. Data collected for the cross-sectional data sheet included dates to calculate age of CF diagnosis, age of first infection with *Pseudomonas aeruginosa* and when possible age of chronic infection with *P. aeruginosa*, age at infection with, and age of chronic colonization with *Burkholderia cepacia* complex (BCC), age of transplantation and death. We also created categorical variables for the presence of chronic infection with *P. aeruginosa* to increase the study sample size as in some cases it was not possible to obtain the age of first or chronic infection with this pathogen, but it was possible to categorize them as chronically infected or not. Additional variables that were collected were CFTR genotype, sweat chloride levels, pancreatic sufficiency status, and presence/absence of liver disease, gastrointestinal reflux and diabetes mellitus. The subjects' most current stable clinical status data were entered in the cross-sectional data sheet as well (included lung function, height, weight and date of clinic encounter). Stable clinical status is defined below.

Our objective was to collect over 2 years of longitudinal data for each subject and ideally up to 10 years of continuous follow-up data. In some cases, for all the Canadian clinics, we have more than 10 years of clinical data collected and this is for CF patients who became transiently or chronically infected with BCC and data was collected for 2 intervals pre- (2 years) and post-acquisition of BCC (to most current encounter). Ultimately we have collected a minimum of seven to ten years of continuous follow-up data for the Vancouver and Victoria centers. Less than seven years of data were available for 11 (range 2.5-6.7 years) out of 107 patients from the Vancouver adult clinic and 5 out of 11 patients 1.0-3.6 years of data) from the Victoria children's clinics due to their young age or transfer of a patient from another center in Canada for whom we could not readily obtain the missing data (2 out of 18 from the Adult clinic in Victoria). Ten years of data were collected from the Vancouver B.C. Children's hospital for 89 out of 122 patients and for the remainder of the study sample due to their young age we were able to collect 2-5 years of clinical data for 25 patients and less than 2 years (1.2-1.9 years) of data for 8 patients. For the Montreal clinic we were able to collect 3-10 years of clinical data. We collected 5-10 years of data on 100 out of the 146 CF patients who volunteered for the study. We collected over 5 years and up to 10 years of clinical data on 61 out of the 92 CF patients who volunteered from the two

Hamilton clinics. The Toronto centers only participated in the study of the alpha-1-antitrypsin gene and longitudinal data (> 2 years) was not collected for the alpha-1-antitrypsin gene analysis.

The following data were obtained for the longitudinal data file: pulmonary function (i.e., FEV_1 and FVC; absolute and percent predicted values when available), height, weight, bacteriology (the primary pathogens *Pseudomonas aeruginosa* and BCC, as well as for *Staphylococcus aureus*, *Hemophilus influenza* and *Xanthomonas maltophilia*) and reason for visit. The date the clinical data were collected was recorded as the date of the encounter. These data were obtained from the patients' medical charts, as was information on frequency and duration of hospitalizations for each year of data collection. In the longitudinal data file, columns were available for entry of the date of admission and discharge from hospital. Hospitalizations were coded based on whether they were for pulmonary infections requiring intravenous (IV) antibiotic therapy, or for other non-pulmonary reasons. In the case of hospitalizations for pulmonary infections, the duration of therapy (start and end date for IV therapy) rather than the duration of hospitalization was collected for the longitudinal data file. Therapy with home IV for pulmonary infections has become the standard mode of treatment in the past seven years, in some cases patients are admitted for 2-5 days and then discharged to continue with IV therapy at home. In other cases patients are started immediately on home IV therapy without a short hospitalization. Clinical data in this instance (i.e., pulmonary function, height and weight) were collected close to the end of IV therapy (1-3 days prior to IV therapy completion) or 2-10 days following completion of IV therapy to represent stable clinical status. Exam date for this encounter is the date of clinical data collection. Bacteriology represents pathogens grown at time of diagnosis with a pulmonary exacerbation.

Outpatient encounters were coded as clinic stable or clinic ill (from a respiratory perspective). Since our main phenotype of study was pulmonary status, clinic visits were coded as stable or ill based on a respiratory perspective. Sometimes CF patients may be seen by the clinic during the interval of IV therapy and these encounters are also coded as such in order to identify them and exclude them as pulmonary function is expected to be adversely affected.

Longitudinal data collected from Hamilton for visits coded as clinic ill did not distinguish between ill from pulmonary causes versus other (such as weight loss, hemoptysis, liver disease, and diabetes). In cases where all the visits within a 12 month period were coded as clinic ill, the clinic visit with the highest pulmonary function data, which was also similar (higher or equal to) to pulmonary function data from the previous and/or subsequent 12 month interval which was coded as clinic stable was used as the encounter characterizing the patient for this year and accepted as the stable value for this year.

2.5.3 Definition of stable clinical status and pulmonary exacerbation in CF

Stable clinical status was defined as the absence of pulmonary exacerbation over the previous 4 weeks, absence of a current mild exacerbation requiring oral antibiotics and the absence of two or more clinical symptoms (increased cough, sputum volume and purulence, increased dyspnea, reduced weight and a fall in FEV_1>10%). Pulmonary exacerbation was defined as a pulmonary infection that required the administration of IV antibiotics based on clinical signs assessed by the CF physician.

2.5.4 Identification and typing of B. cepacia Complex

Organisms suspected of being BCC were sent to the Canadian *B. cepacia* complex Research and Referral Repository where BCC confirmation was performed as described in Henry and associates (16). Isolates were checked for purity, and then screened for growth on BCSA agar (17). Organisms that were BSCA positive were set up to the API Rapid NFT strip (Biomerieux Vitek Inc., Hazelwood MO), supplemented by glucose, maltose, lactose, adonitol, xylose and sucrose oxidation / fermentation sugars and lysine and ornithine decarboxylases. Organisms that matched the phenotypic identification criteria of Henry and associates (16) were confirmed by species specific molecular probes**.** Epidemiologic typing of BCC was performed by randomly amplified polymorphic DNA (RAPD) fingerprinting as described in Mahenthiralingam and coworkers (18). Sequential isolates of BCC recovered from individual patients were typed whenever possible over a minimum period of one year to ensure consistency of colonization with a single strain type.

2.5.4 Sub-study data collection: Measurement of alpha-1-antitrypsin levels during a pulmonary exacerbation episode

The CF patients were recruited exclusively from the St. Paul's Hospital adult CF clinic. In total 31 consecutive patients from the CF clinic (mean age (±SEM) 27.5(1.1) years), who developed an acute pulmonary exacerbation were recruited to participate in the study. We measured α_1-AT levels during the acute phase and 2-3 months later during a stable phase. Details of the experimental procedures, risks and benefits involved were explained to subjects before obtaining written consent, which was approved by the Ethics Committee of the institution in this study (Appendix A).

CF patients with a history of liver disease, or liver/lung transplanted were excluded from the α_1-AT levels study. Patients from the St. Paul's Hospital CF clinic were recruited at the time of admission for an acute pulmonary exacerbation. These patients were characterized into two groups based on %predFEV_1 during clinical stability: mild/moderate pulmonary impairment (%predFEV_1>50% predicted), and severe pulmonary impairment (%predFEV_1≤50% predicted). For this group we calculated Schwachman-Kulczycki (S-K) (19) and Brasfield (20) scores during clinical stability.

S-K scores were calculated for clinical status at the time of testing, which represented stable (non-pulmonary exacerbation) clinical status. Radiographic data were obtained from the patients' medical charts, as was information on activity patterns, pulmonary and nutritional status for S-K scoring. From a maximal score of 100 (maximum 25 per category), points were deducted for the level or degree of:

- Inactivity, fatigue and non-participation in daily living.
- Pulmonary symptoms and finger clubbing.
- Growth and nutritional deficiencies.
- Chest radiographic abnormalities

S-K scoring was performed by me and the results were reviewed with Dr. Wilcox. The guidelines for scoring were followed as described by Schwachman and Kulczycki (19). Chest radiographs were scored utilizing the Brasfield clinical scoring system (20) by Dr.

P.G. Wilcox, who was blinded to the identity of the patient. The severity of lung disease was determined from a maximal score of 25, with points deducted for:

- Degree of hyperinflation.
- Peribronchial thickening.
- Nodular cystic structures (bronchiectasis).
- Areas of atelectasis or pneumonia.
- Assessment of overall severity.

We also obtained from the patients' medical charts the number of days treated for pulmonary infections over a 2-year period (number of hospitalization days). Blood samples for measurement of α_1-AT were obtained within the first two days of hospital admission and repeated at days 4, 7, 10 and 13 of a 14-day therapeutic intervention for most patients. Stable α_1-AT levels were measured 2-3 months post-exacerbation during clinical stability. These patients were also included in the larger study and were genotyped for the α_1-AT polymorphisms.

2.6 GENETIC ANALYSES

2.6.1 Haplotype construction

To calculated haplotypes we used PHASE version 1.0.1. PHASE is a Linux based program for reconstruction of haplotypes (5). The mean error is half that obtained by the expectation-maximization (EM) algorithm. We used this program to infer haplotypes for the 4 MBL gene polymorphisms and for the pulmonary surfactant polymorphisms investigated.

2.6.2 Calculation of linkage disequilibrium and Hardy Weinberg equilibrium

Linkage disequilibrium and Hardy Weinberg equilibrium were calculated using Arlequin version 2.0 software package (21). We used the program PHASE version 2 (5, 6) to infer haplotypes probabilities. We used the chi-square statistic to test for linkage disequilibrium between the MBL and pulmonary surfactant genes.

2.7 STATISTICAL ANALYSIS

Specific statistical analyses used to investigate the proposed modifier genes are explained in detail in section 2.5 for each modifier gene, or family of genes investigated. We used univariate analysis of variance (ANOVA) for group comparisons. Cox Regression was used for survival analysis. Data analysis was performed using SPSS statistical software (SPSS® statistical software (Chicago, IL)) and Splus (Insightful Corporation).General statistical methods are described in this section.

2.7.1 Description of outcome variables

The primary outcome variables used to characterize:

- Pulmonary disease severity was %predFEV_1.
- Pulmonary disease progression was %predFEV_1 over time; with time being the longitudinal interval (prospective and retrospective) collected on the study subjects for a maximum 10 year interval (or more for BCC infected CF patients).
- Survival was the end point of age of death or lung transplantation.
- First and chronic infection with *P. aeruginosa* was age of first and chronic infection with the pathogen *P. aeruginosa*.
- Frequency of pulmonary infections was the frequency of encounters defined as pulmonary infections that were treated with intravenous antibiotics in the longitudinal data set for set time intervals investigated for statistical analyses (2 and 5 years) as described in section 2.7.6.

Forced expiratory volume in one second (FEV_1) was performed in accordance with American Thoracic Society criteria {Anonymous., 1995 #365}. FEV_1 was expressed as percent predicted values based on height, age and gender and calculated from equations derived for adults {Crapo, 1981 #363} and children {Hibbert, 1989 #344}. Pulmonary disease progression was defined as %predFEV_1 over time; with time being the longitudinal interval (prospective and retrospective) collected on the study subjects. Pulmonary function data from clinical encounters coded as stable were used for statistical analyses. A maximum 10 year interval of clinical data was collected for controls used for BCC infection study (N=196) and 5-10 years of clinical data was collected for controls used in the MBL2 study (N=428). For the BCC infected cohort, clinical data was collected retrospectively and prospectively. Specifically for the BCC patients, FEV_1 was retrospectively collected for 2-5

years pre-colonization with BCC and data post-acquisition of BCC were collected to July 2004 (or to time of death/lung transplantation). There were 90 (transient / chronic infection N=20/70) and 39 (chronic infection) CF patients infected with BCC that comprised the study group for the BCC infection and MBL2 study, respectively.

2.7.2 Description of independent variables

CFTR grouping. We used three different groupings for CFTR mutations. In Chapter 3, the CFTR grouping for the α_1-AT gene polymorphisms was fairly simplistic (and related to the commonly occurring CFTR mutation deltaF508) as was the case with studies being published at the time. Chapter 3 presents the α_1-AT gene polymorphism study as published in the journal American Journal of Respiratory and Critical Care Medicine in 1998 (22). During the duration of the thesis study further research in CFTR mutations lead to the categorization of CFTR in the literature into grouping based on the functional affect of the mutations and our groupings for CFTR have incorporated this knowledge. There is still no recognized or established grouping for presenting CFTR mutation classes in research studies.

CFTR genotype was coded as follows for the investigation of the α_1-AT gene:

- Homozygous ΔF508,
- Heterozygous ΔF508 (i.e., ΔF508/ other)
- Other.

Other included all other CFTR mutations; known and unknown. For subsequent genes investigated in this study, the CFTR gene was coded based on the CFTR class of the mutation. The following categories were developed based on the severity of the CFTR mutation:

- CF class homozygous severe: Class 1, 2, or 3 mutations on both chromosomes (homozygous for class 1, 2 or 3 mutations).
- CF class mild: Class 4 or 5 on one chromosome and either:
    - Class 4 or 5 on the second chromosome.
    - Class 1, 2, and 3 on the second chromosome.
    - Unknown or unclassified mutation on the second chromosome.

- CF class heterozygous severe: Class 1, 2, or 3 on one chromosome and unknown or unclassified mutation on the second chromosome.
- CF class unknown/unclassified: Unknown or unclassified mutations on both chromosomes.

When we analyzed the GST and GCLC gene polymorphisms (chapter 5) the grouping CFTR was collapsed into 2 groups in some cases. This secondary CFTR grouping was used where there was a small number of subjects when grouped by GCLC grouping and mild CFTR which otherwise caused estimation problems with the statistical models. In this grouping we excluded subjects who were classified as CF class unknown/unclassified for two reasons: a) small sample size and b) the unknown nature of the mutation(s) prohibited us from including this group with either the severe or mild CFTR groups described below. The following two CFTR groups were created:

- CF class severe: Included CF subjects classified above as CF class homozygous and heterozygous severe.
- CF class mild: Same grouping as defined above (see 2).

Pancreatic sufficiency status. In some cases we were unable to use CFTR class grouping in our statistical models as was the case in chapter 4 for BCC related analyses. In this particular case we found that all BCC infected CF patients carried severe CFTR mutations. In this case we also used the variable pancreatic sufficiency status as a measure of disease severity. CF patients were categorized as pancreatic:

- Insufficient: The exocrine pancreas was affected and CF patients were required to take fat metabolizing enzymes orally,
- Sufficient. The exocrine pancreas was functioning properly.

***P. aeruginosa* and BCC infection.**Categorical variables described the infection status of CF patients for *P. aeruginosa* and BCC. Specifically CF patients were categorized independently for these two pathogens initially into the following four categories:

- Not infected
- Infection with the pathogen only on one encounter.
- Transient (Short-term/sporadic) infection: the CF patient showed infection with the pathogen on more than one occasion but was limited to infection with the pathogen for less than 6 months and thereafter did not show growth of the pathogen.
- Chronic infection with the pathogen.

The grouping was also collapsed for use of the variables in our models into three categories:

- Not infected
- Transient
- Chronic

The age (day-month-year) of first and chronic infection with the pathogens were also collected. In the case of BCC infection the pathogen was further characterized by genomovar grouping and RAPD type (see Table 2.3).

2.7.3 Analysis of pulmonary disease progression

Mixed effects linear regression models were used to model the effect of the candidate modifier gene genotype or haplotype on our primary outcome variable which we defined as pulmonary disease progression, which is %predFEV$_1$ over the longitudinal interval. Pulmonary disease severity was defined as the mean %predFEV$_1$ over the entire longitudinal interval. Parameters also used as independent variables in our equations were current age, sex (categorized as male (0), female (1)), age of CF diagnosis, CFTR genotype or pancreatic sufficiency status (categorized as insufficient (0) and sufficient (1)), center code (categorized as Vancouver (1), Montreal (2), Hamilton (3), Toronto (4), Victoria (5), and Seattle, Washington, U.S.A (6)), body mass index (BMI) and infection with *Pseudomonas aeruginosa*. The latter was used as a categorical variable (*Pseudomonas aeruginosa* (PA) status as infected (1) and not infected (0)) in the mixed effects regression models.

The models investigated for each gene polymorphism are presented in Tables 2.6 -2.12.

α_1-AT gene. We investigated 3 polymorphisms in two independent analyses for the α_1-AT gene. The S and Z polymorphisms were analyzed in one model with α_1-AT deficiency defined by the genotypes MS, SS and MZ. The base group was MM. The 3 prime mutation in the untranslated region was investigated in the second model and the grouping was GG versus GA and AA. We investigated whether there was a difference in the rate of decline in %predFEV$_1$ (2-year interval). The models and variables used in the analyses are presented in Table 2.6.

MBL2 gene. We investigated four polymorphisms in the MBL2 gene. Subjects were grouped for MBL polymorphisms based on the functional effect of the polymorphisms as described in Garred and associates (23) into two groups; deficient and wild-type. We investigated whether there was a difference in the rate of decline in %predFEV$_1$ with *P. aeruginosa* infection controlling for CFTR genotype, *P. aeruginosa* infection status and gender. We next investigated whether there was a difference in the rate of decline in %predFEV$_1$ with BCC infection; whether MBL deficiency and chronic infection with BCC were associated with worse pulmonary disease progression in CF. The pre and post-acquisition interval for BCC infection were included in the models as fixed and random effects. Patients chronically infected with BCC regardless of genomovar group were grouped together, based on our findings that the rate of decline in %predFEV$_1$ with BCC infection was similar regardless of BCC genomovar. The models and variables used in the analyses are presented in Table 2.7.

SPA-1 gene. We investigated the SPA-1 (C655T) polymorphism. The TT genotype was not represented in our study cohort. We investigated whether having the less common CT versus CC genotype was associated with a different rate decline in %predFEV$_1$. The models and variables used in the analyses are presented in Table 2.8.

SPA-2 gene. Inferred haplotypes generated by the PHASE program for the two polymorphisms (A26C and A667C) studied in this gene were used to group the study cohort into three groups based on whether there were zero (SPA-2CA0), one (SPA-2CA1), or two (SPA-2CA2) copies of the inferred haplotype CA and investigate whether there was a difference in the rate of decline in %predFEV$_1$. The base group in our models was the most common diplotype (i.e., two copies of CA: SPA-2CA2). The models and variables used in the analyses are presented in Table 2.9.

SPD gene. The grouping for the gene polymorphism was based on genotype into three groups: CC, TT and CT. We first examined the rate of decline in %predFEV$_1$ based on our SPD genotype grouping. In a reduced model we combined the CT and CC group, that is having one or two copies of the polymorphism versus being homozygous for the common

allele TT, and examined the rate of decline in %predFEV$_1$. The models and variables used in the analyses are presented in Table 2.10.

GST genes and GCLC gene. We investigated whether having the gene deletion or not for GSTM1 and T1 were associated with a different rate of decline in %predFEV$_1$. The grouping for the gene deletion was whether subjects had one or two gene deletions for GSTMI and T1. In the case of the GSTP1 analyses our grouping was whether having one or two G-alleles for the GSTP1 (Ile105Val) polymorphism were associated with different rates of decline in %predFEV$_1$. For the GCLC gene we elucidated the number of GAC repeats in our study population and then investigated the commonly occurring ones. Our grouping for GCLC was:

- Homozygosity for 7 GAC repeats (i.e., GCLC7/7),
- 7/8 GAC repeats (i.e., GCLC7/8),
- 7/9 GAC repeats (i.e., GCLC7/9), and
- Greater than 7 GAC repeats on both chromosomes (i.e., GCLCgr7; includes 8/8, 8/9 and 9/9 GAC repeats).

Genotypes 6/9 and 7/10 GAC repeats which were also observed in our study cohort were rare and were not used in our analyses. We were unable to investigate *P. aeruginosa* infection status and the GST and GCLC polymorphisms by CFTR genotype severity in all cases due to the small number of patients who were carriers of mild CFTR mutations and who were also not infected chronically with *P. aeruginosa*, which caused estimation problems with the models. The models and variables used in the analyses are presented in Table 2.11.

MUC2 and MUC5B genes. CF patients were categorized into 2 groups for the MUC2 polymorphism: a) heterozygous / homozygous for 2 repeats and b) the common genotype in the cohort (homozygous for 1 repeat). CF patients were grouped based on the common genotype for the cohort for the MUC5B polymorphism into two groups: a) common genotype, which was homozygosity for the 7 repeat allele and b) all others. Lastly, we examined the two genes together and categorized our cohort into two groups; the common diplotype observed (MUC2-MUC5B/ MUC2-MUC5B=1-7/1-7, N=162) and all others (MUC2-MUC5B/ MUC2-MUC5B=1-3/1-7, 1-5/1-7, 1-7/1-8, 2-7/1-7, N=110). CFTR genotype was not included in any of our statistical models as it was not well represented in

the mucin gene groupings. We first examined the rate of decline in %predFEV$_1$ based on our individual gene groupings and then for the diplotype grouping. The models and variables used in the analyses are presented in Table 2.12.

2.7.4 Survival

Cox proportional hazards regression was used to investigate survival. Models used for each gene are described in the respective chapter and tables. In our survival models:

Dependent variables were: The time to event (death or lung transplantation (coded 0/1=alive/deceased or lung transplanted) and current age or age of event. Current age was used for CF patients who were still alive and age of event was used for deceased and lung transplanted CF patients.

Main effects were: Modifier gene SNP(s), CFTR class. *P. aeruginosa* (PA) infection status was also used in select models. Other variables also included in select models were: CF diagnosis age, current %predFEV$_1$, BMI and age.

Interactions were: Modifier gene* CFTR class, Modifier gene* PA infection status, Modifier gene* CF diagnosis age.

Other variables used as main effect and for interactions terms were:
PSS instead of CFTR class
BCC infection status and BCC related variables: genomovar group, RAPD group type, and BCC and *P. aeruginosa* co-infection.

2.7.5 Age of first infection and chronic infection with *P. aeruginosa*

Cox proportional hazards regression was used to investigate age of first infection and chronic infection with *P. aeruginosa*. In our models investigating age of first infection (and chronic infection) with *P. aeruginosa*):

Dependent variables were: Age of 1st infection (age of chronic infection), categorical variable PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class + CF diagnosis age + current %predFEV$_1$ + current BMI + current age.

Interactions: Modifier gene* CF diagnosis age, Gene * current %predFEV$_1$.

2.7.6 Pulmonary infections requiring therapy with intravenous antibiotics

The Poisson regression model was used to investigate differences in the frequency of pulmonary infections requiring IV therapy and our candidate modifier genes. The response variable was the number of pulmonary infections requiring IV therapy over time. We used a fixed time point from our longitudinal data collection file of January 1, 2000. To ensure that each patient included in the analysis data set was alive or that we had knowledge of the patient's clinical status (dead or lost to follow-up) we determined that there were clinical data available on the patient after the fixed time point of January 1, 2000. We investigated two longitudinal time intervals:

- Two years length of follow-up (January 1, 1998-January 1, 2000).
- Up to 5 years length of follow-up (January 1, 1995-January 1, 2000).

The total number of subjects who had the full 2 year length of follow-up was 257. For the 5-year interval, cases who met the inclusion criterion for interval A were used. Cases who met inclusion criteria for A were included for analysis of interval B. In our analysis we adjusted the length of follow-up for cases that had less than 5 years of data to the length of follow-up to reflect the shorter interval available.

2.8 HYPOTHESES TESTED

α_1-AT gene

Pulmonary disease severity and progression

1. We hypothesized that heterozygosity for the Z and S alleles, or homozygosity for the S allele of α_1-AT would be associated with more severe pulmonary disease severity and progression in CF.
2. We hypothesized that heterozygosity or homozygosity for the A allele for the 3' $G_{1237} \rightarrow A$ polymorphism of α_1-AT would be associated with more severe pulmonary disease severity and progression in CF.

Frequency of pulmonary infections

3. We hypothesized that MZ, MS, SS genotype for α_1-AT would be associated with a higher frequency of pulmonary infections requiring IV antibiotics over the 2-year interval followed in the CF cohort.
4. We hypothesized that heterozygosity or homozygosity for the A allele for the 3' $G_{1237} \rightarrow A$ polymorphism of α_1-AT would be associated with a higher frequency of pulmonary infections requiring IV antibiotics over the 2-year interval followed in the CF cohort

MBL2 gene

Pulmonary disease severity and progression

1. We hypothesized that CF patients who have an MBL2 deficient genotype would show more severe pulmonary disease severity and progression (steeper decline) than CF patients with a MBL2 wild-type genotype and this association would be further exaggerated when also controlling for *P. aeruginosa* infection and CFTR genotype. That is the rate of decline in %predFEV_1 would be:

$$\text{MBL2 deficient} > \text{MBL2 wild-type}$$

2. We hypothesized that the changes in pulmonary disease severity and progression would occur and be more evident in statistical analyses when distinguishing the time of BCC acquisition and specifically we hypothesized that:
    a. There would be similar pulmonary disease progression (i.e., rate of decline in pulmonary function) during the pre-acquisition with BCC interval for CF patients who later become chronically infected with BCC compared with non-BCC infected CF patients. That is for the 2-year pre-acquisition with BCC interval we hypothesized that the rate of decline in pulmonary function would be:

$$\text{MBL2 deficient} \neq \text{MBL2 wild-type}$$

$$BCC_{\text{MBL2deficient}} = \text{No } BCC_{\text{MBL2deficient}}$$

$$BCC_{\text{MBL2wild-type}} = \text{No } BCC_{\text{MBL2wild-type}}$$

b. Once infected with BCC that CF patients with a deficient MBL genotype would show more severe pulmonary disease severity and progression in the post-acquisition interval than CF patients with a: a) MBL2 wild-type genotype who are chronically infected with BCC and b) MBL2 deficient genotype not infected with BCC. That is the rate of decline in %predFEV_1 would be:

$$BCC_{MBL2deficient} > BCC_{MBL2wild\text{-}type} \text{ and}$$
$$BCC_{MBL2deficient} > \text{No } BCC_{MBL2deficient}$$

Frequency of pulmonary infections

3. We hypothesized that having an MBL2 deficient genotype compared with MBL2 wild-type would be associated with a different frequency of pulmonary infections requiring IV antibiotics over the 2-year and 5-year interval followed in the CF cohort.

Survival (death or lung transplantation)

4. We hypothesized that CF patients with an MBL2 deficient genotype would show worse outcome than MBL wild-type CF patients.
5. We hypothesized that CF patients with an MBL2 deficient genotype who were also chronically infected BCC would show worse outcome.
6. We also hypothesized that CF patients infected chronically with both pathogens (BCC and *P. aeruginosa*) would be more likely to experience an event than CF patients infected with only one of the pathogens or neither of the pathogens.

Susceptibility to BCC and *P. aeruginosa* infection

7. We hypothesized that MBL genotype is not directly associated with susceptibility to BCC chronic infection.
8. We hypothesized that MBL2 genotype is not associated with a different age of first infection and chronic infection with *P. aeruginosa*.

BCC infection, Genomovar and RAPD type grouping

Pulmonary disease severity and progression

1 We hypothesized that infection with BCC would be associated with more severe pulmonary disease severity and steeper pulmonary disease progression compared to CF patients not infected with BCC and specifically that the decline in %predFEV_1 over time would be:

Chronic BCC ≠ Controls (i.e., non BCC infected)

Transient BCC ≠ Controls

Chronic BCC ≠ Transient BCC

2 We hypothesized that chronic infection with BCC and *P. aeruginosa* would be associated with more severe pulmonary disease severity and steeper pulmonary disease progression compared to chronic infection with only one pathogen or infection with neither pathogen. Specifically:

Chronic BCC > Controls (i.e., non BCC and *P. aeruginosa* infected)

Chronic *P. aeruginosa* > Controls (not infected with *P. aeruginosa*)

Chronic BCC and *P. aeruginosa* > Chronic BCC or Chronic *P. aeruginosa* ≠ Controls

3 We hypothesized that infection with BCC genomovar 2 versus BCC genomovar 4 would be associated with different pulmonary disease severity and pulmonary disease progression and specifically:

BCC genomovar 2 ≠ Non-BCC infected

BCC genomovar 4 ≠ Non-BCC infected

BCC genomovar 2 ≠ BCC genomovar 4

4 We hypothesized that infection with BCC genomovar 4 RAPD-type 2 versus BCC genomovar 4 RAPD-type 1,4,6 would be associated with different pulmonary disease severity and pulmonary disease progression.

Survival (death or lung transplantation)

5 We hypothesized that there would be a different outcome (i.e., rapid deterioration to the event that is death or requiring lung transplantation) associated with CF patients who were either transiently or chronically infected with BCC or not infected with the pathogen.

6 We hypothesized that there would be a different outcome in CF patients associated with chronic BCC infection when also considering co-infection in these patients with *P. aeruginosa* (PA). Specifically:

Chronic BCC and PA ≠ Chronic BCC and
No PA ≠ No BCC or PA chronic infection

7 We hypothesized that being chronically infected with BCC genomovar 2 versus BCC genomovar 4 would show a different outcome.

8 We hypothesized that infection with BCC genomovar 4 RAPD-type 2 versus BCC genomovar 4 RAPD-type 1, 4, and 6 would show a different outcome.

SPA-1 gene

Pulmonary disease severity and progression

1. We hypothesized that having the less common CT versus CC genotype for the SPA-1 polymorphism would be associated with different pulmonary disease severity and pulmonary disease progression in CF.
2. We hypothesized that having the less common CT versus CC genotype for the SPA-1 polymorphism and chronic infection with *P. aeruginosa* would be associated with different pulmonary disease severity and pulmonary disease progression in CF.

Susceptibility to *P. aeruginosa* infection

3. We hypothesized that having the less common CT versus CC genotype for the SPA-1 polymorphism would be associated with a different age of first infection and chronic infection with *P. aeruginosa*.

Frequency of pulmonary infections

4. We hypothesized that having the less common CT versus CC genotype for the SPA-1 polymorphism would be associated with a different frequency of pulmonary infections requiring IV antibiotics over the 2-year and 5-year interval followed in the CF cohort.

Survival (death or lung transplantation)

5. We hypothesized that having the less common CT versus CC genotype for the SPA-1 polymorphism and chronic infection with *P. aeruginosa* would show a different outcome (i.e., rapid deterioration to the event that is death or requiring lung transplantation).

SPA-2 gene

Pulmonary disease severity and progression

1. We hypothesized that zero versus one versus two copies of the inferred haplotype CA for the SPA-2 gene would be associated with different pulmonary disease severity and pulmonary disease progression in CF.
2. We hypothesized that zero versus one versus two copies of the inferred haplotype CA for the SPA-2 gene and chronic infection with *P. aeruginosa* would be associated with different pulmonary disease severity and pulmonary disease progression in CF.

Susceptibility to *P. aeruginosa* infection

3. We hypothesized that zero versus one versus two copies of the inferred haplotype CA for the SPA-2 gene would be associated with a different age of first infection and chronic infection with *P. aeruginosa*.

Frequency of pulmonary infections

4. We hypothesized that zero versus one versus two copies of the inferred haplotype CA for the SPA-2 gene would be associated with a different frequency of pulmonary infections requiring IV antibiotics over the 2-year and 5-year interval followed in the CF cohort.

Survival (death or lung transplantation)

5. We hypothesized that zero versus one versus two copies of the inferred haplotype CA for the SPA-2 gene and chronic infection with *P. aeruginosa* would show different outcome (i.e., rapid deterioration to the event that is death or requiring lung transplantation).

SPD gene

Pulmonary disease severity and progression

1. We hypothesized that the three genotypes for the SPD polymorphism would be associated with different pulmonary disease severity and pulmonary disease progression in CF.
2. We hypothesized that the three genotypes for the SPD polymorphism and chronic infection with *P. aeruginosa* would be associated with a different pulmonary disease severity and pulmonary disease progression in CF.

Susceptibility to *P. aeruginosa* infection

3. We hypothesized that the three genotypes for the SPD polymorphism would show different susceptibility to first infection and chronic infection with *P. aeruginosa*.

Frequency of pulmonary infections

4. We hypothesized that the three genotypes for the SPD polymorphism would be associated with a different frequency of pulmonary infections requiring IV antibiotics over the 2-year and 5-year interval followed in the CF cohort.

Survival (death or lung transplantation)

5. We hypothesized that the three genotypes for the SPD polymorphism and chronic infection with *P. aeruginosa* would show a different outcome (i.e., rapid deterioration to the event that is death or requiring lung transplantation).

GST genes and GCLC gene

Pulmonary disease severity and progression

1. We hypothesized that CF patients with some CFTR function (i.e. homozygous or heterozygous for CFTR class IV or V mutations) and who do not have the deletion polymorphism for GSTM1 and GSTT1 will have better pulmonary function and a lower rate of decline in pulmonary function over time compared to CF patients who have the deletion polymorphisms for GSTM1 and GSTT1.
2. We hypothesized those CF patients who are homozygous or heterozygous for isoleucine for the GSTP1 (Ile105Val) polymorphism and have some CFTR function (i.e., mild CFTR genotype) will show increased pulmonary disease severity and a steeper decline in pulmonary function over time compared to CF patients with a mild CFTR genotype who are homozygous for the valine polymorphism for the GSTP1 gene.

3. We hypothesized that GSTT1, GSTM1 and GSTP1 polymorphisms resulting in decreased levels of these enzymes, regardless of CFTR genotype will be associated with more severe pulmonary disease severity and progression.
4. We hypothesize that CF patients who have a lower number of GCLC GAC repeats will show worse pulmonary disease severity and a steeper decline in pulmonary function over time.
5. We hypothesized that these associations will be exaggerated in CF patients who have a severe CFTR genotype.

Susceptibility to pathogen (BCC or *P. aeruginosa*) infection

6. We hypothesized that GSTT1, GSTM1 and GSTP1 polymorphisms resulting in decreased levels of these enzymes, regardless of CFTR genotype will be associated with increased susceptibility to chronic infection with common CF respiratory pathogens *P. aeruginosa* and BCC.
7. We hypothesized that those CF patients who have a lower number of GCLC GAC repeats, regardless of CFTR genotype, will be associated with increased susceptibility to chronic infection with common CF respiratory pathogens.

Susceptibility to liver disease

8. The GSTT1 and GSTM1 deletion polymorphisms and GSTP1 (Ile105Val) polymorphism will be associated with susceptibility to CF liver disease.

MUC2 and MUC5B genes

Pulmonary disease severity and progression

1. We hypothesized that variation in the length of the tandem repeats of MUC2 (i.e., higher number of repeats) would be associated with more severe pulmonary disease severity and progression in CF.
2. We hypothesized that a higher number of repeats of the VNTR of the MUC5B gene would be associated with more severe pulmonary disease severity and progression in CF.

Susceptibility to pathogen (BCC or *P. aeruginosa*) infection

3. We hypothesized that the prevalence of chronic infection with *P. aeruginosa* or BCC is affected by MUC5B or MUC2 genotype. Specifically we hypothesized that:
4. Polymorphisms which may increase the viscosity of mucus such as higher number of repeats of the VNTR of the MUC5B polymorphism may contribute to earlier repeat

infection with common CF pathogens followed by earlier chronic colonization with *P. aeruginosa* or BCC.

5. A longer length of the tandem repeats of the MUC2 polymorphism may contribute to earlier repeat infection with common CF pathogens followed by earlier chronic colonization with *P. aeruginosa* or BCC.

Table 2.1. A description of participating clinics and contribution to sample size for modifier gene analysis in the study.

Center	AAT	MBL2	Surfactant	GSTM1/T1/P1	GCLC	MUC2/5B
Vancouver Adult	97	107	107	108/106/106	106	90/80
Vancouver Children's	97	122	122	122/55/55	69	64/10
Hamilton Health Sciences Adult clinic	46	44	45	45	45	43/45
Hamilton Health Sciences Children's clinic	45	47	47	45	45	43/44
Montreal Adult	146	145	146	146/143/142	146	139/145
Victoria Adult	0	11	11	11	18	0
Victoria Children	0	9	9	9	11	0
Toronto Adult	187	0	0	0	0	0
Toronto Children	106	0	0	0	0	0
Seattle Adult	0	0	0	101	101	0

Table 2.2. Summary of genes and polymorphisms studied and genotyping method utilized.

Gene	Polymorphisms	Genotyping method
AAT	S and Z	multiplex SDM / RFLP
AAT	3 prime (G_{1237}→A)	PCR/RFLP
MBL2	B and C allele	PCR-RFLP
MBL2	D allele	SDM PCR
MBL2-X/Y	Promoter	ASP
SPA-1	Arg219Trp (C655T)	SDM PCR/RFLP
SPA-2	Thr9Asn (A26C)	SDM PCR/RFLP
SPA-2	Lys223Gln (A667C)	SDM PCR/RFLP
SPD	11Thr/Met (32C/T)	SDM PCR/RFLP
GSTP1	Ile105Val (A313G)	PCR/RFLP
GSTM1	Gene deletion	multiplex PCR
GSTT1	Gene deletion	multiplex PCR
GCLC	GAC repeat	PCR
MUC2	Imperfectly conserved repeats (Thr/Ser/Pro rich subdomain) upstream from 69 bp VNTR	PCR
MUC5B	59 bp/repeat in intron 36	PCR

Table 2.3. Demographic spreadsheet variables (Part 1): Clinical parameters.

Name of variable	Number of columns	Data values	Data type
ID code	1	Same value as for longitudinal spreadsheet	Continuous
Center code	1	To be entered by Vancouver center	Continuous
Date of birth	3	Day-month-year	Continuous
Date of CF diagnosis	3	Day-month-year	Continuous
CFTR genotype	2	One column for each allele	String
Sweat chloride level	1	Value in mmol/L	Continuous
Pancreatic function	1	0=insufficient 1=sufficient	Categorical
Sex	1	0=male 1=female	Categorical
Other complicating conditions		0=no disease 1=disease present 2=disease not looked for	
Diabetes mellitus	1	See row above	Categorical
Liver disease	1	See shaded row above	Categorical
GI reflux	1	See shaded row above	Categorical
Meconium ileus	1	0=no 1=yes	Categorical
Transplantation	1	0=no transplant 1=lung transplant 2=heart-lung transplant 3=liver transplant	Categorical
Date of transplant	3	Day-month-year	Continuous
Liver disease enzymes			
Aspartate aminotransferase (AST)	1	In μ/L	Continuous
Alanine aminotransferase (ALT)	1	In μ/L	Continuous
Gamma glutamyltransferase (GGT)	1	In μ/L	Continuous
Alkaline phosphatase (AP)	1	In μ/L	Continuous
Home postal code	1		String

Table 2.4. Demographic spreadsheet variables (Part 2): Pathogen infection parameters and current status (deceased/alive).

Name of variable	Number of columns	Data values	Data type
Date of 1st *P. aeruginosa* infection	3	Day-month-year	Continuous
Date of *P. aeruginosa* chronic infection*	3	Day-month-year	Continuous
P. aeruginosa infection status	1	0=no growth 1=one time growth 2=sporadic growth 3=chronic growth	Categorical
P. aeruginosa infection status	1	0=not chronically infected 1=chronically infected	Categorical
Burkholderia cepacia complex (BCC) infection status	1	0=no growth 1=one time or short-term growth 2=chronic growth	Categorical
BCC genomovar group	1	0=not infected 1=Genomovar I 2=Genomovar II (*B. multivorans*) 3.1= Genomovar IIIa 3.2= Genomovar IIIb 4= Genomovar IV 5= Genomovar V (*B. vietnamesis*)	Categorical
BCC RAPD* group	1	0=not infected RAPD type entered as word	String
Date of BCC colonization	3	Day-month-year	Continuous
Date of death	3	Day-month-year	Continuous
Primary cause of death	1	1= death due to end stage respiratory disease (hemoptysis, right heart failure) 2= death following transplantation 3= death due to BCC colonization 4= death due to non-CF cause (e.g. motor vehicle accident) 5= death due to CF but not respiratory related (liver disease, pancreatitis)	Continuous

** These dates may be the same as for 1st infection with the pathogen in many cases. More recently there has been more aggressive treatment to eradicate this pathogen when patients first grow it and therefore there are cases where the pathogen is successfully eradicated for up to a few years. An attempt was made to elucidate the date of first infection and date of chronic colonization.

*RAPD = Random Amplification of Polymorphic DNA

Table 2.5. Longitudinal data spreadsheet.

Name of variable	Number of columns	Data values	Data type
ID code	1		Continuous
Name or initials	1		String
Status at encounter	1	1=alive, 2=transplant, 3=deceased at this encounter, 4=death post-transplant, 5=BCC acquisition	Categorical
Date of encounter	3	Day-month-year	Continuous
Height (cm)	1	In cm	Continuous
Weight (kg)	1	In kg	Continuous
Pathogen infection at encounter		0=not present in sputum 1=present in sputum	
Staphylococcus aureus	1	0/1	Categorical
P. aeruginosa	1	0/1	Categorical
BCC	1	0/1	Categorical
H. influenza	1	0/1	Categorical
X. maltophilia	1	0/1	Categorical
Pulmonary function			
FEV1	1	In liters	Continuous
% predicted FEV1 (%)	1	In percent	Continuous
FVC in liters	1	In liters	Continuous
% predicted FVC (%)	1	In percent	Continuous
Hospital admission date	3	Day-month-year	Continuous
Hospital discharge date	3	Day-month-year	Continuous
Reason for encounter If coding is different than the one presented in column 3 then please supply us with your codes.	1	1= clinic visit, 2=hospitalization or home IV therapy for pulmonary infection, 3=clinic visit but ill 5=clinic visit, but on Home IV antibiotics, 6= hospitalization for other non-pulmonary complication, 9=specifics of visit / hospitalization not available.	Categorical

Table 2.6. Models for statistical analysis of pulmonary disease severity and progression of α_1-AT polymorphisms for Chapter 3.

Gene polymorphism	Base group in model	Model
S and Z *Taq*I promoter polymorphism in 3' region ($G_{1237}\rightarrow A$)	MM GG	α_1-AT coding: 1=MS, MZ, SS, or for the promoter polymorphism 1=GA and AA %predFEV$_1$= Time + α_1-AT genotype + Sex (0/1=male/female) + CF diagnosis age + CFTR genotype + BMI Alternatively using PSS instead of CFTR genotype %predFEV$_1$= Time + α_1-AT genotype + Sex (0/1=male/female) + CF diagnosis age + Current Age + Pancreatic sufficiency status + BMI Including PA infection status* %predFEV$_1$= Time + α_1-AT genotype + Sex + PA infection status + CF diagnosis age + Current Age +CFTR grouping + BMI %predFEV$_1$= Time + α_1-AT genotype + Sex + Current age + PA infection status + CF diagnosis age + BMI

* PA status- *P. aeruginosa* infection status

+ CFTR genotype is an abbreviation for the CFTR grouping into 3 groups (as described for α_1-AT analyses in section 2.4.1).

Table 2.7. Models for statistical analysis of pulmonary disease severity and progression of MBL2 gene polymorphisms and the effects of chronic BCC infection (on pre and post BCC acquisition) on pulmonary disease progression for chapter 4.

Gene polymorphisms	Base group in model	Model
MBL2	MBL deficient	Model 4.2.3-A %predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + PA infection status[†] (0/1=not infected/chronically infected) + CFTR genotype[+] + MBL2 deficiency * Time Including 2 and 3 way interactions of covariates. Model 4.2.3-B %predFEV$_1$ = Time + Sex + MBL2 deficiency + PA infection status + BCC infection status (0/1=not infected/chronically infected) + MBL2 deficiency * Time + PA infection status * Time + BCC infection status * Time + MBL2 deficiency * PA infection status * Time + MBL2 deficiency * BCC infection status * Time Model 4.2.3-C (pre-acquisition of BCC) %predFEV$_1$ = Time + Sex + MBL2 deficiency + PA infection status + BCC infection status + MBL2 deficiency * Time + PA infection status * Time + BCC infection status * Time Model 4.2.3-D %predFEV$_1$ = Time + Sex + MBL2 deficiency + BCC infection status + AgePP (0/(>0)=pre-colonization/days post-colonization) + MBL2 deficiency * AgePP).

[†] PA status- *P. aeruginosa* infection status

[+] CFTR genotype is an abbreviation for the CFTR class groupings into 4 groups (see 2.4.1).

Table 2.8. Models for statistical analysis of pulmonary disease severity and progression of SPA-1 gene polymorphism for chapter 4.

Gene polymorphism	Base group in model	Model
SPA-1	CC genotype	Model 4.2.3-E %predFEV$_1$= Time + Sex (0/1=male/female) + SPA-1 genotype (0/1=CC/CT) + SPA-1 genotype * Time). Model 4.2.3-F %predFEV$_1$= Time + Sex + SPA-1 genotype + PA infection status† (0/1=not infected/chronically infected) + SPA-1 genotype * PA infection status + SPA-1 genotype * Time + PA infection status * Time + SPA-1 genotype * PA infection status * Time Model 4.2.3-G %predFEV$_1$= Time + Sex + SPA-1 genotype + PA infection status + CFTR genotype^{+} + SPA-1 genotype * PA infection status + CFTR genotype * SPA-1 genotype + SPA-1 genotype * Time + PA infection status * Time + CFTR genotype^{+} * Time + SPA-1 genotype * PA infection status * Time + CFTR genotype * SPA-1 genotype * Time + CFTR genotype * SPA-1 genotype * PA infection status * Time

† PA status- *P. aeruginosa* infection status

$^{+}$ CFTR genotype is an abbreviation for the CFTR class groupings into 4 groups (see 2.4.1).

Table 2.9. Models for statistical analysis of pulmonary disease severity and progression of SPA-2 gene polymorphism for chapter 4.

Gene	Base group in model	Model
SPA-2	2 copies of CA.	Model 4.2.3-H $\%predFEV_1$ = Time + Sex (0/1=male/female) + SPA-2CA0 (0/1=1 or 2 copies of CA/0 copies of CA) + SPA-2CA1 (0/1=0 or 2 copies of CA/1 copy of CA) + SPA-2CA0 * Time + SPA-2CA1 * Time. Model 4.2.3-I $\%predFEV_1$ = Time + Sex (0/1=male/female) + SPA-2CA0 (0/1=1 or 2 copies of CA/0 copies of CA) + SPA-2CA1 (0/1=0 or 2 copies of CA/1 copy of CA) + PA infection status + SPA-2CA0 * PA infection status + SPA-2CA1 * PA infection status + SPA-2CA0 * Time + SPA-2CA1 * Time + PA infection status * Time + SPA-2CA0 * PA infection status * Time+ SPA-2CA1 * PA infection status * Time.

Table 2.10. Models for statistical analysis of pulmonary disease severity and progression of SPD gene polymorphism for chapter 4.

Gene	Base group in model	Model
SPD	TT genotype	Model 4.2.3-J %predFEV$_1$ = Time + Sex (0/1=male/female) + SPD-CC (0/1=CT or TT/CC) + SPD-CT (0/1=CC or TT/CT) + SPD-CC * Time + SPD-CT * Time. Model 4.2.3-K %predFEV$_1$ = Time + Sex + SPD-CC (0/1=CT or TT/CC) + SPD-CT (0/1=CC or TT/CT) + PA infection status + SPD-CC * PA infection status + SPD-CT * PA infection status + SPD-CC * Time + SPD-CT * Time + PA infection status * Time + SPD-CC * PA infection status * Time+ SPD-CT * PA infection status * Time. Reduced model: CT and CC genotypes combined Model 4.2.3-L %predFEV$_1$ = Time + Sex (0/1=male/female) + SPD-CC/CT (0=TT and 1=CC/CT) + SPD-CC/CT * Time.

Table 2.11. Models for statistical analysis of pulmonary disease severity and progression of GSTs and GCLC polymorphisms for chapter 5.

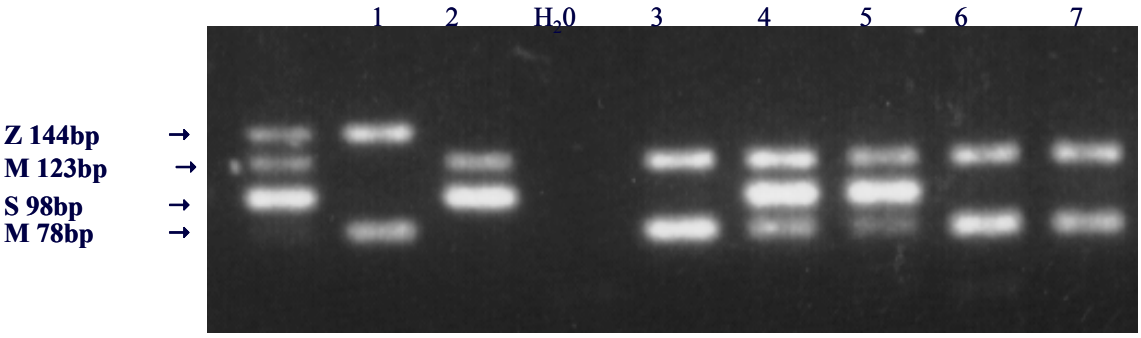
Gene	Base group in model	Model
GSTM1 and T1	Zero null in GSTM1 or T1	Longitudinal data set Model 5.2.3-A %predFEV$_1$ = Time + Sex (0/1=male/female) + 1 null in GSTM1 or T1 (0/1=0 or 2 null copies/1 null copy) + 2 null in GSTM1 and T1 (0/1=0 or 1 null copies/2 null copies) + 1 null in GSTM1 or T1 * Time + 2 null in GSTM1 or T1 * Time Cross-sectional data set Model 5.2.4-A Current %predFEV$_1$ = Current age (yrs) + Sex + CFTR genotype (0/1=severe/mild) + BMI + PA infection status (0/1=not infected /chronically infected) + 1 null in GSTM1 or T1 + 2 null in GSTM1 and T1
GSTP1	AA genotype	Model 5.2.3-B %predFEV$_1$ = Time + Sex (0/1=male/female) + GG genotype for GSTP1 (0/1=0 or 1 copy of G-allele/ GG) + AG genotype for GSTP1 (0/1=0 or 2 copies of G-allele/ AG) + GG genotype for GSTP1 * Time + AG genotype for GSTP1 * Time
GCLC	7/7 GAC repeats	Model 5.2.3-C %predFEV$_1$ = Time + Sex (0/1=male/female) + GCLC7/8 + GCLC7/9 + GCLCgr7 (0/1 1=genotypes with >7GAC repeats on both chromosomes) + GCLC7/8 * Time + GCLC7/9 * Time + GCLCgr7 * Time

Table 2.12. Models for statistical analysis of pulmonary disease severity and progression for MUC2 and MUC5B polymorphisms for chapter 6.

Gene	Base group in model	Model
MUC2	Heterozygous / homozygous for 2 repeats	Model 6.2.3-A %predFEV$_1$ = Time + MUC2 (0/1= homozygous for one repeat) + Sex (0/1=male/female) + MUC2 * Time Model 6.2.3-B %predFEV$_1$ = Time + MUC2 + MUC2 * Time + PA infection status (0/1=not infected /chronically infected) + MUC2 * Time + PA infection status * MUC2 + PA infection status * Time + PA infection status * MUC2 * Time
MUC5B	other than 7-7 VNTR genotype	Model 6.2.3-C %predFEV$_1$ = Time + MUC5B (0/1= 7-7 VNTR genotype) + Sex (0/1=male/female) + MUC5B * Time Model 6.2.3-D %predFEV$_1$ = Time + Sex + MUC5B + PA infection status (0/1=not infected /chronically infected) + MUC5B * Time + PA infection status * MUC5B + PA infection status * Time + PA infection status * MUC5B * Time
MUC2/ MUC5B diplotype	1-7/1-7 repeats	Model 6.2.3-E %predFEV$_1$ = Time + Sex (0/1=male/female) + MUC2/5B diplotype (0/1=1-7/1-7) + MUC2/5B diplotype * Time Model 6.2.3-F %predFEV$_1$ = Time + Sex (0/1=male/female) + MUC2/5B diplotype + PA infection status (0/1=not infected /chronically infected) + MUC2/5B * Time + PA infection status * MUC2/5B + PA infection status * Time + PA infection status * MUC2/5B * Time

Figure 2.1.α_1-AT S and Z alleles; *TaqI* digestion and visualization of PCR product on an agarose gel. On Gel A, positive controls are in columns 1-3 and column 4 is the negative control (i.e., no template DNA). Subjects in columns 4, 6 and 7 are homozygous wild-type for the Z and S alleles. Subjects in columns 4 and 5 are heterozygous for the S allele (i.e., MS MM). On Gel B, the subject in column 2 is heterozygous for the Z allele (MM MZ), the subject in column 9 is homozygous for the S allele (SS MM), and all other subjects are wild-type for both the S and Z polymorphisms.

A.



B.

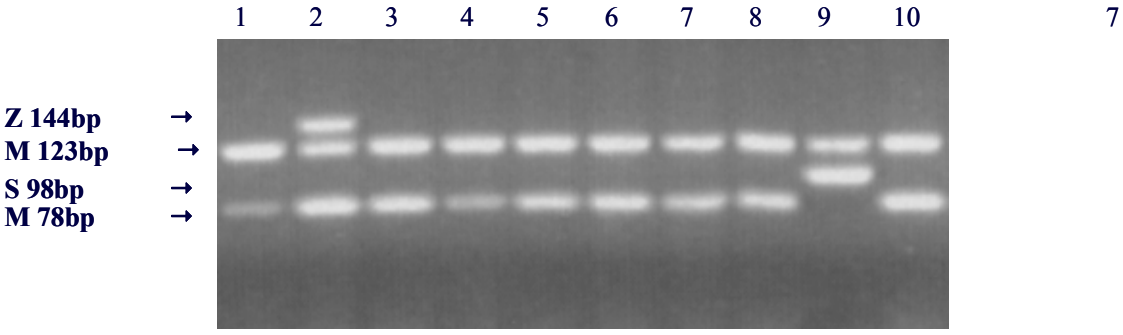
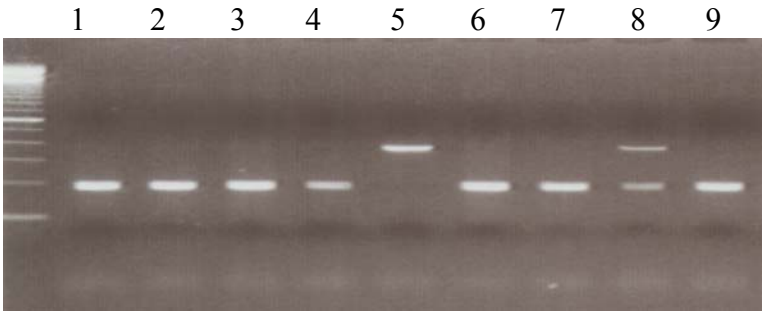


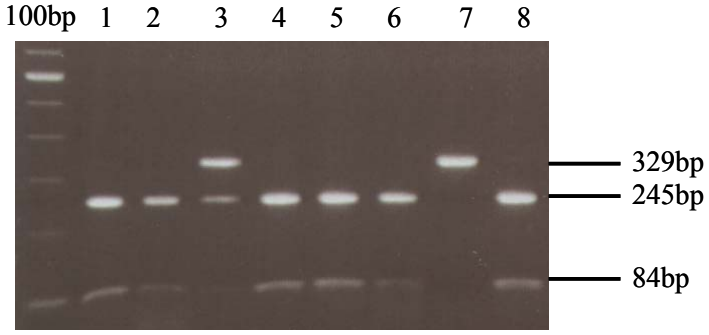
Figure 2.2. α_1-AT 3 prime polymorphism; *TaqI* digestion and visualization of the PCR product on an agarose gel. The *TaqI* enzyme cuts the wild-type G allele into two fragments (of size 191 and 182bp) and the A allele is left uncut (373 bp fragment). In the gel shown, subjects 1-4, 6, 7, 9 are homozygous for the G allele, subject 5 is homozygous for the A allele and subject 8 is heterozygous for the G allele.



← A allele (373 bp)
← G allele:
191 bp
182 bp

Figure 2.3. MBL2 gene B and C alleles; the B (codon 54) and C (codon 57) alleles for MBL2 gene were detected by restriction enzymes *BanI* and *MboII*, respectively. A. *Ban*I cleaves the A allele into two fragments (245 bp and 84 bp) and leaves the B allele undigested. Subjects 1, 2, 4-6 and 8 are homozygous wild-type (AA), subject 3 is heterozygous wild-type (AB) and subject 7 is homozygous for the B allele. B. *Mbo*II cleaves the C allele into two fragments (266bp and 63bp) and leaves the A allele undigested. Subjects 1 and 3-9 are homozygous wild-type (AA), subjects 2, 10 are heterozygous wild-type (AC).

A.



B.

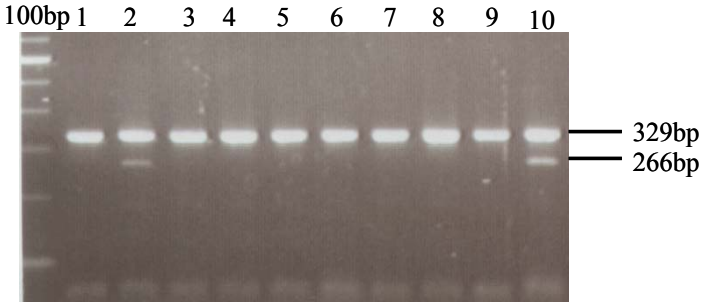
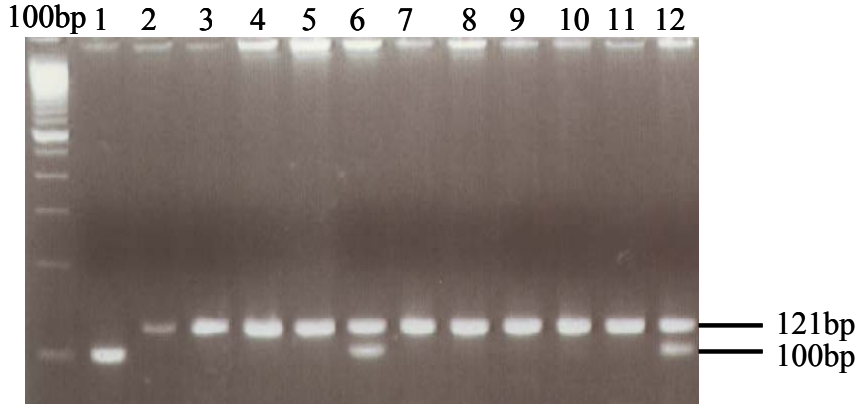


Figure 2.4. The D allele and the XY promoter polymorphisms for MBL2 gene. A. The *MluI* restriction enzyme site was introduced into the amplification product of the D allele (codon 52) by site-directed mutagenesis (SDM)-PCR; *MluI* cleaves D allele into two bands (21 bp and 100 bp) and leaves A allele uncut. Subject 1 is homozygous for the D allele, subjects 2-5, 7-11 are homozygous wild type (AA), subjects 6,12 are heterozygous wild-type (AD) and. B. The MBL promoter polymorphism was investigated by PCR using sequence-specific priming. β-globulin was used as a positive control. The presence of a band on the gel at 154bp indicates the X or Y allele was present. In this gel, subjects 2, 5, 7, 9, 10 have at least one X allele, whereas subjects 1, 3, 4, 6, and 8, who do not show a band at 154bp, are homozygous for the Y allele.

A.



B.

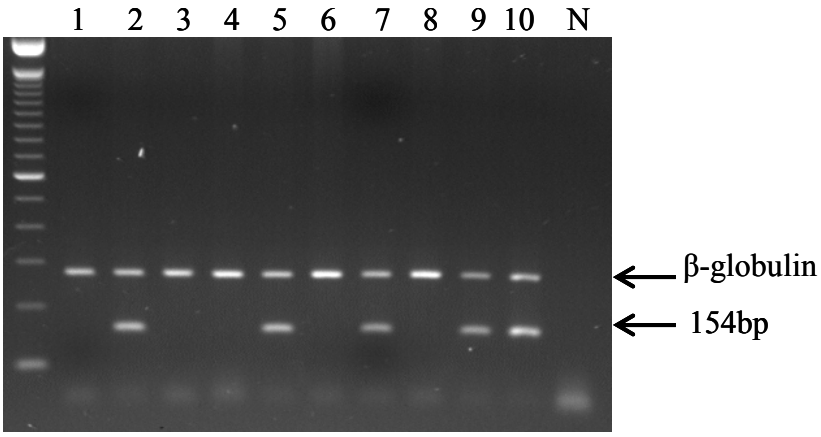


Figure 2.5. Site directed mutagenesis PCR/RFLP for the SPA-1 polymorphism. A 159bp region of the gene was amplified and the restriction enzyme *Taq*1 cuts the C allele into 2 fragments (138bp and 21bp) and leaves the T allele uncut (159bp). Subjects 1, 4, 6, 7 are heterozygous (CT) and subjects 2, 3, 5, 8-10 are homozygous for the C allele.

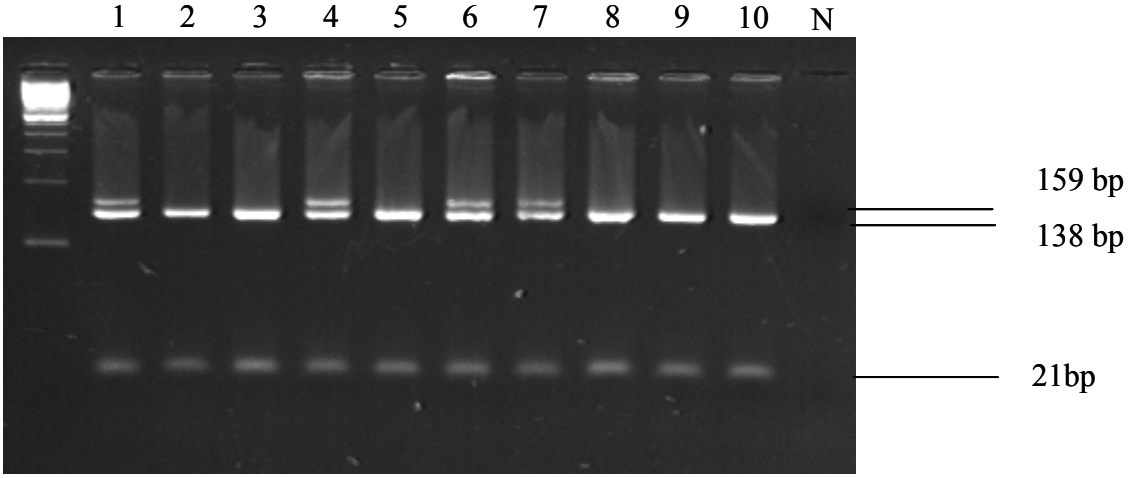
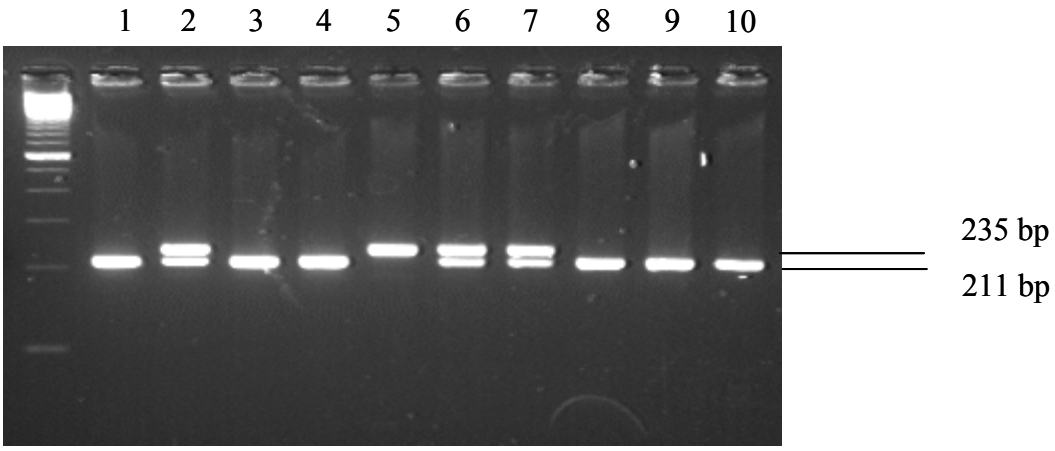


Figure 2.6. Exon 2 and 4 polymorphisms investigated in the SPA-2 gene. A. Site directed mutagenesis PCR/RFLP for SPA-2 (A667C) polymorphism. A 235bp region of exon 4 of the gene was amplified and the restriction enzyme *Hha*I cuts the C allele into 2 fragments (24bp and 211bp) and leaves the A allele uncut (235bp). Subjects 1,3, 4, 8-10 are homozygous for the C allele, subjects 2,6,7 are heterozygous (AC) and subject 5 is homozygous for the A allele. B. A 150bp region of the gene on exon 2 was amplified and the restriction enzyme *Mse*1 cuts the A allele into 2 fragments (18bp and 132bp) and leaves the C allele uncut (150bp). Subjects 1, 5 are homozygous for the A allele, subject 2, 4, and 6 are heterozygous (AC) and subject 3 is homozygous for the C allele.

A.



B.

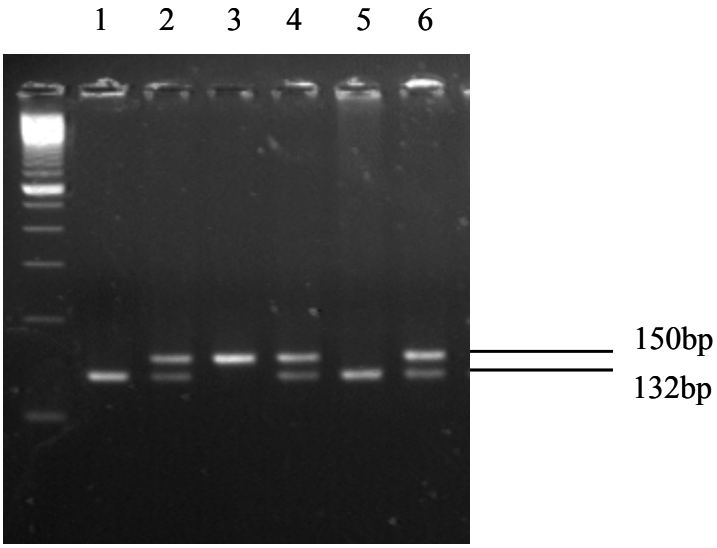


Figure 2.7. Site directed mutagenesis PCR/RFLP of SPD polymorphism in exon 4. A 235bp region of exon 4 of the gene was amplified and the restriction enzyme *Fsp*I cuts the T allele into 2 fragments (82bp and 19bp) and leaves the C allele uncut (101bp). Subjects 2, 6, and 7 are homozygous for the C allele, subjects 1 and 3-5 are heterozygous (CT) and subjects 8 and 9 are homozygous for the T allele.

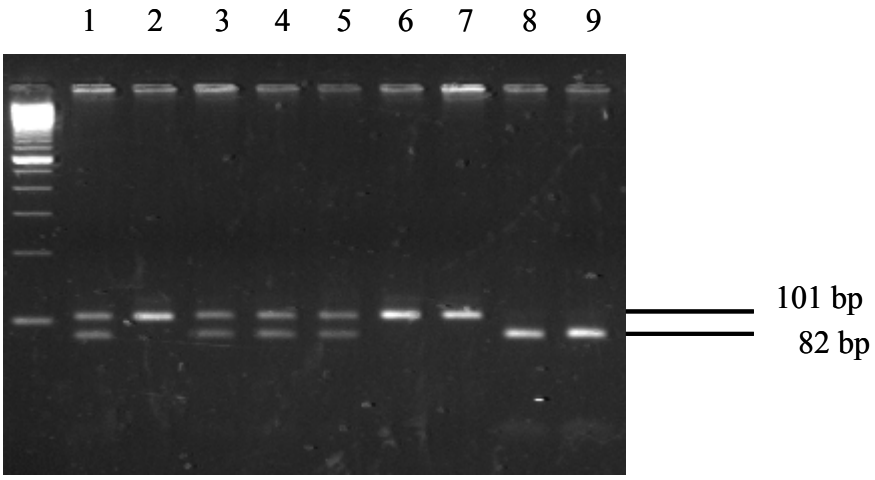


Figure 2.8. MUC2 polymorphism was examined by PCR using primers which amplified the repetitive threonine/serine/proline-rich subdomain. Subjects differ in the number of imperfectly conserved repeats which visualized on an agarose gel as differences in the size of the amplified region (subjects 1 and 3 are heterozygous for 2 long repeats, subjects 2 and 5 are homozygous for 2 long repeats and subject 4 is homozygous for medium repeat). N = negative control.

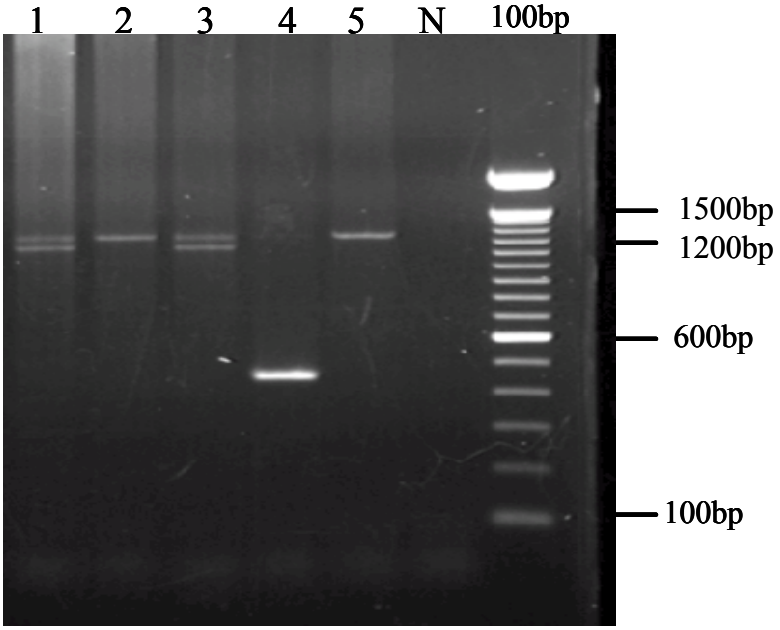


Figure 2.9. MUC5B polymorphism was examined by PCR using primers which amplified the repetitive threonine/serine/proline-rich subdomain. Subjects differ in the number of imperfectly conserved repeats which is visualized on an agarose gel as differences in the size of the amplified region. Subject 1 is homozygous for 7 repeats, subject 2 has 7/5 repeats, subject 3 has 3 repeats, subject 4 has 7/8 repeats, subject 5 has 5 repeats, subject 6 has 8 repeats, subject 7 has 7/6 repeats, subject 8 has 2/3 repeats, subject 9 has 7/4 repeats. N = negative control.

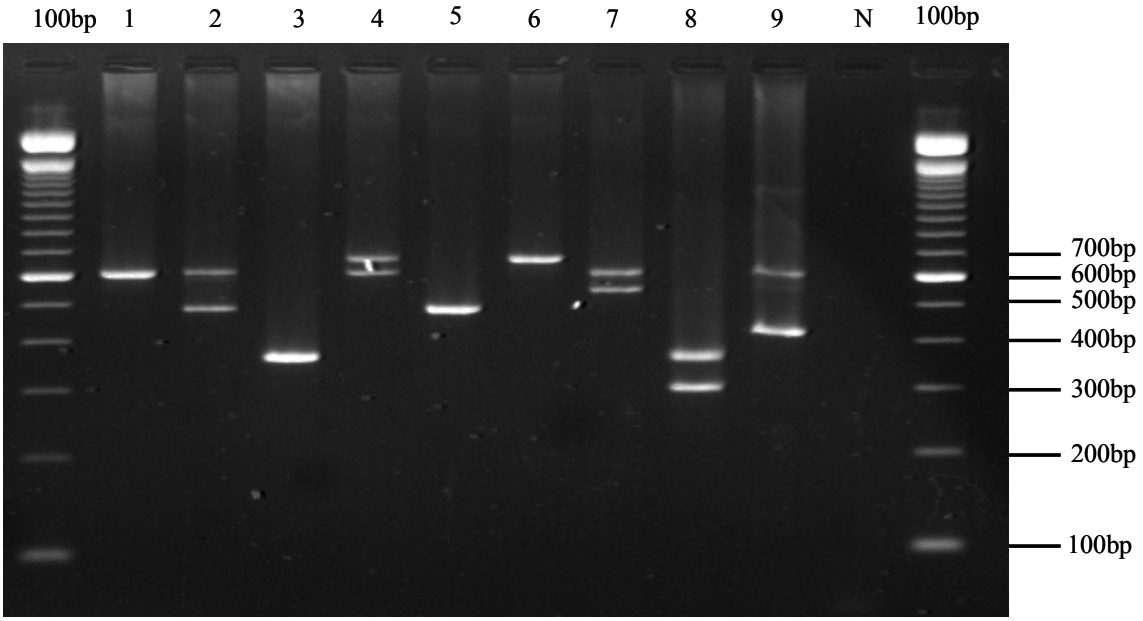
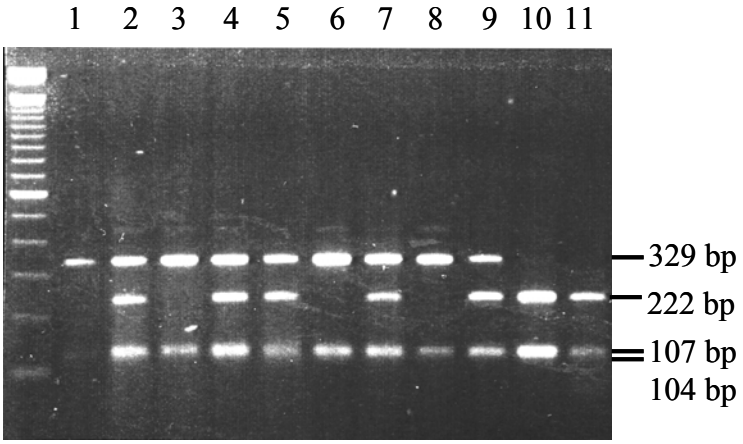
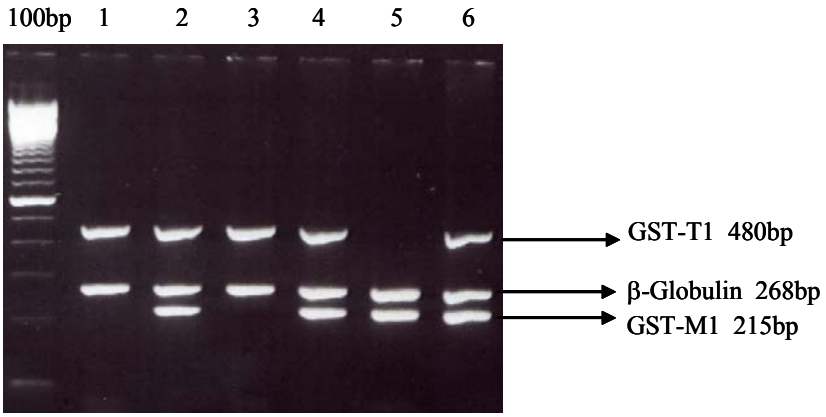


Figure 2.10. Polymorphisms investigated in GST genes P1, T1 and M1. A. PCR/RFLP of the GSTP1 polymorphism in exon 5 (A313G). A 433bp region of exon 5 of the gene was amplified and the restriction enzyme *Bsm*AI cut the A allele into 2 fragments (104bp and 329bp) and the G allele into three fragments (104bp, 222bp and 107bp). Subjects 1, 3, 6, 8 are homozygous for the A allele, subjects 2, 4-5, 7, 9 are heterozygous (AG) and subjects 10 and 11 are homozygous for the G allele. B. Multiplex PCR of the GSTM1 and GSTT1 deletions were examined by multiplex PCR using primers which amplified regions of exon 5 and 3. Genotype is shown in table beside the gel. Subjects 1 and 3 are null for GSTM1 and subject 5 is null for GSTT1.

A.



B.



	GST-T1	GST-M1
1	+	-
2	+	+
3	+	-
4	+	+
5	-	+
6	+	+

2.9 BIBLIOGRAPHY

1.Gum, J. R. J. 1992. Mucin genes and the proteins they encode: structure, diversity, and regulation. *Am J Respir Cell Mol Biol.* 7(6):557-64.

2.Pigny, P., V. Guyonnet-Duperat, A. S. Hill, W. S. Pratt, S. Galiegue-Zouitina, M. C. d'Hooge, A. Laine, I. Van-Seuningen, P. Degand, J. R. Gum, Y. S. Kim, D. M. Swallow, J. P. Aubert, and N. Porchet. 1996. Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 38(3):340-52.

3.Pigny, P., I. Van Seuningen, J. L. Desseyn, S. Nollet, N. Porchet, A. Laine, and J. P. Aubert. 1996. Identification of a 42-kDa nuclear factor (NF1-MUC5B) from HT-29 MTX cells that binds to the 3' region of human mucin gene MUC5B. *Biochemical & Biophysical Research Communications* 220(1):186-91.

4.McKone, E. F., J. Shao, D. D. Frangolias, C. L. Keener, C. A. Shephard, F. M. Farin, M. R. Tonelli, P. D. Pare, A. J. Sandford, M. L. Aitken, and T. J. Kavanagh. 2006. Variants in the glutamate-cysteine-ligase gene are associated with cystic fibrosis lung disease. *Am J Respir Crit Care Med* 174(4):415-9.

5.Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-89.

6.Stephens, M., and P. Donnelly. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73(5):1162-9.

7.Sandford, A. J., T. Chagani, J. J. Spinelli, and P. D. Pare. 1999. alpha1-antitrypsin genotypes and the acute-phase response to open heart surgery. *Am J Respir Crit Care Med* 159(5 Pt 1):1624-8.

8.Madsen, H. O., P. Garred, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, S. Thiel, and A. Svejgaard. 1994. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* 40(1):37-44.

9.Madsen, H. O., P. Garred, S. Thiel, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, and A. Svejgaard. 1995. Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol* 155(6):3013-20.

10.White, R. T., D. Damm, J. Miller, K. Spratt, J. Schilling, S. Hawgood, B. Benson, and B. Cordell. 1985. Isolation and characterization of the human pulmonary surfactant apoprotein gene. *Nature* 317(6035):361-3.

11.Katyal, S. L., G. Singh, and J. Locker. 1992. Characterization of a second human pulmonary surfactant-associated protein SP-A gene. *Am J Respir Cell Mol Biol* 6(4):446-52.

12.Yim, J. J., G. Y. Park, C. T. Lee, Y. W. Kim, S. K. Han, Y. S. Shim, and C. G. Yoo. 2000. Genetic susceptibility to chronic obstructive pulmonary disease in Koreans: combined analysis of polymorphic genotypes for microsomal epoxide hydrolase and glutathione S-transferase M1 and T1. *Thorax* 55(2):121-5.

13.Anonymous. 1995. Standardization of Spirometry, 1994 update. American Thoracic Society. *American Journal of Respiratory and Critical Care Medicine* 152(3):1107-36.

14.Crapo, R. O., A. H. Morris, and R. M. Gardner. 1981. Reference spirometric values using techniques and equipment that meet ATS recommendations. *Am Rev Respir Dis* 123(6):659-64.

15.Hibbert, M. E., A. Lannigan, L. I. Landau, and P. D. Phelan. 1989. Lung function values from a longitudinal study of healthy children and adolescents. *Pediatric Pulmonology* 7(2):101-9.

16.Henry, D. A., E. Mahenthiralingam, P. Vandamme, T. Coenye, and D. P. Speert. 2001. Phenotypic methods for determining genomovar status of the Burkholderia cepacia complex. *J Clin Microbiol* 39(3):1073-8.
17.Henry, D. A., M. E. Campbell, J. J. LiPuma, and D. P. Speert. 1997. Identification of Burkholderia cepacia isolates from patients with cystic fibrosis and use of a simple new selective medium. *J Clin Microbiol* 35(3):614-9.
18.Mahenthiralingam, E., M. E. Campbell, D. A. Henry, and D. P. Speert. 1996. Epidemiology of Burkholderia cepacia infection in patients with cystic fibrosis: analysis by randomly amplified polymorphic DNA fingerprinting. *J Clin Microbiol* 34(12):2914-20.
19.Shwachman, H., and L. Kulczycki. 1958. Long-term study of one hundred five patients with cystic fibrosis. *American Journal of Diseases of Children* 96:6-15.
20.Brasfield, D., G. Hicks, S. Soong, and R. E. Tiller. 1979. The chest roentgenogram in cystic fibrosis: a new scoring system. *Pediatrics* 63(1):24-9.
21.Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin: A software for population genetics data analysis., 2.000 ed. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva., Geneva.
22.Frangolias, D. D., J. Ruan, P. J. Wilcox, A. G. Davidson, L. T. Wong, Y. Berthiaume, R. Hennessey, A. Freitag, L. Pedder, M. Corey, N. Sweezey, J. Zielenski, E. Tullis, and A. J. Sandford. 2003. Alpha 1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *Am J Respir Cell Mol Biol* 29(3 Pt 1):390-6.
23.Garred, P., T. Pressler, H. O. Madsen, B. Frederiksen, A. Svejgaard, N. Hoiby, M. Schwartz, and C. Koch. 1999. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest* 104(4):431-7.

CHAPTER 3: ALPHA-1-ANTITRYPSIN DEFICIENCY ALLELES IN CYSTIC FIBROSIS LUNG DISEASE

A version of this chapter has been published and the citation is:

'Frangolias D.D., Ruan J., Wilcox P.G., Berthiaume Y., Davidson G., Hennessey R., Corey M., Tullis E., Zielenski J., Wilson W.M., Freitag A., Sandford A.. Alpha-1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *American Journal of Respiratory Cell and Molecular Biology*. 29:390-396, 2003.'

3.0. INTRODUCTION

Polymorphisms in the serine protease inhibitor gene, alpha-1-antitrypsin (α_1-AT) were investigated. Polymorphisms in the gene which have been shown to affect either the levels of the protein produced or regulate expression of the gene were investigated. In this study we investigated the associations of these polymorphisms on pulmonary disease severity and progression, survival and succeptibility to infection with common CF pathogens.

3.1. RATIONALE AND MAIN HYPOTHESIS

The primary pathophysiological processes responsible for premature death and disability in patients with cystic fibrosis (CF) are chronic pulmonary infection and inflammation. The inflammatory process in response to pulmonary infection in CF airways is characterized by a massive influx of neutrophils (1). Neutrophils represent less than 5% of the cells recovered in bronchoalveolar lavage fluid (BALF) in normal individuals but in adults and children (1-5 years of age) with CF, neutrophils may comprise up to 95% of the cell population (2). Neutrophils contain a number of proteolytic enzymes one of which, neutrophil elastase (NE), has been implicated in the excessive pulmonary damage observed in cigarette smokers and in patients with CF. Elevated levels of NE have been reported in the sputum of patients who have CF (3, 4). NE is capable of causing direct lung damage by hydrolyzing all the major connective tissue proteins that make up the lung and airway matrix. In addtition, excess NE adversely affects the airways in CF, by enhancing mucous secretion (5, 6), and by interfering with the opsonization and elimination of bacterial pathogens, particularly *P. aeruginosa* (5, 7). In normal hosts, the actions of NE are prevented primarily by alpha-1-antitrypsin (α_1-AT), a serine protease inhibitor that binds to NE and inhibits the breakdown of elastic tissue in the lung. Normal to elevated levels of α_1-AT have been reported in airway secretions (2) and plasma (2, 8) in patients who have CF. Elevated levels of α_1-AT have been reported during pulmonary infections in this patient population (9).

The extremely high levels of NE in the airways of CF patients clearly indicate that there is an imbalance between α_1-AT and elastase in the airways of patients with CF. Since great variation in disease severity and progression exists among CF patients possessing the same

cystic fibrosis transmembrane conductance regulator (CFTR) genotypes (10, 11) it is possible that genes other than the CFTR may contribute to pulmonary disease progression. If this were the case, individuals who have lower than normal levels of α_1-AT may be at increased risk for lung damage. Several mutations of the α_1-AT gene result in a deficiency of this antiprotease. There is also evidence that α_1-AT genotype influences the acute phase response (12). Findings to date are inconclusive concerning the role that α_1-AT may play in pulmonary disease progression in CF (13-16). The main limitation of these studies is their small sample sizes and therefore the high possibility of type 2 error (false negative).

The purpose of this study was twofold. First, to investigate whether the α_1-AT gene (the Z, S deficiency alleles and the 3' $G_{1237} \rightarrow A$ mutation) is a modifier of pulmonary disease progression in a large cohort of CF patients who were characterized by a heterogeneous severity of pulmonary disease? Second, we wanted to measure α_1-AT levels during a stable clinical phase and the acute increases during pulmonary exacerbations in an adult group of CF patients who had varying degrees of pulmonary dysfunction. The rationale was to characterize the acute phase response to pulmonary infection in this patient group to help tailor the possible future administration of anti-proteolytic therapeutic agents.

Our main hypothesis was that heterozygosity for the Z, S and/or the 3' $G_{1237} \rightarrow A$ alleles of α_1-AT would result in earlier onset of pulmonary disease, more rapid deterioration in pulmonary function and consequently more severe pulmonary dysfunction after controlling for other known predictors of pulmonary function decline.

3.2 RESULTS

Tables 3.1 and 3.2 show demographic and clinical measures for our study cohort stratified by α_1-AT genotype. Prevalences of the S and Z alleles were similar to expected frequencies (Z = 1-3% and S = 2-4%) in the normal Caucasian population(17-21). Heterozygosity for the 3' $G_{1237} \rightarrow A$ mutation has been documented from smaller studies to be 5-15% (22, 23). Figures 3.1 and 3.2 show %predFEV_1 and sample size by α_1-AT genotype. Z and S alleles of the α_1-AT gene did not predict pulmonary disease severity and were not included in the final regression equation (R^2adj=0.23, p=0.0001)

Model 3.2-A

*%predFEV_1=70.1 + 0.76*CF diagnosis age – 1.29*Exam Age +7.43*Pancreatic sufficiency status + 0.71*BMI*

Base group for α_1-AT was: MM genotype

The effect of the variable Center was included as a random effect; however, center to center variability did not reach statistical significance (was less than 1% of the total variability). To determine whether inclusion of patients < 18 years of age (who had less severe pulmonary disease) could reduce the power to detect a significant association with a genetic modifier, the analysis was repeated in the subset of patients over 18 years of age. This analysis confirmed the lack of association of the Z and S alleles with disease severity seen in the entire group (α_1-AT genotype p=0.96). To determine whether there was a specific interaction of α_1-AT genotype with ΔF508, which may have been obscured by the presence of other CFTR alleles, we also ran this model selecting only those who were homozygous ΔF508, however α_1-AT genotype was not a significant predictor.

In the second model we included the categorical variable *P. aeruginosa* infection status, available for 555 of the 714 study subjects. This model accounted for 25% of the variability in cross-sectional pulmonary disease severity (R^2adj=0.25, p=0.0001).

Model 3.2-B

*%predFEV$_1$=79.21 – 1.23*Exam Age – 9.56* P. aeruginosa infection status + 0.66*CF diagnosis age + 0.50*BMI*

Burkholderia cepacia complex (BBC) colonization status (i.e., not infected (0) versus chronic infection with the pathogen (1)) was available for 558 patients (from the Vancouver, Hamilton and Toronto centers) and was evaluated as a predictor of pulmonary disease severity. This parameter was not a significant predictor of pulmonary disease in our cohort. The Toronto cohort did show higher prevalence of chronic BCC infection (26% of their sample) compared to the Hamilton and Vancouver cohorts (9.5% and 10.2% of the study samples, respectively). Similar analyses were used to investigate the 3' $G_{1237}\rightarrow A$ mutation in the α_1-AT gene and results showed that this mutation also was not predictive of pulmonary disease severity. The base group used for this polymorphism in the mixed effects regression analysis was GG genotype.

We did not show increased prevalence of death or lung transplantation in CF patients who were homozygous or heterozygous for the Z, S or the 3' $G_{1237}\rightarrow A$ alleles versus wild type (Tables 3.1 and 3.2). CF patients who had died or who were lung transplant recipients were significantly older (mean (±SEM) = 28.6(±1.3) years) than CF patients who were still alive (21.3(±0.4) years; p=0.0001), although both groups had been diagnosed with CF at similar mean (±SEM) ages (3.7(±0.8) and 4.5(±0.3) years respectively, p=0.44). We showed similar frequency of pulmonary exacerbations and duration of IV therapy (over 12 months) in the α_1-AT genotypic groups (Tables 3.1 and 3.2). BCC colonization was not more prevalent in those CF patients who were carriers of the Z, S or 3' alleles (data not shown). Liver disease status was available for 553 of the study subjects. We did not show increased prevalence of the Z or S alleles in patients identified with liver disease (7 out of 76) compared to wild type patients (45 out of 477; p=0.58). Specifically, only 3 of the 16 patients heterozygous for the Z allele were identified with liver disease.

Table 3.3 shows the anthropometric, clinical and lung function data as well as stable and acute phase increases in α_1-AT levels in our sub-study group investigating α_1-AT levels. As expected, there were significant differences between groups in measures of disease severity (i.e., %predFVC, S-K and Brasfield scores). The stable status levels of α_1-AT were within

the normal range for the mild/moderate group and outside the upper limit of the normal range in our laboratory for the severe group (Figure 3.3). The peak values for α_1-AT during pulmonary exacerbation were significantly elevated above normal in both groups, and declined at a similar rate throughout the 14-day intervention. Although the levels in patients with severe pulmonary disease were lower at most time points neither group returned to normal in the 14-day period (Figure 3.3). The percent change in α_1-AT levels from peak to stable was greatest in patients with mild/moderate pulmonary disease, whose mean value was within the upper limit of the normal range (0.95-1.77 g/L) at the time of stable status, whereas that of patients with severe pulmonary disease was not.

There were no significant associations between spirometric and clinical measures and stable or percent changes in α_1-AT levels. The percent change in α_1-AT levels was positively skewed and was log base 10 transformed prior to regression analysis. Stepwise regression identified BMI as a predictor of percent change in α_1-AT (%change in α_1-AT =-50.5+3.5*BMI; adjusted R^2=0.08, p=0.05). Significant correlation coefficients were shown for BMI and stable α_1-AT (r=-0.34, p=0.04) and percent change in α_1-AT (r=0.32, p=0.05). When we ran the above analysis selecting out the five S/Z subjects who were all in the severe group, there was no change in the model.

3.3 DISCUSSION

The results of this study indicate that the α_1-AT gene is not a modifier gene in CF. In the large study group Z and S polymorphisms were not associated with increased pulmonary disease severity as defined by pulmonary function, frequency and duration of pulmonary exacerbations, death or lung transplantation. Similarly, the A allele of the 3 prime mutation in the α_1-AT gene was not protective. We showed a blunted acute phase increase in α_1-AT to pulmonary infection in those CF patients in a malnourished state. Since we recruited consecutive patients for the α_1-AT levels study as they were admitted for a pulmonary exacerbation it is conceivable that our cohort is biased towards more severe CF patients who require more frequent IV antibiotic therapy. However patients with mild pulmonary disease infrequently require admissions for pulmonary exacerbations and therefore such a bias is difficult to avoid.

The α_1-AT MZ genotype has been shown to be a risk factor for COPD (24). In CF cohorts, the association of α_1-AT genotype and pulmonary disease severity is unclear. Doring and colleagues (13) found no association between α_1-AT S and Z alleles and pulmonary disease severity but they did show an earlier age of onset of *P. aeruginosa* infection in individuals with these deficiency alleles (6 out of a total sample of 215). Mahadeva and associates questioned this association and in fact showed that patients who were heterozygous for the S and Z alleles (19 out of a total sample of 147) had higher levels of pulmonary function than the wild type individuals (15). In another study, the same authors showed that the α_1-AT Z and S deficiency alleles were not more prevalent in those CF patients with severe pulmonary disease (dead or lung transplanted CF patients)(14). We also showed no difference in outcome (death or transplant) for our cohort, but our cohort was followed prospectively for a relatively short time interval (5.5 years).

Kalsheker and associates (22) showed that the 3' mutation was associated with COPD (22). Morgan and colleagues provided *in vitro* evidence that the association with COPD may be due to deficiency in the α_1-AT acute phase response (12). However, Sandford and associates (23) did not show that the 3' mutation attenuated the acute phase rise in α_1-AT in patients undergoing open-heart surgery. Similarly, Madadeva et al. showed that the 3'

mutation had no effect on α_1-AT levels in CF patients (15). In a recent study by Henry and associates (16) they showed less severe pulmonary disease and fewer infective pulmonary exacerbations over 2 years in CF patients who were heterozygous for the A allele. These data suggest that heterozygotes may have a slower disease progression. The results of our study (which also included 7 homozygous individuals for the A allele) do not support the findings of Henry and associates (16). None of our measures of disease severity showed the A allele to be associated with less severe pulmonary disease.

In genetic association studies such as this, population stratification based on ethnicity can be a confounding factor. However, as expected the vast majority of our study sample was Caucasian, and therefore it is unlikely that the lack of association in this study represents a type 2 error due to stratification. In a multicenter study such as this there may have been differences in ethnic diversity between centers. To address this, and other potential confounders, we created a categorical variable called CENTER. However, center was not a significant predictor of lung function and center to center variability was small (less than 1% of the total variability, which was less than the variability within centers). This result suggests that there were no large ethnic (or other) differences between centers that could have affected the association of genotype with measures of lung function.

Other possible confounders in our study include social class and environment (i.e., smoking / passive smoking, increased exposure to air pollutants and infectious agents) and differences in center care. Center care is not likely a confounder in Canada as the care received by CF patients across Canadian CF clinics is standardized and therefore it comes as no surprise that the variable Center was not a significant predictor of lung function.

Interestingly, infusions of α_1-AT have been shown to reduce NE to undetectable levels in CF BALF (25). We showed CF patients with severe pulmonary disease (i.e., %predFEV_1<50%) had a blunted acute phase rise in α_1-AT (Figure 3.3). Also noted is a much lower percent change in α_1-AT levels from peak to stable in CF patients with severe disease. Possible reasons for this difference are the higher baseline values in the severe group and also the possibility that milder exacerbations in the severe group led to hospitalization or initiation of home IV therapy compared to the mild/moderate group. CF

patients with severe disease also had poor nutritional status, which has been shown in non-CF cohorts to affect α_1-AT levels.

The mean 20% increase in α_1-AT during an acute infection reflects a relatively small acute phase response. Voulgari and associates reported a 78% increase in α_1-AT levels in response to bacterial infection in a non-CF population (26). Kueppers measured α_1-AT levels in response to a typhoid vaccine injection in otherwise healthy males who were homozygous wild type and heterozygous for the α_1-AT deficiency alleles Z and S (27). While there was a lower baseline α_1-AT levels in heterozygotes, a similar percent rise in levels was seen across the groups. In the study by Kueppers (27), α_1-AT levels were monitored over 15 days; both the homozygous wild type and heterozygotes for the deficiency alleles showed a gradual return to normal values unlike the pattern seen in our group who had severe pulmonary disease. It may be that the high baseline level and blunted acute phase increase in α_1-AT levels in our study is related to the chronic pulmonary bacterial infection in CF. Alternatively, poor dietary intake and malabsorption leading to malnutrition and cachexia could play a role. Morlese and associates measured the acute α_1-AT phase response to infection in nine 10-year old children who were also diagnosed with severe malnutrition (28). These children showed a blunted acute phase increase in α_1-AT to bacterial infection. CF patients who showed heterozygosity for the S and Z polymorphisms also showed BMI values indicative of malnutrition as well as a significantly lower percent increase in α_1-AT levels during pulmonary exacerbation.

In our analyses BMI was an important predictor of α_1-AT levels even when the CF patients who were heterozygous for the S and Z polymorphisms were selected out, suggesting that BMI irrespective of α_1-AT genotype is a predictor of percent change in α_1-AT levels during pulmonary exacerbation. While we did not show significant differences for BMI when our cohort was grouped by pulmonary disease severity, the mean BMI for the severe group was below 20 kg/m^2 (Table 3.3) indicative of poor nutritional status and malnutrition. All the Canadian clinics promote patients to maintain normal body weight (i.e. adult BMI values between 20-25 kg/m^2 and preferably between 22-25 kg/m^2) in order to allow for weight loss that usually occurs during periods of pulmonary exacerbation or other CF-related illnesses. It is likely that a combination of poor nutritional status and chronic pulmonary bacterial

infection, which is common in CF patients with severe pulmonary disease, contributes to the pattern shown in stable and acute rise in α_1-AT levels observed. An exogenous source of α_1-AT may be of potential benefit in malnourished CF patients with more severe pulmonary disease during pulmonary exacerbation, while measures are taken to normalize their body weight.

In this study we used the common lung function parameter %predFEV_1 as a measure of disease severity. This parameter is universally used as a measure of pulmonary disease severity and also correlates with measures of nutritional status (i.e., BMI in adults and percent of ideal weight in children) and pancreatic sufficiency status. Clinical scoring using Schwachman-Kulczycki (S-K) and Brasfield chest radiographic scores are also used, but have not been commonly used in large-scale studies. A benefit of using S-K scores is that this score takes into account pulmonary, nutritional, chest radiographic status and a measure of activity or mobility. Exercise capacity (29) (30), quality of life (31-33) and sputum volume (33), are alternative measures of pulmonary disease but these measures are not commonly utilized in CF patients in clinical assessment. As our study showed that non-pulmonary measures (i.e., BMI, PSS) affect pulmonary disease severity, this suggests that a measure for multi-system CF disease severity such as BMI and PSS is essential.

3.4 CONCLUSIONS

The results of our study show that α_1-AT genotype is not a major contributor to the variability of pulmonary disease severity in CF. Specifically α_1-AT genotype did not correlate with %predFEV_1, pulmonary infections and death or lung transplantation. Our study shows, however, that the levels of α_1-AT during pulmonary infections may be affected by poor nutritional status independent of α_1-AT genotype.

Table 3.1. Clinical characteristics of study subjects stratified by α_1-AT S and Z genotypes. Values are shown as mean (SEM)*.

	Total	MM	MS, SS or MZ	p-value
Sex (M/F)	378/338	317/299	61/39	0.08
Age (years)	22.8(0.4)	23.0(0.5)	21.8(1.1)	0.31
CFTR genotype (ΔF508/ΔF508; ΔF508/other; other/other)	391/295/72	342/254/59	49/41/13	0.44
Age of CF diagnosis (years)	4.7(0.3)	4.6(0.5)	5.5(0.9)	0.32
%predFEV$_1$	64.8(0.9)	64.6(1.0)	65.7(2.7)	0.68
BMI (kg/m^2)	20.3(0.2)	20.4(0.2)	19.7(0.3)	0.22
Pancreatic sufficiency status (sufficient/insufficient)	119/648	101/561	18/87	0.35
# of P. aeruginosa positive/not colonized	478/85	421/67	57/18	0.04
Age of 1st infection with P. aeruginosa	11.1(0.5)	11.4(0.5)	9.1(1.1)	0.10
Age of chronic P. aeruginosa infection	14.4(0.8)	14.9(0.9)	11.3(1.5)	0.16
Frequency of IV Treatment/year	0.9(0.1)	0.9(0.1)	0.7(0.2)	0.22
Days of IV Treatment/year	13.7(1.2)	14.1(1.3)	11.5(3.1)	0.41
Dead or lung transplanted / alive	63/644	54/555	9/89	0.85

*The study sample varied between 713-716 for univariate analyses. For analysis of age of first infection with P. aeruginosa and P. aeruginosa status the study sample size was 461 and 555, respectively. For analysis of age of chronic infection with P. aeruginosa the study sample size was 159.

Table 3.2. Clinical characteristics of study subjects stratified by α1-AT 3' G1237→A genotype. Values are shown as mean (SEM)*.

	Total	G_{1237}→G	G^{1237}→A or A_{1237}→A	p-value
Sex (M/F)	379/337	331/283	48/54	0.20
Age (years)	22.8(0.4)	22.7(0.4)	23.5(1.2)	0.49
CFTR genotype (ΔF508/ΔF508; ΔF508/other; other/other)	391/296/72	343/244/60	48/52/12	0.14
Age of CF diagnosis (years)	4.7(0.3)	4.6(0.3)	5.4(0.9)	0.38
%predFEV_1	64.8(0.9)	64.9(1.0)	63.7(2.3)	0.64
BMI (kg/m^2)	20.3(0.2)	20.4(0.2)	20.0(0.4)	0.49
Pancreatic sufficiency status (sufficient/insufficient)	119/649	95/561	24/88	0.05
# of *P. aeruginosa* positive / not colonized	478/85	416/70	62/15	0.23
Age of 1st infection with *P. aeruginosa*	11.1(0.5)	11.3(0.5)	10.0(1.1)	0.36
Age of chronic P. Aeruginosa infection	14.4(0.8)	14.5(0.9)	13.6(1.6)	0.71
Frequency of IV Treatment/year	0.9(0.1)	0.9(0.1)	0.9(0.2)	0.86
Days of IV Treatment/year	13.7(1.2)	13.5(1.6)	15.4(4.5)	0.60
Dead or lung transplanted / alive	63/644	53/551	10/92	0.71

*The study sample varied between 713-716 for univariate analyses. For analysis of age of first infection with *P. aeruginosa* and *P. aeruginosa* status the study sample size was 461 and 555, respectively. For analysis of age of chronic infection with *P. aeruginosa* the study sample was 159.

Table 3.3. Characteristics of the sub-group of patients in the acute phase α_1-AT level study.

	Total	Mild/Moderate	Severe	P-value
Gender (male/female)	24/15	7/7	17/8	0.22*
Age (yrs)	27.5 (1.1)**	28.3 (2.2)	27.1 (1.3)	0.60
CF diagnosis age (yrs)	3.1 (1.0)	2.6 (1.4)	3.4 (1.4)	0.73
Infection with *P. aeruginosa****	2/34/3	0/12/2	2/22/1	0.43*
Age of chronic infection with *P. aeruginosa*	14.6 (1.4)	17.8 (2.8)	12.8 (1.5)	0.09
BMI	20.5 (0.5)	21.5 (0.7)	19.9 (0.7)	0.11
%predFEV_1 (%)	43.7 (3.5)	67.8 (8.0)	30.2 (2.1)	0.0001
%predFVC (%)	64.0 (3.8)	88.3 (4.2)	50.5 (2.9)	0.0001
S-K score	59.5 (3.0)	75.0 (2.7)	50.8 (3.3)	0.0001
Brasfield score	13.7 (0.7)	16.9 (0.6)	11.9 (0.8)	0.0001
# of hospitalization days	38.7 (5.1)	26.8 (3.5)	45.4 (7.4)	0.08
Stable α_1AT (g/L)	1.84 (0.06)	1.73 (0.07)	1.90 (0.09)	0.19
Acute α_1AT levels (g/L)	2.18 (0.10)	2.38 (0.23)	2.07 (0.09)	0.13
% change in α_1AT	20.5 (5.4)	36.7 (10.4)	11.4 (5.4)	0.02

*- Chi-square analysis.

**- Values are mean (SEM).

***- Values are number (#) of subjects not infected with *P. aeruginosa* / # of subjects chronically infected with *P. aeruginosa* / # of subjects infected with *P. aeruginosa* but not chronically. One of the patients in the severe group not infected with *P. aeruginosa* was chronically infected with *Burkholderia cepacia* complex.

Figure 3.1. Comparison of pulmonary disease severity and Z and S alleles of the α_1-AT gene. Values are presented as mean+SEM.

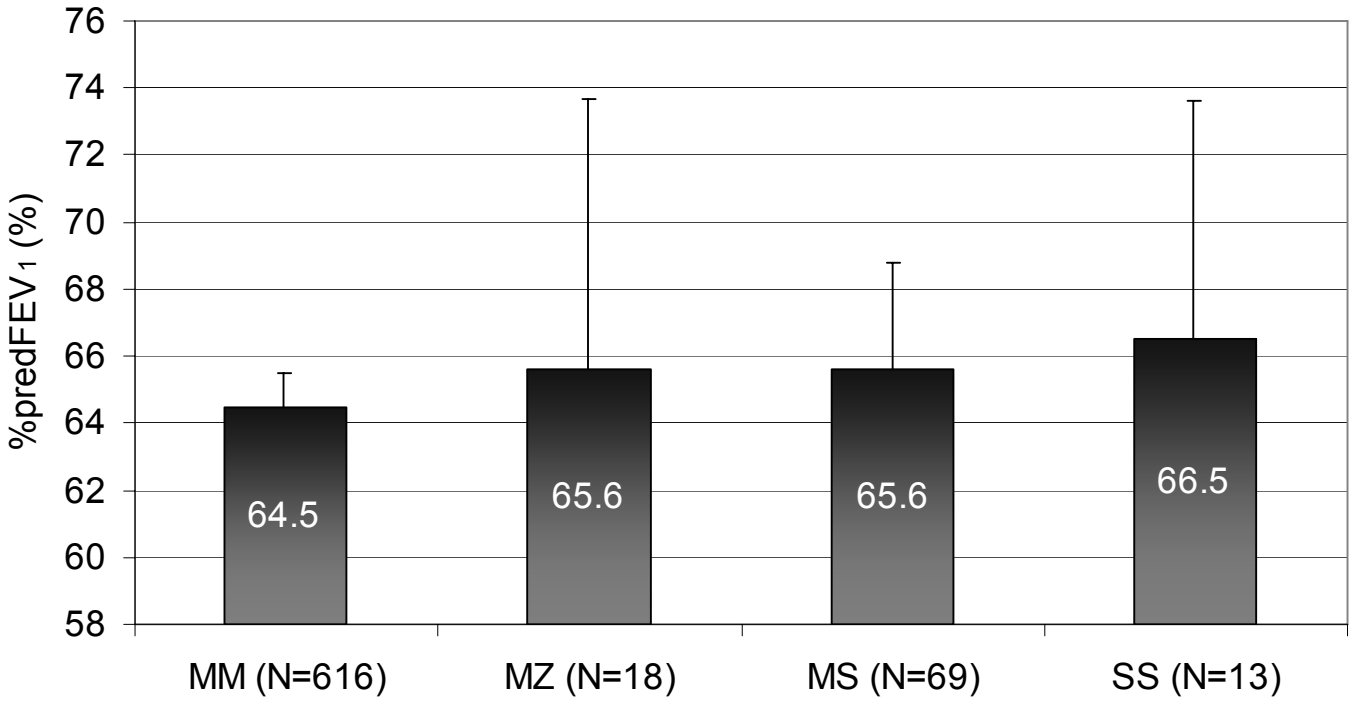


Figure 3.2.Comparison of pulmonary disease severity and the 3' $G_{1237} \rightarrow A$ mutation of the α_1-AT gene. Values are presented as mean+SEM.

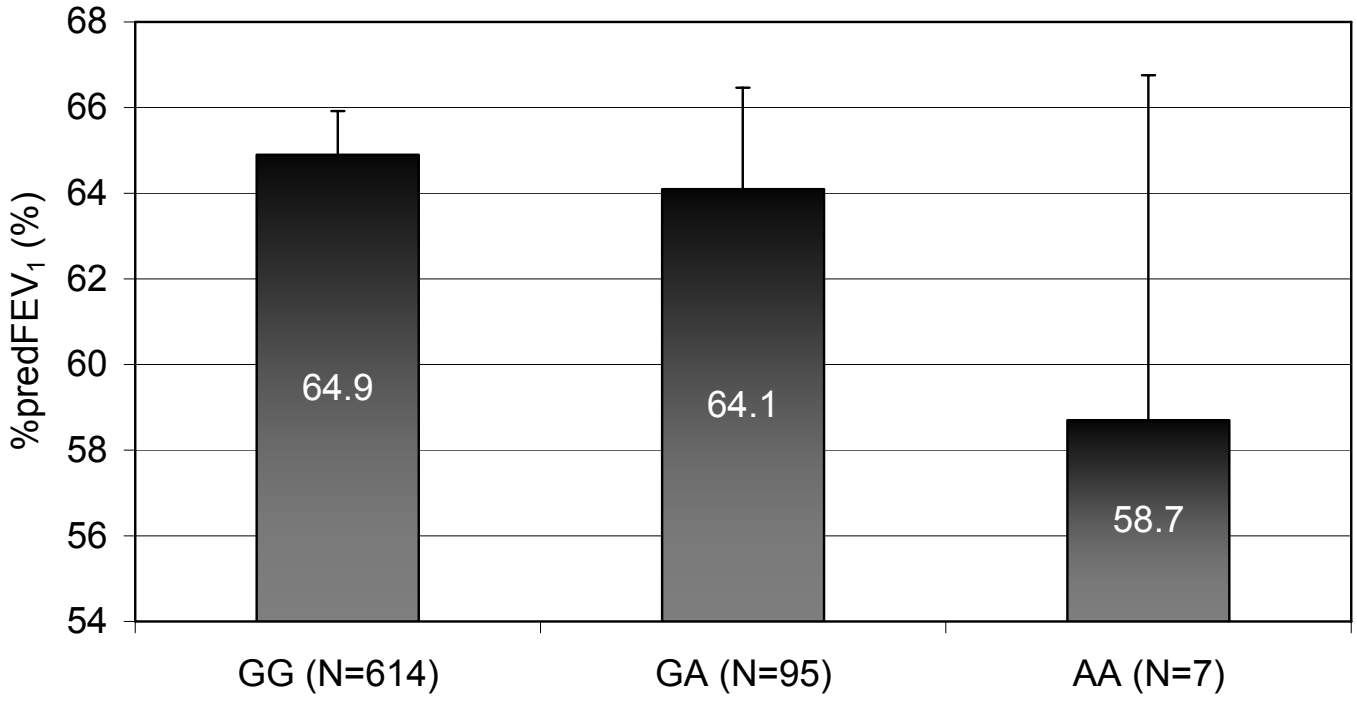
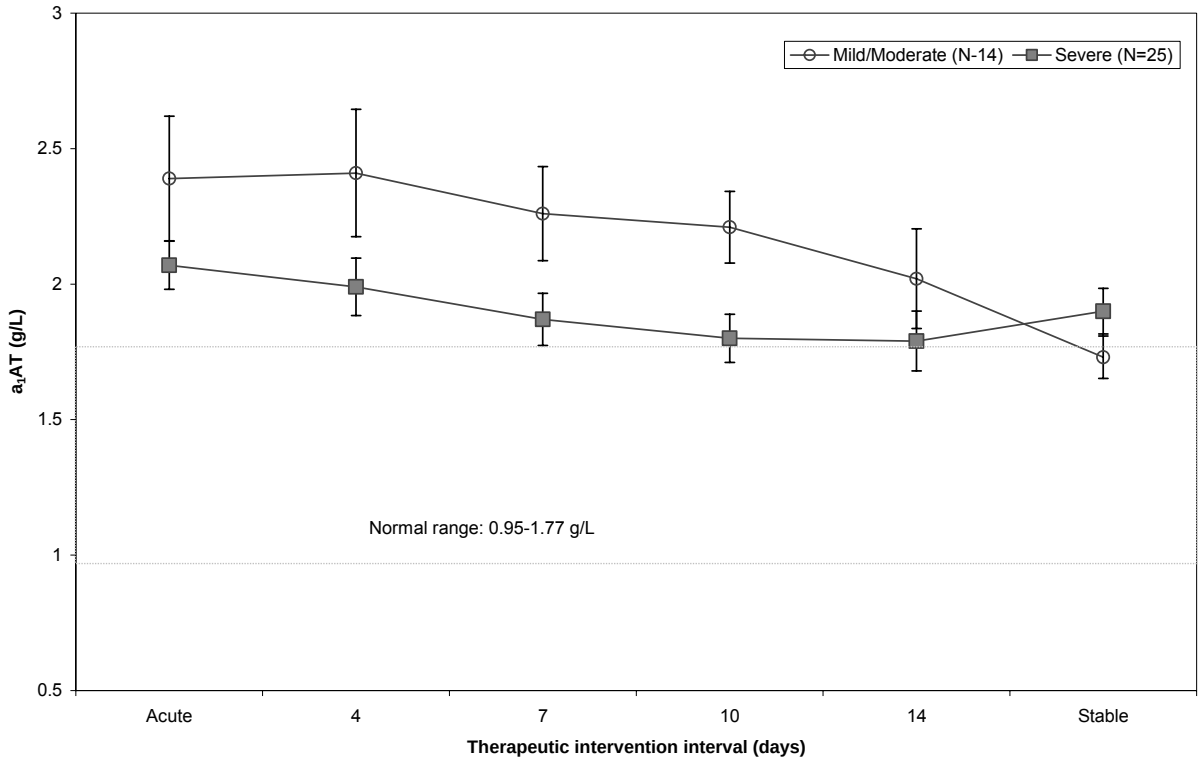


Figure 3.3. Alpha-1-antitrypsin levels during a pulmonary exacerbation and post-exacerbation levels during stable clinical status. Levels are shown during the intravenous antibiotic intervention period (14 day) and at a post-exacerbation (stable) time point by pulmonary disease severity. Dotted lines show the range of normal levels for α_1-AT.



3.5 BIBLIOGRAPHY

1. Cantin, A. M., S. Lafrenaye, and R. O. Begin. 1991. Antineutrophil elastase activity in cystic fibrosis serum. *Pediatric Pulmonology* 11(3):249-53.
2. Birrer, P., N. G. McElvaney, A. Rudeberg, C. W. Sommer, S. Liechti-Gallati, R. Kraemer, R. Hubbard, and R. G. Crystal. 1994. Protease-antiprotease imbalance in the lungs of children with cystic fibrosis. *American Journal of Respiratory & Critical Care Medicine* 150(1):207-13.
3. O'Connor, C. M., K. Gaffney, J. Keane, A. Southey, N. Byrne, S. O'Mahoney, and M. X. Fitzgerald. 1993. alpha 1-Proteinase inhibitor, elastase activity, and lung disease severity in cystic fibrosis. *Am Rev Respir Dis* 148(6 Pt 1):1665-70.
4. Goldstein, W., and G. Doring. 1986. Lysosomal enzymes from polymorphonuclear leukocytes and proteinase inhibitors in patients with cystic fibrosis. *American Review of Respiratory Disease* 134(1):49-56.
5. Plotkowski, M. C., G. Beck, J. M. Tournier, M. Bernardo-Filho, E. A. Marques, and E. Puchelle. 1989. Adherence of Pseudomonas aeruginosa to respiratory epithelium and the effect of leucocyte elastase. *J Med Microbiol* 30(4):285-93.
6. Plotkowski, M. C., A. M. Saliba, S. H. Pereira, M. P. Cervante, and O. Bajolet-Laudinat. 1994. Pseudomonas aeruginosa selective adherence to and entry into human endothelial cells. *Infect Immun* 62(12):5456-63.
7. Hilman, B. C. 1997. Genetic and immunologic aspects of cystic fibrosis. *Ann Allergy Asthma Immunol* 79(5):379-90; quiz 391-4.
8. Brown, M. A., W. J. Morgan, P. R. Finley, and P. Scuderi. 1991. Circulating levels of tumor necrosis factor and interleukin-1 in cystic fibrosis. *Pediatric Pulmonology* 10(2):86-91.
9. Suter, S. 1986. Imbalance between polymorphonuclear leukocyte proteases and antiproteases in chronic pyogenic infections and its relation to the proteolysis of complement component C3. *Complement* 3(1):1-24.
10. Santis, G., L. Osborne, R. Knight, M. E. Hodson, and M. Ramsay. 1990. Genetic influences on pulmonary severity in cystic fibrosis. *Lancet* 335(8684):294.
11. Santis, G., L. Osborne, R. Knight, M. Ramsay, R. Williamson, and M. Hodson. 1990. Cystic fibrosis haplotype association and the delta F508 mutation in adult British CF patients. *Hum Genet* 85(4):424-5.

12. Morgan, K., G. Scobie, P. Marsters, and N. A. Kalsheker. 1997. Mutation in an alpha1-antitrypsin enhancer results in an interleukin-6 deficient acute-phase response due to loss of cooperativity between transcription factors. *Biochim Biophys Acta* 1362(1):67-76.

13. Doring, G., H. Krogh-Johansen, S. Weidinger, and N. Hoiby. 1994. Allotypes of alpha 1-antitrypsin in patients with cystic fibrosis, homozygous and heterozygous for deltaF508. *Pediatric Pulmonology* 18(1):3-7.

14. Mahadeva, R., S. Stewart, D. Bilton, and D. A. Lomas. 1998. Alpha-1 antitrypsin deficiency alleles and severe cystic fibrosis lung disease. *Thorax* 53(12):1022-4.

15. Mahadeva, R., R. C. Westerbeek, D. J. Perry, J. U. Lovegrove, D. B. Whitehouse, N. R. Carroll, R. I. Ross-Russell, A. K. Webb, D. Bilton, and D. A. Lomas. 1998. Alpha1-antitrypsin deficiency alleles and the Taq-I G-->A allele in cystic fibrosis lung disease. *Eur Respir J* 11(4):873-9.

16. Henry, M. T., S. Cave, J. Rendall, C. M. O'Connor, K. Morgan, M. X. FitzGerald, and N. Kalsheker. 2001. An alpha(1)-antitrypsin enhancer polymorphism is a genetic modifier of pulmonary outcome in cystic fibrosis. *Eur J Hum Genet* 9(4):273-8.

17. Crystal, R. G. 1989. The alpha 1-antitrypsin gene and its deficiency states. *Trends Genet* 5(12):411-7.

18. Cox, D. W., S. L. Woo, and T. Mansfield. 1985. DNA restriction fragments associated with alpha 1-antitrypsin indicate a single origin for deficiency allele PI Z. *Nature* 316(6023):79-81.

19. Curiel, D. T., A. Chytil, M. Courtney, and R. G. Crystal. 1989. Serum alpha 1-antitrypsin deficiency associated with the common S-type (Glu264----Val) mutation results from intracellular degradation of alpha 1-antitrypsin prior to secretion. *J Biol Chem* 264(18):10477-86.

20. Carrell, R. W. 1986. alpha 1-Antitrypsin: molecular pathology, leukocytes, and tissue damage. *J Clin Invest* 78(6):1427-31.

21. Carrell, R. W., and D. A. Lomas. 2002. Alpha1-antitrypsin deficiency--a model for conformational diseases. *N Engl J Med* 346(1):45-53.

22. Kalsheker, N. A., G. L. Watkins, S. Hill, K. Morgan, R. A. Stockley, and R. B. Fick. 1990. Independent mutations in the flanking sequence of the alpha-1- antitrypsin gene are associated with chronic obstructive airways disease. *Dis Markers* 8(3):151-7.

23. Sandford, A. J., T. Chagani, J. J. Spinelli, and P. D. Pare. 1999. alpha1-antitrypsin genotypes and the acute-phase response to open heart surgery. *Am J Respir Crit Care Med* 159(5 Pt 1):1624-8.

24. Seersholm, N., J. T. Wilcke, A. Kok-Jensen, and A. Dirksen. 2000. Risk of hospital admission for obstructive pulmonary disease in alpha(1)-antitrypsin heterozygotes of phenotype PiMZ. *Am J Respir Crit Care Med* 161(1):81-4.

25. McElvaney, N. G., R. C. Hubbard, P. Birrer, M. S. Chernick, D. B. Caplan, M. M. Frank, and R. G. Crystal. 1991. Aerosol alpha 1-antitrypsin treatment for cystic fibrosis. *Lancet* 337(8738):392-4.

26. Voulgari, F., P. Cummins, T. I. Gardecki, N. J. Beeching, P. C. Stone, and J. Stuart. 1982. Serum levels of acute phase and cardiac proteins after myocardial infarction, surgery, and infection. *Br Heart J* 48(4):352-6.

27. Kueppers, F. 1968. Genetically determined differences in the response of alpha-antitrypsin levels in human serum to typhoid vaccine. *Humangenetik* 6(3):207-14.

28. Morlese, J. F., T. Forrester, and F. Jahoor. 1998. Acute-phase protein response to infection in severe malnutrition. *Am J Physiol* 275(1 Pt 1):E112-7.

29. Frangolias, D. D., C. L. Holloway, S. Vedal, and P. G. Wilcox. 2003. Role of exercise and lung function in predicting work status in cystic fibrosis. *Am J Respir Crit Care Med* 167(2):150-7.

30. Frangolias, D. D., and P. G. Wilcox. 2001. Predictability of oxygen desaturation during sleep in patients with cystic fibrosis : clinical, spirometric, and exercise parameters. *Chest* 119(2):434-41.

31. Quittner, A. L., and A. Buu. 2002. Effects of tobramycin solution for inhalation on global ratings of quality of life in patients with cystic fibrosis and Pseudomonas aeruginosa infection. *Pediatr Pulmonol* 33(4):269-76.

32. Goldbeck, L., and T. G. Schmitz. 2001. Comparison of three generic questionnaires measuring quality of life in adolescents and adults with cystic fibrosis: the 36-item short form health survey, the quality of life profile for chronic diseases, and the questions on life satisfaction. *Qual Life Res* 10(1):23-36.

33. Bradley, J., O. McAlister, and S. Elborn. 2001. Pulmonary function, inflammation, exercise capacity and quality of life in cystic fibrosis. *Eur Respir J* 17(4):712-5.

CHAPTER 4: INNATE IMMUNITY GENES AS POTENTIAL MODIFIER LOCI IN CYSTIC FIBROSIS

A version of this chapter may be published in the future.

4.0 INTRODUCTION

Polymorphisms in genes involved in innate immunity were investigated. Polymorphisms in the coding region of the mannose binding lectin (MBL2) gene have been shown to affect the levels of the protein produced and polymorphisms in the promoter region have been shown to affect regulation. Numerous polymorphisms have been identified in the pulmonary surfactant genes A1, A2 and D. It is not known at this time whether the polymorphisms identified affect protein function and levels. In this study we investigated the association of amino acid changing polymorphisms in the pulmonary surfactant genes and MBL2 gene polymorphisms on pulmonary disease severity and progression, survival and succeptibility to infection with common CF pathogens.

4.1 RATIONALE FOR THE INVESTIGATION OF INNATE IMMUNITY GENES AS POTENTIAL MODIFIERS IN CF

Chronic colonization with *Pseudomonas aeruginosa* is a common phenotypic characteristic in the CF population and an earlier age of chronic infection is associated with more rapid pulmonary disease progression (1-4). Commonly, repeated infections with *Staphylococcus aureus* and *Haemophilus influenza* describe the clinical course prior to colonization with *P. aeruginosa*. A less common pathogen in the CF population is *Burkholderia cepacia* complex (BCC), and colonization with this pathogen is associated with a more severe clinical outcome. First line innate defenses against pathogens invading the airways are therefore important in preventing colonization and possibly modulating the inflammatory response.

Mannose binding lectin (MBL2) and pulmonary surfactant A (SPA) proteins have been shown to interact with *Staphylococcus aureus*, a pathogen known to cause recurrent early infections in the CF population and also recurrent infections and chronic colonization following chronic colonization with the pathogen *P. aeruginosa* in this population. In addition to binding to *Staphylococcus aureus* (5), MBL2 has been shown to bind to *P. aeruginosa* and BCC (6). Davies and associates, using BCC isolates obtained from chronically infected CF patients, showed that MBL2 binds to BCC and activates complement, as measured by C4 deposition (6). SPA and SPD bind to pathogens, which are known to be responsible for respiratory infections and include common CF respiratory pathogens *Staphylococcus aureus*, *P. aeruginosa* and *Haemophilus influenza* (7). Although

it has been suggested that direct binding of SPA and SPD to the pathogen is not always necessary to trigger microbicidal activity, it is believed that the presence of these proteins may somehow mediate host defense (8-11). Studies have shown that SPA (12) and SPD (13) proteins enhance the alveolar macrophage phagocytosis of *P. aeruginosa*.

In this study we limited our investigation of innate immunity genes to polymorphisms in 4 genes: MBL2, SPA-1, SPA-2 and SPD. Other potential modifiers in the innate immunity pathway which could potentially affect pulmonary disease severity, progression and susceptibility to infection with CF pathogens are numerous. Genetic variations that disrupt innate immune sensing of infectious organisms can affect the ability of the host to respond to infection. Therefore variants in proteins which bind to pathogen, receptors on phagocytic cells, proteins in the complement cascade, signaling molecules are all potential candidate modifier genes for CF. Some potential candidates where variants have been identified include: inflammatory mediators IL-1beta, IL-1Ra, IL-8, IL-10, plasma complement proteins of the classical, lectin and alternate pathway of complement activation, constituents of host cells (alpha defensins and acyoxyacyl hydrolase of neutrophils, beta defensins of epithelial cells), innate immune receptors (CD14, toll-like receptors 2, 4 and 6), serine proteases (Mannose binding lectin associated serine protease, C1q, factor P).

4.1.1 Tissue distribution of innate immunity proteins

MBL2, SPA and SPD proteins are regulated hormonally and developmentally and are influenced by inflammation. MBL2 is characterized as an acute phase protein and levels of the protein are two thirds of adult levels at birth and increase to adult values within a few weeks following birth (14). SPA and SPD are found in amniotic fluid as early as 16-20 weeks of gestation and increase during pregnancy. Levels of SPA and SPD can be increased prematurely artificially with glucocorticoid treatment (15).

MBL2 and its associated serine proteases (MASP-1, 2 and 3 and MAp19) are synthesized by hepatocytes and secreted into the plasma. Alveolar type II cells and unciliated bronchial epithelial cells synthesize SPA and SPD proteins and they are secreted onto the airway surface lining fluid (ASLF) (16). Lin and associates have shown that the SPA genes are also expressed in the small and large intestine (17). SPD is also synthesized and secreted by epithelial cells of the gastrointestinal mucosa and exocrine glands (18-20).

4.1.2 Characteristics of innate immunity proteins

The human MBL2 and pulmonary surfactant A and D genes are located on the long arm of chromosome 10 with MBL2 mapped to 10q11.2-q21 and SPA and SPD mapped to region 10q22-q23. The order of the genes is (21) (Table 4.1):

Centromere _ MBL2 _ SPA-2 _ SPA pseudogene _ SPA-1 _ SPD _ Telomere

There are two functional SPA genes and a SPA pseudogene (22). SPA-1 and SPA-2 genes have similar DNA sequences and the organization of each SPA gene consists of four coding exons which span a region less than 5 kb (23, 24). Sequence analysis of the three SPA genes genomic sequences by Hoover and Floros suggests that an ancestral SPA sequence was duplicated giving rise to SPA-1 and SPA-2. A subsequent duplication of SPA-2 produced the SPA pseudogene (21). The orientations of the SPA-1 and SPA-2 genes are 'head to head' in opposite transcriptional orientations and may share *cis*-acting regulatory elements (21). Of the genes listed above, the MBL2 locus is physically the farthest away from SPA and SPD and is considered more distant in evolutionary terms (25). MBL2 is not in linkage disequilibrium (LD) with any of the surfactant genes (21). SPD is closer from an evolutionary perspective and in terms of physical distance to the SPA locus than MBL2 (21, 25). SPA-1 and SPA-2 are in strong LD (21).

4.1.3 Structure and function of innate immunity proteins

MBL2, SPA and the SPD proteins bind to carbohydrates commonly found on bacterial and viral surfaces but not found in the host. MBL2, SPA-1, SPA2 and SPD proteins share a similar primary structure. They are composed of four structural domains; polypeptide subunits with N terminal segments which interconnect and form the center of the molecule, a collagen-like domain, an acidic hydrophobic neck domain, and a C-type lectin carbohydrate recognition domain that binds carbohydrates in a calcium dependent manner (26, 27). MBL2, SPA-1 and SPA-2 proteins form a three dimensional structure that resembles a 'flower or tulip bouquet', which is similar to the structure of the first component of the classical complement pathway molecule C1q (28, 29). SPD is similar in structure and amino acid composition to MBL2, SPA-1 and SPA-2 proteins, but assembles into an X-like formation.

The MBL2 protein is made up of six 96 kDa subunits and each subunit consists of three identical 32 kDa polypeptide chains. MBL2 is shown to opsonize numerous pathogens, which are summarized in Table 4.2. A comprehensive study by Neth and associates investigated MBL2 binding and complement activation of multiple isolates of clinically obtained pathogens (30). They showed that MBL2 binds to specific pathogens and promotes C4 deposition (30). Davis and associates investigated two pathogens commonly infecting CF patients (mucoid and non-mucoid isolates of *P. aeruginosa* and of BCC obtained from CF patients) and showed that significant amount of MBL2 binds to BCC and that lesser amounts bind to mucoid strains of *P. aeruginosa*. MBL2 binding to BCC specimens resulted in complement activation as measured by C4 deposition (6). Younger and co-workers also showed intermediate binding of MBL2 to *P. aeruginosa*, and they found that the alternative pathway accounted for the majority of opsonized *P. aeruginosa* (31). Younger and co-workers showed that opsonization of *P. aeruginosa* (strain UI-18) in a murine model was preserved, even when the function of C1q and MBL2 were inhibited, suggesting that the alternative pathway was important in initiating opsonization of this pathogen (31).

The MBL2 gene contains four exons and each encodes one of the four regions of the protein. Specifically, exon 1 encodes the N-terminal segment, exon 2 encodes the collagen-like domain, exon 3 encodes the neck region, and exon 4 encodes the carbohydrate recognition domain. Three polymorphisms have been identified in exon 1 that encode a region of the protein that interacts with collectin receptors and the MBL2-associated serine proteases (MASP). These three polymorphisms encode different structural variants of the MBL2 protein (32-35). The polymorphisms result in three different amino acid substitutions (codon 52 (Arg→Cys, D allele), codon 54 (Gly→Asp, B allele) and 57 (Gly→Glu, C allele)) that prevent the correct assembly of the MBL2 protein subunits into a trimeric structure making the molecule unstable and decreasing the binding capacity to ligands (36). Interaction between the variant MBL2 and the MASPs still occurs, however, there is lack of complement activation (as measured by C4 deposition) and it is believed that the reason for this is reduced capacity of the variant MBL2 to bind ligand (36). Additionally variant MBL2 is shown to be more susceptible to matrix metalloproteinase proteolysis (37). The wild type allele is denoted as A for all three polymorphisms. Two promoter polymorphisms have been identified at positions -550 (H/L genotype) and -221 (Y/X genotype) which both involve a

single nucleotide substitution of G to C (29). The promoter haplotypes HY, LY and LX (HX was not observed) lead to high, intermediate and low MBL2 levels, respectively (29, 38, 39). The LX promoter haplotype is only found with the wild-type coding allele for exon 1 (29). A single nucleotide substitution of G to C has also been documented at position +4 but this polymorphism has not been shown to have an affect on MBL2 levels (38). Haplotype reconstruction and LD analysis performed on a sample consisting of 69 Eskimos, 120 Caucasians and 61 native black Africans in the study by Madsen and associates (29) showed LD between promoter variants except in the black African cohort. They also showed strong LD between the promoter and structural polymorphisms. The X promoter allele was only found in haplotypes with the L promoter allele and only with the A allele for the three coding polymorphisms. The B, C and D coding alleles were only observed with the high expression promoter (i.e., HY) haplotype (29).

MBL2 levels increase up to threefold after stimulation. The B, C and D coding alleles are associated with low plasma MBL2 protein concentrations (40-42). Intermediate plasma MBL2 protein concentrations are expressed in those who are heterozygous wild-type. Garred and co-workers (43) measured plasma levels of MBL2 protein in the CF population and showed absent or very low MBL2 levels in CF patients who were homozygous or heterozygous for the coding polymorphisms regardless of promoter haplotype. Low, intermediate and high MBL2 plasma levels were associated with XX, XY and YY promoter genotypes, respectively (in individuals who were wild-type for the exon 1 polymorphisms) (43).

Unlike MBL2, SPA and SPD proteins do not associate with MASPs and do not activate complement. While SPD is believed to have strictly antimicrobial functions, SPA proteins are also required for the structure and / or stability of surfactant aggregates. SPA and SPD are synthesized and released following an inflammatory stimulus and their mode of action is:

- Direct binding of SPA and SPD protein via their carbohydrate recognition domains to mannose carbohydrate on the pathogen's surface (to aggregate and opsonize pathogens directly)(7, 8), or
- Modulation of macrophage function (potentially by increasing the activity of the macrophage mannose receptor and other receptors on immune cells)(7, 11).

It is believed that SPA and SPD proteins bind to a number of different immune cell receptors including the MBL2 macrophage mannose receptor (7). SPA exhibits a similar acute phase response as MBL2 protein after LPS aerosolization (44) and intratracheal instillation (45). SPA binds to the common CF pathogens, *Staphylococcus aureus* and *P. aeruginosa*. SPA functions as an opsonin and binds to *Staphylococcus aureus* and the C1q receptor of the phagocytes and stimulates phagocytosis by alveolar macrophages (8, 46, 47) Macrophage uptake of non-opsonized *Staphylococcus aureus*, as well as *P. aeruginosa,* have also been shown to occur in the presence of SPA (48). SPA protein binds to Gram negative bacteria by interacting with the pathogen's LPS, specifically SPA has been shown to bind to LPS via a lipid A domain (49, 50). It is believed that SPA binding to Gram negative bacteria is dependent on the LPS structure (7, 48). In the case of the pathogen *Mycobacterium tuberculosis*, SPA acts as a 'Trojan horse' for the pathogen. SPA binds to *Mycobacterium tuberculosis* via its carbohydrate recognition domains and then binds to the macrophage mannose receptor stimulating its uptake (51), and the pathogen can multiply within the macrophage.

In the European population, there are at least 153 and 110 polymorphisms in the SPA-1 and SPA-2 genes, respectively (Seattle SNPs, http://pga.gs.washington.edu/). Three (Ala19Val, Leu50Val, Arg219Trp) of the polymorphisms identified in SPA-1 and three (Asn9Thr, Ala91Pro, Gln223Lys) of the polymorphisms identified in SPA-2 are amino acid changing (52). Floros and associates investigated in a cohort of 239 the LD between SPA-1 and SPA-2 and found to be strong (22). Many of the SPA-1 and SPA-2 haplotypes previously reported by Flores and associates (22) are rare and only four SPA-1 and SPA-2 gene haplotypes (i.e., $6A^2$-$1A^0$, $6A^3$-$1A^1$, $6A^3$-$1A^0$ and $6A^3$-$1A^2$) are observed in Caucasians (53). Table 4.3 shows the common single nucleotide polymorphisms that make up the SPA-1 and SPA-2 haplotypes as reported by DiAngelo and associates (52).

SPA and SPD bind to a broad range of pathogens with some overlap (Table 4.2) but differ in their mode of interaction with the pathogen (7). For example in the case of Gram-negative bacteria, Van Iwaarden and associates (50) showed that SPA binds to the lipid moiety of LPS, whereas SPD interacts with the core oligosaccharides of the Gram-negative bacteria (7). One receptor that SPD has been proposed to bind to is GP-340 found on alveolar macrophages, although other receptors have been purported to exist, but have not been

identified at this time. It is also hypothesized that SPD binds to pathogen and somehow mediates modifications in the presentation of pathogen to host defense cells (54). SPD has been shown to bind to pathogens which infect CF patients: *P. aeruginosa* (48) and *Haemophilus influenza* (9, 55, 56). Three polymorphisms in the coding region of the SPD gene have been described; Thr11Met and Thr160Ala (52) and Thr270Ser (52, 57).

The clinical implications of SPA-1, SPA-2 and SPD polymorphisms are unclear and there is no consensus on the normal range for these three proteins, or correlations between concentrations of these proteins and genotypes (28). Lower levels of these proteins have been reported in serum and BALF in patients who have cystic fibrosis, adult respiratory distress syndrome, and chronic smokers and in patients characterized as having an increased risk of pneumonia (58, 59). Lower levels of these proteins are used as clinical indicators for lower respiratory tract infections, type II cell hyperplasia and acute lung injury in respiratory diseases (60-64). Mikerov and associates(65) have shown differences in the abilities of SPA-1 and SPA-2 proteins to increase phagocytosis of *P. aeruginosa* by alveolar macrophages and that SPA-2 variants tested (i.e., $1A^0$ and 1A) were more active than SPA-1 variants ($6A^4$ and $6A^2$). Increased susceptibility to infection and mortality have been shown in SPA and SPD knock-out mice (66-68) and associations shown between respiratory diseases and known polymorphisms (69-73).

4.1.4 Complement activation pathways

The complement system comprises a series of enzymatic and non-enzymatic proteins, which are required for the operation of the innate immune system. The complement pathway is activated in three ways with the latter two involved in first line defense against pathogen infection:

- The classical pathway
- The alternate pathway
- The lectin pathway

In the 1940s, researchers proposed that an antibody-independent pathway, the alternative pathway, existed and that complement could be activated by bacterial surfaces. It was later shown that there are pattern recognition receptors expressed by the host that allow the host to recognize pathogen-associated molecular patterns (PAMPs) and thus distinguish pathogens from self antigens (74, 75). The lectin pathway is a recent discovery dated to

1987 (76, 77) and a better understanding of this pathway came when the MASPs were identified (14, 78). The lectin and alternative pathways provide a non-adaptive first line of defense against microbial infection since complement activation does not require specific antigen to bind to pathogen surfaces as required in the classical pathway. In addition, the response to pathogen invasion is an immediate activation compared with the classical pathway where there is a five to seven day delay before the production of the required antibody. The three pathways merge at the activation step of C3. Figure 4.1 summarizes the alternate and lectin pathways of complement activation and compares them to the classical pathway of complement activation.

The classical pathway is triggered by antibody binding to antigen on the pathogen surface and is part of the adaptive humoral immune response. The binding of immunoglobulin M and G (IgM and IgG) antibodies to the pathogen surface and to C1q activates the complement cascade. The elimination of pathogens is achieved through a series of reactions, which involve plasma proteins that interact with the bound antibodies and are recognized as complement receptors by phagocytes that are stimulated to engulf the pathogen.

The alternative pathway is initiated by structures on the pathogen surfaces and in this pathway SPA and SPD proteins coat the surface of pathogens (opsonization) facilitating the uptake by macrophages and neutrophils. The alternative pathway of complement activation does not require the presence of specific antibody. C3 freely circulates in the plasma and can undergo spontaneous cleavage in plasma to generate fragments C3b and C3a. C3 contains an intramolecular thioester bond, which, once exposed to the molecular surface of invading pathogens, forms a covalent bond with it and facilitates the phagocytosis of pathogens through the C3 receptors on phagocytic cells.

C3b only binds to pathogen; if it does not bind to pathogen, it becomes inactivated by cell surface proteins on host cells. Pathogen cells lack these proteins and C3b covalently binds to the pathogen surface and also binds to additional proteins (i.e., factor B which is cleaved by factor D forming the C3b,Bb complex which is structurally and functionally homologous to C4b,2b of the classical pathway). C3b,Bb is the C3 convertase of the alternative pathway; the complex is further stabilized on the pathogen surface by binding to factor P. The stabilized C3b,Bb then acts in the same way as the C3 convertase of the classical pathway

converting circulating C3 molecules to C3b that coat the pathogen surface and C3a molecules which mediate local inflammation. Some of the C3b molecules bind to existing C3b,Bb complex and form C3b2,Bb. This is the alternative pathway of C5 convertase activation leading to the activation of the terminal complement components through binding and cleavage of C5 by C3b molecules. The C5b that is generated initiates the lytic pathway and C5a is a potent inflammatory mediator.

The C-lectin pathway is also initiated in an antibody-independent manner. MBL2 (and the ficolins) bind to carbohydrates (mannose and N-acetylglucosamine) on the surface of pathogens. MBL2 is structurally similar to complement C1 subcomponent, C1q, and the mode of activation for MBL2 is through associated serine proteases called MBL2-associated serine proteases (MASP). MBL2 associates through its collagen-like domain with serine proteases called MASP-1 and MASP-2 (41) and MASP-3 (79). Current research suggests that MASP-2 is the main enzyme responsible for activation of the lectin pathway (80). The MASP proteases are similar in structure, function and activation to C1r and C1s of the classical pathway. Specifically, binding of MBL2 to mannose containing proteins or carbohydrate groups on the surface of pathogens results in conformational changes to MASP-2; MASP-2 activates C4 and C2 and the C3 convertase C4b2a complexes are generated (81). MBL2 is also associated with a truncated form of MASP-2 called small MBL2-associated protein (sMAP). The MBL2-MASP-sMAP complex circulates in the blood and once MBL2 binds to a pathogen, the MASPs are activated acquiring proteolytic activities. Circulating in the complexed form MASPs are an inactive proenzyme (82, 83). The activated lectin pathway ultimately generates C3b and C3a. The killing of the pathogen via the MBL2 lectin pathway is mediated by the host-mediated terminal lytic complex and by clearance after promoting phagocytosis by macrophages.

4.1.5 Review of the literature for MBL2, SPA and SPD: Clinical correlates

The deficiency polymorphisms in the MBL2 gene have been investigated as potential risk factors for infections. The coding polymorphisms in exon 1 result in mutant MBL2 molecules, which do not participate in opsonization or complement activation (38). The B polymorphism is unable to bind MASP (82) and both the B and C polymorphisms are unable to activate complement (27, 84). MBL2 deficiency has been associated with increased risk of ear infections (85), respiratory infections (86-88) and other infections (87, 89) in non-CF

children . Kielgast and associates showed a potential association between viral infections and low umbilical cord serum MBL2 levels (90). Worse prognosis from viral and bacterial infections has been shown in immuno-compromised adults with low serum MBL2 levels (89, 91-93). Garred and co-workers showed an increased frequency of hospitalizations for infections in children with MBL2 deficient genotypes (91). The MBL2 D allele at codon 52 polymorphism is considered a potential contributor to chronic necrotizing pulmonary aspergillosis (94). Yang and co-workers have shown an association for increased hospitalizations for pulmonary infections in COPD patients with the MBL2 B allele at codon 54 (95). Neth and co-workers showed that children with febrile neutropenia who were undergoing chemotherapy for cancer had on average longer hospitalizations if they had an MBL2 deficient haplotype (5). Peterslund and associates showed a higher frequency of bacteremia and pneumonia in a group of patients undergoing chemotherapy for cancer, who also had low MBL2 levels (96).

MBL2 deficiency is common. It has been suggested that there is an evolutionary selection advantage to explain why these deficiency alleles are maintained in the population. In this respect, Garred and associates (97) have shown that MBL2 protein binds to *Mycobacterium leprae* and patients infected by *Mycobacterium leprae* are more likely to be MBL2 wild-type (based on MBL2 protein levels) compared with a control group of healthy blood donors residing in the same area. Similarly, an MBL2 haplotype associated with low MBL2 protein levels is protective against infections with *Plasmodium falciparum* (98), with *Cryptosporidium parvum* in AIDS patients (99), and *Leishmania chagasi* (100). Garred and colleagues concluded that during infection with intracellular pathogens the opsonic mechanisms of host defense are used by some pathogens to invade host cells (92, 97, 101).

Polymorphisms in the MBL2 gene resulting in reduced levels of this protein have been investigated in multiple CF cohorts to date. In three of these studies, MBL deficiency was associated with more severe pulmonary disease, earlier colonization with *Pseudomonas aeruginosa* (43, 102) and increased susceptibility to BCC colonization(43, 103). Choi and associates(104) only showed an association for the MBL2 structural deficiency alleles and pulmonary disease severity in CF patients who were homozygous for the delta F508 CFTR mutation. While these studies suggest that MBL2 is a modifier gene in CF, the possibility of type one error due to small sample size (105) cannot be dismissed at present. Also of

relevance for the further investigation of these polymorphisms in the CF population are the findings of Davis and colleagues (6), who have shown high levels of MBL2 binding to BCC isolates (obtained from colonized CF patients) and observed that MBL2 protein binding to BCC resulted in complement activation. However, the authors in their recent epidemiological study state that MBL2 may not be important in BCC infection (106), however their study design is limited to investigating susceptibility to BCC infection and not whether infection with BCC affects long term pulmonary disease course. In this study (106)they showed a decline in pulmonary function and oxygen saturation, and a higher frequency of hospital admissions for pulmonary infections requiring IV therapy over the short interval they investigated (one year interval) in CF patients who were MBL2 homozygous for the deficiency alleles. They did not show differences in infection rates with *P. aeruginosa* and BCC and homozygosity for the MBL2 deficiency alleles (106). Carlsson and co-workers (107) investigated deficiency alleles for MBL2 and MASP-2 in a CF cohort. In their study, they did not show an association between MBL2 or MASP-2 deficiency genotypes and pulmonary disease severity, but did show worse pulmonary disease severity in a subgroup of CF patients (N=27) who were colonized with *Staphylococcus aureus*. *Staphylococcus aureus* is a pathogen to which MBL2 protein has been shown to binds to (Table 4.2). Choi and associates showed worse %predFEV$_1$ in CF patients who were carriers of the MBL2 structural deficiency polymorphisms (104). In the study of Trevisiol and associates (102) of a small Italian CF cohort (N=47) showed lower mean %predFEV$_1$ and earlier median age of chronic infection with *P. aeruginosa* in MBL2 deficient patients. The authors state that the association of worse pulmonary function in MBL2 deficient patients was only shown in those CF patients who were infected with *P. aeruginosa* were carriers of a severe CFTR genotype(102). Considering the sample size and distribution of CF patients for MBL2 genotype, *P. aeruginosa* infection status and CFTR genotype their conclusions are based on a very basic descriptive analysis of the data. MBL2 deficiency genotype (0/0, A/0, XA/0, XA/XA) was studied in a larger multicenter study and they did not show an MBL2 deficiency genotype to be associated with severe pulmonary disease even when chronic infection with *P. aeruginosa* was considered in their models (108).

SP-A deficiency and SP-D deficiency are associated with increased risk of infection and death from infection in SPA and SPD knockout mice. LeVine and associates (67) infected mice lacking SPA, or SPD and wild-type mice with group B *Streptococcus* or *H. Influenza*

and showed deficient uptake of both pathogens by alveolar macrophages in the knock-out mice. There were differences in the bacterial killing, degree of inflammation and oxidant production between SPA knock-out mice and SPD knock-out mice. One study showed that SPD deficient mice developed chronic inflammation, emphysema and lung fibrosis (109). This study and others show that SPA and SPD have related but also potentially distinct functions in response to pathogen infection in the lung (67, 110-112). Sano and associates showed that SPA and SPD bind to LPS but differ in the mechanism by which they modulate the LPS-CD14 interaction (113). The SPD protein has been shown to enhance and suppress inflammatory mediator production and the outcome is dependent on the receptor SPD binds to and the binding orientation of the molecule (114). If the carbohydrate recognition domains of SPD protein are complexed with pathogen carbohydrates or lipid ligands then SPD functions to stimulate phagocytosis and proinflammatory responses, but if SPD is not complexed with ligands then it functions to block proinflammatory mediator production (57, 113). Leth-Larsen and colleagues hypothesize that the different molecular weight forms of SPD related to the coding polymorphisms are linked to the different functions associated with SPD (discussed below) (57).

The SP-A loci have been investigated as a site for candidate genes for respiratory distress syndrome (RDS) (69, 115-118). Kala and associates showed a positive association for RDS and the SP-A loci (69). Haataja and co-workers identified an SPA-1-SPA-2 haplotype ($6A^2$, $1A^2$) that was over-represented in RDS infants and showed that this association was dependent on the degree of prematurity and also homozygosity for threonine at the SPB (IIe131Thr) polymorphism (116). The SPA-1-SPA-2 haplotype $6A^3$-$1A^2$ was shown to be protective for RDS (115, 116). Homozygosity for the T allele for the SPD (Thr11Met; C32T) gene was shown to be associated with increased susceptibility to severe respiratory syncytial virus infection (70). Choi and associates investigated SPA gene polymorphisms in a CF cohort and showed significantly worse %predFEV_1 and S-K clinical scores in CF patients with the SPA-1 $6A^3$ allele and SPA-2 $1A^1$ allele, and this association was also present when investigating the SPA-1-SPA-2 haplotype ($6A^3/1A^1$)(104). SPA-1, SPA-2, and SPD gene polymorphisms were investigated as candidate modifier genes for COPD. Guo and associates showed an association between one silent SPA-1 (non amino acid changing) polymorphism at position 60 and increased risk for COPD (72). Lin and colleagues investigated all known polymorphisms in SP-A and SP-D as candidate genes variants in

acute respiratory distress syndrome (ARDS) and showed no significant associations between variant alleles and ARDS (119). Leth-Larsen and associates genotyped 206 Danes for the three coding polymorphisms in SPD and measured serum levels of SPD. They showed higher serum levels of SPD in those subjects who were TT versus CC and CT (mean value was 1035 ng/ml for TT versus 849.6 and 744.1 ng/ml for TC and CC) for the Thr11Met polymorphism and no significant differences in serum levels associated with the other two coding polymorphisms. There were two structurally different variant forms of the protein in serum from TT individuals (a high and low molecular weight protein). The low molecular weight form associated with the TT genotype was similar to the protein found in serum from CC individuals. The average distribution of high to low molecular weight SPD protein in serum was 1 to 1.6 for TT individuals and 1 to 5.1 for CC individuals. The high molecular weight variant was shown to preferentially bind to the Influenza A virus and Gram-positive and Gram-negative pathogens, while the low molecular weight variant bound to simpler ligands like LPS (57). The authors (57) also showed that the high molecular weight variants had differential:

- Binding capacities to pathogens,
- Aggregation,
- Clearance of pathogen.

Similarly Mikerov and associates(65) have shown differences in the abilities of SPA-1 and SPA-2 proteins to increase phagocytosis of *P. aeruginosa* by alveolar macrophages and showed that SPA-2 variants tested (i.e., $1A^0$ and 1A) were more active than SPA-1 variants ($6A^4$ and $6A^2$).

The investigation of polymorphisms in SPA and SPD as potential candidate genes in pulmonary disease pathogenesis is still in its infancy and while variants in these genes have been described, the exact effects these variant alleles have on gene transcription and translation and protein function are still unclear. Variations in the levels of surfactant proteins in patient populations as well as data from mouse knock-out studies of these proteins, and association studies investigating surfactant polymorphisms have shown possible associations with pulmonary diseases. There will be the added difficulty when investigating the SPA gene polymorphisms as this protein is encoded by two genes with several allelic polymorphisms. Current findings show that the surfactant genes are potential targets for investigations into pulmonary disease pathogenesis.

4.1.6 Review of the literature for *Burkholderia cepacia* complex infection in CF

In CF, impaired airway epithelial chloride transport leads to a milieu favoring colonization by bacteria, particularly *Staphylococcus aureus* and *P. aeruginosa*. Over two decades ago, another microbe, originally named *Pseudomonas cepacia* and now known as *Burkholderia cepacia complex* (BCC), came to the foreground as a colonizer of CF patients with a prevalence varying widely between clinics. In the literature this pathogen has been associated with an accelerated decline in clinical status and increased mortality in CF patients colonized with BCC versus *P. aeruginosa* (120-123), while others have noted a varied response with BCC colonization of CF patients with many reporting no difference in outcome (124-126). In our earlier case-control study, we showed increased mortality and pulmonary exacerbations in BCC infected CF patients, but no differences in pulmonary disease progression (change in %predFEV_1 and FVC in pre- and post-acquisition interval). Lewin and associates specifically showed in their retrospective study a higher mortality in the first year post-acquisition in 124 CF patients infected with BCC compared with a similar number of patients colonized with *P. aeruginosa* but, of interest, not in the second year following acquisition of BCC (124). Taylor and colleagues (127)showed a more rapid deterioration and increased mortality associated with CF patients exhibiting severe pulmonary disease (%predFEV_1<40%) at the time of BCC acquisition, while those patients infected with BCC exhibiting mild and moderate disease at the time of acquisition maintained a stable clinical status in that first year of monitoring. Co-infection with *P. aeruginosa* and BCC has been shown to result in a more severe decline in pulmonary function (123, 128, 129).

A number of more recent studies have reported worse outcomes for CF patients who acquire BCC. Chaparro and associates (130) reviewed the Toronto experience for lung transplantation in CF patients chronically colonized with BCC GEN IIIA, RAPD-type 2 (GEN=genomovar and RAPD= random amplification of polymorphic DNA) and showed reduced 1, 2 and 3 year survival compared to CF patients who received a lung transplant and were chronically infected with *P. aeruginosa*. De Boeck and colleagues reported on the Belgian experience of poor 5-year survival in CF patients infected with either GEN IIIA or GEN II (131). Aris and associates showned higher mortality in post-lung transplant CF patients who were chronically infected with GEN III compared to CF patients in their study

who were chronically infected with either GEN II, *P. aeruginosa* or other CF respiratory pathogens (132). CF patients infected with GEN IIIA strains in their study were all negative for the cable pilin gene (132), and therefore were not RAPD group 2. Infection with genomovar II is postulated to be associated with a more benign clinical course, based mostly on clinical experience, but this association has not been directly studied.

Infection with BCC has a number of implications for patient management in CF. Patient to patient transmission of BCC infection, at least of specific genomovar and RAPD groups, has been clearly shown (133-136). With the overall increase in mortality associated with BCC infection, clinics have instituted rigorous cohorting of CF patients infected with BCC from non-BCC infected CF patients. The possibility of variable pathogenicity of different BCC genomovars provides support for an extension of this cohorting by segregating BCC infected CF patients based on genomovar and RAPD grouping; however, this has yet to be investigated.

The heterogeneity in clinical course in CF patients infected with BCC may be related to variability in BCC species. Application of molecular techniques has established the complexity of BCC, with a number of distinct genomovars described. Currently, BCC has been ordered into ten genomovar groups identified and named as distinct species (137-146). Limited information is available to characterize any epidemiological BCC genomovar differences. The possibility of increased virulence of specific genomovars has been extrapolated from the observations of spread of common strains amongst CF patients with environmental contacts. BCC has been shown to adhere to respiratory epithelial cells (147, 148), and BCC secretory products induce the release of proinflammatory cytokines (interleukin 6 and 8) and prostaglandin E2 from these cells which is proposed to contribute to the excessive inflammation that characterizes this infection (149). Sajjan and co-workers(147) have shown that while CF airway mucus still functions in trapping BCC, it has impaired ability to kill the pathogen or prevent BCC related epithelial damage. It is also suggested that CF airway mucus may lack bactericidal agents (147). The GEN IIIA RAPD-type 2 (also known as ET12 lineage) has been found by Sajjan and colleagues (150) in the epithelial cells in the terminal and respiratory bronchioles, in hyperplastic epithelia, within inflamed alveolar septae and in luminal and parenchymal macrophages. BCC has been shown to form biofilms and co-colonize biofilms with *P. aeruginosa* (151). Others have also

shown that BCC persists in macrophages under in vitro conditions and is able to replicate and resist killing by oxidative burst (152, 153)**.** The pattern of distribution of BCC in the CF airways appears different from the common pathogens that colonize CF patients; *P. aeruginosa* is confined to the airway lumen. BCC, particularly in more severe cases, has been shown to cross the epithelial barrier into the parenchyma and into capillaries (154). With these advances, we can now examine whether variability in clinical course can be explained by genomovar differences.

MBL2 binding to BCC should result in complement activation but also in the activation of alveolar macrophages and therefore should modulate the lung defense system through the regulation of inflammatory cytokines. Garred has implied that susceptibility to BCC infection is associated with an MBL2 deficiency status (43). In their experience, they reported that seven of the 10 patients who became chronically infected with BCC during an 8-year data collection interval had an MBL2 deficient genotype.

In order to address the effect of modifier genes on BCC susceptibility and infection we first have to address a number of questions, which have not been previously investigated. At this time, there is only one study, a case control study, which addressed pulmonary disease progression pre-and post-colonization with BCC (155). Based on the BCC taxonomy information that is available it is important to elucidate whether the differences observed in pulmonary disease progression and survival are related to genomovar group and/or RAPD type group. At this time we cannot rule out that co-infection with *P. aeruginosa* may explain most of the heterogeneity observed in pulmonary disease progression and survival. Therefore, before addressing the specific questions related to MBL2 as a potential modifier gene in CF, these questions need to be addressed and will be in this study.

4.2 RESULTS

4.2.1 Hardy Weinberg equilibrium and linkage disequilibrium

Table 4.4 shows Hardy Weinberg equilibrium (HWE) results and the frequencies of the MBL2 and pulmonary surfactant SPA-1, SPA-2 and SPD gene polymorphisms in the study cohort. We showed deviations from HWE for two of the polymorphisms investigated in the MBL2 gene: MBL2-B and D where we showed a deficiency of heterozygotes (18% versus 27% for B-allele, p = 0.00001 and 12% versus 14% for D-allele, p=0.02; Table 4.4). The prevalence of the MBL2 promoter and structural polymorphisms were within the ranges reported in the literature for CF (43) and non-CF cohorts (29, 86, 87). Turner (81) summarized the literature concerning MBL2 allele frequencies and found the B allele varied between 0-16%, the C allele between 0-29% and the D allele between 0-6%. The frequency for the SPD (Thr11Met) allele was within the reported range for Caucasian populations (52.8-70.8%) (53, 57) with similar distribution of genotypes (TT-35.4%, CT-47.1% and CC-17.5%(57)). The SPA-1 polymorphism was also not in HWE and we showed a deficiency in the T-allele, specifically our cohort did not have any CF patients with a TT genotype (Table 4.4). We also showed a deficiency in heterozygotes for the SPA-2 (Thr9Asn) polymorphism (Table 4.4). Reported frequencies for the SPA genes SNPs are based on commonly occurring haplotypes based on known polymorphisms in the two genes as described in DiAngelo and associates (52), but we were unable to extrapolate the frequency of the alleles for individual SNPs as the percentages for the haplotypes are given without reference to counts.

In order to address the possibility that there may be preferential death in a genotype, we also examined HWE after stratifying the subjects by age (<25 and $\geq$ 25 years of age). Our results are shown in Tables 4.4 and 4.5. The deviation from HWE for the MBL2-B and MBL2-D alleles occurred, respectively, in the adult and children's group (Table 4.5 and 4.6). However, the values (frequency of alleles) were within reported ranges in the literature (81). We did not show deviations from HWE for the pulmonary surfactant polymorphisms when stratified by age (Table 4.6).

The results for LD are presented in Table 4.7. The MBL2 gene polymorphisms were not all in LD. The MBL2 gene polymorphisms (except for the B-allele) were not in LD with the

pulmonary surfactant polymorphisms investigated, and this observation concurs with literature findings (21). In our study, we showed weak LD between the MBL2 B-allele and the SPA-2 (A26C) polymorphism. We showed strong LD between the SPA-1 (C655T) and SPA-2 gene polymorphisms investigated in this study (A26C, A667C).

Inferred haplotype probabilities using the software package PHASE (156, 157) are presented in Tables 4.7 and 4.8. The inferred haplotypes for the four MBL2 SNPs inferred using PHASE (156, 157) are shown in Table 4.8. The low expression promoter for MBL2 was shown to always be found with the wild-type allele for the three exon 1 coding polymorphisms (i.e., for haplotype XAAA N=215, 22%) and six inferred haplotypes were shown with the high expression promoter allele and the coding polymorphisms, as previously documented in the literature (158).

Haplotypes probabilities were inferred for the four pulmonary surfactant gene polymorphisms investigated (Table 4.9), for SPA-1 with the two SPA-2 polymorphisms (Table 4.10), and for the two SPA-2 polymorphisms (Table 4.11). The CA haplotype was the most common inferred haplotype for the SPA-2 SNPs (frequency=61%). The CAC haplotype was the most frequent inferred haplotype across the three SPA SNPs investigated (frequency=44%). The inferred haplotype results were used for grouping our cohort for the SPA-2 SNPs into three groups based on having zero; one or two copies of the CA haplotype. Because of the number of inferred haplotypes for the three pulmonary surfactant genes and the low LD across SPA-1 and SPA-2 with SPD, we did not use the inferred haplotypes information to group the study cohort, instead we individually investigated the SPD polymorphism.

4.2.2 Descriptive data results and study cohort grouping

Table 4.12 presents the data for the cohort used for investigating MBL2 deficiency. In addition, the characteristics of the BCC (transiently and chronically infected) and control groups are presented in Section 4.2.7 and in Table 4.20. Susceptibility to BCC infection in MBL2 deficient versus wild-type CF patients was investigated using logistic regression. We did not show differences in BCC colonization in CF patients who were MBL2 deficient (5.7%) compared to MBL2 wild-type (5.3%, p=0.50).

The T allele for the SPA-1 gene polymorphism was rare and only occurred in heterozygous form in the study cohort. The descriptive clinical characteristics of the study cohort presented in Table 4.13 for the SPA-1 polymorphism were based on whether the T allele was present (i.e., CT genotype) or not (CC genotype). Inferred haplotypes generated by the PHASE program for the two polymorphisms (A26C and A667C) studied in the SPA-2 gene were used to group the study cohort based on whether there were zero, one, or two copies of the most common inferred haplotype CA (Table 4.14). SPD genotype was used to group the study cohort into three groups (Table 4.15).

4.2.3 Pulmonary disease progression: mixed effects regression on %predFEV$_1$

MBL2 gene: The subjects were grouped for MBL2 polymorphisms based on the functional effect of the polymorphisms as described in Garred and associates (43). We first investigated whether there was a difference in the rate of decline in %predFEV$_1$ with *P. aeruginosa* infection. We showed no association with rate of decline in %predFEV$_1$ and MBL2 deficiency status when controlling for CFTR genotype, *P. aeruginosa* infection status and gender.

Model 4.2.3-A

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + PA infection status (0/1=not infected/chronically infected) + CFTR class mild+ CFTR class heterozygous severe + CFTR unknown/unclassified + MBL2 deficiency * Time*
Base group was MBL deficient

In the reduced model (removing non-significant terms), we showed a similar rate of decline in %predFEV$_1$ irrespective of MBL2 gene grouping. The p-values for the interaction terms were not significant: MBL2 deficiency * Time (p=0.77), MBL2 deficiency * *P. aeruginosa* infection status (p=0.74), MBL2 deficiency * *P. aeruginosa* infection status * Time (p=0.66). The only significant interaction term in the model was *P. aeruginosa* infection status * Time (p=0.001), indicating a steeper decline in %predFEV$_1$ when chronically infected with *P. aeruginosa*.

Model 4.2.3-A (reduced model)

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + PA infection status (0/1=not infected/chronically infected) + MBL2 deficiency * Time*
Base group was MBL deficient

We next investigated whether there was a difference in the rate of decline in %predFEV$_1$ with BCC infection. Since there is the belief that different genomovars of BCC have a variable effect on pulmonary disease progression and survival, we first investigated whether this was the case and examined if there were differences in pulmonary disease progression post-acquisition of BCC, and whether there were differences based on the BCC genomovar. These analyses are presented at the end of the results section entitled 'BCC infection and pulmonary disease progression and survival' (in Section 4.2.7). Based on these results, which showed no differences in pulmonary disease progression in CF patients chronically infected with BCC GENII versus GEN IIIA, we did not stratify by BCC genomovar for the MBL2 analyses. Descriptive data on the study cohort available for statistical analysis are presented in Table 4.22.

In our linear mixed effects models we investigated whether MBL2 deficiency and chronic infection with BCC were associated with more rapid pulmonary disease progression in CF. In the first model we used a dummy variable to categorize patients into CF patients not infected with BCC and CF patients chronically infected with BCC (regardless of the genomovar group they were infected with). In this model, we did not distinguish the pre- and post-acquisition interval with BCC. Our model was similar to the model described above including all two-way, three-way and four-way interactions with the variable Time.

Model 4.2.3-B

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + PA infection status (0/1=not infected/chronically infected) + BCC infection status (0/1=not infected/chronically infected) + MBL2 deficiency * Time + PA infection status * Time + BCC infection status * Time).*
Base group was MBL deficient

The p-values for the interaction terms with MBL deficiency (i.e., MBL2 deficiency * Time, p=0.78; MBL2 deficiency * PA infection status * Time, p=0.46; MBL2 deficiency * BCC infection status * Time, p=0.81 and MBL2 deficiency * BCC infection status * PA infection status * Time, p=0.90) were not significant. The following interaction terms, which did not include MBL deficiency status were significant (PA infection status * Time, p=0.003; BCC infection status * Time, p=0.0006) showing a steeper decline in %predFEV$_1$ over time in CF patients who are chronically infected with either *P. aeruginosa* or BCC.

In the next model, we used the longitudinal data collected during the pre-acquisition with BCC interval for the CF patients chronically infected with BCC and compared to the longitudinal data collected from CF patients not infected with the BCC pathogen (Table 4.16, model 1). Patients chronically infected with BCC regardless of genomovar group were grouped together. We showed no significant differences (p=0.63) in the rate of decline in %predFEV$_1$ over time in MBL2 deficient versus wild-type CF patients.

Model 4.2.3-C (pre-acquisition of BCC)

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + PA infection status (0/1=not infected/chronically infected) + BCC infection status (0/1=not infected/chronically infected) + MBL2 deficiency * Time + PA infection status * Time + BCC infection status * Time).*

Base group for MBL2 was MBL deficient. Time for the BCC infected patients was BCC pre-acquisition interval

In the third model, the pre and post-acquisition intervals for BCC infection were included in the models as fixed and random effects. Patients chronically infected with BCC regardless of genomovar group were grouped together. In this model, we examined if the rates of decline in %predFEV$_1$ were different for MBL2 deficient and wild-type groups following infection with BCC (Table 4.16, model 2).

Model 4.2.3-D

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MBL2 deficiency (0/1=deficient/wild-type) + BCC infection status (0/1=no BCC infection/ BCC colonization) + AgePP (0/(1-+∞)=pre-colonization/days post-colonization) + MBL2 deficiency * AgePP).*

Base group for MBL2 is MBL deficient

Our results showed that significant differences existed between MBL deficient and wild-type groups and rate of decline in %predFEV$_1$ in CF patients chronically infected with BCC (p=0.001).

Additional comparisons are shown in Table 4.16 model 2 (labeled comparisons). In model 2 comparison 1, we have not distinguished BCC infection status (i.e., BCC infected or not infected) and simply examined pre and post-acquisition. Our results show a significantly different rate of decline in %predFEV$_1$ in CF patients who are MBL deficient versus wild-type (p=0.02). In comparison 2, we examined only controls (i.e., not infected with BCC) and showed no significant difference in the rate of decline in %predFEV$_1$ in CF patients who are MBL deficient versus wild-type (p=0.45). In comparison 3, we examined only CF patients chronically infected with BCC (not distinguishing genomovar group) and showed a significantly steeper rate of decline in %predFEV$_1$ in CF patients who were MBL2 wild-type (p=0.02). In comparison 4, we examined only CF patients who were MBL2 deficient and looked at the rate of decline in %predFEV$_1$ in those CF patients who were chronically infected with BCC versus controls (not infected with BCC) and showed a significantly steeper rate of decline in %predFEV$_1$ in CF patients who were MBL2 deficient and chronically infected with BCC (p=0.01). In our last comparison 5, we examined only CF patients who were MBL2 wild-type and looked at the rate of decline in %predFEV$_1$ in those CF patients who were chronically infected with BCC versus controls (not infected with BCC) and showed a significantly steeper rate of decline in %predFEV$_1$ in CF patients who were MBL2 wild-type and chronically infected with BCC (p<0.0001).

SPA-1 gene: In our linear mixed effects models we investigated whether having the less common CT versus CC genotype for the SPA-1 (C655T) polymorphism was associated with a different rate of decline in %predFEV$_1$. We showed a similar rate of decline in %predFEV$_1$ irrespective of SPA-1 genotype

Model 4.2.3-E

*%predFEV$_1$= Time + Sex (0/1=male/female) + SPA-1 genotype (0/1=CC/CT) + SPA-1 genotype * Time).*

Base group for SPA-1 was CC

The p-value for the interaction term SPA-1 genotype * Time was 0.57. In a second model we included *P. aeruginosa* infection status (0/1=not infected/chronically infected) including interaction terms (i.e., SPA-1 genotype * *P. aeruginosa* infection status (p=0.82), *P. aeruginosa* infection status * Time (p=0.13), SPA-1 genotype * *P. aeruginosa* infection status * Time (p=0.73)), but did not show significance.

Model 4.2.3-F

*%predFEV$_1$= Time + Sex (0/1=male/female) + SPA-1 genotype (0/1=CC/CT) + PA infection status (0/1=not infected/chronically infected) + SPA-1 genotype * PA infection status + SPA-1 genotype * Time + PA infection status * Time + SPA-1 genotype * PA infection status * Time*

Base group for SPA-1 was CC

In our final model (i.e., model 4.2.3-G) we included CFTR genotype with *P. aeruginosa* infection status (0/1=not infected/chronically infected) including the following interaction terms in our final reduced model: SPA-1 genotype * Time (p=0.005), SPA-1 genotype * *P. aeruginosa* infection status * Time (p<0.0001), and SPA-1 genotype * *P. aeruginosa* infection status * CF class (base group is homozygous for 2 severe mutations, p=0.02). In this model we showed that

- The slope is steeper if a patient has a CT genotype compared to a CC genotype.
- The mean %predFEV$_1$ is lower if they have at least one mild CFTR mutation, are chronically infected with *P. aeruginosa* and have a SPA-1 genotype of CC compared to any other 3-way combination (i.e., for the interaction term: SPA-1 genotype * *P. aeruginosa* infection status * CF class).

Model 4.2.3-G

*%predFEV$_1$= Time + Sex (0/1=male/female) + SPA-1 genotype (0/1=CC/CT) + PA infection status (0/1=not infected/chronically infected) + CFTR genotype + SPA-1 genotype * PA infection status + CFTR genotype * SPA-1 genotype + SPA-1 genotype * Time + PA infection status * Time + CFTR genotype * Time + SPA-1 genotype * PA infection status * Time + CFTR genotype * SPA-1 genotype * Time + CFTR genotype * SPA-1 genotype * PA infection status * Time*

Base group SPA-1 was CC. CFTR genotype is an abbreviation for the CFTR class groupings.

SPA-2 gene: Inferred haplotypes generated by the PHASE program for the two polymorphisms (A26C and A667C) studied in the SPA-2 gene were used to group the study cohort based on whether there were zero (SPA-2CA0), one (SPA-2CA1), or two (SPA-2CA2) copies of the most common inferred haplotype CA. The clinical characteristics of the cohort based on this grouping are presented in Table 4.14. We showed a similar rate of decline in %predFEV$_1$ based on our SPA-2 grouping, that is between SPA-2CA0 and SPA-2CA2 (p=0.37), between SPA-2CA1 and SPA-2CA2 (p=0.21) and between SPA-2CA0 versus SPA-2CA1 (p=1.00).

Model 4.2.3-H

*%predFEV$_1$ = Time + Sex (0/1=male/female) + SPA-2CA0 (0/1=1 or 2 copies of CA/0 copies of CA) + SPA-2CA1 (0/1=0 or 2 copies of CA/1 copy of CA) + SPA-2CA0 * Time + SPA-2CA1 * Time.*

Base group for SPA-2 was 2 copies of CA (i.e., SPA-2CA2).

In a second model we included *P. aeruginosa* infection status (0/1=not infected/chronically infected) including interaction terms (i.e., SPA-2CA0 and SPA-2CA1) * *P. aeruginosa* infection status (p=0.40 and 0.25), *P. aeruginosa* infection status * Time (p<0.0001), (SPA-2CA0 and SPA-2CA1) * *P. aeruginosa* infection status * Time (p=0.17 and 0.15)), but did not show significance with our gene polymorphism of interest.

Model 4.2.3-I

*%predFEV$_1$ = Time + Sex (0/1=male/female) + SPA-2CA0 (0/1=1 or 2 copies of CA/0 copies of CA) + SPA-2CA1 (0/1=0 or 2 copies of CA/1 copy of CA) + PA infection status + SPA-2CA0 * PA infection status + SPA-2CA1 * PA infection status + SPA-2CA0 * Time + SPA-2CA1 * Time + PA infection status * Time + SPA-2CA0 * PA infection status * Time+ SPA-2CA1 * PA infection status * Time.*

Base group for SPA-2 was 2 copies of CA (i.e., SPA-2CA2).

The model was reduced but did not show any significant results for the gene polymorphism (p-values for main effects SPA-2CA0 and SPA-2CA1 were 0.34 and 0.32, respectively), except for a significant *P. aeruginosa* infection status * Time interaction ($p<0.0001$) meaning a steeper decline in %predFEV$_1$ in those CF patients who were also chronically infected with *P. aeruginosa* versus not infected with the pathogen.

SPD gene: The grouping for the SPD gene polymorphism (C32T) was based on genotype into three groups: CC, TT and CT. We showed a similar rate of decline in %predFEV$_1$ based on our SPD genotype grouping.

Model 4.2.3-J

*%predFEV$_1$ = Time + Sex (0/1=male/female) + SPD-CC (0/1=CT or TT/CC) + SPD-CT (0/1=CC or TT/CT) + SPD-CC * Time + SPD-CT * Time.*

Base group was SPD-TT.

The p-value for the interaction terms SPD-CC * Time + SPD-CT * Time were 0.93 and 0.80, respectively, showing no difference in the rate of change in %predFEV$_1$ between SPD-CC and SPD-TT and between SPD-CT and SPD-TT. There was a similar rate of decline in %predFEV$_1$ for SPD-CC versus SPD-CT (p=0.86). In a second model we included *P. aeruginosa* infection status (0/1=not infected/chronically infected) including interaction terms (i.e., (SPD-CC and SPD-CT) * *P. aeruginosa* infection status (p=0.95 and 0.20), *P. aeruginosa* infection status * Time ($p<0.0001$), (SPD-CC and SPD-CT) * *P. aeruginosa*

infection status * Time (p=0.49 and 0.17)), but did not show significance with our gene polymorphism of interest.

Model 4.2.3-K

*%predFEV$_1$ = Time + Sex (0/1=male/female) + SPD-CC (0/1=CT or TT/CC) + SPD-CT (0/1=CC or TT/CT) + PA infection status + SPD-CC * PA infection status + SPD-CT * PA infection status + SPD-CC * Time + SPD-CT * Time + PA infection status * Time + SPD-CC * PA infection status * Time+ SPD-CT * PA infection status * Time.*
Base group was SPD-TT.

The model was reduced, but we did not show any significant results for the gene polymorphism (p-values for main effects SPD-CC and SPD-CT were 0.59 and 0.49, respectively), except for a significant *P. aeruginosa* infection status * Time interaction (p<0.0001) meaning a steeper decline in %predFEV$_1$ over time in those CF patients who were chronically infected with *P. aeruginosa* versus not infected with the pathogen.

Lastly, in a reduced model we combined the CT and CC group that is having one or two copies of the polymorphism versus being homozygous for the common allele (i.e., TT).

Model 4.2.3-L

*%predFEV$_1$ = Time + Sex (0/1=male/female) + SPD-CC/CT (0=TT and 1=CC/CT) + SPD-CC/CT * Time.*
Base group was SPD-TT.

In this model we showed similar rate of decline in %predFEV$_1$ for SPD-TT versus SPD-CC/CT (p=0.86). Then we included CFTR genotype and *P. aeruginosa* infection status (0/1=not infected/chronically infected) in the model including interaction terms (i.e., SPD-CC/CT * *P. aeruginosa* infection status (p=0.25), *P. aeruginosa* infection status * Time (p=0.0001), SPD-CC/CT * *P. aeruginosa* infection status * Time (p=0.14)). CFTR genotype was not a significant predictor and interaction terms were dropped from the final reduced

model above, leaving the variable in the model only as a main effect. We did not show significance with our gene polymorphism of interest.

4.2.4 Survival analysis

Our survival models were run using CFTR class grouping and alternatively using PSS. The reason for this was that the number of CF patients with a mild CFTR genotype (i.e., homozygosity or heterozygosity for class 4 or 5 mutation) was small and in some cases mild CFTR genotype was not represented in our gene polymorphism groupings and this caused estimation problems with the statistical models.

MBL2 deficiency status was not a significant predictor of survival. The first question we addressed was:

Is there a difference in survival (i.e., time to death or lung transplantation) between MBL2 deficient versus MBL2 wild type CF patients? MBL2 deficiency was not a significant predictor of survival (Table 4.17).

The second question we addressed was:

Is there a difference in survival based on MBL2 genotype and whether the CF patient is chronically infected with BCC? MBL2 deficiency was not significantly associated with age of chronic BCC infection (RR=0.93, $CI_{0.95}$: (0.5-1.75), p=0.83) or age of transient or chronic infection with BCC (RR=0.98, $CI_{0.95}$: (0.60-1.62), p=0.95). CF patients used in the investigation of the effect of BCC infection and genomovar species on pulmonary disease progression and survival were not included in entirety in the MBL2 analyses as some of the patients died prior to collection of blood for MBL2 genotyping.

To explore the possibility of a potential bias in our cohort who consented to the study and blood was obtained for genotyping (versus BCC infected CF patients for whom we were unable to obtain blood for MBL2 genotyping before their death), we looked at the number of CF patients whom we had genotyped for MBL2 and had died or had received a lung transplant for any major differences in patient numbers based on MBL2 genotype. Of the CF patients chronically infected with BCC who were genotyped for MBL2 there were six deaths

in the MBL2 wild-type and six deaths in the MBL2 deficient groups. In the group transiently infected with BCC there were four and three deaths in the MBL2 wild-type and deficient groups, respectively. There were no BCC infected patients who had undergone lung transplantation and were MBL2 deficient. There were three and two patients, respectively, who were lung transplantation recipients and infected either chronically or transiently with BCC, and had a MBL2 wild-type genotype. Of these transplanted patients noted above only one of them chronically infected with BCC was still alive. It would appear that a descriptive examination of the data does not reveal differences in the number of deaths in BCC infected CF patients who were MBL2 wild-type or deficient in our study cohort. However, of the lung-transplanted patients an MBL2 wild-type genotype was found in all cases in our study cohort. It appears that our study cohort was not biased for MBL2 genotype. MBL2 deficiency was not predictive of survival (Table 4.17).

SPA-1 genotype, SPA-2 haplotype, and SPD genotype were not predictive of survival. CFTR class, *P. aeruginosa* infection status, sex and modifier gene * *P. aeruginosa* infection status were included in the models. The results are presented in Table 4.17.

4.2.5. Effect of modifier genes on *P. aeruginosa* infection status

We investigated whether polymorphisms in the pulmonary surfactant genes and MBL2 deficiency contributed to infection with *P. aeruginosa*. We separately investigated age of first infection and age of chronic infection with the pathogen and results are presented in Tables 4.17 and 4.18, respectively. CFTR class was included in the models. We did not show a significant effect of MBL2 deficiency and genotype for SPA-1, SPA-2 and SPD on age of first and chronic infection with *P. aeruginosa*.

4.2.6 Frequency of pulmonary infections requiring intravenous antibiotic therapy

We used Poisson regression to investigate the effect of the modifier gene polymorphisms on the mean frequencies of pulmonary infections requiring IV therapy. We investigated two time intervals:

- Two years retrospectively; the interval was January 1, 1998 to January 1, 2000 (N=357),

- Up to 5 years retrospectively; the interval was January 1, 1995 (is the earliest possible start interval) to January 1, 2000 (N=357).

We only analyzed patients who had data within the 2 year interval labeled A. The inclusion criteria were:

- Patients in the data set had to have a clinic visit after January 1, 2000.
- Patients had to have data for at least the 24-month interval prior to January 2000.

In the case of the 5-year interval, the patients who had more than 2 years but less than 5 years of data were used in this analysis and the analysis was adjusted for the time interval available for the patients. All patients included in the 2-year interval had data to be included as well in the 5-year interval. The response variable was number of pulmonary infections requiring IV therapy with separate models used for intervals A and B. We accounted for gender, *P. aeruginosa* infection (BCC infection status for MBL2 deficiency only) status and CFTR class. In the case where not all four CFTR classes were represented in the study cohort, we used the variable PSS.

We showed a significant difference in the mean frequency of IV therapy for pulmonary infections over 5 years (32% higher) in CF patients who were CT versus CC for the SPA-1 polymorphism (p=0.005, N=357). For the SPA-2 polymorphisms investigated, we used the number of copies of the CA haplotypes as the grouping variable. We showed that CF patients who had zero copies of the CA haplotype had significantly lower mean frequency of pulmonary infections over 5 years by 26.5% than CF patients who had 2 copies of the CA haplotype (p=0.01, N=357). We did not show significant differences in the mean frequency of pulmonary infections in CF patients who had one versus two copies of the CA haplotype (p=0.69). The 2-year analyses were not significant for either the SPA-1 or SPA-2 polymorphisms investigated.

For the SPD polymorphism we showed that having a CT genotype was associated with a lower mean frequency of pulmonary infections over the 2-year (21%, p=0.04) and 5-year (21%, p=0.002) intervals compared to CF patients with a TT genotype. There were no significant differences in the mean frequency of pulmonary infections over the 2-year and 5-year interval investigated for CF patients who were CC versus TT. Our analyses also showed that CF patients who were chronically infected with *P. aeruginosa* had significantly

higher mean number of pulmonary infections compared to those CF patients not infected with the pathogen for both the 2 (28.0%, p<0.0001) and 5 (33.1%, p<0.0001) year time intervals.

MBL2 deficiency was not associated with significant differences in the mean frequency of pulmonary infections over the two (p=0.32, N=356) and 5-year (p=0.89, N=356) interval. To examine MBL2 deficiency including the variable BCC infection status we used the variable PSS as a measure of disease severity instead of CF class since all CF patients who were chronically infected with BCC in our cohort were homozygous for severe CFTR mutations. In our first model we used as covariates: gender, BCC infection status and PSS. We did not show significant differences in the mean frequency of pulmonary infections over two (p=0.14, N=356) and five years (p=0.59, N=356) for MBL2 deficient CF patients, however chronic infection with BCC was associated with significant differences in mean frequency of pulmonary infections over two (p<0.0001, N=356) and five years (p<0.0001, N=356). In the second model, we also included as a covariate *P. aeruginosa* infection status. We did not show significant differences in the mean frequency of pulmonary infections over two (p=0.43, N=356) and five years (p=0.72, N=356) for MBL2 deficient CF patients.

4.2.7BCC infection and pulmonary disease progression and survival

4.2.7.1 BCC infection and pulmonary disease progression

The characteristics of the BCC (transiently and chronically) infected and control groups are presented in Table 4.20. We identified 93 patients who were transiently or chronically infected with BCC and longitudinal data were available on 90 of these individuals. The majority of CF patients infected with BCC were infected with genomovar II and IIIA. There were 16 (10) and 48 (2) CF patients chronically (transiently) infected with GEN II and GEN IIIA, respectively. There were 1(0), 0(4), 1(1) and 5(3) CF patients chronically (transiently) infected with GEN I, IIIB, IV and V, respectively. The small numbers prevented us from including these latter groups in our models. There were no CF patients who carried known mild CFTR genotypes (homozygous or heterozygous for class 4 or 5 mutations) and were either transiently or chronically infected with BCC (p=0.0001). The remainder of the cohort consisted of:

- Fifty five patients who were homozygous for CFTR class 1, 2, or 3 mutations,

- Eighteen patients who were heterozygous for class 1, 2 or 3 mutation and one mutation classified as unknown or unclassified and
- Twenty one CF patients who were homozygous for two unknown or unclassified CFTR mutations.

The mixed effects regression models were initially run including CFTR genotype as a fixed effect. Given that BCC infected patients were characterized only with severe CFTR genotypes, we alternatively used pancreatic sufficiency status (PSS) as a measure of CF disease severity. PSS proved not to be a significant predictor in our models and thus both variables (CFTR genotype and PSS) were excluded from the final models (Table 4.21). The effect of the center variable was included as a random effect; however, center-to-center variability was small (less than 1% of the total variability).

We have previously shown that CF patients who become chronically infected with BCC do not differ in clinical course from CF controls (without BCC infection) prior to infection with BCC (155). In the present study, we investigated the estimated rate of decline in $\%predFEV_1$ over time for up to 5 years pre-acquisition of BCC using linear mixed effects models to validate our previous results. Our analysis showed similar rate of decline in $\%predFEV_1$ in the pre-acquisition interval for transiently (-2.09 %/year) and chronically (-1.51 %/year) BCC infected groups and for the controls (-1.38 %/year). Table 4.21 summarizes our linear mixed models results for the post-BCC acquisition interval. In model 1, we showed a more rapid decline in $\%predFEV_1$ in CF patients chronically infected with BCC compared with CF controls not infected with BCC ($p=0.04$). Comparisons were also made to the non-infected control group (Table 4.21 model 1). In this model we treated the transiently BCC infected group as a separate group (not shown in Table 4.21). Our results show the transiently BCC infected group has an estimated rate of decline in $\%predFEV_1$ over the follow up time which is intermediate between the non-infected controls ($p=0.13$) and the chronic BCC infected patients ($p=0.43$). The rate of decline in $\%predFEV_1$ in the transiently BCC infected group is not statistically significantly different from either the control or the chronically infected with BCC groups. The rate of decline in $\%predFEV_1$ was similar when comparing different BCC genomovars and RAPD types in chronically infected CF patients (Table 4.21, model 2). In this model, we further investigated the rate of change in $\%predFEV_1$ between CF patients infected with GEN II and GEN IIIA. The rate of decline in $\%predFEV_1$ in the post-acquisition interval with BCC was not statistically different for GEN

IIIA RAPD-type groups, these groups were collapsed and the analysis rerun for the truncated GEN IIIA group and the results are shown in Table 4.21 (model 2- labeled collapsed GEN IIIA).

4.2.7.2 BCC and *P. aeruginosa* infection and pulmonary disease progression

We next investigated whether concurrent chronic infection with *P. aeruginosa* would further contribute to decline in %predFEV$_1$. These results are shown in Table 4.21 model 3. The *P. aeruginosa* infected CF patients were treated to have a single rate of decline in %predFEV$_1$ prior to their infection with BCC and we are assuming that this single line is a good representation of the patients' rate of decline in %predFEV$_1$ for the BCC post-acquisition interval. While the majority of co-infected patients first harbored *P. aeruginosa*, seven acquired the organism after BCC infection. Because of the small numbers, however, we were unable to investigate the order of co-infection in our regression models, although our main interest was whether there was a cumulative effect of infection with both pathogens. Of the CF patients colonized with *P. aeruginosa*, 12 were transiently infected and 36 were chronically infected with BCC. The mean time from age of first infection with *P. aeruginosa* to co-infection with BCC was 3.7(1.8) years for transiently infected and 5.2(1.2) years for patients chronically infected with BCC. There were seven patients who were chronically infected with BCC and subsequently became infected with *P. aeruginosa* (time to co-infection with *P. aeruginosa* 0.6-13.5 years post-acquisition of BCC). CF patients who were already infected (or were identified within 8 months post-BCC infection to be infected with *P. aeruginosa*) were included for analyses in model 3 in the co-infection group. In model 3, we investigated the rate of decline in %predFEV$_1$ for CF patients infected chronically with *P. aeruginosa* or BCC, or both pathogens versus those infected with neither pathogen. Our results show that concurrent chronic infection with BCC (regardless of genomovar) and *P. aeruginosa* contributed to a greater deterioration in %predFEV$_1$ compared to the decline in %predFEV$_1$ exhibited by BCC or *P. aeruginosa* infection alone or CF controls infected with neither pathogen.

4.2.7.3 BCC infection and survival

We addressed three questions in our survival analysis with Cox Proportional Hazards modeling:

Question 1: Are there survival differences between transient and chronic BCC infected patients and patients not infected with BCC? Gender and PSS (but not CFTR class groups) were significant effects and were therefore included in the model. We showed that BCC infection, whether chronic (RR =3.93, $CI_{0.95}$(2.4, 6.4), p<0.0001) or transient (RR=6.46, $CI_{0.95}$(2.9, 14.4), p<0.0001) (relative to controls), adversely affected survival.

Question 2: Are there survival differences among CF patients chronically or transiently infected with BCC and controls with and without co-infection with *P. aeruginosa*? In this model, we compared transiently and chronically BCC infected groups against our BCC uninfected cohort but in this case divided the uninfected with BCC group into 2 groups; those who were chronically infected and those who had never been infected with *P. aeruginosa*. The BCC infected group was also divided for a further analysis in BCC infected CF patients with and without co-infection with *P. aeruginosa*. We were unable to explore the question of whether there are survival differences with co-infection versus infection with *P. aeruginosa* or no pathogen including the BCC transiently infected group due to estimation problems fitting the model (small N for BCC transiently infected patients with none of the patients chronically infected with *P. aeruginosa*). We did not adjust for the order of infection with *P. aeruginosa*, i.e., whether *P. aeruginosa* preceded or followed BCC infection, as we were interested in whether there was an additive effect of infection with both pathogens. We showed:

- Chronic infection with *P. aeruginosa* in the control group was associated with a significantly higher likelihood of death or lung transplant compared with non-infection with both pathogens (RR=4.1, $CI_{0.95}$(1.4, 12.0), p=0.01).
- CF patients who were chronically (RR=13.6, $CI_{0.95}$(4.4, 42.0), p=0.0001) or transiently (RR=22.7, $CI_{0.95}$(6.2, 83.0), p=0.0001) infected with BCC showed a significantly higher likelihood of death or lung transplant relative to non-infected controls.
- CF patients chronically infected with BCC and co-infected with *P. aeruginosa* showed a significantly higher likelihood of death or lung transplant (RR=6.7, $CI_{0.95}$(1.7, 26.8), p=0.007) than non-infected controls.
- There were no significant differences in death or lung transplant between CF patients chronically infected with BCC only and controls infected with *P. aeruginosa* (RR=2.0, $CI_{0.95}$(0.8, 4.7), p=0.12).

- There were no significant differences in death or lung transplant between CF patients chronically infected with BCC only and chronically infected with BCC with co-infection with *P. aeruginosa* (RR=1.6, $CI_{0.95}$(0.6, 3.8), p=0.33).
- CF patients chronically infected with BCC with co-infection with *P. aeruginosa* showed a significantly higher likelihood of death or lung transplant (RR=3.1, $CI_{0.95}$(1.7, 5.4), p<0.0001) than controls infected with *P. aeruginosa*.

Question 3: Are there survival differences between genomovar IIIA and genomovar II infected patients? We showed that genomovar IIIA infected patients have a higher risk of death or lung transplant compared to genomovar II infected patients (RR=7.60, $CI_{0.95}$(0.9, 57.5), p=0.05). Our analysis did not show differential survival of genomovar IIIA RAPD-type 1, 4, 6 infected patients compared to genomovar IIIA RAPD-type 2 (RR=2.0, p=0. 18).

4.2.7.4 Frequency of pulmonary infections requiring IV therapy 12 and 24 months pre- and post-colonization with BCC

We did not show differences in the frequency of pulmonary infections treated with IV antibiotic therapy in the BCC GEN IIIA versus GEN II group, nor in the GEN IIIA RAPD-type 2 versus GEN IIIA RAPD-type 1, 4, 6 groups for:

- The 12 and 24 month pre-colonization (p>0.05) interval investigated, and
- The 12 and 24 month post-colonization intervals investigated (p>0.05) (results not shown).

4.3 DISCUSSION

4.3.1 Innate immunity genes: MBL2, SPA and SPD

The innate immunity genes investigated in this study produce proteins, which are involved in the host response to bacterial infection. The pulmonary surfactant SPA-1, SPA-2 and SPD genes produce proteins that are predominately involved in innate immunity in the lung, whereas the MBL2 protein is identified as an acute phase serum protein produced by the liver. In this study we investigated polymorphisms in these four genes as potential modifiers of pulmonary disease severity and progression in CF. These genes are located on the long arm of chromosome 10 and are closely related in structure and function (26, 27). The MBL2 polymorphisms have been investigated, although the effects of these polymorphisms have not been clearly elucidated. Polymorphisms on the pulmonary surfactant genes have been identified and their effects are being investigated. Our study showed that CF patients who were homozygous for the C allele for the SPA-1 gene polymorphism were more likely to become chronically infected with *P. aeruginosa* at an earlier age and to experience a steeper decline over time in %predFEV_1. However, the CT genotype was associated with a steeper decline in %predFEV_1 over time and a higher mean frequency of pulmonary infections compared with the CC genotype. We used the inferred haplotypes generated by the program PHASE and grouped our subjects based on the number of copies of the common CA haplotype present for the SPA-2 gene polymorphisms investigated. We showed no differences in survival, age of first and chronic infection with *P. aeruginosa* and pulmonary disease progression. A 26.5% lower mean frequency of pulmonary infections over 5 years was shown for CF patients who had zero versus two copies of the CA haplotype. Our subjects were grouped based on genotype for the SPD polymorphism investigated and we showed a significant finding for frequency of pulmonary infections (i.e., higher frequency) requiring IV antibiotic therapy in patients who were TT versus CT genotype (and no differences for CC versus either CT or TT).

The polymorphisms investigated in the MBL2 gene are not in LD (except MBL2-B with SPA-2) with the pulmonary surfactant gene polymorphisms investigated in this study. Therefore the MBL2 gene polymorphisms were investigated as independent potential modifiers in CF. The functional effects of the promoter polymorphism and the three exon 1 polymorphisms are to reduce levels of circulating MBL2 protein; thus the grouping was

based on these functional effects (i.e. MBL2 wild-type and deficient groups). We showed that MBL2 deficiency was associated with pulmonary disease progression in CF, but only in those patients who became infected chronically with BCC. Our analyses revealed that MBL2 deficiency is not associated with susceptibility to BCC infection, but that once infected with the pathogen CF patients who have a wild-type genotype experience a more rapid decline in %predFEV$_1$ over time. MBL2 deficiency was not associated with earlier infection with *P. aeruginosa.* MBL2 deficient CF patients who were chronically infected with *P. aeruginosa* did not show a higher frequency of pulmonary infections requiring IV antibiotic therapy, nor was infection with this pathogen associated with a different (i.e., more rapid) pulmonary disease progression in MBL2 deficient CF patients.

The MBL2 genotype resulting in normal levels of the protein was shown in the current study to be predictive of a greater rate of decline in %predFEV$_1$ following chronic colonization with BCC. This result was contrary to our hypothesis and differs from studies which have explored the association between MBL2 genotype and pulmonary disease in CF. Garred and associates(43) reported increased susceptibility to BCC infection in CF. In their experience 7 of the 10 patients who became chronically infected with BCC during the study interval had an MBL2 deficient genotype (43). They reported in their CF cohort better pulmonary function in their MBL2 wild type group (defined as YA/0 and A/A) (N=130) and showed an improvement in pulmonary function over the 8 years that their cohort was retrospectively followed (43). Gabolde and associates (103) matched the 11 MBL2 deficient patients out of 169 CF patients who were genotyped as MBL2 wild-type on sex and age and showed a lower mean %predFEV$_1$ in the MBL2 deficient and a trend towards higher frequency of colonization with *P. aeruginosa.*

There were limitations, however, in some previously published studies that question the associations between susceptibility to infection, disease progression and MBL2 deficiency. The study of susceptibility to infection in Garred and associates (43) was merely a descriptive presentation of affected CF patients with BCC infection in their cohort. In this study, they show a progressive improvement in pulmonary function over the study interval in the MBL2 wild-type group, but is not addressed in the article; instead they discuss mean pulmonary function over the study interval. In the case-control study by Gabolde and associates(103) a lower mean pulmonary function in age and sex matched MBL2 wild-type

and deficient CF patients was shown. The sample size in this study was small. In addition, in the study by Garred and associates (43) differences at baseline were noted between MBL2 wild-type and deficient groups. The mean %predFEV$_1$ for their MBL2 wild-type group was approximately 68% at the start of the data collection interval compared to approximately 53% in their MBL2 deficient group (defined as XA/0 and 0/0). The longitudinal data did not show a significant decline over time. The initial difference in %predFEV$_1$ in the two MBL2 groups could reflect a coincidental higher rate of *P. aeruginosa* colonization in the MBL2 deficient group, a factor not controlled for in their longitudinal analysis. Factors other than environment and modifier genes, which could affect pulmonary disease severity and progression in CF patients, is CFTR genotype (mild versus severe genotype). While these factors alone do not explain pulmonary disease severity and progression, they do contribute and therefore should be controlled for in the statistical models, especially when examining a small cohort. Their results were, however, corroborated in a later study on an independent cohort of similar size by Davies and associates (106) who showed lower lung function and oxygen saturation, and a higher frequency of hospital admissions for pulmonary exacerbations requiring IV therapy over a one year interval investigated in MBL2 deficient CF patients.

In a more recent study Choi and associates (104), they only showed more severe pulmonary disease in a sub-group of CF patients, who were MBL2 deficient and carriers of two severe CFTR mutations (homozygous delta F508). In another study, this one by Carlsson and co-workers (107), they showed worse pulmonary disease severity, but only in those CF patients (N=27) who were colonized with *Staphylococcus aureus*; a pathogen documented to be recognized by the MBL2 protein. Others and we have shown that there is a steeper decline in %predFEV$_1$ over time when CF patients are chronically infected with *P. aeruginosa*. There is no doubt that repeated infections with *Staphylococcus aureus* and *Hemophilus influenza*, which in the early stages of CF pulmonary disease are the most common CF infecting pathogens, set up the environment in the lungs conducive to subsequent chronic infection with *P. aeruginosa* (1, 2). This does not necessarily mean that MBL2 genotype accounts for some of this variability. Chronic infection with *P. aeruginosa* should be controlled for when investigating the effect of MBL2 deficiency on pulmonary disease and survival. Therefore, a positive association between MBL2 deficiency alleles and *P. aeruginosa* infection may be indirectly due to other factors such as early recurrent infections

with other pathogens, or recurrent or chronic infection with the pathogens *BCC or Staphylococcus aureus* at the time of investigation, or biased sample. However, whether this difference in disease progression is directly related to MBL2 deficiency is not definitively answered by previous studies (43). In summary, previous findings on CF cohorts were based on small cohort sizes (i.e., 146-164) with a small number of BCC infected patients (6.8%), BCC genotype was not documented, *P. aeruginosa* infection status was not accounted for in their models, and pre- and post-colonization with BCC clinical status was not investigated, study interval and analyses did not reflect long-term CF pulmonary disease progression.

In our study, MBL2 deficiency and wild-type groups were similarly defined as in previous studies (43, 103). Unlike previous studies, we stringently classified our BCC infected patients according to BCC infection status, BCC genomovar grouping, and collected %predFEV$_1$ data spanning the pre-and post-colonization interval (and only included values from stable encounters) for use in our linear mixed effects regression models. By doing so, we showed MBL2 wild-type genotype to be associated with more rapid decline in %predFEV$_1$ in CF patients infected with BCC. There was a similar rate of decline in %predFEV$_1$ for controls and for BCC infected patients in the pre-acquisition interval. Our findings show that once colonized with this pathogen, pulmonary disease progression proceeds at a more accelerated rate in MBL2 wild-type CF patients. Our results shed light into the contradictory results in previous studies. It is reasonable to show an increased frequency of infections in non-BCC infected MBL2 deficient CF patients, and a steeper rate of decline in pulmonary function limited to MBL2 deficient CF patients infected with *Staphylococcus aureus*. Worse pulmonary disease severity in MBL2 deficient CF patients infected with *Staphylococcus aureus* was shown by Carlsson and associates(107). The ability of BCC to survive and multiply intercellularly would then result in a steeper rate of decline in pulmonary function in BCC infected CF patients who are MBL2 wild-type. We also looked at survival and found that MBL2 genotype was not a significant predictor of survival when age at BCC acquisition and BCC infection status were included in the model. The possibility exists that our study subjects infected with BCC may be biased for MBL2 genotype as we only have MBL2 genotype data on CF patients recruited for the study from 1997 onwards. Therefore, selection bias may be a confounder if the MBL2 gene is a

modifier gene in CF. Follow-up for a longer duration will ultimately help to elucidate the relationship between MBL2 genotype and survival in BCC infected CF patients.

The design of our study did not permit us to elucidate mechanisms responsible for increased pulmonary disease progression in MBL2 wild-type genotype CF patients infected with BCC. One potential mechanism is a heightened immune response in the airways of MBL2 wild-type patients. This more exuberant inflammatory response could contribute to airway damage accelerating the manifestation of structural changes such as bronchiectasis and further the cycle of bacterial proliferation and fixed airflow obstruction seen in CF. An environment favorable to BCC colonization exists in the CF airways: mucus hypersecretion and dehydration leading to inspissation of secretions, increased adherence of thick mucus to airway cells, ciliary dysfunction and potentially reduced or inactivated bactericidal agents (147) all contribute to colonization. We propose that the uptake of opsonized BCC by macrophages may serve as a vehicle for intracellular survival of this organism. Davies and associates (6) showed that MBL2 protein binds in significant amounts to BCC isolates obtained from colonized CF patients and in intermediate levels with mucoid strains of *P. aeruginosa* obtained from colonized CF patients. MBL2 binding to BCC specimens resulted in complement activation as measured by C4 deposition (6). A heightened immune response in MBL2 wild-type CF patients may also contribute to lung damage. Younger and co-workers have also shown that MBL2 opsonizes *P. aeruginosa* (strain UI-18), although their results show that the alternative pathway and the surfactant proteins account for the majority of opsonization of *P. aeruginosa* (31).

A precedent for MBL2 sustaining infection has been reported for other infections in both normal and immuno-compromised hosts. In this respect, wild-type MBL2 protein levels are positively associated with infection with *M. tuberculosis* or *Mycobacterium leprae* (92, 97, 101), *Plasmodium falciparum* (98), *Leishmania chagasi* (100), and *Cryptosporidium parvum* in AIDS patients (99). Garred and colleagues concluded that during infection with intracellular pathogens the opsonic mechanisms of host defense are used by some pathogens to invade host cells (97, 101). BCC has also been shown *in vivo* to survive and replicate within macrophages (152, 153). Furthermore BCC (GEN IIIA RAPD-type 2) has been found to adhere to respiratory epithelial cells (147, 148) and has been found in respiratory epithelial cells and other airway cells from CF lung explants (150). Garred and associates

showed a high incidence of *M. tuberculosis* and *M. leprae* in Ethiopians with MBL levels associated with wild-type MBL2 genotype. These pathogens seem to depend on phagocytosis as their mode of entry into host cells (92, 97). While there have been no previous data to link BCC infection and MBL levels and genotype, our data and the previous literature support our proposal that BCC may be another pathogen capable of intracellular survival and replication in macrophages and possibly other cells. The pattern of distribution of BCC in the CF airways seems to differ from the common pathogen that colonize CF patients, *P. aeruginosa* which is confined to the airway lumen. BCC has been shown to form biofilms and co-colonize biofilms with *P. aeruginosa* (151), but has also been shown to cross the epithelial barrier into the parenchyma and into capillaries (154). BCC secretory products induce the release of proinflammatory cytokines (interleukin 6 and 8) and prostaglandin E2 from respiratory epithelial cells which are proposed to contribute to the excessive inflammation that characterize BCC infection (149).

Given the discrepancy between our results and previous work on CF cohorts (43, 103, 106), the relationship between pathogen susceptibility, disease progression and MBL2 genotype needs to be examined in a larger independent CF cohort. This will necessitate characterization and approximation of disease severity at the onset of BCC acquisition and a significant duration of follow-up to properly address important disease outcome measures. In addition, a record of infection with the pathogen .*Staphylococcus aureus* during the study interval would also seem prudent since it is a pathogen that MBL protein recognizes and binds to. Such work is important in gaining a better understanding of the reasons for heterogeneity of bacteriology and pulmonary disease progression in CF and in ultimately designing intervention strategies.

The T-allele in the SPA-1 polymorphism (Arg219Trp) was not found in homozygous form in our study cohort and was not in Hardy Weinberg equilibrium (Table 4.4). The expected count was five (Table 4.4). To ensure that the deviation from HWE was not due to genotyping error, we sequenced patients who were identified as having a CT and CC genotype and sequencing confirmed the genotyping assay. Of relevance to this present study is the differential survival of CF patients with a specific genotype for the modifier gene investigated which would then affect the distribution of genotypes for this gene polymorphism in the surviving CF population. There were no CF patients with a SPA-1

genotype TT for the polymorphism investigated. Deviations from HWE occur when there is consanguineous or assortative mating, natural selection, population stratification, genotyping error or differential survival in the presence of a contributing risk factor; in this case, CF. This deviation in HWE is unlikely to be due to population stratification, as our study sample was predominantly Caucasian CF patients. Consanguineous or assortative mating is also unlikely. Genotyping error due to error in reading the gels or entering the genotype onto the data sheet was kept low by having two individuals read the gels and the data entered was triple checked, although by one person. To validate the genotyping, a number of samples were sequenced, which confirmed our genotyping results.

A logical explanation for this deviation from Hardy-Weinberg equilibrium could be natural selection, which would mean that homozygosity for the T allele (or some other polymorphism in the SPA-1 or SPA-2 genes, which is in LD with the SPA-1 (Arg219Trp) polymorphism, has a significant effect on survival. If this is limited to CF patients then the deviation in Hardy-Weinberg equilibrium could be due to differential survival in CF patients. Differential survival is the most probable reason that explains our results. CF patients with a TT genotype may either die at a very young age; we did not include CF patients in our study sample who were less than 5 years old, as we would not be able to investigate our primary outcome variable, that is, pulmonary disease severity and progression. It is also possible that a CF fetus with a TT genotype is aborted during the gestation period or dies within the first years of life and may be either diagnosed or not as having CF at the time of death. We did not show a significant survival advantage in our CF cohort for CF patients with a CC genotype relative to CF patients who carried the T-allele at this locus. It appears that homozygosity for the T-allele may have a deleterious effect on survival and therefore is rare.

A number of polymorphisms have been reported in the SPA-1 gene and the SPA-2 gene that are in LD (22). There are additional polymorphisms in the SPA-1 gene that are amino acid changing (i.e., Ala19Val and Leu50Val) or silent (non-amino acid changing; Pro62Pro). The SPA-1 and SPA-2 gene haplotypes formed by the known polymorphisms across the two genes have been defined and are denoted as alleles 6A-$6A^4$ for SPA-1 and 1A-$1A^6$ for SPA-2 (52). There are four SPA-1 and SPA-2 gene haplotypes (i.e., $6A^2$-$1A^0$, $6A^3$-$1A^1$, $6A^3$-$1A^0$ and $6A^3$-$1A^2$) which are observed in Caucasians (53). The T-allele for the SPA-1

(Arg219Trp) polymorphism investigated in this study is only observed in the less common SPA-1 haplotype, $6A^4$. A survival effect therefore could potentially explain the deviations from Hardy-Weinberg equilibrium observed in the study cohort. In our study, haplotypes using PHASE were reconstructed for the two SPA-2 polymorphisms (Table 4.11) and for the SPA-1 and SPA-2 polymorphisms (Table 4.10). We identified the T-allele in two of the six reconstructed diplotypes.

We also investigated one of the three known amino acid changing polymorphisms in the SPD gene (C32T; Thr11Met). We showed LD between the two SPA-2 polymorphisms and the SPD (C32T) polymorphism, the SPA-1 polymorphism was not in LD with SPD polymorphism investigated.

There have been limited studies investigating SPA and SPD polymorphisms in pulmonary disease. In this study, select polymorphisms in these genes were investigated as potential modifiers of pulmonary disease progression and disease severity in CF. A recent study by Choi and associates investigated SPA gene polymorphisms in a CF cohort and showed significantly worse %predFEV_1 and S-K clinical scores in CF patients with the SPA-1 $6A^3$ allele and SPA-2 $1A^1$ allele ($6A^3/1A^1$)(104). The SPA-1, SPA-2 and SPD proteins have antimicrobial functions and are synthesized, predominantly by alveolar type II cells and Clara cells within the lung and are secreted into the airway surface lining fluid (16). SPA is also required for the structure and / or stability of surfactant aggregates. SPD is currently thought to have a predominantly antimicrobial role and has been shown to bind to glycoconjugates expressed by Gram negative bacteria which include known CF colonizers *P. aeruginosa* and *Hemophilus influenza* (54, 159) and has also been shown to bind to lipopolysacharide from *P. aeruginosa* (31). SPA has been shown to act as an opsonin binding directly to *Staphylococcus aureus* (48) and *Hemophilus influenza* type A (46) and also shown to enhance the uptake of non-opsonized *Staphylococcus aureus* and *P. aeruginosa* by alveolar macrophages (48). Leth-Larsen (57) showed that serum levels of SPD were affected by the SPD (Thr11Met; C32T) polymorphism investigated in this study; the levels of the SPD protein were highest in subjects with a TT genotype and the TT genotype showed two structurally different forms of SPD in the serum; a high and a low molecular weight protein compared to subjects with a CC genotype who showed predominately the low molecular weight form. These different molecular weight SPD forms

were shown to have different binding characteristics to pathogens; the high molecular weight form was critical in the binding, aggregation and clearing of pathogen (Gram positive and negative bacteria), whereas the low molecular weight form bound to simpler ligands like LPS (57). Wu and co-workers (160) showed SPA and SPD proteins also function to increase microbial cell membrane permeability likely in a similar manner to lysozyme and defensins (i.e. involving ionic interactions between the positively charged residues of the host defense protein and the negative charges of the microbial cell membrane). Further work is needed to investigate the polymorphisms in the SPA-1, SPA-2 and SPD genes and whether the polymorphisms affect the proteins' structure and function, but based on our current knowledge of their function and the CF phenotype the pulmonary surfactant genes are good candidates to further investigate as potential modifier genes in CF.

We showed associations for the SPA-1 gene polymorphism (Arg219Trp) and pulmonary disease severity, susceptibility to *P. aeruginosa* infection and pulmonary disease progression. Having a CT genotype was associated with a steeper decline in pulmonary function compared to the decline exhibited by the patients with a CC genotype. The SPA protein has been shown to bind to this pathogen (48). Data from mouse knock-out studies of the SPA and SPD proteins have shown deficient uptake of pathogens by alveolar macrophages, in addition to increased inflammation and inflammatory cell recruitment in the lung following pathogen infection (67, 110-112). Haataja and co-workers (116) identified a SPA-1-SPA-2 haplotype ($6A^2$, $1A^2$) that was over-represented in RDS infants. Karinch and associates (161) showed low levels of SPA mRNA associated with the $1A^0$:$6A^2$ haplotype. Low SPA levels have been shown in the lungs of infants who have died from RDS (117, 162). Least investigated is the SPD gene, where three coding polymorphisms (Thr11Met, Thr160Ala and Thr270Ser) have been identified (52, 57). Lahti and coworkers showed a positive association between severe respiratory syncytial virus infection and homozygosity for the T allele for the SPD (Thr11Met; C32T) gene polymorphism (70). Leth-Larsen and associates showed that the polymorphism at position 11 (Thr11Met; C32T) influenced the structure, function and concentration of the protein in the serum (57). The SPA-1, SPA-2, and SPD gene polymorphisms were investigated as potential candidate modifier genes for COPD by Guo and associates (72) and they showed one positive association with the silent SPA-1 (non amino acid changing) polymorphism at position 62

and increased risk for COPD. Their study investigated each polymorphism independently as a potential disease modifier and the power of this study is questionable.

Our study adds to the accumulating evidence for the importance of these genes in pulmonary disease pathogenesis. These genes may potentially play a role in preventing pathogen colonization or extending the interval to chronic colonization in CF airways, in addition to mounting an inflammatory response. Our data suggest that the T-allele in SPA-1 (Arg219Trp), or other polymorphisms in the SPA-1 or SPA-2 genes with which it is in strong LD, is associated with differential rate of pulmonary disease progression in CF as well as risk of infection and colonization with common CF pathogens. While we did not show differences in pulmonary disease progression, and age of colonization with *P. aeruginosa*, the TT genotype for the SPD (Thr11Met) gene polymorphism was associated with a higher mean frequency of pulmonary infections compared to the CT genotype. However for SPD, we did not show differences in the mean frequency of pulmonary infections over the two time intervals for CF patients who were CC versus TT. While variants in these three genes have been described, the exact effects these variant alleles have on gene transcription, translation and protein function are unclear. There is added difficulty when investigating SPA as there are two proteins encoded by two genes (i.e., SPA-1 and SPA-2) with several allelic polymorphisms, which are not differentiated in functional studies.

.

Preliminary results from our study show a difference in pulmonary disease progression for at least one of the polymorphisms investigated and therefore suggest that the SPA proteins produced from the variant alleles somehow affect the protein structure, function, or both. We also showed differences in the frequency of pulmonary infections over time, which were not necessarily in concordance with results regarding, rate of decline in pulmonary function for the SPA-2 polymorphisms. For the SPD polymorphism, we also showed differences in the frequency of pulmonary infections over time, but no differences in pulmonary disease progression or survival.

There is much research still needed to elucidate the functions of the pulmonary surfactant proteins and the potential effects of the polymorphisms. It is unclear whether homozygosity or heterozygosity for the polymorphisms in the case of the pulmonary surfactant genes A

and D affect protein levels, or result in low levels of the proteins, or defective protein that is degraded and results in the complete absence of these proteins from the lungs. The variant alleles in exon 1 of the MBL2 gene result in defective MBL2 protein. MBL2 deficiency has been shown to increase susceptibility to infection in immuno-compromised patients, but may also have a protective effect. This may also be the case with the polymorphisms in the surfactant genes. Heterozygosity for the deficiency alleles for the MBL2 gene results in low levels of the protein that is insufficient for the acute phase response to pathogen invasion. Defective MBL2 protein caused by the exon 1 polymorphisms may have a protective effect and deficiency is argued to be advantageous as it prevents invasion of pathogens that may use this avenue as a vehicle for intracellular entry. While for healthy individuals this deficiency may not pose a problem, in the case of an immuno-compromised patient this may significantly compromise the normal innate immunity response to pathogen infection and colonization.

Our results for the polymorphisms investigated for the pulmonary surfactant genes are more difficult to interpret. While there are many polymorphisms, documented in the three pulmonary surfactant genes, we only investigated a fraction of these. We showed a steeper decline in pulmonary function over time for the CT genotype for the SPA-1 polymorphism. We also showed a higher frequency of pulmonary infections requiring IV therapy in CF patients who showed a CT genotype. We hypothesize that the latter finding implies an inadequate innate immune response by CF patients with a CT genotype to clear the pathogen effectively resulting in more frequent pulmonary infections requiring IV therapy to maintain a stable clinical status. The steeper decline in %predFEV_1 over time in CF patients with a CT genotype shows that CF patients with this genotype experience a heightened immune response to the pathogen, but the inability to effectively clear the secretions or reduce the pathogen burden in CF contributes to a more exuberant chronic inflammatory response and consequent lung damage

Until the function of the altered pulmonary surfactant proteins is elucidated, it will remain unclear whether these polymorphisms and the translated proteins produced contribute to the heterogeneity in CF pulmonary disease and are protective or accelerate pulmonary damage. We showed an association with pulmonary disease progression and BCC infection in MBL2 wild-type CF patients. In our study, wild-type MBL2 genotype was associated with more

rapid decline in %predFEV$_1$ in CF patients once infected chronically with BCC. The function of innate immunity proteins in chronic diseases such as CF and their effect or contribution to disease progression may be different than is predicted or observed in the otherwise healthy population during an acute infection. Normal response to inflammation and infection in the case of chronic disease may contribute to tissue damage and permanent deterioration of clinical status. The potential role of SPA gene polymorphisms in pulmonary disease progression will require the elucidation of the effects of the variant alleles on protein structure and function.

4.3.2 BCC infection in CF

In this study, we confirmed that chronic infection with BCC has an adverse effect on pulmonary disease progression, with a steeper decline in %predFEV$_1$ shown in the post-acquisition interval. Although we showed comparable rates of decline in %predFEV$_1$ for the different BCC genomovar groups, chronic infection with GEN IIIA was associated with increased mortality compared to GEN II. We did not show survival differences among GEN IIIA infected patients based on RAPD-type. Co-infection with *P. aeruginosa* was associated with more rapid pulmonary disease progression than BCC or *P. aeruginosa* infection alone. Chronic infection with BCC was associated with worse survival than transient BCC infection or *P. aeruginosa* infection alone.

Differences in chronic pathogen infection in the lungs are considered one of the important explanations for the heterogeneous clinical course in CF. *P. aeruginosa* is the most common respiratory pathogen infecting adult CF patients and co-infection with BCC has been shown previously to result in a deterioration in clinical status (123, 128, 129). Previous studies, however, were not based on investigating longitudinal data. In our study, we controlled for clinical status, co-infection with *P. aeruginosa* and identified the time of chronic infection with BCC, and then compared pulmonary disease progression pre and post-acquisition of BCC. In our study, we showed that infection with *P. aeruginosa* was associated with a decline in %predFEV$_1$ of 2% per year compared to 0.5% per year if CF patients were not infected with this pathogen. Co-infection with *P. aeruginosa* and BCC was characterized by more rapid pulmonary disease progression (4% per year). CF patients who were infected solely with BCC showed a similar yearly rate of decline in %predFEV$_1$ (i.e., 2% per year) as CF patients harboring only *P. aeruginosa*. Additionally co-infection with both pathogens

was associated with worse survival compared with CF patients who were infected with neither *P. aeruginosa* nor BCC. Of interest was our finding that even transient growth of BCC was associated with significantly reduced survival. We interpret our results to suggest that susceptibility to BCC infection is a potential marker of disease severity, as it seems that CF patients who are already characterized with moderate or severe pulmonary disease severity are more likely to become infected with the pathogen. Once infected chronically with BCC this pathogen may augment pulmonary disease progression.

A number of studies have reported worse outcomes for CF patients who acquire BCC. Chaparro and associates (130) reviewed the Toronto experience with lung transplantation in CF patients chronically infected with GEN IIIA, RAPD-type 2 and showed reduced 1, 2 and 3 year survival compared to CF patients who received a lung transplant and were chronically infected with *P. aeruginosa*. De Boeck and collegues (131) reported on the Belgian experience of poor 5-year survival in CF patients infected with either GEN IIIA or GEN II. Lewin and associates previously showed in their retrospective study a higher mortality in the first year post-acquisition in 124 CF patients infected with BCC compared with a similar number of patients colonized with *P. aeruginosa* but, of interest, not in the second year following acquisition of BCC (124). Taylor and colleagues (127) showed a more rapid deterioration and increased mortality associated with CF patients with advanced lung dysfunction (%predFEV_1<40%) at the time of BCC acquisition, while those patients infected with BCC who had mild and moderate disease at the time of acquisition maintained a stable clinical status in that first year of monitoring. While the interpretation of this latter study is limited by the absence of a control group, their data are in agreement with our results. As we show, the level of pulmonary disease at the time of acquisition of BCC as well as co-infection with other pathogens such as *P. aeruginosa* could explain some variability in clinical course and survival. In our study, there was only one death in our chronically infected BCC cohort within the first twelve months post-colonization.

Application of molecular techniques has established the complexity of BCC, with a number of distinct genomovars described. Limited information is available to characterize any epidemiological BCC genomovar differences. The possibility of increased virulence of specific genomovars has been extrapolated from the clinic observations of spread of common strains amongst patients with environmental contacts. Furthermore, a very

discrepant clinical course is described for CF patients with BCC, ranging from relentless progressive infection ('cepacia syndrome') to a benign outcome. By collaborating with centers across Canada we were able to gain adequate representation of CF patients infected with the most common BCC genomovars infecting this population; genomovar II and IIIA. We also examined if there were differences in virulence and clinical course between BCC GEN IIIA species RAPD-type 1, 2, 4, and 6. In our study, we showed higher mortality for GEN IIIA patients compared to GEN II. Sufficient representation of other genomovars was not achieved to enable a broader survival comparison. In our cohort, 65% of CF patients chronically infected with GEN IIIA had died or received a lung transplant compared to only 12% of CF patients not infected with BCC. These findings suggest an increased pathogenicity for this genomovar in CF patients and are consistent with the observations of Aris and co-workers (132). These investigators showed higher mortality in post-lung transplant CF patients who were chronically infected with GEN III compared to CF patients in their study who were chronically infected with either GEN II, *P. aeruginosa* or other CF respiratory pathogens (132). CF patients who were infected with GEN IIIA strains in their study and the sputum samples collected from the patients were all negative for the cable pilin gene (132), and therefore were not RAPD group 2. Infection with genomovar II is postulated to be associated with a more benign clinical course, based mostly on clinical experience, but has not been directly studied. There are several caveats from our work that question this purported survival disadvantage for patients infected with GEN IIIA. First is our finding that the decline in %predFEV$_1$ post-acquisition of BCC was similar in CF patients infected with GEN IIIA and GEN II. Spirometric measurements are a reasonable predictor of outcome in CF, with most mortality attributed to progressive lung disease. Nonetheless, the estimated rate of decline in %predFEV$_1$ may not discriminate key mortality factors and perhaps other parameters such as measures of systemic and local inflammation, not measured in this study, would have been more predictive. Arguably most important was our observation that CF patients infected with genomovar IIIA had more severe pulmonary disease at the time of acquisition of BCC. We therefore acknowledge the possibility that worsening survival may merely reflect more advanced disease at the time of infection with GEN IIIA rather than this organism being the driver of increased mortality. This question can be answered by the further study of BCC genomovar cohorts matched for pulmonary disease severity at the time of BCC acquisition.

In our study, we were unable to investigate GEN IIIB and GEN I, IV and V due to small sample sizes. In our cohort, we had CF patients who became infected with GEN IIIB and grew the pathogen for a short term, suggesting that this pathogen has different characteristics from GEN IIIA and does not seem to be a likely chronic colonizer of CF patients. We do however concede that the number of patients with GEN IIIB was small (N=4) and this organism has been shown to chronically colonize CF patients in the United States (163) (the GEN III strain designated in this reference article as PHDC is classified as GEN IIIB (personal communication from D. Henry)). Further investigation on a large cohort of patients infected with GEN IIIB is necessary. There were 3 and 5 CF patients who were infected transiently or chronically with GEN V, respectively. As there were no deaths or lung transplants in these genomovar groups (i.e., GEN, I, IIIB, IV and V) we were unable to investigate them in our survival models.

Infection with BCC has a number of implications for patient management in CF. Transmission of infection, at least of specific genomovar and RAPD groups, has been clearly shown (133-136, 164). With the overall increase in mortality, clinics have instituted rigorous cohorting of patients with BCC infection. The possibility of variable pathogenicity of different BCC genomovars provides support for an extension of this cohorting separating all BCC infected CF patients; however, this needs further investigation. These recommendations must be balanced against the social costs of further isolating CF patients and the further burden that this would place on our clinics and hospitals.

4.4 CONCLUSIONS

In this study, we have investigated polymorphisms in the pulmonary surfactant genes SPA-1, SPA-2, SPD and MBL2 that have been shown to have antimicrobial functions and are involved in innate immunity as possible candidate modifier genes in Cystic Fibrosis. MBL2 deficiency appears to be a predictor of pulmonary disease progression following chronic infection with BCC in CF patients. Contrary to the previous literature, wild-type MBL2 genotype was associated with more severe decline in %predFEV$_1$ over time following colonization with BCC, but was not associated with increased susceptibility to BCC infection. Further studies incorporating greater representation of BCC genomovars, larger numbers with characterization of MBL2 genotypes and longer prospective follow-up are required to expand upon our findings.

We showed a varied response for the pulmonary surfactant gene polymorphisms depending on the response variable investigated. Our results show that in CF, these genes and their polymorphisms impact on the effect that CFTR dysfunction has on the respiratory system. The interaction between the CFTR mutations (mild and severe) and variant alleles in the innate immune genes have a complex effect on the clinical course of CF patients. The measure of disease severity used to study modifier genes in CF will also potentially influence the significance of such associations.

We showed an augmented decline in clinical course with acquisition of BCC which was further augmented by co-infection with *P. aeruginosa*. Our survival results showed similar mortality with infection with either BCC or *P. aeruginosa*. We showed, however, higher mortality in the BCC infected group who was co-infected with *P. aeruginosa* compared to non-infected controls or controls infected only with *P. aeruginosa*. CF patients infected with both pathogens did not show a significant difference in risk of death or lung transplant compared to infection with BCC only. Although mortality was higher in CF patients infected with GEN IIIA, the estimated decline in %predFEV$_1$ over time was similar among CF patients chronically infected with BCC genomovar II and IIIA. These findings provide a clear rationale for segregating BCC infected CF patients irrespective of BCC genotype.

Table 4.1. The chromosomal organization of the MBL2 and the pulmonary surfactant genes (SPA-1, SPA-2, and SPD) on chromosome 10.

Gene	Location (bp), in the chromosome 10 reference sequence (NC 000010.9)
MBL2	54,195,146 – 54, 201,466
SPA-2	81,305,608 – 81,310,101
SPA-1	81,360,664 – 81,363,921
SPD	81,687,476 – 81,698,841

Table 4.2. Summary of pathogens that MBL2, SPA and SPD proteins have been shown to bind to.

Pathogen	MBL2	SPA	SPD
Staphylococcus aureus	(30)	(8, 47, 48)	
Haemophilus influenza type A	(30, 56, 165)	(46, 56, 166)	(55, 56)
Haemophilus influenza type B	(30)		
P. aeruginosa	(6)	(31, 48, 167)	(31, 54)
Burkholderia cepacia complex	(6)		
Listeria monocytogene	(165)		
Neisseria meningitides, Neisseria cinera, Neisseria suflava	(165, 168)		
Mycoplasma pulmonis		(10)	
Mycobacterium tuberculosis		(51, 169)	
Escherichia coli		(9, 48)	(9)
Candida albicans	(30, 170)		
Cryptococcus neoformans	(171)		
Herpes simplex virus type 1		(50, 172)	
HIV-1 and HIV-2	(173)		
Aspergillus fumigatus	(30)	(174, 175)	(54, 174-176)
Klebsiella pneumoniae	(30)		(54)
Pneumocystis carinii		(177, 178)	(54, 174-176)
Mycobacterium and Mycobacterium leprae	(179)		
Respiratory syncytial virus		(167)	(180)
Pollen grains (pollen from: Lombardy poplar (*Populus nigra italica*), Kentucky blue grass (*Poa pratensis*), cultivated rye (*Secale cerale*), and short ragweed (*Ambrosia elatior*).		(181)	
Whole mite extracts (*Dermatophagoides pteronyssinus*) and allergens (from faecal pellets from house dust mites		(182)	(182)
Escherichia coli	(30)		(54)

Table 4.3. SPA-1-SPA-2 haplotypes are presented as reported in DiAngelo and associates (73). Three SNPs were investigated in the present study and are identified in the table.

	SPA-2					SPA-1					
Alleles	**C—**	**SPA-2**				**SPA-1**				**—T**	Alleles
1A	C	C	C	C		C	C	G	G	C	6A
$1A^0$	C	C	G	A		T	G	A	A	C	$6A^2$
$1A^1$	A	T	G	C		T	C	A	A	C	$6A^3$
$1A^2$	C	C	G	C		T	C	G	A	T	$6A^4$
$1A^3$	A	T	G	A							
$1A^5$	C	T	C	C							
Nucleotide change	**C**/**A**AG	TC **C**/**T**	**C**/**G** CT	A**A**/**C**C		G**C**/**T**G	**C**/**G** TC	CC**G**/**A**	AC **G**/**A**	**C**/**T** GG	
AA number	223	140	91	9		19	50	62	133	219	
AA change	Gln/Lys			Asn/Thr		Ala/Val	Leu/Val			Arg/Trp	
SNP investigated *	+			+						+	

* The SNPs that were investigated in the present study are identified by the symbol +.

Table 4.4. Distribution of genotypes, allele frequencies and Hardy Weinberg equilibrium for MBL2, SPA-1, SPA-2 and SPD gene polymorphisms.

MBL2 PROMOTER	YY	YX	XX	Y	X	P-value
Observed	289(59.7%)	173 (36.0%)	22 (5.0%)	751(78.0%)	217 (22.0%)	0.83
Expected	291(60.0%)	168 (35.0%)	24 (5.0%)			
MBL2-B Gly54Asp)	AA	AB	BB	A	B	
Observed	364(75.0%)	87 (18.0%)	35 (7.0%)	815(84.0%)	157 (16.0%)	0.00001
Expected	342(70.0%)	132 (27.0%)	13 (2.6%)			
MBL2-C (Gly57Glu)	AA	AC	CC	A	C	
Observed	467(96.0%)	19 (4.0%)	0 (0%)	953(98.0%)	19(2.0%)	0.91
Expected	467(96.0%)	19 (4.0%)	0 (0%)			
MBL2-D (Arg52Cys)	AA	AD	DD	A	D	
Observed	420(86.0%)	59 (12.0%)	7 (1.4%)	899(92.5%)	73 (7.5%)	0.02
Expected	416(85.0%)	68 (14.0%)	3 (0.6%)			
SPA-1 (Arg219Trp)	TT	CT	CC	T	C	
Observed	0(0%)	102 (20.6%)	392(79.0%)	102(10.0%)	886 (90.0%)	0.04
Expected	5(1.0%)	91 (18.4%)	397(80.4%)			
SPA-2 (Thr9Asn)	AA	AC	CC	A	C	
Observed	232(47.9%)	185 (38.2%)	67 (13.8%)	649(67.0%)	319 (33.0%)	0.01
Expected	218(45.0%)	214 (44.2%)	53 (10.9%)			
SPA-2 (Lys223Gln)	AA	AC	CC	A	C	
Observed	19 (3.9%)	146 (30.0%)	321(66.0%)	184(19.0%)	788 (81.0%)	0.89
Expected	17 (3.6%)	149 (30.7%)	319(65.6%)			
SPD	TT	CT	CC	T	C	

(Thr11Met)						
Observed	182(37.0%)	233 (48.0%)	71 (14.6%)	597 (61%)	375 (39%)	0.97
Expected	183(37.4%)	230 (47.0%)	72 (15.0%)			

Table 4.5. Distribution of genotypes, allele frequencies and Hardy Weinberg equilibrium for MBL2 gene polymorphisms based on age grouping (i.e., <25 and ≥25 years of age).

<25 yrs of age						
MBL2 PROMOTER	YY	YX	XX	Y	X	P-value
Observed	171 (60.2%)	100(35.2%)	13 (4.6%)	442 (78.0%)	126 (22.0%)	0.94
Expected	172 (60.5%)	98 (34.5%)	14 (4.9%)			
MBL2-B (Gly54Asp)	AA	AB	BB	A	B	
Observed	214 (75.1%)	65 (22.8%)	6 (2.0%)	493 (86.0%)	77 (14.0%)	0.92
Expected	213 (74.8%)	67 (23.4%)	5 (1.8%)			
MBL2-C (Gly57Glu)	AA	AC	CC	A	C	
Observed	275 (96.5%)	10 (3.5%)	0 (0%)	560 (98%)	10 (2.0%)	0.96
Expected	275 (96.5%)	10 (3.5%)	0 (0%)			
MBL2-D (Arg52Cys)	AA	AD	DD	A	D	
Observed	429 (92.3%)	32 (6.9%)	4 (0.9%)	890 (96%)	40 (0.4%)	0.002
Expected	426 (92.6%)	38 (8.2%)	1 (0.2%)			
≥25 yrs of age						
MBL2 PROMOTER	YY	YX	XX	Y	X	P-value
Observed	118 (59.0%)	73 (36.5%)	9 (4.5%)	309 (77.0%)	91 (23.0%)	0.86
Expected	119 (59.7%)	70 (35.1%)	10 (5.2%)			
MBL2-B (Gly54Asp)	AA	AB	BB	A	B	
Observed	151(75.1%)	40 (19.9%)	10 (5.0%)	342 (85.0%)	60 (15.0%)	0.009
Expected	145(72.4%)	51 (25.4%)	4 (2.2%)			
MBL2-C (Gly57Glu)	AA	AC	CC	A	C	
Observed	193(96.2%)	8 (4.0%)	0 (0%)	394 (98%)	8 (2.0%)	0.96

Expected	193 (96.1%)	8 (3.9%)	0 (0%)			
MBL2-D (Arg52Cys)	AA	AD	DD	A	D	
Observed	172 (85.6%)	28 (13.9%)	1 (0.5%)	372 (93%)	30 (0.7%)	0.99
Expected	172 (85.6%)	28 (13.8%)	1 (0.6%)			

Table 4.6. Distribution of genotypes, allele frequencies and Hardy Weinberg equilibrium for SPA-1, SPA-2 and SPD gene polymorphisms based on age grouping (i.e., <25 and ≥25 years of age).

<25 yrs of age						
SPA-1 (Arg219Trp)	TT	CT	CC	T	C	P-value
Observed	0 (0%)	48 (16.8%)	237 (83.2%)	48 (8.0%)	522 (92%)	0.30
Expected	2 (0.7%)	44 (15.4%)	239(83.9%)			
SPA-2 (Thr9Asn)	AA	AC	CC	A	C	
Observed	140 (49.3%)	107 (37.7%)	37(13.0%)	387(68%)	181 (32%)	0.08
Expected	132 (46.4%)	123 (43.4%)	29(10.2%)			
SPA-2 (Lys223Gln)	AA	AC	CC	A	C	
Observed	10 (3.5%)	97 (34.0%)	178 (62.5%)	117(21%)	483 (79%)	0.77
Expected	12 (4.2%)	93 (32.6%)	180 (63.2%)			
SPD (Thr11Met)	TT	CT	CC	T	C	
Observed	108 (37.9%)	131 (46.0%)	46(16.1%)	347(61%)	223 (39%)	0.84
Expected	106 (37.1%)	136 (47.6%)	44(15.3%)			
≥25 yrs of age						
SPA-1 (Arg219Trp)	TT	CT	CC	T	C	
Observed	0 (0%)	46 (22.9%)	155 (77.1%)	46(11.0%)	356 (89%)	0.19
Expected	3 (1.3%)	41 (20.3%)	158 (78.4%)			
SPA-2 (Thr9Asn)	AA	AC	CC	A	C	
Observed	92 (46.0%)	78 (39.0%)	30 (15.0%)	262 (66%)	138 (35%)	0.15
Expected	86 (42.9%)	90 (45.2%)	24 (11.9%)			
SPA-2 (Lys223Gln)	AA	AC	CC	A	C	

Observed	9 (4.5%)	49 (24.4%)	143 (71.1%)	67 (17%)	335 (83%)	0.22
Expected	6 (2.8%)	56 (27.8%)	140 (69.4%)			
SPD (Thr11Met)	TT	CT	CC	T	C	
Observed	74 (36.8%)	102 (50.7%)	25 (12.4%)	250 (62%)	152 (38%)	0.5336
Expected	78 (38.7%)	95 (47.0%)	29 (14.3%)			

Table 4.7. Pairwise linkage disequilibrium results for MBL2 and pulmonary surfactant gene polymorphisms.

Pairwise Linkage Disequilibrium	D' (D/Dmax)	p-value	R^2
MBL2 gene			
MBL2PROM and MBL2-B (Gly54Asp)	0.998	<0.0001	0.05
MBL2PROM and MBL2-C (Gly57Glu)	0.987	0.02	0.006
MBL2PROM and MBL2-D (Arg52Cys)	0.997	<0.0001	0.02
MBL2-B (Gly54Asp) and MBL2-C (Gly57Glu)	0.204	0.69	0.0002
MBL2-B (Gly54Asp) and MBL2-D (Arg52Cys)	0.879	0.0006	0.01
MBL2-C (Gly57Glu) and MBL2-D (Arg52Cys)	0.961	0.23	0.001
MBL2 gene with pulmonary surfactant genes			
SPA-1 (Arg219Trp) and MBL2PROM	0.067	0.71	0.0001
SPA-1 (Arg219Trp) and MBL2-B (Gly54Asp)	0.11	0.61	0.0003
SPA-1 (Arg219Trp) and MBL2-C (Gly57Glu)	0.09	0.23	0.001
SPA-1 (Arg219Trp) and MBL2-D (Arg52Cys)	0.03	0.41	0.0007
SPA-2 (Thr9Asn) and MBL2PROM	0.02	0.87	0.004
SPA-2 (Thr9Asn) and MBL2-B (Gly54Asp)	0.21	0.05	0.0003
SPA-2 (Thr9Asn) and MBL2-C (Gly57Glu)	0.18	0.58	0.002
SPA-2 (Thr9Asn) and MBL2-D (Arg52Cys)	0.01	0.90	0.005
SPA-2 (Lys223Gln) and MBL2PROM	0.02	0.56	0.001
SPA-2 (Lys223Gln) and MBL2-B (Gly54Asp)	0.03	0.43	0.001
SPA-2 (Lys223Gln) and MBL2-C (Gly57Glu)	0.45	0.33	0.001
SPA-2 (Lys223Gln) and MBL2-D (Arg52Cys)	0.27	0.22	0.001
SPD (Thr11Met) and MBL2PROM	0.04	0.57	0.0007
SPD (Thr11Met) and MBL2-B (Gly54Asp)	0.05	0.41	0.007
SPD (Thr11Met) and MBL2-C (Gly57Glu)	0.10	0.57	0.0003
SPD (Thr11Met) and MBL2-D (Arg52Cys)	0.08	0.40	0.0007
Pulmonary surfactant genes			
SPA-1 (Arg219Trp) and SPA-2 (Thr9Asn)	0.83	<0.0001	0.15
SPA-1 (Arg219Trp) and SPA-2 (Lys223Gln)	0.99	<0.0001	0.025
SPA-1 (Arg219Trp) and SPD (Thr11Met)	0.14	0.07	0.0035

SPA-2 (Thr9Asn) and SPA-2 (Lys223Gln)	0.40	<0.0001	0.077
SPA-2 (Thr9Asn) and SPD (Thr11Met)	0.19	<0.0001	0.028
SPA-2 (Lys223Gln)and SPD (Thr11Met)	0.28	<0.0001	0.029

Table 4.8. List of haplotypes for MBL2 gene polymorphisms.

MBL2	N (%)	%
YAAA	530	55
XAAA	215	22
YBAA	133	14
YACA	16	2
YAAD	69	7
YBAD	1	0.1
YBCA	2	0.2

Table 4.9. List of haplotypes for SPA-1, SPA-2 and SPD gene polymorphisms found using PHASE.

SPD(C32T)-SPA-1(C655T)-SPA-2(A667C)-SPA-2(A26C)	N	%
TCAA	436	44.9
TCCT	47	4.8
TAAA	19	2.0
CCAA	138	14.2
CAAA	52	5.3
TCCA	48	4.9
CACA	71	7.3
CCCA	72	7.4
CCCT	39	4.0
TACA	42	4.3
TCAT	5	0.5
CCAT	3	0.3

Table 4.10. List of haplotypes for SPA-1 and SPA-2 gene polymorphisms found using PHASE.

SPA-2:SPA-1 (SPA-2(A26C)–SPA-2(A667C)–SPA-1(C655T))	N	%
CCT	73	7.5
CCC	82	8.4
CAT	75	7.7
CAC	432	44.4
ACT	4	0.4
ACC	15	1.5
AAT	15	1.5
AAC	276	28.4

Table 4.11. List of haplotypes for SPA-2 polymorphisms found using PHASE.

SPA-2 (SPA-2(A26C)-SPA-2(A667C)	N (%)	%
CA	591	61
CC	197	20.3
AA	60	6.2
AC	124	12.8

Table 4.12. Clinical characteristics of study cohort by MBL2 genotype grouping.

	Wild-type	Deficient	p-value**
N	252	200	
Sex (Male/Female) [+]	126/126	104/96	0.37
Age of CF diagnosis (yrs)	4.4(0.5)	4.8(0.6)	0.54
Current age (yrs)	23.5(0.7)	25.0(0.8)	0.13
%predFEV_1 (% of predicted)	63.1(1.8)	63.4(1.9)	0.92
BMI (kg/m^2)	19.8(0.2)	20.6(0.2)	0.009
CFTR genotype*[+]	174/53/17/8	133/49/11/7	0.80
PSS (sufficient/insufficient)	33/219	21/179	0.24
PA infection status (not infected /chronically infected) [+ ++]	103/149	61/138	0.02
PA1 age (yrs) [++]	13.5(0.8) N=146	13.8(1.1) N=120	0.80
PAC age (yrs) [++]	14.6(1.0) N=94	14.7(1.1) N=85	0.92
Current status (alive/deceased or transplanted) [+]	209/43	168/32	0.43

* CFTR genotype: homozygous for severe mutations (class 1,2, or 3) / Heterozygous for severe mutation (class 1,2 or 3) and the other mutation is classified as unknown or unclassified / homozygous or heterozygous for mild mutation class 4, or 5 / homozygous or heterozygous for 2 unknown or unclassified mutations.

** P-value denotes univariate ANOVA or chi square results.

+ Values presented are number of patients.

++ PA denotes *P. aeruginosa* and PA1 and PAC denote first infection and chronic infection with *P. aeruginosa*.

Table 4.13. Clinical characteristics of study cohort by SPA-1 genotype grouping.

	CC	CT	p-value**
N	364	89	
Sex (Male/Female) [+]	182/182	48/41	0.29
Age of CF diagnosis (yrs)	4.2(0.4)	6.2(1.0)	0.04
Current age (yrs)	23.8(0.6)	25.5(1.2)	0.19
%predFEV$_1$ (% of predicted)	62.9(1.5)	65.1(2.6)	0.50
BMI (kg/m^2)	20.0(0.2)	20.6(0.4)	0.14
CFTR genotype*[+]	248/18/85/13	60/10/17/2	0.09
PSS (sufficient/insufficient) [+]	36/328	18/71	0.01
PA infection status (not infected /chronically infected) [+ ++]	131/233	33/55	0.81
PA1 age (yrs) [++]	13.4(0.8) N=215	14.2(1.5) N=52	0.64
PAC age (yrs) [++]	14.5(0.8) N=145	15.0(1.8) N=35	0.80
Current status (alive/deceased or transplanted) [+]	303/61	81/8	0.07

* CFTR genotype: homozygous for severe mutations (class 1,2, or 3) / Heterozygous for severe mutation (class 1,2 or 3) and the other mutation is classified as unknown or unclassified / homozygous or heterozygous for mild mutation class 4, or 5 / homozygous or heterozygous for 2 unknown or unclassified mutations.

** p-value denotes univariate ANOVA or chi square results.

+ Values presented are number of patients.

++ PA denotes *P. aeruginosa* and PA1 and PAC denote first infection and chronic infection with *P. aeruginosa*.

Table 4.14. Clinical characteristics of study cohort by SPA-2 genotype grouping. The grouping is based on whether patients have zero, one, or two copies of the CA haplotype.

	2 copies of CA	1 copy of CA	0 copies of CA	p-value**
N	179	233	74	486
Sex (Male/Female) [+]	86/93	124/109	39/35	0.48
Age of CF diagnosis (yrs)	3.9(0.5)	4.5(0.6)	5.2(1.0)	0.51
Current age (yrs)	22.9(0.8)	23.6(0.8)	25.2(1.2)	0.32
%predFEV$_1$ (% of predicted)	60.3(2.2)	65.4(2.0)	64.3(3.2)	0.21
BMI (kg/m^2)	19.9(0.2)	20.1(0.2)	20.3(0.4)	0.68
CFTR genotype*[+]	120/45/6/8	159/51/16/7	53/13/7/1	0.31
PSS (sufficient/insufficient) [+]	17/162	24/209	14/60	0.08
PA infection status (not infected /chronically infected) [+ ++]	65/112	88/141	27/44	0.94
PA1 age (yrs) [++]	12.4(1.0) N=106	13.5(1.0) N=131	15.1(2.1) N=39	0.39
PAC age (yrs) [++]	14.0(1.0) N=77	15.0(1.2) N=83	13.7(2.1) N=24	0.78
Current status (alive/deceased or transplanted) [+]	148/31	186/37	65/9	0.59

* CFTR genotype: homozygous for severe mutations (class 1,2, or 3) / Heterozygous for severe mutation (class 1,2 or 3) and the other mutation is classified as unknown or unclassified / homozygous or heterozygous for mild mutation class 4, or 5 / homozygous or heterozygous for 2 unknown or unclassified mutations.

** P-value denotes univariate ANOVA or chi square results.

+ Values presented are number of patients.

++ PA denotes *P. aeruginosa* and PA1 and PAC denote first infection and chronic infection with *P. aeruginosa*.

Table 4.15. Clinical characteristics of study cohort by SPD genotype grouping.

	TT	CT	CC	p-value**
N	174	217	62	
Sex (Male/Female) [+]	86/88	109/108	35/27	0.62
Age of CF diagnosis (yrs)	4.3(0.5)	4.6(0.5)	5.0(1.3)	0.84
Current age (yrs)	23.8(0.8)	24.3(0.7)	24.7(1.5)	0.80
%predFEV$_1$ (% of predicted)	62.2(2.2)	64.9(1.9)	61.1(3.5)	0.52
BMI (kg/m^2)	19.7(0.2)	20.4(0.2)	20.5(0.4)	0.10
CFTR genotype*[+]	118/44/4/8	139/20/52/6	51/4/6/1	0.01
PSS (sufficient/insufficient) [+]	15/159	32/185	7/55	0.18
PA infection status (not infected /chronically infected) [+ ++]	64/110	78/138	22/40	0.98
PA1 age (yrs) [++]	12.7(0.9) N=105	13.7(1.0) N=122	15.7(2.3) N=40	0.35
PAC age (yrs) [++]	14.2(1.1) N=71	14.9(1.1) N=84	15.0(2.3) N=25	0.88
Current status (alive/deceased or transplanted) [+]	144/30	183/34	52/11	0.88

* CFTR genotype: homozygous for severe mutations (class 1,2, or 3) / Heterozygous for severe mutation (class 1,2 or 3) and the other mutation is classified as unknown or unclassified / homozygous or heterozygous for mild mutation class 4, or 5 / homozygous or heterozygous for 2 unknown or unclassified mutations.

** P-value denotes univariate ANOVA or chi square results.

+ Values presented are number of patients.

++ PA denotes *P. aeruginosa* and PA1 and PAC denote first infection and chronic infection with *P. aeruginosa*.

Table 4.16. Mixed effects models for pulmonary disease progression considering MBL2 deficiency and BCC infection status.

MBL2	Mixed model equation
Pre-acquisition interval	
Model 1	Difference in pre-acquisition slope for MBL2 deficient and wild-type.
MBL2 deficient MBL2 wild-type Comparison of MBL2 deficient and wild-type**	Rate of change* =-1.02 SE=0.13 CI: (-1.26, -0.77) Rate of change =-2.54 SE=0.59 CI: (-5.75, -3.24) p-value=0.63
Model 2	Difference in slopes in MBL2 deficient and wild type for BCC infected.
Chronic BCC and MBL2 deficient Chronic BCC and MBL2 wild-type Pre-Post change in %predFEV$_1$ comparisons: MBL2 deficient Vs wild-type Controls$^+$: MBL2 deficient Vs wild-type Chronic BCC: MBL2 deficient Vs wild-type MBL2 deficient: Chronic BCC Vs Controls MBL2 wild-type: Chronic BCC Vs Controls	Rate of change =-2.54 SE=0.59 CI: (-5.75, -3.24) Rate of change =-4.50 SE=0.64 CI: (-3.69, -1.39) p-value=0.02 p-value=0.45 p-value=0.02 p-value=0.01 p-value <0.0001

* Rate of change=change in %predFEV$_1$/year.

** Includes BCC infected (pre-acquisition interval only) and controls.

$^+$ Control group includes CF patients who were transiently infected with BCC. The slopes for controls are presented in model 1.

Table 4.17. Survival analysis of the time to event (death or lung transplantation) for CF patients.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MBL2 (N=476)				
MBL2 (Deficient=0/Wild-type=1)	1.42	0.64	0.33	6.02
Sex (Male=0/Female=1)	1.47	0.10	0.93	2.32
CFTR class (mild=1)	0.27	0.21	0.03	2.11
CFTR class (1,2,3/unknown=1)	0.64	0.15	0.35	1.18
CFTR class (unknown/unknown=1)	0.32	0.27	0.04	2.36
PA infection status (chronically infected=1)	0.84	0.05	0.99	11.32
MBL2 * PA status	0.84	0.83	0.18	3.86
SPA-1 Model (N=477)				
SPA-1 (CC=0/CT=1)	0.82	0.81	0.16	4.15
Sex (Male=0/Female=1)	1.36	0.20	0.85	2.15
CFTR class (mild=1)	0.27	0.21	0.03	2.12
CFTR class (1,2,3/unknown=1)	0.61	0.12	0.33	1.13
CFTR class (unknown/unknown=1)	0.29	0.22	0.04	2.10
PA infection status (chronically infected=1)	1.23	0.80	0.24	6.32
SPA-1* PA status	1.05	0.26	0.47	17.35
SPA-2 Model (N=477)				
SPA-2 (1 copies of CA)	1.87	0.46	0.36	9.78
SPA-2 (0 copies of CA)	0.65	0.73	0.59	7.26
Sex (Male=0/Female=1)	1.38	0.18	0.86	2.21
CFTR class (mild=1)	0.25	0.19	0.03	1.99
CFTR class (1,2,3/unknown=1)	0.61	0.11	0.33	1.12
CFTR class (unknown/unknown=1)	0.29	0.05	0.04	2.13
PA infection status (chronically infected=1)	4.29	0.05	1.00	18.31
SPA-2 (1 copies of CA)* PA status	0.43	0.34	0.08	2.45
SPA-2 (0 copies of CA)* PA status	1.28	0.85	0.10	16.22
SPD Model (N=477)				

SPD (CC genotype)	0.54	0.48	0.10	2.98
SPD (CT genotype)	0.24	0.11	0.04	1.38
Sex (Male=0/Female=1)	1.48	0.09	0.94	2.35
CFTR class (mild=1)	0.38	0.36	0.05	3.00
CFTR class (1,2,3/unknown=1)	0.65	0.17	0.35	1.20
CFTR class (unknown/unknown=1)	0.32	0.26	0.04	2.32
PA infection status (chronically infected=1)	1.43	0.51	0.49	4.14
SPD (CC genotype)* PA status	2.22	0.41	0.34	14.58
SPD (CT genotype)* PA status	4.84	0.09	0.79	29.58

Variables used in the Survival models:

Dependent variables: Event=death or lung transplanted coded as 1 ((0=alive), current age or age of event.

Main effects: Gene, CFTR class, PA infection status,.

Interactions: Gene * PA infection status.

Table 4.18. Effect of innate immunity gene polymorphisms on age of first *P. aeruginosa* infection.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MBL2 (N=179)				
MBL2 (deficient=0/Wild-type=1)	0.89	0.44	0.66	1.20
Sex (Male=0/Female=1)	1.21	0.19	0.91	1.63
CFTR class (mild=1)	0.06	0.006	0.01	0.45
CFTR class (1,2,3/unknown=1)	0.64	0.01	0.45	0.91
CFTR class (unknown/unknown=1)	0.62	0.26	0.27	1.42
SPA-1 (N=267)				
SPA-1 (CC=0/CT=1)	1.01	0.96	0.71	1.44
Sex (Male=0/Female=1)	1.24	0.15	0.93	1.66
CFTR class (mild=1)	0.06	0.006	0.01	0.45
CFTR class (1,2,3/unknown=1)	0.64	0.01	0.45	0.92
CFTR class (unknown/unknown=1)	0.64	0.28	0.28	1.45
SPA-2 (N=276)				
SPA-2 (1=1 copies of CA), base group is 2CA	0.74	0.20	0.47	1.17
SPA-2 (1=0 copies of CA)	0.89	0.48	0.65	1.22
Sex (Male=0/Female=1)	1.22	0.19	0.91	1.63
CFTR class (mild=1)	0.07	0.006	0.01	0.46
CFTR class (1,2,3/unknown=1)	0.65	0.02	0.45	0.91
CFTR class (unknown/unknown=1)	0.66	0.32	0.29	1.51
SPD (N=267)				
SPD (1=CC), case group is TT	0.74	0.18	0.47	1.15
SPD (1=CT)	1.05	0.74	0.77	1.44
Sex (Male=0/Female=1)	1.26	0.12	0.94	1.68
CFTR class (mild=1)	0.06	0.005	0.01	0.43
CFTR class (1,2,3/unknown=1)	0.62	0.008	0.43	0.88
CFTR class (unknown/unknown=1)	0.61	0.24	0.27	1.39

* Formula:

Dependent variables: Age of 1st infection (age of chronic infection), PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class

Interactions: Modifier gene* CFTR class

+ CFTR class: The comparison is CFTR class noted in table versus base group of homozygous severe CFTR class 1, 2, or 3.

Table 4.19. Effect of innate immunity gene polymorphisms on age of chronic *P. aeruginosa* infection.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MBL2 (N=183)				
MBL2 (CC=0/CT=1)	0.94	0.66	0.69	1.26
Sex (Male=0/Female=1)	1.19	0.24	0.89	1.61
CFTR class (mild=1)	0.47	0.45	0.07	3.37
CFTR class (1,2,3/unknown=1)	0.75	0.11	0.53	1.06
CFTR class (unknown/unknown=1)	0.57	0.18	0.25	1.30
SPA-1 (N=184)				
SPA-1 (CC=0/CT=1)	0.98	0.91	0.68	1.40
Sex (Male=0/Female=1)	1.20	0.23	0.89	1.61
CFTR class (mild=1)	0.49	0.48	0.07	3.49
CFTR class (1,2,3/unknown=1)	0.75	0.11	0.53	1.07
CFTR class (unknown/unknown=1)	0.58	0.19	0.25	1.32
SPA-2 (N=184)				
SPA-2 (1=1 copies of CA), base group is 2CA	0.96	0.81	0.69	1.33
SPA-2 (1=0 copies of CA)	1.20	0.45	0.75	1.93
Sex (Male=0/Female=1)	1.19	0.25	0.88	1.62
CFTR class (mild=1)	0.51	0.50	0.07	3.62
CFTR class (1,2,3/unknown=1)	0.74	0.10	0.52	1.06
CFTR class (unknown/unknown=1)	0.53	0.14	0.23	1.24
SPD (N=184)				
SPD (1=CC), case group is TT	0.87	0.56	0.55	1.39
SPD (1=CT)	0.96	0.80	0.70	1.32
Sex (Male=0/Female=1)	1.19	0.26	0.88	1.60
CFTR class (mild=1)	0.49	0.47	0.07	3.48
CFTR class (1,2,3/unknown=1)	0.74	0.10	0.52	1.06
CFTR class (unknown/unknown=1)	0.56	0.17	0.24	1.29

* Formula

Dependent variables: Age of chronic infection with PA, PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class

Interactions: Modifier gene* CFTR class.

Table 4.20. Clinical characteristics of study subjects used for investigating BCC genomovar and pulmonary disease progression (chronically and transiently infected with BCC and the control group). Values are shown as mean (SEM) and the p-values are from univariate analysis of variance. The p-values from chi-square analysis of categorical variables are also presented.

	BCC infected (sporadic)-1	BCC infected (chronic)-2	Control -3	P-value
Sex (M/F)	6/14	36/34	105/88	0.11
Age (years)	16.3(1.5)	24.7(1.0)	22.2(0.8)	0.005* 2>3>1; 3>1
Age of CF diagnosis (years)	1.5(0.5)	2.4(0.8)	5.4(0.7)	0.01 1>2>3
%predFEV$_1$	67.9(6.9)	42.7(3.0)	69.5(2.1)	0.005 3>2>1; 2>1
BMI (kg/m^2)	18.3(0.7)	19.2(0.3)	19.9(0.2)	0.03
PSS (insufficient/sufficient)	0/20	65/5	165/28	0.06
# of *P. aeruginosa* positive/not colonized	8/12	18/52	84/109	0.03
Age of 1st infection with *P. aeruginosa*	7.9(1.2) N=20	11.4(1.2) N=47	12.6(0.8) N=149	0.09
Age of chronic *P. aeruginosa* infection	9.6(1.3) N=11	11.6(1.3) N=35	13.9(0.9) N=102	0.16
Dead or lung transplanted / alive	14/6	35/35	167/26	0.0001
Age of BCC infection (yrs)	11.6(1.3)	18.6(1.0)		0.001
%predFEV$_1$ at BCC acquisition	76.2(5.5)	62.4(3.0)		0.04

* P-value and significant post-hoc comparisons are presented. 1, 2, 3 represent transient BCC infection, chronic BCC infection and control group, respectively.

Table 4.21. Linear mixed effects models for BCC infection and pulmonary disease progression.

Model 1	Difference in post-acquisition slope in chronic BCC Vs transient BCC Vs non-infected controls
Controls	Rate of change* =-1.44 SE=0.15 CI: (-1.74,-1.13)
Transient BCC infection	Rate of change =-2.77 SE=0.86 CI: (-4.46,-1.08)
Chronic BCC infection	Rate of change =-3.52 SE=0.41 CI: (-4.33,-2.72)
<u>Comparisons:</u>	
Chronic Vs controls	p-value=0.04
Transient Vs chronic BCC	p-value=0.43
Transient Vs controls	p-value=0.13
Model 2	**Difference in post-acquisition slope in chronic BCC GEN IIIA Vs chronic BCC GEN II versus non-infected controls**
Controls	
Chronic BCC GEN II	Rate of change =-1.42 SE=0.15 CI: (-1.72,-1.11)
Chronic BCC GEN IIIA (RAPD-type 1,4,6)	Rate of change=-3.56 SE=0.91 CI: (-5.34,-1.78)
Chronic BCC GEN IIIA (RAPD-type 2)	Rate of change =-4.19 SE=0.71 CI: (-5.57,-2.80)
Chronic BCC GEN I/IV/V	Rate of change =-2.80 SE=0.70 CI: (-4.17,-1.43)
<u>Comparisons to non-infected controls:</u>	
Controls Vs GEN II	p-value =0.02
Controls Vs GEN IIIA (RAPD-type 1,4,6)	p-value=0.0001
Controls Vs GEN IIIA (RAPD-type 2)	p-value=0.05
<u>Comparison of GEN IIIA RAPD-type</u>	p-value=0.16
<u>RAPD-type 1,4,6 Vs RAPD-type 2</u>	
<u>Comparison of GEN IIIA to GEN II:</u>	
GEN IIIA (RAPD-type 1,4,6) Vs GEN II	p-value=0.59
GEN IIIA (RAPD-type 2) Vs GEN II	p-value=0.51
Collapsed GEN IIIA	
Chronic BCC GEN IIIA(RAPD-type	Rate of change =-3.51 SE=0.50 CI: (-4.48,-2.53)

1,2,4,6) Comparisons: GEN IIIA Vs controls GEN IIIA Vs BCC GEN II	Slopes for GEN II and controls are as above p-value=0.00004 p-value=0.96
Model 3: Concurrent *P. aeruginosa* infection	Difference in post-acquisition slope in chronic BCC Vs non-infected controls based on whether also infected with *P. aeruginosa* (PA)
Controls no PA Controls with PA Chronic BCC no PA Chronic BCC with PA Comparisons: Chronic BCC no PA Vs Controls with PA Chronic BCC with PA Vs Controls with PA	Rate of change =-0.46 SE=0.24 CI: (-0.93, 0.01) Rate of change =-2.01 SE=0.19 CI: (-2.37, -1.64) Rate of change =-2.20 SE=0.66 CI: (-3.50,-0.91) Rate of change =-4.28 SE=0.51 CI: (-5.28,-3.28) p-value=0.02 p-value <0.0001

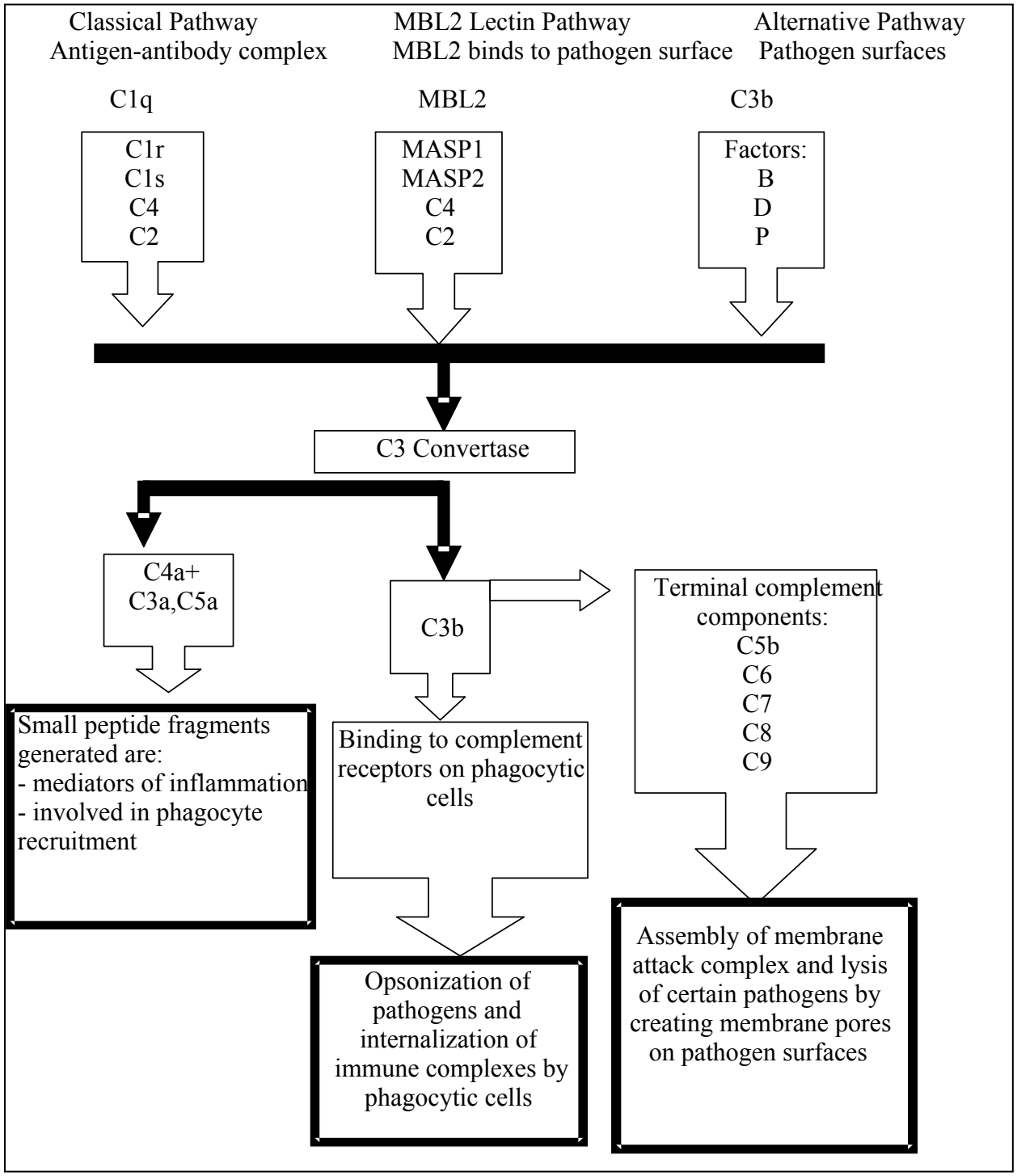
* Rate of change=change in %predFEV_1/year

Table 4.22. Clinical characteristics of study subjects used for investigating BCC genomovar and MBL2 deficiency and pulmonary disease progression. Values are shown as mean (SEM) and the p-values from univariate analysis of variance. The p-values from chi-square analysis of categorical variables are also presented.

	Total BCC infected used for MBL2	Total control used for MBL2	P-value
Sex (M/F)	20/19	219/208	0.50
Age (years)	26.0(1.5)	24.0(0.5)	0.28
Age of CF diagnosis (years)	3.2(1.3)	5.0(0.4)	0.21
%predFEV_1	47.7(4.2)	65.1(1.4)	0.001
BMI (kg/m^2)	20.1(0.5)	20.2(0.2)	0.90
PSS (insufficient/sufficient)	3/35	54/373	0.29
# of *P. aeruginosa* positive/not colonized*	12/26	159/267	0.30
Age of 1st infection with *P. aeruginosa**	13.8(2.0)	13.6(0.7)	0.94
Age of chronic *P. aeruginosa* infection*	12.7(1.7)	14.7(0.8)	0.38
Dead or lung transplanted / alive	25/13	371/56	0.001
Age of BCC infection (yrs)	18.8(1.5)		
%predFEV_1 at BCC acquisition	68.1(4.3)		

*For longitudinal analysis of MBL2 and pulmonary disease progression the total sample sizes available for calculation of age of 1st infection and chronic infection with *P. aeruginosa* were 30/246 and 22/162 (chronic/control), respectively.

Figure 4.1. Schematic overview of the three pathways that activate complement. The MBL2 protein activates complement through the lectin pathway. The activated plasma complement proteins can directly activate complement cascade through the alternative pathway. Their effect is immediate compared to the classical pathway, which requires 5-7 days for antibody production. Diagram adapted from Janeway and Travers (183).


Classical Pathway
Antigen-antibody complex
MBL2 Lectin Pathway
MBL2 binds to pathogen surface
Alternative Pathway
Pathogen surfaces
C1q
MBL2
C3b
C1r
C1s
C4
C2
MASP1
MASP2
C4
C2
Factors:
B
D
P
C3 Convertase
C4a+
C3a,C5a
C3b
Terminal complement components:
C5b
C6
C7
C8
C9
Small peptide fragments generated are:
- mediators of inflammation
- involved in phagocyte recruitment
Binding to complement receptors on phagocytic cells
Opsonization of pathogens and internalization of immune complexes by phagocytic cells
Assembly of membrane attack complex and lysis of certain pathogens by creating membrane pores on pathogen surfaces

4.5 BIBLIOGRAPHY

1.Boucher, R. C. 2001. Pathogenesis of cystic fibrosis airways disease. *Trans Am Clin Climatol Assoc* 112:99-107.

2.Boucher, R. C. 2002. An overview of the pathogenesis of cystic fibrosis lung disease. *Adv Drug Deliv Rev* 54(11):1359-71.

3.Boucher, R. C. 2003. Regulation of airway surface liquid volume by human airway epithelia. *Pflugers Arch* 445(4):495-8.

4.Perez-Vilar, J., and R. C. Boucher. 2004. Reevaluating gel-forming mucins' roles in cystic fibrosis lung disease. *Free Radic Biol Med* 37(10):1564-77.

5.Neth, O., I. Hann, M. W. Turner, and N. J. Klein. 2001. Deficiency of mannose-binding lectin and burden of infection in children with malignancy: a prospective study. *Lancet* 358(9282):614-8.

6.Davies, J., O. Neth, E. Alton, N. Klein, and M. Turner. 2000. Differential binding of mannose-binding lectin to respiratory pathogens in cystic fibrosis. *Lancet* 355(9218):1885-6.

7.Clark, H. W., K. B. Reid, and R. B. Sim. 2000. Collectins and innate immunity in the lung. *Microbes Infect* 2(3):273-8.

8.Geertsma, M. F., P. H. Nibbering, H. P. Haagsman, M. R. Daha, and R. van Furth. 1994. Binding of surfactant protein A to C1q receptors mediates phagocytosis of Staphylococcus aureus by monocytes. *Am J Physiol* 267(5 Pt 1):L578-84.

9.Hartshorn, K. L., E. Crouch, M. R. White, M. L. Colamussi, A. Kakkanatt, B. Tauber, V. Shepherd, and K. N. Sastry. 1998. Pulmonary surfactant proteins A and D enhance neutrophil uptake of bacteria. *Am J Physiol* 274(6 Pt 1):L958-69.

10.Hickman-Davis, J. M., J. R. Lindsey, S. Zhu, and S. Matalon. 1998. Surfactant protein A mediates mycoplasmacidal activity of alveolar macrophages. *Am J Physiol* 274(2 Pt 1):L270-7.

11.Kabha, K., J. Schmegner, Y. Keisari, H. Parolis, J. Schlepper-Schaeffer, and I. Ofek. 1997. SP-A enhances phagocytosis of Klebsiella by interaction with capsular polysaccharides and alveolar macrophages. *Am J Physiol* 272(2 Pt 1):L344-52.

12.Mariencheck, W. I., J. Savov, Q. Dong, M. J. Tino, and J. R. Wright. 1999. Surfactant protein A enhances alveolar macrophage phagocytosis of a live, mucoid strain of P. aeruginosa. *Am J Physiol* 277(4 Pt 1):L777-86.

13.Restrepo, C. I., Q. Dong, J. Savov, W. I. Mariencheck, and J. R. Wright. 1999. Surfactant protein D stimulates phagocytosis of Pseudomonas aeruginosa by alveolar macrophages. *Am J Respir Cell Mol Biol* 21(5):576-85.

14.Thiel, S., T. Bjerke, D. Hansen, L. K. Poulsen, P. O. Schiotz, and J. C. Jensenius. 1995. Ontogeny of human mannan-binding protein, a lectin of the innate immune system. *Pediatr Allergy Immunol* 6(1):20-3.

15.Ballard, P. L. 2000. Scientific rationale for the use of antenatal glucocorticoids to promote fetal development. *Pediatr Rev* 1(5):E83-90.

16.Wong, C. J., J. Akiyama, L. Allen, and S. Hawgood. 1996. Localization and developmental expression of surfactant proteins D and A in the respiratory tract of the mouse. *Pediatr Res* 39(6):930-7.

17.Lin, Z., D. deMello, D. S. Phelps, W. A. Koltun, M. Page, and J. Floros. 2001. Both human SP-A1 and Sp-A2 genes are expressed in small and large intestine. *Pediatr Pathol Mol Med* 20(5):367-86.

18.Madsen, J., A. Kliem, I. Tornoe, K. Skjodt, C. Koch, and U. Holmskov. 2000. Localization of lung surfactant protein D on mucosal surfaces in human tissues. *J Immunol* 164(11):5866-70.

19.Stahlman, M. T., M. E. Gray, W. M. Hull, and J. A. Whitsett. 2002. Immunolocalization of surfactant protein-D (SP-D) in human fetal, newborn, and adult tissues. *J Histochem Cytochem* 50(5):651-60.

20.Fisher, J. H., and R. Mason. 1995. Expression of pulmonary surfactant protein D in rat gastric mucosa. *Am J Respir Cell Mol Biol* 12(1):13-8.

21.Hoover, R. R., and J. Floros. 1998. Organization of the human SP-A and SP-D loci at 10q22-q23. Physical and radiation hybrid mapping reveal gene order and orientation. *Am J Respir Cell Mol Biol* 18(3):353-62.

22.Floros, J., S. DiAngelo, M. Koptides, A. M. Karinch, P. K. Rogan, H. Nielsen, R. G. Spragg, K. Watterberg, and G. Deiter. 1996. Human SP-A locus: allele frequencies and linkage disequilibrium between the two surfactant protein A genes. *Am J Respir Cell Mol Biol* 15(4):489-98.

23.Katyal, S. L., G. Singh, and J. Locker. 1992. Characterization of a second human pulmonary surfactant-associated protein SP-A gene. *Am J Respir Cell Mol Biol* 6(4):446-52.

24.White, R. T., D. Damm, J. Miller, K. Spratt, J. Schilling, S. Hawgood, B. Benson, and B. Cordell. 1985. Isolation and characterization of the human pulmonary surfactant apoprotein gene. *Nature* 317(6035):361-3.

25.Kolble, K., J. Lu, S. E. Mole, S. Kaluz, and K. B. Reid. 1993. Assignment of the human pulmonary surfactant protein D gene (SFTP4) to 10q22-q23 close to the surfactant protein A gene cluster. *Genomics* 17(2):294-8.

26.Drickamer, K. 1988. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J Biol Chem* 263(20):9557-60.

27.Super, M., S. D. Gillies, S. Foley, K. Sastry, J. E. Schweinle, V. J. Silverman, and R. A. Ezekowitz. 1992. Distinct and overlapping functions of allelic forms of human mannose binding protein. *Nat Genet* 2(1):50-5.

28.Holmskov, U., S. Thiel, and J. C. Jensenius. 2003. Collections and ficolins: humoral lectins of the innate immune defense. *Annu Rev Immunol* 21:547-78.

29.Madsen, H. O., P. Garred, S. Thiel, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, and A. Svejgaard. 1995. Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol* 155(6):3013-20.

30.Neth, O., D. L. Jack, A. W. Dodds, H. Holzel, N. J. Klein, and M. W. Turner. 2000. Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. *Infect Immun* 68(2):688-93.

31.Younger, J. G., S. Shankar-Sinha, M. Mickiewicz, A. S. Brinkman, G. A. Valencia, J. V. Sarma, E. M. Younkin, T. J. Standiford, F. S. Zetoune, and P. A. Ward. 2003. Murine complement interactions with Pseudomonas aeruginosa and their consequences during pneumonia. *Am J Respir Cell Mol Biol* 29(4):432-8.

32.Sumiya, M., M. Super, P. Tabona, R. J. Levinsky, T. Arai, M. W. Turner, and J. A. Summerfield. 1991. Molecular basis of opsonic defect in immunodeficient children. *Lancet* 337(8757):1569-70.

33.Lipscombe, R. J., Y. L. Lau, R. J. Levinsky, M. Sumiya, J. A. Summerfield, and M. W. Turner. 1992. Identical point mutation leading to low levels of mannose binding protein and poor C3b mediated opsonisation in Chinese and Caucasian populations. *Immunol Lett* 32(3):253-7.

34.Lipscombe, R. J., M. Sumiya, A. V. Hill, Y. L. Lau, R. J. Levinsky, J. A. Summerfield, and M. W. Turner. 1992. High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum Mol Genet* 1(9):709-15.

35.Madsen, H. O., P. Garred, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, S. Thiel, and A. Svejgaard. 1994. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* 40(1):37-44.

36.Larsen, F., H. O. Madsen, R. B. Sim, C. Koch, and P. Garred. 2004. Disease-associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein. *J Biol Chem* 279(20):21302-11.

37.Butler, G. S., D. Sim, E. Tam, D. Devine, and C. M. Overall. 2002. Mannose-binding lectin (MBL) mutants are susceptible to matrix metalloproteinase proteolysis: potential role in human MBL deficiency. *J Biol Chem* 277(20):17511-9.

38.Terai, I., K. Kobayashi, M. Matsushita, H. Miyakawa, N. Mafune, and H. Kikuta. 2003. Relationship between gene polymorphisms of mannose-binding lectin (MBL) and two molecular forms of MBL. *Eur J Immunol* 33(10):2755-63.

39.Crosdale, D. J., W. E. Ollier, W. Thomson, P. A. Dyer, J. Jensenious, R. W. Johnson, and K. V. Poulton. 2000. Mannose binding lectin (MBL) genotype distributions with relation to serum levels in UK Caucasoids. *Eur J Immunogenet* 27(3):111-7.

40.Jack, D. L., G. A. Jarvis, C. L. Booth, M. W. Turner, and N. J. Klein. 2001. Mannose-binding lectin accelerates complement activation and increases serum killing of Neisseria meningitidis serogroup C. *J Infect Dis* 184(7):836-45.

41.Jack, D. L., N. J. Klein, and M. W. Turner. 2001. Mannose-binding lectin: targeting the microbial world for complement attack and opsonophagocytosis. *Immunol Rev* 180:86-99.

42.Jack, D. L., R. C. Read, A. J. Tenner, M. Frosch, M. W. Turner, and N. J. Klein. 2001. Mannose-binding lectin regulates the inflammatory response of human professional phagocytes to Neisseria meningitidis serogroup B. *J Infect Dis* 184(9):1152-62.

43.Garred, P., T. Pressler, H. O. Madsen, B. Frederiksen, A. Svejgaard, N. Hoiby, M. Schwartz, and C. Koch. 1999. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest* 104(4):431-7.

44.van Helden, H. P., W. C. Kuijpers, D. Steenvoorden, C. Go, P. L. Bruijnzeel, M. van Eijk, and H. P. Haagsman. 1997. Intratracheal aerosolization of endotoxin (LPS) in the rat: a comprehensive animal model to study adult (acute) respiratory distress syndrome. *Exp Lung Res* 23(4):297-316.

45.McIntosh, J. C., A. H. Swyers, J. H. Fisher, and J. R. Wright. 1996. Surfactant proteins A and D increase in response to intratracheal lipopolysaccharide. *Am J Respir Cell Mol Biol* 15(4):509-19.

46.McNeely, T. B., and J. D. Coonrod. 1994. Aggregation and opsonization of type A but not type B Hemophilus influenzae by surfactant protein A. *Am J Respir Cell Mol Biol* 11(1):114-22.

47.McNeely, T. B., and J. D. Coonrod. 1993. Comparison of the opsonic activity of human surfactant protein A for Staphylococcus aureus and Streptococcus pneumoniae with rabbit and human macrophages. *J Infect Dis* 167(1):91-7.

48.Manz-Keinke, H., H. Plattner, and J. Schlepper-Schafer. 1992. Lung surfactant protein A (SP-A) enhances serum-independent phagocytosis of bacteria by alveolar macrophages. *Eur J Cell Biol* 57(1):95-100.

49.Kalina, M., H. Blau, S. Riklis, and V. Kravtsov. 1995. Interaction of surfactant protein A with bacterial lipopolysaccharide may affect some biological functions. *Am J Physiol* 268(1 Pt 1):L144-51.

50.Van Iwaarden, J. F., J. C. Pikaar, J. Storm, E. Brouwer, J. Verhoef, R. S. Oosting, L. M. van Golde, and J. A. van Strijp. 1994. Binding of surfactant protein A to the lipid A moiety of bacterial lipopolysaccharides. *Biochem J* 303 (Pt 2):407-11.

51.Gaynor, C. D., F. X. McCormack, D. R. Voelker, S. E. McGowan, and L. S. Schlesinger. 1995. Pulmonary surfactant protein A mediates enhanced phagocytosis of Mycobacterium tuberculosis by a direct interaction with human macrophages. *J Immunol* 155(11):5343-51.

52.DiAngelo, S., Z. Lin, G. Wang, S. Phillips, M. Ramet, J. Luo, and J. Floros. 1999. Novel, non-radioactive, simple and multiplex PCR-cRFLP methods for genotyping human SP-A and SP-D marker alleles. *Dis Markers* 15(4):269-81.

53.Liu, W., C. M. Bentley, and J. Floros. 2003. Study of human SP-A, SP-B and SP-D loci: allele frequencies, linkage disequilibrium and heterozygosity in different races and ethnic groups. *BMC Genet* 4(1):13.

54.Crouch, E. C. 1998. Structure, biologic properties, and expression of surfactant protein D (SP-D). *Biochim Biophys Acta* 1408(2-3):278-89.

55.Motwani, M., R. A. White, N. Guo, L. L. Dowler, A. I. Tauber, and K. N. Sastry. 1995. Mouse surfactant protein-D. cDNA cloning, characterization, and gene localization to chromosome 14. *J Immunol* 155(12):5671-7.

56.Hartshorn, K. L., E. C. Crouch, M. R. White, P. Eggleton, A. I. Tauber, D. Chang, and K. Sastry. 1994. Evidence for a protective role of pulmonary surfactant protein D (SP-D) against influenza A viruses. *J Clin Invest* 94(1):311-9.

57.Leth-Larsen, R., P. Garred, H. Jensenius, J. Meschi, K. Hartshorn, J. Madsen, I. Tornoe, H. O. Madsen, G. Sorensen, E. Crouch, and U. Holmskov. 2005. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 174(3):1532-8.

58.Hermans, C., and A. Bernard. 1999. Lung epithelium-specific proteins: characteristics and potential applications as markers. *Am J Respir Crit Care Med* 159(2):646-78.

59.Tobin, M. J. 2002. Pediatrics, surfactant, and cystic fibrosis in AJRCCM 2001. *Am J Respir Crit Care Med* 165(5):619-30.

60.Takahashi, H., T. Fujishima, H. Koba, S. Murakami, K. Kurokawa, Y. Shibuya, M. Shiratori, Y. Kuroki, and S. Abe. 2000. Serum surfactant proteins A and D as prognostic factors in idiopathic pulmonary fibrosis and their relationship to disease extent. *Am J Respir Crit Care Med* 162(3 Pt 1):1109-14.

61.Takahashi, H., Y. Kuroki, H. Tanaka, T. Saito, K. Kurokawa, H. Chiba, A. Sagawa, H. Nagae, and S. Abe. 2000. Serum levels of surfactant proteins A and D are useful biomarkers for interstitial lung disease in patients with progressive systemic sclerosis. *Am J Respir Crit Care Med* 162(1):258-63.

62.Asano, Y., H. Ihn, K. Yamane, N. Yazawa, M. Kubo, M. Fujimoto, and K. Tamaki. 2001. Clinical significance of surfactant protein D as a serum marker for evaluating pulmonary fibrosis in patients with systemic sclerosis. *Arthritis Rheum* 44(6):1363-9.

63.Ohnishi, H., A. Yokoyama, K. Kondo, H. Hamada, M. Abe, K. Nishimura, K. Hiwada, and N. Kohno. 2002. Comparative study of KL-6, surfactant protein-A, surfactant protein-D, and monocyte chemoattractant protein-1 as serum markers for interstitial lung diseases. *Am J Respir Crit Care Med* 165(3):378-81.

64.Greene, K. E., T. E. King, Jr., Y. Kuroki, B. Bucher-Bartelson, G. W. Hunninghake, L. S. Newman, H. Nagae, and R. J. Mason. 2002. Serum surfactant proteins-A and -D as biomarkers in idiopathic pulmonary fibrosis. *Eur Respir J* 19(3):439-46.

65.Mikerov, A. N., T. M. Umstead, W. Huang, W. Liu, D. S. Phelps, and J. Floros. 2005. SP-A1 and SP-A2 variants differentially enhance association of Pseudomonas aeruginosa with rat alveolar macrophages. *Am J Physiol Lung Cell Mol Physiol* 288(1):L150-8.

66.LeVine, A. M., M. D. Bruno, K. M. Huelsman, G. F. Ross, J. A. Whitsett, and T. R. Korfhagen. 1997. Surfactant protein A-deficient mice are susceptible to group B streptococcal infection. *J Immunol* 158(9):4336-40.

67.LeVine, A. M., J. A. Whitsett, J. A. Gwozdz, T. R. Richardson, J. H. Fisher, M. S. Burhans, and T. R. Korfhagen. 2000. Distinct effects of surfactant protein A or D deficiency during bacterial infection on the lung. *J Immunol* 165(7):3934-40.

68.LeVine, A. M., J. A. Whitsett, K. L. Hartshorn, E. C. Crouch, and T. R. Korfhagen. 2001. Surfactant protein D enhances clearance of influenza A virus from the lung in vivo. *J Immunol* 167(10):5868-73.

69.Kala, P., T. Ten Have, H. Nielsen, M. Dunn, and J. Floros. 1998. Association of pulmonary surfactant protein A (SP-A) gene and respiratory distress syndrome: interaction with SP-B. *Pediatr Res* 43(2):169-77.

70.Lahti, M., J. Lofgren, R. Marttila, M. Renko, T. Klaavuniemi, R. Haataja, M. Ramet, and M. Hallman. 2002. Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 51(6):696-9.

71.Guo, X., H. M. Lin, Z. Lin, M. Montano, R. Sansores, G. Wang, S. DiAngelo, A. Pardo, M. Selman, and J. Floros. 2000. Polymorphisms of surfactant protein gene A, B, D, and of SP-B-linked microsatellite markers in COPD of a Mexican population. *Chest* 117(5 Suppl 1):249S-50S.

72.Guo, X., H. M. Lin, Z. Lin, M. Montano, R. Sansores, G. Wang, S. DiAngelo, A. Pardo, M. Selman, and J. Floros. 2001. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 18(3):482-90.

73.Floros, J., R. Fan, S. Diangelo, X. Guo, J. Wert, and J. Luo. 2001. Surfactant protein (SP) B associations and interactions with SP-A in white and black subjects with respiratory distress syndrome. *Pediatr Int* 43(6):567-76.

74.Janeway, C. A., Jr. 2002. A trip through my life with an immunological theme. *Annu Rev Immunol* 20:1-28.

75.Janeway, C. A., Jr., and R. Medzhitov. 2002. Innate immune recognition. *Annu Rev Immunol* 20:197-216.

76.Ikeda, K., T. Sannoh, N. Kawasaki, T. Kawasaki, and I. Yamashina. 1987. Serum lectin with known structure activates complement through the classical pathway. *J Biol Chem* 262(16):7451-4.

77.Fujita, T., M. Matsushita, and Y. Endo. 2004. The lectin-complement pathway--its role in innate immunity and evolution. *Immunol Rev* 198:185-202.

78.Thiel, S. 1992. Mannan-binding protein, a complement activating animal lectin. *Immunopharmacology* 24(2):91-9.

79.Dahl, M. R., S. Thiel, M. Matsushita, T. Fujita, A. C. Willis, T. Christensen, T. Vorup-Jensen, and J. C. Jensenius. 2001. MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway. *Immunity* 15(1):127-35.

80.Ezekowitz, R. A. 2003. Role of the mannose-binding lectin in innate immunity. *J Infect Dis* 187 Suppl 2:S335-9.

81.Turner, M. W. 1996. Mannose-binding lectin: the pluripotent molecule of the innate immune system. *Immunol Today* 17(11):532-40.

82.Matsushita, M., R. A. Ezekowitz, and T. Fujita. 1995. The Gly-54-->Asp allelic form of human mannose-binding protein (MBP) fails to bind MBP-associated serine protease. *Biochem J* 311 (Pt 3):1021-3.

83.Matsushita, M., S. Thiel, J. C. Jensenius, I. Terai, and T. Fujita. 2000. Proteolytic activities of two types of mannose-binding lectin-associated serine protease. *J Immunol* 165(5):2637-42.

84.Kurata, H., H. M. Cheng, Y. Kozutsumi, Y. Yokota, and T. Kawasaki. 1993. Role of the collagen-like domain of the human serum mannan-binding protein in the activation of complement and the secretion of this lectin. *Biochem Biophys Res Commun* 191(3):1204-10.

85.Garred, P., K. Brygge, C. H. Sorensen, H. O. Madsen, S. Thiel, and A. Svejgaard. 1993. Mannan-binding protein--levels in plasma and upper-airways secretions and frequency of genotypes in children with recurrence of otitis media. *Clin Exp Immunol* 94(1):99-104.

86.Koch, A., M. Melbye, P. Sorensen, P. Homoe, H. O. Madsen, K. Molbak, C. H. Hansen, L. H. Andersen, G. W. Hahn, and P. Garred. 2001. Acute respiratory tract infections and mannose-binding lectin insufficiency during early childhood. *Jama* 285(10):1316-21.

87.Summerfield, J. A., M. Sumiya, M. Levin, and M. W. Turner. 1997. Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series. *Bmj* 314(7089):1229-32.

88.Garred, P., A. Voss, H. O. Madsen, and P. Junker. 2001. Association of mannose-binding lectin gene variation with disease severity and infections in a population-based cohort of systemic lupus erythematosus patients. *Genes Immun* 2(8):442-50.

89.Summerfield, J. A., S. Ryder, M. Sumiya, M. Thursz, A. Gorchein, M. A. Monteil, and M. W. Turner. 1995. Mannose binding protein gene mutations associated with unusual and severe infections in adults. *Lancet* 345(8954):886-9.

90.Kielgast, S., S. Thiel, T. B. Henriksen, T. Bjerke, J. Olsen, and J. C. Jensenius. 2003. Umbilical cord mannan-binding lectin and infections in early childhood. *Scand J Immunol* 57(2):167-72.

91.Garred, P., H. O. Madsen, P. Halberg, J. Petersen, G. Kronborg, A. Svejgaard, V. Andersen, and S. Jacobsen. 1999. Mannose-binding lectin polymorphisms and susceptibility to infection in systemic lupus erythematosus. *Arthritis Rheum* 42(10):2145-52.

92.Garred, P., C. Richter, A. B. Andersen, H. O. Madsen, I. Mtoni, A. Svejgaard, and J. Shao. 1997. Mannan-binding lectin in the sub-Saharan HIV and tuberculosis epidemics. *Scand J Immunol* 46(2):204-8.

93.Garred, P., H. O. Madsen, U. Balslev, B. Hofmann, C. Pedersen, J. Gerstoft, and A. Svejgaard. 1997. Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of mannose-binding lectin. *Lancet* 349(9047):236-40.

94.Crosdale, D. J., K. V. Poulton, W. E. Ollier, W. Thomson, and D. W. Denning. 2001. Mannose-binding lectin gene polymorphisms as a susceptibility factor for chronic necrotizing pulmonary aspergillosis. *J Infect Dis* 184(5):653-6.

95.Yang, I. A., S. L. Seeney, J. M. Wolter, E. M. Anders, J. G. McCormack, A. M. Tunnicliffe, G. C. Rabnott, J. G. Shaw, A. G. Dent, S. T. Kim, P. V. Zimmerman, and K. M. Fong. 2003. Mannose-binding lectin gene polymorphism predicts hospital admissions for COPD infections. *Genes Immun* 4(4):269-74.

96.Peterslund, N. A., C. Koch, J. C. Jensenius, and S. Thiel. 2001. Association between deficiency of mannose-binding lectin and severe infections after chemotherapy. *Lancet* 358(9282):637-8.

97.Garred, P., M. Harboe, T. Oettinger, C. Koch, and A. Svejgaard. 1994. Dual role of mannan-binding protein in infections: another case of heterosis? *Eur J Immunogenet* 21(2):125-31.

98.Luty, A. J., J. F. Kun, and P. G. Kremsner. 1998. Mannose-binding lectin plasma levels and gene polymorphisms in Plasmodium falciparum malaria. *J Infect Dis* 178(4):1221-4.

99.Kelly, P., D. L. Jack, A. Naeem, B. Mandanda, R. C. Pollok, N. J. Klein, M. W. Turner, and M. J. Farthing. 2000. Mannose-binding lectin is a component of innate mucosal defense against Cryptosporidium parvum in AIDS. *Gastroenterology* 119(5):1236-42.

100.Santos, I. K., C. H. Costa, H. Krieger, M. F. Feitosa, D. Zurakowski, B. Fardin, R. B. Gomes, D. L. Weiner, D. A. Harn, R. A. Ezekowitz, and J. E. Epstein. 2001. Mannan-binding lectin enhances susceptibility to visceral leishmaniasis. *Infect Immun* 69(8):5212-5.

101. Garred, P., F. Larsen, J. Seyfarth, R. Fujita, and H. O. Madsen. 2006. Mannose-binding lectin and its genetic variants. *Genes Immun* 7(2):85-94.
102. Trevisiol, C., M. Boniotto, L. Giglio, F. Poli, M. Morgutti, and S. Crovella. 2005. MBL2 polymorphisms screening in a regional Italian CF Center. *J Cyst Fibros* 4(3):189-91.
103. Gabolde, M., M. Guilloud-Bataille, J. Feingold, and C. Besmond. 1999. Association of variant alleles of mannose binding lectin with severity of pulmonary disease in cystic fibrosis: cohort study. *Bmj* 319(7218):1166-7.
104. Choi, E. H., M. Ehrmantraut, C. B. Foster, J. Moss, and S. J. Chanock. 2006. Association of common haplotypes of surfactant protein A1 and A2 (SFTPA1 and SFTPA2) genes with severity of lung disease in cystic fibrosis. *Pediatr Pulmonol* 41(3):255-62.
105. Wacholder, S., S. Chanock, M. Garcia-Closas, L. El Ghormli, and N. Rothman. 2004. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 96(6):434-42.
106. Davies, J. C., M. W. Turner, and N. Klein. 2004. Impaired pulmonary status in cystic fibrosis adults with two mutated MBL-2 alleles. *Eur Respir J* 24(5):798-804.
107. Carlsson, M., A. G. Sjoholm, L. Eriksson, S. Thiel, J. C. Jensenius, M. Segelmark, and L. Truedsson. 2005. Deficiency of the mannan-binding lectin pathway of complement and poor outcome in cystic fibrosis: bacterial colonization may be decisive for a relationship. *Clin Exp Immunol* 139(2):306-13.
108. Drumm, M. L., M. W. Konstan, M. D. Schluchter, A. Handler, R. Pace, F. Zou, M. Zariwala, D. Fargo, A. Xu, J. M. Dunn, R. J. Darrah, R. Dorfman, A. J. Sandford, M. Corey, J. Zielenski, P. Durie, K. Goddard, J. R. Yankaskas, F. A. Wright, and M. R. Knowles. 2005. Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* 353(14):1443-53.
109. Wert, S. E., M. Yoshida, A. M. LeVine, M. Ikegami, T. Jones, G. F. Ross, J. H. Fisher, T. R. Korfhagen, and J. A. Whitsett. 2000. Increased metalloprotenase activity, oxidant production, and emphysema in surfactant protein D gene-inactivated mice. *Proceeds of the National Academy of Science* 97:5972-.
110. Bufler, P., B. Schmidt, D. Schikor, A. Bauernfeind, E. C. Crouch, and M. Griese. 2003. Surfactant protein A and D differently regulate the immune response to nonmucoid Pseudomonas aeruginosa and its lipopolysaccharide. *Am J Respir Cell Mol Biol* 28(2):249-56.
111. Chabot, S., L. Salez, F. X. McCormack, L. Touqui, and M. Chignard. 2003. Surfactant protein A inhibits lipopolysaccharide-induced in vivo production of interleukin-10 by

mononuclear phagocytes during lung inflammation. *Am J Respir Cell Mol Biol* 28(3):347-53.

112. Hussain, S., J. R. Wright, and W. J. Martin, 2nd. 2003. Surfactant protein A decreases nitric oxide production by macrophages in a tumor necrosis factor-alpha-dependent mechanism. *Am J Respir Cell Mol Biol* 28(4):520-7.

113. Sano, H., H. Chiba, D. Iwaki, H. Sohma, D. R. Voelker, and Y. Kuroki. 2000. Surfactant proteins A and D bind CD14 by different mechanisms. *J Biol Chem* 275:22442.

114. Gardai, S. J., Q. Xiao, M. Dickinson, J. A. Nick, D. R. Voelker, K. E. Greene, and P. M. Henson. 2003. By binding SIRPa or calreticulin/CD91, lung collectins act as dual function surveillance to suppress or enhance inflammation. *Cell* 115:13.

115. Haataja, R., R. Marttila, P. Uimari, J. Lofgren, M. Ramet, and M. Hallman. 2001. Respiratory distress syndrome: evaluation of genetic susceptibility and protection by transmission disequilibrium test. *Hum Genet* 109(3):351-5.

116. Haataja, R., M. Ramet, R. Marttila, and M. Hallman. 2000. Surfactant proteins A and B as interactive genetic determinants of neonatal respiratory distress syndrome. *Hum Mol Genet* 9(18):2751-60.

117. deMello, D. E., S. Heyman, D. S. Phelps, and J. Floros. 1993. Immunogold localization of SP-A in lungs of infants dying from respiratory distress syndrome. *Am J Pathol* 142(5):1631-40.

118. Moya, F. R., H. F. Montes, V. L. Thomas, A. M. Mouzinho, J. F. Smith, and C. R. Rosenfeld. 1994. Surfactant protein A and saturated phosphatidylcholine in respiratory distress syndrome. *Am J Respir Crit Care Med* 150(6 Pt 1):1672-7.

119. Lin, Z., C. Pearson, V. Chinchilli, S. M. Pietschmann, J. Luo, U. Pison, and J. Floros. 2000. Polymorphisms of human SP-A, SP-B, and SP-D genes: association of SP-B Thr131Ile with ARDS. *Clin Genet* 58(3):181-91.

120. Muhdi, K., F. P. Edenborough, L. Gumery, S. O'Hickey, E. G. Smith, D. L. Smith, and D. E. Stableforth. 1996. Outcome for patients colonised with Burkholderia cepacia in a Birmingham adult cystic fibrosis clinic and the end of an epidemic. *Thorax* 51(4):374-7.

121. Tablan, O. C., T. L. Chorba, D. V. Schidlow, J. W. White, K. A. Hardy, P. H. Gilligan, W. M. Morgan, L. A. Carson, W. J. Martone, J. M. Jason, and et al. 1985. Pseudomonas cepacia colonization in patients with cystic fibrosis: risk factors and clinical outcome. *J Pediatr* 107(3):382-7.

122. Thomassen, M. J., C. A. Demko, J. D. Klinger, and R. C. Stern. 1985. Pseudomonas cepacia colonization among patients with cystic fibrosis. A new opportunist. *Am Rev Respir Dis* 131(5):791-6.

123. Isles, A., I. Maclusky, M. Corey, R. Gold, C. Prober, P. Fleming, and H. Levison. 1984. Pseudomonas cepacia infection in cystic fibrosis: an emerging problem. *J Pediatr* 104(2):206-10.

124. Lewin, L. O., P. J. Byard, and P. B. Davis. 1990. Effect of Pseudomonas cepacia colonization on survival and pulmonary function of cystic fibrosis patients. *J Clin Epidemiol* 43(2):125-31.

125. Gladman, G., P. J. Connor, R. F. Williams, and T. J. David. 1992. Controlled study of Pseudomonas cepacia and Pseudomonas maltophilia in cystic fibrosis. *Arch Dis Child* 67(2):192-5.

126. Thomassen, M. J., C. A. Demko, C. F. Doershuk, R. C. Stern, and J. D. Klinger. 1986. Pseudomonas cepacia: decrease in colonization in patients with cystic fibrosis. *Am Rev Respir Dis* 134(4):669-71.

127. Taylor, R. F., H. Gaya, and M. E. Hodson. 1993. Pseudomonas cepacia: pulmonary infection in patients with cystic fibrosis. *Respir Med* 87(3):187-92.

128. Jacques, I., J. Derelle, M. Weber, and M. Vidailhet. 1998. Pulmonary evolution of cystic fibrosis patients colonized by Pseudomonas aeruginosa and/or Burkholderia cepacia. *Eur J Pediatr* 157(5):427-31.

129. Geisenberger, O., M. Givskov, K. Riedel, N. Hoiby, B. Tummler, and L. Eberl. 2000. Production of N-acyl-L-homoserine lactones by P. aeruginosa isolates from chronic lung infections associated with cystic fibrosis. *FEMS Microbiol Lett* 184(2):273-8.

130. Chaparro, C., J. Maurer, C. Gutierrez, M. Krajden, C. Chan, T. Winton, S. Keshavjee, M. Scavuzzo, E. Tullis, M. Hutcheon, and S. Kesten. 2001. Infection with Burkholderia cepacia in cystic fibrosis: outcome following lung transplantation. *Am J Respir Crit Care Med* 163(1):43-8.

131. De Boeck, K., A. Malfroot, L. Van Schil, P. Lebecque, C. Knoop, J. R. Govan, C. Doherty, S. Laevens, and P. Vandamme. 2004. Epidemiology of Burkholderia cepacia complex colonisation in cystic fibrosis patients. *Eur Respir J* 23(6):851-6.

132. Aris, R. M., J. C. Routh, J. J. LiPuma, D. G. Heath, and P. H. Gilligan. 2001. Lung transplantation for cystic fibrosis patients with Burkholderia cepacia complex. Survival linked to genomovar type. *Am J Respir Crit Care Med* 164(11):2102-6.

133. LiPuma, J. J. 1998. Burkholderia cepacia. Management issues and new insights. *Clin Chest Med* 19(3):473-86, vi.

134. Lipuma, J. J., K. A. Marks-Austin, D. S. Holsclaw, Jr., G. B. Winnie, P. H. Gilligan, and T. L. Stull. 1994. Inapparent transmission of Pseudomonas (Burkholderia) cepacia among patients with cystic fibrosis. *Pediatr Infect Dis J* 13(8):716-9.

135. Mahenthiralingam, E., P. Vandamme, M. E. Campbell, D. A. Henry, A. M. Gravelle, L. T. Wong, A. G. Davidson, P. G. Wilcox, B. Nakielna, and D. P. Speert. 2001. Infection with Burkholderia cepacia complex genomovars in patients with cystic fibrosis: virulent transmissible strains of genomovar III can replace Burkholderia multivorans. *Clin Infect Dis* 33(9):1469-75.

136. Mahenthiralingam, E., A. Baldwin, and P. Vandamme. 2002. Burkholderia cepacia complex infection in patients with cystic fibrosis. *J Med Microbiol* 51(7):533-8.

137. Coenye, T., J. J. LiPuma, D. Henry, B. Hoste, K. Vandemeulebroecke, M. Gillis, D. P. Speert, and P. Vandamme. 2001. Burkholderia cepacia genomovar VI, a new member of the Burkholderia cepacia complex isolated from cystic fibrosis patients. *Int J Syst Evol Microbiol* 51(Pt 2):271-9.

138. Coenye, T., E. Mahenthiralingam, D. Henry, J. J. LiPuma, S. Laevens, M. Gillis, D. P. Speert, and P. Vandamme. 2001. Burkholderia ambifaria sp. nov., a novel member of the Burkholderia cepacia complex including biocontrol and cystic fibrosis-related isolates. *Int J Syst Evol Microbiol* 51(Pt 4):1481-90.

139. Coenye, T., and J. J. LiPuma. 2003. Population structure analysis of Burkholderia cepacia genomovar III: varying degrees of genetic recombination characterize major clonal complexes. *Microbiology* 149(Pt 1):77-88.

140. Coenye, T., P. Vandamme, J. J. LiPuma, J. R. Govan, and E. Mahenthiralingam. 2003. Updated version of the Burkholderia cepacia complex experimental strain panel. *J Clin Microbiol* 41(6):2797-8.

141. Coenye, T., and J. J. LiPuma. 2002. Multilocus restriction typing: a novel tool for studying global epidemiology of Burkholderia cepacia complex infection in cystic fibrosis. *J Infect Dis* 185(10):1454-62.

142. Henry, D. A., E. Mahenthiralingam, P. Vandamme, T. Coenye, and D. P. Speert. 2001. Phenotypic methods for determining genomovar status of the Burkholderia cepacia complex. *J Clin Microbiol* 39(3):1073-8.

143. Vandamme, P., D. Henry, T. Coenye, S. Nzula, M. Vancanneyt, J. J. LiPuma, D. P. Speert, J. R. Govan, and E. Mahenthiralingam. 2002. Burkholderia anthina sp. nov. and Burkholderia pyrrocinia, two additional Burkholderia cepacia complex bacteria, may confound results of new molecular diagnostic tools. *FEMS Immunol Med Microbiol* 33(2):143-9.

144. Vandamme, P., B. Holmes, T. Coenye, J. Goris, E. Mahenthiralingam, J. J. LiPuma, and J. R. Govan. 2003. Burkholderia cenocepacia sp. nov.--a new twist to an old story. *Res Microbiol* 154(2):91-6.

145. Vandamme, P., B. Holmes, M. Vancanneyt, T. Coenye, B. Hoste, R. Coopman, H. Revets, S. Lauwers, M. Gillis, K. Kersters, and J. R. Govan. 1997. Occurrence of multiple genomovars of Burkholderia cepacia in cystic fibrosis patients and proposal of Burkholderia multivorans sp. nov. *Int J Syst Bacteriol* 47(4):1188-200.

146. Vandamme, P., E. Mahenthiralingam, B. Holmes, T. Coenye, B. Hoste, P. De Vos, D. Henry, and D. P. Speert. 2000. Identification and population structure of Burkholderia stabilis sp. nov. (formerly Burkholderia cepacia genomovar IV). *J Clin Microbiol* 38(3):1042-7.

147. Sajjan, U., S. Keshavjee, and J. Forstner. 2004. Responses of well-differentiated airway epithelial cell cultures from healthy donors and patients with cystic fibrosis to Burkholderia cenocepacia infection. *Infect Immun* 72(7):4188-99.

148. Chiu, C. H., S. Wong, R. E. Hancock, and D. P. Speert. 2001. Adherence of Burkholderia cepacia to respiratory tract epithelial cells and inhibition with dextrans. *Microbiology* 147(Pt 10):2651-8.

149. Fink, J., J. H. Steer, D. A. Joyce, A. S. McWilliam, and G. A. Stewart. 2003. Pro-inflammatory effects of Burkholderia cepacia on cystic fibrosis respiratory epithelium. *FEMS Immunol Med Microbiol* 38(3):273-82.

150. Sajjan, U., M. Corey, A. Humar, E. Tullis, E. Cutz, C. Ackerley, and J. Forstner. 2001. Immunolocalisation of Burkholderia cepacia in the lungs of cystic fibrosis patients. *J Med Microbiol* 50(6):535-46.

151. Tomlin, K. L., O. P. Coll, and H. Ceri. 2001. Interspecies biofilms of Pseudomonas aeruginosa and Burkholderia cepacia. *Can J Microbiol* 47(10):949-54.

152. Saini, L. S., S. B. Galsworthy, M. A. John, and M. A. Valvano. 1999. Intracellular survival of Burkholderia cepacia complex isolates in the presence of macrophage cell activation. *Microbiology* 145 (Pt 12):3465-75.

153. Martin, D. W., and C. D. Mohr. 2000. Invasion and intracellular survival of Burkholderia cepacia. *Infect Immun* 68(1):24-9.

154. Tomashefski, J. F., M. J. Thomassen, M. C. Bruce, H. I. Goldberg, M. W. Konstan, and R. C. Stern. 1988. Pseudomonas cepacia-associated pneumonia in cystic fibrosis. Relation of clinical features to histopathologic patterns of pneumonia. *Arch Pathol Lab Med* 112(2):166-72.

155. Frangolias, D. D., E. Mahenthiralingam, S. Rae, J. M. Raboud, A. G. Davidson, R. Wittmann, and P. G. Wilcox. 1999. Burkholderia cepacia in cystic fibrosis. Variable disease course. *Am J Respir Crit Care Med* 160(5 Pt 1):1572-7.

156. Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68(4):978-89.

157. Stephens, M., and P. Donnelly. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 73(5):1162-9.

158. Casanova, J.-L., and L. Abel. 2004. Human mannose-binding lectin in immunity: Friend, foe, or both? *J. Exp. Med* 199(10):1295-99.

159. Crouch, E., K. Hartshorn, and I. Ofek. 2000. Collectins and pulmonary innate immunity. *Immunol Rev* 173:52-65.

160. Wu, H., A. Kuzmenko, S. Wan, L. Schaffer, A. Weiss, J. H. Fisher, K. S. Kim, and F. X. McCormack. 2003. Surfactant proteins A and D inhibit the growth of Gram-negative bacteria by increasing membrane permeability. *J Clin Invest* 111(10):1589-602.

161. Karinch, A. M., D. E. deMello, and J. Floros. 1997. Effect of genotype on the levels of surfactant protein A mRNA and on the SP-A2 splice variants in adult humans. *Biochem J* 321 (Pt 1):39-47.

162. deMello, D. E., D. S. Phelps, G. Patel, J. Floros, and D. Lagunoff. 1989. Expression of the 35kDa and low molecular weight surfactant-associated proteins in the lungs of infants dying with respiratory distress syndrome. *Am J Pathol* 134(6):1285-93.

163. Chen, J. S., K. A. Witzmann, T. Spilker, R. J. Fink, and J. J. LiPuma. 2001. Endemicity and inter-city spread of Burkholderia cepacia genomovar III in cystic fibrosis. *J Pediatr* 139(5):643-9.

164. LiPuma, J. J. 1998. Burkholderia cepacia epidemiology and pathogenesis: implications for infection control. *Curr Opin Pulm Med* 4(6):337-41.

165. van Emmerik, L. C., E. J. Kuijper, C. A. Fijen, J. Dankert, and S. Thiel. 1994. Binding of mannan-binding protein to various bacterial pathogens of meningitis. *Clin Exp Immunol* 97(3):411-6.

166. Benne, C. A., B. Benaissa-Trouw, J. A. van Strijp, C. A. Kraaijeveld, and J. F. van Iwaarden. 1997. Surfactant protein A, but not surfactant protein D, is an opsonin for influenza A virus phagocytosis by rat alveolar macrophages. *Eur J Immunol* 27(4):886-90.

167. LeVine, A. M., K. E. Kurak, M. D. Bruno, J. M. Stark, J. A. Whitsett, and T. R. Korfhagen. 1998. Surfactant protein-A-deficient mice are succeptible to *Pseudomonas aeruginosa* infection. *Am J Respir Cell Mol Biol* 19:700-708.

168. Jack, D. L., and M. W. Turner. 2003. Anti-microbial activities of mannose-binding lectin. *Biochem Soc Trans* 31(Pt 4):753-7.

169. Beharka, A. A., C. D. Gaynor, B. K. Kang, D. R. Voelker, F. X. McCormack, and L. S. Schlesinger. 2002. Pulmonary surfactant protein A up-regulates activity of the mannose receptor, a pattern recognition receptor expressed on human macrophages. *J Immunol* 169(7):3565-73.

170. Tabona, P., A. Mellor, and J. A. Summerfield. 1995. Mannose binding protein is involved in first-line host defence: evidence from transgenic mice. *Immunology* 85(1):153-9.

171. Schelenz, S., R. Malhotra, R. B. Sim, U. Holmskov, and G. J. Bancroft. 1995. Binding of host collectins to the pathogenic yeast Cryptococcus neoformans: human surfactant protein D acts as an agglutinin for acapsular yeast cells. *Infect Immun* 63(9):3360-6.

172. Van Iwaarden, J. F., H. Shimizu, P. H. Van Golde, D. R. Voelker, and L. M. Van Golde. 1992. Rat surfactant protein D enhances the production of oxygen radicals by rat alveolar macrophages. *Biochem J* 286 (Pt 1):5-8.

173. Haurum, J. S., S. Thiel, I. M. Jones, P. B. Fischer, S. B. Laursen, and J. C. Jensenius. 1993. Complement activation upon binding of mannan-binding protein to HIV envelope glycoproteins. *Aids* 7(10):1307-13.

174. Madan, T., U. Kishore, A. Shah, P. Eggleton, P. Strong, J. Y. Wang, S. S. Aggrawal, P. U. Sarma, and K. B. Reid. 1997. Lung surfactant proteins A and D can inhibit specific IgE binding to the allergens of Aspergillus fumigatus and block allergen-induced histamine release from human basophils. *Clin Exp Immunol* 110(2):241-9.

175. Rosseau, S., A. Guenther, W. Seeger, and J. Lohmeyer. 1997. Phagocytosis of viable Candida albicans by alveolar macrophages: lack of opsonin function of surfactant protein A. *J Infect Dis* 175(2):421-8.

176. Madan, T., U. Kishore, M. Singh, P. Strong, H. Clark, E. M. Hussain, K. B. Reid, and P. U. Sarma. 2001. Surfactant proteins A and D protect mice against pulmonary hypersensitivity induced by Aspergillus fumigatus antigens and allergens. *J Clin Invest* 107(4):467-75.

177. McCormack, F. X., A. L. Festa, R. P. Andrews, M. Linke, and P. D. Walzer. 1997. The carbohydrate recognition domain of surfactant protein A mediates binding to the major surface glycoprotein of Pneumocystis carinii. *Biochemistry* 36(26):8092-9.

178. Zimmerman, P. E., D. R. Voelker, F. X. McCormack, J. R. Paulsrud, and W. J. Martin, 2nd. 1992. 120-kD surface glycoprotein of Pneumocystis carinii is a ligand for surfactant protein A. *J Clin Invest* 89(1):143-9.

179. Hoppe, H. C., B. J. de Wet, C. Cywes, M. Daffe, and M. R. Ehlers. 1997. Identification of phosphatidylinositol mannoside as a mycobacterial adhesin mediating both direct and opsonic binding to nonphagocytic mammalian cells. *Infect Immun* 65(9):3896-905.

180. Hickling, T. P., H. Bright, S. L. Martin, D. Gower, R. B. Sim, and R. Malhotra. 1999. A recombinant trimeric surfactant protein-D carbohydrate recognition domain inhibits RSV infection in vitro and in vivo. *Eur J Immunol* 29:3478-3484.

181. Malhotra, R., J. Haurum, S. Thiel, J. C. Jensenius, and R. B. Sim. 1993. Pollen grains bind to lung alveolar type II cells (A549) via lung surfactant protein A (SP-A). *Biosci Rep* 13(2):79-90.

182. Wang, J. Y., U. Kishore, B. L. Lim, P. Strong, and K. B. Reid. 1996. Interaction of human lung surfactant proteins A and D with mite (Dermatophagoides pteronyssinus) allergens. *Clin Exp Immunol* 106(2):367-73.

183. Janeway, C. A., Jr. and Travers, P. 1996. Immunobiology. The immune system in health and disease., 2ond ed. Current Biology Ltd./Garland Publishing Inc., London, San Fransisco, Philadelphia/New York and London.

CHAPTER 5: GLUTATHIONE METABOLISM ASSOCIATED GENES AS POTENTIAL MODIFIER GENES IN CYSTIC FIBROSIS

A version of this chapter may be published in the future.

5.0 INTRODUCTION

In this chapter we investigated genes which are involved in the removal of reactive oxygen species by glutathione peroxidases; genes that can influence glutathione synthesis and function (glutathione-S-transferases and the rate limiting enzyme glutamate cysteine ligase catalytic subunit). The effect of an amino acid changing polymorphism in glutathione-S-transferases P1 and gene deletions in glutathione-S-transferases M1 and T1, and a repeat polymorphism in the glutamate cysteine ligase catalytic subunit gene and their effects on pulmonary disease severity and progression, survival and succeptibility to infection with common CF pathogens were investigated.

5.1 RATIONALE FOR THE INVESTIGATION OF GLUTATHIONE METABOLISM GENES AS POTENTIAL MODIFIERS IN CF

The respiratory epithelium is a protective barrier that modulates inflammatory and metabolic responses to oxidative stress. The respiratory epithelium is exposed daily to oxidants which are generated as part of normal metabolism by the body as well oxidants which are inhaled and include: mineral dust, ozone, nitrogen oxides, ultraviolet and ionizing radiation, car exhaust (1) and cigarette smoke (2). Reactive oxygen species (ROS) are generated as part of:

- Cellular respiration in the mitochondrial electron transport chain by the cyclooxygenase pathway and other cellular enzymes (xanthine oxidase, cytochrome P450 oxidase),
- Normal host defense and chronic inflammation by inflammatory cells (i.e., neutrophils, eosinophils and alveolar macrophages (3)) and respiratory tissue (alveolar and bronchial epithelium and endothelium (4, 5)).

Most ROS are due to reduction of molecular oxygen and generation of superoxide anions (O_2^-). Other ROS include hydrogen peroxide (H_2O_2), nitric oxide, and hydroxyl radicals. ROS are removed by antioxidants and a complex system of antioxidant defense has evolved to maintain cellular homeostasis and includes numerous antioxidant enzymes such as glutathione peroxidases, catalase, superoxide dismutases, hemeoxygenase and others.

Oxidative stress results when there is an imbalance between oxidants and antioxidants. A build up of ROS has been shown to lead to peroxidation of cell membrane lipids, depletion of nicotinamide nucleotides, increases in intracellular calcium ions, cytoskeleton disruption and DNA damage (6). Oxidative stress has been implicated in contributing to tissue injury in numerous respiratory disorders (7).

In this chapter we explore genes which are involved in the removal of ROS by glutathione peroxidases. Reduced glutathione (GSH) is involved in metabolism, amino acid transport and protection of the cell from reactive molecules and toxic compounds of endogenous and exogenous origin (8). GSH, also called L-γ-glutamyl-L-cysteinyl-glycine, is a ubiquitous tripeptide which is produced by both animals and plants from the amino acids glutamine, glycine and cysteine. Cysteine is the rate-limiting amino acid in the reaction. The sulfur-hydrogen (thiol) group is a potent reducing agent and GSH is one of the two most important water-soluble antioxidants available (ascorbic acid is the other). GSH exists in two forms: the antioxidant reduced form (GSH) and the oxidized form which is a sulfur-linked compound and is also known as glutathione disulfide (GSSG) (Figure 5.1).

Water-soluble antioxidants are present in the cell cytosol and are also found extracellularly. GSH can be recycled. The CFTR protein is implicated in modulating epithelial reduced glutathione transport. The CFTR channel is permeable to chloride, but has also been shown to be permeable to other larger organic anions including reduced glutathione (9, 10). GSH is present in the airway surface lining (ASL) fluid, blood plasma, semen and saliva. In recent years abnormal transport of reduced glutathione in CF has been implicated in the lung pathology that has been documented in this disease. In addition, there are similarities in the lung pathology seen in CF with that in other diseases where GSH deficiency has been implicated. GSH is an important antioxidant and helps protect the lung from repeated and chronic exposure to reactive oxygen species.

GSH deficiency in the airway surface lining (ASL) fluid has been reported in: chronic obstructive pulmonary disease (COPD) (11), acute respiratory distress syndrome (12), idiopathic pulmonary fibrosis in non-smokers (13), idiopathic respiratory distress syndrome (14) and in HIV-positive patients (15-18). Konstan and coworkers (19) and others (20-22) have shown the presence of inflammation in the absence of infection in CF patients. In

addition, others have shown constitutively elevated levels of inflammatory mediators and cytokines in CF patients (23, 24). High levels of neutrophils have been reported in the bronchoalveolar lavage (BAL) fluid of CF patients, and have even been shown in CF patients who have stable and mild pulmonary disease during initial and early pulmonary infections (25). While CF patients in the early stages of pulmonary disease are capable of clearing pathogens without the use of antibiotic intervention (25), this is not the case over longer follow-up. Abman and colleagues have shown that as early as 2 years of age the capacity of CF patients to clear pathogens without adjuvant therapy is reduced and chronic colonization is common despite the continuous high influx of neutrophils into the CF airways (26).

GSH regulates inflammation and the immune response and promotes mucolysis. GSH deficiency leads to an enhanced inflammatory response and elevated levels of inflammatory cytokines which in turn leads to macrophage and neutrophil recruitment into the airways (27, 28). Depletion of extracellular GSH has been shown to compromise antioxidant capability and specifically oxidant-derived inactivation of the antiprotease system has been shown to result in unchecked activity of neutrophil elastase (29), which can contribute to lung damage. GSH can reduce the viscoelasticity of mucus by cleaving the disulfide bonds of secreted mucins (30). An increase in mucin secretion in the presence of *P. aeruginosa* has also been shown in GSH deficiency (31). Indirectly, GSH depletion has also been shown to lead to increased neutrophil elastase levels which also stimulate mucin secretion (32). Considering that CF is characterized by high levels of neutrophil elastase, mucus hypersecretion and increased mucus viscosity, deficiency in extracellular GSH will likely contribute to the pattern of lung damage shown in CF. Most likely some of this lung damage can be attributed to further oxidation of oxidized glutathione (GSSG) into long-lasting oxidants (chloramines) which have been shown to damage lung tissue (33, 34).

5.1.1 Tissue distribution and function of GSH

GSH is made by virtually every cell of the body. Redox equilibrium for GSH is established within each of these cells and the efflux of GSH from the cells establishes redox equilibrium outside the cells as well. The extracellular redox equilibrium varies with each body system and for the ASL fluid the level of GSH is 140 times the normal levels of GSH in blood plasma. GSH is synthesized in the cytoplasm and levels have been reported to range from 1-

10 mM (35), but up to 400 mM have been reported in ASL fluid of healthy non-CF patients (36). The lung is the major importer of GSH and ASL fluid contains high levels of GSH (36, 37). However, in CF, GSH levels are low in the ASL fluid (10) and also in the plasma (38).

5.1.2 GSH deficiency is a common characteristic in CF

Deficit in GSH levels in ASL fluid to 5-10% of normal and plasma levels to 50% of normal have been reported in CF patients (39). A systemic deficiency in GSH has been shown in CF and other chronic diseases. This extracellular deficiency has been purported to lead to oxidant damage of airway epithelial cells and lung fibrosis, and is associated with reduced pulmonary function and increased adhesion of bacteria to airway epithelial cells (11, 40-42). Additionally, reduced intracellular GSH levels have been shown in the peripheral blood lymphocytes of CF patients (43). Boxer and associates (44) showed increased release of hydrogen peroxide by leukocytes in the case of intracellular GSH deficiency. GSH directly neutralizes hydrogen peroxide (HOCl; the reaction is: 2 GSH + HOCl $\rightarrow$GSSG + H_20 + HCl) (44). Specifically, defective CFTR is hypothesized to create an antioxidant imbalance. The chronic oxidative stress in CF is exacerbated by decreased antioxidant capacity in the ASL fluid and plasma. GSH deficiency also causes oxidant injury to, and impairs the function of, the liver (45, 46) and the pancreas (47) and is associated with a number of clinical disorders (48-50).

Gao and co-workers have shown low levels of GSH in the apical fluid from CF airway epithelial cells and they showed that this decreased GSH content is due to defective GSH efflux (10). Lindsell and Hanrahan showed that chemical clamping of the CFTR channel in baby hamster kidney cells resulted in cessation of GSH efflux (9). Velsor and associates in their *in vivo* work demonstrated that absence of CFTR gene product:

- Affects (reduces) GSH and oxidized GSH concentrations in ASL fluid,
- Increases activities of airway antioxidant enzymes and,
- May stimulate oxidative stress in CFTR knock-out mice (51).

In their experiments (51) they investigated levels of GSH and oxidized GSH, glutathione reductase (GR), glutathione peroxidase (GPx) and γ-glutamyltransferase (γ-GT). They showed increased activity of GR and GPx and unchanged activity for γ-GT in ASL fluid and increased levels of some oxidative stress markers (i.e., thiobarbituric acid reactive

substances and 8-hydroxy-2-deoxyguanosine) in the lung tissue from CFTR knock-out mice (51).

Gao and associates (52) investigated restoring apical GSH content by pharmacologically increasing chloride secretion to increase GSH secretion. They used a synthetic peptide which could induce transepithelial chloride and fluid transport in a number of cell types including CF airway epithelial cells (IB3-1). They used this synthetic chloride channel-forming peptide ($N\text{-}K_4\text{-}M_2GlyR$) to bypass CFTR-dependent chloride transport. They illustrated that GSH efflux was dependent on chloride permeability and not CFTR protein per se. The synthetic peptide formed another chloride channel on the apical surface of airway epithelial cells and allowed for GSH transport. They also increased potassium movement through the basolateral membrane with the use of chlorzoxazone (a basolateral potassium channel activator) which had two effects which ultimately led to more GSH secretion to the apical surface; a) increased potassium recycling through basolateral potassium channels (i.e., K^+ channel and $Na^+/K^+/Cl^-$ co-transporter) and thus increased apical chloride secretion, b) chlorzoxazone hyperpolarized the cells allowing increased potassium leakage to the basolateral fluid (52).

The presence of residual levels of GSH in CF shows that there are some cells in the CF patient that are capable of GSH efflux (43). To be capable of GSH secretion these cells would therefore possess alternate channels redundant to the CFTR channel and CFTR co-stimulation. One likely candidate for this function is the ABC transporter family protein, multidrug resistance protein (MRP), and cells that normally express MRP are cells involved in transport and/or detoxification and include hepatic cells and cells derived from hemocytoblasts such as red blood cells, neutrophils, monocytes and macrophages, and lymphocytes (T and B cells). The low levels of GSH reported by Lands and associates (43) suggest that cells capable of GSH transport through alternate channels are GSH depleted in CF. Neutrophils produce superoxide and hydrogen peroxide. It seems, however that the neutrophils even in the excess numbers seen in CF airways are inefficient at fighting infections. The stepwise killing process of pathogen for neutrophils involves pathogen internalization (phagocytosis) and then fusion of the phagosome with intracellular granules where the pathogen is killed via oxygen dependent and independent pathways. The intracellular granules contain preformed antimicrobials in concentrated form and enzymes

which convert oxygen into ROS. Depleted intracellular GSH decreases the effectiveness of these cells in superoxide production (44).

Chronic and severe GSH deficiency reduces the effectiveness of the antiprotease system. CF is characterized by an imbalance between α_1-AT and neutrophil elastase (NE) in the airways of patients with CF. Low levels of extracellular GSH do not prevent small water soluble oxidants from inactivating the major antiproteases of the lower and upper respiratory tract i.e. α_1-AT and secretory leukoprotease inhibitor (SLPI), respectively. Therefore an imbalance with unbound neutrophil elastase is created (25, 53) contributing to injury of the connective tissue of CF airways. Unbound NE has been shown to cleave lung elastin, fibronectin, immunoglobulins, complement receptors on neutrophils and receptors on T and B-cells (54). High levels of NE have also been associated with inhibition of ciliary beating, stimulation of mucus secretion from goblet cells and with facilitation of adhesion of *P. aeruginosa* to airway cells (54).GSH is also implicated in lung surfactant maintenance; a) oxidant damage may degrade surfactant (55), b) GSH deficiency is associated with failure to maintain adequate levels of phosphatidylcholine (56); consequently pulmonary surfactant levels may be insufficient and therefore contribute to pathogen invasion.

Kokura and colleagues have shown neutrophil-endothelial cell adhesion is adversely affected by GSH depletion (57). The reduced extracellular GSH also results in a negative feedback for resynthesis of oxidized GSH (GSSG) to GSH by glutathione reductase and reduced nicotinamide adenine dinucleotide phosphate (NADPH). Instead GSSG is further oxidized to chloramines which are oxidants with a long half life. Studies have shown diminished reduction of GSSG in a GSH deficient environment (33, 34) and elevated levels of glutathione reductase implying low use of this enzyme in CF patients (58). In this case there is likely further oxidation of GSSG (59, 60) and this shift to oxidation has been shown to decrease glutamate cysteine ligase activity (61). Alterations in cell structure which include injury to epithelial and endothelial cells have been shown to result from GSH depletion (62). These changes compromise the integrity of the lung tissue and facilitate adhesion of pathogens (63). GSH deficiency also directly stimulates mucin hypersecretion (31, 64) and contributes to the increased viscosity of mucus (35, 65). Extracellular GSH is used to breakdown mucus; GSH facilitates the cleavage of disulfide bonds in mucus and

makes mucus "thinner" (i.e., less viscous). Additional consequences of GSH deficiency include: a) increased lipid peroxidation, and b) potential effects on the level and diffusion of nitric oxide species (65).

Given the variable clinical course in CF patients with the same CFTR genotype and the role of GSH in antioxidant defenses, differences in the ability to synthesize GSH may help explain some of the heterogeneity observed in pulmonary disease course. Glutathione is likely an important factor in the modulation of oxidant-induced lung injury in CF. Genetic determinants of glutathione synthesis and function therefore may modify CF lung injury. Genes that can influence glutathione synthesis and function can therefore be potential modifier genes in CF lung and liver disease.

5.1.3 Modifier genes to explain the heterogeneity in pulmonary and liver disease severity in CF

5.1.3.1 Pathway for glutathione synthesis and metabolism and potential modifier genes for pulmonary disease severity in CF

Figure 5.1 shows the pathway of glutathione synthesis and metabolism. GSH is synthesized in the cytoplasm and this synthesis is catalyzed by two enzymes which utilize ATP. The rate limiting step is glutamate cysteine ligase which catalyzes the ligation of glutamic acid to cysteine. The availability of cysteine is the rate limiting factor in the reaction. Cysteine is generated from the essential amino acid methionine. An essential amino acid is defined as one which cannot be synthesized by the body but is required by the body and needs to be obtained from foods or from the turnover of endogenous protein. When cells are confronted with GSH depletion or oxidative stress, glutamate-cysteine ligase catalytic subunit (GCLC) gene expression is upregulated (11, 66, 67). In mammalian species the enzyme glutamate cysteine ligase is composed of two catalytic subunits. The larger catalytic subunit (73 kDa) is encoded by the GCLC subunit on chromosome 6p12 (68). The smaller modifier subunit (31 kDa) is encoded by the gene glutamate cysteine ligase regulatory subunit (GCLR) on chromosome 1p22 and is thought to have a regulatory function on the kinetic efficiency of GCLC (69). GCLC can carry out all the catalytic activity required and is feedback inhibited by GSH, thereby ensuring homeostatic control over GSH synthesis (11, 66, 67). Upstream of the GCLC translation start codon (10 base pairs from the 5' start codon) a GAC

trinucleotide repeat has been reported (70). Common repeats shown in the Caucasian population are 7, 8, and 9 GAC repeats (71). Variation in the size of the repeat is hypothesized to influence transcription rate, transcription efficiency, possibly RNA stability and may also potentially result in nucleotide differences in the regulatory sequences or within the coding region (68). Walsh and co-workers have shown a positive association between the number of GAC repeats and GSH levels (68). McKone and associates(72) investigated the association of GCLC GAC repeat polymorphisms in a CF cohort of patients recruited from the Seattle adult CF clinic and also used our Vancouver adult and children's CF clinic study cohort to increase their sample size. They used similar grouping, as in our study, for GAC repeat genotype but varied in their grouping for CFTR genotype from grouping used in the present study. Using cross-sectional data they showed a trend between GAC repeat genotype and %predFEV_1 ($p=0.10$), and a significant association ($p=0.001$) within the CF patients with a mild CFTR genotype (i.e., homozygous or heterozygous for at least one class 4, or 5 mutation). Their mild CFTR genotype group also included CF patients with mutations classified as unknown or unclassified (72). Additionally, two single nucleotide polymorphisms in the GCLC gene have been recently identified and shown to result in GSH deficiency (73, 74).

The second reaction leading to reduced glutathione synthesis involves the combining of γ-glutamylcysteine with glycine to generate GSH and this reaction is catalyzed by glutathione synthase (Figure 5.1). The enzyme glutathione synthase does not have a regulatory role (11). One of the applications of GSH is for free radical scavenging. Therefore, other possible candidates in the GSH pathway are the glutathione S-transferase (GST) isoenzymes. GST conjugates reactive compounds to GSH prior to excretion from the body. GST also acts as a glutathione peroxidase and converts lipid peroxides to a hydroxyl form and oxidizes GSH to GSSG. GSSG can be recycled back to GSH again by the enzyme glutathione reductase, which uses coenzyme NADPH as a source of electrons. Oxidant stress causes alterations in redox balance by affecting changes in the GSH:GSSG ratio which in turn influence the expression of transcription factors including NFκB.

There are two supergene families that encode proteins with glutathione S-transferase activity. There are 16 genes that encode proteins expressed in tissue cytosol and six genes

that encode proteins expressed in cell membranes. There are eight distinct gene families that encode cytosolic soluble GST enzymes:

- Alpha on chromosome 6p
- Mu on chromosome 1p (GSTM)
- Theta on chromosome 22q (GSTT)
- Pi on chromosome 11q (GSTP)
- Zeta on chromosome 14q
- Sigma on chromosome 4q
- Kappa (expressed in mitochondria, unknown chromosomal location)
- Chi or omega on chromosome 10q

GSTM1 and GSTP1 are expressed in airway epithelial cells. Polymorphisms have been described for the mu, pi and theta class clusters. There are 5 mu genes situated in tandem (telomere-GSTM4-GSTM2-GSTM1-GSTM5-GSTM3-centromere) and spaced 20 kb apart on chromosome 1p13.3 (75). GSTM enzymes detoxify alpha and beta unsaturated carbonyls which are products of cigarette smoke, lipid peroxidation, damaged DNA and quinine-containing oxidation products of catecholamines (76-79). Individuals homozygous for the deletion of GSTM1 express no protein. Homozygosity for the null GSTM1 genotype is common, with reported frequency of 40-60% in Europeans, Japanese and white Americans (and 25% in black Americans) (80) and has been associated with increased risk for some cancers (81, 82), pulmonary emphysema in combination with lung cancer (83), severe chronic bronchitis in heavy smokers (84) and lung cancer in heavy smokers (85, 86). Deletion of GSTM1 has been investigated in a few studies in small CF cohorts.

Baravov and co-workers showed that homozygosity for the GSTM1 deletion was associated with earlier CF diagnosis and death prior to the age of 5 years (87). Hull and Thomson (88) investigated homozygosity for the GSTM1 deletion in 53 CF children. Prevalence for homozygosity of the GSTM1 deletion genotype in their cohort was 49%. In their study they showed worse chest radiographic scores and S-K clinical scores in CF patients who were homozygous for the GSTM1 deletion polymorphism. There sample was homogeneous for CFTR genotype severity with all their patients homozygous for severe CFTR mutations.

GSTP1 is located on chromosome 11q13. The GSTP1 enzyme is involved in detoxification of carcinogens and mutagenic compounds; GSTP1 metabolizes, low molecular weight

halogenated compounds and reactive epoxides (thymine hydroperoxide, arachidonic acid hydroperoxide). The gene is composed of 7 exons and in exon 5 there is a polymorphism at codon 105 which results in an amino acid change from isoleucine (ATC) to valine (GTC). The consequence of this amino acid substitution is a change in catalytic activity. The prevalence of the valine polymorphism is 31% (89). A seven fold higher efficiency for diol epoxides of polycyclic aromatic hydrocarbons has been shown for GSTP1 enzymes containing valine at codon 105 rather than isoleucine (90). The GSTP1 (Ile105Val) polymorphism has been associated with bladder and testicular cancer (91) and lung cancer in smokers (92). He and co-workers investigated the GSTP1 Ile105Val polymorphism and the null mutations for GSTM1 and T1 in COPD patients who smoked and were part of the Lung Health Study and they showed a more rapid decline in %predFEV_1 in patients who had the cumulative effect of the presence of the null polymorphisms in GSTM1 and GSTT1 and the Ile/Ile genotype for GSTP1 at codon 105 (93). Others have shown homozygosity for valine for GSTP1 at codon 105 to be associated with reduced airway hyperresponsiveness and atopy in adults (94, 95) and children (96).

There are 2 GSTT genes that have similar structures with identical intron and exon boundaries and are separated by 50 kb (97). The chromosomal location of GSTT1 is 22q11.2. GSTT1 is expressed in the liver and red blood cells. The frequency of the GSTT1 gene deletion is 38% and the prevalence of homozygosity for the deletion is 14% in Caucasians (98). Individuals homozygous for the deletion express no protein (98). Individuals who have a null genotype have been shown to have reduced ability to conjugate GSH with environmental carcinogens and toxins (dibromoethane, dichloromethane, ethylene oxide, methylene chloride and methyl bromide) (98-100). Decreased metabolism of these compounds in human blood has been attributed to the GSTT1 gene deletion (98, 101, 102). Wiebel and associates showed that having two functional GSTT1 alleles was associated with double enzyme expression (103). Deletion of GSTT1 has been previously investigated and has been associated with more rapid decline in pulmonary function in smoking COPD patients (93), susceptibility to cancer (104), and worse prognosis from cancer (104, 105).

5.1.3.2 Liver disease in CF and modifier genes

The liver is a major site of GSH synthesis and GSH export and the hypothesis postulated is that alterations in GSH transport contribute to CF liver disease. GSTT1 and GSTM1 are

expressed in the liver and were investigated in this study for their role in contributing to CF liver disease. Studies have shown an association between GSH depletion in liver cells and cirrhosis of the liver and end-stage liver disease (45, 46, 106, 107). GSTM1 and GSTP1 are expressed in biliary epithelial cells and hepatic stellate cells and GSTM1 is also expressed in hepatocytes (108, 109). CFTR is expressed in biliary epithelial cells and contributes to ductal bile secretion and in CF the CFTR defect has been shown to impair ductal bile secretion (110). The prevalence of liver disease in CF is 1% and it is the second leading cause of death in CF patients after pulmonary disease and accounts for 2% of overall mortality (111). The GST genes are potential candidate modifiers of CF liver disease as low levels or low expression of different isoforms of these enzymes would affect (reduce) conjugation of various electrophiles with GSH (112). The GSTM class accounts for more than 50% of the cytosolic GST detoxifying activity (113). Henrion-Caude and associates (112) genotyped 106 Caucasian CF patients attending 2 CF clinics in France for the GSTM1 deletion polymorphism and the GSTP1 (Ile105Val) polymorphism. They showed an association with liver disease and the GSTP1 Ile/Ile genotype and this relationship persisted in CF patients who were homozygous or heterozygous for the CFTR delta F508 mutation. Henrion-Caude and associates postulated that GSTP1 Ile/Ile genotype may be a modifier gene in CF-related bile duct-type liver disease, as GSTP1 is expressed at high levels in biliary epithelial cells (108). They also posed an alternative hypothesis that GSTP1 is a modifier gene of liver fibrosis.

5.1.3.3 Summary of literature findings

The findings to date show that the CFTR is a potential channel which is used to export GSH extracellularly. The defective CFTR channel in CF would explain the low extracellular levels of GSH and normal intracellular levels of GSH. In CF pulmonary disease, CF patients who have some CFTR function, that is carry at least one mild CFTR mutation belonging to class IV or V would allow some chloride conductance and would also have the ability to transport GSH out of the cell. Whether partially functioning CFTR would be adequate for GSH transport is unclear. CF patients who have two severe CFTR mutations (i.e., class I, II or III) where there is no chloride conductance will also show no GSH transport through the mutant CFTR. We thus hypothesize that patients with some CFTR function and who do not have the deletion polymorphism for GSTM1 and GSTT1 will have better pulmonary function and a lower rate of decline in pulmonary function over time compared with CF

patients who have the deletion polymorphisms for GSTM1 and GSTT1. We also hypothesize those CF patients who are homozygous or heterozygous for isoleucine for the GSTP1 (Ile105Val) polymorphism will show increased pulmonary disease severity and progression. GSTT1, GSTM1 and GSTP1 polymorphisms resulting in decreased levels of these enzymes, regardless of CFTR genotype will be associated with earlier chronic colonization with respiratory pathogens, more severe pulmonary disease severity and progression. CF patients who have a lower number of GCLC GAC repeats will show worse pulmonary disease severity and progression. The GSTT1 and GSTM1 deletion polymorphisms and GSTP1 (Ile105Val) polymorphism will be associated with susceptibility to CF liver disease.

5.2 RESULTS

5.2.1 Hardy Weinberg equilibrium

Table 5.1 shows Hardy Weinberg equilibrium (HWE) results and the frequencies of the GSTP1 (Ile105Val) polymorphism on the cross-sectional and the longitudinal study cohort. Table 5.2 shows HWE results and the frequencies of the GCLC gene GAC repeats on the cross-sectional and longitudinal study cohort. All polymorphisms were in Hardy Weinberg equilibrium.

5.2.2 Descriptive data results and study cohort grouping

The clinical characteristics of the study cohort studied are presented in Tables 5.3 -5.5. Data are presented for the entire cohort used in the cross sectional analysis and for the cohort used for mixed effects regression modeling for the longitudinal data. Table 5.3 presents the clinical characteristics of the study cohort stratified by the GSTM1 and T1 polymorphisms, and Table 5.4 presents these characteristics for the group stratified by the GCLC polymorphism. Table 5.5 presents the clinical characteristics for the group stratified by the GSTP1 polymorphism.

5.2.3 Pulmonary disease progression: mixed effects regression on %predFEV$_1$

In our linear mixed effects models we first investigated whether having the gene deletion or not for GSTM1 and T1 and in a separate analysis whether having one or two G-alleles for the GSTP1 (Ile105Val) polymorphism were associated with different rates of decline in %predFEV$_1$. For the GCLC gene we elucidated the number of GAC repeats in our study population and then investigated the commonly occurring ones. We were unable to investigate *P. aeruginosa* infection status and the GST and GCLC polymorphisms by CFTR genotype severity in all cases due to the small number of patients who were carriers of mild CFTR mutations and who were also not infected chronically with *P. aeruginosa*, which caused estimation problems with the models. We did not show significant associations for having one (p=0.40) or two gene deletions (p=0.71) for GSTMI and T1 and pulmonary disease severity (p=0.40 for one deletion and p=0.71 and for two deletions) and pulmonary disease progression (p=0.41 for one deletion and p=0.95 and for two deletions).

Model 5.2.3-A

*%predFEV$_1$ = Time + Sex (0/1=male/female) + 1 null in GSTM1 or T1 (0/1=0 or 2 null copies/1 null copy) + 2 null in GSTM1 and T1 (0/1=0 or 1 null copies/2 null copies) + 1 null in GSTM1 or T1 * Time + 2 null in GSTM1 or T1 * Time*
Base group for GSTM1 and T1 was: zero null in GSTM1 or T1

We also investigated the GSTP1 (Ile105Val) polymorphism and showed no differences in pulmonary disease severity (p=0.45 for AG and p=0.92 for GG versus AA) and pulmonary disease progression (p=0.60 for AG and p=0.65 for GG versus AA) with GSTP1 genotype.

Model 5.2.3-B

*%predFEV$_1$ = Time + Sex (0/1=male/female) + GG genotype for GSTP1 (0/1=0 or 1 copy of G-allele/ GG) + AG genotype for GSTP1 (0/1=0 or 2 copies of G-allele/ AG) + GG genotype for GSTP1 * Time + AG genotype for GSTP1 * Time*
Base group was AA for GSTP1).

We showed the presence of uncommon GAC repeat alleles for GCLC in our study cohort, however, we were unable to investigate these uncommon polymorphisms in our models below due to limited patient numbers. We investigated GAC repeats for GCLC which were represented in our cohort and clinical data for these groupings are presented in Table 5.4. CFTR genotype was used as a covariate in our study. CFTR genotype for CF patients in our cohort who were genotyped for CFTR was arranged into two groups for our models in this section; a severe (homozygous for class 1, 2, 3 CFTR mutation, or class 1, 2, 3 and unknown/unclassified CFTR mutation) and mild (one or two chromosomes with a class 4, or 5 mutation) CFTR severity genotype group. For pulmonary disease progression our grouping for GCLC was:

- Homozygosity for 7 GAC repeats (i.e., GCLC7/7),
- Heterozygosity for 7 GAC repeats: 7/8 GAC repeats (i.e., GCLC7/8),
- Heterozygosity for 7 GAC repeats: 7/9 GAC repeats (i.e., GCLC7/9), and
- Greater than 7 GAC repeats on both chromosomes (i.e., GCLCgr7; includes 8/8, 8/9 and 9/9 GAC repeats).

All other genotypes were rare and were not used in our analyses (i.e., 6/9 and 7/10 GAC repeats).

We showed a significant main effect (i.e., pulmonary disease severity, p=0.014) for having the genotype 7/9 versus 7/7 GAC repeats for GCLC, but no difference in pulmonary disease progression for the comparison (i.e., the slope GCLC 7/9 * Time). However, comparisons of GCLC 7/9 to 7/8 and to GCLCgr7 of the mean levels (comparing the main effects) and their slopes (comparing the interaction terms gene polymorphism * Time) gave insignificant results for GCLC7/9 versus GCLC7/8 and GCLC7/9 versus GCLCgr7 with respect to pulmonary disease severity and progression.

Model 5.2.3-C

*%predFEV$_1$ = Time + Sex (0/1=male/female) + GCLC7/8 + GCLC7/9 + GCLCgr7 (0/1 1=genotypes with >7GAC repeats on both chromosomes) + GCLC7/8 * Time + GCLC7/9 * Time + GCLCgr7 * Time*

Base group for GCLC was 7/7 GAC repeats

5.2.4 Pulmonary disease severity: mixed effects regression on current %predFEV$_1$

In this section we used our cross-sectional data set to enlarge our sample size to up to 434, which included subjects from Canadian clinics included in this study but for whom we did not have longitudinal data available as well as the CF cohort from one adult clinic from Seattle Washington, U.S.A., which is the U.S. cohort described in McKone et al(72). In our model we included CFTR class genotype, nutritional status (BMI), current age, *P. aeruginosa* infection status and sex. The cohort was grouped into two categories for CFTR grouping as described in section 5.2.3,

Model 5.2.4-A

Current %predFEV$_1$ = Current age (yrs) + Sex (0/1=male/female) + CFTR genotype (0/1=severe/mild) + BMI + PA infection status (0/1=not infected /chronically infected) + 1 null in GSTM1 or T1 (0/1=0 or 2 null copies/1 null copy) + 2 null in GSTM1 and T1 (0/1=0 or 1 null copies/2 null copies).

Base group was zero null in GSTM1 or T1

Model 5.2.4-B

Current %predFEV$_1$ = Current age (yrs) + Sex (0/1=male/female) +CFTR genotype (0/1=severe/mild) + BMI + PA infection status (0/1=not infected /chronically infected) + GG genotype for GSTP1 (0/1=0 or 1 copy of G-allele/ GG) + AG genotype for GSTP1 (0/1=0 or 2 copies of G-allele/ AG).

Base group was AA for GSTP1.

Model 5.2.4-C

Current %predFEV$_1$ = Current age (yrs) + Sex (0/1=male/female) +CFTR genotype (0/1=severe/mild) + BMI + PA infection status (0/1=not infected /chronically infected) + GCLC7/8 + GCLC7/9 + GCLCgr7 (0/1 1=genotypes with >7GAC repeats on both chromosomes).

Base group for GCLC was 7/7 GAC repeats

Overall we did not show significant associations between our candidate gene polymorphisms and subjects' most current collected %predFEV$_1$, characterizing their current pulmonary disease severity status, although other cross-sectional variables were associated with %predFEV$_1$ (i.e., current age, BMI, *P. aeruginosa* infection status and sex). We showed a p-value of 0.08 for having the genotype 7/9 versus 7/7 GAC repeats for the GCLC gene polymorphism and pulmonary disease severity.

5.2.5 Survival analysis

Our survival models were run using CFTR class grouping and alternatively using PSS (or the continuous variable sweat chloride levels) as not all CF genotype classes were found in each candidate modifier gene category when including other covariates. The number of CF patients with a mild CFTR genotype (i.e., homozygosity or heterozygosity for class 4 or 5 mutation) was small and therefore we ran our survival models using the alternate predictors when necessary. Our results are presented in Table 5.6. We showed that being null for both genes (i.e., GSTM1 and T1) was associated with a more rapid progression to our endpoint (time to death or lung transplantation) compared with CF patients who have neither gene deletion. However the RR and confidence intervals are very large and our result is suspect

and is likely due too few subjects in the cell with an event for this genotype. For the same reason the interaction with chronic infection with *P. aeruginosa* is suspect and will both be considered as inconclusive for the study. We showed that being diagnosed with CF at an earlier age and having either GSTM1 or T1 gene deletion was associated with a more rapid progression. We showed that being diagnosed with CF at an earlier age and having either GSTM1 or T1 gene deletion were associated with a more rapid progression to our endpoint compared with CF patients who have neither gene deletion and who were diagnosed with CF at a similar age. Our analysis did not show differential survival for GSTP1 genotype. We showed a marginal significant association (p=0.08) for the GCLC gene; a more rapid progression to our endpoint (time to death or lung transplantation) in CF patients who had a GCLC genotype of 7/8 GAC repeats compared with our base group 7/7 GAC repeats (Table 5.6).

5.2.6 Effect of modifier genes on *P. aeruginosa* infection status

We investigated whether polymorphisms in the GST genes and the GCLC gene contributed to infection with *P. aeruginosa*. We separately investigated age of first infection and age of chronic infection with the pathogen and results are presented in Tables 5.7 and 5.8, respectively. CFTR class was included in the GCLC models and GSTM1 and T1 deletion polymorphisms. For the GSTM1 and T1 deletion polymorphisms, we showed that fixing %predFEV$_1$ (i.e., two CF patients having the same %predFEV$_1$) those patients with GSTM1 and T1 gene deletion are more likely to become chronically infected with *P. aeruginosa* at an earlier age than those CF patients who have neither a M1 or T1 gene deletion, but there were no differences for age of first infection with *P. aeruginosa*. We did not show significant associations between age of first infection and chronic infection *P. aeruginosa* and candidate modifier genes; GCLC and GSTP1.

5.2.7 Liver Disease and CF

Simple logistic regression was used to investigate the association of liver disease and GSTM1 and GSTP1 genotypes, while controlling for known confounders (chronic infection with *P. aeruginosa* or BCC, CFTR severity genotype and current age). We show that deletion of the GSTM1 gene is a modifier gene for CF liver disease (p=0.05). Homozygosity or heterozygosity for the GSTP1 (Ile105Val) Ile allele was not associated with liver disease in CF (p=0.55-0.75). The variability explained by this model is however small (adjusted

R^2=0.06, p=0.05). The combination GST M1 and P1 analysis did not show a significant association for liver disease (p=0.32-0.59). Liver disease measures are presented for the CF cohort in Table 5.9.

Liver disease (present/not present) = Sex (0/1=male/female) + CFTR genotype + BMI + GG genotype for GSTP1 (0/1=0 or 1 copy of G-allele/ GG) + AG genotype for GSTP1 (0/1=0 or 2 copies of G-allele/ AG)+ GSTM1 (0/1=0 or null copy) + GSTM1 GSTP1.*
Base group was AA for GSTP1 and homozygous for class 1 mutations for CFTR genotype.

5.3 DISCUSSION

In this study we investigated gene polymorphisms which would affect glutathione production and cycling of reduced glutathione (GSH). We showed more severe pulmonary disease in patients with 7/8 versus 7/7 GAC repeats in our cross sectional study cohort, but no association of the GCLC gene GAC repeat with pulmonary disease progression. GST transferases were investigated, specifically GSTM1 and GSTT1 gene deletions were investigated as potential modifiers of pulmonary disease and we showed no association with GSTM1 and GSTT1 gene deletions and pulmonary disease severity and pulmonary disease progression. Lastly, the GSTP1 (Ile105Val) polymorphism was investigated and we showed no association with pulmonary disease severity and pulmonary disease progression. We also examined age of CF diagnosis, *P. aeruginosa* infection status and age of first and chronic infection with this pathogen and showed no significant differences with the candidate gene polymorphisms. We lastly showed that deletion of the GSTM1 gene was a significant modifier gene for CF liver disease, although the variability explained by this association was small. In this study we did not show an additional contribution of the GSTP1 gene polymorphism as a significant modifier gene for CF liver disease.

Earlier research have shown that homozygosity for the GSTM1 deletion in CF patients is associated with earlier age of CF diagnosis and death prior to the age of 5 years (87), worse chest radiographic scores and S-K clinical scores (88). Flamant and associates (114) showed lung function at age nine years was better in CF children who carried the B-allele for GSTM3, but found no associations of lung function and the deletion polymorphisms for GSTM1 and T1 or the GSTP1 polymorphism. We investigated gene deletions of GSTM1 and T1 and did not show associations with pulmonary disease severity and progression. We controlled for CFTR genotype, sex and age in both our cross-sectional and longitudinal models. In our study cohort we categorized CF patients based on their CFTR genotype into severe and mild CFTR genotype based on CFTR class. We excluded patients who had two unknown or unclassified CFTR mutations, as the number of patients was small to consider as a separate group and we did not want to dilute the other CFTR severity genotype groups. Hull and Thomson (88) reported that their study cohort consisted of CF patients who carried two severe CFTR mutations. The sample sizes of these two previous studies were small and the associations reported may be due to type one error.

We would expect that CF patients with mild CFTR genotype may show variable disease course based on GSTM1 and GSTT1 genotype. We hypothesized that those CF patients with mild CFTR genotype who were positive for the GSTM1 and GSTT1 deletions would have more severe pulmonary disease and pulmonary disease progression. The number of patients with mild CFTR genotype was rather small and it may be that an association exists but due to our small sample size for this group we were unable to observe a significant association. It is also equally likely that our larger overall sample size prevented us from observing false positive associations, which may have been the case in previous studies (87, 88).

Imboden and associates showed a significant decline in %predFEV_1 over the 11 year study interval investigated for GSTT1 gene deletion alone or in combination with GSTM1 gene deletion in male smokers and non-smokers (89). He and associates (93) showed a positive association for combined gene deletions of GSTM1 and GSTT1 and the GSTP1 polymorphism genotype Ile/Ile at codon 105 in COPD patients who smoked and were categorized by a rapid decline in pulmonary function. The GSTP1 polymorphism has also been associated with lung cancer in smokers (92). We did not show a positive association with the GSTP1 polymorphism and pulmonary disease severity and pulmonary disease progression in our study cohort. We controlled for CFTR class, age and sex. As with the GSTM1 and GSTT1 models we were unable to include *P. aeruginosa* infection status as an interaction term as there were estimation problems with the model that prevented analyses including this variable. Chronic infection with *P. aeruginosa* is associated with more severe pulmonary disease in CF (115-117).

The GCLC gene is characterized by a region of GAC repeats. In our study cohort the frequency of the most common alleles were similar to frequencies reported for non-CF populations. As our study cohort was large compared with other studies we also found more rare alleles (i.e., low or high GAC repeats: 6 and 10 repeats). A positive association has been shown between the number of GAC repeats and GSH levels (68). Variation in the size of the repeat is hypothesized to influence transcription rate, transcription efficiency, possibly RNA stability and may also potentially be associated with nucleotide differences in the regulatory sequences or within the coding region (68). Walsh and co-workers and others have shown GCLC gene expression to be up regulated when GSH levels are low and during oxidative

stress (11, 66-68). Low GSH levels have been shown in CF (9, 10, 51). We therefore hypothesized that CF patients who had a larger number of GAC repeats in the GCLC gene would be protected. This may still be the case, however, the majority of our study cohort was characterized by 7 and 8 GAC repeats and these 2 groups did not show differences in pulmonary disease severity and progression compared with CF patients who were characterized by at least one 9 or 10 GAC repeats. We showed a significant main effect for %predFEV$_1$ (i.e., pulmonary disease severity for GCLC 7/9 versus GCLC 7/7 genotype, but no difference in pulmonary disease progression (i.e., the slope GCLC 7/9 * Time compared to the base group GCLC 7/7 genotype). Further comparisons of mean levels of %predFEV$_1$ and their slopes showed similar slopes for GCLC7/9 versus GCLC7/8 and GCLC7/9 versus GCLCgr7 with respect to pulmonary disease severity and progression. It was unclear whether this polymorphism would be important in CF patients with severe CFTR genotype, as most studies have shown depleted GSH levels not only in serum and extracellular levels, but also in immune cells (neutrophils and macrophages) which use an alternate channel for GSH extrusion. We expected to see a significant difference in CF patients with a mild CFTR genotype, where there is some functional CFTR protein and therefore ability of GSH to traffic into the extracellular space. However our sample size for CF patients with mild CFTR as defined in our study was a limitation. In a different analysis McKone and associates showed a significant association between GAC repeats and cross-sectional %predFEV$_1$ in the subgroup of CF patients with mild CFTR mutations (72). Our results incorporating their study cohort into our study group and looking at the cohort's current %predFEV$_1$ using a different statistical technique (mixed effects regression) showed a similar trend to our longitudinal results and the results of McKone and associates(72) but with a more significant p-value of 0.08. The other main difference for cross-sectional analysis between this study and the study by McKone and associates (72) was CFTR grouping; our grouping for mild CFTR excluded CF patients with unclassified or unknown mutations. Although immune cells have an alternate channel which is used for GSH trafficking, these cells have been shown to be depleted as well in CF. It is unclear at this time whether CF patients with mild CFTR genotype may have intermediate levels of extracellular GSH.

Liver disease is the second leading cause of death in CF patients. CFTR genotype does not correlate with liver disease in CF. Since CFTR modulates GSH transport and in CF CFTR dysfunction creates an antioxidant imbalance, liver detoxifying enzymes such as GSTP1 and

GSTM1, both expressed by biliary epithelium, may be potential modifiers of CF liver disease. CFTR is expressed in the biliary epithelium and it is assumed that in CF there is an impairment of ductal bile secretion leading to obstruction of the bile ducts ultimately giving rise to cirrhosis and end-stage liver disease. Henrion-Caude showed an association between CF liver disease and homozygosity for the Ile allele for GSTP1 (Ile105Val) (112). They also investigated the gene deletion polymorphism and another polymorphism on exon 7 for GSTM1, but these two polymorphisms were not predictors of CF liver disease.

In our study we showed a significant association between liver disease and the GSTM1 gene deletion and no significant association with the GSTP1 (Ile105Val) polymorphism. Liver biopsy together with clinical examination, abdominal ultrasonography (hepatomegaly, liver hyperechogenicity, nodularity of the liver edge and signs of portal hypertension) and liver enzyme levels were used to diagnose liver disease for 50% of the study cohort, and clinical examination and liver enzymes levels used to diagnose liver disease for the remainder of the study cohort. Although serum levels of alanine transaminase, aspartate transaminase, alkaline phosphates and gamma-glutamyltransferase were measured on our patient cohort, aspartate transaminase and gamma-glutamyltransferase levels were most predictive of liver disease. Our findings do not concur with the findings of Henrion-Caude and co-workers. Our results are based on a much larger study cohort (N=285), and CF patients with unknown CFTR genotype were excluded from the study cohort analyzed. The study by Henrion-Caude consists of 106 children (age range 4-24 years). The genotyping methods we used did not allow us to determine whether the patients were homozygous or heterozygous for the GSTM1 gene deletion, whereas in the study by Henrion-Caude and associates they distinguished all genotypes. Methods to diagnose liver disease were similar in both studies, other than for liver biopsies performed in a section of our study cohort as an additional parameter for diagnosis of liver disease. Current findings suggest that glutathione S-transferases are candidate modifier genes which explain some of the risk for developing liver disease in CF.

5.4 CONCLUSIONS

We conclude that polymorphisms which affect levels of glutathione produced, released and utilized are potential modifier genes in CF pulmonary disease. Our results show that the deletion polymorphisms (GSTM1 and T1) contribute to adverse outcomes (death/lung transplantation). For the GCLC gene GAC repeat polymorphisms we showed a significant difference in pulmonary disease severity and survival in CF patients with a 7/8 GAC repeat genotype versus 7/7 GAC repeat genotype, but no difference in pulmonary disease progression. The GCLC GAC repeat polymorphism needs to be further investigated in an independent CF cohort of larger sample size and more CF patients with mild CFTR genotype. The deletion of the GSTM1 gene was shown to be a modifier gene for liver disease in Cystic Fibrosis. It is probable that the oxidant burden in CF overwhelms the ability of the immune system to keep up with production of GSH. Polymorphisms in genes involved in glutathione production (GCLC) and in antioxidant response (GSTs) have an effect on pulmonary disease in CF. Polymorphisms leading to reduced levels of GSTs (deletion polymorphism in GSTM1 gene) may contribute to CF liver disease.

Figure 5.1. Pathway of glutathione synthesis and metabolism.

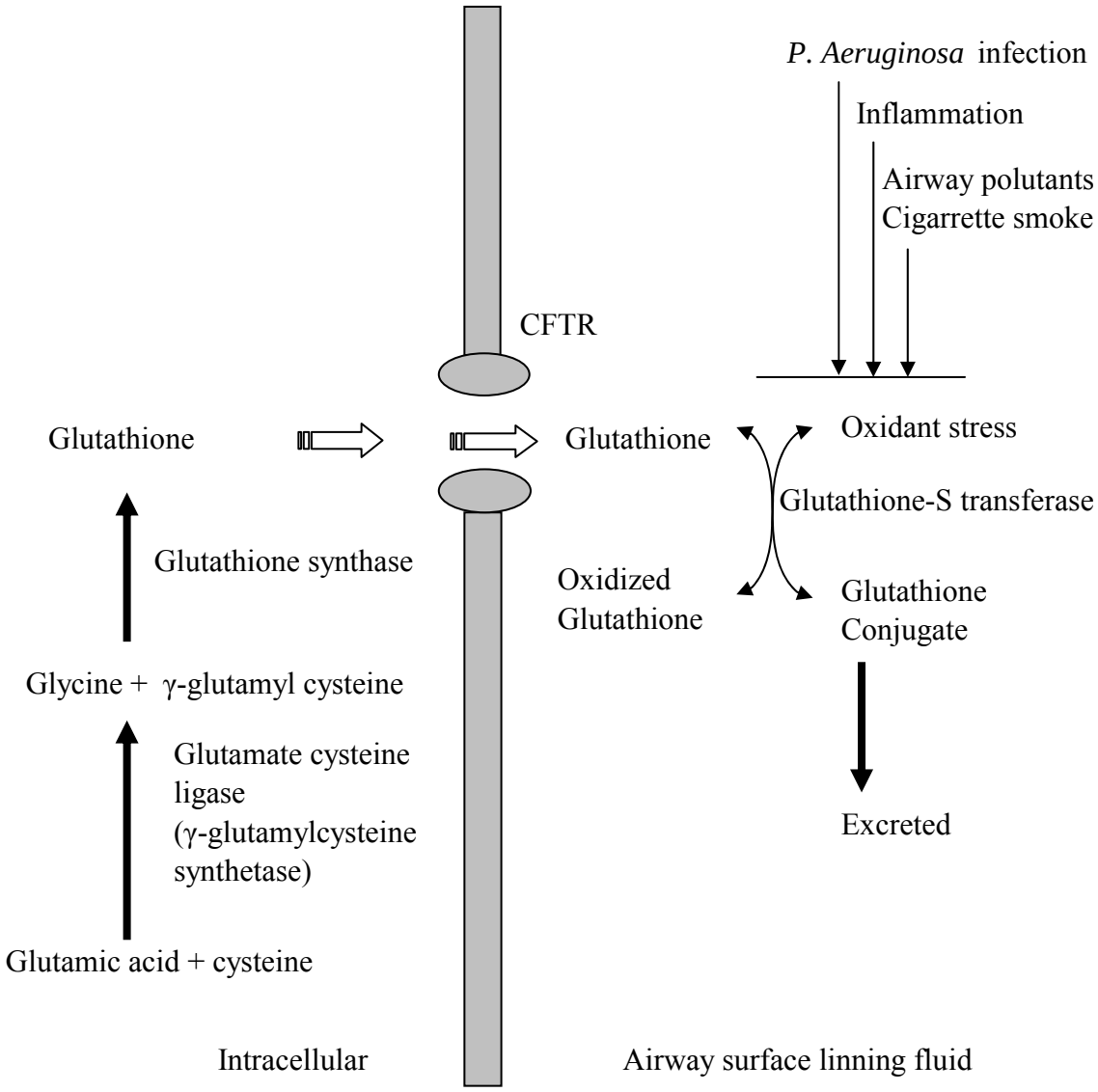


Table 5.1. Frequency of alleles and Hardy Weinberg equilibrium for GSTP1 gene polymorphism in the cross-sectional and longitudinal study cohort.

	AA	AG	GG	A	G	P-value
Cross-sectional						
Observed	233 (47.0%)	208 (42.0%)	54 (11.0%)	674 (68.0%)	316 (32.0%)	0.83
Expected	229 (46.0%)	215 (43.0%)	50 (10.0%)			
Longitudinal						
Observed	191 (48.0%)	167 (42.0%)	40 (10.0%)	549 (69.0%)	247 (31.0%)	0.83
Expected	189 (48.0%)	170 (43.0%)	38 (10.0%)			

Table 5.2. Frequency of alleles and Hardy Weinberg equilibrium for GCLC gene polymorphism in the cross-sectional study and longitudinal study cohort. Genotypes noted in shaded rows were not used in statistical analyses of clinical measures (i.e., pulmonary disease severity and progression, survival, etc) due to low subject numbers.

	Observed	Expected	Repeat	Frequency of repeat	P-value
Cross-sectional					
9/9	29 (5.3%)	5.3%	7	672 (62.0%)	0.67
8/9	38 (7.0%)	7.0%	8	158 (15.0%)	
7/9	154 (28.4%)	28.5%	9	251 (23.0%)	
7/8	92 (17.0%)	18.6%	6	1 (0.09%)	
7/7	212 (39.1%)	38.4%)	1	2 (0.2%)	
8/8	14 (2.6%)	2.2%			
7/1	2 (0.4%)	0.2%			
9/6	1 (0.2%)	0.04%			
Longitudinal					
9/9	26 (5.9%)	5.3%	7	560 (64.0%)	0.43
8/9	30 (6.8%)	6.0%	8	117(13.0%)	
7/9	117 (26.6%)	29.4%	9	200 (23.0%)	
7/8	69 (15.7%)	16.6%	6	1 (0.1%)	
7/7	186 (42.3%)	40.1%	1	2 (0.2%)	
8/8	9 (2.0%)	1.5%			
7/1	2 (0.4%)	0.3%			
9/6	1 (0.2%)	0.05%			

Table 5.3. Clinical characteristics of study cohort by GSTM1 and T1 grouping based on having zero, 1 or 2 null alleles across the 2 gene polymorphisms.

	Total	No null	One null	2 null	p-value
Longitudinal study cohort					
Sample size	398	140(35%)	215(54%)	43(11%)	
Current age (years)	25.6 (0.5)	25.8 (1.0)	25.3 (0.7)	26.1 (1.6)	0.88
Current %predFEV_1 (% of predicted)	61.0(1.4)	59.7(2.4)	61.8(1.9)	61.4(4.1)	0.80
Current BMI	20.3(0.2)	20.4(0.3)	20.3(0.2)	20.2(0.5)	0.93
Age of CF diagnosis (years)	4.7(0.4)	5.1(0.8)	4.5(0.5)	4.4(1.3)	0.80
Sex (male/female)	203/195	70/70	115/100	18/25	0.36
PSS (insufficient/sufficient)	354/44	124/16	194/21	36/7	0.45
CFTR genotype*	272/91/24/11	96/34/5/5	151/44/14/6	25/13/5/0	0.25
P. aeruginosa infection status (not /chronically infected)	117/208	39/78	62/113	16/17	0.27
PA1 age	14.6(0.8) N=223	14.6(1.2) N=90	14.0(1.0) N=113	18.1(2.9) N=20	0.32
PAC age	15.3(0.8) N=152	15.2(1.3) N=62	14.6(1.1) N=78	20.4(3.6) N=12	0.18
Cross-sectional study cohort					
Sample size	498	180(36%)	271(54%)	47(9%)	
Current age	26.9 (0.5)	26.8 (0.8)	27.0 (0.7)	26.8 (1.5)	0.98
Current %predFEV_1	60.4(1.2)	59.4(2.0)	60.9(1.7)	61.4(3.8)	0.83
Sex	254/244	91/89	144/127	19/28	0.27

PSS	442/55	162/18	240/30	40/7	0.63
CFTR genotype*	341/116/29/12	127/41/6/6	186/61/18/6	28/14/5/0	0.36
P. aeruginosa infection status	142/283	45/112	79/152	18/19	0.06

..* CFTR genotype: homozygous for severe mutations (class 1, 2, or 3) / Heterozygous for severe mutation (class 1, 2 or 3) and the other mutation is classified as unknown or unclassified / homozygous or heterozygous for mild mutation class 4, or 5 / homozygous or heterozygous for 2 unknown or unclassified mutations.

** p-value denotes univariate ANOVA or chi square results.

Table 5.4. Clinical characteristics of study cohort by GCLC ligase gene grouping based on number of GAC repeats genotype. In mixed effects regression and survival analyses the groups 8/8, 8/9 and 9/9 are combined into the GCLCgr7 group.

	Total	7/7	7/8	7/9	8/8	8/9	9/9+	P-value**
Longitudinal study cohort+								
Sample size	416	173 (42%)	67 (16%)	110 (26%)	9 (2.0%)	28 (7%)	26 (6%)	
Current age (years)	25.2 (0.5)	25.4 (0.8)	24.3 (1.2)	24.6 (1.1)	25.8 (9.4)	27.3 (2.2)	25.5 (2.1)	0.85
Current %predFEV$_1$ (% of predicted)	62.0 (1.4)	61.2 (2.2)	61.8 (3.5)	65.2 (2.5)	48.5 (7.2)	60.7 (5.6)	56.8 (5.3)	0.18
Sex (male/female)	209/195	90/83	36/31	54/56	5/4	14/14	15/11	0.94
PSS (insufficient/sufficient)	358/46	155/18	58/9	96/14	8/1	24/4	25/1	0.62
CFTR genotype*	272/92/25/15	120/37/11/5	44/14/3/6	71/27/8/4	6/3/0/0	19/6/3/0	18/8/0/0	0.88
P. aeruginosa infection status (not / chronically infected)	128/205	69/77	16/42	32/57	2/5	7/13	4/16	0.08
Cross-sectional study cohort +								
Sample size	516	199	90	146	14	35	29	

		(39%)	(17%)	(28%)	(3.0%)	(7%)	(6%)	
Current age	25.2 (0.5)	26.0 (0.8)	27.3 (1.2)	26.3 (0.9)	28.5 (2.5)	29.2 (2.0)	25.5 (1.9)	0.65
Current %predFEV_1	60.8 (1.4)	60.5 (2.0)	60.7 (3.0)	64.1 (2.1)	57.4 (6.5)	59.8 (4.9)	56.1 (4.7)	0.25
Age of CF diagnosis	26.6 (0.5)	26.0 (0.8)	27.3 (1.2)	26.3 (0.9)	28.5 (2.5)	29.2 (2.0)	25.5 (1.9)	0.65
Sex	257/ 242	105/94	52/38	67/79	8/6	17/18	16/13	0.62
PSS	442/ 56	180/18	79/11	127/19	12/2	28/7	28/1	0.31
CFTR genotype*	338/ 115/ 30/1 6	136/45/ 13/5	63/18/3/ 6	96/36/9/ 5	9/5/0/0	23/7/5/ 0	20/9/0/ 0	0.74
P. aeruginosa infection status	151/ 277	76/96	24/57	39/86	4/8	8/19	4/19	0.07

* and ** are explained in Table 5.3.

[+]Additional genotypes observed in the:

Cross-sectional data set were: 6/9 (N=1), and 7/10 (N=2).

Longitudinal data set: 6/9 (N=1), and 7/10 (N=2).

Table 5.5. Clinical characteristics of study cohort by GSTP1 (Ile105Val) gene polymorphism grouping based on genotype.

	Total	AA	AG	GG	p-value**
Cross-sectional study cohort					
Sample size	495	233(47%)	208(42%)	54(11%)	
Current age (years)	26.9 (0.5)	26.6 (0.7)	27.5 (0.7)	26.4 (1.4)	0.59
Current %predFEV_1 (% of predicted)	60.4(1.2)	59.5(1.8)	61.3(1.9)	61.2(3.6)	0.78
Sex (male/female)	252/243	114/119	113/95	25/29	0.41
PSS (insufficient/sufficient)	440/54	208/25	184/23	48/6	0.99
CFTR genotype*	339/116/28/12	167/49/13/4	130/58/12/8	42/9/3/0	0.29
P. aeruginosa infection status (not infected/chronically infected)	142/280	69/134	58/115	17/31	0.96
Longitudinal study cohort					
Sample size	398	191(48%)	167(42%)	40(10%)	
Current age (years)	25.6 (0.5)	25.4 (0.8)	26.0 (0.8)	24.5 (1.6)	0.69
Current %predFEV_1 (% of predicted)	61.0(1.4)	58.6(2.0)	63.2(2.2)	63.7(4.4)	0.31
Current BMI	20.0(0.2)	19.8(0.2)	20.2(0.3)	20.2(0.6)	0.40
Age of CF diagnosis (years)	4.8(0.9)	4.4(0.5)	5.3(0.7)	4.4(0.9)	0.54
Sex (male/female)	203/195	98/93	86/81	19/21	0.90
PSS (insufficient/sufficient)	354/44	178/20	148/19	35/5	0.92
CFTR genotype*	272/91/24/11	138/37/12/4	103/47/10/7	31/7/2/0	0.24

P. aeruginosa infection status (not infected/chronically infected)	14/258	67/125	58/110	16/24	0.80
PA1 age	14.6(0.8) N=222	15.0(1.1) N=114	14.4(1.3) N=80	13.7(1.8) N=28	0.84
PAC age	15.3(0.8) N=152	15.9(1.3) N=76	14.5(1.2) N=60	14.9(1.9) N=18	0.72
Liver disease (present/not present)	341/57	161/30	145/22	35/5	0.75

* and ** are explained in Table 5.3.

Table 5.6. Survival analysis of the time to event (death or lung transplantation) for CF patients.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
GSTM1 and T1 (N=325)				
GSTM1/T1 (one null=1), base group is 0 null	0.28	0.39	0.40	2.43
GSTM1/T1 (two null=1)	$2.60x10^{3}$	0.03	2.36	$2.86x10^{6}$
Sex (Male=0/Female=1)	2.67	0.02	1.19	6.00
PSS (0/1=insufficient/sufficient)	1.2	0.90	0.18	$6.95\ x10^{6}$
CF diagnosis age	0.92	0.10	0.83	1.02
P. aeruginosa infection status (0/1=not infected/chronically infected)	1.22	0.77	0.32	4.60
Current %predFEV_1	0.90	<0.0001	0.86	0.94
Current BMI	0.94	0.35	0.83	1.07
GSTM1/T1 (one null=1)* CF diagnosis age	1.19	0.03	1.02	1.40
GSTM1/T1 (two null=1)* *P. aeruginosa* infection status	0.004	0.02	$3.33x10^{-5}$	0.46
GSTP1 (N=398)				
GSTP1 (AG=0/GG=1), base group is AA	0.11	0.92	0.04	33.06
GSTP1 (GG=0/AG=1)	0.29	0.46	0.08	2.26
Sex (Male=0/Female=1)	3.29	0.0006	1.67	6.52
CFTR class 3 (Heterozygous class 1,2 or 3/unknown or unclassified vs. homozygous severe 1,2,3)	0.29	0.01	0.11	0.75
P. aeruginosa infection status (0/1=not infected/chronically infected)				
Diagnosis age	0.95	0.24	0.88	1.03
Current %predFEV_1	0.87	<0.0001	0.83	0.91
Current BMI	1.01	0.85	0.88	1.14
GCLC (N=340)				

GCLC 7/8 (1=7/8 GAC repeats), base group is 7/7 GAC repeats	2.19	0.08	0.91	5.22
GCLC 7/9 (1=7/9 GAC repeats)	1.36	0.51	0.54	3.41
GCLCgr7 (1=>7 GAC repeats on both chromosomes	0.73	0.55	0.26	2.04
Sex (Male=0/Female=1)	2.83	0.007	1.33	5.98
PSS	0.61	0.63	0.08	4.50
CF diagnosis age	0.96	0.39	0.89	1.05
P. aeruginosa infection status (0/1=not infected/chronically infected)	0.91	0.84	0.32	2.53
Current %predFEV$_1$	0.89	<0.00001	0.86	0.93
Current BMI	0.94	0.35	0.83	1.07

* Formula

Dependent variables: Age of 1st infection (age of chronic infection), PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class + CF diagnosis age + current %predFEV$_1$ + current BMI.

Interactions: Modifier gene* CF diagnosis age, Gene * current %predFEV$_1$.

Table 5.7. Age of first infection with pathogen *P. aeruginosa*.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
GSTM1 and T1 (N=213)				
GSTM1/T1 (one null=1), base group is 0 null	0.99	0.98	0.40	2.43
GSTM1/T1 (two null=1)	0.37	0.19	0.08	1.62
Sex (Male=0/Female=1)	1.70	0.003	1.19	2.41
CFTR class (Mild CFTR class vs. base group homozygous severe 1,2,3)	0.12	0.04	0.02	0.91
CF diagnosis age	0.98	0.25	0.94	1.02
Current %predFEV_1	0.98	0.007	0.97	1.00
Current BMI	1.01	0.84	0.94	1.07
GSTP1 (N=213)				
GSTP1 (AG=0/GG=1), base group is AA	1.24	0.76	0.31	4.97
GSTP1 (GG=0/AG=1)	1.05	0.92	0.41	2.68
Sex (Male=0/Female=1)	1.65	0.004	1.17	2.33
CFTR class (Mild CFTR class vs. homozygous severe 1,2,3)	0.11	0.03	0.01	0.85
Diagnosis age	0.98	0.30	0.94	1.00
Current %predFEV_1	0.98	0.003	0.98	1.00
Current BMI	1.01	0.67	0.95	1.08
GCLC (N=222)				
GCLC 7/8 (1=7/8 GAC repeats), base group is 7/7 GAC repeats	1.46	0.55	0.42	5.02
GCLC 7/9 (1=7/9 GAC repeats)	1.34	0.57	0.49	3.65
GCLCgr7 (1=>7 GAC repeats on both chromosomes	0.75	0.62	0.24	2.33
Sex (Male=0/Female=1)	1.74	0.005	1.24	2.45
CFTR class (Mild CFTR class vs. homozygous severe 1,2,3)	0.13	0.06	0.02	1.06
CF diagnosis age	0.95	0.09	0.90	1.01
Current %predFEV_1	0.98	0.003	0.97	1.00

Current BMI	1.01	0.85	0.94	1.08

* Formula

Dependent variables: Age of 1st infection (age of chronic infection), PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class + CF diagnosis age + current %predFEV$_1$ + current BMI. CFTR class categories results which were significant are only presented in the table.

Interactions: Modifier gene* CF diagnosis age, Gene * current %predFEV$_1$. Interaction results are only presented in the table if significant ($p<0.05$).

Table 5.8. Age of chronic infection with pathogen *P. aeruginosa*.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
GSTM1 and T1 (N=213)				
GSTM1/T1 (one null=1), base group is 0 null	1.22	0.67	0.49	3.05
GSTM1/T1 (two null=1)	0.22	0.07	0.04	1.12
Sex (Male=0/Female=1)	1.18	0.40	0.81	1.73
CF diagnosis age	0.96	0.10	0.92	1.01
Current %predFEV_1	0.99	0.02	0.97	1.00
Current BMI	0.94	0.07	0.87	1.00
GSTM1/T1 (two null=1)* Current %predFEV_1	1.03	0.05	0.99	1.05
GSTP1 (N=148)				
GSTP1 (AG=0/GG=1), base group is AA	0.57	0.55	0.09	3.59
GSTP1 (GG=0/AG=1)	0.91	0.85	0.34	2.45
Sex (Male=0/Female=1)	1.14	0.49	0.79	1.65
Diagnosis age	0.95	0.06	0.90	1.00
Current %predFEV_1	0.99	0.02	0.98	1.00
Current BMI	0.95	0.15	0.88	1.02
GCLC (N=150)				
GCLC 7/8 (1=7/8 GAC repeats), base group is 7/7 GAC repeats	1.13	0.84	0.32	4.02
GCLC 7/9 (1=7/9 GAC repeats)	1.41	0.56	0.44	4.45
GCLCgr7 (1=>7 GAC repeats on both chromosomes	0.77	0.66	0.23	2.51
Sex (Male=0/Female=1)	1.04	0.83	0.72	1.49
CF diagnosis age	0.98	0.58	0.91	1.05
Current %predFEV_1	0.99	0.11	0.98	1.00
Current BMI	0.94	0.12	0.88	1.02

***** Formula

Dependent variables: Age of 1st infection (age of chronic infection), PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class + CF diagnosis age + current %predFEV$_1$ + current BMI. CFTR class categories results which were significant are only presented in the table.

Interactions: Modifier gene* CF diagnosis age, Gene * current %predFEV$_1$. Interaction results are only presented in the table if significant ($p<0.05$).

Table 5.9. Clinical characteristics of CF patients based on liver disease status.

	No liver disease	Liver disease	p-value
Sample size	393	73	
Age of CF diagnosis (years)	5.1(0.5)	2.6(0.6)	0.03
Current age (years)	23.7(0.6)	23.4(0.9)	0.85
Current %predFEV_1 (% of predicted)	64.4(1.5)	60.5(3.1)	0.29
Current BMI (kg/m^2)	20.1(0.2)	20.2(0.3)	0.76
Sex (male/female)	203/190	39/34	0.44
PSS (insufficient/sufficient)	343/50	71/2	0.006
CFTR genotype*	272/95/26	54/111/3	0.60
P. aeruginosa or BCC infections status (not infected/chronically infected)	151/238	14/59	0.001
P. aeruginosa infection status (not infected/chronically infected)	160/228	15/58	<0.0001
PA1 age (years)	14.0(0.8) N=218	10.0(1.2) N=46	0.03
PAC age (years)	14.9(0.9) N=139	11.7(1.5) N=34	0.10
GSTM1 (null/gene)	187/156	38/19	0.06
GSTP1 (Ile-Ile/ Ile-Val/Val-Val)	161/145/36	30/22/5	0.75

* CFTR genotype grouped by class severity into: 2 severe mutations (class 1,2 or 3), one severe mutation and one mutation categorized as unknown or unclassified, one or two mild CFTR mutation (i.e., class 4, or 5).

5.5 BIBLIOGRAPHY

1. Janssen, Y. M., B. Van Houten, P. J. Borm, and B. T. Mossman. 1993. Cell and tissue responses to oxidative damage. *Lab Invest* 69(3):261-74.
2. Church, D. F., and W. A. Pryor. 1985. Free-radical chemistry of cigarette smoke and its toxicological implications. *Environ Health Perspect* 64:111-26.
3. Kinnula, V. L., J. D. Crapo, and K. O. Raivio. 1995. Generation and disposal of reactive oxygen metabolites in the lung. *Lab Invest* 73(1):3-19.
4. Kinnula, V. L., A. R. Whorton, L. Y. Chang, and J. D. Crapo. 1992. Regulation of hydrogen peroxide generation in cultured endothelial cells. *Am J Respir Cell Mol Biol* 6(2):175-82.
5. Kinnula, V. L., L. Chang, J. I. Everitt, and J. D. Crapo. 1992. Oxidants and antioxidants in alveolar epithelial type II cells: in situ, freshly isolated, and cultured cells. *Am J Physiol* 262(1 Pt 1):L69-77.
6. Halliwell, B., and O. I. Aruoma. 1991. DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. *FEBS Lett* 281(1-2):9-19.
7. Crystal, R. G. 1991. Oxidants and respiratory tract epithelial injury: pathogenesis and strategies for therapeutic intervention. *Am J Med* 91(3C):39S-44S.
8. Meister, A. 1988. Glutathione metabolism and its selective modification. *J Biol Chem* 263(33):17205-8.
9. Linsdell, P., and J. W. Hanrahan. 1998. Glutathione permeability of CFTR. *Am J Physiol* 275(1 Pt 1):C323-6.
10. Gao, L., K. J. Kim, J. R. Yankaskas, and H. J. Forman. 1999. Abnormal glutathione transport in cystic fibrosis airway epithelia. *Am J Physiol* 277(1 Pt 1):L113-8.
11. Rahman, I., and W. MacNee. 1999. Lung glutathione and oxidative stress: implications in cigarette smoke-induced airway disease. *Am J Physiol* 277(6 Pt 1):L1067-88.
12. Moss, M., D. M. Guidot, M. Wong-Lambertina, T. Ten Hoor, R. L. Perez, and L. A. Brown. 2000. The effects of chronic alcohol abuse on pulmonary glutathione homeostasis. *Am J Respir Crit Care Med* 161(2 Pt 1):414-9.
13. Rahman, I., E. Skwarska, M. Henry, M. Davis, C. M. O'Connor, M. X. FitzGerald, A. Greening, and W. MacNee. 1999. Systemic and pulmonary oxidative stress in idiopathic pulmonary fibrosis. *Free Radic Biol Med* 27(1-2):60-8.

14. Boda, D., I. Nemeth, and S. Pinter. 1998. Surface tension, glutathione content and redox ratio of the tracheal aspirate fluid of premature infants with IRDS. *Biol Neonate* 74(4):281-8.

15. Droge, W., and E. Holm. 1997. Role of cysteine and glutathione in HIV infection and other diseases associated with muscle wasting and immunological dysfunction. *Faseb J* 11(13):1077-89.

16. Mihm, S., J. Ennen, U. Pessara, R. Kurth, and W. Droge. 1991. Inhibition of HIV-1 replication and NF-kappa B activity by cysteine and cysteine derivatives. *Aids* 5(5):497-503.

17. Pacht, E. R., P. Diaz, T. Clanton, J. Hart, and J. E. Gadek. 1997. Alveolar fluid glutathione decreases in asymptomatic HIV-seropositive subjects over time. *Chest* 112(3):785-8.

18. Staal, F. J., S. W. Ela, M. Roederer, M. T. Anderson, and L. A. Herzenberg. 1992. Glutathione deficiency and human immunodeficiency virus infection. *Lancet* 339(8798):909-12.

19. Konstan, M. W., and M. Berger. 1997. Current understanding of the inflammatory process in cystic fibrosis: onset and etiology. *Pediatr Pulmonol* 24(2):137-42; discussion 159-61.

20. Balough, K., M. McCubbin, M. Weinberger, W. Smits, R. Ahrens, and R. Fick. 1995. The relationship between infection and inflammation in the early stages of lung disease from cystic fibrosis. *Pediatr Pulmonol* 20(2):63-70.

21. Tirouvanziam, R. 1999. [Cystic fibrosis: a complex disease and a paradigm for biomedical research]. *J Soc Biol* 193(2):189-209.

22. Khan, T. Z., J. S. Wagener, T. Bost, J. Martinez, F. J. Accurso, and D. W. Riches. 1995. Early pulmonary inflammation in infants with cystic fibrosis. *Am J Respir Crit Care Med* 151(4):1075-82.

23. Noah, T. L., H. R. Black, P. W. Cheng, R. E. Wood, and M. W. Leigh. 1997. Nasal and bronchoalveolar lavage fluid cytokines in early cystic fibrosis. *J Infect Dis* 175(3):638-47.

24. Osika, E., J. M. Cavaillon, K. Chadelat, M. Boule, C. Fitting, G. Tournier, and A. Clement. 1999. Distinct sputum cytokine profiles in cystic fibrosis and other chronic inflammatory airway disease. *Eur Respir J* 14(2):339-46.

25. Konstan, M. W., K. A. Hilliard, T. M. Norvell, and M. Berger. 1994. Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. *Am J Respir Crit Care Med* 150(2):448-54.

26. Abman, S. H., J. W. Ogle, R. J. Harbeck, N. Butler-Simon, K. B. Hammond, and F. J. Accurso. 1991. Early bacteriologic, immunologic, and clinical courses of young infants with cystic fibrosis identified by neonatal screening. *J Pediatr* 119(2):211-7.

27. Cuzzocrea, S., G. Costantino, B. Zingarelli, E. Mazzon, A. Micali, and A. P. Caputi. 1999. The protective role of endogenous glutathione in carrageenan-induced pleurisy in the rat. *Eur J Pharmacol* 372(2):187-97.

28. Gosset, P., B. Wallaert, A. B. Tonnel, and C. Fourneau. 1999. Thiol regulation of the production of TNF-alpha, IL-6 and IL-8 by human alveolar macrophages. *Eur Respir J* 14(1):98-105.

29. Buhl, R., A. Meyer, and C. Vogelmeier. 1996. Oxidant-protease interaction in the lung. Prospects for antioxidant therapy. *Chest* 110(6 Suppl):267S-272S.

30. Noszal, B., D. Visky, and M. Kraszni. 2000. Population, acid-base, and redox properties of N-acetylcysteine conformers. *J Med Chem* 43(11):2176-82.

31. Li, J. D., W. Feng, M. Gallup, J. H. Kim, J. Gum, Y. Kim, and C. Basbaum. 1998. Activation of NF-kappaB via a Src-dependent Ras-MAPK-pp90rsk pathway is required for Pseudomonas aeruginosa-induced mucin overproduction in epithelial cells. *Proc Natl Acad Sci U S A* 95(10):5718-23.

32. Fischer, B., and J. Voynow. 2000. Neutrophil elastase induces MUC5AC messenger RNA expression by an oxidant-dependent mechanism. *Chest* 117(5 Suppl 1):317S-20S.

33. Bilzer, M., and B. H. Lauterburg. 1991. Glutathione metabolism in activated human neutrophils: stimulation of glutathione synthesis and consumption of glutathione by reactive oxygen species. *Eur J Clin Invest* 21(3):316-22.

34. Ogino, T., L. Packer, and J. J. Maguire. 1997. Neutrophil antioxidant capacity during the respiratory burst: loss of glutathione induced by chloramines. *Free Radic Biol Med* 23(3):445-52.

35. Smith, C. V., D. P. Jones, T. M. Guenthner, L. H. Lash, and B. H. Lauterburg. 1996. Compartmentation of glutathione: implications for the study of toxicity and disease. *Toxicol Appl Pharmacol* 140(1):1-12.

36. Cantin, A. M., S. L. North, R. C. Hubbard, and R. G. Crystal. 1987. Normal alveolar epithelial lining fluid contains high levels of glutathione. *J Appl Physiol* 63(1):152-7.

37. Sutherland, M. W., J. Nelson, G. Harrison, and H. J. Forman. 1985. Effects of t-butyl hydroperoxide on NADPH, glutathione, and the respiratory burst of rat alveolar macrophages. *Arch Biochem Biophys* 243(2):325-31.

38. Roum, J. H., Z. Borok, N. G. McElvaney, G. J. Grimes, A. D. Bokser, R. Buhl, and R. G. Crystal. 1999. Glutathione aerosol suppresses lung epithelial surface inflammatory cell-derived oxidants in cystic fibrosis. *J Appl Physiol* 87(1):438-43.

39. Lands, L. C., V. L. Grey, and C. Grenier. 2000. Total plasma antioxidant capacity in cystic fibrosis. *Pediatr Pulmonol* 29(2):81-7.

40. Rahman, Q., P. Abidi, F. Afaq, D. Schiffmann, B. T. Mossman, D. W. Kamp, and M. Athar. 1999. Glutathione redox system in oxidative lung injury. *Crit Rev Toxicol* 29(6):543-68.

41. Smith, L. J., and J. Anderson. 1992. Oxygen-induced lung damage. Relationship to lung mitochondrial glutathione levels. *Am Rev Respir Dis* 146(6):1452-7.

42. Kelly, F. J. 1999. Gluthathione: in defence of the lung. *Food Chem Toxicol* 37(9-10):963-6.

43. Lands, L. C., V. Grey, A. A. Smountas, V. G. Kramer, and D. McKenna. 1999. Lymphocyte glutathione levels in children with cystic fibrosis. *Chest* 116(1):201-5.

44. Boxer, L. A., J. M. Oliver, S. P. Spielberg, J. M. Allen, and J. D. Schulman. 1979. Protection of granulocytes by vitamin E in glutathione synthetase deficiency. *N Engl J Med* 301(17):901-5.

45. Lauterburg, B. H., and M. E. Velez. 1988. Glutathione deficiency in alcoholics: risk factor for paracetamol hepatotoxicity. *Gut* 29(9):1153-7.

46. Loguercio, C., D. Taranto, L. M. Vitale, F. Beneduce, and C. Del Vecchio Blanco. 1996. Effect of liver cirrhosis and age on the glutathione concentration in the plasma, erythrocytes, and gastric mucosa of man. *Free Radic Biol Med* 20(3):483-8.

47. Wallig, M. A. 1998. Xenobiotic metabolism, oxidant stress and chronic pancreatitis. Focus on glutathione. *Digestion* 59 Suppl 4:13-24.

48. Vijayalingam, S., A. Parthiban, K. R. Shanmugasundaram, and V. Mohan. 1996. Abnormal antioxidant status in impaired glucose tolerance and non-insulin-dependent diabetes mellitus. *Diabet Med* 13(8):715-9.

49. Ushmorov, A., V. Hack, and W. Droge. 1999. Differential reconstitution of mitochondrial respiratory chain activity and plasma redox state by cysteine and ornithine in a model of cancer cachexia. *Cancer Res* 59(14):3527-34.

50. Ardite, E., M. Sans, J. Panes, F. J. Romero, J. M. Pique, and J. C. Fernandez-Checa. 2000. Replenishment of glutathione levels improves mucosal function in experimental acute colitis. *Lab Invest* 80(5):735-44.

51. Velsor, L. W., A. van Heeckeren, and B. J. Day. 2001. Antioxidant imbalance in the lungs of cystic fibrosis transmembrane conductance regulator protein mutant mice. *Am J Physiol Lung Cell Mol Physiol* 281(1):L31-8.

52. Gao, L., J. R. Broughman, T. Iwamoto, J. M. Tomich, C. J. Venglarik, and H. J. Forman. 2001. Synthetic chloride channel restores glutathione secretion in cystic fibrosis airway epithelia. *Am J Physiol Lung Cell Mol Physiol* 281(1):L24-30.

53. Birrer, P., N. G. McElvaney, A. Rudeberg, C. W. Sommer, S. Liechti-Gallati, R. Kraemer, R. Hubbard, and R. G. Crystal. 1994. Protease-antiprotease imbalance in the lungs of children with cystic fibrosis. *Am J Respir Crit Care Med* 150(1):207-13.

54. 1994. Therapeutic approaches to cystic fibrosis: memorandum from a joint WHO/ICF(M)A meeting. *Bull World Health Organ* 72(3):341-52.

55. Jain, A., J. Martensson, T. Mehta, A. N. Krauss, P. A. Auld, and A. Meister. 1992. Ascorbic acid prevents oxidative stress in glutathione-deficient mice: effects on lung type 2 cell lamellar bodies, lung surfactant, and skeletal muscle. *Proc Natl Acad Sci U S A* 89(11):5093-7.

56. Jain, A., T. Mehta, P. A. Auld, J. Rodrigues, R. F. Ward, M. K. Schwartz, and J. Martensson. 1995. Glutathione metabolism in newborns: evidence for glutathione deficiency in plasma, bronchoalveolar lavage fluid, and lymphocytes in prematures. *Pediatr Pulmonol* 20(3):160-6.

57. Kokura, S., R. E. Wolf, T. Yoshikawa, D. N. Granger, and T. Y. Aw. 1999. Molecular mechanisms of neutrophil-endothelial cell adhesion induced by redox imbalance. *Circ Res* 84(5):516-24.

58. Dominguez, C., S. Gartner, S. Linan, N. Cobos, and A. Moreno. 1998. Enhanced oxidative damage in cystic fibrosis patients. *Biofactors* 8(1-2):149-53.

59. Barker, J. E., S. J. Heales, A. Cassidy, J. P. Bolanos, J. M. Land, and J. B. Clark. 1996. Depletion of brain glutathione results in a decrease of glutathione reductase activity; an enzyme susceptible to oxidative damage. *Brain Res* 716(1-2):118-22.

60. Remiao, F., H. Carmo, F. D. Carvalho, and M. L. Bastos. 2000. Inhibition of glutathione reductase by isoproterenol oxidation products. *J Enzyme Inhib* 15(1):47-61.

61. van Klaveren, R. J., P. H. Hoet, J. L. Pype, M. Demedts, and B. Nemery. 1997. Increase in gamma-glutamyltransferase by glutathione depletion in rat type II pneumocytes. *Free Radic Biol Med* 22(3):525-34.

62. Gilmont, R. R., A. Dardano, M. Young, J. S. Engle, B. S. Adamson, D. J. Smith, Jr., and R. S. Rees. 1998. Effects of glutathione depletion on oxidant-induced endothelial cell injury. *J Surg Res* 80(1):62-8.

63. Li, X. Y., K. Donaldson, I. Rahman, and W. MacNee. 1994. An investigation of the role of glutathione in increased epithelial permeability induced by cigarette smoke in vivo and in vitro. *Am J Respir Crit Care Med* 149(6):1518-25.

64. Jiang, N., K. L. Dreher, J. A. Dye, Y. Li, J. H. Richards, L. D. Martin, and K. B. Adler. 2000. Residual oil fly ash induces cytotoxicity and mucin secretion by guinea pig tracheal epithelial cells via an oxidant-mediated mechanism. *Toxicol Appl Pharmacol* 163(3):221-30.

65. van der Vliet, A., J. P. Eiserich, G. P. Marelich, B. Halliwell, and C. E. Cross. 1997. Oxidative stress in cystic fibrosis: does it occur and does it matter? *Adv Pharmacol* 38:491-513.

66. Cho, S., M. Hazama, Y. Urata, S. Goto, S. Horiuchi, K. Sumikawa, and T. Kondo. 1999. Protective role of glutathione synthesis in response to oxidized low density lipoprotein in human vascular endothelial cells. *Free Radic Biol Med* 26(5-6):589-602.

67. Galloway, D. C., D. G. Blake, A. G. Shepherd, and L. I. McLellan. 1997. Regulation of human gamma-glutamylcysteine synthetase: co-ordinate induction of the catalytic and regulatory subunits in HepG2 cells. *Biochem J* 328 (Pt 1):99-104.

68. Walsh, A. C., J. A. Feulner, and A. Reilly. 2001. Evidence for functionally significant polymorphism of human glutamate cysteine ligase catalytic subunit: association with glutathione levels and drug resistance in the National Cancer Institute tumor cell line panel. *Toxicol Sci* 61(2):218-23.

69. Huang, C. S., M. E. Anderson, and A. Meister. 1993. Amino acid sequence and function of the light subunit of rat kidney gamma-glutamylcysteine synthetase. *J Biol Chem* 268(27):20578-83.

70. Walsh, A. C., W. Li, D. R. Rosen, and D. A. Lawrence. 1996. Genetic mapping of GLCLC, the human gene encoding the catalytic subunit of gamma-glutamyl-cysteine synthetase, to chromosome band 6p12 and characterization of a polymorphic trinucleotide repeat within its 5' untranslated region. *Cytogenet Cell Genet* 75(1):14-6.

71. Mulcahy, R. T., and J. J. Gipp. 1995. Identification of a putative antioxidant response element in the 5'-flanking region of the human gamma-glutamylcysteine synthetase heavy subunit gene. *Biochem Biophys Res Commun* 209(1):227-33.

72. McKone, E. F., J. Shao, D. D. Frangolias, C. L. Keener, C. A. Shephard, F. M. Farin, M. R. Tonelli, P. D. Pare, A. J. Sandford, M. L. Aitken, and T. J. Kavanagh. 2006. Variants in the glutamate-cysteine-ligase gene are associated with cystic fibrosis lung disease. *Am J Respir Crit Care Med* 174(4):415-9.

73. Ristoff, E., C. Augustson, J. Geissler, T. de Rijk, K. Carlsson, J. L. Luo, K. Andersson, R. S. Weening, R. van Zwieten, A. Larsson, and D. Roos. 2000. A missense mutation in the heavy subunit of gamma-glutamylcysteine synthetase gene causes hemolytic anemia. *Blood* 95(7):2193-6.

74. Beutler, E., T. Gelbart, T. Kondo, and A. T. Matsunaga. 1999. The molecular basis of a case of gamma-glutamylcysteine synthetase deficiency. *Blood* 94(8):2890-4.

75. Xu, S., Y. Wang, B. Roe, and W. R. Pearson. 1998. Characterization of the human class Mu glutathione S-transferase gene cluster and the GSTM1 deletion. *J Biol Chem* 273(6):3517-27.

76. Baez, S., J. Segura-Aguilar, M. Widersten, A. S. Johansson, and B. Mannervik. 1997. Glutathione transferases catalyse the detoxication of oxidized metabolites (o-quinones) of catecholamines and may serve as an antioxidant system preventing degenerative cellular processes. *Biochem J* 324 (Pt 1):25-8.

77. Berhane, K., M. Widersten, A. Engstrom, J. W. Kozarich, and B. Mannervik. 1994. Detoxication of base propenals and other alpha, beta-unsaturated aldehyde products of radical reactions and lipid peroxidation by human glutathione transferases. *Proc Natl Acad Sci U S A* 91(4):1480-4.

78. Hubatsch, I., M. Ridderstrom, and B. Mannervik. 1998. Human glutathione transferase A4-4: an alpha class enzyme with high catalytic efficiency in the conjugation of 4-hydroxynonenal and other genotoxic products of lipid peroxidation. *Biochem J* 330 (Pt 1):175-9.

79. Hayes, J. D., and L. I. McLellan. 1999. Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. *Free Radic Res* 31(4):273-300.

80. Cotton, S. C., L. Sharp, J. Little, and N. Brockton. 2000. Glutathione S-transferase polymorphisms and colorectal cancer: a HuGE review. *Am J Epidemiol* 151(1):7-32.

81. Strange, R. C., B. Matharoo, G. C. Faulder, P. Jones, W. Cotton, J. B. Elder, and M. Deakin. 1991. The human glutathione S-transferases: a case-control study of the incidence of the GST1 0 phenotype in patients with adenocarcinoma. *Carcinogenesis* 12(1):25-8.

82. Heagerty, A. H., D. Fitzgerald, A. Smith, B. Bowers, P. Jones, A. A. Fryer, L. Zhao, J. Alldersea, and R. C. Strange. 1994. Glutathione S-transferase GSTM1 phenotypes and protection against cutaneous tumours. *Lancet* 343(8892):266-8.

83. Harrison, D. J., A. M. Cantlay, F. Rae, D. Lamb, and C. A. Smith. 1997. Frequency of glutathione S-transferase M1 deletion in smokers with emphysema and lung cancer. *Hum Exp Toxicol* 16(7):356-60.

84. Baranova, H., J. Perriot, E. Albuisson, T. Ivaschenko, V. S. Baranov, B. Hemery, P. Mouraire, N. Riol, and P. Malet. 1997. Peculiarities of the GSTM1 0/0 genotype in French heavy smokers with various types of chronic bronchitis. *Hum Genet* 99(6):822-6.

85. Mohr, L. C., J. K. Rodgers, and G. A. Silvestri. 2003. Glutathione S-transferase M1 polymorphism and the risk of lung cancer. *Anticancer Res* 23(3A):2111-24.

86. Nazar-Stewart, V., T. L. Vaughan, P. Stapleton, J. Van Loo, B. Nicol-Blades, and D. L. Eaton. 2003. A population-based study of glutathione S-transferase M1, T1 and P1 genotypes and risk for lung cancer. *Lung Cancer* 40(3):247-58.

87. Baranov, V. S., T. Ivaschenko, B. Bakay, M. Aseev, R. Belotserkovskaya, H. Baranova, P. Malet, J. Perriot, P. Mouraire, V. N. Baskakov, G. A. Savitskyi, S. Gorbushin, S. I. Deyneka, E. Michnin, A. Barchuck, V. Vakharlovsky, G. Pavlov, V. I. Shilko, T. Guembitzkaya, and L. Kovaleva. 1996. Proportion of the GSTM1 0/0 genotype in some Slavic populations and its correlation with cystic fibrosis and some multifactorial diseases. *Hum Genet* 97(4):516-20.

88. Hull, J., and A. H. Thomson. 1998. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. *Thorax* 53(12):1018-21.

89. Imboden, M., S. H. Downs, O. Senn, G. Matyas, O. Brandli, E. W. Russi, C. Schindler, U. Ackermann-Liebrich, W. Berhger, N. M. Probst-Hensch, and S. Team. 2007. Glutathione S-transferase genotypes modify lung function decline in the general population: SAPALDIA cohort study. *Respiratory Research* 8(2):January 11.

90. Watson, M. A., R. K. Stewart, G. B. Smith, T. E. Massey, and D. A. Bell. 1998. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. *Carcinogenesis* 19(2):275-80.

91. Harries, L. W., M. J. Stubbins, D. Forman, G. C. Howard, and C. R. Wolf. 1997. Identification of genetic polymorphisms at the glutathione S-transferase Pi locus and association with susceptibility to bladder, testicular and prostate cancer. *Carcinogenesis* 18(4):641-4.

92. Miller, D. P., D. Neuberg, I. de Vivo, J. C. Wain, T. J. Lynch, L. Su, and D. C. Christiani. 2003. Smoking and the risk of lung cancer: susceptibility with GSTP1 polymorphisms. *Epidemiology* 14(5):545-51.

93. He, J. Q., J. Ruan, J. E. Connett, N. R. Anthonisen, P. D. Pare, and A. J. Sandford. 2002. Antioxidant gene polymorphisms and susceptibility to a rapid decline in lung function in smokers. *Am J Respir Crit Care Med* 166(3):323-8.

94. Strange, R. C., M. A. Spiteri, S. Ramachandran, and A. A. Fryer. 2001. Glutathione-S-transferase family of enzymes. *Mutat Res* 482(1-2):21-6.

95. Fryer, A. A., A. Bianco, M. Hepple, P. W. Jones, R. C. Strange, and M. A. Spiteri. 2000. Polymorphism at the glutathione S-transferase GSTP1 locus. A new marker for bronchial hyperresponsiveness and asthma. *Am J Respir Crit Care Med* 161(5):1437-42.

96. Child, F., W. Lenney, S. Clayton, S. Davies, P. W. Jones, J. E. Alldersea, R. C. Strange, and A. A. Fryer. 2003. The association of maternal but not paternal genetic variation in GSTP1 with asthma phenotypes in children. *Respir Med* 97(12):1247-56.

97. Coggan, M., L. Whitbread, A. Whittington, and P. Board. 1998. Structure and organization of the human theta-class glutathione S-transferase and D-dopachrome tautomerase gene complex. *Biochem J* 334 (Pt 3):617-23.

98. Pemble, S., K. R. Schroeder, S. R. Spencer, D. J. Meyer, E. Hallier, H. M. Bolt, B. Ketterer, and J. B. Taylor. 1994. Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J* 300 (Pt 1):271-6.

99. Sherratt, P. J., D. J. Pulford, D. J. Harrison, T. Green, and J. D. Hayes. 1997. Evidence that human class Theta glutathione S-transferase T1-1 can catalyse the activation of dichloromethane, a liver and lung carcinogen in the mouse. Comparison of the tissue distribution of GST T1-1 with that of classes Alpha, Mu and Pi GST in human. *Biochem J* 326 (Pt 3):837-46.

100. Guengerich, F. P., R. Thier, M. Persmark, J. B. Taylor, S. E. Pemble, and B. Ketterer. 1995. Conjugation of carcinogens by theta class glutathione s-transferases: mechanisms and relevance to variations in human risk. *Pharmacogenetics* 5 Spec No:S103-7.

101. Schroder, K. R., F. A. Wiebel, S. Reich, D. Dannappel, H. M. Bolt, and E. Hallier. 1995. Glutathione-S-transferase (GST) theta polymorphism influences background SCE rate. *Arch Toxicol* 69(7):505-7.

102. Peter, H., S. Deutschmann, C. Reichel, and E. Hallier. 1989. Metabolism of methyl chloride by human erythrocytes. *Arch Toxicol* 63(5):351-5.

103. Wiebel, F. A., A. Dommermuth, and R. Thier. 1999. The hereditary transmission of the glutathione transferase hGSTT1-1 conjugator phenotype in a large family. *Pharmacogenetics* 9(2):251-6.

104. Rebbeck, T. R. 1997. Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility. *Cancer Epidemiol Biomarkers Prev* 6(9):733-43.

105. Matthias, C., V. Jahnke, P. W. Jones, P. R. Hoban, J. E. Alldersea, S. F. Worrall, A. A. Fryer, and R. C. Strange. 1999. Cyclin D1, glutathione S-transferase, and cytochrome P450 genotypes and outcome in patients with upper aerodigestive tract cancers: assessment of the importance of individual genes using multivariate analysis. *Cancer Epidemiol Biomarkers Prev* 8(9):815-23.

106. Colombo, C., A. Crosignani, and P. M. Battezzati. 1999. Liver involvement in cystic fibrosis. *J Hepatol* 31(5):946-54.

107. Lindblad, A., H. Glaumann, and B. Strandvik. 1999. Natural history of liver disease in cystic fibrosis. *Hepatology* 30(5):1151-8.

108. Lakehal, F., D. Wendum, V. Barbu, L. Becquemont, R. Poupon, P. Balladur, L. Hannoun, F. Ballet, P. H. Beaune, and C. Housset. 1999. Phase I and phase II drug-metabolizing enzymes are expressed and heterogeneously distributed in the biliary epithelium. *Hepatology* 30(6):1498-506.

109. Whalen, R., D. C. Rockey, S. L. Friedman, and T. D. Boyer. 1999. Activation of rat hepatic stellate cells leads to loss of glutathione S-transferases and their enzymatic activity against products of oxidative stress. *Hepatology* 30(4):927-33.

110. Zsembery, A., W. Jessner, G. Sitter, C. Spirli, M. Strazzabosco, and J. Graf. 2002. Correction of CFTR malfunction and stimulation of Ca-activated Cl channels restore HCO3-secretion in cystic fibrosis bile ductular cells. *Hepatology* 35(1):95-104.

111. Fitzsimmons, S. C. 1995. Cystic fibrosis foundation, Patient registry 1994 annual data report. MD Cystic Fibrosis Foundation, Bethesda.

112. Henrion-Caude, A., C. Flamant, M. Roussey, C. Housset, A. Flahault, A. A. Fryer, K. Chadelat, R. C. Strange, and A. Clement. 2002. Liver disease in pediatric patients with cystic fibrosis is associated with glutathione S-transferase P1 polymorphism. *Hepatology* 36(4 Pt 1):913-7.

113. Buetler, T. M. 1998. Identification of glutathione S-transferase isozymes and gamma-glutamylcysteine synthetase as negative acute-phase proteins in rat liver. *Hepatology* 28(6):1551-60.

114. Flamant, C., A. Henrion-Caude, P. Y. Boelle, F. Bremont, J. Brouard, B. Delaisi, J. F. Duhamel, C. Marguet, M. Roussey, M. C. Miesch, M. Boule, R. C. Strange, and A. Clement. 2004. Glutathione-S-transferase M1, M3, P1 and T1 polymorphisms and severity of lung disease in children with cystic fibrosis. *Pharmacogenetics* 14(5):295-301.

115. Boucher, R. C. 2001. Pathogenesis of cystic fibrosis airways disease. *Trans Am Clin Climatol Assoc* 112:99-107.

116. Boucher, R. C. 2002. An overview of the pathogenesis of cystic fibrosis lung disease. *Adv Drug Deliv Rev* 54(11):1359-71.

117. Perez-Vilar, J., and R. C. Boucher. 2004. Reevaluating gel-forming mucins' roles in cystic fibrosis lung disease. *Free Radic Biol Med* 37(10):1564-77.

CHAPTER 6: RESPIRATORY MUCIN GENES AS POTENTIAL MODIFIER GENES IN CYSTIC FIBROSIS

A version of this chapter may be published in the future.

6.0 INTRODUCTION

In this chapter we investigated mucin protein producing genes. We investigated the effect of variable number tandem repeat polymorphisms in mucin genes MUC2 and MUC5B and the association between repeat number and pulmonary disease severity and progression, survival and succeptibility to infection with common CF pathogens.

6.1 RATIONALE FOR THE INVESTIGATION OF MUCIN GENES AS POTENTIAL MODIFIER GENES IN CF

A major characteristic of CF pulmonary disease is mucus hypersecretion and increased viscosity of mucus and in this chapter we have investigated polymorphisms in mucin genes expressed in the pulmonary system as potential modifier genes for CF pulmonary disease. The respiratory tract epithelial responses associated with CF but also observed in patients diagnosed as having asthma and chronic bronchitis include mucus hypersecretion, changes in mucociliary escalator activity and release of inflammatory mediators (cytokines and lipid-derived mediators). The CFTR defect is hypothesized to contribute to dehydration of the airway mucus which leads to poor mucociliary clearance and creation of mucus plugs (1, 2). *P. aeruginosa* lipopolysaccharide (LPS) (3, 4), proinflammatory cytokines (IL-1beta, TNF-alpha) and neutrophil elastase (4) have been shown to increase transcription of mucin mRNA levels via activation of the NF-κB pathway (5, 6). Mucin glycoproteins are the major constituent of mucus and are implicated in the process of chronic infection and inflammation in CF airways (7, 8). Pathogens commonly colonizing CF patients have all been shown to adhere to respiratory mucins. A number of receptors on mucins have been identified that bind to pathogens commonly infecting CF patients such as: *P. aeruginosa* (9-13), *Staphyloccocus aureus* (12, 14), *H. influenzae* (12, 15) and *Burkholderia cepacia* complex (BCC) (12, 16). Li and associates showed that bacterial exo-products of *P. aeruginosa* can upregulate mucin gene transcription via activation of the NF-κB pathway (5). The murine model of CF (CFTR knock-out mice) has been shown to be a poor model for the human pulmonary disease manifestations of CF. Two reasons for this are the lack of mucus producing cells (i.e., goblet cells) in the murine airways and expression of an alternate chloride channel as the major chloride ion channel in mice. Mucus obstruction and colonization of the airways with common pathogens colonizing human CF patients are therefore not observed in the CFTR knock-out mouse. Mucins have a complex structure and

their ultimate composition and viscosity are influenced by the respiratory milieu. In this chapter we review the composition of airway secretions and structure of mucins and investigate the role they play in CF pulmonary disease progression by exploring polymorphisms in two mucin genes expressed in the respiratory tract as potential candidate modifier genes in pulmonary disease severity and progression.

6.1.1 Composition of airway secretions and mucuciliary clearance

Airway secretions consist of surfactant, periciliary fluid and mucus. Mucus consists of two phases, a superficial layer referred to as the *gel phase* or mucous layer and a liquid or periciliary fluid layer which is called the *sol phase*. A thin layer of surfactant is sandwiched in between the gel and sol phase. Surfactant is secreted by submucosal glands. Surfactant is believed to promote spreading of the gel phase, preserving the integrity of the gel and sol phases and contributes to energy transfer from the beating cilia to the mucus without the cilia becoming entangled in the mucus (7, 17-20). The gel phase extends from the bronchioles to the upper airways and forms a 5-20 μm thick layer which rests on the tips of the cilia. Macromolecules can be found to be densely distributed in the gel phase. The sol phase appears as a continuous blanket which lies between the cell surface and the gel phase. The height of this layer extends just short of the fully extended cilium (0.5-3.5μm in thickness)(7, 18). Figure 6.1 shows the gel and sol phase lining respiratory epithelia in normal and CF airways.

Airway mucus predominately consists of water, and secondarily of carbohydrates, proteins and lipids (21, 22). The carbohydrates, lipids, and proteins make up 5-7% of the mucus (21, 22), and originate predominately from secreted mucins and secondarily from secreted antimicrobial proteins and peptides, phospholipids, inhaled particles and cellular debris. Cellular debris includes inflammatory cells (macrophages, neutrophils, eosinophils and lymphocytes) and sloughed off epithelial cells.

Mucins are densely packed in the secretory granules of goblet cells and undergo massive swelling due to hydration during exocytosis. Mucus is formed within the Golgi apparatus of mucus producing cells and is secreted through vesicles into the airway lumen. Exocytosis is characterized by fusion of the secretory granules with the apical membrane of the goblet cell (23-26). Once the secretory granules fuse with the cell membrane, it is believed that there is

an exchange of calcium ions in the vesicles for extracellular sodium ions and water resulting in the hydration and release of the mucins from the secretory vesicles (27). The water for the hydration of mucins is donated from the periciliary or sol phase (28).

The main cell type in the small airways, which is also the site of mucus obstruction and plugging in CF, is the goblet cell. The massive swelling of released mucins from the secretory cells is dependent on the hydration and osmolarlity of the surrounding milieu (29). This mucin gel expansion is ultimately constrained by the active ion and water transport of the airway ciliated epithelial cells. Active sodium ion transport (absorption) across the airway epithelium regulates water absorption and chloride secretion and this process is mainly regulated through the CFTR channel and promotes fluid secretion. Water flow is transcellular and regulates the depth of the sol phase and the hydration of the gel phase. ATP and UTP stimulate chloride secretion from normal and CF airway epithelial cells (30-32).

Cilia are responsible for the cranial movement of mucus and foreign particles out of the airways. Ciliated cells line the surface of airways and by coordinated beating of these cilia clear mucus and entrapped material from the pulmonary system are transported up to the oropharynx and from there are either expectorated or swallowed. Mucociliary clearance is dependent on:

- The viscoelasticity of secretions (33), and
- Properly functioning airway epithelial cilia (25, 28, 33).

6.1.2 Rheological properties of CF airway mucus and mucociliary transport

In CF, mucus is more viscous and adhesive. The CFTR defect contributes both directly and indirectly to this increased viscosity and adhesiveness of airway mucus, although it is unclear whether airway mucus is abnormal at birth and contributes to the early inflammation and infection in CF. There is no chloride secretion by the CFTR under resting conditions and absorption of sodium through the sodium epithelial channel is dominant. In the absence of CFTR inhibition of sodium absorption, as occurs in CF, there is reduced airway fluid volume. Specifically CF mucus becomes dehydrated once excreted by goblet cells and other mucus-secreting cells and the observed outcome is thick and viscous mucus. CF sputum has been shown to be more adhesive than sputum from patients diagnosed with other respiratory diseases (i.e., fatal asthma and airway mucus from normals) and comparable with mucus

from patients diagnosed with chronic bronchitis (19). An inverse relationship has been shown between sputum adhesion and the ability to clear the secretions (19). The CFTR defect contributes to the dehydration of the gel phase and a point is reached when the gel phase can no longer donate liquid to the periciliary liquid phase. Mucus clearance is abolished when the cilia are ultimately unable to fully extend and beat because of the absorption of the periciliary liquid or sol phase and the increased airway mucus viscosity, which contribute to the physical adhesion of mucus to airway epithelia surfaces (1, 2).

Considering both Welsh's high salt hypothesis and more acidic respiratory milieu (34) and Boucher's low airway surface fluid(2), the post-translational structure of mucin proteins would be affected and potentially contribute to the viscosity observed in CF mucus and its impaired clearance. Another hypothesis set forward to explain CF airway disease and mucus viscosity and impaired clearance considers the mucin polymer structure and the alterations shown in CF mucus and the CF airway environment (dehydration and abnormal airway pH)(35). They suggest that(35):

1. If CF mucus is more acidic than mucus from normal healthy individuals and mucin polymerization is pH driven, this would result in an increased formation of disulfide bonds between mucin molecules resulting in the formation of disulfide-linked multimeric molecules.
2. Reduced levels of reduced glutathione (GSH) would favor the formation of inter-chain disulfide bonds among non-O-glycocylated, cysteine rich domains and produce mucin aggregation: MUC-SH + MUC-SH + GSSG→MUC-S-S-MUC + GSH
3. Elevated levels of hypochlorous acid (HOCl) would favor the formation of interchain disulfide bonds and sulfonamide bonds among non-O-glycocylated, cysteine rich domains and produce mucin aggregation:
    a. Sulfide bonds: MUC-SH + MUC-SH +HOCl→ MUC-S-S-MUC + H_2O + HCl
    b. Sulfonamide bonds: MUC-SH + MUC-NH +HOCl→ MUC-SO_2-NH-MUC + H_2O + HCl
4. Increased levels of myeloperoxidase may catalyze protein cross-links through oxidation of tyrosine residues.

Also contributing to the viscosity and viscoelasticity of mucus are products of inflammation. CF chronic inflammation is characterized by high neutrophil influx and lysis of these cells (and other inflammatory cells) within the airway lumen and the consequent release of cytoplasmic proteins, neutrophil derived DNA and filamentous actin; all identified in CF secretions (36, 37). Necrotic neutrophils release proinflammatory mediators which damage the epithelium and recruit more inflammatory cells (38, 39). The high levels of neutrophil elastase may also impair removal of apoptotic neutrophils by macrophages (39). The neutrophil derived DNA and filamentous actin have been shown to increase the viscosity of airway secretions by interweaving in the mucin gel and forming a rigid network within the airway secretions (40, 41). In summary, defective CFTR affects the airway mileu and contributes to mucus viscosity and impaired mucociliary clearance, but also contributes to the process of inflammation. The products of inflammation and infection affect mucus viscosity, viscoelasticity and mucociliary clearance.

6.1.3 Structure of mucin

Mucins are large molecular mass glycoconjugates (154 to $\geq$7000kDa) which contain 50-85 percent carbohydrates and a protein backbone. (7, 42). There are two types of mucins:

- Secreted gel forming mucins which form the mucous gel layer and
- Membrane tethered mucins or transmembrane mucins.

The membrane tethered mucins are synthesized by most epithelial cells, and are present on epithelial cell surfaces and act as cell surface receptors. The protein is anchored to the apical surface of epithelial cells by a transmembrane domain. Their role on airway epithelium appears to be to protect the cells/host by binding to pathogens. High levels of expression have been shown and they are hypothesized to function as activators of signaling pathways resulting in innate non-immune defense and epithelial repair/proliferation or differentiation during injury or stress (43). A number of common features have been shown for membrane bound mucins; trasmembrane domain, short cytoplasmic tail and a highly glycocylated extracellular domain(44).

In this study we investigated two secreted gel mucins, and this review will be limited to the structure of secreted gel mucins. The protein backbone of secreted gel forming mucins is called apomucin. The mucin protein is composed of tandemly repeating sequences of amino acids; rich in serine and threonine which serve as the sites of O-linked glycosylation. The

mucin oligosaccharides bind to the protein core initially through an alpha-O-glycosidic linkage of N-acetyl-galactosamine to the hydroxyl end of the serine or threonine and this reaction is catalyzed by a uridine diphosphate- N-acetyl-galactosamine-alpha- N-acetyl-galactosamine transferase. Additional oligosaccharides (fucose, galactose, N-acetylglucosamine, N-acetylneuraminic acid and carbohydrates) then bind to the N-acetyl-galactosamine creating the oligosaccharide side chains of mucin proteins. The ratio of protein and oligosaccharides in mucins is 4 to 1. Each oligosaccharide side chain linked via O-linked glycosylation is described as having a core, backbone and periphery. The peptide linked N-acetyl-galactosamine and the oligosaccharide directly attached to it form the core region. Four core types have been identified based on the oligosaccharide attached to the N-acetyl-galactosamine. The backbone, which can be linear or branched, is made up of disaccharides formed by alternating galactose and N-acetylglucosamine. The periphery contains acidic molecules (sulfate or peripheral sugars: fucose, N-acetylneuraminic acid, N-acetylgalactosamine, galactose) which are the sites of hydration during exocytosis from the secretory granules (45). The secreted gel-forming mucins form long polymers by linking end to end with other mucin molecules by terminal disulphide bonds via the cysteine rich domains on their amino and terminal carboxyl non-repetitive domains, thus forming large disulfide linked polymers. Other intermolecular bonds (hydrogen bonds, ionic bonds, Van der Waal's forces) contribute to the formation of additional connections with other mucin molecules, as well as with DNA from inflammatory cells and exoproducts from bacteria (46). The disulfide bonds are believed to contribute to mucous viscosity, as reduction of these bonds results in the loss of viscosity of mucous gels (47). Figure 6.2 shows the major domains of secreting gel mucins and Figure 6.3 shows the protein backbone and oligosaccharide structure of these mucins.

The final translated mucin products are complex macromolecules which undergo further post-translational modification by glycosyltransferases, sialotransferases and sulfotransferases in the Golgi compartments and are also further influenced by disease which can also affect glycosylation and post-translational modifications to the mucin protein backbone (7). The CFTR deltaF508 mutation and likely other CFTR mutation proteins belonging to class 1 and 2 where the CFTR protein is either not produced or the protein produced is degraded in the endoplasmic reticulum may lead to modifications in post-translational sulfation and glycosylation processes for mucin (48, 49). Zhang and associates

showed higher sulfate content in CF respiratory mucins and additionally showed that sulfate content was related to CFTR genotype (50). The authors also showed differences in glycosylation in secreted mucins from CF patients with severe pulmonary disease and that bacterial infection may affect the expression of glycosyltransferases and sialotransferases in airway mucus cells (50).

6.1.4 Origin of mucin secreting cells and mucin genes

Twenty mucin genes have been identified to date(44, 51). Secreted mucins include: MUC2, MUC5AC, MUC5B, MUC6, MUC8, and MUC19. Transmembrane mucins include: MUC1, MUC3A, MUC3B, MUC4, MUC11-18, and MUC19. There are a few mucins which are not easily classified; MUC9 is classified as a secreted and membrane bound mucin(44, 51) and MUC7 is found in secretions, but is considered a secreted but not gel forming mucin, as it exists in monomer form (44). Mucins found in the respiratory tract include MUC2, MUC5AC, MUC5B, and MUC6 of the secreted mucins and MUC 1, MUC4 and MUC18 of the membrane bound mucins(44, 51). The principal secreted and membrane bound mucins in the respiratory tract are MUC2, MUC5AC and MUC5B of the secreted mucins and MUC1 and MUC4 of the membrane mucins(44). Other mucins which are found in secretions include MUC4, MUC6, MUC7, and MUC8, however the specific cells which secrete the latter three mucins have not been elucidated (52-54). MUC5AC mucin is secreted exclusively by goblet cells in surface epithelium (53) and MUC5B mucin is secreted by submucosal glands located in the lamina propria (52). MUC2 and MUC4 are secreted by both cell types (52, 55) and MUC2 is also expressed in non-respiratory tissues (jejunum, ileum and colon) (54). Increased levels of MUC2, MUC5AC and MUC5B gene products have been shown in CF compared with normal airways. A cluster of four gel-secreted mucin genes spanning a region of more than 400 kb map to chromosome 11p15.4 to 15.5 and their order is (56-58):

Telomere-MUC6-MUC2-MUC5AC-MUC5B-Centromere

It has been suggested that this gene cluster arose by a series of gene duplications and intragenic duplication events (56, 59). This cluster of mucin genes shows significant homology in their proline-threonine-serine residue regions (PTS regions)(44, 51). Also

noted is that a cluster of membrane bound mucins (MUC3A, MUC3B, MUC11, MUC12 and MUC17) map to chromosome 7q22(44, 51).

Vinall and co-workers investigated these four genes (i.e., MUC2, MUC5AC, MUC5B and MUC6) using the 40 core Centre d'Etude du Polymorphisme Humain (CEPH) families (58). The non-glycosylated protein region of these four genes has been shown to be conserved across these genes and across species and this is believed to reflect the importance of the cysteine residues in inter and intra-molecular cross-linking of mucin proteins (58). MUC5AC and MUC5B showed small inter-individual differences in length variation. A number of variable number tandem repeat (VNTR) polymorphisms were observed within the coding region of the MUC5AC gene. There was no evidence of VNTR polymorphisms in the MUC5B gene (58). MUC5AC and MUC5B have uninterrupted cysteine rich domains (60, 61) and this pattern is hypothesized to stabilize the structure (58) and limit VNTR polymorphisms. MUC2 and MUC6, which are closer to the telomeric side of the cluster, showed large tandem repeat regions that consist of VNTRs and these long stretches of repeats have been hypothesized to be prone to VNTR polymorphisms (58). The authors (58) showed that there was a bimodal distribution of MUC2 alleles with the large alleles being double the size of the small alleles observed in their study group. In one family trio in their study they documented a doubling in MUC2 allele size. It was theorized that the bimodal distribution observed indicates that these polymorphisms arose from duplications of the tandem repeat regions of the genes (58, 62). Recombinations were detected between MUC2 and MUC6 in the CEPH cohort. An intronic VNTR polymorphism in MUC5B (intron 36) together with the large and small alleles of MUC5AC were investigated in the CEPH families in a follow-up study by Vinall and colleagues (63). They showed no linkage disequilibrium between these gene polymorphisms (63). Results from the above studies suggest that polymorphisms in these secreting mucin genes should be considered independent genetic markers (63).

6.1.5 Mucin genes as modifier genes of CF pulmonary disease

In this study we wanted to investigate whether polymorphisms in respiratory secreted gel mucin genes which would lead to increased viscosity were associated with pulmonary disease severity and progression in CF. MUC2, MUC5AC and MUC5B have been shown in

CF airway secretions (64). Mucin genes 2 and 5B (MUC2, MUC5B) were investigated in this study. The main characteristic of all mucin genes is the presence of a central domain composed of a variable number of tandem repeats. Each repeat unit contains multiple sites for O-linked glycosylation. Increased glycosylation would be expected to increase mucus viscosity and slow mucociliary clearance of purulent broncho-pulmonary secretions. The increased viscosity of the mucus can be partially explained by the CFTR defect, changes in mucus producing cells (hypertrophy of submucosal glands, goblet cell hyperplasia and metaplasia), chronic inflammation and bacterial colonization.

The MUC2 gene contains a 69 bp VNTR previously investigated by Vinall and co-workers (63). This gene also contains a region of imperfectly conserved repeats upstream from this VNTR which is a Thr/Ser/Pro rich subdomain (mRNA sequence accession number: NM_002457) (65). These repeats contain multiple glycosylation sites and were previously believed to be monomorphic, but preliminary work in our laboratory has shown that this region is polymorphic. Length variation in the coding sequence of the gene will affect the size of the mucin protein produced (prior to glycosylation) and the larger number repeats will result in the transcription and translation of a larger mucin protein with an increased number of glycosylation sites which will likely have an impact on the properties of the mucin gel. A mucin with a higher number of repeats may have a higher number of glycosylation sites and this may alter the viscoelastic properties of the molecule. Additionally, a consequence of this polymorphism may be to influence the severity of pulmonary disease in CF. Therefore, in the early stages of pulmonary disease progression having a larger mucin protein with increased number of glycosylation sites may help in the removal of pathogen. As severity of pulmonary disease increases this same mucin protein may contribute to disease progression and contribute to stagnant mucus and compromised clearance.

The MUC5B gene did not show VNTR variation in the coding region (58, 60). The MUC5B gene has been shown to contain a highly variable number of tandem repeat region (59 bp/repeat) in intron 36 (genomic DNA accession number: Y09788) (56, 57). This region contains a nuclear factor binding site and may be involved in the regulation and level of expression of this gene (56, 57).

The MUC5B polymorphism and MUC2 VNTR polymorphism were investigated by Vinall and associates in atopic patients with and without asthma, smokers with and without COPD and healthy volunteers (63). In their study they showed an association with longer mean allele length for MUC2 in atopic patients without asthma and concluded that the longer allele is protective against developing asthma (63). The purpose of this study was to investigate the association of two MUC polymorphisms (a) the VNTR (59 bp/repeat) in intron 36 of the MUC5B gene and b) the region of imperfectly conserved repeats rich in Thr/Ser/Pro of the MUC2 gene in CF pulmonary disease severity and progression.

6.2 RESULTS

6.2.1 Hardy Weinberg equilibrium and frequency of genotypes

The polymorphisms investigated in MUC2 and MUC5B were not in HWE equilibrium (p-values were 0.002 and 0.007, respectively). For MUC2, CF patients were characterized by long, medium, or short imperfect repeat alleles and the genotypes observed in our cohort are presented in Figure 6.4. For MUC5B, CF patients are characterized by genotype for the number of repeats on both chromosomes and the data are presented in Figure 6.5.

6.2.2 Descriptive data results and study cohort grouping

This was a multicenter study (N=311) with participating children's (M/F=24/29) and adult CF clinics (M/F=130/128). Demographic and clinical measures for our study cohort are presented in Tables 6.1 and 6.2 grouped for the MUC2 and MUC5B polymorphisms, respectively. Variable measures are presented for the whole study group and as grouped for analysis in our mixed effects models. For MUC2, the CF patients were characterized by long, medium, or short imperfect repeat alleles and were grouped for mixed effects regression analyses based on whether the subjects were homozygous for the long imperfect repeat (N=305, 79.4%) or heterozygous for this repeat with a shorter allele (N=74, 19.2%) (Table 6.1). For MUC5B, CF patients were characterized by genotype for the number of repeats on both chromosomes collapsed into (Table 6.2):

- Two low repeat alleles (i.e., the following genotypes: 2/2 (N=1, 0.3%), 2/7 (N=1, 0.3%), 3/7 (N=15, 4.4%), 3/5 (N=2, 0.6%), 5/5 (N=4, 1.2%), 5/7 (N=39, 11.4%), 5/8 (N=5, 1.5%)), and
- Two long VNTR alleles into two additional groups:
    - Genotype 7/7 (N=223, 65.4%) and
    - Genotypes 7/8 (N=45, 13.2%) or 8/8 (N=1, 0.3%).

There were four CF patients with a 6/9 genotype (N=1, 0.3%) or a 6/7 (N=3, 0.9%) genotype and they were excluded from the mixed effects regression analyses.

6.2.3 Pulmonary disease progression: mixed effects regression on %predFEV$_1$

MUC2 gene. CF patients were grouped into: a) heterozygous / homozygous for 2 repeats and b) the common genotype in the cohort (homozygous for 1 repeat).

Model 6.2.3-A

*%predFEV$_1$ = Time + MUC2 (0/1= homozygous for one repeat) + Sex (0/1=male/female) + MUC2 * Time*

Base group for MUC2 was heterozygous / homozygous for 2 repeats

We showed no difference in pulmonary disease severity (p=0.94) and a similar rate of decline in %predFEV$_1$ based on our MUC2 grouping (i.e., MUC2 * Time, p=0.52), controlling for sex (0/1=male/female, p=0.51). The model was also run including *P. aeruginosa* infection status.

Model 6.2.3-B

*%predFEV$_1$ = Time + MUC2 (0/1= homozygous for one repeat) + MUC2 * Time + PA infection status (0/1=not infected /chronically infected) + MUC2 * Time + PA infection status * MUC2 + PA infection status * Time + PA infection status * MUC2 * Time)*

We did not show a significant association with pulmonary disease severity (p=0.57) or pulmonary disease progression (i.e., MUC2 * Time, p=0.83). We also did not show significant two and three way interactions with *P. aeruginosa* infection status (i.e., PA infection status * MUC2, p=0.46 and PA infection status * MUC2 * Time, p=0.53).

MUC5B gene. CF patients were ultimately grouped based on the common genotype for the cohort into two groups: a) common genotype, which was homozygosity for 7 repeat allele and b) all others. CFTR genotype was not included as it was not well represented in the mucin gene grouping. We showed no difference in pulmonary disease severity and MUC5B genotype (p=0.12).

Model 6.2.3-C

*%predFEV$_1$ = Time + MUC5B (0/1= 7-7 VNTR genotype) + Sex (0/1=male/female) + MUC5B * Time*

Base group for MUC5B was genotype other than 7-7 VNTR.

We did, however, show significant two (PA infection status * MUC5B, p=0.02; PA infection status * Time, p=0.004; MUC5B * Time, p=0.05) and three way interactions (PA

infection status * MUC5B * Time; p=0.02) for MUC5B 7/7 genotype versus our base group defined as other. Further examination of the slopes (i.e., rate of decline in %predFEV$_1$ per year) showed no significant differences (MUB5B other and no PA infection=0.06%/year, MUC5B 7/7 and no PA infection=-1.18%/year, MUC5B other with PA infection=-1.71%/year and MUC5B 7/7 with PA infection=-1.26%/year); the slopes for rate of decline in %predFEV1 were not significantly different from zero.

Model 6.2.3-D

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MUC5B (0/1= 7-7 VNTR genotype) + PA infection status (0/1=not infected /chronically infected) + MUC5B * Time + PA infection status * MUC5B + PA infection status * Time + PA infection status * MUC5B * Time*

MUC2 and MUC5B diplotype. We grouped our cohort into two groups; the common diplotype observed (MUC2-MUC5B/ MUC2-MUC5B=1-7/1-7, N=162) and all others (MUC2-MUC5B/ MUC2-MUC5B=1-3/1-7, 1-5/1-7, 1-7/1-8, 2-7/1-7, N=110). We showed a similar rate of decline in %predFEV$_1$ based on our diplotype grouping. We showed similar level of pulmonary disease severity (i.e., MUC2/5B diplotype main effect, p=0.86) and pulmonary disease progression (i.e., MUC2/5B diplotype * Time, p=0.87) by MUC2/5 diplotype grouping.

Model 6.2.3-E

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MUC2/5B5 diplotype (0/1=1-7/1-7) + MUC2/55 diplotype * Time*

Base group for MUC2/5B diplotype was: 1-7/1-7 repeats

In a second model we also included *P. aeruginosa* infection status, however we did not show a significant association with pulmonary disease severity (i.e. MUC2/5B diplotype * PA status, p=0.23) and progression (i.e., MUC2/5B diplotype * PA infection status * Time, p=0.53).

Model 6.2.3-F

*%predFEV$_1$ = Time + Sex (0/1=male/female) + MUC2/5B diplotype (0/1=1-7/1-7) + PA infection status (0/1=not infected /chronically infected) + MUC2/5B * Time + PA infection status * MUC2/5B + PA infection status * Time + PA infection status * MUC2/5B * Time*

6.2.4. Survival analysis

MUC2 and MUC5B were not predictive of survival. CFTR class, *P. aeruginosa* infection status, sex, current age and modifier gene * *P. aeruginosa* infection status were included in the models. Results are presented in Table 6.3. Lastly we included the MUC2/MUC5B diplotype, as described above, in the model but did not show any differences based on diplotype grouping and progression to our endpoint (time to death or lung transplantation) (RR=1.40, p=0.70; $CI_{0.05}$=0.26 and $CI_{0.95}$=7.39). Also being chronically infected with *P. aeruginosa* (MUC2/MUC5B diplotype * *P. aeruginosa* infection status; RR=1.10, p=0.92; $CI_{0.05}$=0.18 and $CI_{0.95}$=6.65) was not a significant interaction term.

6.2.5 Effect of modifier genes on *P. aeruginosa* infection status

We did not show significant differences in susceptibility to chronic infection with *P. aeruginosa* and *P. aeruginosa* or BCC associated with the MUC2 and MUC5B gene polymorphisms (Tables 6.4 and 6.5). We then investigated whether the polymorphisms in the mucin genes contributed to age of first infection and age of chronic infection with *P. aeruginosa* and the results are presented in Tables 6.6 and 6.7, respectively. CFTR class was included in the models. If there were empty cells for CFTR grouping then the variable PSS was used instead.

We showed that homozygosity for one repeat for the MUC2 polymorphism was associated with earlier first infection with the pathogen *P. aeruginosa*. In addition, fixing %predFEV$_1$ (i.e., two CF patients having the same %predFEV$_1$) those CF patients who were homozygous for one repeat for the MUC2 polymorphism became infected with *P. aeruginosa* for the first time at a later age than CF patients who were heterozygous or homozygous for 2 repeats (Table 6.6, p=0.03).

In the case of the MUC5B polymorphism, we showed that holding CF diagnosis age fixed, being homozygous for 7 repeats was associated with later age of first infection with *P.*

aeruginosa than when the CF patient was grouped other (p=0.02). We did not show a significant effect of MUC2 and MUC5B genotype on age of chronic infection with *P. aeruginosa*. Also we did not show a significant effect of MUC2/5B diplotype on age of first and chronic infection with *P. aeruginosa*.

6.3 DISCUSSION

In this study we investigated variations in the length of the intronic VNTR in the MUC5B gene as a potential modifier gene in CF. We also showed that the MUC2 gene contains a region of imperfectly conserved repeats in the coding region and was investigated as a potential modifier gene in CF pulmonary disease progression. Both genes are expressed in the respiratory tree. Our hypothesis was that variation in the length of the tandem repeats of MUC2 may potentially be associated with more severe pulmonary disease severity and progression in CF. We showed that variations in the number of repeats was not associated in our study cohort with pulmonary disease severity and progression or with survival, even after also considering potential covariates in our models. We showed significant difference in pulmonary disease severity and progression for the MUC5B polymorphism (i.e., common genotype versus others), but the individual slopes were not significantly different. We also investigated whether MUC gene polymorphisms were associated with infection with common CF pathogens, as a measure of disease severity. We did not show differences in the prevalence of chronic infection with *P. aeruginosa* or BCC associated with MUC5B or MUC2 genotype. We did show a later age of first infection in CF patients who had a 7-7 VNTR genotype for MUC5B versus other less common shorter or longer VNTR genotypes. For MUC2, we showed that CF patients who have the common genotype (homozygous for one repeat) had an earlier age of first infection with *P. aeruginosa*, but no difference in age of chronic infection with the pathogen. In addition, CF patients who were homozygous for one repeat for the MUC2 polymorphism who also had similar pulmonary disease severity became infected with the pathogen at a later age.

Increased viscosity of mucus can be partially explained by the CFTR defect, the chronic inflammation and the subsequent changes in mucus producing cells (hypertrophy of submucosal glands, goblet cell hyperplasia and metaplasia). Milder CFTR mutations belonging to class 4 and 5, where there is some functional CFTR protein produced show a milder pulmonary phenotype compared with patients who have CFTR mutations belonging to classes 1, 2 and 3 where there is either no CFTR protein produced or there is defective maturation or blocked activation of the protein (66). CF patients with severe CFTR genotypes were investigated to see if a combination of a longer length VNTR for MUC5B and larger number of incomplete repeats of MUC2 would be associated with more severe

pulmonary disease progression and also affect survival. We hypothesized that polymorphisms which may increase the viscosity of mucus may contribute to earlier repeat infection with common CF pathogens followed by earlier chronic colonization with *P. aeruginosa* or BCC and more rapid decline in %predFEV$_1$. We did not find differences in chronic colonization by *P. aeruginosa* or BCC and MUC2 and MUC5B genotype. We also did not show a significant association of mucin gene polymorphisms when CFTR genotype severity, infection status (i.e., chronic infection) with common CF pathogens (*P. aeruginosa*, BCC) and gender were included in the model to investigate pulmonary disease progression. In our models the milder CFTR genotypes were associated with later age of first infection with *P. aeruginosa* (Table 6.6, MUC 2 analysis)

MUC2, MUC5AC and MUC5B have been shown in respiratory secretions (58, 63, 64, 67). Vinall and colleagues (63) showed that atopic non-asthmatic individuals had a longer MUC2 mean allele length in comparison with asthmatic and COPD patients, and healthy volunteers. The same investigators (63) and others (5, 68) have shown upregulation of MUC2 in response to pathogen invasion and in response to inflammatory cytokines. In CF, Cheng and associates (69) have shown altered sulfation/glycosylation of mucus in cultured CF nasal epithelial cells. Voynow and colleagues (70) investigated MUC5AC, MUC2 and MUC1 mRNA levels in normal and CF nasal epithelial cells and showed altered mucin gene expression in non-inflamed CF nasal cells. It has been hypothesized that MUC2 may be upregulated during inflammatory disease and the length of the MUC2 protein somehow interacts with the more abundant MUC5AC and 5B producing a mucin gel with altered physicochemical properties (63).

Pulmonary disease in CF is characterized by chronic *P. aeruginosa* infection, neutrophilic inflammation, mucus hypersecretion and obstruction of the airways. Viscous mucus and impaired clearance of mucus may contribute to a favorable environment for the creation of biofilms by *P. aeruginosa* leading to chronic colonization with this pathogen (71). *P. aeruginosa* contributes to mucus viscosity and impairment to mucociliary clearance. Li and colleagues have shown that *P. aeruginosa* lipopolysacharide (LPS) upregulates transcription of MUC2 in epithelial cells through inducible enhancer elements and showed that this upregulation can be blocked by tyrosine kinase inhibitors genistein and tyrphostin AG 126 (3). It is likely that soluble *P. aeruginosa* products (e.g., pyocyanin) may impair mucociliary

clearance in part through a pathway that involves neutrophils and is mediated by reactive oxygen species (37, 72). Extracellular GSH depletion, which is documented in CF, contributes to mucin hypersecretion directly and indirectly through the effect of GSH depletion on neutrophils. GSH depletion also contributes to ineffective mucolysis and thus contributes to increased mucus viscosity. The lipid A portion of *P. aeruginosa* LPS and the LPS of other Gram-negative bacteria are all effective in stimulating upregulation of MUC2 and MUC5AC (73). Infection with *P. aeruginosa* would also potentially increase mucus secretion and viscosity. Kishioka and associates (74) showed that *P. aeruginosa* alginate stimulates mucin and lysozyme secretion in ferret tracheal tissue to similar levels as shown in CF sputum. They showed alginate stimulated secretion of mucin was greater than of lysozyme and concluded that this meant that mucin secretion was predominately from mucous versus serous cells (74). This would mean that the mucus secreted would have a higher mechanical impedance (i.e. would be more rigid) and would be more difficult for the CF patient to clear. The viscosity and post-transcriptional modification of secreted mucin is therefore affected by various factors operating in the CF patient. Although the size of the MUC2 repeat will affect the degree of glycosylation, this may have an insignificant effect on mucin viscosity compared with other factors, as described above.

In theory, mucin with a higher number of repeats may have a higher number of glycosylation sites and could have two effects on pulmonary disease progression in CF. Having longer MUC2 alleles would result in a larger protein with an increased number of glycosylation sites that may be beneficial in removal of pathogen in early disease. However a mucin protein with a higher number of glycosylation sites may alter the viscoelastic properties of the molecule and contribute to stagnant mucus and compromised clearance. In more severe pulmonary disease, the effect of more viscid mucus may outweigh any benefits. In our cohort we did not show differences in pulmonary disease progression or *P. aeruginosa* infection status for MUC2. We showed variable pulmonary disease course for CF patients based on MUC5B genotype, when controlling for *P. aeruginosa* infection status. We also showed a variable effect based on MUC2 and independently for MUC5B genotype and age of first infection with *P. aeruginosa*, but no differences for age of chronic infection with *P. aeruginosa*. Our results indicate that the polymorphisms investigated for MUC2 and 5B or polymorphisms which are in LD with these polymorphisms affect in some way pulmonary disease severity and progression. In some cases due to missing data on clinical

variables our study cohort was considerably reduced, therefore study sample size may have been a limitation in our study. Variables for which there was appreciable data unavailable were age of first and chronic infection with *P. aeruginosa*, in addition there were limited patients with mild CFTR genotypes. Therefore, when we controlled for *P. aeruginosa* infection status as well as CFTR genotype, empty cells were created and we were unable to investigate interactions with our gene polymorphisms of interest. Vinall and associates hypothesized that the VNTR in intron 36 may have a functional effect; this region contains a protein binding site and may be involved in the regulation and level of expression of the gene (56, 63). Our preliminary findings suggest that this polymorphism, as well as the MUC2 VNTR should be investigated in a larger CF cohort.

6.4. CONCLUSIONS

We hypothesized that variation in the length of the tandem repeats of the MUC2 gene polymorphism and a higher number of repeats of the VNTR of the MUC5B gene would be associated with more severe pulmonary disease severity and progression in CF. The number of repeats for the MUC5B gene polymorphism was associated in our study cohort with pulmonary disease progression. Although the polymorphisms investigated did not affect survival or the prevalence of chronic infection with *P. aeruginosa* or *Burkholderia cepacia* complex, we did show variability in age of first infection with *P. aeruginosa*. The polymorphisms investigated in this study may have an affect on mucus viscosity, however compared with other factors operating in the CF airways their contribution to the CF pulmonary disease outcome variables investigated is likely minor. Nonetheless, our data suggest that these polymorphisms augment disease course in CF and further examination controlling for CFTR genotype on a larger CF cohort with a greater representation of mild CFTR genotypes is worthy of further investigation.

Figure 6.1. The effect of normal and abnormal ion transport on airway mucociliary clearance. Excess epithelial absorption of sodium and consequently fluid results in decreased volume of the sol phase and ultimately a more concentrated gel phase. Ciliated cells line the surface of airways and by coordinated beating of these cilia transport mucus and entrapped material cranially up to the oropharynx. When the sol phase is reduced cilia are unable to fully extend and beat and therefore mucociliary clearance is compromised. Additionally in CF there is hypersecretion of mucus which further contributes to obstruction.

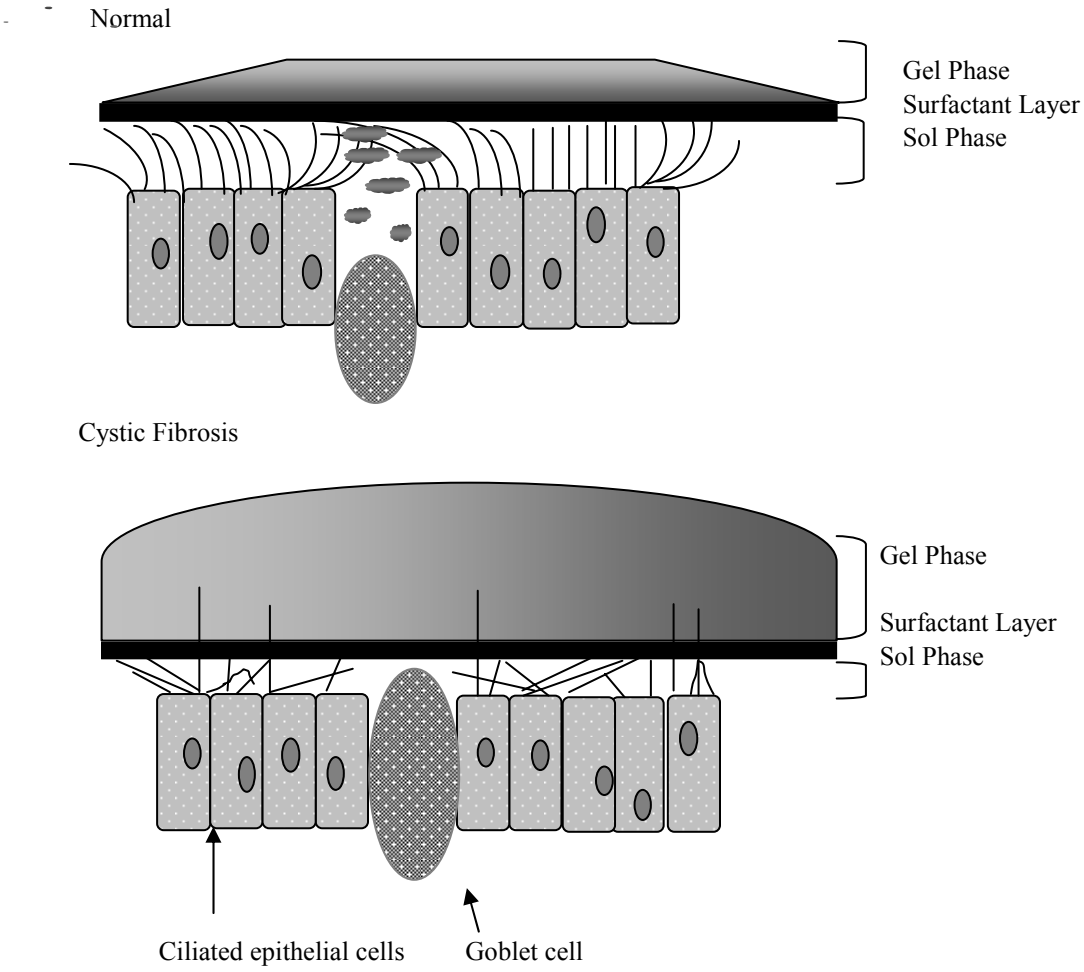


Figure 6.1. The effect of normal and abnormal ion transport on airway mucociliary clearance. Excess epithelial absorption of sodium and consequently fluid results in decreased volume of the sol phase and ultimately a more concentrated gel phase. Ciliated cells line the surface of airways and by coordinated beating of these cilia transport mucus and entrapped material cranially up to the oropharynx. When the sol phase is reduced cilia are unable to fully extend and beat and therefore mucociliary clearance is compromised. Additionally in CF there is hypersecretion of mucus which further contributes to obstruction.

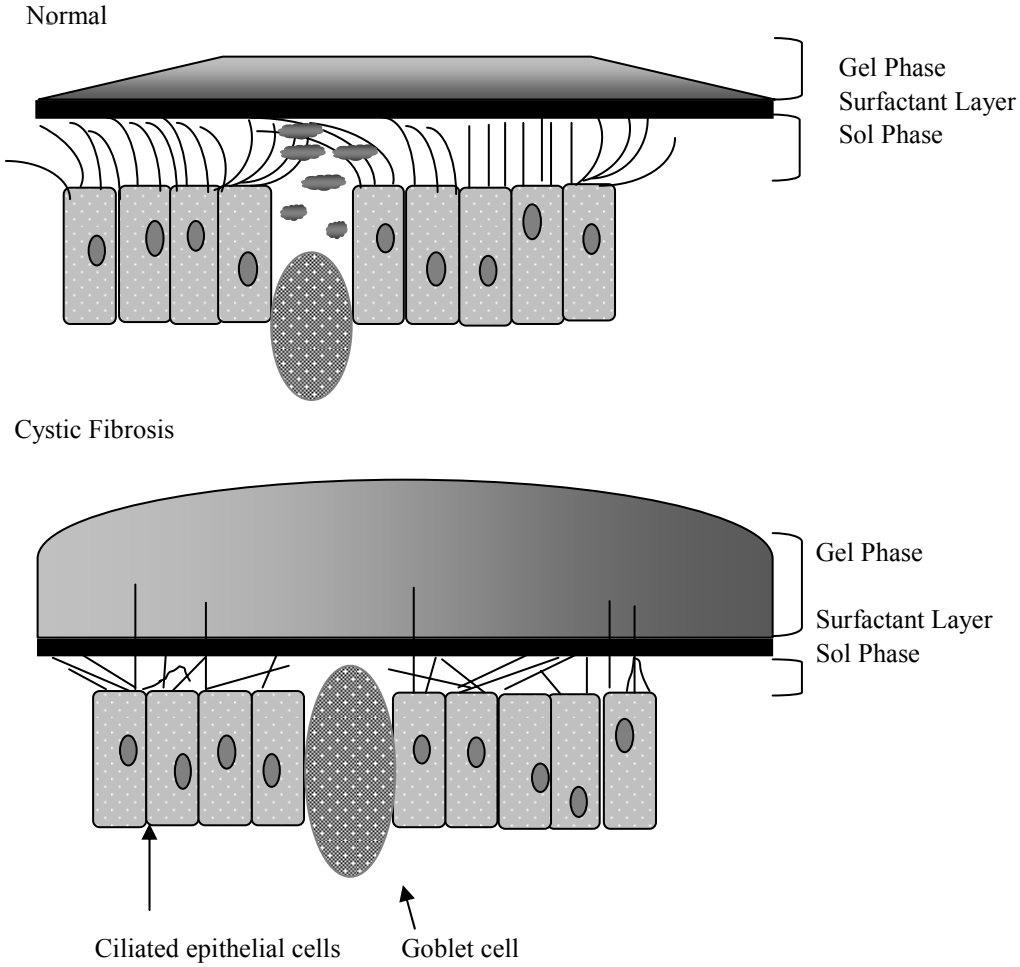


Figure 6.2. Schematic diagram of the four major protein domains in secreted gel-forming mucins. There is a signal sequence in the amino terminus followed by a domain which participates in oligomerization of mucins and a similar domain located further down the structure which also participates in oligomerization of mucins (regions highlighted with asterisks). The next domain consists of a large tandem repeating sequence of amino acid residues that are rich in serine and threonine and O-linked glycosylation occurs at these sites. The remainder of the mucin protein structure includes two unique non-repeating sequences. The depicted regions are not to scale. Individual mucin gene proteins vary within these domains.

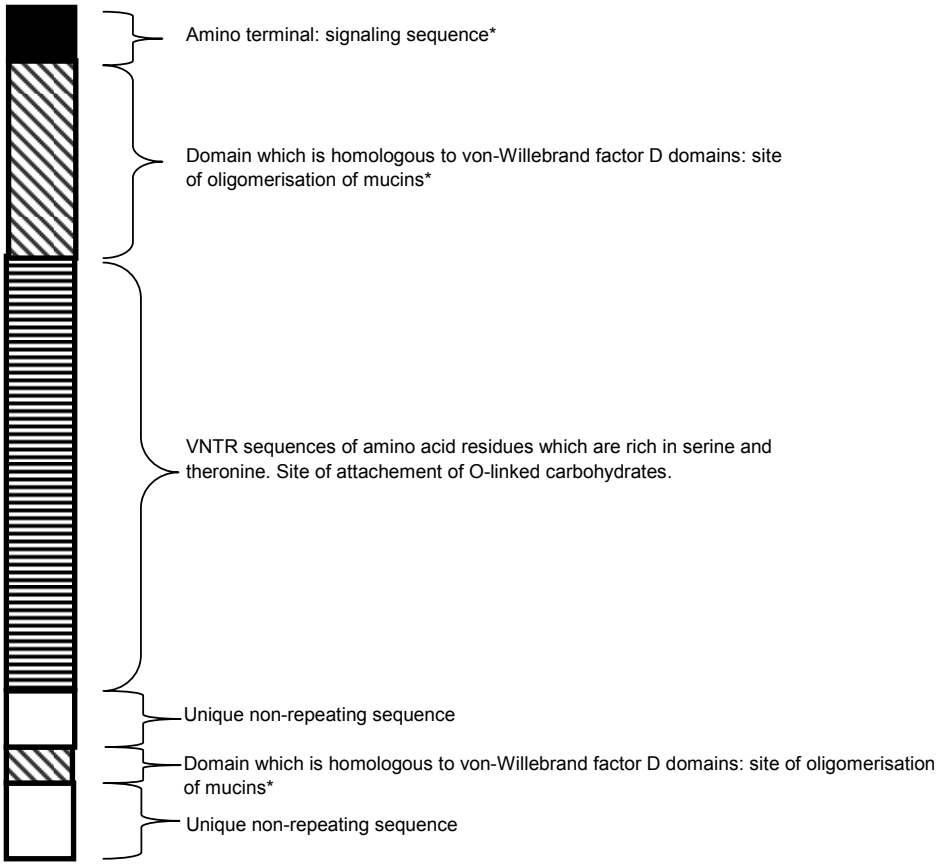


Figure 6.3. Schematic diagram of a mucin molecule. The general structure of a mucin molecule is of a bottle-brush with many (hundreds of) oligosaccharides chains attached to the protein core called apomucin. The oligosaccharide chains link to serine and threonine of the apomucin via N-acetylgalactosamine (GaN) in an alpha-O-glycosidic linkage to the hydroxyl oxygen of the serine or threonine. Each O-linked carbohydrate is defined by a core, a backbone and a periphery (shown in the inset of the figure).F=fucose; G=galactose; Gn=N-acetylglucosamine.

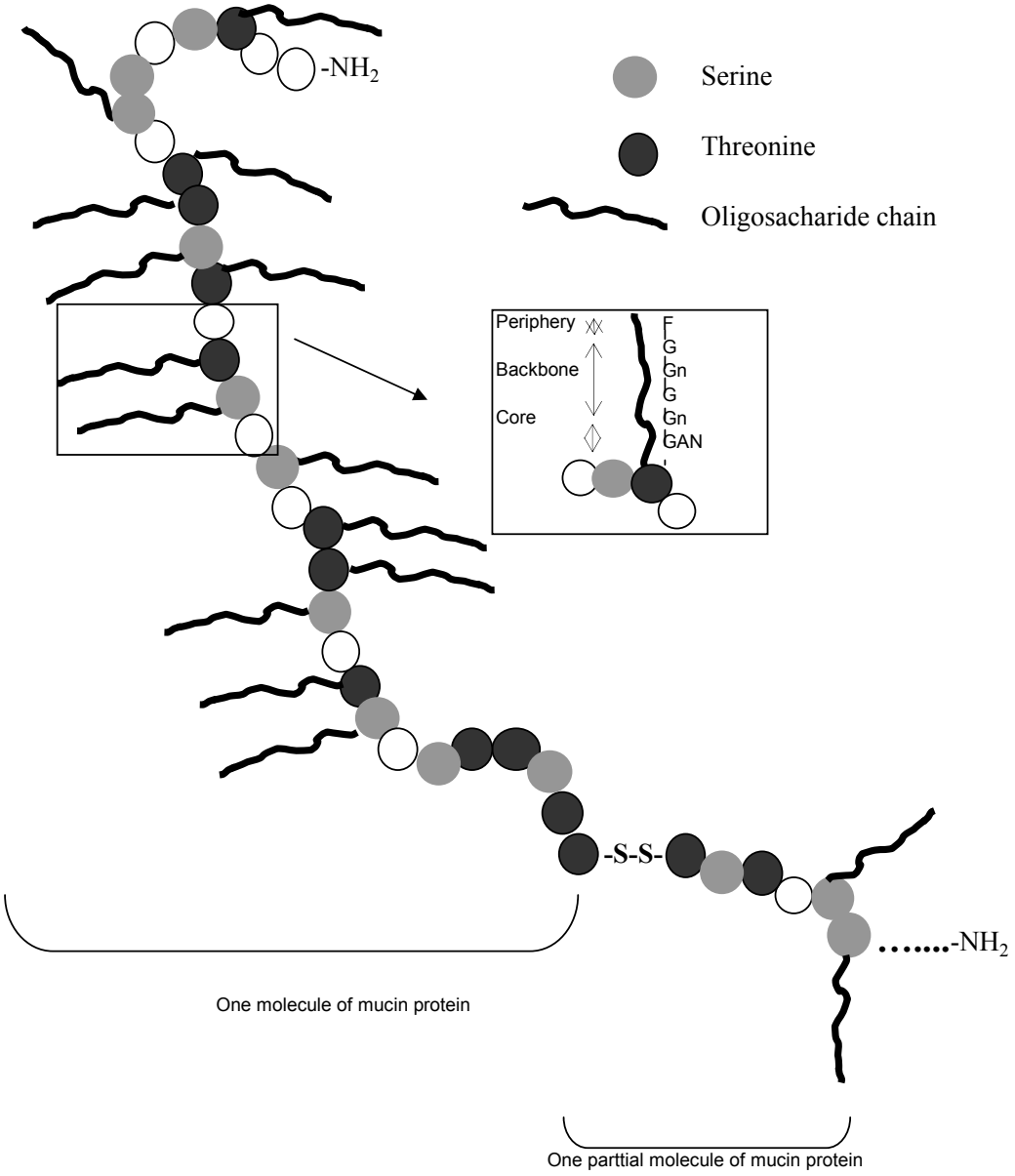


Figure 6.4. Distribution of genotypes for the MUC2 gene polymorphism. The numbers represent the total cohort which was genotyped for this gene and includes patients who were excluded from statistical analyses based on limited numbers of MUC2 genotype or CFTR genotype.

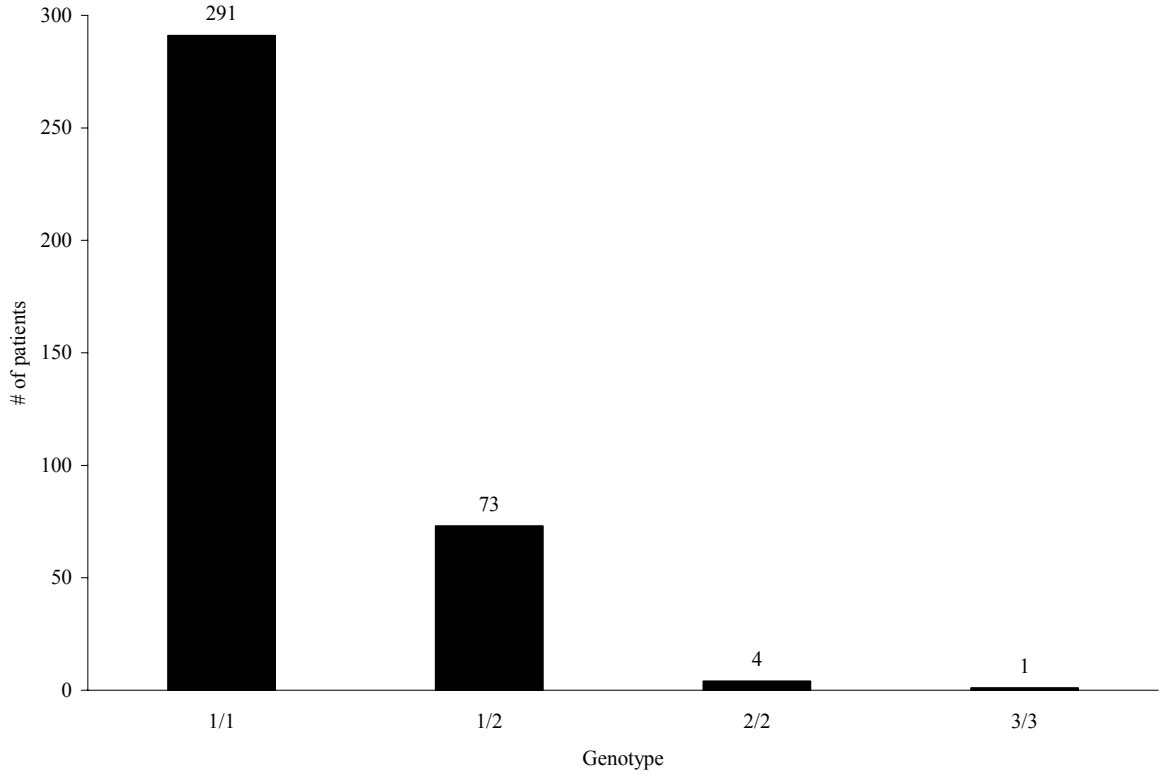


Figure 6.5. Distribution of genotypes for the MUC5B gene polymorphism. The numbers represent the total cohort which was genotyped for this gene and includes patients who were excluded from statistical analyses based on limited numbers of MUC5B genotype or CFTR genotype.

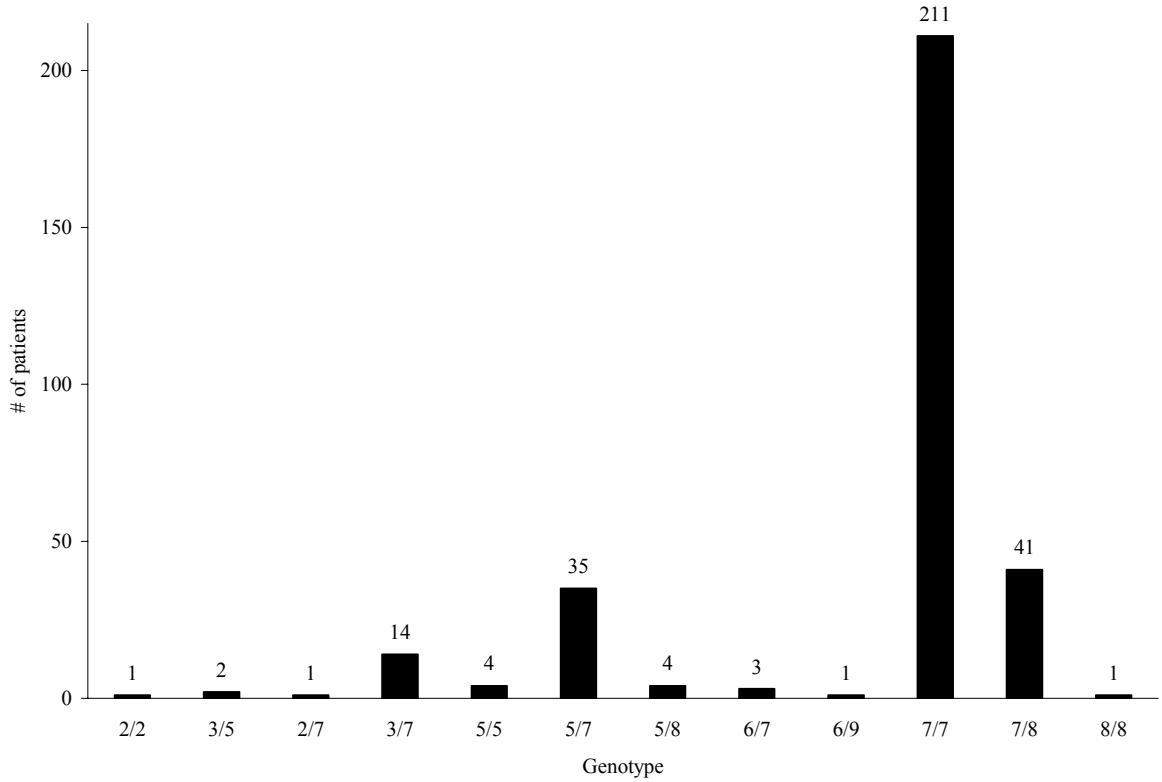


Table 6.1. Clinical characteristics of study cohort by MUC2 as grouped for analyses. The grouping was based on having one or 2 long alleles.

	Total*	2 long alleles	1 long allele	P-value
Current age (yrs)	24.8 (0.5)	25.0(0.6)	23.5(1.2)	0.82
Male/Female	196/180	151/146	40/34	0.09
CF diagnosis age (yrs)	4.5(0.4)	4.6(0.4)	4.2(0.8)	0.21
CFTR genotype*	259/81/28/11	204/65/22/9	51/16/5/2	
PSS (insufficient/sufficient)	332/47	262/38	66/8	0.80
PA status (not infected/infected)	133/238	106/187	25/48	0.94
BCC status (not infected/infected)	338/14	275/25	70/2	0.42
Current status (alive/deceased or transplanted)	316/63	248/52	63/11	0.53
PA1 age (yrs)	14.5(0.8) N=205	14.9(0.9) N=158	13.3(1.7) N=46	0.63
PAC age (yrs)	15.3(0.8) N=140	15.6(0.9) N=106	14.5(2.0) N=33	0.61
Current %predFEV_1 (%)	61.3(1.5)	61.8(1.7)	59.7(3.2)	0.48
Current BMI (kg/m^2)	20.3(0.2)	20.2(0.2)	20.3(0.4)	0.24

*Total measures include subjects with infrequent alleles as well.

** CFTR genotype: class 1, 2, 3 on both chromosomes/ class 1, 2, 3 on one chromosome and unknown or unclassified on the other chromosome/ class 4,5 on one chromosome and 4,5 or 1,2,3 or unknown or unclassified on the other chromosome/unknown or unclassified on both chromosomes. This last group was not included in univariate or mixed effects regression analyses.

Table 6.2. Clinical characteristics of study cohort by MUC5B genotype. Grouping was based the three most common genotypes based on number of 59 bp repeats

	Total*	5/7	7/7	7/8	P-value
Current age (yrs)	26.8 (0.6)	27.9(1.7)	27.2(0.7)	25.4(1.5)	0.41
Male/Female	161/158	19/24	108/103	21/21	0.26
CF diagnosis age (yrs)	4.9(0.5)	5.6(1.2)	4.9(0.6)	4.2(1.0)	0.76
CFTR genotype**	223/67/23/9	29/14/1/0	147/41/16/8	29/9/3/1	
PSS (insufficient/sufficient)	284/38	42/2	183/29	38/4	0.29
PA status (not infected/infected)	91/223	9/34	65/145	11/29	0.70
BCC status (not infected/infected)	298/24	42/2	194/18	40/2	0.69
Current status (alive/deceased or transplanted)	269/53	39/5	173/39	36/6	0.56
PA1 age (yrs)	16.5(0.9) N=164	18.7(2.9) N=22	16.7(1.2) N=108	15.6(2.0) N=26	0.71
PAC age (yrs)	16.6(1.0) N=120	22.5(3.5) N=17	15.6(1.1) N=77	16.6(2.3) N=20	0.32
Current %predFEV_1 (%)	58.1(1.5)	58.1(3.9)	57.4(4.6)	59.6(4.1)	0.93
Current BMI (kg/m^2)	20.6(0.2)	20.6(0.5)	20.6(0.2)	20.6(0.2)	0.85

*Total measures include subjects with infrequent alleles as well.

** CFTR genotype: class 1, 2, 3 on both chromosomes/ class 1, 2, 3 on one chromosome and unknown or unclassified on the other chromosome/ class 4,5 on one chromosome and 4,5 or 1,2,3 or unknown or unclassified on the other chromosome/unknown or unclassified on both chromosomes. This last group was not included in univariate or mixed effects regression analyses.

Table 6.3. Survival analysis of the time to event (death or lung transplantation) for CF patients classified by MUC2 and MUC5B grouping.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MUC2 (N=366)				
MUC2 (2 long alleles=1)	1.14	0.91	0.13	9.63
Sex (Male=0/Female=1)	1.28	0.34	0.77	2.13
Current age	0.49	0.001	0.42	0.56
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown*	0.23	0.17	0.03	1.88
CFTR 1= class 1,2,3 / unknown*	0.54	0.11	0.26	1.15
CFTR 1= class unknown / unknown*	0.17	0.09	0.02	1.32
PA infection status	1.18	0.88	0.14	9.94
MUC2 * PA status	1.66	0.66	0.17	15.99
MUC5B Model (N=314)				
MUC5B (0/1=all others/common genotype 7/7 repeats)	2.05	0.51	0.24	17.53
Sex (Male=0/Female=1)	1.34	0.74	0.77	2.35
Current age	0.48	<0.0001	2.08	0.57
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown*	0.22	0.16	0.03	1.81
CFTR 1= class 1,2,3 / unknown*	0.41	0.05	0.17	0.98
CFTR 1= class unknown / unknown*	0.16	0.09	0.02	1.28
PA infection status	1.42	0.74	0.18	11.59
MUC5B (7/7 and 7/8=0 / 5/7=1) * PA status	1.02	0.99	0.11	9.86

* Base group for CFTR was homozygous class 1, 2, 3.

Table 6.4. MUC2 and MUC5B do not influence susceptibility to chronic infection with *P. aeruginosa* infection.

Model	β	SE(β)	P-value
MUC2 (N=366)			
MUC2 (2 long alleles=1)	-0.07	0.29	0.82
Sex (Male=0/Female=1)	-0.04	0.23	0.85
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown*	-2.70	0.56	1.41×10^{-6}
CFTR 1= class 1,2,3 / unknown*	-0.64	0.27	0.02
CFTR 1= class unknown / unknown*	-0.95	0.65	0.14
MUC5B Model (N=314)			
MUC5B (0/1=all others/common genotype 7/7 repeats)	-0.31	0.29	0.28
Sex (Male=0/Female=1)	0.34	0.27	0.19
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	-2.56	0.53	1.66×10^{-6}
CFTR 1= class 1,2,3 / unknown	-0.75	0.31	0.02
CFTR 1= class unknown / unknown	-0.69	0.76	0.36

* Base group for CFTR was homozygous class 1, 2, 3.

Table 6.5. MUC2 and MUC5B do not influence susceptibility to chronic infection with *P. aeruginosa* or BCC infection.

Model	β	SE(β)	P-value
MUC2 (N=346)			
MUC2 (2 long alleles=1)	0.10	0.29	0.72
Sex (Male=0/Female=1)	-0.09	0.24	0.69
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown*	-2.83	0.56	4.66×10^{-7}
CFTR 1= class 1,2,3 / unknown*	-0.60	0.27	0.03
CFTR 1= class unknown / unknown*	-0.67	0.66	0.31
MUC5B Model (N=314)			
MUC5B (0/1=all others/common genotype 7/7 repeats)	-0.18	0.29	0.54
Sex (Male=0/Female=1)	-0.40	0.27	0.15
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	-2.73	0.54	3.69×10^{-7}
CFTR 1= class 1,2,3 / unknown	-0.78	0.32	0.01
CFTR 1= class unknown / unknown	-0.31	0.84	0.71

* Base group for CFTR was homozygous class 1, 2, 3.

Table 6.6. Effect of MUC2 and MUC5B polymorphisms on age of first *P. aeruginosa* infection.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MUC2 (N=196)				
MUC2 (1= homozygous for 1 repeat)	3.29	0.03	1.13	9.59
Sex (Male=0/Female=1)	1.33	0.13	0.92	1.92
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	0.08	0.01	0.009	0.60
CFTR 1= class 1,2,3 / unknown	0.74	0.18	0.48	1.15
CFTR 1= class unknown / unknown	0.57	0.29	0.20	1.63
CF diagnosis age	1.02	0.62	0.95	1.09
Current %predFEV_1	0.99	0.82	0.98	1.01
Current BMI	1.04	0.17	0.98	1.12
Current age	0.87	$2.2x10^{-16}$	0.84	0.90
MUC2 * Current %predFEV_1	0.98	0.03	0.96	1.00
MUC2 * CF diagnosis age	0.95	0.17	0.88	1.02
MUC5B (N=159)				
MUC5B (1=homozygous for 7 repeats)	2.49	0.11	0.81	7.59
Sex (Male=0/Female=1)	1.25	0.28	0.83	1.89
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	0.06	0.008	0.008	0.49
CFTR 1= class 1,2,3 / unknown	0.74	0.23	0.46	1.20
CFTR 1= class unknown / unknown	0.89	0.83	0.31	2.55
CF diagnosis age	1.02	0.37	0.97	1.08
Current %predFEV_1	1.00	0.90	0.98	1.01
Current BMI	0.99	0.84	0.92	1.07
Current age	0.85	<0.0001	0.82	0.89
MUC5B * Current %predFEV_1	0.99	0.26	0.97	1.01
MUC5B * CF diagnosis age	0.93	0.02	0.88	0.99
MUC2 and 5B diplotype (N=134)				

MUC2/5B diplotype (1=homozygous for 7 repeats)	1.47	0.52	0.45	4.76
Sex (Male=0/Female=1)	1.45	0.11	0.92	2.30
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	0.002	0.42	$9x10^{-10}$	6315.66
CFTR 1= class 1,2,3 / unknown	0.77	0.33	0.45	1.31
CFTR 1= class unknown / unknown	0.86	0.79	0.29	2.58
CF diagnosis age	0.94	0.10	0.88	1.01
Current %predFEV$_1$	0.99	0.10	0.98	1.01
Current BMI	0.98	0.68	0.89	1.08
Current age	0.86	$6.8x10^{-13}$	0.83	0.90
MUC2/5B * Current %predFEV$_1$	-0.009	0.36	0.97	1.01
MUC2/5B * CF diagnosis age	1.01	0.74	0.94	1.01

* Formula

Dependent variables: Age of 1st infection (age of chronic infection), PA infection status (0/1=not infected/chronically infected).

Main effects: Sex + Modifier gene SNP(s) + CFTR class + CF diagnosis age + current %predFEV$_1$ + current BMI + current age

Interactions: Modifier gene* CF diagnosis age, Modifier gene * current %predFEV$_1$.

Table 6.7. Effect of MUC2 and MUC5B polymorphisms on age of chronic *P. aeruginosa* infection.

Model	RR	P-value	Lower $CI_{.95}$	Upper $CI_{.95}$
MUC2 (N=136)				
MUC2 (1= homozygous for 1 repeat)	1.89	0.22	0.68	5.22
Sex (Male=0/Female=1)	0.99	0.96	0.67	1.46
PSS (1=sufficient)$^{+}$	2.01	0.18	0.72	5.58
CF diagnosis age	0.97	0.38	0.92	1.03
Current %predFEV_1	1.00	0.83	0.99	1.04
Current BMI	0.97	0.37	0.90	1.04
Current age	0.87	0.87	0.84	0.90
MUC2 * Current %predFEV_1	0.99	0.15	0.97	1.00
MUC2 * CF diagnosis age	0.98	0.61	0.92	1.05
MUC5B (N=118)				
MUC5B (1=homozygous for 7 repeats)	1.78	0.30	0.60	5.24
Sex (Male=0/Female=1)	1.05	0.84	0.68	1.61
CFTR: Mild 1= homozygous class 4,5 or 4,5/class 1,2,3 or unknown	0.73	0.79	0.07	7.43
CFTR 1= class 1,2,3 / unknown	0.78	0.31	0.48	1.27
CFTR 1= class unknown / unknown	0.80	0.68	0.27	2.34
CF diagnosis age	0.97	0.27	0.91	1.03
Current %predFEV_1	0.99	0.30	0.98	1.01
Current BMI	0.96	0.31	0.88	1.04
Current age	0.85	$1.4x10^{-15}$	0.82	0.88
MUC5B * Current %predFEV_1	1.00	0.63	0.98	1.01
MUC5B * CF diagnosis age	0.99	0.80	0.92	1.06
MUC2 and 5B diplotype (N=95)				
MUC2/5B diplotype (1=homozygous for 7 repeats)	0.69	0.53	0.22	2.19
Sex (Male=0/Female=1)	1.07	0.79	0.67	1.72

PSS	1.58	0.47	0.46	5.56
CF diagnosis age	0.89	0.17	0.82	0.98
Current %predFEV$_1$	0.98	0.07	0.97	1.00
Current BMI	0.95	0.34	0.85	1.06
Current age	0.85	$3.8x10^{-10}$	0.81	0.90
MUC2/5B * Current %predFEV$_1$	1.00	0.97	0.98	1.02
MUC2/5B * CF diagnosis age	1.06	0.29	0.95	1.17

* See previous table for formula

+ PSS used instead of CFTR genotype.

6.5. BIBLIOGRAPHY

1.Boucher, R. C. 2002. An overview of the pathogenesis of cystic fibrosis lung disease. *Adv Drug Deliv Rev* 54(11):1359-71.

2.Boucher, R. C. 2003. Regulation of airway surface liquid volume by human airway epithelia. *Pflugers Arch* 445(4):495-8.

3.Li, J. D., A. F. Dohrman, M. Gallup, S. Miyata, J. R. Gum, Y. S. Kim, J. A. Nadel, A. Prince, and C. B. Basbaum. 1997. Transcriptional activation of mucin by Pseudomonas aeruginosa lipopolysaccharide in the pathogenesis of cystic fibrosis lung disease. *Proceedings of the National Academy of Sciences of the United States of America* 94(3):967-72.

4.Koo, J. S., Y. D. Kim, A. M. Jetten, P. Belloni, and P. Nettesheim. 2002. Overexpression of mucin genes induced by interleukin-1 beta, tumor necrosis factor-alpha, lipopolysaccharide, and neutrophil elastase is inhibited by a retinoic acid receptor alpha antagonist. *Experimental Lung Research* 28(4):315-32.

5.Li, J. D., W. Feng, M. Gallup, J. H. Kim, J. Gum, Y. Kim, and C. Basbaum. 1998. Activation of NF-kappaB via a Src-dependent Ras-MAPK-pp90rsk pathway is required for Pseudomonas aeruginosa-induced mucin overproduction in epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* 95(10):5718-23.

6.Perrais, M., P. Pigny, M. C. Copin, J. P. Aubert, and I. Van Seuningen. 2002. Induction of MUC2 and MUC5AC mucins by factors of the epidermal growth factor (EGF) family is mediated by EGF receptor/Ras/Raf/extracellular signal-regulated kinase cascade and Sp1. *Journal of Biological Chemistry* 277(35):32258-67.

7.Rose, M. C. 1992. Mucins: structure, function, and role in pulmonary diseases.[see comment]. *American Journal of Physiology* 263(4 Pt 1).

8.Gendler, S. J., and A. P. Spicer. 1995. Epithelial mucin genes. *Annual Review of Physiology* 57:607-34.

9.Ramphal, R., C. Carnoy, S. Fievre, J. C. Michalski, N. Houdret, G. Lamblin, G. Strecker, and P. Roussel. 1991. Pseudomonas aeruginosa recognizes carbohydrate chains containing type 1 (Gal beta 1-3GlcNAc) or type 2 (Gal beta 1-4GlcNAc) disaccharide units. *Infection & Immunity* 59(2):700-4.

10.Ramphal, R., and M. Pyle. 1983. Evidence for mucins and sialic acid as receptors for Pseudomonas aeruginosa in the lower respiratory tract. *Infection & Immunity* 41(1):339-44.

11.Scharfman, A., S. Degroote, J. Beau, G. Lamblin, P. Roussel, and J. Mazurier. 1999. Pseudomonas aeruginosa binds to neoglycoconjugates bearing mucin carbohydrate determinants and predominantly to sialyl-Lewis x conjugates. *Glycobiology* 9(8):757-64.

12.Krivan, H. C., D. D. Roberts, and V. Ginsburg. 1988. Many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence GalNAc beta 1-4Gal found in some glycolipids. *Proceedings of the National Academy of Sciences of the United States of America* 85(16):6157-61.

13.Lamblin, G., M. Lhermitte, A. Klein, N. Houdret, A. Scharfman, R. Ramphal, and P. Roussel. 1991. The carbohydrate diversity of human respiratory mucins: a protection of the underlying mucosa? *American Review of Respiratory Disease* 144(3 Pt 2).

14.Ramphal, R. 1990. The role of bacterial adhesion in cystic fibrosis including the staphylococcal aspect. *Infection* 18(1):61-4.

15.Reddy, M. S., J. M. Bernstein, T. F. Murphy, and H. S. Faden. 1996. Binding between outer membrane proteins of nontypeable Haemophilus influenzae and human nasopharyngeal mucin. *Infection & Immunity* 64(4):1477-9.

16.Krivan, H. C., V. Ginsburg, and D. D. Roberts. 1988. Pseudomonas aeruginosa and Pseudomonas cepacia isolated from cystic fibrosis patients bind specifically to gangliotetraosylceramide (asialo GM1) and gangliotriaosylceramide (asialo GM2). *Archives of Biochemistry & Biophysics* 260(1):493-6.

17.Lillehoj, E. R., and K. C. Kim. 2002. Airway mucus: its components and function. *Arch Pharm Res* 25(6):770-80.

18.Rubin, B. K. 2002. Physiology of airway mucus clearance. *Respir Care* 47(7):761-8.

19.Rubin, B. K. 1998. Surface Pproperties of Respiratory Secretions: Relationship to Mucus Transport. *In* G. L. Baum, Z. Priel, V. Roth, N. Liron and E. J. Ostfeld, editors. Cilia, Mucus, and Mucociliary Interactions. Marcel Dekker Inc., New York. 317-24.

20.Morgenroth, K., and J. Bolz. 1985. Morphological features of the interaction between mucus and surfactant on the bronchial mucosa. *Respiration* 47(3):225-31.

21.Hayez, R., and P. Roussel. 1976. Bronchial mucus: physical and biochemical features. *In* E. B. Weiss and S. M.S., editors. Bronchial Asthma: Mechanisms and Therapeutics. Little, Brown and Company, Boston. 409-22.

22.Lopez-Vidriero, M. T. 1981. Airway mucus; production and composition. *Chest* 80(6 Suppl):799-804.

23.Kim, S. K., C. E. Nasjleti, and S. S. Han. 1972. The secretion processes in mucous and serous secretory cells of the rat sublingual gland. *Journal of Ultrastructure Research* 38(3):371-89.

24.Puchelle, E., A. Beorchia, M. Menager, J. M. Zahm, and D. Ploton. 1991. Three-dimensional imaging of the mucus secretory process in the cryofixed frog respiratory epithelium. *Biology of the Cell* 72(1-2):159-66.

25.Puchelle, E., S. de Bentzmann, and J. M. Zahm. 1995. Physical and functional properties of airway secretions in cystic fibrosis--therapeutic approaches. *Respiration* 1:2-12.

26.Verdugo, P. 1990. Goblet cells secretion and mucogenesis. *Annual Review of Physiology* 52:157-76.

27.Verdugo, P. 1991. Mucin exocytosis. *Am Rev Respir Dis* 144:S33-S37.

28.Feldman, C. 2001. Nonspecific Host Defenses: Mucociliary clearance and cough. *In* M. S. Niederman, G. A. Sarosi and J. Glassroth, editors. Respiratory Infections, 2 ed. Lippitcott Williams and Wilkins, Philadelphia. 13-26.

29.Widdicombe, J. H., and J. G. Widdicombe. 1995. Regulation of human airway surface liquid. *Respiration Physiology* 99(1):3-12.

30.Brown, H. A., E. R. Lazarowski, R. C. Boucher, and T. K. Harden. 1991. Evidence that UTP and ATP regulate phospholipase C through a common extracellular 5'-nucleotide receptor in human airway epithelial cells. *Molecular Pharmacology* 40(5):648-55.

31.Knowles, M. R., L. L. Clarke, and R. C. Boucher. 1991. Activation by extracellular nucleotides of chloride secretion in the airway epithelia of patients with cystic fibrosis.[see comment]. *New England Journal of Medicine* 325(8):533-8.

32.Mason, S. J., A. M. Paradiso, and R. C. Boucher. 1991. Regulation of transepithelial ion transport and intracellular calcium by extracellular ATP in human normal and cystic fibrosis airway epithelium. *British Journal of Pharmacology* 103(3):1649-56.

33.Puchelle, E., J. M. Polu, J. M. Zahm, and P. Sadoul. 1980. Role of the rheological properties of bronchial secretions in the mucociliary transport at the bronchial surface. *European Journal of Respiratory Diseases Supplement* 111:29-34.

34.Smith, J. J., P. H. Karp, and M. J. Welsh. 1994. Defective fluid transport by cystic fibrosis airway epithelia. *J Clin Invest* 93(3):1307-11.

35.Perez-Vilar, J., and R. C. Boucher. 2004. Reevaluating gel-forming mucins' roles in cystic fibrosis lung disease. *Free Radic Biol Med* 37(10):1564-77.

36.Salathe, M., M. M. Pratt, and A. Wanner. 1993. Protein kinase C-dependent phosphorylation of a ciliary membrane protein and inhibition of ciliary beating. *Journal of Cell Science* 106(Pt 4):1211-20.

37.Seybold, Z. V., W. M. Abraham, H. Gazeroglu, and A. Wanner. 1992. Impairment of airway mucociliary transport by Pseudomonas aeruginosa products. Role of oxygen radicals. *Am Rev Respir Dis* 146(5 Pt 1):1173-6.

38.Vandivier, R. W., V. A. Fadok, C. A. Ogden, P. R. Hoffmann, J. D. Brain, F. J. Accurso, J. H. Fisher, K. E. Greene, and P. M. Henson. 2002. Impaired clearance of apoptotic cells from cystic fibrosis airways. *Chest* 121(3 Suppl).

39.Vandivier, R. W., V. A. Fadok, P. R. Hoffmann, D. L. Bratton, C. Penvari, K. K. Brown, J. D. Brain, F. J. Accurso, and P. M. Henson. 2002. Elastase-mediated phosphatidylserine receptor cleavage impairs apoptotic cell clearance in cystic fibrosis and bronchiectasis. *Journal of Clinical Investigation* 109(5):661-70.

40.Shak, S., D. J. Capon, R. Hellmiss, S. A. Marsters, and C. L. Baker. 1990. Recombinant human DNase I reduces the viscosity of cystic fibrosis sputum. *Proceedings of the National Academy of Sciences of the United States of America* 87(23):9188-92.

41.Vasconcellos, C. A., P. G. Allen, M. E. Wohl, J. M. Drazen, P. A. Janmey, and T. P. Stossel. 1994. Reduction in viscosity of cystic fibrosis sputum in vitro by gelsolin. *Science* 263(5149):969-71.

42.Lamblin, G., J. P. Aubert, J. M. Perini, A. Klein, N. Porchet, P. Degand, and P. Roussel. 1992. Human respiratory mucins. *European Respiratory Journal* 5(2):247-56.

43.Voynow, J. A. 2002. What does mucin have to do with lung disease? *Paediatric Respiratory Reviews* 3(2):98-103.

44.Ali, M. S., and J. P. Pearson. 2007. Upper airway mucin gene expression: a review. *Laryngoscope* 117(5):932-8.

45.Lamblin, G., S. Degroote, J. M. Perini, P. Delmotte, A. Scharfman, M. Davril, J.-M. Lo-Guidice, N. Houdret, V. Dumur, A. Klein, and P. Roussel. 2001. Human airway mucin glycosylation: A combinatory of carbohydrate determinants which vary in cystic fibrosis. *Glycoconjugate J.* 18:661-684.

46.Majima, Y. 2002. Mucoactive medications and airway disease. *Paediatric Respiratory Reviews* 3(2):104-9.

47.Strous, G. J., and J. Dekker. 1992. Mucin-type glycoproteins. *Critical Reviews in Biochemistry & Molecular Biology* 27(1-2):57-92.

48.Scharfman, A., H. Kroczynski, C. Carnoy, E. Van Brussel, G. Lamblin, R. Ramphal, and P. Roussel. 1996. Adhesion of Pseudomonas aeruginosa to respiratory mucins and expression of mucin-binding proteins are increased by limiting iron during growth. *Infection & Immunity* 64(12):5417-20.

49.Scharfman, A., E. Van Brussel, N. Houdret, G. Lamblin, and P. Roussel. 1996. Interactions between glycoconjugates from human respiratory airways and Pseudomonas aeruginosa. *American Journal of Respiratory & Critical Care Medicine* 154(4 Pt 2).

50.Zhang, Y., B. Doranz, J. R. Yankaskas, and J. F. Engelhardt. 1995. Genotypic analysis of respiratory mucous sulfation defects in cystic fibrosis. *Journal of Clinical Investigation* 96(6):2997-3004.

51.Dekker, J., J. W. Rossen, H. A. Buller, and A. W. Einerhand. 2002. The MUC family: an obituary. *Trends Biochem Sci* 27(3):126-31.

52.Audie, J. P., A. Janin, N. Porchet, M. C. Copin, B. Gosselin, and J. P. Aubert. 1993. Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization. *Journal of Histochemistry & Cytochemistry* 41(10):1479-85.

53.Jany, B. H., M. W. Gallup, P. S. Yan, J. R. Gum, Y. S. Kim, and C. B. Basbaum. 1991. Human bronchus and intestine express the same mucin gene. *Journal of Clinical Investigation* 87(1):77-82.

54.Ho, S. B. 1998. Tissue Distribution of Secretory Mucins and Alterations of Mucin Gene Expression in Chronic H. Pylori Infection of the Stomach. *In* G. L. Baum, Z. Priel, V. Roth, N. Liron and E. J. Ostfeld, editors. Cilia, Mucus, and Mucociliary Interactions. Marcel Dekker Inc., New York. 353-66.

55.Copin, M. C., M. P. Buisine, L. Devisme, X. Leroy, F. Escande, B. Gosselin, J. P. Aubert, and N. Porchet. 2001. Normal respiratory mucosa, precursor lesions and lung carcinomas: differential expression of human mucin genes. *Frontiers in Bioscience* 6(75):1.

56.Pigny, P., V. Guyonnet-Duperat, A. S. Hill, W. S. Pratt, S. Galiegue-Zouitina, M. C. d'Hooge, A. Laine, I. Van-Seuningen, P. Degand, J. R. Gum, Y. S. Kim, D. M. Swallow, J. P. Aubert, and N. Porchet. 1996. Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* 38(3):340-52.

57.Pigny, P., I. Van Seuningen, J. L. Desseyn, S. Nollet, N. Porchet, A. Laine, and J. P. Aubert. 1996. Identification of a 42-kDa nuclear factor (NF1-MUC5B) from HT-29 MTX

cells that binds to the 3' region of human mucin gene MUC5B. *Biochemical & Biophysical Research Communications* 220(1):186-91.

58.Vinall, L. E., A. S. Hill, P. Pigny, W. S. Pratt, N. Toribara, J. R. Gum, Y. S. Kim, N. Porchet, J. P. Aubert, and D. M. Swallow. 1998. Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Human Genetics* 102(3):357-66.

59.Desseyn, J. L., M. P. Buisine, N. Porchet, J. P. Aubert, P. Degand, and A. Laine. 1998. Evolutionary history of the 11p15 human mucin gene family. *Journal of Molecular Evolution* 46(1):102-6.

60.Desseyn, J. L., V. Guyonnet-Duperat, N. Porchet, J. P. Aubert, and A. Laine. 1997. Human mucin gene MUC5B, the 10.7-kb large central exon encodes various alternate subdomains resulting in a super-repeat. Structural evidence for a 11p15.5 gene family. *Journal of Biological Chemistry* 272(6):3168-78.

61.Desseyn, J. L., J. P. Aubert, I. Van Seuningen, N. Porchet, and A. Laine. 1997. Genomic organization of the 3' region of the human mucin gene MUC5B. *Journal of Biological Chemistry* 272(27):16873-83.

62.Pratt, W. S., I. Islam, and D. M. Swallow. 1996. Two additional polymorphisms within the hypervariable MUC1 gene: association of alleles either side of the VNTR region. *Annals of Human Genetics* 60(Pt 1):21-28.

63.Vinall, L. E., J. C. Fowler, A. L. Jones, H. J. Kirkbride, C. de Bolos, A. Laine, N. Porchet, J. R. Gum, Y. S. Kim, F. M. Moss, D. M. Mitchell, and D. M. Swallow. 2000. Polymorphism of human mucin genes in chest disease: possible significance of MUC2. *American Journal of Respiratory Cell & Molecular Biology* 23(5):678-86.

64.Davies, J. R., N. Svitacheva, I. Lannefors, R. Kornfolt, and I. Carlstedt. 1999. Identification of MUC5B, MUC5AC and small amounts of MUC2 mucins in cystic fibroisis airway secretions. *Biochem. J.* 344:321-330.

65.Gum, J. R. J. 1992. Mucin genes and the proteins they encode: structure, diversity, and regulation. *Am J Respir Cell Mol Biol.* 7(6):557-64.

66.McKone, E. F., S. S. Emerson, K. L. Edwards, and M. L. Aitken. 2003. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet* 361:1671-76.

67.Hovenberg, H. W., J. R. Davies, C. J. Hermann, C. J. Linden, and I. Carlstedt. 1996. MUC5B, but not MUC2, is a prominent mucin in respiratory secretions. *Glycoconjugate J.* 13:839-847.

68.Dabbagh, K., K. Takeyama, H. M. Lee, I. F. Ueki, J. A. Lausier, and J. Nadel. 1999. Il-4 induces mucin gene expression and goblet cell metaplasia in vitro and in vivo. *Journal of Immunology* 162:6233-6237.

69.Boat, T. F., and P. W. Cheng. 1989. Epithelial cell dysfunction in cystic fibrosis: implications for airways disease. *Acta Paediatr Scand Suppl.* 363:25-30.

70.Voynow, J. A., D. M. Selby, and M. C. Rose. 1998. Mucin gene expression (MUC1, MUC2, and MUC5/AC) in nasal epithelial cells of cystic fibrosis, allergic rhinitis, and normal individuals. *Lung* 176:345-354.

71.Boucher, R. C. 2001. Pathogenesis of cystic fibrosis airways disease. *Trans Am Clin Climatol Assoc* 112:99-107.

72.Jackowski, J. T., Z. Szepfalusi, D. A. Wanner, Z. Seybold, M. W. Sielczak, I. T. Lauredo, T. Adams, W. M. Abraham, and A. Wanner. 1991. Effects of P. aeruginosa-derived bacterial products on tracheal ciliary function: role of O2 radicals. *Am J Physiol* 260(2 Pt 1):L61-7.

73.Basbaum, C. B., and J.-D. Li. 1998. Mucin Synthesis is Upregulated at the Transcriptional Level by Pseudomonas Aeruginosa LPS. *In* G. L. Baum, Z. Priel, V. Roth, N. Liron and E. J. Ostfeld, editors. Cilia, Mucus, and Mucociliary Interactions. Marcel Dekker Inc., New York. 349-52.

74.Kishioka, C., R. P. Tomkiewicz, D. J. Hassett, and B. K. Rubin. 1998. Effect of Pseudomonas aeruginosa Alginate on Airway Mucus Secretion. *In* G. L. Baum, Z. Priel, V. Roth, N. Liron and E. J. Ostfeld, editors. Cilia, Mucus, and Mucociliary Interactions. Marcel Dekker Inc., New York. 367-72.

CHAPTER 7: CONCLUDING CHAPTER

7.0 OVERALL CONCLUSIONS

The chapter summarizes the study's findings, strength and weaknesses of the study, implications to therapy and future directions. There have been a number of studies published during this write-up and they are reviewed and discussed, where relevant, with this study's findings.

7.1 CONCLUSIONS FOR MODIFIER GENES INVESTIGATED

Cystic fibrosis is a systemic disorder transmitted as an autosomal recessive trait (1). Knowledge of the gene which causes the disorder has helped in the understanding of the function of the gene product and advances in therapeutic regimens have increased the life expectancy of CF patients, but have not cured them of the genetic disorder. The classic CF phenotype includes pulmonary disease, pancreatic exocrine insufficiency and abnormal sweat gland function. The CF gene is expressed in the epithelial cells of many organs including the sweat glands, pancreas, lungs, gastrointestinal, and reproductive tracts. While the CFTR mutation affects numerous organs, it is mainly the effect of the disorder on the pulmonary system that affects survival. Patients who have CF show a heterogeneous pulmonary disease course and only a portion of this variability can be explained by the individual's CFTR genotype. Although some CFTR genotypes are associated with less severe disease, patients possessing the same genotype show great variation in pulmonary disease severity and progression (1). There is no doubt that environment plays a role in explaining some of the variability in disease severity and progression. We hypothesized that other genes may play a significant role in explaining this wide range of clinical outcomes and pulmonary disease progression. Our main hypothesis was that genes other than CFTR contribute to the heterogeneity we observe in pulmonary disease severity and progression. Polymorphisms in secondary genes may act in concert to positively or negatively contribute to disease severity and progression in CF. We selected genes to investigate which had known polymorphisms and which, based on their protein function, could be involved in the pathogenesis of pulmonary disease in CF. We selected genes which modulate the inflammatory response to infection; innate immunity genes (MBL2, SPA-1 SPA-2 and

SPD), the antiprotease gene alpha-1-antitrypsin (α_1-AT), genes that affect levels of glutathione (the GAC trinucleotide repeat of the GCLC gene) and antioxidant response (gene deletion polymorphisms in GSTM1 and T1, GSTP1 (Ile105Val)) and genes that could affect viscosity of CF mucus (MUC2 and MUC5B). We will summarize our findings below for each gene category following a brief introduction to our rationale for selecting these genes.

α_1-AT. In CF, infection provokes a vigorous local inflammatory response that prevents the spread of the infection beyond the lung, however by doing so it likely contributes to lung destruction (2-4). The inflammatory process in response to pulmonary infection in CF airways is characterized by a massive influx of neutrophils (5). Neutrophils represent 95% of the cells recovered in bronchoalveolar lavage fluid (BALF) in CF adults and children compared to 5% in normal subjects (6). Neutrophils contain a number of proteolytic enzymes one of which, neutrophil elastase (NE), has been implicated in the excessive pulmonary damage observed in cigarette smokers and in patients with CF. NE is capable of causing direct lung damage but also indirect damage by enhancing mucous secretion (7, 8), and by interfering with the opsonization and elimination of bacterial pathogens, particularly *P. aeruginosa* (7, 9). In normal hosts, the actions of NE are prevented primarily by alpha-1-antitrypsin (α_1-AT), a serine protease inhibitor that binds to NE and inhibits the breakdown of elastic tissue in the lung. Three polymorphisms in the α_1-AT gene which affect the plasma levels of this protein were investigated in this study. We showed that α_1-AT genotype is not a major contributor to the variability of pulmonary disease severity in CF. Specifically α_1-AT genotype did not correlate with %predFEV_1, pulmonary infections and risk of death or lung transplantation.

At first glance it would seem that low levels of this antiprotease do not contribute to disease progression. Pancreatic insufficiency and recurrent infections do however take a toll on CF patients and poor nutritional status is characteristic. On a sub cohort of adult CF patients we monitored α_1-AT plasma levels during a pulmonary infection requiring IV therapy and levels were also measured during a period of stable clinical status. CF patients showed a high baseline α_1-AT level (stable clinical status) and a blunted acute phase increase in α_1-AT levels. Our analysis showed that α_1-AT levels are affected by nutritional status (as measured by body mass index); CF patients with poor nutritional status had low α_1-AT

plasma levels during the acute phase response and for the duration of the IV therapy interval/pulmonary infection. It is likely that a combination of poor nutritional status and chronic pulmonary bacterial infection in CF patients with severe pulmonary disease contributes to the pattern shown in the stable clinical periods and the blunted acute rise in α_1-AT levels and likely negates any effect we would see and attribute to α_1-AT deficiency genotype.

We therefore conclude that α_1-AT may be a modifier of CF pulmonary disease, but that differences in levels of α_1-AT are also dependent on other factors which likely obscure the effect of genotype on α_1-AT levels in CF.

Innate immunity genes. CF pulmonary disease is characterized early on by recurrent infections with the pathogens *Staphylococcus aureus* and *H. influenzae* followed by chronic infection with *Pseudomonas aeruginosa* (2, 3) and in some cases with *Burkholderia cepacia* complex (BCC). Innate first defenses are involved in eradicating pathogens invading the airways and therefore are important in preventing colonization and possibly modulating the inflammatory response. We investigated four innate immunity genes (mannose binding lectin (MBL2) and pulmonary surfactant A (SPA-1 and 2) and D) which have been shown: a) to bind to CF pathogens, b) have polymorphisms resulting in reduced levels of the protein, and/or c) literature showing a deficient state affecting pulmonary status.

In many of our models investigating pulmonary disease progression we reported a significant *P. aeruginosa* x Time interaction, showing a significantly steeper decline in pulmonary function over time in CF patients chronically infected with *P. aeruginosa* versus non-infected CF patients. We also showed a significantly steeper decline in pulmonary function over time in CF patients chronically infected with both BCC and *P. aeruginosa* versus being infected with only one pathogen or neither pathogen. Therefore preventing or delaying chronic infection with these pathogens would potentially slow down pulmonary disease progression in CF. The MBL2 protein has been shown to bind to two known CF pathogens *Staphylococcus aureus* and BCC. The pulmonary surfactant genes A and D produce proteins that bind to the common CF colonizer, *P. aeruginosa*. The scenario for the progression of pulmonary disease in CF considering the knowledge obtained from our

findings and other studies(10, 11) would suggest that early on CF patients who are MBL2 deficient would be prone to experience more frequent infections with *Staphylococcus aureus*. These infections would, in turn, contribute to the chronic inflammatory state characterized by a massive influx of neutrophils (5) and release of proteolytic enzymes including neutrophil elastase (NE).

We did not show differences in pulmonary disease progression with MBL2 deficiency when controlling only for CFTR genotype, gender and *P. aeruginosa* infection status. When the pre- and post-acquisition intervals for BCC were distinguished in the model, MBL2 deficiency was shown to be a predictor of pulmonary disease progression following chronic infection with BCC in CF patients. Contrary to the previous literature, wild-type MBL2 genotype was associated with more severe decline in %predFEV_1 over time following colonization with BCC, but was not associated with increased susceptibility to BCC infection. Interestingly, when including BCC infection status in the model (as a categorical variable: chronically infected versus not infected) and not distinguishing the pre- and post-acquisition interval for BCC infection we did not show differences in pulmonary disease progression with MBL2 deficiency. Further studies incorporating greater representation of BCC genomovars, larger numbers of subjects in whom there is characterization of MBL2 genotypes and longer prospective follow-up are required to expand upon our findings.

We showed a varied response for the pulmonary surfactant gene polymorphisms depending on the response variable investigated. Our results show that in CF, these genes and their polymorphisms impact on the effect that CFTR dysfunction has on the respiratory system. The interaction between the CFTR mutations (mild and severe) and variant alleles in the innate immune genes had a complex effect on the clinical course of CF patients. The measure of disease severity used to study modifier genes in CF will also potentially influence the significance of such associations.

GSTs and GCLC genes. Deficits in reduced glutathione (GSH) levels in ASL fluid (5-10% of normal) and plasma levels (50% of normal) have been reported in CF patients (12). A systemic deficiency in GSH has been shown in CF and other chronic diseases. This extracellular deficiency has been purported to lead to oxidant damage of airway epithelial cells and lung fibrosis, and is associated with reduced pulmonary function and increased

adhesion of bacteria to airway epithelial cells (13-15). In this study we investigated gene polymorphisms which would affect levels of glutathione produced (the GAC trinucleotide repeat of the GCLC gene) and antioxidant response (gene deletion polymorphisms in GSTM1 and T1) and a coding variant GSTP1 (Ile105Val)) as potential modifier genes in CF. We showed that polymorphisms in genes which affect levels of glutathione produced, released and utilized are potential modifier genes in CF pulmonary disease. We showed that the deletion polymorphisms (GSTM1 and T1) contribute to an earlier event (death/lung transplantation). However, we did not show differences in pulmonary disease severity or progression. We showed similar pulmonary disease severity and progression in CF patients who are homozygous or heterozygous for isoleucine for the GSTP1 (Ile105Val) polymorphism. One of the limitations of this study may be the small number of patients with mild CFTR mutations, which prevented us from exploring this group in our analyses. Whether partially functioning CFTR would be adequate for GSH transport is unclear in the literature. We hypothesized that patients with some CFTR function who do not have the deletion polymorphism for GSTM1 and GSTT1 would have better pulmonary function and a lower rate of decline in pulmonary function over time compared with CF patients who have the deletion polymorphisms for GSTM1 and GSTT1. CF patients who have two severe CFTR mutations (i.e., class I, II or III) where there is no chloride conductance will also show no GSH transport through the mutant CFTR, therefore the relevance of the GST polymorphisms would likely be realized only in the mild CFTR group. The potential effect of these polymorphisms in CF patients with mild CFTR is discussed further below.

We hypothesized that CF patients who have a lower number of GCLC GAC repeats will show worse pulmonary disease severity and progression. We showed a significant difference in pulmonary disease severity and survival in CF patients with a 7/8 GAC repeat genotype versus 7/7 GAC repeat genotype, but no difference in pulmonary disease progression. The GCLC GAC repeat polymorphism needs to be further investigated in an independent CF cohort of larger sample size which includes a larger number of CF patients with mild CFTR genotype. The deletion of the GSTM1 gene was shown to be a modifier gene for liver disease in CF.

It is probable that the oxidant burden in CF overwhelms the ability of the immune system to keep up with production of GSH. We conclude that polymorphisms in genes involved in

glutathione production (GCLC) and in antioxidant response (GSTs) have an effect on the pulmonary disease process in CF, but the severity of the effect of the CF disease process may make it difficult to investigate these genes as modifiers of pulmonary disease in CF patients with severe CFTR genotypes. Polymorphisms leading to reduced levels of GSTs (deletion polymorphism in GSTM1 gene) may contribute to CF liver disease.

Mucin genes. In CF, mucus has been shown to be more viscous and adhesive. The CFTR defect contributes both directly and indirectly to this increased viscosity and adhesiveness of airway mucus; however, the products of inflammation and infection also affect mucus viscosity, viscoelasticity and mucociliary clearance. In this study we considered polymorphisms in two mucin genes which may also contribute to the viscosity of CF mucus, MUC2 and MUC5B. We hypothesized that a longer length of the tandem repeats of the MUC2 gene polymorphism and a higher number of repeats of the VNTR of the MUC5B gene would be associated with more severe pulmonary disease severity and progression in CF and earlier onset of infection with CF pathogens. The number of repeats for the MUC5B gene polymorphism was associated in our study cohort with pulmonary disease progression. Although the polymorphisms investigated did not affect survival or the prevalence of chronic infection with *P. aeruginosa* or *Burkholderia cepacia* complex, we did show variability in age of first infection with *P. aeruginosa*. The polymorphisms investigated in this study may have an effect on mucous viscosity, however compared with other factors operating in the CF airways, their contribution to the CF pulmonary disease outcome variables investigated, is likely minor. Nonetheless, our data suggest that these polymorphisms contribute to pulmonary disease severity and progression in CF and further examination controlling for CFTR genotype on a larger CF cohort with a greater representation of mild CFTR genotypes is worthy of further investigation.

7.2 CONCLUSIONS FOR BCC INFECTION SUB-STUDY

We showed an more rapid decline in lung function following acquisition of BCC which was further augmented by co-infection with *P. aeruginosa*. There were no differences in pulmonary disease progression when we compared the rate of decline in %predFEV_1 in the

pre- BCC- acquisition interval between the BCC group and non-infected controls, but showed a steeper decline in %predFEV$_1$ over time in the BCC group following chronic infection with BCC. A closer examination of the rate of decline in %predFEV$_1$ for the post-acquisition of BCC interval showed a similar rate of decline (2% per year) in BCC infected CF patients as shown for the rate of decline in CF patients infected chronically with *P. aeruginosa*. In addition CF patients chronically infected with both BCC and *P. aeruginosa* showed a steeper rate of decline in %predFEV$_1$ over time, which seemed in our cohort to be additive (i.e., 4% per year) compared to rate of decline observed in CF patients infected chronically with either BCC or *P. aeruginosa* (i.e., 2% per year).

The mortality rate was similar in patients infected with either BCC or *P. aeruginosa*, but higher in the BCC infected group that was co-infected with *P. aeruginosa* compared with non-infected CF patients or patients infected only with *P. aeruginosa*. CF patients infected with both pathogens did not show a significant difference in risk of death or lung transplant compared with infection with BCC only. We also showed higher mortality in CF patients transiently infected with BCC compared with non-infected CF controls. Although mortality was higher in CF patients infected with genomovar IIIA versus genomovar II, the estimated decline in %predFEV$_1$ over time was similar between CF patients chronically infected with BCC genomovar II and IIIA. The rate of decline in %predFEV$_1$ was similar for CF patients infected with BCC genomovar IIIA across RAPD types.

In our study we were unable to control for CFTR genotype as all patients infected chronically with BCC had a severe CFTR genotype. It is unclear whether this finding is confined to our study cohort or a common finding in the CF population, since previous literature has not always reported or controlled for CFTR genotype or grouped CFTR genotype by CFTR class mutations.

Our findings provide a clear rationale for segregating BCC infected CF patients irrespective of BCC genotype and RAPD type. The time of acquisition of BCC was defined in our study and we were therefore able to show that BCC infection is not necessarily a marker of disease severity, but that chronic infection or even transient infection with BCC affect pulmonary disease progression and mortality and co-infection with *P. aeruginosa* has an additive effect on both pulmonary disease progression and risk of death or lung transplant. We therefore

conclude that worse prognosis usually associated with chronic BCC infection is not directly related to BCC infection but to the overall clinical state of the CF patient.

7. 3 RECENT RELAVENT RESEARCH PUBLISHED DURING THIS WRITE-UP

There have been a number of studies published during this write-up. In this section, these studies will be reviewed and discussed.

There have been 2 recent studies that have been published investigating BCC infection. Kalish and associates (16) grouped their BCC infected CF patients into BCC infected with GEN II versus GEN VI. They showed higher mortality in GEN II and GEN VI groups compared to matched CF controls infected with *P. aeruginosa*. Jones and associates investigated two common BCC colonizers in the United Kingdom (GEN IIIA and GEN II). This group showed higher mortality in the GEN IIIA group compared to matched CF controls infected with *P. aeruginosa* only (17). There were, however, no comparisons made in these two studies (16, 17) between the BCC cohorts investigated. Both these studies showed that infection with BCC has a more severe outcome (survival and decline in pulmonary function) than infection with *P. aeruginosa*. Our study confirms their findings and also showed that infection with pathogens, BCC and *P. aeruginosa*, has an additive effect; more rapid progression to an event (death or lung transplantation) and a steeper decline in %predFEV_1 over time compared to infection with one pathogen. Our study design also allowed us to investigate their clinical status pre and post-acquisition of BCC and to further investigate differences based on RAPD type for genomovar IIIA. In our study we also investigated the modifier gene MBL2 and showed no difference in susceptibility to BCC acquisition among MBL2 wild-type and deficient CF patients and a steeper decline in %predFEV_1 over time but only following chronic infection with BCC (no differences in the pre-acquisition interval).

MBL2 deficiency has been investigated in a number of different CF cohorts. There is inconstancy in the definition of MBL2 deficiency. In early studies including our own classification of MBL2 wild-type included homozygosity or heterozygosity for the high expression promoter allele Y at -221 and the wild-type allele (A) for the 3 codon polymorphisms on exon 1. More recent studies have grouped their CF patients for the three

exon 1 polymorphisms as A/A (homozygous wildtype for all three exon 1 polymorphisms) and A/0, and 0/0 (homozygous for condon 52, 54 and 57 polymorphisms). This grouping was also used by some of the studies already presented in this thesis(18). In the study by Buranawuti and coworkers they showed worse survival for CF patients (total N=164) with a 0/0 genotype, and no significant differences in (cross-sectional) %predFEV$_1$ (61% Vs 58%) and succeptibility to infection with BCC (2.6% Vs 2.8%) or *P. aeruginosa* (94% Vs 89%) for MBL2 A/A or A/0 versus 0/0 genotype, respectively. Muhlebach and colleagues measured MBL2 serum levels in children and adult CF patients (total N=102 with pulmonary function data)(19). They showed higher levels of serum MBL2 in CF patients than expected in healthy controls. They suggest that there is an age-dependent effect of MBL2 levels with CF phenotype as defined by %predFEV$_1$; CF children with low MBL2 levels are somehow protected from lung damage caused by excessive complement activation, but experience more pulmonary infections. With the onset of puberty there is a switch and CF patients with high MBL2 levels seem to show better pulmonary function, which the authors suggest is related to changes in hormone levels (growth hormone, thyroxin, glucocorticoids) with puberty which affect circulating levels of MBL2(19). Another possibility is that low MBL2 levels possibly associated with a MBL2 deficient genotype are protective in the early years as they suggest, and this could be also because these children are free of infections at the time. With increasing age MBL2 deficiency contributes to pulmonary disease progression and the reason for this could be that there is an inadequate response to recurrent and chronic infection with *P. aeruginosa* and *Staphylococcus aureus* leading to increased lung damage due to chronic inflammation and recurrent and chronic infection. Our study showed that in the case of BCC infection MBL2 deficiency is protective and MBL2 wildtype genotype may be associated with intracellular invasion by the BCC pathogen and more advanced pulmonary disease over time with chronic infection with BCC. Olesen and associates showed that a polymorphism in the serine protease MBL2-associated serine protease (MASP-2) gene (Asp120Gly, A359G) was associated with variable levels of the protease (20). MASP2 levels were reduced in heterozygotes (211 Vs 348ng/mL) and homozygosity for the MASP2 polymorphism was associated with the lowest levels (43ng/mL) and also showed no functional activity for this latter protein. The study, however, did not show an association between the MASP2 polymorphism and pulmonary function in their CF cohort (20). CF patients were also genotyped for the three exon 1 polymorphisms and the three promoter polymorphisms for

MBL2 and serum levels measured. Multiple analyses were done with various pulmonary function parameters and MBL2 groupings. Their total sample size was small (100 patients with pulmonary function values) and associations were shown for various grouping of MBL2 genotype and pulmonary function (20). This study does show the importance of defining CF phenotype in one's study and controlling for confounding variables to avoid spurious associations. Interestingly the latter study supports our grouping for MBL2 based on levels of the protein as presented by Garred and associates. Infection with multiple pathogens needs to be considered in the study model and also precisely defining pulmonary function when only investigating cross-sectional %predFEV$_1$ is important for BCC as we have shown that infection with the pathogen affects pulmonary disease progression, which may be hidden if the study design is not properly designed.

Choi and associates(21) investigated the MBL2 promoter (-221X/Y) and exon 1 polymorphisms and SPA-1 and 2 haplotypes in 137 CF patients as reported by Flores et al(22). They showed lower mean pulmonary function values for CF subjects homozygous for the delta F508 mutation who were homozygous or heterozygous for exon 1 polymorphisms (mean %predFEV$_1$ of 64% versus 48% of predicted, p=0.03). For the pulmonary surfactant A1 and A2 genes the common haplotypes were investigated. They showed lower mean %predFEV$_1$ in CF patients with at least one 6A^3 allele of SPA1 (55% versus 71% of predicted, p=0.01). They also showed lower mean %predFEV$_1$ in CF patients with at least one 1A^1 allele of SPA2 (54% versus 68% of predicted, p=0.01) compared to CF patients carrying any other SPA-2 haplotype. In addition, Choi and associates(21) showed CF patients with at least one SPA-1/SPA-2 haplotype 6A^3/1A^1 to have lower mean %predFEV$_1$ (53% versus 67% of predicted, p=0.009) compared to other SPA-1/SPA-2 haplotypes. The authors showed similar results for other measures of disease severity such as modified Schwachman-Kulczycki scores(21).

The 3 prime enhancer polymorphism in the α_1-AT gene (G 1237 A) was investigated by Courtney and associates (23). Three hundred and twenty CF patients were genotyped for the polymorphism. They showed a small significant association between the variant allele and %predFEV$_1$. In our larger study this difference was not statistically significant and this has been confirmed by the more recent study by Drumm and associates (24). Acute α_1-AT

levels and levels during a pulmonary infection requiring IV therapy were collected on a sub-sample (N=16), following a similar protocol to our study and subjects grouped based on the polymorphisms into two groups (variant allele: AA or AG genotype and wild-type: GG genotype). They did not show an association with the polymorphism and levels of the protein, which are in agreement with our study findings (25).

The human beta defensin 2 gene located on chromosome 8p23.1 was investigated by Hollox and associates on a CF sample (N=355)(26). A 260kb repeat unit which varies in copy number in people from 2-12 was investigated in a CF cohort (N=355). Their hypothesis was that copy number variation was associated with pulmonary disease(26). There were no significant associations shown in this study(26).

A repeat polymorphism in the promoter region of the macrophage migration inhibitory factor (MIP) gene was shown as a potential modifier gene in CF by Plant and associates (27). They showed that homozygosity or heterozygosity for the low repeat (i.e., 5 copies of CATT repeat), which is associated with lower MIP promoter activity in *vitro* (28), was associated with decreased incidence of infection with *P. aeruginosa* and they noted a trend for milder pulmonary disease based on cross-sectional measurement of %predFEV_1 . While there is functional data to suggest a biological plausible role for this modifier in CF pulmonary disease progression, their study design and sample size are a limitation and this study needs to be further studied and replicated in independent CF cohorts.

Blaisdell and colleagues investigated an alternate chloride channel protein (CLC-2) with known polymorphisms in the promoter region and intron 1 (N=4), which is expressed in the respiratory epithelia (29). The study compared CF patients with mild and severe pulmonary disease and the frequency of the polymorphisms in these two groups (29).

A dinucleotide (GT) repeat polymorphism in the 5 prime untranslated region of the nitric oxide synthase 1 gene was investigated CF (N=59) and healthy controls (N=59) (30). The study showed a positive association between number of repeats and exhaled NO levels (30). Subjects were grouped for analysis of pulmonary function over 5 years based on having 0, 1, or 2 alleles with more than 27 repeats and showed a slower decline in pulmonary function over 5 years in CF patients who had two alleles with greater than 27 repeats (30).Grasemann

and associates (31) investigated a polymorphism in the nitric oxide synthase 3 gene (G894T) in a CF cohort (N=70). Although they showed no significant differences for this modifier gene and measured parameters (%predFEV$_1$ and frequency of *P. aeruginosa* infection) for the total group, when stratified for gender they showed higher exhaled nitric oxide levels in female CF patients with the T-allele (TT or GT genotype). These results were not confirmed by the larger study by Drumm and associates (24).

Yarden and coworkers (32)and Buranawuti and associates (33) investigated tumor necrosis factor alpha (TNFα) promoter (C851T, G308A and G238A) and intron (+691 G ins/del) gene polymorphisms in CF populations. Buranawuti and associates showed improved survival (GG Vs GA) for the TNFα (G238A) polymorphism (33). Yarden and coworkers did not show differences in pulmonary function and TNFα (G238A) polymorphism, but showed better pulmonary function associated with heterozygosity for the intron polymorphism (32). These findings are interesting but need to be investigated in larger subsets of patients and better defined CF phenotypes, including CFTR class genotype grouping.

In a large multicenter study, which included our study cohort, polymorphisms in ten modifier genes previously investigated as modifiers in CF cohorts were investigated in this study (α_1-AT, MBL2, angiotensin-converting enzyme, transforming growth factor 1 (TGFβ1), beta-2 adrenergic receptor, GSTM1, GSTP1, interleukin 10, TNFα and nitric oxide synthase 3)(24). A significant association with severe pulmonary disease was shown for the TGFβ1 gene polymorphism at codon 10 (CC genotype)(24).

The results have been published from numerous CF sibling and twin studies and have shown that shared genes (CF and other genes) and environment explain a large part of the variability observed in pulmonary disease independent of CFTR genotype, therefore the search for modifier genes is validated. Picard and associates (34) showed high familial concordance with among CF siblings and pancreatic insufficiency, meconeum ilieus, age of CF diagnosis, infection with CF related pathogens and respiratory symptoms. They showed that usually because of the first sibling already diagnosed with CF, the younger sibling tended to be diagnosed earlier, but also tended to be come infected with respiratory

pathogens earlier, which they attributed to cross-infection from the already colonized CF sibling(34). Earlier exposure to specifically to *Staphylococus aureus* was associated to earlier exposure of the second younger CF sibling to the hospital environment (34, 35). Vanscoy and coworkers (36) have assembled a study cohort of monozygotic (MZ) and dizygotic twins (DZ), and affected with CF siblings in the U.S.A. and have shown high familial concordance for pancreatic sufficiency, nutritional status (BMI) and pulmonary function (%predFEV$_1$). Stratifying for gender also increased concordance and they showed that same sex DZ twins and same sex siblings produced higher correlations for nutritional and pulmonary function parameters than values shown for the entire group of siblings (DZ and sibships) (36). MZ and DZ twins have been shown to demonstrate similar levels of concordance for %predFEV$_1$ suggesting that increased gene sharing does not contribute to substantially increase similarity for this parameter, but that other shared factors contribute to explanation of variability (37). To that effect, preliminary results presented in a review paper by Cutting (37) shows that environment contributes to variability in %predFEV$_1$; lower concordance was shown for MZ twins who live apart (r=0.73) versus living together (r=0.96). Twin studies have added to the rationale for investigating other genes to explain for the variability shown in pulmonary function and disease progression in CF and that environment also contributes to this variability.

7.4 STRENGTHS AND WEAKNESSES OF THE THESIS STUDY

7.4.1 Study strengths

The major strength of our study was the collection of clinical data on a relatively large multi-center CF cohort encompassing CF patients with variable disease course and age (5 years and greater) and collection of a substantial longitudinal time interval for study. We were able to control for numerous parameters which would contribute to pulmonary disease severity and progression such as: a) chronic infection with pathogens *P. aeruginosa* and BCC, b) CFTR genotype with patients classified by CFTR class mutations, c) PSS and nutritional status (BMI) and d) clinical status at encounters used for the longitudinal time interval.

The clinical status at each encounter was coded and thus were able to identify encounters where the patient was clinically stable or ill, and whether an encounter was part of a hospitalization (which was further classified) or home IV therapy. For the investigation of

pulmonary disease progression we selected only encounters classified as stable thereby reducing variability in %predFEV$_1$ for each subject's longitudinal data due to poor clinical status. From the clinical status variable we also identified periods of IV therapy for pulmonary infections and used this data to calculate the frequency of pulmonary infections requiring IV therapy over 2 and 5 years for the study cohort.

A major objective of the study was to examine a significantly long time interval in order to be able to observe changes in clinical status for the study cohort. Since CF patients with mild (%predFEV$_1$>70%) and severe pulmonary disease (%predFEV$_1$<40%) show small changes in pulmonary disease progression over short intervals (1 or 2 years) the goal was to collect clinical data prospectively and retrospectively for up to 10 years. In addition we collected descriptive data by reviewing the patient charts for age of first infection and chronic infection with *P. aeruginosa* and BCC and recorded the infection status for each encounter and the patients' clinical status at each encounter. We used stable clinical status encounters to investigate pulmonary disease progression, thus reducing variability in pulmonary function due to poor health status. In our investigation of BCC infection on pulmonary disease progression, the time point of BCC acquisition was determined and encounter data collected both for the pre-acquisition and post-acquisition with BCC intervals.

Many of the findings in the investigation of pulmonary disease progression and survival and MBL2 deficiency status in BCC infected CF patients would not have been observed if we had not allowed for the following considerations in the study design and data collection process:

- Considered the time of BCC acquisition.
- Grouped subjects for BCC infection into transient and chronic BCC infected and non-infected control CF patients and in addition considered infection status for the pathogen *P. aeruginosa* for these groups.
- Distinguished between the pre- and post-acquisition with BCC longitudinal interval and collected clinical data for the pre and post-acquisition interval.
- Further classified CF patients infected with BCC for genomovar and RAPD typing of BCC.

7.4.2 Study Weaknesses

A potential limitation of this study was overlooking the potential importance of controlling for *Staphylococcus aureus* infection status for pulmonary disease progression models. Inclusion of *Staphylococcus aureus* infection status may have been relevant for MBL2 models investigating pulmonary disease progression, age of onset and age of chronic infection with *P. aeruginosa* and mortality. Collecting data for age of first infection of *Staphylococcus aureus*, and infection status for the pathogen for the longitudinal interval collected would potentially have allowed for the investigation of the effect of this organism on pulmonary disease progression and age of onset of chronic infection with *P. aeruginosa* and relationship with MBL2 deficiency status. It may have also been important to include *Staphylococcus aureus* infection status for other modifier genes investigated (eg., mucin genes) as well.

CF patients with mild CFTR mutations were not represented sufficiently in this study to investigate as a subgroup in many analyses. Some of the problems encountered included:

- CF patients with specific CFTR functional types were not represented in sufficient for individual modifier gene groupings.
- CF patients with specific CFTR types were not represented in sufficient numbers to correlate with other study variables included in the models, such as *P. aeruginosa* infection status, survival endpoints (i.e., death or lung transplantation). When modifier gene grouping was considered there were empty cells and the statistical model could not be investigated due to estimation errors.

Since patients who are homozygous or heterozygous for class 4 or 5 CFTR mutations, represent a very small percentage of the Canadian CF population, it is important for future research to try and recruit this group aggressively. In addition, grouping these patients consistently across studies is important. Commonly these patients are either excluded from study or if used in a study cohort are grouped with patients who are characterized with one or 2 unknown or unclassified CFTR mutations. The assumption that CF patients who possess one or 2 unknown or unclassified CFTR mutations have mild CFTR mutations (class 4 or 5) is not assured and may bias the results. In this study, CF patients who possessed 2 unknown or unclassified CFTR mutations were either included in the analyses and studied as a group or excluded if their numbers were small for study variables including

modifier gene groupings. CF patients with class 4 or 5 CFTR mutations are an important group to study (i.e., partially functioning CFTR) since the effect of the modifier gene(s) on the clinical variables may be potentially more apparent in this group.

The study cohort was genotyped for a relatively small number of polymorphisms in the pulmonary surfactant genes. Since little was known about the functional significance of the polymorphisms in these genes, the purpose of our study was merely to explore 1 or 2 amino acid changing polymorphisms per gene. Our preliminary results show that further research is warranted to investigate these genes as potential modifiers of CF pulmonary disease.

7.5 APPLICATIONS OF OUR RESEARCH FINDINGS

- An exogenous source of α_1-AT may be of potential benefit in malnourished CF patients who have more severe pulmonary disease during pulmonary infections, while measures are taken to normalize their body weight.
- Aggressive treatment of infections with *Staphylococcus aureus*, especially in CF patients who are not yet chronically infected by *P. aeruginosa* or BCC, may be of potential benefit by prolonging time to acquisition and chronic infection with *P. aeruginosa* or BCC. Therefore genotyping for MBL2 deficiency alleles in CF children may be of potential benefit for identifying CF patients at increased risk for *Staphylococcus aureus* recurrent infections, especially in early childhood, and BCC infection.
- CF patients who are chronically infected with BCC regardless of genomovar group should be cohorted separately to prevent cross infection with a different BCC species.

7.6 FUTURE DIRECTIONS

Further studies investigating MBL2 genotypes in CF are warranted and the following directions are proposed:

- A study replicating our design to study pulmonary disease progression and mortality with BCC infection and MBL2 deficiency in an independent CF cohort to confirm our findings.
- A study incorporating follow-up of 10 or more years controlling for BCC infection status and time of BCC acquisition in the model, but also including *Staphylococcus aureus* infection status for the interval investigated to assess whether recurrent or chronic co-infection with *Staphylococcus aureus* with and without co-infection with *P. aeruginosa* is associated with more severe pulmonary disease progression and increased risk of death or lung transplantation.
- Study the susceptibility of CF patients to *Staphylococcus aureus* infection based on MBL2 genotype and to investigate the association of frequency of infection with this pathogen with age of first and chronic infection with *P. aeruginosa* and/or BCC.
- Further studies investigating known polymorphisms in pulmonary surfactant genes A-1, A-2 and D. Using our study design to investigate pulmonary disease severity and progression, mortality and age of onset of first and chronic *P. aeruginosa* infection, with regard to SPA1-SPA2 haplotypes as presented by DiAngelo and associates(38) and known SPD polymorphisms.
- Further studies investigating α_1-AT deficiency. We propose that a cohort of young CF patients be followed prospectively from birth (or time of CF diagnosis) to when they become chronically infected with CF pathogens recording yearly baseline α_1-AT levels and acute phase response levels to pulmonary infections. A similar longitudinal design to our study would be used and patients followed for pulmonary disease progression controlling for CFTR genotype, CF pathogen infection status, and body mass index. With this design we may be able to evaluate whether α_1-AT deficiency genotype is a modifier of pulmonary disease in the early stages of CF disease.
- Further study of the GCLC GAC repeats polymorphism with greater representation of CFTR class mutation groups, especially CF patients classified with mild CFTR genotype and replicate our study design.

7.7 BIBLIOGRAPHY

1.Zielenski, J., and L. C. Tsui. 1995. Cystic fibrosis: genotypic and phenotypic variations. *Annu Rev Genet* 29:777-807.

2.Boucher, R. C. 2002. An overview of the pathogenesis of cystic fibrosis lung disease. *Adv Drug Deliv Rev* 54(11):1359-71.

3.Boucher, R. C. 2001. Pathogenesis of cystic fibrosis airways disease. *Trans Am Clin Climatol Assoc* 112:99-107.

4.Hoiby, N., E. W. Flensborg, B. Beck, B. Friis, S. V. Jacobsen, and L. Jacobsen. 1977. Pseudomonas aeruginosa infection in cystic fibrosis. Diagnostic and prognostic significance of Pseudomonas aeruginosa precipitins determined by means of crossed immunoelectrophoresis. *Scand J Respir Dis* 58(2):65-79.

5.Cantin, A. M., S. Lafrenaye, and R. O. Begin. 1991. Antineutrophil elastase activity in cystic fibrosis serum. *Pediatric Pulmonology* 11(3):249-53.

6.Birrer, P., N. G. McElvaney, A. Rudeberg, C. W. Sommer, S. Liechti-Gallati, R. Kraemer, R. Hubbard, and R. G. Crystal. 1994. Protease-antiprotease imbalance in the lungs of children with cystic fibrosis. *American Journal of Respiratory & Critical Care Medicine* 150(1):207-13.

7.Plotkowski, M. C., G. Beck, J. M. Tournier, M. Bernardo-Filho, E. A. Marques, and E. Puchelle. 1989. Adherence of Pseudomonas aeruginosa to respiratory epithelium and the effect of leucocyte elastase. *J Med Microbiol* 30(4):285-93.

8.Plotkowski, M. C., A. M. Saliba, S. H. Pereira, M. P. Cervante, and O. Bajolet-Laudinat. 1994. Pseudomonas aeruginosa selective adherence to and entry into human endothelial cells. *Infect Immun* 62(12):5456-63.

9.Hilman, B. C. 1997. Genetic and immunologic aspects of cystic fibrosis. *Ann Allergy Asthma Immunol* 79(5):379-90; quiz 391-4.

10.Gabolde, M., M. Guilloud-Bataille, J. Feingold, and C. Besmond. 1999. Association of variant alleles of mannose binding lectin with severity of pulmonary disease in cystic fibrosis: cohort study. *Bmj* 319(7218):1166-7.

11.Carlsson, M., A. G. Sjoholm, L. Eriksson, S. Thiel, J. C. Jensenius, M. Segelmark, and L. Truedsson. 2005. Deficiency of the mannan-binding lectin pathway of complement and poor

outcome in cystic fibrosis: bacterial colonization may be decisive for a relationship. *Clin Exp Immunol* 139(2):306-13.

12.Lands, L. C., V. L. Grey, and C. Grenier. 2000. Total plasma antioxidant capacity in cystic fibrosis. *Pediatr Pulmonol* 29(2):81-7.

13.Rahman, I., E. Skwarska, M. Henry, M. Davis, C. M. O'Connor, M. X. FitzGerald, A. Greening, and W. MacNee. 1999. Systemic and pulmonary oxidative stress in idiopathic pulmonary fibrosis. *Free Radic Biol Med* 27(1-2):60-8.

14.Smith, L. J., and J. Anderson. 1992. Oxygen-induced lung damage. Relationship to lung mitochondrial glutathione levels. *Am Rev Respir Dis* 146(6):1452-7.

15.Kelly, F. J. 1999. Gluthathione: in defence of the lung. *Food Chem Toxicol* 37(9-10):963-6.

16.Kalish, L. A., D. A. Waltz, M. Dovey, G. Potter-Bynoe, A. J. McAdam, J. J. Lipuma, C. Gerard, and D. Goldmann. 2006. Impact of Burkholderia dolosa on lung function and survival in cystic fibrosis. *Am J Respir Crit Care Med* 173(4):421-5.

17.Jones, A. M., M. E. Dodd, J. R. Govan, V. Barcus, C. J. Doherty, J. Morris, and A. K. Webb. 2004. Burkholderia cenocepacia and Burkholderia multivorans: influence on survival in cystic fibrosis. *Thorax* 59(11):948-51.

18.Davies, J. C., M. W. Turner, and N. Klein. 2004. Impaired pulmonary status in cystic fibrosis adults with two mutated MBL-2 alleles. *Eur Respir J* 24(5):798-804.

19.Muhlebach, M. S., S. L. MacDonald, B. Button, J. J. Hubbard, M. L. Turner, R. C. Boucher, and D. C. Kilpatrick. 2006. Association between mannan-binding lectin and impaired lung function in cystic fibrosis may be age-dependent. *Clin Exp Immunol* 145(2):302-7.

20.Olesen, H. V., J. C. Jensenius, R. Steffensen, S. Thiel, and P. O. Schiotz. 2006. The mannan-binding lectin pathway and lung disease in cystic fibrosis--disfunction of mannan-binding lectin-associated serine protease 2 (MASP-2) may be a major modifier. *Clin Immunol* 121(3):324-31.

21.Choi, E. H., M. Ehrmantraut, C. B. Foster, J. Moss, and S. J. Chanock. 2006. Association of common haplotypes of surfactant protein A1 and A2 (SFTPA1 and SFTPA2) genes with severity of lung disease in cystic fibrosis. *Pediatr Pulmonol* 41(3):255-62.

22.Floros, J., S. DiAngelo, M. Koptides, A. M. Karinch, P. K. Rogan, H. Nielsen, R. G. Spragg, K. Watterberg, and G. Deiter. 1996. Human SP-A locus: allele frequencies and

linkage disequilibrium between the two surfactant protein A genes. *Am J Respir Cell Mol Biol* 15(4):489-98.

23.Courtney, J. M., B. J. Plant, K. Morgan, J. Rendall, C. Gallagher, M. Ennis, N. Kalsheker, S. Elborn, and C. M. O'Connor. 2006. Association of improved pulmonary phenotype in Irish cystic fibrosis patients with a 3' enhancer polymorphism in alpha-1-antitrypsin. *Pediatr Pulmonol* 41(6):584-91.

24.Drumm, M. L., M. W. Konstan, M. D. Schluchter, A. Handler, R. Pace, F. Zou, M. Zariwala, D. Fargo, A. Xu, J. M. Dunn, R. J. Darrah, R. Dorfman, A. J. Sandford, M. Corey, J. Zielenski, P. Durie, K. Goddard, J. R. Yankaskas, F. A. Wright, and M. R. Knowles. 2005. Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* 353(14):1443-53.

25.Frangolias, D. D., J. Ruan, P. J. Wilcox, A. G. Davidson, L. T. Wong, Y. Berthiaume, R. Hennessey, A. Freitag, L. Pedder, M. Corey, N. Sweezey, J. Zielenski, E. Tullis, and A. J. Sandford. 2003. Alpha 1-antitrypsin deficiency alleles in cystic fibrosis lung disease. *Am J Respir Cell Mol Biol* 29(3 Pt 1):390-6.

26.Hollox, E. J., J. Davies, U. Griesenbach, J. Burgess, E. W. Alton, and J. A. Armour. 2005. Beta-defensin genomic copy number is not a modifier locus for cystic fibrosis. *J Negat Results Biomed* 4:9.

27.Plant, B. J., C. G. Gallagher, R. Bucala, J. A. Baugh, S. Chappell, L. Morgan, C. M. O'Connor, K. Morgan, and S. C. Donnelly. 2005. Cystic fibrosis, disease severity, and a macrophage migration inhibitory factor polymorphism. *Am J Respir Crit Care Med* 172(11):1412-5.

28.Baugh, J. A., S. Chitnis, S. C. Donnelly, J. Monteiro, X. Lin, B. J. Plant, F. Wolfe, P. K. Gregersen, and R. Bucala. 2002. A functional promoter polymorphism in the macrophage migration inhibitory factor (MIF) gene associated with disease severity in rheumatoid arthritis. *Genes Immun* 3(3):170-6.

29.Blaisdell, C. J., T. D. Howard, A. Stern, P. Bamford, E. R. Bleecker, and O. C. Stine. 2004. CLC-2 single nucleotide polymorphisms (SNPs) as potential modifiers of cystic fibrosis disease severity. *BMC Med Genet* 5:26.

30.Texereau, J., S. Marullo, D. Hubert, J. Coste, D. J. Dusser, J. Dall'Ava-Santucci, and A. T. Dinh-Xuan. 2004. Nitric oxide synthase 1 as a potential modifier gene of decline in lung function in patients with cystic fibrosis. *Thorax* 59(2):156-8.

31.Grasemann, H., K. Storm van's Gravesande, R. Buscher, N. Knauer, E. S. Silverman, L. J. Palmer, J. M. Drazen, and F. Ratjen. 2003. Endothelial nitric oxide synthase variants in cystic fibrosis lung disease. *Am J Respir Crit Care Med* 167(3):390-4.

32.Yarden, J., D. Radojkovic, K. De Boeck, M. Macek, Jr., D. Zemkova, V. Vavrova, R. Vlietinck, J. J. Cassiman, and H. Cuppens. 2005. Association of tumour necrosis factor alpha variants with the CF pulmonary phenotype. *Thorax* 60(4):320-5.

33.Buranawuti, K., M. P. Boyle, S. Cheng, L. L. Steiner, K. McDougal, M. D. Fallin, C. Merlo, P. L. Zeitlin, B. J. Rosenstein, P. J. Mogayzel, Jr., X. Wang, and G. R. Cutting. 2007. Variants in mannose-binding lectin and tumour necrosis factor alpha affect survival in cystic fibrosis. *J Med Genet* 44(3):209-14.

34.Picard, E., M. Aviram, Y. Yahav, J. Rivlin, H. Blau, L. Bentur, A. Avital, Y. Villa, S. Schwartz, B. Kerem, and E. Kerem. 2004. Familial concordance of phenotype and microbial variation among siblings with CF. *Pediatr Pulmonol* 38(4):292-7.

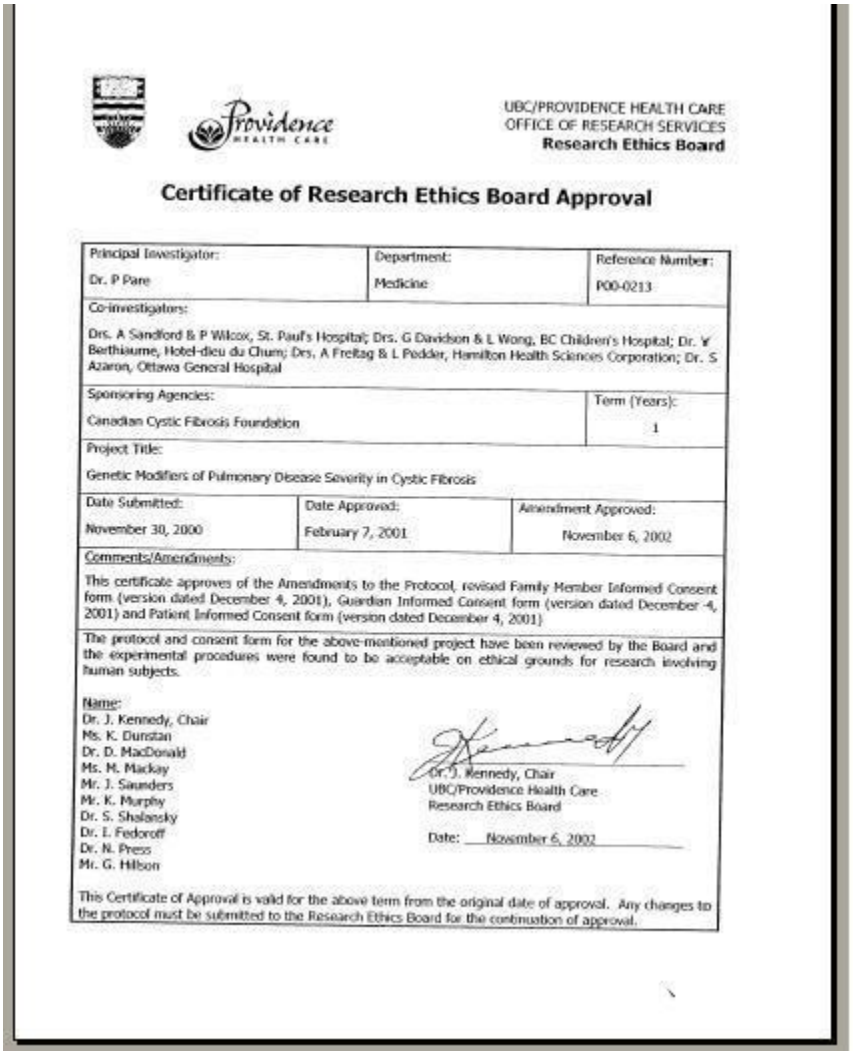
35.Givney, R., A. Vickery, A. Holliday, M. Pegler, and R. Benn. 1997. Methicillin-resistant Staphylococcus aureus in a cystic fibrosis unit. *J Hosp Infect* 35(1):27-36.

36.Vanscoy, L. L., S. M. Blackman, J. M. Collaco, A. Bowers, T. Lai, K. Naughton, M. Algire, R. McWilliams, S. Beck, J. Hoover-Fong, A. Hamosh, D. Cutler, and G. R. Cutting. 2007. Heritability of lung disease severity in cystic fibrosis. *Am J Respir Crit Care Med* 175(10):1036-43.

37.Cutting, G. R. 2005. Modifier genetics: cystic fibrosis. *Annu Rev Genomics Hum Genet* 6:237-60.

38.DiAngelo, S., Z. Lin, G. Wang, S. Phillips, M. Ramet, J. Luo, and J. Floros. 1999. Novel, non-radioactive, simple and multiplex PCR-cRFLP methods for genotyping human SP-A and SP-D marker alleles. *Dis Markers* 15(4):269-81.

APPENDIX A: ETHICS APPROVAL FORMS



UBC/PROVIDENCE HEALTH CARE
OFFICE OF RESEARCH SERVICES
Research Ethics Board

Certificate of Research Ethics Board Approval

Principal Investigator:	Department:	Reference Number:
Dr. P Pare	Medicine	P00-0213

Co-investigators:

Drs. A Sandford & P Wilcox, St. Paul's Hospital; Drs. G Davidson & L Wong, BC Children's Hospital; Dr. Y Berthiaume, Hotel-dieu du Chum; Drs. A Freitag & L Pedder, Hamilton Health Sciences Corporation; Dr. S Azaron, Ottawa General Hospital

Sponsoring Agencies:	Term (Years):
Canadian Cystic Fibrosis Foundation	1

Project Title:

Genetic Modifiers of Pulmonary Disease Severity in Cystic Fibrosis

Date Submitted:	Date Approved:	Amendment Approved:
November 30, 2000	February 7, 2001	November 6, 2002

Comments/Amendments:

This certificate approves of the Amendments to the Protocol, revised Family Member Informed Consent form (version dated December 4, 2001), Guardian Informed Consent form (version dated December 4, 2001) and Patient Informed Consent form (version dated December 4, 2001)

The protocol and consent form for the above-mentioned project have been reviewed by the Board and the experimental procedures were found to be acceptable on ethical grounds for research involving human subjects.

Name:
Dr. J. Kennedy, Chair
Ms. K. Dunstan
Dr. D. MacDonald
Ms. M. Mackay
Mr. J. Saunders
Mr. K. Murphy
Dr. S. Shalansky
Dr. I. Fedoroff
Dr. N. Press
Mr. G. Hilson

Dr. J. Kennedy, Chair
UBC/Providence Health Care
Research Ethics Board

Date: November 6, 2002

This Certificate of Approval is valid for the above term from the original date of approval. Any changes to the protocol must be submitted to the Research Ethics Board for the continuation of approval.

UBC/PROVIDENCE HEALTH CARE
OFFICE OF RESEARCH SERVICES

Certificate of Final Approval

Principal Investigator:	Department:	Reference Number:
Dr. P. Pare	Medicine	P00-0213
Co-investigators: Dr. Sandford		
Sponsoring Agencies: Red Cross		Term (Years): 3
Project Title: Contribution of genes other then CFTR to disease severity in cystic fibrosis		
Date Submitted: November 30, 2000	Date Ethical Approval: February 7, 2001	Date Final Approval: February 12, 2001

The above-mentioned study has recently been approved by the UBC/PHC Research Ethics Board. All other necessary departmental approvals (*Medical Records, Laboratory & Nursing*) are now in place and I am pleased to inform you that you have the permission of the hospital to begin your study.

Dr. M.V. O'Shaughnessy
Vice President, Research
Providence Health Care

Date: February 12, 2001