Polymorphisms of CF modifier genes: Their relationship to Pseudomonas aeruginosa infection and severity of disease in CF patients

by

Rossitta Pui Ki Yung

B.Sc., University of British Columbia, 2002
BMLSc., University of British Columbia, 2003

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

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Experimental Medicine)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2008

© Rossitta Pui Ki Yung, 2008
Abstract

Cystic Fibrosis is one of the most common genetic recessive diseases among Caucasians and is caused by mutations in the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) gene on chromosome 7. There are different classes of CFTR mutation, leading to differences in disease severity among patients. In addition to the CFTR genotype, secondary genetic factors, modifier genes, also influence CF phenotypes. Due to the dysfunction of CFTR protein and production of thickened mucus, bacterial infection in the lungs is favored and can lead to further clinical complications in CF patients. *Pseudomonas aeruginosa* is one of the most common bacteria detected among patients. The aim of this project was to investigate four candidate modifier genes, Factor B, Complement Factor 3, Toll-like Receptor 4 and Heme oxygenase-1, which might affect the status of *Pseudomonas aeruginosa* infection. A total of 22 single nucleotide polymorphisms (SNPs) were selected in these four genes and they were tested against five phenotypic traits, including age of diagnosis, FEV1% predicted and standard deviation value, age of first *Pseudomonas aeruginosa* infection and *Pseudomonas aeruginosa* infection status. Among the selected SNPs, both case-control studies and family-based analysis were performed in order to establish any correlation between the genotypes and the phenotypes. In addition, haplotype analysis was performed to determine whether there was interaction between SNPs or whether there were unidentified SNPs in the vicinity of the selected ones that might contribute to the observed phenotypic traits. Among the 22 chosen SNPs, 13 of them were found to be significantly linked to one or more of the tested phenotypes. The three most significant associations were BF_2557 with lung function, HMOX1_9531 with lung function and BF_7202 with age of diagnosis. Several haplotypes were significantly associated with one of the five phenotypes. There was no evidence for the presence of unidentified SNPs or interaction between SNPs. Most of haplotype associations were likely due to the presence of a single SNP which was found to be significantly linked to the phenotype. Conclusively, both SNPs and haplotype analyses suggest that the four candidate genes are modifiers of disease severity in CF.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xvi</td>
</tr>
<tr>
<td>Chapter 1: Introduction</td>
<td></td>
</tr>
<tr>
<td>1.1 Cystic fibrosis transmembrane regulator (CFTR) mutation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 CFTR protein and its function</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Diagnosis and clinical symptoms of CF disease</td>
<td>4</td>
</tr>
<tr>
<td>1.4 Pseudomonas aeruginosa infection in CF patients</td>
<td>5</td>
</tr>
<tr>
<td>1.5 Candidate modifier genes</td>
<td>6</td>
</tr>
<tr>
<td>1.5.1 Factor B and Complement factor 3</td>
<td>7</td>
</tr>
<tr>
<td>1.5.2 Toll-like receptor 4</td>
<td>10</td>
</tr>
<tr>
<td>1.5.3 Heme oxygenase-1</td>
<td>11</td>
</tr>
<tr>
<td>1.6 Single nucleotide polymorphism (SNP)</td>
<td>12</td>
</tr>
<tr>
<td>1.7 Association study</td>
<td>13</td>
</tr>
<tr>
<td>1.8 Thesis objectives</td>
<td>14</td>
</tr>
<tr>
<td>Chapter 2: Materials and Methods</td>
<td></td>
</tr>
<tr>
<td>2.1 Patients recruitment</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Quality control</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Selection of SNPs</td>
<td>16</td>
</tr>
<tr>
<td>2.4 TaqMan assays</td>
<td>21</td>
</tr>
<tr>
<td>2.5 Primers and probes for genotyping</td>
<td>23</td>
</tr>
<tr>
<td>2.6 Real time PCR reactions</td>
<td>25</td>
</tr>
<tr>
<td>2.7 Sequencing</td>
<td>28</td>
</tr>
<tr>
<td>2.8 Dilution of samples</td>
<td>30</td>
</tr>
<tr>
<td>2.9 Preparation for polymerase chain reaction and genotyping</td>
<td>30</td>
</tr>
<tr>
<td>2.10 Polymerase chain reaction</td>
<td>30</td>
</tr>
<tr>
<td>2.11 Genotyping</td>
<td>31</td>
</tr>
<tr>
<td>2.12 Quantification of DNA samples</td>
<td>32</td>
</tr>
<tr>
<td>2.13 Re-genotyping</td>
<td>35</td>
</tr>
<tr>
<td>2.14 Statistical data analysis</td>
<td>36</td>
</tr>
</tbody>
</table>
Chapter 3: Results

3.1 PicoGreen reaction ................................................................. 43
3.2 Analysis of the genotypic data for Mendelian inconsistencies .................................................. 44
3.3 Sequencing result by the University of British Columbia ................................................................ 44
3.4 Genotypes of the participating individuals .................................................................................. 45
3.5 Phenotypic characteristics of the study subjects ........................................................................ 46
3.6 Determination of Hardy-Weinberg equilibrium in the patient population .................................... 46
3.7 Comparison of genotype frequencies in the parent population and in online database .................. 49
3.8 ANOVA analysis of the influence of genotype on the phenotypes ................................................ 51
  (a) Factor B
    - (I) Age of diagnosis ........................................................................ 51
    - (II) FEV1 predicted value ................................................................. 52
    - (III) FEV1 standard deviation value .................................................. 53
  (b) Complement factor 3
    - (I) Age of diagnosis ......................................................................... 53
    - (II) FEV1 predicted value ................................................................. 53
    - (III) FEV1 standard deviation value .................................................. 53
  (c) Toll-like receptor 4
    - (I) Age of diagnosis ......................................................................... 54
    - (II) FEV1 predicted value ................................................................. 54
    - (III) FEV1 standard deviation value .................................................. 54
  (d) Heme oxygenase-1
    - (I) Age of diagnosis ......................................................................... 54
    - (II) FEV1 predicted value ................................................................. 55
    - (III) FEV1 standard deviation value .................................................. 55
3.9 Regression analysis ........................................................................... 56
  (a) BF_2557 ...................................................................................... 57
  (b) BF_7202 ...................................................................................... 58
  (c) TLR4_1859 .................................................................................... 59
  (d) HMOX1_2790 ............................................................................... 60
  (e) HMOX1_9531 ............................................................................... 61
3.10 Age of onset analysis for the age of first *Pseudomonas aeruginosa* infection

(a) Factor B
(b) Complement factor 3
(c) Toll-like receptor 4
(d) Heme oxygenase-1

3.11 *Pseudomonas aeruginosa* infection status

(a) Factor B
(b) Complement factor 3
(c) Toll-like receptor 4
(d) Heme oxygenase-1

3.12 Re-genotyping

3.13 FBAT analysis of phenotypic characteristics

(a) Age of diagnosis
(b) FEV1 predicted value
(c) FEV1 standard deviation value

3.14 Haplotype analysis by GRui program

(a) Factor B
   - (I) Age of diagnosis
   - (II) FEV1 predicted value
   - (III) FEV1 standard deviation value

(b) Complement factor 3
   - (I) Age of diagnosis
   - (II) FEV1 predicted value
   - (III) FEV1 standard deviation value

(c) Toll-like receptor 4
   - (I) Age of diagnosis
   - (II) FEV1 predicted value
   - (III) FEV1 standard deviation value

(e) Heme oxygenase-1
   - (I) Age of diagnosis
   - (II) FEV1 predicted value
   - (III) FEV1 standard deviation value

3.15 Haplotype analysis by the FBAT program

(a) Factor B
   - (I) Age of diagnosis
   - (II) FEV1 predicted value
   - (III) FEV1 standard deviation value
(b) Complement factor 3
- (I) Age of diagnosis..................................................82
- (II) FEV1 predicted value..............................................82
- (III) FEV1 standard deviation value..............................83
(c) Toll-like receptor 4
- (I) Age of diagnosis..................................................83
- (II) FEV1 predicted value..............................................84
- (III) FEV1 standard deviation value..............................84
(d) Heme oxygenase-1
- (I) Age of diagnosis..................................................85
- (II) FEV1 predicted value..............................................85
- (III) FEV1 standard deviation value..............................86

3.16 Haplotype analysis of the age of first
Pseudomonas aeruginosa infection by Hapstat.................86
(a) Factor B................................................................86
(b) Complement factor 3..................................................86
(c) Toll-like receptor 4....................................................87
(d) Heme oxygenase-1......................................................87

Chapter 4: Discussion
4.1 Analysis of the genotypic data for Mendelian
inconsistencies..................................................................88
4.2 Genotypes of all participating individuals...................89
4.3 Genotyping analysis of the parental population.............89
4.4 ANOVA, regression analysis and FBAT.........................91
4.5 Age of onset analysis...................................................95
4.6 Pseudomonas aeruginosa infection status......................96
4.7 Haplotype analysis by the GRui and FBAT programs........97
(a) Haplotype analysis by RGui program...........................97
(b) Haplotype analysis by the FBAT program.....................98
4.8 Haplotype analysis of the age of first Pseudomonas
aeruginosa infection by Hapstat.....................................100
4.9 Position of SNPs with significant association and
their effect of gene function............................................101
4.10 Summary.................................................................106
4.11 Future studies..........................................................106
**List of Tables**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Factor B polymorphism selection</td>
<td>18</td>
</tr>
<tr>
<td>Table 2</td>
<td>Complement factor 3 polymorphism selection</td>
<td>19</td>
</tr>
<tr>
<td>Table 3</td>
<td>Toll-like receptor 4 polymorphism selection</td>
<td>20</td>
</tr>
<tr>
<td>Table 4</td>
<td>Heme oxygenase-1 polymorphism selection</td>
<td>21</td>
</tr>
<tr>
<td>Table 5</td>
<td>The sequence of all primers used for SNP assays</td>
<td>24</td>
</tr>
<tr>
<td>Table 6</td>
<td>The sequence of all probes used for SNP assays</td>
<td>25</td>
</tr>
<tr>
<td>Table 7</td>
<td>Genotype discrepancies between the real time PCR results in this study and those posted by the Seattle SNPs website</td>
<td>28</td>
</tr>
<tr>
<td>Table 8</td>
<td>The sequence of both left and right primers used for sequencing of BF_7202 for Coriell samples</td>
<td>28</td>
</tr>
<tr>
<td>Table 9</td>
<td>Layout of a 96-well plate with standard DNA samples and DNA samples from source plates sent by Toronto</td>
<td>35</td>
</tr>
<tr>
<td>Table 10</td>
<td>Hardy-Weinberg equilibrium among the parental population</td>
<td>47</td>
</tr>
<tr>
<td>Table 11</td>
<td>Comparison of allele frequencies between the genotyping results and the reported values on either IIPGA or Seattle SNPs websites</td>
<td>50</td>
</tr>
<tr>
<td>Table 12</td>
<td>The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Factor B</td>
<td>52</td>
</tr>
<tr>
<td>Table 13</td>
<td>The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Factor B</td>
<td>52</td>
</tr>
<tr>
<td>Table 14</td>
<td>The ANOVA result of examining FEV1 standard deviation value among different genotypes of the selected SNPs in Factor B</td>
<td>53</td>
</tr>
<tr>
<td>Table 15</td>
<td>The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Toll-like receptor 4</td>
<td>54</td>
</tr>
<tr>
<td>Table 16</td>
<td>The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Heme oxygenase-1</td>
<td>55</td>
</tr>
<tr>
<td>Table 17</td>
<td>The ANOVA result of examining FEV1 standard deviation value among different genotypes of the selected SNPs in Heme oxygenase-1</td>
<td>55</td>
</tr>
<tr>
<td>Table 18</td>
<td>Regression analysis of the association of BF_2557 and FEV1 predicted value with confounding factors</td>
<td>57</td>
</tr>
<tr>
<td>Table 19</td>
<td>Regression analysis of the association of BF_2557 and FEV1 standard deviation value with confounding factors</td>
<td>57</td>
</tr>
<tr>
<td>Table 20</td>
<td>Regression analysis of the association of BF_7202 and age of diagnosis with confounding factors</td>
<td>58</td>
</tr>
<tr>
<td>Table 21</td>
<td>Regression analysis of the association of BF_7202 and FEV1 predicted value with confounding factors</td>
<td>58</td>
</tr>
<tr>
<td>Table 22</td>
<td>Regression analysis of the association of BF_7202 and FEV1 standard deviation value with confounding factors</td>
<td>59</td>
</tr>
<tr>
<td>Table 23</td>
<td>Regression analysis of the association of TLR4_1859 and age of diagnosis with</td>
<td></td>
</tr>
</tbody>
</table>
Table 24  Regression analysis of the association of HMOX1_2790 and FEV1 predicted value with confounding factors

Table 25  Regression analysis of the association of HMOX1_2790 and FEV1 standard deviation value with confounding factors

Table 26  Regression analysis of the association of HMOX1_9531 and FEV1 predicted value with confounding factors

Table 27  Regression analysis of the association of HMOX1_9531 and FEV1 standard deviation value with confounding factors

Table 28  Age of onset analysis investigating association between age of first Pseudomonas aeruginosa infection and selected SNPs in Toll-like receptor 4

Table 29  Chi square test for investigating the relationship between different genotypes of the selected SNPs in Complement factor 3 and Pseudomonas aeruginosa infection status

Table 30  Comparing the result of genotyping and re-genotyping

Table 31  Detailed results of FBAT analysis of age of diagnosis and SNPs under additive model

Table 32  Detailed results of FBAT analysis of age of diagnosis and SNPs under dominant model

Table 33  Detailed results of FBAT analysis of FEV1 predicted value and SNPs under additive model

Table 34  Detailed results of FBAT analysis of FEV1 predicted value and SNPs under dominant model

Table 35  Detailed results of FBAT analysis of FEV1 standard deviation value and SNPs under additive model

Table 36  Detailed results of FBAT analysis of FEV1 standard deviation value and SNPs under dominant model

Table 37  Haplotype analysis for investigation of correlation relationship between haplotype of Factor B and age of diagnosis with no adjustment for confounding factors

Table 38  Haplotype analysis for investigation of correlation relationship between haplotype of Factor B and FEV1 standard deviation value with no adjustment for confounding factors

Table 39  Haplotype analysis for investigation of correlation relationship between haplotype of Factor B and FEV1 standard deviation value with adjustment for confounding factors

Table 40  Haplotype analysis for investigation of correlation relationship between haplotype of complement factor 3 and age of diagnosis with no adjustment for confounding factors

Table 41  Haplotype analysis for investigation of correlation relationship between haplotype of Complement factor 3 and age of diagnosis with adjustment for confounding factors
Table 43  Haplotype analysis for investigation of correlation relationship between haplotype of Toll-like receptor 4 and age of diagnosis with no adjustment for confounding factors

Table 44  Haplotype analysis for investigation of correlation relationship between haplotype of Toll-like receptor 4 and age of diagnosis with adjustment for confounding factors

Table 45  Haplotype analysis for investigation of correlation relationship between haplotype of Heme oxygenase-1 and FEV1 predicted value with no adjustment for confounding factors

Table 46  Haplotype analysis for investigation of correlation relationship between haplotype of Heme oxygenase-1 and FEV1 predicted value with adjustment for confounding factors

Table 47  Haplotype analysis for investigation of correlation relationship between haplotype of Heme oxygenase-1 and FEV1 standard deviation value with no adjustment for confounding factors

Table 48  Haplotype analysis for investigation of correlation relationship between haplotype of Heme oxygenase-1 and FEV1 standard deviation value with adjustment for confounding factors

Table 49  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Factor B and FEV1 predicted value by the FBAT program

Table 50  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Factor B and FEV1 standard deviation value by the FBAT program

Table 51  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Complement factor 3 and age of diagnosis by the FBAT program

Table 52  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Complement factor 3 and FEV1 standard deviation value by the FBAT program

Table 53  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Toll-like receptor 4 and FEV1 predicted value by the FBAT program

Table 54  Haplotype analysis for investigation of correlation between combinations of selected SNPs in Heme oxygenase-1 and FEV1 predicted value by the FBAT program

Table 55  Summary of all significant SNP-phenotype association by ANOVA and FEAT analyses

Table 56  Summary of all significant haplotype-phenotype association when testing by RGui and FBAT analyses

Table 57  Summary Table of SNPs which revealed a significant association with the Pseudomonas aeruginosa infection status and the age of first Pseudomonas aeruginosa infection

Table 58  Summary table of the conservation score of the selected SNPs in the four candidate genes
TABLE A1 | Concentration of a subset of the DNA samples in the original source plates

TABLE A2 | Families with non-Mendelian inheritance

TABLE A3 | Genotype frequency of each of the SNPs examined in the gene of Factor B

TABLE A4 | Genotype frequency of each of the SNPs examined in the gene of Complement factor 3

TABLE A5 | Genotype frequency of each of the SNPs examined in the gene of Toll-like receptor 4

TABLE A6 | Genotype and allele frequency of each of the SNP examined in the gene of Heme oxygenase-1

TABLE A7 | Phenotypic characteristics of the CF patients

TABLE A8 | The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Complement factor 3

TABLE A9 | The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Complement factor 3

TABLE A10 | The ANOVA result of examining FEV1 sd value among different genotypes of the selected SNPs in Complement factor 3

TABLE A11 | The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Toll-like receptor 4

TABLE A12 | The ANOVA result of examining FEV1 sd value among different genotypes of the selected SNPs in Toll-like receptor 4

TABLE A13 | The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Heme oxygenase-1

TABLE A14 | Age of Onset Analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Factor B

TABLE A15 | Age of Onset Analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Complement factor 3

TABLE A16 | Age of Onset Analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Heme oxygenase 1

TABLE A17 | Chi squared test for investigating the relationship between different genotypes of the selected SNPs in Factor B and *Pseudomonas aeruginosa* infection status

TABLE A18 | Chi square test for investigating the relationship between different genotypes of the selected SNPs in Toll-like receptor 4 and *Pseudomonas aeruginosa* infection status

TABLE A19 | Chi square test for investigating the relationship between different genotypes of the selected SNPs in Heme oxygenase-1 and *Pseudomonas aeruginosa* infection status

TABLE A20 | FBAT analysis of the age of diagnosis under the additive model

TABLE A21 | FBAT analysis of the age of diagnosis under the dominant model
Table A22  FBAT analysis of FEV1 predicted value under
the additive model.................................................................132
Table A23  FBAT analysis of FEV1 predicted value under
the dominant model...............................................................133
Table A24  FBAT analysis of FEV1 standard deviation value
under the additive model..........................................................134
Table A25  FBAT analysis of FEV1 standard deviation value
under the dominant model........................................................135
Table A26  Frequencies of possible haplotypes generated for the
Factor B gene when determining the presence of any correlation between the haplotypes and age of
Diagnosis..................................................................................136
Table A27  Frequencies of possible haplotypes generated for the
Factor B gene when determining the presence of any correlation between the haplotypes and
FEV1 predicted value.................................................................136
Table A28  Haplotype analysis for investigation of correlation
relationship between combinations of selected SNPs in
Factor B and FEV1 predicted value, with no adjustment for the confounding factors...........................................136
Table A29  Haplotype analysis for investigation of possible correlation between combinations of selected SNPs
in Factor B and FEV1 predicted value, with adjustment for the confounding factors.................................................137
Table A30  Frequencies of possible haplotypes generated for the
Factor B gene when determining the presence of any correlation between the haplotypes and FEV1 standard
deviation value........................................................................137
Table A31  Frequencies of possible haplotypes generated for the
Complement factor 3 gene when determining the presence of any correlation between the haplotypes and age of
diagnosis..................................................................................138
Table A32  Frequencies of possible haplotypes generated for the
Complement factor 3 gene when determining the presence of any correlation between the haplotypes and FEV1
predicted value........................................................................138
Table A33  Haplotype analysis for investigation of correlation
relationship between combinations of selected SNPs in
Complement factor 3 and FEV1 predicted value, with no adjustment for the confounding factors.................................139
Table A34  Haplotype analysis for investigation of correlation
relationship between combinations of selected SNPs in
Complement factor 3 and FEV1 predicted value, with adjustment for the confounding factors...........................................139
Table A35  Frequencies of possible haplotypes generated for the
Complement factor 3 gene when determining the presence of any correlation between the haplotypes and FEV1 standard
deviation value........................................................................140
Table A36  Haplotype analysis for investigation of correlation
relationship between combinations of selected SNPs in
Complement factor 3 and FEV1 standard deviation value, with no adjustment for the confounding factors.................................140
Table A37  Haplotype analysis for investigation of correlation
relationship between combinations of selected SNPs in
Complement factor 3 and FEV1 standard deviation value, with adjustment for the confounding factors...........................................141
Table A38 Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and age of diagnosis.

Table A39 Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value.

Table A40 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 predicted value, with no adjustment for the confounding factors.

Table A41 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 predicted value, with adjustment for the confounding factors.

Table A42 Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and FEV1 standard deviation value.

Table A43 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 standard deviation value, with no adjustment for the confounding factors.

Table A44 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 standard deviation value, with adjustment for the confounding factors.

Table A45 Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and age of diagnosis.

Table A46 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Heme oxygenase-1 and the age of diagnosis, with no adjustment for the confounding factors.

Table A47 Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Heme oxygenase-1 and the age of diagnosis, with adjustment for the confounding factors.

Table A48 Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value.

Table A49 Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value.

Table A50 Frequencies of possible haplotypes generated for the Factor B gene by the FBAT program.

Table A51 Haplotype analysis for investigation of correlation between combinations of selected SNPs in Factor B and age of diagnosis by the FBAT program.

Table A52 Frequencies of possible haplotypes generated for the Complement factor 3 gene by the FBAT program.

Table A53 Haplotype analysis for investigation of correlation between combinations of selected SNPs in Complement.
factor 3 and FEV1 predicted value by FBAT program

Table A54

Frequencies of possible haplotypes generated for
the gene of Toll-like receptor 4 by the FBAT program

Table A55

Haplotype analysis for investigation of correlation
between combinations of selected SNPs in
Toll-like receptor 4 and age of diagnosis by
the FBAT program

Table A56

Haplotype analysis for investigation of correlation
between combinations of selected SNPs in
Toll-like receptor 4 and FEV1 standard deviation value
by the FBAT program

Table A57

Frequencies of possible haplotypes generated for the
Heme-oxygenase-1 gene by the FBAT program

Table A58

Haplotype analysis for investigation of correlation
between combinations of selected SNPs in
Heme oxygenase-1 and age of diagnosis by the
FBAT program

Table A59

Haplotype analysis for investigation of correlation
between combinations of selected SNPs in
Heme oxygenase-1 and FEV1 standard deviation value
by the FBAT program

Table A60

Haplotype analysis between the haplotypes formed
by the five selected SNPs in the Factor B gene and
age of Pseudomonas aeruginosa infection by Hapstat

Table A61

Haplotype analysis between the haplotypes formed by
the four selected SNPs in the Complement factor 3 gene
and age of Pseudomonas aeruginosa infection by Hapstat

Table A62

Haplotype analysis between the haplotypes formed
by the seven selected SNPs in the Toll-like receptor 4
gene and age of Pseudomonas aeruginosa infection
by Hapstat

Table A63

Haplotype analysis between the haplotypes formed by
the six selected SNPs in the Heme oxygenase-1 gene
and age of Pseudomonas aeruginosa infection by Hapstat
### List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Structure of the CFTR protein</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Function of CFTR protein</td>
<td>4</td>
</tr>
<tr>
<td>Figure 3</td>
<td>The classic complement pathway</td>
<td>8</td>
</tr>
<tr>
<td>Figure 4</td>
<td>The alternative pathway</td>
<td>9</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Function of Heme oxygenase-1 in lung disease</td>
<td>11</td>
</tr>
<tr>
<td>Figure 6</td>
<td>The mechanism underlying TaqMan assays</td>
<td>23</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Sample genotyping result of Factor B SNP 2557</td>
<td>31</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Example genotyping result for SNP 149 of HMOX1</td>
<td>33</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Normalization of the FEV1 percent predicted value by age</td>
<td>38</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Standard curve of optical density versus DNA concentration of the serially diluted λ DNA samples</td>
<td>43</td>
</tr>
</tbody>
</table>
Acknowledgments

I would like to take this opportunity to express my appreciation for those who had given me their valuable help and support:

Dr. Sandford for his supervision, patience, and precious time for guiding me and helping me, starting from the brain-storming of this project to the execution of the laboratory work and finally to the write-up of this thesis. The successful completion of this thesis would have been impossible without his guidance and supportive efforts.

Dr. Peter Paré, Dr. Pearce Wilcox and Dr. John Hill for being members of my supervisory committee and providing me with valuable opinions on my thesis.

The CF Modifier Gene Project team at UBC especially Roxanne Rousseau and at The Hospital for Sick Children in Toronto for supplying the DNA samples and phenotypic data of the participating patients and providing relevant information.

Loubna Akhabir and Dorota Stefanowicz for helping me in the laboratory work and providing me their professional technical advice.

Karey Shumansky for her expert knowledge and patient guidance in Statistics.

My family and friends for their understanding and support.
Chapter 1 Introduction

Cystic Fibrosis (CF) is one of the most common genetic recessive disorders among Caucasians. In 2000, the estimated rate of CF in Canada was 1 in 3608 births although the rate may decline if antenatal carrier screening of the general population is implemented. [1] Mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene located on chromosome 7 have been identified to be the cause of the disease by the technique of positional cloning. [2] A wide diversity of CF pulmonary phenotypes is observed among CF patients. Furthermore, they may suffer from other CF-related medical conditions, for example, obstructive azoospermia and pancreatitis. [3] It is clear that CFTR genotype alone does not account for the broad spectrum of severity of the disease. It is likely that secondary genetic factors separate from the CFTR locus significantly influence CF phenotypes, with these loci termed modifier genes. [4] A number of putative modifier genes have been identified, for example, mannose-binding lectin, glutathione S-transferases, transforming growth factor beta-1, tumor necrosis factor-alpha, beta-2 adrenergic receptor and HLA class II. [5] The investigation of CF modifier genes may offer new insights into the pathophysiology of CF and provide leads for new CF therapeutic interventions.

1. Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Mutations

Cystic Fibrosis is caused by a defect in the CFTR protein and the function of this chloride channel will be discussed in detail in the next section. At the molecular level, the CFTR gene is located at chromosome 7q31 and it includes 27 exons for a total size of 230kb. There are more than 1000 CFTR mutations [6] and five categories have been established to describe CFTR mutations with regard to their effect on CFTR function as a chloride channel. [7, 8] This classification scheme is summarized in the following table:
<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>Mutations that cause defective synthesis and thus total or partial absence of CFTR protein. Consequently, there is a loss of conductance of Cl⁻ by the channel.</td>
</tr>
<tr>
<td>Class 2</td>
<td>Mutations that affect maturation of CFTR protein. CFTR mRNA is formed, but the protein fails to mature and does not traffic correctly to the cell membrane. The protein is either absent or present in very small quantities in the membrane.</td>
</tr>
<tr>
<td>Class 3</td>
<td>Mutations that disturb the regulation of the Cl⁻ channel and these usually occur in the ATP binding domains. CFTR protein is produced and traffics to the apical cell membrane.</td>
</tr>
<tr>
<td>Class 4</td>
<td>Mutations that affect the conductance of the Cl⁻ channel which result in the reduction of ion flux and/or modification of selectivity. CFTR protein is produced and traffics to the apical cell membrane.</td>
</tr>
<tr>
<td>Class 5</td>
<td>Mutations that alter the stability of mRNA or the stability of mature CFTR protein. Lower levels of CFTR protein are produced and traffic to the apical cell membrane is reduced.</td>
</tr>
</tbody>
</table>

Among the more than 1000 CFTR mutations, delta F508 is the most common and it is detected in about 70% of CF genes in North America and North Europe. [9] Delta F508 is a mutation that results in the deletion of a phenylalanine amino acid at position 508 of the hydrophilic region for ATP binding which is called the nucleotide binding fold (NBF). This mutation is categorized as a Class 2 mutation. For patients with this type of CFTR mutation, no or a small amount of CFTR protein can be produced in the lungs, pancreas or other organs. The transport of chlorides will be aborted or reduced which results in the thickening of mucus in the affected organs. This condition favors bacterial infection and eventually, CF patients are more susceptible to infection in the respiratory tract because of their compromised defense mechanism. [10]

2. CFTR Protein and its Function

The CFTR protein is a membrane associated, N-linked glycoprotein located at the apical membrane (Figure 1). CFTR is composed of two repeated motifs with a total size of 1480 amino acids. Each motif consists of a hydrophilic membrane-spanning domain with six helices and a hydrophilic NBF region which is responsible for ATP binding. The two motifs are connected by a cytoplasmic regulatory region (R domain) with a number of charged residues. [9]
Figure 1: Structure of the CFTR protein. There are three major regions in a CFTR protein: (1) two membrane-spanning domains (TMD1 and TMD2) each composed of six transmembrane segments; (2) two nucleotide binding domains where binding of ATP occurs (NBD1 and NBD2); and (3) a regulatory region (R Domain) which is responsible for the phosphorylation by protein kinase A. As indicated in the diagram, some of the common CFTR mutations are located in the NBDs. [11]

CFTR has been classified as a member of the family of ATP-binding cassette (ABC) transporters. The CFTR mainly functions as a chloride channel in the lungs, pancreas, liver, digestive tract, reproductive tract and skin. Upon phosphorylation by protein kinase A in the regulatory region, a conformational change takes place for transporting the chloride ions across the membrane with the consumption of ATP as the energy source. In addition to working as a cAMP-induced chloride channel, CFTR is also capable of regulating other ion channels, for example, it regulates an outwardly rectifying chloride channel (ORCC), an epithelial sodium ion channel (eNaC), and at least two outwardly rectifying potassium ion channels (RMOK1 and RMOK2), as shown in Figure 2. [12]
Figure 2: Function of the CFTR protein. CFTR protein can be found at the apical membrane of epithelial cells in the lungs, pancreas, liver, digestive tract, reproductive tract and skin. With normally functioning CFTR protein, chloride ions are transported across the membrane with ATP as the energy source upon phosphorylation by protein kinase A. In addition, CFTR can then regulate the transport of chloride ions through an outwardly rectifying chloride channel (ORCC) in the presence of ATP; and the transport of sodium ions through an epithelial sodium ion channel (eNaC) in the absence of ATP. CFTR also regulates the pumping out of potassium ions through an outwardly rectifying potassium ion channel (ROMK). [13]

3. Diagnosis and Clinical Symptoms of CF Disease

Detection of CFTR mutations by genotyping is a powerful method for diagnosing CF patients, and it is especially important for infants where both parents are known to be CF carriers. Clinical complications (for example, meconium ileus and failure to thrive) may present in some of the new-born babies which can be clues for the diagnosis of CF in those infants. Testing the sweat chloride level can confirm the diagnosis in such situations, since CFTR is the only channel for
reabsorption of chloride ions in the sweat glands. CF patients with abnormal chloride channels are found to have a sweat chloride level which is five times higher than unaffected individuals. [14] However, patients with some forms of CFTR mutation show milder clinical symptoms and a borderline sweat test result. Consequently, DNA screening should be used for confirmation of the disease.

In addition to the lungs, CFTR is expressed in other organs, for example, liver, pancreas, intestines and reproductive tract. Consequently, CF patients commonly suffer from problems in those organs. Approximately 85% of CF patients have pancreatic insufficiency. [15] Complications can result due to the lack of functional CFTR protein in the ductules of the pancreas, including the failure to produce adequate bicarbonate ions and water; the blockage of ductules further results in inability of the pancreatic enzymes to reach the intestine in order to facilitate normal digestion. However, exogenous enzymes can be prescribed to patients in order to correct for problems in this exocrine organ. [16] Intestinal obstructions, both meconium ileus at birth and distal intestinal obstruction syndrome (DIOS) later in life, are another frequently seen manifestation in CF patients. Meconium ileus and DIOS patients fail to pass meconium at birth and suffer from excessive mucus production, respectively. Constipation is a common outcome. [17] Some CF patients may suffer from CF-related liver disease. Without functional CFTR in the lining of the epithelial cells in the biliary ductules, the ductules become blocked with mucous secretion and the accumulation of eosinophils. [18] Further symptoms include obstructive cirrhosis, splenomegaly and hypersplenism. A total of 1-2% of deaths in CF patients are liver-related. [19] Other CF complications include azoospermia in males [20] and chronic sinusitis.

4. *Pseudomonas aeruginosa* Infection in CF Patients

*Staphylococcus aureus, Pseudomonas aeruginosa, Burkholderia cepacia, Hemophilus influenzae and Xanthomonas maltophilia* are common bacteria which can be found in the lungs of many CF patients. Among them, *Pseudomonas aeruginosa* is found in nearly 80% of CF patients and the morbidity and the mortality rates increase when individuals are chronically colonized with this pathogen. [21] The bacteria adhere to
host tissues via their lipopolysaccharide layer, which also protects them from digestion and lysing by leukocytes. The organisms mainly grow in the bronchioles of CF patients as a biofilm and they are relatively resistant to the host's defense system and antibiotics by causing a low phagocytic response. [22] Furthermore, *Pseudomonas aeruginosa* can secrete several toxins and chemicals upon infection, for example, lipopolysaccharide, exotoxin A, exoenzyme S, elastase, alkaline proteinase, and phospholipase C protein. [23] The body's immune system tries to eradicate the bacteria; however, CF patients may suffer from lung infection due to failure of recruitment of innate defense mechanisms and involvement of immune complex formation. [24] This inflammatory situation induces a phenotypic shift of the pathogen from non-mucoid to mucoid type, which leads to exacerbation of the disease [25] The reason why antibiotics or other therapeutic methods cannot eradicate *Pseudomonas aeruginosa* infection in CF patients is still unknown. [26]

Some individuals are detected to be CF patients shortly after birth due to a positive family history of CF and/or acute clinical symptoms, while others are diagnosed at a later age. Usually the former group of patients is found to exhibit more serious CF-related problems, and the type of CFTR mutations present may be one of the causes of this increased severity, in addition to modifier genes. Not every CF patient becomes infected by *Pseudomonas aeruginosa*, however, the rates of morbidity and mortality increase significantly once the individual is chronically colonized with this pathogen. We hypothesized that the particular genotype of any of the polymorphisms in the four candidate genes that we investigated might contribute to the difference between patients with regard to the status of *Pseudomonas aeruginosa* infection and therefore the severity of the disease.

5. Candidate Modifier Genes

Previous studies have identified some putative modifier genes which may contribute to the diversity of disease severity in CF patients. The aim of this project was to investigate specific genes that are involved in *Pseudomonas aeruginosa* infection in CF patients and to investigate relationships between polymorphisms in those genes
and disease severity. Four candidate modifier genes were chosen for experiment in this project, namely, Factor B, Complement Factor 3, Toll-like Receptor 4 and Heme Oxygenase-1.

(a) Factor B and Complement Factor 3

The human immune system is responsible for providing different degrees of defenses which protect the body from being harmed by foreign substances or antigens. There are many components in the immune system and one of them is the cascade of complement reactions which helps in fighting infections. Complement reactions are a series of non-specific reactions in the host defense system to destroy invading microorganisms. Complement reactions can be sub-divided into two pathways: the classical pathway and the alternative pathway. The classical pathway is initiated by the binding of antibodies to complement protein Clq which results in cleavage of Clq into Cls and Clr. The complete set of reactions in the classical pathway is shown in Figure 3:
Figure 3: The classical complement pathway. Upon binding of the microbe to the antibody, Clq is cleaved intoClr and Cls which both catalyze the formation of C4b2b complex. After combining with C3b, the C4b2b3b complex catalyzes the cleavage of C5 into C5b which eventually results in the synthesis of membrane attack complex (MAC). (Note: each dotted line shows a catalytic effect on the corresponding reaction in box)

In comparison, the alternative pathway is antibody-independent. Consequently, the presence of the microbes can directly trigger the alternative pathway. However, the two pathways share some of the same proteins. The complete set of reactions in the alternative pathway is shown in Figure 4:
Figure 4: The alternative pathway. The presence of the microbe (light beige ellipses) can directly trigger the formation of C3b-Factor B complex. After modification of the complex by Factor D and joining of another C3b unit, the resulting complex of C3bBb3b cleaves C5 into C5b and subsequently leads to the synthesis of the membrane attack complex (MAC). (Note: each dotted line shows a catalytic effect on the corresponding reaction in box)

Both complement pathways result in the generation of the membrane attack complex (MAC) which leads to neutrophil activation and bacterial cell lysis. In addition to the major triggering factors of the two pathways, they also differ in the reactions for the generation of complement C3. The key player is C4 and factor B in the classical pathway and alternative pathways, respectively. It has been found that the alternative pathway is critical for the generation of a protective response against *Pseudomonas aeruginosa* in a murine model of pneumonia. [27] In that study, C3 and factor B deficient mice were infected with *Pseudomonas aeruginosa* via intranasal inoculation and they were found
to have a higher mortality rate when compared with wild-type mice. However, the same outcome was not seen for C4 deficient mice. This analysis indicated that alternative pathway was critical for the eradication of *Pseudomonas aeruginosa* infection from the host. This result motivated us to investigate the role of the alternative pathway in host defense against *Pseudomonas aeruginosa* infection in CF patients. Consequently, we hypothesized that the alternative pathway also plays a key role in eradication of *Pseudomonas aeruginosa* from the lungs of CF patients and that polymorphisms in genes encoding for factor B and C3 are associated with severity of disease in CF patients.

(b) Toll-like 4 Receptor

The lipopolysaccharide (LPS) molecules of *Pseudomonas aeruginosa* are composed of three main parts, namely lipid A, core and O-polysaccharide. [28] Modifications of lipid A are observed in *Pseudomonas aeruginosa* from CF patients, for example, the pathogen synthesizes a unique hexa-acylated lipid A containing palmitate and aminoarabinose during adaptation to the CF airway while it produces a novel hepta-acylated lipid A in CF patients with severe pulmonary disease. [29] Lipid A of *Pseudomonas aeruginosa* is recognized by toll-like receptor 4 on the epithelial surface of the lungs. A 222 amino acid region in the extracellular portion of human toll-like receptor 4 has been identified using *Pseudomonas aeruginosa* lipid A with different levels of acylation and this region has been found to be crucial for the recognition of the lipid A which is specifically found in CF patients. [30] Toll-like receptor 4 is a transmembrane LPS receptor. Release of antimicrobial peptides, inflammatory cytokines and chemokines, and co-stimulatory molecules is triggered upon binding of *Pseudomonas aeruginosa*. All of these result in the activation of the innate immune response to the pathogens. Therefore, we hypothesized that polymorphisms in the gene encoding toll-like receptor 4 could lead to failure in recognition of *Pseudomonas aeruginosa* via its lipid A, the accumulation of the pathogens in lungs of CF patients and therefore increased disease severity.
(c) Heme Oxygenase-1

Heme oxygenase-1 (HO-1) is an enzyme which is responsible for the catalysis of the oxidative reaction of heme to anti-oxidant molecules, bilirubin and carbon monoxide. [31] There are two isoenzymes of heme oxygenase-1: an inducible HO-1 and a constitutively expressed HO-2. [32] HO-1 is induced by a wide variety of stimuli including conditions of oxidative stress, inflammatory agents, transforming growth factor beta and heat shock. [33] Because of their defective immune defense system, CF patients usually develop secondary acute or chronic bacterial infections and recruitment of the inflammatory defense mechanisms. Excessive inflammation can result in tissue damage which makes the condition of CF patients more severe. However, HO-1 plays a major role in resolving this situation. The major roles of HO-1 have been shown to include anti-inflammatory, anti-apoptotic, and anti-proliferative effects [34] which are illustrated in Figure 5:

Figure 5: Function of heme oxygenase-1 in lung diseases. Patients with lung diseases like cystic fibrosis may suffer from inflammation, oxidative stress and apoptosis. Tissue injury can be the final result. However, these complications can be modulated by the production of heme oxygenase which can catalytically break down heme into carbon monoxide, iron and biliverdin-IXα. All those components have a negative effect on the detrimental processes mentioned above and therefore can reduce tissue injury. Heme oxygenase has been conclusively shown to be anti-inflammatory, anti-apoptotic, and anti-proliferative.
As mentioned before, the recruitment of the inflammatory defense mechanisms in CF patients against *Pseudomonas aeruginosa* infection can lead to tissue damage in lungs. Due to the anti-inflammatory, anti-apoptotic and anti-proliferative effect, HO-1 can modulate these adverse outcomes. [35] It has been demonstrated that HO-1 level is raised in CF patients and it is responsible for cytoprotective effects against *Pseudomonas aeruginosa* infection. [36] The levels of HO-1 and related by-products in CF patients and controls were examined by different methods [36]: (1) heme expression in the lungs was measured by both immunochemistry and quantitative reverse transcription PCR; (2) the level of acute inflammation and an increase in oxidant stress was measured by myelo-peroxidase staining; and (3) iron status was assessed by ferritin staining. All the above tests revealed an increase in HO-1 level in diseased lungs when compared with normal controls. Furthermore, investigation was performed to determine whether *Pseudomonas aeruginosa* infection was a direct cause of upregulation in HO-1 expression. A cell line of human CF bronchial epithelial cells (IB3.1) was treated with either *Pseudomonas aeruginosa* or LPS, however, no significant increase of HO-1 protein could be detected in either sample. The same authors also evaluated the survival of IB3.1 cells overpressing HO-1 in response to *Pseudomonas aeruginosa* infection by transfecting the IB3.1 cells with either pc DNA3.1 empty vector or pc 3.1 DNA 3.1/HO-1 vector [36]. Cells transfected with HO-1 vectors were shown to have a higher survival rate when compared with those with empty vectors. As a result of this study, we hypothesized that polymorphisms in the HO-1 gene could result in defective synthesis of HO-1, which in turn could lead to failure in inhibition of tissue injury due to *Pseudomonas aeruginosa* infection and increased severity of disease in CF patients.

6. Single Nucleotide Polymorphisms

This project includes the above four candidate genes, i.e., Factor B, Complement factor 3, Toll-like receptor 4 and Heme Oxygenase-1. A total of 22 single nucleotide polymorphisms (SNPs) in these genes were selected. A detailed description of the SNP selection is presented in the Materials and Methods section. A SNP is defined as a single-base variation in DNA sequence with an allele frequency of at least 1% in a
population. Not all SNPs result in alteration of protein structure and function. For example, a synonymous SNP usually occurs when the third nucleotide in the codon is altered but both the synthesis and function of the resulting protein remain unchanged. However, the structure and function of the protein are often changed if either a non-sense SNP or non-synonymous SNP is present. A non-sense SNP leads to premature termination in the elongation of the polynucleotide chain while a non-synonymous SNP results in substitution of one amino acid for another. In addition, the position of the SNP in the entire polynucleotide chain is also important when analyzing its effect. SNPs in coding regions are a common cause of many monogenic disorders, however, SNPs in promoters or other regulatory regions can also show their significance in determination of timing, location and level of gene expression. Even if no impact is found for SNPs in non-coding regions, they may be important in the investigation of the genetic aspects of a disease since they can be used as disease-markers for further investigation and/or analysis.

7. Association Studies

Linkage analysis can be used to determine the genetic location of a disease gene. The goal of the analysis is to identify an allele of a gene that is co-inherited with the disease. Linkage analysis is a powerful technique for the identification of genes that cause monogenic disorders. For example, it was used for the identification of the CFTR as the causal gene in cystic fibrosis. However, this approach is less powerful when a more complex phenotype is to be studied i.e. a trait influenced by several genetic and environmental factors. Therefore, we used an association approach in this project since severity of the disease is likely a complex phenotype. Case-control studies and family based trios (i.e. offspring and both his/her parents) are two commonly used types of association studies. An example of an analysis that uses trios is the transmission disequilibrium test (TDT). The TDT requires genetic information from both parents and the affected individual since it tests for the transmission from heterozygous parents to the offspring. The TDT provides a better approach in the investigation and establishment of genetic associations under certain circumstances. If a case-control method is used for analysis, false-positive results may be
generated due to insufficient matching of the case and control groups for genetic background. However, there are still some drawbacks associated with the TDT method. First, the marker has to be in close proximity to the disease-causing gene. Otherwise, no association can be identified. Second, at least one of the parents has to be heterozygous, because if both parents are homozygous, the child will be homozygous as well and nothing can be concluded from the analysis. Therefore, the case-control approach was also included in the analysis of this project. Case-control studies test for the prevalence of a specific allele in the two groups and two approaches can be taken. First, the cases are defined as CF patients with severe lung disease and the controls are defined as CF patients with mild lung disease. The division of mild and severe lung disease is based on pulmonary function test data. Second, the patients are stratified by genotype and differences in the means of the outcome variables are investigated. In this study we utilized the latter case-control approach and supplemented the analysis with the family-based design.

8. Thesis Objectives

1. To perform a literature review to select candidate CF modifier genes. The following genes were selected:
   (a) Complement factor 3 and factor B protein - deficiency in recruitment of complement reactions against *Pseudomonas aeruginosa* infection in CF patients.
   (b) Toll-like receptor 4 - attachment of *Pseudomonas aeruginosa* onto host cells via binding of its LPS in CF patients.
   (c) Heme oxygenase-1 - cytoprotective effects against *Pseudomonas aeruginosa* infection in CF patients.
2. To pick potential SNPs within each candidate modifier gene utilizing the LD Select software.
3. To genotype DNA samples of participating individuals (trio members) at the selected SNPs by TaqMan assays.
4. To analyze the genotyping results with respect to phenotypic traits in an attempt to determine the presence of any association with disease severity.
Chapter 2 Materials and Methods

1. Patient Recruitment

This is a sub-study of a large, Canada-wide and international endeavor: the Canadian Consortium for Cystic Fibrosis Modifiers (http://www.cfmod.ca). There were 1674 individuals from 558 trios (558 patients and both of their parents) in this sub-study who were recruited from Cystic Fibrosis clinics in participating hospitals in Canada. Patients with a diagnosis of CF on the basis of clinical signs, elevated sweat chloride values and/or positive genotyping for the CFTR gene were recruited for the study. CF patients who had received a lung transplant were also recruited and pulmonary function data for these individuals were collected prior to transplantation. The recruitment of patients and their family members was based upon several criteria: (a) satisfying ethical requirements (for example, willingness to give informed consent); (b) willingness to provide blood/DNA sample for the study; (c) availability of verified clinical information. Blood samples from patients in different provinces were sent to the Hospital for Sick Children in Toronto for extraction of DNA and, for the patient samples, establishment of cell lines.

2. Quality Control

As mentioned before, this project was a sub-study of the Canadian Consortium for Cystic Fibrosis Modifiers. Genotyping was also carried out for all samples for other analyses by our colleagues in Toronto. Non-Mendelian inheritance was found in some of the families by both Toronto and by the use of the FBAT program in this sub-study. Mendelian errors were detected for many SNPs in some of the families. This implied that there had been sample mislabeling or there was non-paternity. Consequently, a total of 23 families were deleted from this study. However, when Mendelian inconsistency was detected only at one SNP in a given family, analysis was continued with those families except that genotypes for those particular non-Mendelian SNPs were excluded. In total, DNA samples from 1605 individuals/535 trios (535 patients and both of their parents) were used for the analysis.
3. Selection of SNPs

Several polymorphic sites were picked for each candidate modifier gene. Information about polymorphisms for each candidate gene was obtained from the websites of Innate Immunity Programs for Genomic Applications or IIPGA (http://innateimmunity.net/) and UW-FHCRC Variation Discovery Resource or Seattle SNPS (http://pga.mbt.washington.edu/). By using the LD Select software (http://droog.gs.washington.edu/ldSelect.html)[37], all sites of polymorphism in each gene were divided into groups/bins according to specific settings as listed below. Only one SNP was selected and examined for each group. An assumption was made that phenotypic characteristics were more likely be affected if the site of polymorphism brought about a change of amino acid. Based upon this assumption, two stages were applied for the selection of a particular SNP in each group: (a) if one of the SNPs in the group resulted in a change of amino acid, it would be selected for experimentation; (b) if none of the SNPs in the group resulted in a change of amino acid, an arbitrary tagSNP was picked for experimentation. A tagSNP is a representative SNP in each bin which shows a high degree of linkage disequilibrium with all the other SNPs in the bin.

Linkage disequilibrium (LD) is a measure for detecting the presence of any non-random association of alleles at two or more loci on the same or different chromosomes. A strong allelic association is termed to be complete LD, which means the involved alleles are found together at a high frequency in the population. In general, linkage disequilibrium can be represented by the formula \( P_{A1}P_{B1} \neq P_{A1B1} \); whereas linkage equilibrium can be shown by the equation \( P_{A1}P_{B1} = P_{A1B1} \) (\( P_{A1} \) and \( P_{B1} \) are the probabilities of allele A1 and allele B1, respectively in the population and \( P_{A1B1} \) is the probability of the A1B1 haplotype). There are several parameters for representing the degree of linkage disequilibrium but the two most commonly used are Lewontin's D' and a measure that can be interpreted as a squared correlation coefficient, \( r^2 \). D is determined by the difference between the observed haplotype frequency and the expected haplotype frequency. Then D' is calculated by the equation \( D' = (D / D_{\text{max}}) \), where \( D_{\text{max}} \) is the maximum difference between the observed haplotype frequency and the expected haplotype frequency.
frequency that can be observed in a given population. The value of $D'$ ranges from 0 to 1: 0 represents linkage equilibrium and 1 represents complete linkage disequilibrium i.e. one or more haplotypes are not present in the population. A disadvantage of the $D'$ measure is that it overestimates LD in small samples.

The evaluation of $r^2$ depends on the value of $D$ and the allele frequency at every locus involved. It is calculated by dividing $D^2$ by the product of all allele frequencies, as illustrated by the equation $r^2 = (D^2 / a_1^*a_2^*b_1^*b_2)$. $r^2$ also lies between 0 and 1; with 0 indicating linkage equilibrium and 1 indicating perfect linkage disequilibrium. This measure takes value 1 if only two haplotypes are present. For low allele frequencies $r^2$ is a more reliable measure of LD than $D'$. $r^2$ measures statistical association and there is a simple inverse relationship between this measure and the sample size required to detect association between a susceptibility locus and a marker SNP.

The SNPs in each of the candidate genes were grouped according to the set conditions listed below. It would be time- and effort-consuming if all the SNPs in a gene were required to be genotyped. By arranging them according to the specific settings listed below, it was possible to achieve the tasks of this project without genotyping every single SNP.

Using the approach described above we selected a panel of SNPs for genotyping in each of our candidate genes. For the factor B gene five SNPs were selected including a tagSNP for an amino acid changing polymorphism (Table 1).
Table 1: Factor B polymorphism selection (the criteria for grouping all SNPs by the LD Select software package: $r^2=0.8$ and minor allele frequency $>0.1$). The positions of the SNPs are described in reference to the coding sequence of the factor B gene (accession number AF551848). No TaqMan assay could be created for testing of SNPs in Bin3. Therefore, only five SNPs were included in the analysis for Factor B. Although SNP1802 in Bin1 results in an amino acid change from arginine to glutamine, SNP8311 was selected due to the failure of production of an appropriate TaqMan assay for SNP1802.

<table>
<thead>
<tr>
<th>Bin</th>
<th>Position of SNP</th>
<th>Amino acid change</th>
<th>Selected SNP for the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1802</td>
<td>Arg to Gln</td>
<td>8311</td>
</tr>
<tr>
<td></td>
<td>4573</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5180</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7541</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7580</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8311</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4022</td>
<td>n/a</td>
<td>4022</td>
</tr>
<tr>
<td></td>
<td>9878</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5162</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>9099</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2557</td>
<td>n/a</td>
<td>2557</td>
</tr>
<tr>
<td>5</td>
<td>6484</td>
<td>n/a</td>
<td>6484</td>
</tr>
<tr>
<td>6</td>
<td>7202</td>
<td>n/a</td>
<td>7202</td>
</tr>
</tbody>
</table>

The same tagSNP approach was applied to the complement factor 3 gene. However, using the parameters of $r^2=0.8$ and minor allele frequency $>0.1$ a total of more than 40 bins resulted. This large number of SNPs was beyond the scope of this study. Therefore, the minor allele frequency cutoff was increased to $>0.45$ in order to study only common SNPs. This resulted in the selection of four polymorphisms (Table 2).
Table 2: Complement factor 3 polymorphism selection (the criteria for grouping all SNPs by the LD Select software package: $r^2 = 0.8$ and minor allele frequency $> 0.45$). The positions of the SNPs are described in reference to the coding sequence of the complement factor 3 gene (accession number AY513239). Due to the large size of this gene (46kb) and consequently large number of SNPs, a higher minor allele filter was chosen so that a feasible number of tagSNPs could be identified. No SNPs were picked for the analysis due to failure of the manufacture of TagMan assays for SNPs in Bins 2 and 5.

<table>
<thead>
<tr>
<th>Bin</th>
<th>Position of SNP</th>
<th>Amino acid change</th>
<th>Selected SNP for the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43118</td>
<td>n/a</td>
<td>43118</td>
</tr>
<tr>
<td></td>
<td>43179</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>43928</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44692</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>27159</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>27678</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>28433</td>
<td>n/a</td>
<td>28795</td>
</tr>
<tr>
<td></td>
<td>28795</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>963</td>
<td>n/a</td>
<td>963</td>
</tr>
<tr>
<td>5</td>
<td>25884</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>6</td>
<td>36735</td>
<td>n/a</td>
<td>26735</td>
</tr>
</tbody>
</table>

The original selection criteria i.e. those utilized for the factor B gene resulted in the identification of seven tagSNPs in the TLR4 gene and these are detailed in Table 3.
Table 3: TLR4 polymorphism selection (the criteria for grouping all SNPs by the LD Select software package: $r^2=0.8$ and minor allele frequency $>0.1$). The positions of the SNPs are described in reference to the coding sequence of the TLR4 gene (accession number NM_138556).

<table>
<thead>
<tr>
<th>Bin</th>
<th>Position of SNP</th>
<th>Amino acid change</th>
<th>Selected SNP for the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1893</td>
<td>n/a</td>
<td>11912</td>
</tr>
<tr>
<td></td>
<td>2032</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2437</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7764</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11912</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>16649</td>
<td>n/a</td>
<td>17050</td>
</tr>
<tr>
<td></td>
<td>17050</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17447</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17923</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2856</td>
<td>n/a</td>
<td>2856</td>
</tr>
<tr>
<td></td>
<td>10478</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11541</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>851</td>
<td>n/a</td>
<td>851</td>
</tr>
<tr>
<td></td>
<td>11995</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1859</td>
<td>n/a</td>
<td>1859</td>
</tr>
<tr>
<td></td>
<td>10329</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>9263</td>
<td>n/a</td>
<td>9263</td>
</tr>
<tr>
<td>7</td>
<td>15844</td>
<td>n/a</td>
<td>15844</td>
</tr>
</tbody>
</table>

Finally, we identified eight tagSNP bins in the heme oxygenase-1 gene and of these six SNPs could be assayed by the TaqMan methodology (Table 4).
Table 4: Heme oxygenase-1 polymorphism selection (the criteria for grouping all SNPs by the LD Select software package: \( r^2 = 0.8 \) and minor allele frequency > 0.1). The positions of the SNPs are described in reference to the coding sequence of the heme oxygenase-1 gene (accession number AY460337). Only six SNPs were chosen for this project due to failure in the synthesis of TaqMan assays for SNPs in Bins 6 and 7.

<table>
<thead>
<tr>
<th>Bin</th>
<th>Position of SNP</th>
<th>Amino acid change</th>
<th>Selected SNP for the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>149</td>
<td>n/a</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>7325</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12825</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12832</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12860</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12992</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13286</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13354</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15028</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3303</td>
<td>n/a</td>
<td>3303</td>
</tr>
<tr>
<td></td>
<td>5079</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13449</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2007</td>
<td>n/a</td>
<td>9531</td>
</tr>
<tr>
<td></td>
<td>9531</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16442</td>
<td>n/a</td>
<td>16442</td>
</tr>
<tr>
<td></td>
<td>17893</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2790</td>
<td>n/a</td>
<td>2790</td>
</tr>
<tr>
<td>6</td>
<td>15382</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>7</td>
<td>17922</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>8</td>
<td>1038</td>
<td>n/a</td>
<td>1038</td>
</tr>
</tbody>
</table>

4. TaqMan Assays

The TaqMan allelic discrimination assay is a qualitative analytic method for detection of the presence of different alleles at any nucleotide in a gene. The assay utilizes fluorescently labeled allele-specific probes and the 5'-> 3' exonuclease activity of Taq DNA polymerase (Figure 6). A DNA sample is first amplified by the Polymerase Chain Reaction (PCR)(discussed in a later section). In addition to the TaqMan Universal PCR master mix, a mixture of specific primer/probe solution is added into the DNA sample for the PCR. After
the double-stranded template DNA is denatured into single strands, the complementary PCR primers hybridize to the template DNA 150-200 nucleotides apart and encompassing the SNP site. Furthermore, the allele-specific probes that are complementary to the SNP site hybridize to the template at this step.

Generally, the probes are molecules of about 20-nucleotides long with a reporter dye attached to the 5’ end and a quencher dye attached to the 3’ end. Two types of probes are included, one for each allele of the SNP being assayed. The two probes are differentiated by their reporter dyes: one probe contains VIC and the other FAM. The two reporter molecules absorb light and then release it at different wavelengths. Before polymerization, energy is transferred to the quencher molecule when light hits the reporter molecule since they are located close to each other. From this process, called Fluorescence Resource Energy Transfer, no fluorescence can be detected since the energy is then released in the form of heat. However, the probe is degraded by the 5’-> 3’ exonuclease activity of Taq DNA polymerase as polymerization is initiated from the primer and proceeds along the strand. Since the reporter molecule is no longer linked to the quencher molecule, it absorbs light energy and emits energy as light of a specific wavelength.

Therefore different fluorescence is detected depending on whether the VIC-labeled probe the FAM-labeled probe is degraded or both are degraded. From this, the presence of both alleles can be deduced for each individual. In addition to the two probes, there is also a third dye called ROX that is included in the TaqMan assay. ROX is free in the master mix solution and is used for normalizing the well to well variation in volume. Therefore, the signal intensity detected for either the FAM or VIC molecules was corrected by the ROX signal (although this is less of a concern than in quantitative real-time PCR applications).
Figure 6: The mechanism underlying TaqMan assays. Before the reaction, the primer is complementary base-pairing to the segment which is about 150-200 nucleotides away from the SNP, while the probe is attached to the site of SNP. As elongation of nucleotides proceeds, the probe is degraded which separates the dye molecule from the quencher.

5. Primers and Probes for Genotyping

Primers and probes for each SNP were ordered from Applied Biosystems (Foster City, CA). The sequences of all primers and probes are listed in tables 5 and 6:
Table 5: The sequence of all primers used for SNP assays

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene_SNP</th>
<th>Left Primer</th>
<th>Right Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>TCGCACCTGCCAAGTGAA</td>
<td>CTCACCTCCGTTGTCACAGATC</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>GCGACAGGGAGGACCAC</td>
<td>CTCCATGCCCACAGGATTCCT</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>CAGAACCTAGCTTAGAGG</td>
<td>GTCTCACCACACTGTCATTC</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>TGGGTCCCTAGTCTGATTC</td>
<td>TGCTGGAAGCATGGGCTGTT</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>GGTAGCCTTGGCCTCCACACTC</td>
<td>GTGGGAGGGAGGAAGGGAAGA</td>
<td></td>
</tr>
<tr>
<td><strong>Complement Factor 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>CTGCAGTGAGCTGTGATTGC</td>
<td>CCATCCCTGTTGACGACATATT</td>
<td>CTTT</td>
</tr>
<tr>
<td>C3_28795</td>
<td>CGAGACCTGCTCTCTCTCTCA</td>
<td>GGGTGCGCTTGAGAAGATG</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>TGGAGTGTGAGGAGCTTATTTC</td>
<td>AGAAGGTCAACACACAGCTT</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>CTCCAGCTGTGAGACTGAGACGAGA</td>
<td>GCCAGATGTGATGAGAGAGAATG</td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>TGTTTGGACACATGCTGCGCTTT</td>
<td>GGGATACACACTCAATCATTTC</td>
<td>CTGACCT</td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>AGGGATAGGACTGGGCTTATTGGAAT</td>
<td>CTGGCTTCTACACCAAGTGACCA</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>GCAAGCTTCTGCTATGATTAACAACTGA</td>
<td>CAAGATGAGGTTGACAGGACCTTCA</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>CTGGAAACTGATATAAGTGTGAGACAGACTACA</td>
<td>CTTGACCTCCACCACACAGAGAAG</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>GCTGTGCTGATGAGACACGTACACATTCTAGGAC</td>
<td>AGCAGATGAGGTTGACAGGACCTTCA</td>
<td></td>
</tr>
<tr>
<td>TLR4_15844</td>
<td>GTGGGGGCAATGCTCTTGGACGCTGCATTGGTA</td>
<td>ACCCATATTTCACACACAGACAT</td>
<td>TCTGAAG</td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>CTGGGTTGCTTTCCATGCTCATTCTCTGAGG</td>
<td>TTCAAATACACACAGGCTCTGAT</td>
<td></td>
</tr>
<tr>
<td><strong>Toll-like receptor 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>GGGCCCTGGCCCTTGCA</td>
<td>GGAGCCCCAGACCTTGGTAG</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CGACAGAGGAGGAGGAGAGA</td>
<td>CTGGTCTAGAAAGGAAAAAGGACCTTAACAA</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>GGGTGTGCTAAGTTCTGCTGT</td>
<td>CCAGAAAGGTGGAGAGGCA</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CACCGGCCGATGGA</td>
<td>AAGGCGCCGCTCCC</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>CACCTACTCTGATAGAGGAT</td>
<td>CCTCCTCTGTCAGCAGACACT</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>TCGGTTAGGAGAAGTGGTAGATGG</td>
<td>CCTGCGGTCAGCAGAGGACCT</td>
<td></td>
</tr>
</tbody>
</table>
Table 6: The sequences of all probes used for the SNP assays

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene_SNP</th>
<th>VIC</th>
<th>FAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td>BF_2557</td>
<td>TGGCCGGTGAGGTG</td>
<td>TGGCCGGATGGAGTG</td>
</tr>
<tr>
<td></td>
<td>BF_4022</td>
<td>TTTGTAGTCAAAGGTGAAC</td>
<td>TTTGTAGTCAAAAGTTGAAC</td>
</tr>
<tr>
<td></td>
<td>BF_6484</td>
<td>CATTGCCCTTTGTCACTC</td>
<td>ATTGCCCTTCGTCACCTC</td>
</tr>
<tr>
<td></td>
<td>BF_7202</td>
<td>CAGCTAAGACCAAGCA</td>
<td>CAGCTAAGACCAAGCA</td>
</tr>
<tr>
<td></td>
<td>BF_8311</td>
<td>CTAGAGGCTTGAGAGAGA</td>
<td>CTAGAGGCTTGAGAGAGA</td>
</tr>
<tr>
<td>Complement Factor 3</td>
<td>C3_963</td>
<td>CTGAGTGACAGAATGA</td>
<td>TGACTGACAGAATGA</td>
</tr>
<tr>
<td></td>
<td>C3_28795</td>
<td>TCCCCGTAGTCCCCCA</td>
<td>CCCCTGGGTCCCCA</td>
</tr>
<tr>
<td></td>
<td>C3_36735</td>
<td>AATACAAATCTGGTACTCC</td>
<td>ACAATCTGGTACTCC</td>
</tr>
<tr>
<td></td>
<td>C3_43118</td>
<td>CAGCCCGTGGGTCTT</td>
<td>CAGCCCGTGGGTCTT</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_851</td>
<td>CTGGAAGAGCACAACATGA</td>
<td>TGGAAGAGCACAACATGA</td>
</tr>
<tr>
<td></td>
<td>TLR4_1859</td>
<td>CAGGTACCCAGACAAC</td>
<td>CGGTACCCAGACAAC</td>
</tr>
<tr>
<td></td>
<td>TLR4_2856</td>
<td>CTTCCACCCACACTTTATT</td>
<td>TCCACCAACCGTTATT</td>
</tr>
<tr>
<td></td>
<td>TLR4_9263</td>
<td>TTTAAAAGGTAATTTTG</td>
<td>TTTAAAAGGTAATTTTG</td>
</tr>
<tr>
<td></td>
<td>TLR4_11912</td>
<td>ACTTATGATTTATTTTCG</td>
<td>TTTATGATTTATTTTCG</td>
</tr>
<tr>
<td></td>
<td>TLR4_15844</td>
<td>ACATCCACCTTTCCC</td>
<td>ACATCCACCTTTCCC</td>
</tr>
<tr>
<td></td>
<td>TLR4_17050</td>
<td>CACAAATGCACACATC</td>
<td>CACAAATGCACACATC</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_149</td>
<td>CAGGCCCGCAGCAGCAG</td>
<td>CAGGCCCGCAGCAGCAG</td>
</tr>
<tr>
<td></td>
<td>HMOX1_1038</td>
<td>CTTATGCTGACAAGAAC</td>
<td>CTTATGCTGACAAGAAC</td>
</tr>
<tr>
<td></td>
<td>HMOX1_2790</td>
<td>ACCAGGCTATTGCTCT</td>
<td>ACCAGGCTATTGCTCT</td>
</tr>
<tr>
<td></td>
<td>HMOX1_3303</td>
<td>CAAACCGAAGCCGCAA</td>
<td>CAAACCGAAGCCGCAA</td>
</tr>
<tr>
<td></td>
<td>HMOX1_9531</td>
<td>TTACTGCTGTAATCCTC</td>
<td>TTACTGCTGTAATCCTC</td>
</tr>
<tr>
<td></td>
<td>HMOX1_16442</td>
<td>TCACCTTCGTATTCTCAA</td>
<td>TCACCTTCGTATTCTCAA</td>
</tr>
</tbody>
</table>

6. Real-time Polymerase Chain Reaction

Testing of the new real-time PCR assays created by Applied Biosystems was done utilizing the 7900HT Sequence Detection System (Applied Biosystems). Stock Coriell samples at a concentration of 100ng/μL were diluted to 1ng/μL using 1X TE buffer (Tris-EDTA buffer). TE is a commonly used buffer that helps to inactivate DNA nucleases: Tris adjusts the pH of DNA samples to around 8 since DNA nucleases are less active at this pH; EDTA chelates to metal cations which are
required by these enzymes. Coriell samples are DNA samples from individuals of both African American and European descent that have been employed by the IIPGA and SeattleSNPs groups for gene sequencing. These samples are available from a central repository (http://ccr.coriell.org/). Since the genotypes of those African and European individuals were known, their DNA samples were suitable for quality controls to test the accuracy of the ordered TaqMan assays. 5μL of each diluted Coriell sample was put into a 384-well clear optical reaction plate (Applied Biosystems) utilizing a Biomek FX Laboratory Automation Workstation (Beckham Coulter, Fullerton, CA). The 1ng/μL Coriell samples were prepared in a 0.8mL 96-well storage plate (ABgene House, Rochester, NY) and the plate was put onto the designated position on the workstation. The appropriate volume of solution was pipetted out by the automated syringes from the 96-well storage plate into the designated well of the 384-well plates. The plates were then placed in a fume hood for lyophylization of the DNA samples. In total, 24 sample samples of African American DNA and 23 samples of DNA from the Centre d'Etude du Polymorphisme Humain (CEPH) were tested for each assay. The CEPH samples are Utah residents with ancestry from Northern and Western Europe. In addition to the Coriell samples, wells of 1X TE buffer were included as negative controls.

175μL of 2X TaqMan Universal PCR master mix - No AmpErase UNG (Applied Biosystems), 8.75μL of 40X primer/probe solution (Applied Biosystems) and 166.25μL of water were mixed to form the master mix for real-time PCR with a total volume of 350μL. During amplification by TaqMan assay, the nucleotide uracil (U) was incorporated instead of thymine (T) into the DNA molecules. AmpErase is an enzyme (uracil-N-glycosylase) that can be used to cleave DNA molecules at sites containing a U, therefore it can be used for elimination of any carry-over contaminating PCR products from a previous PCR reaction. However, the additional expense of the AmpErase was not justified in the reactions here because no contamination was detected in the negative controls. For each assay, a 5μL aliquot of master mix was added into the appropriate wells in the 384-well plate and the plate was sealed with clear optical adhesive cover. The plate was placed into an Allegra 6 Centrifuge (Beckham Coulter) for centrifugation for about 5 minutes at a speed of about 2500rpm.
Real-time PCR was performed by the 7900HT Sequence Detection System (Applied Biosystems). In the linked computer, the program SDS2.1 was selected and the following settings were chosen: assay = absolute quantification (standard curve); container = 384-wells clear plate; template = blank template. The file containing the Coriell samples layout was imported into the program and an equivalent layout was shown on the screen. All wells that were assigned for testing an assay were highlighted and the appropriate testing marker and detector were chosen from the list. The blank and TE wells were then highlighted and their setting altered to "NTC" (non-template control). Under the instrument tab, the following settings were amended: sample volume = 5μL; the box indicating 9600 emulation was unchecked; temperature in stage 3 = 92°C and repeats = 50. The modified program was connected to the 7900HT Sequence Detection System and the door of the machine was opened. The 384-well plate was placed correctly according to the markings on the tray. The door was then closed and the reaction was initiated.

The real-time PCR was finished in about 2 hours. An allelic discrimination analysis was started once the real-time PCR was done. The allelic discrimination analysis was used to determine the presence of different alleles in the DNA samples by releasing a light source onto the plates and then detecting the quantity of light of different wavelengths being emitted by the reporter dye molecules. Since the genotypes of all Coriell samples were known, the reporter molecules corresponding to each of the alleles could be found. The SDS2.1 program was opened and the following settings were selected: assay = allelic discrimination; container = 384-well clear plates; template = blank template. The file of the layout of the Coriell samples was imported into the program. All wells that were assigned for testing an assay were highlighted and the appropriate testing marker (the name of the SNP) and detector (FAM/VIC) was chosen from the list. The blank and TE wells were then highlighted and their setting altered to "NTC". The program was connected to the 7900HT Sequence Detection System and the analysis was begun. After the analysis was completed, a graph of the fluorescence from allele Y (VIC) plotted against that from allele X (FAM) was generated. Each point on the graph was compared with the genotyping of the corresponding Coriell sample provided by either the
IIPGA or SeattleSNPs websites. The alleles corresponding to both FAM and VIC could be determined and the names of the both detectors were changed to the appropriate alleles.

7. Sequencing

Discrepancies were detected for the Factor B gene SNP 7202 which are indicated in table 7:

Table 7: Genotype discrepancies between the real-time PCR results in this study and those posted by the Seattle SNPs website (http://pga.mbt.washington.edu/)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Result by real-time PCR</th>
<th>Result posted by Seattle SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>D008</td>
<td>GG</td>
<td>AG</td>
</tr>
<tr>
<td>D016</td>
<td>AG</td>
<td>GG</td>
</tr>
<tr>
<td>E016</td>
<td>AG</td>
<td>GG</td>
</tr>
</tbody>
</table>

In order to determine the genotypes of the above three Coriell samples with regard to SNP 7202, sequencing was done at the University of British Columbia sequencing facility. PCR products of around 500 nucleotides were submitted to the UBC Oligonucleotide Synthesis Facility with the 7202 SNP approximately in the middle of each product. Both left and right primers were used for sequencing and they are listed in table 8.

Table 8: The sequence of both left and right primers used for sequencing of BF 7202 for Coriell samples D008, D016 and E016

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_7202L</td>
<td>TGG GTC CCT AGT CTG ATT TTA G</td>
</tr>
<tr>
<td>BF_7202R</td>
<td>TCC TGG AAG CAT GGC TGT TC</td>
</tr>
</tbody>
</table>

The D008, D016 and E016 Coriell samples were diluted from a concentration of 2ng/μL to 1ng/μL with 1X TE buffer. A solution with a total volume of 20μL was prepared for each Coriell sample by mixing 2μL of 10X PCR mix, 2μL dNTP (200 μM each of dGTP, dCTP, dTTP and dATP),
1μL of primer L (900μM), 1μL of primer R (900μM), 0.1μL of TAQHOTSTART (QIAGEN Inc., Mississauga, Ontario), 1μL of Coriell sample and 12.9μL of distilled water. The three solutions were placed in three separate 0.5ml Eppendorf tubes and labeled correctly. The PCR was performed utilizing PCR Express thermal cycler (Thermo Hybrid, Ashford, Middlesex, UK) by setting the cycling conditions to be 95°C for 10 minutes, 40 cycles at 92°C for 15 seconds and 55°C for 30 seconds and 72°C for 30 seconds, 72°C for 5 minutes. The PCR products were stored at -20°C.

Gel electrophoresis was performed to detect the presence of PCR product for each Coriell sample. 2% agarose in 0.5% TBE (Tris Borate EDTA) was prepared by mixing 2g of agarose powder into 100mL TBE buffer and the mixture of solute and solution was heated in a microwave until the agarose powder was completely dissolved. 10μL of ethidium bromide was added into the solution and mixed well. The gel was poured into the tray of a Horizon 20.25 GIBCO BRL Horizontal Gel Electrophoresis Apparatus (Life Technologies, St. Paul, MN) with the comb in place. 5μL of each PCR product and 2μL of sample buffer were mixed together, and the mixture was then loaded into separate wells after the gel solidified. 8μL of DNA ladder was loaded into another well. The gel electrophoresis apparatus was connected to a Power-PAC 300 (BIO-RAD, Hercules, CA) power supply. The gel was run for about 45 minutes under 150V. The presence and separation of bands in the gel was checked by a Spectroline hand held UV lamp (Model EF-140C, Spectronics Corporation, Westbury, NY). The gel was inspected carefully by an Eagle Eye II imager (Stratagene, La Jolla, CA) and a picture of the gel was taken.

The PCR products were then purified according to the QIAquick PCR Purification Kit Protocol as outlined on page 18 of the QIAquick Spin Handbook (QIAGEN Inc., Mississauga, Ontario). The purified PCR products and sequencing primers as listed in table 8 were sent to the Oligonucleotide Synthesis Facility of the University of British Columbia and used for sequencing.
8. Dilution of Samples

A total of 1605 DNA samples from 535 trios (two parents and the patient) at a concentration of 10ng/μL was received from our collaborators in the Hospital for Sick Children in Toronto. All samples were diluted to a final concentration of 1ng/μL with 1X TE buffer. The diluted samples were pipetted into 0.8mL 96-well storage plates (ABgene House), sealed and stored at -20°C.

9. Preparation for Polymerase Chain Reaction and Genotyping

5μL of diluted samples were plated from 96-well plates into 384-well clear optical reaction plates (Applied Biosystems) using a Biomek FX (Beckham Coulter). In addition to the diluted samples, 5μL of Coriell samples were included as the positive controls and TE buffer as the negative controls in each 384-well plate. All 384-well plates were dried in a fume hood overnight and then wrapped with aluminum foil and stored at 4°C.

10. Polymerase Chain Reaction

5000μL of 2X TaqMan Universal PCR master mix - NO AmpErase UNG (Applied Biosystems), 250μL of 40X primer/probe solution (Applied Biosystems) and 4750μL of water were mixed to form the assay for the PCR. A total of 10000μL mixture was sufficient for performing PCR for five 384-well plates. 5μL of this mixture was added into each well in the 384-well plates and they were sealed with clear optical adhesive cover. Then they were placed into an Allegra 6 Centrifuge (Beckham Coulter) for centrifuged for about 5 minutes at a speed of approximately 2500rpm. The plates were then put into a GeneAmp PCR System 9700 (Applied Biosystems) for polymerase chain reaction. The PCR cycles were set to be a 10-minute period at 95°C for denaturation of double stranded DNA and also the activation of the enzyme, 40 repeats of 15-second period at 92°C and 1-minute period at 60°C for annealing of primers and extension of the DNA strands, and another final 10-minute period at 72°C for extension of the DNA strands.
11. Genotyping

After the PCR, the plates were placed into a 7900HT Sequence Detection System (Applied Biosystems). In the linked computer, the SDS2.1 program was opened, and the settings were set as follows: assay = allelic discrimination; container = 384-well clear plates and template = blank template. The appropriate plate layout file was imported into the program and the testing detector and markers were chosen. The wells containing TE buffer were assigned as "NTC". The adjusted program was connected to the detection machine and the door was opened to place the plate in correct orientation as marked on the tray. The door was closed and the analysis was started. The genotyping was revealed in both table and diagram forms. Referring to the diagram, the three types of genotype (2 forms of homozygotes and 1 form of heterozygote) were designated by the auto-caller in the program and a sample diagram is shown in Figure 7. The results were saved and later used for analysis.

Figure 7: Sample genotyping result of Factor B SNP 2557. An allelic discrimination plot of BF_2557 A versus BF_2557 G is shown. Each of the points on the diagram represents a study sample, a Coriell sample or TE buffer. As can be seen in the figure, a cluster of several points was detected near the origin which symbolized wells with TE buffer since none of the two types of fluorescent light (from FAM or VIC) could be measured. In addition, three clusters of points were observed. Those points in the top left of the plot with a lower emission of light by the reporter molecule denoted BF_2557G and a higher emission of light by the reporter molecule denoted BF_2557A represent those individuals who had a homozygous AA genotype at SNP BF_2557. Those points in the middle with detection of emission by both the reporter molecules denoted BF_2557 A and BF_2557 G represent those individuals who were heterozygous AG at SNP BF_2557. Those points in the lower right of the plot with a lower emission of light by the reporter molecule denoted BF_2557A and a higher emission of light by the reporter molecule denoted BF_2557G represent those individuals who had a homozygous GG genotype at SNP BF_2557. There were two points between the clusters of AG and GG for which the genotyping could not be determined by the auto-caller.
12. Quantification of DNA Samples

Unsatisfactory results were obtained during the first round of genotyping samples from the first two 384-well plates for SNP 149 of heme oxygenase-1, as shown in Figure 8.

Figure 8: Example genotyping result for SNP 149 of HMOX1. An allelic discrimination plot of HMOX1_149 A versus HMOX1_149 G is shown. Each of the points on the diagram represents a study sample, Coriell sample or
TE buffer. As can be seen in the figure, no distinct clusters could be observed to classify the individuals into one of the three types of genotype since most of the points were gathered at the lower left hand corner of the plot. This might indicate that most of the DNA samples were not be amplified during PCR and therefore the probes were still intact. Therefore all light absorbed by the reporter molecule was transmitted to the quencher molecule since they were still located in close proximity to each other. Therefore, no emission of fluorescent light was detected.

Genotyping reactions were repeated for those samples in the first two 384-well plates, however, unsatisfactory results were again obtained. One explanation for these results was that the concentration of the samples was lower than expected. Therefore, the PicoGreen assay (Invitrogen, Burlington, Ontario) was used to ensure the concentration
of the original DNA samples from Toronto was 10ng/µL. The PicoGreen dsDNA Quantitation Reagent is an ultra-sensitive fluorescent nucleic acid stain for quantification of double-stranded DNA. The advantage of the PicoGreen reagent over conventional absorbance measurement at 260nm is that nucleotide, single-stranded nucleic acids and proteins all contribute to absorbance at 260nm. These are commonly found contaminants in DNA preparations. The PicoGreen reagent is more selective for double-stranded DNA and does not fluorescence in the presence of protein in addition to being more sensitive. The amount of fluorescence detected is directly proportional to the concentration of DNA, and therefore the concentration unknown DNA sample can be deduced once a standard curve is achieved.

A standard DNA sample (λ DNA) with a concentration of 100ng/µL was diluted to 2ng/µL with 1X TE buffer. 50µL of diluted λ DNA was put into well A1 of a black polystyrene 96-well plate (Corning Incorporation, Corning, NY). 25µL of the solution in well A1 was pipetted into well B1 and mixed with another 25µL of 1X TE buffer. This serial dilution was repeated for the entire column, except the last well. In well H1, only 25µL of 1X TE buffer was added with no standard DNA sample.

Each DNA sample from the source plates sent by Toronto was diluted to three different concentrations (1/20, 1/40 and 1/80). For each concentration of source DNA samples, a serial dilution into other columns of the black 96-well plates was done as outlined above for the standard DNA samples. The layout of the black 96-well plates is shown in Table 9.
Table 9: Layout of a 96-well plate with standard DNA samples and DNA samples from source plates sent by Toronto

<table>
<thead>
<tr>
<th>Plate Layout</th>
<th>Lambda DNA</th>
<th>1/20</th>
<th>1/20</th>
<th>1/20</th>
<th>1/40</th>
<th>1/40</th>
<th>1/40</th>
<th>1/80</th>
<th>1/80</th>
<th>1/80</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0003-01</td>
<td>0003-02</td>
<td>0003-03</td>
<td>0003-01</td>
<td>0003-02</td>
<td>0003-03</td>
<td>0003-01</td>
<td>0003-02</td>
<td>0003-03</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0019-01</td>
<td>0019-02</td>
<td>0019-03</td>
<td>0019-01</td>
<td>0019-02</td>
<td>0019-03</td>
<td>0019-01</td>
<td>0019-02</td>
<td>0019-03</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0052-01</td>
<td>0052-02</td>
<td>0052-03</td>
<td>0052-01</td>
<td>0052-02</td>
<td>0052-03</td>
<td>0052-01</td>
<td>0052-02</td>
<td>0052-03</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>0063-01</td>
<td>0063-02</td>
<td>0063-03</td>
<td>0063-01</td>
<td>0063-02</td>
<td>0063-03</td>
<td>0063-01</td>
<td>0063-02</td>
<td>0063-03</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>0069-01</td>
<td>0069-02</td>
<td>0069-03</td>
<td>0069-01</td>
<td>0069-02</td>
<td>0069-03</td>
<td>0069-01</td>
<td>0069-02</td>
<td>0069-03</td>
</tr>
<tr>
<td>0.125</td>
<td>0.125</td>
<td>0083-01</td>
<td>0083-02</td>
<td>0083-03</td>
<td>0083-01</td>
<td>0083-02</td>
<td>0083-03</td>
<td>0083-01</td>
<td>0083-02</td>
<td>0083-03</td>
</tr>
<tr>
<td>0.063</td>
<td>0.063</td>
<td>0101-01</td>
<td>0101-02</td>
<td>0101-03</td>
<td>0101-01</td>
<td>0101-02</td>
<td>0101-03</td>
<td>0101-01</td>
<td>0101-02</td>
<td>0101-03</td>
</tr>
<tr>
<td>0.031</td>
<td>0.031</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
<td>0150-01</td>
<td>0150-02</td>
<td>0150-03</td>
</tr>
</tbody>
</table>

200X PicoGreen reagent was diluted to 1X with TE buffer. 25μL of 1X PicoGreen reagent was added to each well in the black 96-well plate. The plate was covered with aluminum foil since the PicoGreen reagent is sensitive to light.

The black 96-well plate was placed into a GENios fluorimeter (Tecan Group Ltd., Durham, NC) for the detection of fluorescence emitted by the solution of PicoGreen reagent. In the linked computer, the Tecan-XFFluor4 program was opened. Under the edit measurement parameter tab, the following settings were chosen: general = fluorescence; plate = GRE96fb,pdf; excitation λ = 485nm; emission λ = 535nm, gain manual = 60, integration time = 40μs and the box indicating fluorescein was checked. The measurement was started and the results were saved.

13. Re-genotyping

As a quality control measure for the genotyping results obtained from TaqMan assays, re-genotyping was performed for about 30% of the samples. In this project, the concentration of the DNA samples in source plates #1-8 was found to be lower than the expected value of 0.125.
source plates #1-8 was found to be lower than the expected value of 10ng/µL and unsatisfactory results were obtained from the TaqMan assays. While waiting for a new batch of samples to be sent from Toronto, genotyping of plates #9-20 was performed and the data were used for quality control. Specifically, the data from the first round of genotyping of plates #9-20 were compared with the data from the new DNA samples.

14. Statistical Data Analysis

Before performing any statistical analysis of the genotyping data, the pattern of Mendelian inheritance was first examined for all participating trio members by the Family Based Association Test (FBAT) program (www.biostat.harvard.edu/~fbat/default.html). The presence of Mendelian errors in any one of the families for any one of the SNPs being tested would identify potential problems with the samples and/or genotyping assays, which would lead to misinterpretation of the data.

After the families with Mendelian inconsistencies were deleted from the rest of the analysis, chi square tests were used to determine the presence of Hardy-Weinberg Equilibrium among the parents and to compare the genotype frequencies of all SNPs between the parental population and the online databases. These two procedures helped to identify any inaccurate genotyping assays and/or whether the study population was genetically heterogeneous.

Analysis of Variance (ANOVA) was employed in attempting to establish any relationship between the genotypes and the severity of the disease. Three phenotypic characteristics (age of diagnosis, Forced Expiratory Volume in one second (FEV1) percent predicted and FEV1 standard deviation (see below) were selected as measurements which indicated the disease status of each patient.

FEV1 is defined as the volume exhaled during the first second of a forced expiratory maneuver. FEV1 values were expressed as a percentage of the expected values for age, sex and height calculated using the formula of Knudson et al [38] for those over the age of 10 years and according to the Hospital for Sick Children formula [39] for
those 6 to 10 years of age. In obstructive lung diseases such as CF, a
decrease in FEV1 will be evident and therefore it is a common index for
assessing pulmonary function. Consequently it was chosen as one of the
phenotypic traits in this study in order to measure the severity of the
disease.

As an alternate measure of lung function we used the difference
between a patient’s lung function and the average expected for a cohort
of Canadian CF patients. A regression analysis of age versus FEV1
percent predicted was conducted using the longitudinal data of patients
homozygous for the ΔF508 mutation from the Canadian CF Patient Data
Registry (2002) (see Figure 9). The FEV1 standard deviation score
(FEV1_SD) was calculated by comparing each patient’s FEV1 percent
predicted value with that from the regression curve for that age and
gender. A positive value for FEV1_SD indicates the number of standard
deviations that the individual’s FEV percent predicted is above that
predicted from this relationship and a negative value is the number of
standard deviations the subject’s value is below that predicted for
that age and gender.
Figure 9: Normalization of the FEV1 percent predicted values by age. Longitudinal data from the Canadian CF Patient Data Registry was used to calculate the linear regression curve for patients homozygous for the deltaF508 mutation. The standard deviation scores were calculated for all patients according to the FEV1 percent predicted values and age. There is a deficiency of patients in age groups over 30 with low
scores due to mortality.

By performing the ANOVA test using the program JMP v5.1 (SAS Institute Inc. Cary, NC), the sample should be normally distributed and the patients were unrelated (i.e. were not from the same extended pedigree and thus the observations would be independent). The normality of a sample can be checked by the goodness of fit test in the JMP program. A normal quantile plot was done by plotting the phenotypic variables (i.e. age of diagnosis, FEV1 predicted and standard deviation values) as the y-axis according to the SNP genotype. Although t-tests could be used for identifying significant differences between two genotypic groups, ANOVA was superior over t-test in this analysis since multiple t-tests would be required for the necessary comparisons. In contrast, by using ANOVA we could compare more than two genotypic groups and thus generate a single P value.

Before we concluded that a significant association had been established between one of the three traits and the genotype of any SNP of the four candidate genes, regression analysis was performed to adjust the data for potentially confounding factors such as sex, age and CFTR genotype.

One of the traits, the age of first Pseudomonas aeruginosa infection, was examined by survival analysis or age of onset analysis. Although about 80% of CF patients were infected with Pseudomonas aeruginosa, not every participant recruited in this study had been colonized with this bacterium. However, it is likely that these individuals will become colonized by Pseudomonas aeruginosa in the future. Consequently it was not appropriate to ignore these patients in this part of the study as would be the case if the analysis was done by ANOVA. When performing the age of onset analysis using JMP, these individuals were assigned their current age as their age of first Pseudomonas aeruginosa infection. However, since the age of first Pseudomonas aeruginosa infection of these patients was incorrect, a censor column was added which stated 0=no Pseudomonas aeruginosa infection and 1=Pseudomonas aeruginosa infection. Therefore these individuals would be accounted for in the calculation by the program automatically according to the value in the censor column.
As mentioned before, *Pseudomonas aeruginosa* infection was not found in all patients recruited in this study. One of four categories was assigned to each CF patient with respect to their *Pseudomonas aeruginosa* infection status, namely no *Pseudomonas aeruginosa* infection (code 0), grew *Pseudomonas aeruginosa* once (code 1), sporadic *Pseudomonas aeruginosa* growth (code 2) and chronic *Pseudomonas aeruginosa* infection (code 3). Susceptibility to *Pseudomonas aeruginosa* infection was hypothesized to be influenced by one or more the SNPs of the four candidate modifier genes, and the Chi-test was used to confirm the existence of such a relationship.

Another analytical method, the Family Based Association Test (FBAT) [40], was employed for examining the presence of association between the selected phenotypes and SNPs. The FBAT software implements the TDT and other family-based tests of association. Although FBAT is also a test of genetic association, this test is markedly different from the ANOVA. FBAT tests for association within the family trio members while ANOVA only considered the patients in the analysis. Both phenotype and genotype files needed to be prepared in the correct format for analysis by the FBAT program. Both files were prepared in the Notepad program, with the genotype file saved as a .ped file whereas the phenotype file was saved as a .phe file. In the genotype file, the names of all 22 SNPs were placed in the first row. There were also restrictions for the columns. The first column was the pedigree ID and the second column was the individual ID. Father ID and mother ID were shown in the third and fourth columns, respectively. Gender was indicated in the fifth column with 1=male and 2=female. The next column represented the affection status. The remaining columns showed the alleles at each of the SNPs.

Haplotype analysis was the last part of the analysis in this study. Possible combinations of all the SNPs in each candidate gene were tested as a group. This was performed to determine whether there was any interaction effect between SNPs, which then may have resulted in a new association with the outcome variable that had not been detected when each SNP was analyzed individually. In addition, examination of haplotypes may have identified causal SNPs that were not
detected in the resequencing programs such as Seattle SNPs. These causal SNPs could be outside the region that was resequenced but found preferentially on a specific haplotype. Thus, the haplotype would act as a marker for the causal SNP. However, some of the haplotypes were found to have a low frequency and those were pooled together as a group. Also, the confounding factors, e.g. sex and CFTR genotype, were put into the analysis for adjustment of the data.

The haplotype analysis was first done by the RGui program version 2.5.0, provided by The R Foundation for Statistical Computing. Both genotyping and phenotypic input files were required to be in correct format recognized by the program and saved as a .csv file in Notepad. In the genotypic file, the first row contained the heading of each column: family ID, individual ID, father ID, mother ID and genotype of each allele of every SNP. Then the data of each participant was arranged in a row below according to the headings listed. In the phenotypic file, the first row showed the name of each column: person ID and each of the phenotypic characteristics being collected (including the target phenotypic traits and confounding factors in this study: age of diagnosis, FEV1 predicted and standard deviation values, gender, age and CFTR mutation). The patients' data were put below in the order of the headings named in the first row. The phenotypic trait of age of diagnosis was also logarithmically transformed in order to correct for normality. All SNPs within a single gene were grouped together as a haplotype for determination of any association and analysis was done twice, one with and the other without adjustment for confounding factors such as sex and CFTR genotype. Results from both analyses could then be evaluated in the presence and absence of those confounding factors.

Haplotype analysis between the genotype and phenotype (age of diagnosis, FEV1 predicted and standard deviation values) was also performed by FBAT program. The generation of possible haplotypes was determined according to implementation of the TDT and other family-based tests of association offered by the program. The format of both genotypic and phenotypic input files were exactly the same as those used in previous FBAT calculations. Only the additive model was included in this part of the study. Due to the different computational
basis of both programs, direct comparison between results obtained from the RGui and FBAT programs could not be done. However if consistent results were obtained by both programs this would suggest stronger association.

For the age of first *Pseudomonas aeruginosa* infection, the haplotype analysis was done by the program Hapstat ([http://www.bios.unc.edu/~lin/hapstat/](http://www.bios.unc.edu/~lin/hapstat/)) [41]. The input file was saved as .txt file in Notepad. The first row represented the headings of data listed below: age of first *Pseudomonas aeruginosa* infection, status of *Pseudomonas aeruginosa* infection, gender, CFTR genotype and the genotype of each of the involved SNPs. As the file was loaded as a cohort file into the program the following settings were utilized: observation time = age of first *Pseudomonas aeruginosa* infection; event indicator = status of *Pseudomonas aeruginosa* infection; genotype = all the SNPs in each of the candidate modifier genes; environment = confounding factors (gender and CFTR genotype: deltaF508/deltaF508, deltaF508/other mutation; and other mutation/other mutation). Only the additive model was used in this part of the study. Both boxes of haplotype frequency and effects had to be checked, and the haplotype frequency was calculated based on the number of samples and was set to default as 0.009; while the haplotype effect needed to be changed to 0.05 as the P value.
Chapter 3 Results

1. PicoGreen Reaction

In the initial experiments for this project, unsatisfactory results were obtained from the genotyping of HMOX1 SNP149 for the original source plates #1-8. Therefore, it was suspected that the initial concentration in each of the wells was less than the expected 10ng/μL. Consequently, there may not have been enough DNA present for PCR and amplification could not be done to the expected level for genotyping. Therefore, PicoGreen reactions were performed in order to determine the concentration of DNA from source plates #1-8. A standard curve was plotted with serially diluted λ DNA samples as indicated in the table layout in the "Materials and Methods" section. The standard curve derived from this serial dilution is shown in Figure 10.

Figure 10: Standard curve of optical density (OD) versus DNA concentration of the serially diluted λ DNA samples

With the aid of the above standard curve, the concentration of DNA in the original source plates #1-8 was determined and a representative subset of the results is summarized in Table 1 in the Appendix.

As observed from the Table A1 in the Appendix, the concentration of samples provided by Toronto ranged from -0.03ng/μL to 9.77ng/μL and
most of the concentrations were lower than expected. This was presumably the major reason for the failed genotyping assays. However, this inconsistent concentration of DNA samples was not detected in source plates #9-20. DNA samples in those plates were genotyped for most of the SNPs while waiting for a new batch of DNA samples of all the participants to be sent from Toronto. These results from plates #9-20 would be used as a quality control measure to compare with genotypes from the new DNA samples. (see section 12-Re-genotyping)

2. Analysis of the Genotypic Data for Mendelian Inconsistencies

Analysis of the parent-offspring trios for Mendelian inconsistencies was performed using the Family Based Association Test (FBAT) program [42]. This program tests the inheritance pattern from both parents to their child. Each of the parents should contribute one of their two alleles to the child in accordance with the theory of heredity stated by the father of genetics, Gregor Mendel. The presence of Mendelian errors in any one of the families for any one of the SNPs being tested will identify potential problems with the samples and/or genotyping assays. Mendelian errors were detected in some of the families when testing for particular SNPs and the results are summarized in Table A2 in the Appendix.

There were many potential reasons leading to the observed Mendelian errors, which included contamination of samples by other DNA molecules during sample collection or genotyping, incorrect labeling of the identities of the individuals, non-paternity and random genotyping errors. In this study, those families with more than one error in the Mendelian inheritance patterns were deleted from the study while those with only one error were excluded for the analysis of the particular SNP for which the error was detected.

3. Sequencing Results

Due to discrepancies between genotypes of three of the control individuals when testing the assay for BF_7202, DNA samples from Coriell controls D0008, D016 and E016 were sent to the University of British Columbia for sequencing, together with the sequence of the
primers used. However, these sequencing reactions were not successful for unknown technical reasons. However, the genotypes of these three individuals were repeated two more times by the TaqMan assay designed for BF_7202, and they were found to match those on the SeattleSNPs website. Therefore, the genotyping result was confirmed as reliable regardless of the failure of the sequencing by UBC and the analyses of the BF_7202 SNP were continued.

4. Genotypes of the Participating Individuals

This was a sub-study of a large, Canada-wide and international project: the Canadian Consortium for CF Modifiers. DNA samples from 535 families with a total of 1605 subjects (2 parents and 1 patient) across Canada were included in this sub-study. A DNA sample of each participant was sent from Toronto at a concentration of 10ng/µL. Four candidate genes with a total of 22 SNPs were chosen and all samples were genotyped for those SNPs by TaqMan assays according to the procedures detailed in the "Materials and Methods" section.

Five SNPs were selected for Factor B gene. Genotyping results for this gene are summarized in Table 3 in the Appendix. As noted in the table, non call rates ("undetermined" in Table A3 in the Appendix) for all SNPs were about 2% or less. One of the possible reasons for genotyping failure was a low concentration of DNA (less than the expected concentration of 1ng/µL after dilution) and therefore not enough DNA molecules were amplified for those individuals and their genotypes were undetectable. Two families were deleted from the analysis of SNP 2557 since errors were detected in these two families when investigating the pattern of inheritance of the alleles from parents to offspring utilizing the FBAT program.

Four SNPs were selected for genotyping in the Complement factor 3 gene (Table A4 in the Appendix). Genotyping was not successful for less than 2% of the samples for each SNP. Furthermore, three and one families were deleted from the analysis of SNP 963 and 28795, respectively because of the non-Mendelian inheritance pattern observed at the two SNPs in those families.
Seven SNPs were selected for analysis in the toll-like receptor 4 gene (Table A5 in the Appendix). In the TLR4 gene, only one family was detected to have an error in the inheritance pattern (for SNP 11912) and it was removed from all analyses regarding this SNP. The genotypes of less than 2% of the samples were found to be undetermined.

Six SNPs were included for the analysis of the heme oxygenase-1 gene (Table A6 in the Appendix). For SNP 149, 17 families were excluded due to Mendelian errors and a high number of individuals were detected to have an undetermined genotype. With the exception of SNP 149, only about 1% of genotypes could not be determined. Furthermore, only 2 families were deleted from the analysis due to errors in the inheritance pattern (one each for SNP 2790 and 9531).

5. Phenotypic Characteristics of the Study Subjects

The study consisted of a total of 535 families recruited at CF clinics across Canada. Unfortunately, we were not able to collect characteristics of every single individual enrolled in the study. The available variables are summarized in Table A7 in the Appendix.


Hardy-Weinberg analysis determines how the allele frequency of a given SNP corresponds to the genotype distribution. The Hardy-Weinberg Law states that the frequency of an allele should remain constant over time unless there are outside driving forces acting on the population. A Hardy-Weinberg analysis was performed on the parent population since any deviation from equilibrium could indicate an inaccurate genotyping assay or a genetically heterogeneous study population. A summary table of the investigation of Hardy-Weinberg equilibrium among the parent population is shown in Table 10.
Table 10: Hardy-Weinberg Equilibrium among the parental population

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th>Allele Frequency</th>
<th>Genotype</th>
<th>Observed Frequency</th>
<th>Expected Frequency</th>
<th>X² Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2557</td>
<td>A</td>
<td>0.18</td>
<td>AA</td>
<td>36</td>
<td>34</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.82</td>
<td>AG, GG</td>
<td>314, 707</td>
<td>312, 711</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1057</td>
<td>1057</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4022</td>
<td>A</td>
<td>0.32</td>
<td>AA</td>
<td>107</td>
<td>108</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.68</td>
<td>AG, GG</td>
<td>458, 489</td>
<td>459, 487</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1054</td>
<td>1054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6484</td>
<td>A</td>
<td>0.89</td>
<td>AA</td>
<td>858</td>
<td>844</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.11</td>
<td>AG, GG</td>
<td>189, 19</td>
<td>209, 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1066</td>
<td>1066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7202</td>
<td>A</td>
<td>0.43</td>
<td>AA</td>
<td>197</td>
<td>194</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.57</td>
<td>AG, GG</td>
<td>508, 344</td>
<td>514, 341</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1049</td>
<td>1049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8311</td>
<td>C</td>
<td>0.91</td>
<td>CC, CT, TT</td>
<td>862, 182</td>
<td>870, 172</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0.09</td>
<td>CG, TT</td>
<td>7</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1051</td>
<td>1051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement factor 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28795</td>
<td>G</td>
<td>0.42</td>
<td>AA</td>
<td>355</td>
<td>354</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AG, GG</td>
<td>521, 177</td>
<td>513, 186</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1053</td>
<td>1053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36735</td>
<td>A</td>
<td>0.44</td>
<td>AA</td>
<td>202</td>
<td>203</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.56</td>
<td>AG, GG</td>
<td>519, 330</td>
<td>518, 330</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1051</td>
<td>1051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43118</td>
<td>A</td>
<td>0.48</td>
<td>AA</td>
<td>239</td>
<td>244</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>0.52</td>
<td>AG, GG</td>
<td>529, 292</td>
<td>529, 287</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total</td>
<td>1060</td>
<td>1060</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 10: Hardy-Weinberg Equilibrium among the parental population

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>851</td>
<td>0.72</td>
<td>0.28</td>
<td>1.00</td>
</tr>
<tr>
<td>1859</td>
<td>0.37</td>
<td>0.63</td>
<td>1.00</td>
</tr>
<tr>
<td>2856</td>
<td>0.15</td>
<td>0.85</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**TLR 4**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>9263</td>
<td>0.11</td>
<td>0.89</td>
<td>1.00</td>
</tr>
<tr>
<td>11912</td>
<td>0.33</td>
<td>0.67</td>
<td>1.00</td>
</tr>
<tr>
<td>15884</td>
<td>0.15</td>
<td>0.85</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Heme oxygenase-1**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>149</td>
<td>0.32</td>
<td>0.68</td>
<td>1.00</td>
</tr>
<tr>
<td>1038</td>
<td>0.96</td>
<td>0.04</td>
<td>1.00</td>
</tr>
<tr>
<td>2790</td>
<td>0.57</td>
<td>0.43</td>
<td>1.00</td>
</tr>
<tr>
<td>3303</td>
<td>0.95</td>
<td>0.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.46</td>
<td>0.54</td>
<td>0.95</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95</td>
<td>0.05</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Summary:
- **TLR 4**:
  - A: 0.72, G: 0.28, Total: 1.00
  - A: 0.37, G: 0.63, Total: 1.00
  - A: 0.15, G: 0.85, Total: 1.00

- **Heme oxygenase-1**:
  - A: 0.32, G: 0.68, Total: 1.00
  - A: 0.96, G: 0.04, Total: 1.00
  - A: 0.57, G: 0.43, Total: 1.00
  - A: 0.95, G: 0.05, Total: 1.00

**Herne oxygenase-1**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>G</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.32</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>0.96</td>
<td>0.04</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>0.57</td>
<td>0.43</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>0.05</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

**Summary**:
- A: 0.95, G: 0.05, Total: 1.00
- A: 0.96, G: 0.04, Total: 1.00
- A: 0.57, G: 0.43, Total: 1.00
- A: 0.95, G: 0.05, Total: 1.00
Hardy-Weinberg Equilibrium was established for all the selected SNPs among the parental population, except SNP HMOX1_149.

7. **Comparison of Genotype Frequencies in the Parent Population and in Online Databases**

As another check of the genotyping assays the allele frequencies obtained in this study were compared with those recorded for Caucasian individuals in the Innate Immunity Programs for Genomic Applications (IIPGA) and UW-FHCRC Variation Discovery Resource (Seattle SNPs) website (Table 11). Wide disparities between these two datasets could indicate an inaccurate genotyping assay in this study.
Table 11: Comparison of allele frequencies between the genotyping results and the reported values on either the IIPGA or Seattle SNPs websites

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th>Frequency of the allele in this study</th>
<th>Reported frequency of the allele</th>
<th>X² Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td>2557</td>
<td>A</td>
<td>0.18</td>
<td>0.15</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.82</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4022</td>
<td>A</td>
<td>0.32</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.68</td>
<td>0.71</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>6484</td>
<td>A</td>
<td>0.89</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.11</td>
<td>0.11</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>7202</td>
<td>A</td>
<td>0.43</td>
<td>0.35</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.57</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8311</td>
<td>C</td>
<td>0.91</td>
<td>0.89</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.09</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>963</td>
<td>G</td>
<td>0.51</td>
<td>0.55</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.49</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>28795</td>
<td>A</td>
<td>0.58</td>
<td>0.5</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.42</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36735</td>
<td>A</td>
<td>0.44</td>
<td>0.45</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.56</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>43118</td>
<td>A</td>
<td>0.48</td>
<td>0.47</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.52</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>TLR-4</td>
<td>851</td>
<td>A</td>
<td>0.72</td>
<td>0.72</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.28</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1859</td>
<td>A</td>
<td>0.37</td>
<td>0.39</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.63</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2856</td>
<td>C</td>
<td>0.15</td>
<td>0.13</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.85</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9263</td>
<td>A</td>
<td>0.11</td>
<td>0.13</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C</td>
<td>0.89</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11912</td>
<td>G</td>
<td>0.67</td>
<td>0.70</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.33</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15884</td>
<td>C</td>
<td>0.15</td>
<td>0.11</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.85</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>17050</td>
<td>C</td>
<td>0.15</td>
<td>0.17</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.85</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>HMOX1</td>
<td>149</td>
<td>A</td>
<td>0.32</td>
<td>0.28</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.68</td>
<td>0.72</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>1038</td>
<td>C</td>
<td>0.96</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.04</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2790</td>
<td>A</td>
<td>0.57</td>
<td>0.61</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.43</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3303</td>
<td>C</td>
<td>0.05</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.95</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9531</td>
<td>A</td>
<td>0.46</td>
<td>0.48</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G</td>
<td>0.54</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>16442</td>
<td>A</td>
<td>0.95</td>
<td>0.87</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T</td>
<td>0.05</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>
Inspection of Table 11 shows that the allele frequencies from this study and the reported allele frequencies were generally similar. However, statistical analysis by 2x2 Chi-square test demonstrated significant differences for the heme oxygenase-1 SNPs 1038 and 16442. There were discrepancies of approximately 7% in the allele frequencies for each of those two SNPs. This may indicate that the assays for these SNPs are unreliable. However, the fact that the discrepancies are small and the SNPs are in Hardy-Weinberg equilibrium argues against this.

Samples with different genetic backgrounds would be another possible reason for the discrepancy. The reported heme oxygenase-1 data shown in Table 11 are from the Seattle SNPs website. All the DNA samples used in this study and those used by Seattle SNPs were from the Caucasian population. However, the Caucasian individuals genotyped by Seattle SNPs were Utah residents with ancestry from Northern and Western Europe while those in this study may be representative of a different subset of the Caucasian samples. Another possible reason for the discrepancy could be the small sample size used by Seattle SNPs, i.e. only 23 European samples were sequenced.

8. ANOVA Analysis of the Influence of Genotype on the Phenotypes

Three phenotypic characteristics: age of diagnosis, FEV1 predicted value and FEV1 standard deviation were included in the one-way ANOVA to determine if any of the genotypes contributed to the variation in the selected traits. Each trait provides a measure of the degree of severity of disease in the patients. The use of ANOVA assumes that the dependent variable is normally distributed, therefore, various transformations were performed (including quadratic, cubic, common logarithm and natural logarithm) in order to normalize the distributions if needed. In this study, only the data set of age of diagnosis was slightly skewed and it was then transformed by common logarithm before ANOVA was performed.

(a) Factor B

(I) Age of diagnosis
Table 12: ANOVA of age of diagnosis among different genotypes of the selected SNPs in Factor B. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF 2557</td>
<td>AA</td>
<td>18</td>
<td>-0.27</td>
<td>0.20</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>111</td>
<td>-0.16</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>307</td>
<td>-0.23</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>BF 4022</td>
<td>AA</td>
<td>55</td>
<td>-0.20</td>
<td>0.12</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>167</td>
<td>-0.30</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>217</td>
<td>-0.15</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>BF 6484</td>
<td>AA</td>
<td>353</td>
<td>-0.19</td>
<td>0.05</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>83</td>
<td>-0.31</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>7</td>
<td>-0.13</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>BF 7202</td>
<td>AA</td>
<td>85</td>
<td>-0.02</td>
<td>0.09</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>210</td>
<td>-0.34</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>146</td>
<td>-0.13</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>BF 8311</td>
<td>CC</td>
<td>360</td>
<td>-0.25</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>73</td>
<td>-0.04</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>4</td>
<td>0.11</td>
<td>0.43</td>
<td></td>
</tr>
</tbody>
</table>

(II) FEV1 predicted value

Table 13: ANOVA of FEV1 predicted value among different genotypes of the selected SNPs in Factor B

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF 2557</td>
<td>AA</td>
<td>14</td>
<td>78.31</td>
<td>6.98</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>104</td>
<td>68.90</td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>279</td>
<td>77.86</td>
<td>1.56</td>
<td></td>
</tr>
<tr>
<td>BF 4022</td>
<td>AA</td>
<td>48</td>
<td>69.89</td>
<td>3.79</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>150</td>
<td>74.86</td>
<td>2.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>205</td>
<td>77.72</td>
<td>1.83</td>
<td></td>
</tr>
<tr>
<td>BF 6484</td>
<td>AA</td>
<td>329</td>
<td>75.22</td>
<td>1.45</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>70</td>
<td>79.19</td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>7</td>
<td>68.97</td>
<td>9.92</td>
<td></td>
</tr>
<tr>
<td>BF 7202</td>
<td>AA</td>
<td>82</td>
<td>69.43</td>
<td>2.89</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>185</td>
<td>76.30</td>
<td>1.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>137</td>
<td>78.47</td>
<td>2.24</td>
<td></td>
</tr>
<tr>
<td>BF 8311</td>
<td>CC</td>
<td>321</td>
<td>75.77</td>
<td>1.47</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>72</td>
<td>74.58</td>
<td>3.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>5</td>
<td>72.67</td>
<td>11.79</td>
<td></td>
</tr>
</tbody>
</table>
(III) FEV1 standard deviation value

Table 14: ANOVA of FEV1 sd value among different genotypes of the selected SNPs in Factor B

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF 2557</td>
<td>AA</td>
<td>14</td>
<td>0.70</td>
<td>0.25</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>104</td>
<td>0.23</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>279</td>
<td>0.53</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>BF 4022</td>
<td>AA</td>
<td>48</td>
<td>0.19</td>
<td>0.14</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>150</td>
<td>0.46</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>205</td>
<td>0.53</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>BF 6484</td>
<td>AA</td>
<td>329</td>
<td>0.45</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>70</td>
<td>0.57</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>7</td>
<td>-0.18</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>BF 7202</td>
<td>AA</td>
<td>82</td>
<td>0.23</td>
<td>0.10</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>185</td>
<td>0.48</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>137</td>
<td>0.58</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>BF 8311</td>
<td>CC</td>
<td>321</td>
<td>0.46</td>
<td>0.05</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>72</td>
<td>0.47</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>5</td>
<td>0.08</td>
<td>0.42</td>
<td></td>
</tr>
</tbody>
</table>

From the above tables, the age of diagnosis was log-transformed in order to achieve normally distributed set of data. Three of the five SNPs (BF_4022, BF_6484 and BF_8311) had P values of greater than 0.05 when performing ANOVA for all the phenotypes. Therefore, it was unlikely that there was a relationship between these SNPs and CF disease severity. However, for the two SNPs (BF_2557 and BF_7202) with P values less than 0.05 the analyses were repeated with adjustment for confounding factors in order to confirm a significant association between the tested SNPs and phenotypic traits.

(b) Complement Factor 3

ANOVA tests were completed for investigating the selected three phenotypic traits among the SNPs of the candidate CF modifier gene of Complement factor 3 and the results are shown in Tables A8, A9 and A10 in the Appendix. No adjustment with respect to those confounding factors was required for any of the analyses in this part. None of the analyses revealed a relationship with the measured factors since all of the P values were greater than 0.05.
(c) Toll-like Receptor 4

(I) Age of diagnosis

Table 15: ANOVA of age of diagnosis among different genotypes of the selected SNPs in Toll-like receptor 4. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR4_851</td>
<td>AA</td>
<td>242</td>
<td>-0.19</td>
<td>0.06</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>168</td>
<td>-0.25</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>31</td>
<td>-0.08</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>AA</td>
<td>62</td>
<td>-0.29</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>206</td>
<td>-0.10</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>175</td>
<td>-0.31</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>CC</td>
<td>9</td>
<td>0.03</td>
<td>0.29</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>121</td>
<td>-0.32</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>314</td>
<td>-0.17</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>AA</td>
<td>7</td>
<td>0.19</td>
<td>0.33</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>86</td>
<td>-0.21</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>351</td>
<td>-0.22</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>GG</td>
<td>190</td>
<td>-0.19</td>
<td>0.06</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>191</td>
<td>-0.19</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>59</td>
<td>-0.34</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>CC</td>
<td>9</td>
<td>-0.25</td>
<td>0.29</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>102</td>
<td>-0.24</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>325</td>
<td>-0.21</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>CC</td>
<td>12</td>
<td>-0.15</td>
<td>0.25</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>109</td>
<td>-0.26</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>322</td>
<td>-0.20</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

ANOVA tests were completed for examining the selected three phenotypic traits among the SNPs of the candidate CF modifier gene of Toll-like receptor 4. As shown in the table above and Tables A11 and A12 in the Appendix, no adjustment for confounding factors was performed since all the P values were greater than 0.05, except for the ANOVA of SNP TLR4_1859 and the age of diagnosis. Consequently, regression analysis was performed for this association.

(d) Heme Oxygenase-1

(I) FEV1 predicted value
Table 16: The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in heme oxygenase-1

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_149</td>
<td>AA</td>
<td>49</td>
<td>71.14</td>
<td>3.81</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>159</td>
<td>74.08</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>169</td>
<td>77.64</td>
<td>2.05</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CC</td>
<td>367</td>
<td>75.16</td>
<td>1.37</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>37</td>
<td>81.07</td>
<td>4.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>2</td>
<td>75.30</td>
<td>18.55</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>AA</td>
<td>128</td>
<td>78.59</td>
<td>2.31</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>193</td>
<td>76.39</td>
<td>1.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>83</td>
<td>69.57</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CC</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>39</td>
<td>79.13</td>
<td>4.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>366</td>
<td>75.37</td>
<td>1.37</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>AA</td>
<td>87</td>
<td>79.82</td>
<td>2.80</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>188</td>
<td>77.55</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>126</td>
<td>70.43</td>
<td>2.33</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>AA</td>
<td>366</td>
<td>75.46</td>
<td>1.37</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>40</td>
<td>77.09</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>

(II) FEV1 standard deviation value

Table 17: The ANOVA result of examining FEV1 sd value among different genotypes of the selected SNPs in heme oxygenase-1

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_149</td>
<td>AA</td>
<td>49</td>
<td>0.27</td>
<td>0.14</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>159</td>
<td>0.44</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>169</td>
<td>0.48</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CC</td>
<td>367</td>
<td>0.45</td>
<td>0.05</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>37</td>
<td>0.59</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>2</td>
<td>0.20</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>AA</td>
<td>128</td>
<td>0.49</td>
<td>0.08</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>193</td>
<td>0.54</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>83</td>
<td>0.22</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CC</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>39</td>
<td>0.46</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>366</td>
<td>0.46</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>AA</td>
<td>87</td>
<td>0.55</td>
<td>0.10</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>188</td>
<td>0.59</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>126</td>
<td>0.24</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>AA</td>
<td>366</td>
<td>0.46</td>
<td>0.05</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>40</td>
<td>0.42</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>
From the two tables above and Table A13 in the Appendix, it is unlikely that any of the six SNPs were associated with the three measured phenotypic traits, except for HMOX1_2790 and HMOX1_9531. These SNP were detected to have a P-value of less than 0.05 when performing ANOVA with FEV1 predicted value and FEV1 sd value. Before a significant relationship could be concluded, adjustment for confounding factors was required.

9. Regression Analysis

As indicated by the ANOVA results, a significant association was found for the following SNPs: (1) BF 2557 and FEV1 predicted value/sd value; (2) BF7202 and age of diagnosis; (3) BF 7202 and FEV1 predicted value/sd value; (4) TLR4_1859 and age of diagnosis; (5) HMOX1_2790 and FEV1 predicted value/sd value; and (6) HMOX1_9531 and FEV1 predicted value. The ANOVA of those five SNPs with the specified phenotypes had a P value of less than 0.05 which indicated that there might be an association between the SNP and trait. However, before a final conclusion that a correlation could be established, the association was corrected for confounding factors which included sex, age and CFTR genotype. These factors were hypothesized to have effect on those measured traits, for example, different genotype at the CFTR mutation would lead to a wide variety of clinical symptoms. Regression analysis was performed with those confounding factors and the results are illustrated in the following tables 18-27.
Table 18: Regression analysis of the association of BF_2557 and FEV1 predicted value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557/FEV1 predicted value</td>
<td>BF 2557</td>
<td>AA</td>
<td>77.40</td>
<td>6.63</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>69.25</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>78.53</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>76.19</td>
<td>3.01</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>75.98</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>AF508/AF508</td>
<td>75.61</td>
<td>2.51</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>77.92</td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>72.05</td>
<td>4.68</td>
<td></td>
</tr>
</tbody>
</table>

Table 19: Regression analysis of the association of BF_2557 and the FEV1 sd value with confounding factor

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557/FEV1 sd value</td>
<td>BF 2557</td>
<td>AA</td>
<td>0.72</td>
<td>0.27</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.25</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>0.55</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>0.43</td>
<td>0.12</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>0.51</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>AF508/AF508</td>
<td>0.44</td>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>0.57</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>0.37</td>
<td>0.19</td>
<td></td>
</tr>
</tbody>
</table>

From the above two tables, it can be seen that there were associations between the Factor B SNP 2557 and both FEV1 predicted and standard deviation values. The associations were independent of the potential confounding factors of sex and CFTR genotype. However, a P value of less than 0.0001 was found on evaluation of the effect of age on the FEV1 predicted value. This observation could be explained by the fact that lung function measured by percent predicted would be decreased with increasing age.
### Table 20: Regression analysis of the association of BF_7202 and the age of diagnosis with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_7202/age of diagnosis</td>
<td>BF 7202</td>
<td>AA</td>
<td>2.75</td>
<td>0.63</td>
<td>0.0274</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>2.00</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>3.70</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>2.31</td>
<td>0.45</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>3.04</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>1.59</td>
<td>0.34</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>3.80</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>6.11</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

### Table 21: Regression analysis of the association of BF_7202 and the FEV1 predicted value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_7202/FEV1 predicted value</td>
<td>BF 7202</td>
<td>AA</td>
<td>70.12</td>
<td>2.92</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>76.73</td>
<td>2.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>79.73</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>76.75</td>
<td>2.07</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>76.07</td>
<td>2.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>75.91</td>
<td>1.59</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>78.31</td>
<td>2.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>72.05</td>
<td>4.16</td>
<td></td>
</tr>
</tbody>
</table>
Table 22: Regression analysis of the association of BF_7202 and the FEV1 standard deviation value

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_7202/FEV1 sd value</td>
<td>BF 7202</td>
<td>AA</td>
<td>0.26</td>
<td>0.12</td>
<td>0.0215</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.49</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>0.61</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>0.45</td>
<td>0.08</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>0.52</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>0.45</td>
<td>0.06</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>0.58</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>0.37</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>

A P value of 0.0274 was found, which confirmed the presence of a significant relationship for the regression analysis of BF_7202 and age of diagnosis. Also, the associations between both the FEV1 predicted and standard deviation values and BF_7202 were verified.

(c) TLR4_1859

Table 23: Regression analysis of the association of TLR4_1859 and the age of diagnosis with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR4/age of diagnosis</td>
<td>TLR4_1859</td>
<td>AA</td>
<td>2.17</td>
<td>0.73</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>2.95</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>2.59</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>2.31</td>
<td>0.45</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>3.05</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>1.59</td>
<td>0.36</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>3.80</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>6.11</td>
<td>0.90</td>
<td></td>
</tr>
</tbody>
</table>

The presence of an association between TLR4_1859 and age of diagnosis was not confirmed by this analysis, as there was a p value for TLR4_1859 of 0.68 after adjustment for confounding factors.
Table 24: Regression analysis of the association of HMOX1_2790 and the FEV1 predicted value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_2790/FEV1 predicted value</td>
<td>HMOX1_2790</td>
<td>AA</td>
<td>78.39</td>
<td>2.41</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>77.08</td>
<td>2.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>70.74</td>
<td>2.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>76.47</td>
<td>2.09</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>76.08</td>
<td>2.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>75.87</td>
<td>1.62</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>78.19</td>
<td>2.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>71.14</td>
<td>4.21</td>
<td></td>
</tr>
</tbody>
</table>

Table 25: Regression analysis of the association of HMOX1_2790 and the FEV1 standard deviation value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_2790/FEV1 sd value</td>
<td>HMOX1_2790</td>
<td>AA</td>
<td>0.49</td>
<td>0.097</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>0.57</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0.24</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>0.44</td>
<td>0.08</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>0.51</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>0.44</td>
<td>0.07</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>0.58</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>0.36</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>

For HMOX1_2790, both analyses indicated the correlation between the SNP and the two phenotypic traits. However, the association only showed a borderline P value of 0.05 and 0.045 when testing for FEV1 predicted value and FEV1 sd value respectively, and therefore must be viewed with caution.
Table 26: Regression analysis of the association of HMOX1_9531 and the FEV1 predicted value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_9531/FEV1 predicted value</td>
<td>HMOX1_9531</td>
<td>AA</td>
<td>79.37</td>
<td>2.83</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>78.36</td>
<td>2.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>71.26</td>
<td>2.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>76.68</td>
<td>2.09</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>76.13</td>
<td>2.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>75.92</td>
<td>1.61</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>78.32</td>
<td>2.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>72.05</td>
<td>4.15</td>
<td></td>
</tr>
</tbody>
</table>

Table 27: Regression analysis of the association of HMOX1_9531 and the FEV1 standard deviation value with confounding factors

<table>
<thead>
<tr>
<th>SNP/phenotypic trait</th>
<th>Confounding factor</th>
<th>Sub-groups of confounding factor</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_9531/FEV1 sd value</td>
<td>HMOX1_9531</td>
<td>AA</td>
<td>0.54</td>
<td>0.11</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.61</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>0.26</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>F</td>
<td>0.45</td>
<td>0.08</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>0.52</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR genotype</td>
<td>ΔF508/ΔF508</td>
<td>0.45</td>
<td>0.06</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>0.59</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>other/other</td>
<td>0.37</td>
<td>0.17</td>
<td></td>
</tr>
</tbody>
</table>

For HMOX1_9531, both analyses indicated a strong correlation between the SNP and the two phenotypic traits, as indicated by the P value of 0.007 and 0.006.

10. Age of Onset Analysis for the Age of First Pseudomonas aeruginosa Infection

Approximately 80% of CF patients have either acute or chronic Pseudomonas aeruginosa infection. This was true for most of the
patients recruited in this study. However, there were individuals who had never been infected with \textit{Pseudomonas aeruginosa}. Assuming that there were no errors in the data collection i.e. that the age of first \textit{Pseudomonas aeruginosa} infection was missing; it was not possible to predict if these individuals would have \textit{Pseudomonas aeruginosa} infection in the future. Survival analysis was performed to investigate the association between particular SNP and age of first \textit{Pseudomonas aeruginosa} infection.

(a) Factor B

From Table A14 in the Appendix, one subgroup of each variable is not indicated, for example, only two (AA and AG) out of three possible genotypes were present in the table. The missing subgroup was the one with the greatest proportion. It was used as the baseline for comparing with other subgroups in order to reveal the effect of the genotype of selected SNPs on the phenotypic trait. These analyses indicated that none of the selected SNPs in the candidate gene of Factor B played a role in affecting the age of first \textit{Pseudomonas aeruginosa} infection.

(b) Complement Factor 3

From Table A15 in the Appendix, none of the tested SNPs illustrated a correlation with the age of first \textit{Pseudomonas aeruginosa} infection.
Table 28: Age of onset analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Toll-like receptor 4

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Variable</th>
<th>Sub-group</th>
<th>Estimated Value</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TLR4_851</strong></td>
<td>TLR4_851</td>
<td>AA</td>
<td>0.169</td>
<td>0.155</td>
<td>0.5269</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.0614</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.155</td>
<td>0.086</td>
<td>0.0712</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0192</td>
<td>0.133</td>
<td>0.9773</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.0245</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_1859</strong></td>
<td>TLR4_1859</td>
<td>AA</td>
<td>0.196</td>
<td>0.154</td>
<td>0.4581</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>-0.069</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.139</td>
<td>0.087</td>
<td>0.1109</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.036</td>
<td>0.134</td>
<td>0.9603</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.017</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_2856</strong></td>
<td>TLR4_2856</td>
<td>CC</td>
<td>0.566</td>
<td>0.393</td>
<td>0.286</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CT</td>
<td>-0.198</td>
<td>0.225</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.163</td>
<td>0.086</td>
<td>0.0598</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0293</td>
<td>0.133</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.0104</td>
<td>-0.287</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_9263</strong></td>
<td>TLR4_9263</td>
<td>AA</td>
<td>-0.292</td>
<td>0.693</td>
<td>0.0425</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AC</td>
<td>-0.132</td>
<td>0.376</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.157</td>
<td>0.087</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.066</td>
<td>0.136</td>
<td>0.8834</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.028</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_11912</strong></td>
<td>TLR4_11912</td>
<td>GG</td>
<td>0.073</td>
<td>0.128</td>
<td>0.5244</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GT</td>
<td>-0.131</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.141</td>
<td>0.088</td>
<td>0.1102</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.032</td>
<td>0.135</td>
<td>0.9537</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.031</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_15884</strong></td>
<td>TLR4_15884</td>
<td>CC</td>
<td>0.444</td>
<td>0.397</td>
<td>0.1965</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CG</td>
<td>-0.077</td>
<td>0.231</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.146</td>
<td>0.088</td>
<td>0.0983</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>0.0042</td>
<td>0.135</td>
<td>0.9515</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.0455</td>
<td>0.146</td>
<td></td>
</tr>
<tr>
<td><strong>TLR4_17050</strong></td>
<td>TLR4_17050</td>
<td>CC</td>
<td>-0.065</td>
<td>0.397</td>
<td>0.3172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CT</td>
<td>-0.121</td>
<td>0.233</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.157</td>
<td>0.087</td>
<td>0.0714</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.037</td>
<td>0.135</td>
<td>0.9143</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.051</td>
<td>0.144</td>
<td></td>
</tr>
</tbody>
</table>
Among the seven selected SNPs in the candidate gene of toll-like receptor 4, one of them (TLR4_9263) had a P value of 0.0425. This meant that this SNP was significantly associated to the age of first *Pseudomonas aeruginosa* infection.

(d) Heme oxygenase-1

From the data indicated in Table A16 in the Appendix, it can be seen that none of the six selected SNPs of heme oxygenase-1 was statistically associated to the age of first *Pseudomonas aeruginosa* infection.

11. *Pseudomonas aeruginosa* Infection Status

Not all patients recruited in this study had *Pseudomonas aeruginosa* infection. All the patients had been classified into one of the following categories according to their *Pseudomonas aeruginosa* infection status: no *Pseudomonas aeruginosa* infection (code 0), grew *Pseudomonas aeruginosa* once (code 1), sporadic *Pseudomonas aeruginosa* growth (code 2) and chronic *Pseudomonas aeruginosa* infection (code 3). It was hypothesized that one or more of the SNPs of the four candidate modifier genes might contribute to susceptibility to *Pseudomonas aeruginosa* infection of CF patients. Chi-square tests were performed to confirm the existence of such a relationship.

(a) Factor B

The data in Table A17 in the Appendix show the chi square test results for all SNPs of the Factor B gene. From the statistical analysis, the P-value of each of the Chi square tests carried out for each SNP was greater than 0.05, which indicated that *Pseudomonas aeruginosa* infection status was not related to one of five selected SNPs in the Factor B gene.

(b) Complement factor 3
Table 29: Chi square test for investigating the relationship between different genotypes of the selected SNPs in Complement factor 3 and Pseudomonas aeruginosa infection status

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>PA status (in percentage)</th>
<th>Chi Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>C3_963</td>
<td>GG</td>
<td>37.31</td>
<td>11.94</td>
<td>22.39</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>26.55</td>
<td>17.26</td>
<td>17.70</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>25.93</td>
<td>15.74</td>
<td>28.70</td>
</tr>
<tr>
<td>C3_28795</td>
<td>AA</td>
<td>28.33</td>
<td>17.78</td>
<td>21.67</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>33.81</td>
<td>13.33</td>
<td>20.48</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>21.69</td>
<td>14.46</td>
<td>22.89</td>
</tr>
<tr>
<td>C3_36735</td>
<td>AA</td>
<td>28.24</td>
<td>15.29</td>
<td>16.47</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>31.38</td>
<td>15.48</td>
<td>21.34</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>27.27</td>
<td>13.39</td>
<td>25.17</td>
</tr>
<tr>
<td>C3_43118</td>
<td>AA</td>
<td>34.55</td>
<td>14.55</td>
<td>18.18</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>31.72</td>
<td>13.66</td>
<td>19.82</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>22.06</td>
<td>18.38</td>
<td>26.47</td>
</tr>
</tbody>
</table>

The above table demonstrates that none of the C3_28795, C3_36735 and C3_43118 SNPs was associated with Pseudomonas aeruginosa infection status in CF patients. However, a borderline association was observed with the C3_963 SNP. Thus, genotype at the C3_963 polymorphism might contribute to Pseudomonas aeruginosa infection status in CF patients.

(c) Toll-like receptor 4

From the result Table A18 in the Appendix, it can be seen that all the SNPs in TLR4 had a P-value greater than 0.05 and therefore it was not likely that any of the SNPs influence Pseudomonas aeruginosa infection status among the CF patients.

(d) Heme oxygenase-1

The chi-square test was carried out for each of the SNPs in the HMOX1 gene and the results are summarized in Table A19 in the Appendix. As with the TLR4 gene, there was no relationship between genotypes and
the *Pseudomonas aeruginosa* infection status. However, 0% was detected for some SNPs in some categories of status of *Pseudomonas aeruginosa* infection. For example, all patients with the TT genotype for SNP HMOX1_1038 were either classified to be "grew *Pseudomonas aeruginosa* infection" or "chronic *Pseudomonas aeruginosa* infection". Also, all patients with the TT genotype for SNP HMOX1_16442 were found to be "no *Pseudomonas aeruginosa* infection". Logically one might think that those genotypes influenced the observed *Pseudomonas aeruginosa* infection status. However, this was not proved by the statistical test. Since the number of individuals with genotype TT at HMOX1_9531 and genotype TT at HMOX1_16442 were only 2 and 1, respectively it was not possible to draw valid conclusions regarding the influence of these SNPs on *Pseudomonas aeruginosa* infection status.

12. **Re-genotyping**

Re-genotyping was performed as a quality control measure. The selection of SNPs and individuals for re-genotyping are described in the "Materials and Methods". The comparison is shown in Table 30.

Table 30: Comparing the result of genotyping and re-genotyping

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPs</th>
<th>Number of Individuals</th>
<th>Errors found</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2557</td>
<td>618</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4022</td>
<td>618</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6484</td>
<td>618</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Complement Factor 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>963</td>
<td>609</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Toll-like Receptor 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>851</td>
<td>950</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1859</td>
<td>618</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2856</td>
<td>618</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9263</td>
<td>950</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>11912</td>
<td>950</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>15884</td>
<td>661</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>17050</td>
<td>618</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Heme oxygenase-1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1038</td>
<td>621</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2790</td>
<td>657</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3303</td>
<td>949</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9531</td>
<td>947</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16442</td>
<td>618</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Although not all SNPs and individuals were included for re-genotyping, the above selected SNPs and individuals were sufficient to
represent the whole population in this study. As can be seen from the table, the number of errors found ranged from 0 to 4. Such a low frequency of errors was acceptable and the possible reasons for the mismatch in genotypes included inconsistent concentration, contamination of the different DNA aliquots or random errors.

13. **FBAT Analysis of Phenotypic Characteristics**

FBAT was another analytical method for determining the association between a particular SNP and a phenotypic trait. FBAT differs from ANOVA in that it determines the inheritance pattern within a family, i.e., from parents to their child. FBAT employs a type of transmission disequilibrium test (TDT). Two of the three models (additive, dominant and recessive) were performed, since both dominant and recessive should give the same result. For those SNPs with a P value of less than 0.05, a detailed summary of the test is shown below. The complete FBAT analysis is shown in Tables A20-A25 of the Appendix.

(a) Age of diagnosis

Table 31: Detailed results of FBAT analysis of age of diagnosis and SNPs which had a P value of less than 0.05 under the additive model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>Allele Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3_28795</td>
<td>1</td>
<td>0.583</td>
<td>332</td>
<td>-40.89</td>
<td>-60.10</td>
<td>82.74</td>
<td>2.11</td>
<td>0.0347</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.417</td>
<td>332</td>
<td>-73.47</td>
<td>-54.26</td>
<td>82.74</td>
<td>-2.11</td>
<td>0.0347</td>
</tr>
<tr>
<td>C3_43118</td>
<td>1</td>
<td>0.474</td>
<td>342</td>
<td>-58.60</td>
<td>-77.78</td>
<td>94.39</td>
<td>1.98</td>
<td>0.0484</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.526</td>
<td>342</td>
<td>-87.27</td>
<td>-68.09</td>
<td>94.39</td>
<td>-1.98</td>
<td>0.0484</td>
</tr>
</tbody>
</table>

From the data table, only two out of the selected 22 SNPs (C3_28795 and C3_43118) were shown to have a significant association with the age of diagnosis. All SNPs were tested again under the dominant model.
Table 32: Detailed results of FBAT analysis of age of diagnosis and SNPs which had a P value of less than 0.05 under the dominant model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_7202 1</td>
<td>0.431</td>
<td>231</td>
<td>-41.85</td>
<td>-40.94</td>
<td>43.19</td>
<td>-0.14</td>
<td>0.889</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.569</td>
<td>196</td>
<td>-40.07</td>
<td>-27.23</td>
<td>34.55</td>
<td>-2.19</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>C3_28795 1</td>
<td>0.583</td>
<td>189</td>
<td>-15.82</td>
<td>-20.84</td>
<td>30.73</td>
<td>0.90</td>
<td>0.366</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.417</td>
<td>244</td>
<td>-37.96</td>
<td>-23.75</td>
<td>42.45</td>
<td>-2.18</td>
<td>0.029</td>
<td></td>
</tr>
</tbody>
</table>

By comparing with the results obtained when analyzing by the additive model, C3_28795 was also found to have an association with the age of diagnosis. However, another SNP, BF_7202 instead of C3_43118, was associated under the dominant model.

(b) FEV1 predicted value

Table 33: Detailed results of FBAT analysis of FEV1 predicted value and SNPs which had a P value of less than 0.05 under the additive model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>Allele Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557 1</td>
<td>0.184</td>
<td>201</td>
<td>8083.95</td>
<td>9512.89</td>
<td>371921.59</td>
<td>-2.34</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.816</td>
<td>201</td>
<td>21452.06</td>
<td>20023.12</td>
<td>371921.59</td>
<td>2.34</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884 1</td>
<td>0.149</td>
<td>170</td>
<td>6723.98</td>
<td>7927.42</td>
<td>327605.14</td>
<td>-2.10</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.851</td>
<td>170</td>
<td>20027.89</td>
<td>18824.45</td>
<td>327605.14</td>
<td>2.10</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038 1</td>
<td>0.959</td>
<td>59</td>
<td>6037.98</td>
<td>6846.47</td>
<td>110788.10</td>
<td>-2.43</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.041</td>
<td>59</td>
<td>3324.07</td>
<td>2515.58</td>
<td>110788.10</td>
<td>2.43</td>
<td>0.015</td>
<td></td>
</tr>
</tbody>
</table>

Among the selected 22 SNPs, three (BF_2557, TLR4_15884 and HMOX1_1038) were found to have an association with the age of diagnosis, when analyzed under the additive model by the FBAT program.
Table 34: Detailed results of FBAT analysis of FEV1 predicted value and SNPs which had a P value of less than 0.05 under the dominant model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>Allele Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557</td>
<td>1</td>
<td>0.184</td>
<td>193</td>
<td>6329.50</td>
<td>7829.51</td>
<td>279688.25</td>
<td>-2.84</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.816</td>
<td>49</td>
<td>2503.32</td>
<td>2574.39</td>
<td>60601.62</td>
<td>-0.29</td>
<td>0.773</td>
</tr>
<tr>
<td>BF_4022</td>
<td>1</td>
<td>0.319</td>
<td>232</td>
<td>8853.45</td>
<td>10359.97</td>
<td>341446.47</td>
<td>-2.58</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.681</td>
<td>118</td>
<td>5761.40</td>
<td>5993.83</td>
<td>156520.15</td>
<td>-0.59</td>
<td>0.556</td>
</tr>
<tr>
<td>C3_28795</td>
<td>1</td>
<td>0.583</td>
<td>166</td>
<td>8337.34</td>
<td>8171.95</td>
<td>240014.45</td>
<td>0.34</td>
<td>0.716</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.417</td>
<td>233</td>
<td>9551.50</td>
<td>10781.25</td>
<td>350400.00</td>
<td>-2.08</td>
<td>0.038</td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>1</td>
<td>0.149</td>
<td>165</td>
<td>5859.52</td>
<td>6945.67</td>
<td>271528.68</td>
<td>-2.08</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.851</td>
<td>27</td>
<td>1614.44</td>
<td>1497.16</td>
<td>36498.35</td>
<td>0.61</td>
<td>0.539</td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>1</td>
<td>0.959</td>
<td>4</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.041</td>
<td>59</td>
<td>3173.48</td>
<td>2428.05</td>
<td>99750.38</td>
<td>2.36</td>
<td>0.018</td>
</tr>
</tbody>
</table>

In addition to the three SNPs (BF_2557, TLR4_15884 and HMOX1_1038) which were determined have a relationship with the FEV1 predicted value when analyzed under the additive model, two more SNPs (BF_4022 and C3_28795) also achieved a P value of less than 0.05 when analyzed under the dominant model.

(c) FEV1 standard deviation value

Table 35: Detailed results of FBAT analysis of FEV1 standard deviation value and SNPs which had a P value of less than 0.05 under the additive model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele</th>
<th>Allele Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_4022</td>
<td>1</td>
<td>0.319</td>
<td>264</td>
<td>66.89</td>
<td>90.96</td>
<td>92.05</td>
<td>-2.51</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.681</td>
<td>264</td>
<td>165.27</td>
<td>141.19</td>
<td>92.05</td>
<td>2.51</td>
<td>0.012</td>
</tr>
<tr>
<td>BF_7202</td>
<td>1</td>
<td>0.431</td>
<td>293</td>
<td>99.02</td>
<td>122.02</td>
<td>106.17</td>
<td>-2.23</td>
<td>0.0256</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.569</td>
<td>293</td>
<td>164.93</td>
<td>141.92</td>
<td>106.17</td>
<td>2.23</td>
<td>0.0256</td>
</tr>
</tbody>
</table>

When investigating by the FBAT program under the additive model, two SNPs, BF_4022 and BF_7202, had a P value of less than 0.05.
Table 36: Detailed results of FBAT analysis of FEV1 standard deviation value and SNPs which had a P value of less than 0.05 under the dominant model

<table>
<thead>
<tr>
<th>Marker</th>
<th>Allele Frequency</th>
<th># of families</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557</td>
<td>1 0.184</td>
<td>193</td>
<td>19.32</td>
<td>36.10</td>
<td>48.58</td>
<td>-2.41</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>2 0.816</td>
<td>49</td>
<td>11.60</td>
<td>14.89</td>
<td>9.63</td>
<td>-1.06</td>
<td>0.290</td>
</tr>
<tr>
<td>BF_4022</td>
<td>1 0.319</td>
<td>231</td>
<td>39.14</td>
<td>56.00</td>
<td>54.16</td>
<td>-2.29</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>2 0.681</td>
<td>118</td>
<td>38.09</td>
<td>30.88</td>
<td>25.43</td>
<td>1.43</td>
<td>0.153</td>
</tr>
<tr>
<td>BF_6484</td>
<td>1 0.893</td>
<td>22</td>
<td>8.41</td>
<td>4.24</td>
<td>4.70</td>
<td>1.93</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>2 0.107</td>
<td>131</td>
<td>34.39</td>
<td>32.62</td>
<td>40.18</td>
<td>0.28</td>
<td>0.779</td>
</tr>
<tr>
<td>BF_7202</td>
<td>1 0.431</td>
<td>217</td>
<td>48.80</td>
<td>64.80</td>
<td>52.25</td>
<td>-2.21</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>2 0.569</td>
<td>179</td>
<td>61.86</td>
<td>54.85</td>
<td>38.46</td>
<td>1.13</td>
<td>0.258</td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>1 0.463</td>
<td>215</td>
<td>78.88</td>
<td>55.91</td>
<td>53.01</td>
<td>3.16</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>2 0.537</td>
<td>194</td>
<td>56.87</td>
<td>52.56</td>
<td>54.17</td>
<td>0.59</td>
<td>0.558</td>
</tr>
</tbody>
</table>

In addition to the two SNPs found to have a relationship with the phenotypic characteristic of FEV1 sd value, three more SNPs (BF_2557, BF_6484 and HMOX1_9531) were also determined to have a P value of less than 0.05 when analyzed by the FBAT program under the dominant model.

14. **Haplotype Analysis by the RGui Program**

In the haplotype analysis, combinations of alleles of the selected SNPs of each candidate gene were examined as a group, in order to investigate any association between the haplotype and the phenotypic characteristics. Single SNP analysis had been done by both ANOVA and FBAT, as described in the previous sections. However, combinations of SNPs might offer additional insights into the correlation between the genotype and the phenotype, since the SNPs might interact with one another.
(a) Factor B

(I) Age of diagnosis

Table 37: Analysis of the correlation between haplotypes of selected SNPs in Factor B and the age of diagnosis, with no adjustment for confounding factors. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h11111</td>
<td>0.343</td>
<td>0.168</td>
<td>2.045</td>
<td>0.0409</td>
</tr>
<tr>
<td>h21211</td>
<td>-0.019</td>
<td>0.217</td>
<td>-0.088</td>
<td>0.9299</td>
</tr>
<tr>
<td>h22112</td>
<td>0.552</td>
<td>0.230</td>
<td>2.405</td>
<td>0.0162</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0506</td>
<td>0.2632</td>
<td>-0.1922</td>
<td>0.8476</td>
</tr>
</tbody>
</table>

32 haplotypes were theoretically possible when analyzing the Factor B gene with the five selected SNPs. However, only 8 haplotypes were reported by the program as shown in Table A26 in the Appendix. Since the frequencies of five of the 8 haplotypes were too low, they were combined as a pooled haplotype for the analysis. Two haplotypes (h11111 and h22112) were found to have a P value of less than 0.05, suggesting that an association was present. The investigation was continued with further analysis including confounding factors to confirm the association.

Table 38: Analysis of the correlation between haplotypes of selected SNPs in Factor B and the age of diagnosis, with adjustment for confounding factors. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h11111</td>
<td>0.3386</td>
<td>0.1675</td>
<td>2.0210</td>
<td>0.0433</td>
</tr>
<tr>
<td>h21211</td>
<td>-0.0169</td>
<td>0.2158</td>
<td>-0.0784</td>
<td>0.9375</td>
</tr>
<tr>
<td>h22112</td>
<td>0.5353</td>
<td>0.2287</td>
<td>2.3405</td>
<td>0.0193</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0926</td>
<td>0.2287</td>
<td>-0.3472</td>
<td>0.7284</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0247</td>
<td>0.1820</td>
<td>0.1359</td>
<td>0.8919</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.4342</td>
<td>0.2077</td>
<td>2.0910</td>
<td>0.0365</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>0.4385</td>
<td>0.2903</td>
<td>1.5106</td>
<td>0.1309</td>
</tr>
</tbody>
</table>
The same eight haplotypes were found when analyzing with confounding factors. In addition, both h11111 and h22112 were found to have a P value of less than 0.05, which confirmed the association between these two haplotypes in the Factor B gene and age of diagnosis.

(II) FEV1 predicted value

Among the eight reported haplotypes shown in Table A27 in the Appendix, none demonstrated a relationship with the FEV1 predicted value. Similar results were found when examining with the confounding factors. From Tables A28 and A29 in the Appendix, none of the eight haplotypes in the Factor B gene showed an association with the FEV1 predicted value.

(III) FEV1 standard deviation value

Table 39: Analysis of the correlation between haplotypes of selected SNPs in Factor B and FEV1 standard deviation value, with no adjustment for confounding factors

<table>
<thead>
<tr>
<th></th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h11111</td>
<td>-0.1350</td>
<td>0.0914</td>
<td>-1.4772</td>
<td>0.1396</td>
</tr>
<tr>
<td>h21211</td>
<td>-0.1512</td>
<td>0.1163</td>
<td>-1.3002</td>
<td>0.1935</td>
</tr>
<tr>
<td>h22112</td>
<td>-0.1050</td>
<td>0.1156</td>
<td>-0.9084</td>
<td>0.3636</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.3546</td>
<td>0.1366</td>
<td>-2.5957</td>
<td>0.0094</td>
</tr>
</tbody>
</table>

As indicated in the above table, the pooled haplotypes had a P value of 0.0094, which suggested an association with the phenotype. However, the pooled group included five low frequency haplotypes (table A30 of the Appendix). Therefore, it was not possible to determine which haplotype in the group generated the observed result.
Table 40: Analysis of the correlation between haplotypes of selected SNPs in Factor B and FEV1 standard deviation value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1111</td>
<td>-0.1421</td>
<td>0.0915</td>
<td>-1.5538</td>
<td>0.1202</td>
</tr>
<tr>
<td>h2121</td>
<td>-0.1574</td>
<td>0.1160</td>
<td>-1.3567</td>
<td>0.1749</td>
</tr>
<tr>
<td>h2212</td>
<td>-0.1225</td>
<td>0.1159</td>
<td>-1.0569</td>
<td>0.2906</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.3946</td>
<td>0.1389</td>
<td>-2.8402</td>
<td>0.0045</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0053</td>
<td>0.0981</td>
<td>0.0537</td>
<td>0.9572</td>
</tr>
<tr>
<td>genotypePO</td>
<td>0.1348</td>
<td>0.1128</td>
<td>1.1943</td>
<td>0.2323</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>-0.1104</td>
<td>0.1537</td>
<td>-0.7182</td>
<td>0.4726</td>
</tr>
</tbody>
</table>

As shown in Table 40, a significant P value was found for the pooled haplotypes. This confirmed the results when analyzing without confounding factors. However, this analysis was not able to narrow down the relationship to just one of the five haplotypes in the group.

(b) Complement Factor 3

(I) Age of diagnosis

Table 41: Analysis of the correlation between haplotypes of selected SNPs in Complement factor 3 and the age of diagnosis, with no adjustment for confounding factors. The age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-0.3883</td>
<td>0.3514</td>
<td>-1.1051</td>
<td>0.2691</td>
</tr>
<tr>
<td>h1122</td>
<td>0.0310</td>
<td>0.3892</td>
<td>0.0798</td>
<td>0.9364</td>
</tr>
<tr>
<td>h1212</td>
<td>-0.9269</td>
<td>0.4435</td>
<td>-2.0899</td>
<td>0.0366</td>
</tr>
<tr>
<td>h2121</td>
<td>-0.1406</td>
<td>0.4514</td>
<td>-0.3114</td>
<td>0.7555</td>
</tr>
<tr>
<td>h2122</td>
<td>0.0846</td>
<td>0.3576</td>
<td>0.2366</td>
<td>0.8130</td>
</tr>
<tr>
<td>h2212</td>
<td>-0.1365</td>
<td>0.3439</td>
<td>-0.3969</td>
<td>0.6915</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.1417</td>
<td>0.2889</td>
<td>-0.4905</td>
<td>0.6238</td>
</tr>
</tbody>
</table>

Four SNPs were included in the analysis of the Complement factor 3 gene. There were a possible 16 haplotypes and all were detected in the participants (Table A31 of the Appendix). However, only six haplotypes with high frequencies were examined individually while the
others were grouped together. As indicated in the table, only h1212 revealed a P value of less than 0.05, and therefore a significant relationship between this haplotype and age of diagnosis was indicated. The analysis was repeated taking the confounding factors into consideration.

Table 42: Analysis of the correlation between combinations of selected SNPs in Complement factor 3 and the age of diagnosis, with adjustment for confounding factors. The age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th></th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-0.3764</td>
<td>0.3484</td>
<td>-1.0804</td>
<td>0.2800</td>
</tr>
<tr>
<td>h1122</td>
<td>0.0112</td>
<td>0.3849</td>
<td>0.0290</td>
<td>0.9769</td>
</tr>
<tr>
<td>h1212</td>
<td>-1.1090</td>
<td>0.4397</td>
<td>-2.5223</td>
<td>0.0117</td>
</tr>
<tr>
<td>h2121</td>
<td>-0.1822</td>
<td>0.4379</td>
<td>-0.4160</td>
<td>0.6774</td>
</tr>
<tr>
<td>h2122</td>
<td>0.0093</td>
<td>0.3570</td>
<td>0.0261</td>
<td>0.9792</td>
</tr>
<tr>
<td>h2212</td>
<td>-0.1949</td>
<td>0.3399</td>
<td>-0.5735</td>
<td>0.5663</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.1463</td>
<td>0.2859</td>
<td>-0.5117</td>
<td>0.6088</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.1086</td>
<td>0.1812</td>
<td>0.5994</td>
<td>0.5489</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.4907</td>
<td>0.2076</td>
<td>2.3639</td>
<td>0.0181</td>
</tr>
<tr>
<td>genotype00</td>
<td>0.4984</td>
<td>0.2899</td>
<td>1.7190</td>
<td>0.0856</td>
</tr>
</tbody>
</table>

A similar result was obtained when examining the haplotypes correcting for the confounding factors, as shown in the above table. This indicated an association of h1212 with the age of diagnosis.

(II) FEV1 predicted value

In Table A32 in the Appendix, 16 haplotypes were detected when determining the presence of any correlation between the haplotypes in Complement Factor 3 and FEV1 predicted value. As shown in Table A33 in the Appendix, none of the six haplotypes and the pooled group demonstrated an association with the phenotype of FEV1 predicted when no adjustments were made for confounding factors. The conclusion was confirmed when analyzing with the confounding factors

Table A34 in the Appendix matched the one in the preceding section, which suggested that none of the haplotypes in Complement
factor 3 was related to the observed FEV1 predicted value after adjustments were made for confounding factors.

(III) FEV1 standard deviation value

16 haplotypes were found when determining the presence of any correlation between the haplotypes in Complement Factor 3 and FEV1 standard deviation, as indicated in Table A35 in the Appendix. In Table A36 in the Appendix, all of the SNPs were found to have a P value of greater than 0.05 when no adjustments were made for confounding factors. That is, none of the SNPs were related to the FEV1 standard deviation value observed in the patients.

From Table A37 in the Appendix, none of the haplotypes in Complement factor 3 showed a significant relationship with the FEV1 standard deviation value, even when the confounding factors were included.

(c) Toll-like Receptor 4

(I) Age of diagnosis

Table 43: Analysis of the correlation between haplotypes of selected SNPs in Toll-like receptor 4 and the age of diagnosis, with no adjustment for confounding factors. The age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1122112</td>
<td>-0.1485</td>
<td>0.2048</td>
<td>-0.7250</td>
<td>0.4684</td>
</tr>
<tr>
<td>h1122122</td>
<td>0.1282</td>
<td>0.1835</td>
<td>0.6990</td>
<td>0.4846</td>
</tr>
<tr>
<td>h1212222</td>
<td>-0.3101</td>
<td>0.2334</td>
<td>-1.3284</td>
<td>0.1840</td>
</tr>
<tr>
<td>h1221221</td>
<td>-0.0963</td>
<td>0.2422</td>
<td>-0.3977</td>
<td>0.6908</td>
</tr>
<tr>
<td>h1222222</td>
<td>-0.3622</td>
<td>0.2673</td>
<td>-1.3554</td>
<td>0.1753</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.6202</td>
<td>0.2912</td>
<td>-2.1296</td>
<td>0.0332</td>
</tr>
</tbody>
</table>

From the seven SNPs included in the Toll-like receptor 4 gene, 128 possible haplotypes were theoretically possible. However, only 15 haplotypes were detected in the sample (Table A38 in the Appendix). Five out of the seven haplotypes were analyzed individually whereas the
rest were grouped. Only this pooled group of haplotypes showed a $P$ value of less than 0.05, however, it is not possible to determine which in the pooled group was the driving force behind the observed result.

Table 44: Analysis of the correlation between haplotypes of selected SNPs in Toll-like receptor 4 and the age of diagnosis, with adjustment for the confounding factors. The age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1122112</td>
<td>-0.1676</td>
<td>0.2049</td>
<td>-0.8178</td>
<td>0.4135</td>
</tr>
<tr>
<td>h1122122</td>
<td>0.1098</td>
<td>0.1828</td>
<td>0.6006</td>
<td>0.5481</td>
</tr>
<tr>
<td>h1212222</td>
<td>-0.3163</td>
<td>0.2325</td>
<td>-1.3604</td>
<td>0.1737</td>
</tr>
<tr>
<td>h1221221</td>
<td>-0.1469</td>
<td>0.2428</td>
<td>-0.6052</td>
<td>0.5451</td>
</tr>
<tr>
<td>h1222222</td>
<td>-0.4295</td>
<td>0.2667</td>
<td>-1.6105</td>
<td>0.1073</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.6439</td>
<td>0.2887</td>
<td>-2.2302</td>
<td>0.0257</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0433</td>
<td>0.1817</td>
<td>0.2381</td>
<td>0.8118</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.4723</td>
<td>0.2067</td>
<td>2.2852</td>
<td>0.0223</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>0.4487</td>
<td>0.2894</td>
<td>1.5505</td>
<td>0.1210</td>
</tr>
</tbody>
</table>

As indicated in the above table, the pooled group demonstrated a $P$ value of 0.0257. That is, one or more of the haplotypes in the group might be related to the observed age of diagnosis in the patients.

(II) FEV1 predicted value

15 haplotypes were detected when determining the presence of any correlation between haplotypes in Toll-like receptor 4 and FEV1 predicted value, as shown in Table A39 in the Appendix. Among the five reported haplotypes and the pooled group shown in Table A40 in the Appendix, none of them revealed a $P$ value of less than 0.05 and therefore a significant relationship with the FEV1 predicted value.

From Table A41 in the Appendix, the same conclusion could be deduced when compared with the results analyzed in the absence of confounding factors. None of the haplotypes were shown to be related to the FEV1 predicted value.
(III) FEV1 standard deviation value

Indicated in Table A42 in the Appendix, 15 haplotypes were included when determining the presence of any correlation between these haplotypes in Toll-like receptor 4 and FEV1 standard deviation value. When compared to the analysis of the phenotype of FEV1 predicted, the same conclusion was found since none of the haplotypes were related to the FEV1 standard deviation value, as shown in Table A43 in the Appendix.

Table A44 in the Appendix further strengthened the conclusion in the preceding section, which indicated that none of the haplotypes was found to have a significant association with the FEV1 standard deviation value after adjustments were made for confounding factors.

(d) Heme Oxygenase-1

(I) Age of diagnosis

Indicated in Table A45 in the Appendix, 64 haplotypes were theoretically possible from the six SNPs selected in the heme oxygenase-1 gene. However, only 15 were observed in the sample. Four of the haplotypes with high frequency were considered separately while the rest were grouped together. None of them were found to be associated with the age of diagnosis when no adjustments were made for confounding factors (Table A46 of the Appendix).

In Table A47 in the Appendix, no haplotype was found to be related to the observed age of diagnosis in the sample after confounding factors were adjusted for.
Table 45: Analysis of the correlation between haplotypes of selected SNPs in heme oxygenase-1 and FEV1 predicted value, with no adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-6.3390</td>
<td>2.2781</td>
<td>-2.7826</td>
<td>0.0054</td>
</tr>
<tr>
<td>h211221</td>
<td>-4.2824</td>
<td>3.3487</td>
<td>-1.2788</td>
<td>0.2010</td>
</tr>
<tr>
<td>h212221</td>
<td>-4.9749</td>
<td>4.4609</td>
<td>-1.1152</td>
<td>0.2648</td>
</tr>
<tr>
<td>pooled</td>
<td>2.3963</td>
<td>3.1362</td>
<td>0.7641</td>
<td>0.4448</td>
</tr>
</tbody>
</table>

14 haplotypes were detected when determining the presence of any correlation between haplotypes in Heme oxygenase-1 and FEV1 predicted value, as shown in Table A48 in the Appendix. The above table showed that h112221 had a P value of 0.0054, which indicated that this haplotype was associated with the FEV1 predicted value.

Table 46: Analysis for of the correlation between haplotypes of selected SNPs in heme oxygenase-1 and FEV1 predicted value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-6.3269</td>
<td>2.2775</td>
<td>-2.7780</td>
<td>0.0055</td>
</tr>
<tr>
<td>h211221</td>
<td>-3.9819</td>
<td>3.3483</td>
<td>-1.1893</td>
<td>0.2343</td>
</tr>
<tr>
<td>h212221</td>
<td>-5.0514</td>
<td>4.4708</td>
<td>-1.1298</td>
<td>0.2585</td>
</tr>
<tr>
<td>pooled</td>
<td>2.4056</td>
<td>3.1343</td>
<td>0.7675</td>
<td>0.4428</td>
</tr>
<tr>
<td>SEXM</td>
<td>-3.1139</td>
<td>2.7183</td>
<td>-1.1455</td>
<td>0.2520</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>2.0136</td>
<td>3.1297</td>
<td>0.6434</td>
<td>0.5200</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>0.2597</td>
<td>4.2367</td>
<td>0.0613</td>
<td>0.9511</td>
</tr>
</tbody>
</table>

When taking the confounding factors into account, the h112221 haplotype also demonstrated a P value of less than 0.05. Therefore, the association between the haplotype h112221 and the FEV1 predicted value was confirmed.
Table 47: Analysis of the correlation between haplotypes of selected SNPs in heme oxygenase-1 and FEV1 standard deviation value, with no adjustment for the confounding factors.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-0.2351</td>
<td>0.0815</td>
<td>-2.8838</td>
<td>0.0039</td>
</tr>
<tr>
<td>h211221</td>
<td>-0.2813</td>
<td>0.1198</td>
<td>-2.3488</td>
<td>0.0188</td>
</tr>
<tr>
<td>h212221</td>
<td>-0.0639</td>
<td>0.1611</td>
<td>-0.3965</td>
<td>0.6917</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0584</td>
<td>0.1124</td>
<td>-0.5196</td>
<td>0.6033</td>
</tr>
</tbody>
</table>

As shown in Table A49 in the Appendix, 14 haplotypes were included when determining the presence of any correlation between haplotypes in Heme oxygenase-1 and FEV1 standard deviation value. Two of the SNPs, h112221 and h211221, showed a P value of less than 0.05. Therefore, these two haplotypes had a significant association with the FEV1 standard deviation value.

Table 48: Analysis of the correlation between haplotypes of selected SNPs in heme oxygenase-1 and FEV1 standard deviation value, with adjustment for the confounding factors.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-0.2336</td>
<td>0.0816</td>
<td>-2.8635</td>
<td>0.0042</td>
</tr>
<tr>
<td>h211221</td>
<td>-0.2796</td>
<td>0.1200</td>
<td>-2.3300</td>
<td>0.0198</td>
</tr>
<tr>
<td>h212221</td>
<td>-0.0537</td>
<td>0.1614</td>
<td>-0.3325</td>
<td>0.7395</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0644</td>
<td>0.1125</td>
<td>-0.5720</td>
<td>0.5673</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0203</td>
<td>0.0973</td>
<td>0.2086</td>
<td>0.8347</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.0576</td>
<td>0.1120</td>
<td>0.5143</td>
<td>0.6070</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>-0.0837</td>
<td>0.1516</td>
<td>-0.5520</td>
<td>0.5810</td>
</tr>
</tbody>
</table>

The above table revealed the same conclusion as in Table 47: both h112221 and h211221 were shown to have a significant relationship to the FEV1 standard deviation value.
15. **Haplotype Analysis by the FBAT Program**

Although haplotype analysis was repeated utilizing another program, the FBAT program, different results might have resulted when compared with the results obtained from the RGui program. The FBAT analysis was a transmission disequilibrium test and it considered the inheritance pattern within a family. However, a stronger conclusion could be deduced if consistent results were obtained in these two parts of the project.

(a) **Factor B**

With the five SNPs selected in the Factor B gene, only ten of the possible 32 haplotypes were detected in the sample (Table A50 in the Appendix).

(I) **Age of diagnosis**

As can be seen from Table A51 in the Appendix, none of the haplotypes in Factor B was found to be significantly associated with the age of diagnosis.

(II) **FEV1 predicted value**

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>233</td>
<td>24199.545</td>
<td>23181.74</td>
<td>571383.658</td>
<td>1.346</td>
<td>0.178</td>
</tr>
<tr>
<td>h2</td>
<td>162</td>
<td>7280.643</td>
<td>8812.768</td>
<td>360007.335</td>
<td>-2.554</td>
<td>0.011</td>
</tr>
<tr>
<td>h3</td>
<td>108</td>
<td>5376.241</td>
<td>5192.011</td>
<td>197300.266</td>
<td>0.415</td>
<td>0.678</td>
</tr>
<tr>
<td>h4</td>
<td>114</td>
<td>5409.311</td>
<td>5022.368</td>
<td>203605.092</td>
<td>0.858</td>
<td>0.391</td>
</tr>
<tr>
<td>h5</td>
<td>50</td>
<td>2192.113</td>
<td>2266.254</td>
<td>89422.653</td>
<td>-0.248</td>
<td>0.804</td>
</tr>
<tr>
<td>h6</td>
<td>26</td>
<td>1098.942</td>
<td>1109.629</td>
<td>44874.197</td>
<td>-0.050</td>
<td>0.960</td>
</tr>
<tr>
<td>h7</td>
<td>19</td>
<td>692.77</td>
<td>703.602</td>
<td>28218.472</td>
<td>-0.064</td>
<td>0.949</td>
</tr>
<tr>
<td>h8</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h9</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Among the eight haplotypes, h2 (11111) was shown to have a P value of 0.011. This indicated that there was a significant relationship with the FEV1 predicted value.

(III) FEV1 standard deviation value

Table 50: Analysis of the correlation between haplotypes of selected SNPs in Factor B and FEV1 standard deviation value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>233</td>
<td>157.872</td>
<td>131.288</td>
<td>86.494</td>
<td>2.858</td>
<td>0.004</td>
</tr>
<tr>
<td>h2</td>
<td>162</td>
<td>23.583</td>
<td>39.301</td>
<td>59.438</td>
<td>-2.039</td>
<td>0.041</td>
</tr>
<tr>
<td>h3</td>
<td>107</td>
<td>28.777</td>
<td>31.268</td>
<td>39.812</td>
<td>-0.395</td>
<td>0.693</td>
</tr>
<tr>
<td>h4</td>
<td>113</td>
<td>27.559</td>
<td>28.826</td>
<td>27.411</td>
<td>-0.242</td>
<td>0.809</td>
</tr>
<tr>
<td>h5</td>
<td>50</td>
<td>5.966</td>
<td>11.568</td>
<td>11.986</td>
<td>-1.618</td>
<td>0.106</td>
</tr>
<tr>
<td>h6</td>
<td>26</td>
<td>0.443</td>
<td>3.711</td>
<td>8.222</td>
<td>-1.140</td>
<td>0.254</td>
</tr>
<tr>
<td>h7</td>
<td>19</td>
<td>4.849</td>
<td>3.14</td>
<td>5.123</td>
<td>0.755</td>
<td>0.450</td>
</tr>
<tr>
<td>h8</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h9</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Two haplotypes, h11111 and h22121, were revealed to have a P value of less than 0.05, which showed that both were associated with the FEV1 standard deviation value.

(b) Complement Factor 3

Of the 16 possible haplotypes due to the four selected SNPs in the complement factor 3 gene, all were observed in the sample (Table A52 in the Appendix).
(I) Age of diagnosis

Table 51: Analysis of the correlation between haplotypes of selected SNPs in Complement factor 3 and age of diagnosis by the FBAT program. Age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>164</td>
<td>-20.120</td>
<td>-24.035</td>
<td>32.335</td>
<td>0.689</td>
<td>0.491</td>
</tr>
<tr>
<td>h2</td>
<td>139</td>
<td>-13.746</td>
<td>-18.993</td>
<td>31.238</td>
<td>0.939</td>
<td>0.348</td>
</tr>
<tr>
<td>h3</td>
<td>129</td>
<td>-11.276</td>
<td>-11.547</td>
<td>20.088</td>
<td>0.06</td>
<td>0.952</td>
</tr>
<tr>
<td>h4</td>
<td>121</td>
<td>-16.908</td>
<td>-15.696</td>
<td>22.35</td>
<td>-0.256</td>
<td>0.798</td>
</tr>
<tr>
<td>h5</td>
<td>100</td>
<td>-11.562</td>
<td>-17.531</td>
<td>19.37</td>
<td>1.356</td>
<td>0.175</td>
</tr>
<tr>
<td>h6</td>
<td>98</td>
<td>-21.861</td>
<td>-14.178</td>
<td>18.548</td>
<td>-1.784</td>
<td>0.074</td>
</tr>
<tr>
<td>h7</td>
<td>105</td>
<td>-13.969</td>
<td>-13.781</td>
<td>15.621</td>
<td>-0.048</td>
<td>0.962</td>
</tr>
<tr>
<td>h8</td>
<td>76</td>
<td>-4.641</td>
<td>-5.385</td>
<td>13.185</td>
<td>0.205</td>
<td>0.838</td>
</tr>
<tr>
<td>h9</td>
<td>77</td>
<td>-9.517</td>
<td>-9.925</td>
<td>14.561</td>
<td>0.107</td>
<td>0.915</td>
</tr>
<tr>
<td>h10</td>
<td>71</td>
<td>-9.672</td>
<td>-8.116</td>
<td>12.435</td>
<td>-0.441</td>
<td>0.659</td>
</tr>
<tr>
<td>h11</td>
<td>72</td>
<td>-19.591</td>
<td>-12.871</td>
<td>15.543</td>
<td>-1.704</td>
<td>0.088</td>
</tr>
<tr>
<td>h12</td>
<td>69</td>
<td>-3.127</td>
<td>-4.449</td>
<td>12.926</td>
<td>0.358</td>
<td>0.713</td>
</tr>
<tr>
<td>h13</td>
<td>65</td>
<td>-3.385</td>
<td>-9.696</td>
<td>8.873</td>
<td>2.119</td>
<td>0.034</td>
</tr>
<tr>
<td>h14</td>
<td>37</td>
<td>-6.114</td>
<td>-1.435</td>
<td>11.447</td>
<td>-1.383</td>
<td>0.167</td>
</tr>
<tr>
<td>h15</td>
<td>30</td>
<td>-6.596</td>
<td>-3.471</td>
<td>4.687</td>
<td>-1.443</td>
<td>0.149</td>
</tr>
<tr>
<td>h16</td>
<td>34</td>
<td>-2.501</td>
<td>-3.474</td>
<td>4.957</td>
<td>0.437</td>
<td>0.662</td>
</tr>
</tbody>
</table>

Among the 16 observed haplotypes, h13 (1111) was found to have a P value of 0.034, indicating a significant relationship with age of diagnosis.

(II) FEV1 predicted value

As indicated in Table A53 in the Appendix, none of the 16 haplotypes showed a P value which was less than 0.05 when analyzing with FEV1 predicted value.
Table 52: Analysis of the correlation between haplotypes of selected SNPs in Complement factor 3 and FEV1 standard deviation value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>146</td>
<td>54.271</td>
<td>53.023</td>
<td>37.776</td>
<td>0.203</td>
<td>0.839</td>
</tr>
<tr>
<td>h2</td>
<td>132</td>
<td>46.542</td>
<td>42.171</td>
<td>52.292</td>
<td>0.605</td>
<td>0.545</td>
</tr>
<tr>
<td>h3</td>
<td>124</td>
<td>35.346</td>
<td>30.546</td>
<td>27.929</td>
<td>0.908</td>
<td>0.364</td>
</tr>
<tr>
<td>h4</td>
<td>107</td>
<td>24.055</td>
<td>28.920</td>
<td>35.015</td>
<td>-0.822</td>
<td>0.411</td>
</tr>
<tr>
<td>h5</td>
<td>92</td>
<td>32.645</td>
<td>31.652</td>
<td>23.609</td>
<td>0.204</td>
<td>0.838</td>
</tr>
<tr>
<td>h6</td>
<td>89</td>
<td>12.940</td>
<td>24.835</td>
<td>24.024</td>
<td>-2.427</td>
<td>0.015</td>
</tr>
<tr>
<td>h7</td>
<td>97</td>
<td>24.051</td>
<td>22.509</td>
<td>19.450</td>
<td>0.350</td>
<td>0.727</td>
</tr>
<tr>
<td>h8</td>
<td>63</td>
<td>27.133</td>
<td>23.455</td>
<td>30.673</td>
<td>0.664</td>
<td>0.507</td>
</tr>
<tr>
<td>h9</td>
<td>69</td>
<td>15.999</td>
<td>14.696</td>
<td>17.247</td>
<td>0.314</td>
<td>0.754</td>
</tr>
<tr>
<td>h10</td>
<td>66</td>
<td>14.908</td>
<td>15.584</td>
<td>16.326</td>
<td>-0.167</td>
<td>0.867</td>
</tr>
<tr>
<td>h11</td>
<td>63</td>
<td>14.001</td>
<td>16.883</td>
<td>15.587</td>
<td>-0.730</td>
<td>0.465</td>
</tr>
<tr>
<td>h12</td>
<td>61</td>
<td>21.521</td>
<td>17.252</td>
<td>16.101</td>
<td>1.064</td>
<td>0.287</td>
</tr>
<tr>
<td>h13</td>
<td>61</td>
<td>21.365</td>
<td>19.758</td>
<td>15.987</td>
<td>0.402</td>
<td>0.688</td>
</tr>
<tr>
<td>h14</td>
<td>33</td>
<td>5.338</td>
<td>8.541</td>
<td>10.861</td>
<td>-0.972</td>
<td>0.331</td>
</tr>
<tr>
<td>h15</td>
<td>30</td>
<td>6.126</td>
<td>5.778</td>
<td>5.522</td>
<td>0.148</td>
<td>0.882</td>
</tr>
<tr>
<td>h16</td>
<td>29</td>
<td>4.254</td>
<td>4.891</td>
<td>4.121</td>
<td>-0.313</td>
<td>0.754</td>
</tr>
</tbody>
</table>

Haplotype #6 (1212) was shown to have a P value of 0.015, suggesting an association between the haplotype 1212 and the FEV1 standard deviation value.

(c) Toll-like Receptor 4

With the seven selected SNPs in the gene of Toll-like receptor 4, 17 out of 128 possible haplotypes were reported by the FBAT program (Table A54 in the Appendix).

(I) Age of diagnosis

From Table A55 in the Appendix, none of the haplotypes revealed an association with the age of diagnosis.
Table 53: Analysis of the correlation between haplotypes of selected SNPs in Toll-like receptor 4 and FEV1 predicted value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>215</td>
<td>13935.028</td>
<td>14499.250</td>
<td>503036.881</td>
<td>-0.796</td>
<td>0.426</td>
</tr>
<tr>
<td>h2</td>
<td>198</td>
<td>13166.095</td>
<td>12221.084</td>
<td>434711.069</td>
<td>1.433</td>
<td>0.152</td>
</tr>
<tr>
<td>h3</td>
<td>149</td>
<td>6730.848</td>
<td>7850.667</td>
<td>302636.978</td>
<td>-2.036</td>
<td>0.042</td>
</tr>
<tr>
<td>h4</td>
<td>114</td>
<td>5983.613</td>
<td>5480.313</td>
<td>181054.614</td>
<td>1.183</td>
<td>0.237</td>
</tr>
<tr>
<td>h5</td>
<td>112</td>
<td>6168.453</td>
<td>5826.104</td>
<td>212860.982</td>
<td>0.742</td>
<td>0.458</td>
</tr>
<tr>
<td>h6</td>
<td>76</td>
<td>4058.668</td>
<td>3499.135</td>
<td>122261.301</td>
<td>1.600</td>
<td>0.110</td>
</tr>
<tr>
<td>h7</td>
<td>57</td>
<td>2226.345</td>
<td>2409.595</td>
<td>98033.212</td>
<td>-0.585</td>
<td>0.558</td>
</tr>
<tr>
<td>h8</td>
<td>18</td>
<td>406.307</td>
<td>727.964</td>
<td>33417.196</td>
<td>-1.760</td>
<td>0.078</td>
</tr>
<tr>
<td>h9</td>
<td>10</td>
<td>172.570</td>
<td>330.288</td>
<td>12434.211</td>
<td>-1.414</td>
<td>0.157</td>
</tr>
<tr>
<td>h10</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h22</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h23</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

One of the haplotypes, h3 (1122112), was demonstrated to have a P value of less than 0.05, indicating that this haplotype was associated with the FEV1 predicted value observed in patients.

(III) FEV1 standard deviation value

As indicated in Table A56 in the Appendix, none of the observed haplotypes in Toll-like receptor 4 revealed a significant relationship with the FEV1 standard deviation value.
(d) **Heme Oxygenase-1**

With the six selected SNPs in the heme oxygenase-1 gene, 64 haplotypes were possible. Only 20 of them were observed in the participants as indicated in Table A57 in the Appendix.

(I) **Age of diagnosis**

Among the detected haplotypes seen in the sample, none of them had a $P$ value of smaller than 0.05 as shown in Table A58 in the Appendix.

(II) **FEV1 predicted value**

Table 54: Analysis of the correlation between haplotypes of selected SNPs in Heme oxygenase-1 and FEV1 predicted value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>$E(S)$</th>
<th>$Var(S)$</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>234</td>
<td>17514.369</td>
<td>18099.875</td>
<td>516761.732</td>
<td>-0.814</td>
<td>0.415</td>
</tr>
<tr>
<td>h2</td>
<td>200</td>
<td>14639.825</td>
<td>14051.769</td>
<td>445759.360</td>
<td>0.881</td>
<td>0.378</td>
</tr>
<tr>
<td>h3</td>
<td>114</td>
<td>4934.857</td>
<td>5071.520</td>
<td>205314.256</td>
<td>-0.302</td>
<td>0.763</td>
</tr>
<tr>
<td>h4</td>
<td>69</td>
<td>2971.849</td>
<td>2932.472</td>
<td>114032.044</td>
<td>0.117</td>
<td>0.907</td>
</tr>
<tr>
<td>h5</td>
<td>44</td>
<td>2769.310</td>
<td>2171.419</td>
<td>78544.271</td>
<td>2.133</td>
<td>0.033</td>
</tr>
<tr>
<td>h6</td>
<td>47</td>
<td>1833.378</td>
<td>2138.615</td>
<td>73738.049</td>
<td>-1.124</td>
<td>0.261</td>
</tr>
<tr>
<td>h7</td>
<td>15</td>
<td>613.016</td>
<td>631.542</td>
<td>20264.602</td>
<td>-0.130</td>
<td>0.896</td>
</tr>
<tr>
<td>h8</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h9</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>7</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

One haplotype, h3 (222111), showed a $P$ value of 0.033. Therefore, this haplotype was determined to have a significant relationship with the FEV1 predicted value.
None of the observed haplotypes revealed a P value of less than 0.05 as shown in Table A59 in the Appendix, consequently, no haplotypes in the heme oxygenase-1 gene were found to be significantly associated with the FEV1 standard deviation value.

16. **Haplotype Analysis of the Age of First *Pseudomonas aeruginosa* Infection Utilizing Hapstat**

In addition to examining single locus association between the 22 selected SNPs and the age of first *Pseudomonas aeruginosa* infection, the relationship between the haplotypes formed by the SNPs within each gene and this phenotypic trait was investigated. This test could not be done by either the RGui program or PBAT, due to the fact that the age of first *Pseudomonas aeruginosa* infection was not available if the organism had not colonized the patients during data collection period. Therefore, this analysis was performed utilizing the Hapstat program.

(a) **Factor B**

Among the five selected SNPs in the Factor B gene, 32 possible haplotypes were theoretically possible. However, only 3 haplotypes were analyzed utilizing a cutoff haplotype frequency of greater than 0.05. As indicated in Table A60 in the Appendix, none of the haplotypes showed a relationship to the age of first *Pseudomonas aeruginosa* infection.

(b) **Complement Factor 3**

Among the four selected SNPs in Complement factor 3 gene, 16 haplotypes were possible. However, only 8 haplotypes were analyzed utilizing a cutoff haplotype frequency of greater than 0.05. As indicated in Table A61 in the Appendix, none of the haplotypes showed a relationship to the age of first *Pseudomonas aeruginosa* infection.
(c) Toll-like Receptor 4

Among the seven selected SNPs in Toll-like receptor 4 gene, 128 haplotypes were possible. However, only 5 haplotypes were analyzed utilizing a cutoff haplotype frequency of greater than 0.05. As indicated in Table A62 in the Appendix, none of the haplotypes showed a relationship to the age of first *Pseudomonas aeruginosa* infection.

(d) Heme Oxygenase-1

Among the six selected SNPs in the Heme oxygenase-1 gene, 64 possible haplotypes were possible. However, only 4 haplotypes were analyzed utilizing a cutoff haplotype frequency of greater than 0.05. As indicated in Table A63 in the Appendix, none of the haplotypes showed a relationship to the age of first *Pseudomonas aeruginosa* infection.
Chapter 4 Discussion

This project is a sub-study of a large, Canada-wide and international endeavor: the Canadian Consortium for Cystic Fibrosis Modifiers (http://www.cfmod.ca). It is well-established that Cystic Fibrosis is an autosomal recessive disease which is caused by mutations in the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene located on chromosome 7. There is a wide range of clinical symptoms among affected individuals, for example, deterioration of the lungs, pancreatic insufficiency, and liver disease. Heterogeneity in symptoms is not only seen in patients from different families, as siblings from the same lineage can have very different clinical presentations. As mentioned before, there are various CFTR mutations which can be classified into five categories, and these help to explain some of the differences in phenotypic characteristics of CF patients. However, as the previous literature illustrates there are other factors, both environmental and genetic, which can affect the course of the disease. The main aim of this project is to identify these secondary genetic factors or so called modifier genes, which are influencing the severity of the disease.

1. Analysis of the Genotypic Data for Mendelian Inconsistencies

There were 1605 individuals from 535 trios (535 patients and both of their parents) that participated in this sub-study. According to Gregor Mendel's theory of heredity, each of the parents should contribute one of their alleles to their child. Since the major purpose of this study was to investigate the relationship between polymorphisms of the candidate genes and the disease status, this inheritance pattern had to be tested in order to avoid the possibility of having Mendelian errors which might have led to misinterpretation of the results. Errors might be caused by contamination of samples during sample collection and/or genotyping, incorrect labeling of the identities of the individuals, non-paternity and random genotyping errors.

Mendelian inconsistency was detected in some of the SNPs for some of the families as indicated in the Results section, which indicated that one or more of the above suggested factors were present in the
study. Genotyping results of those families were compared with the analysis done by our colleagues in Toronto and the same conclusion was obtained. Since our colleagues in Toronto used genotyping protocols (Illumina and Luminex) different than the TaqMan assays used in this study and the same outcome was concluded, this helped to minimize the possibility of errors from genotyping due to incorrect methodology. Out of the 23 families that had been excluded completely for this study, 22 of them were also found to have Mendelian inconsistencies by our colleagues in Toronto. For example, for pedigree #9 Toronto had found 16 and 26 errors by Illumina and Luminex respectively, while we detected 7 errors by TaqMan assays. Inclusion of those families in the study would have reduced the accuracy of the results and led to weaker statistical power, therefore those families were deleted from the rest of the study.

2. Genotypes of all Participating Individuals

Genotyping of the selected SNPs in the four candidate modifier genes was done with the appropriate TaqMan assays. The genotype of some individuals could not be detected and those were indicated as "undetermined" in the tables in the Results section. One of the possible explanations for the undetermined genotypes was inadequate concentration of DNA in the source plates for amplification by PCR. This reason seemed likely since the same individuals often had an "undetermined" genotype for several of the selected SNPs. Fortunately, this problem only occurred in a small portion of all individuals.

A total of 22 SNPs in the four candidate modifier genes were selected for testing. There were three genotype groups observed for each SNP and a coherent pattern was observed in both parents and patients for the same SNP (as illustrated in Table A3-A6), and this was a further indication of the reliability of the genotyping results.

3. Genotyping Analysis of the Parental Population

In addition to using the Family Based Association Test (FBAT) program to check for Mendelian inheritance, two additional analyses
were performed on the parental population as quality control procedures.

The first approach was to test for the existence of Hardy-Weinberg Equilibrium among the parental samples. A population is said to be in Hardy-Weinberg Equilibrium if an allele frequency remains constant through successive generations. This will be the case if there are no outside driving forces acting upon the population. There are generally five assumptions for the establishment of Hardy-Weinberg Equilibrium: (1) a large population; (2) no mutation at the locus of interest; (3) no migration; (4) random mating and (5) no natural selection at the locus of interest. Comparison between the observed and the expected genotype frequencies was done by determining Hardy-Weinberg Equilibrium in the parental population. With the establishment of Hardy-Weinberg Equilibrium among parental individuals, we confirmed that the data were consistent with the above five assumptions and the study population was not subject to any major bias due to an inappropriate sampling method. In addition, the presence of Hardy-Weinberg Equilibrium suggested that the genotyping assays were accurate. Among all 22 SNPs, Hardy-Weinberg Equilibrium was found except for SNP 149 of the HMOX1 gene. As observed in both the parental and patient populations, the number of successful genotypings for SNP 149 was lower when compared with other SNPs. This suggested that the assay was not as robust as the others and might be one of the explanations for the failure of Hardy-Weinberg Equilibrium in the parental samples for this SNP. As a result, the observed genotyping results of SNP 149 should be viewed with caution. The departure from Hardy-Weinberg Equilibrium was due to an excess of heterozygotes and therefore was unlikely to be due to unidentified population stratification in the study population. This conclusion is derived from the Wahlund Effect, which was established in 1928 by the Swedish scientist, Sten Wahlund. This theory states that the number of heterozygotes is lower than expected when populations with different allele frequencies are mixed together for analysis.[43]

The second approach was to compare the allele frequency of each SNP in the parents with the values listed on the IIPGA and Seattle SNPs websites. Both websites show genotyping results of 23 Caucasian
individuals from the State of Utah. The allele frequency of each SNP on these websites was compared with the detected genotypes of the parental group in this study by the 2x2 Chi-square test, and this served as another tool to check for the reliability of the TaqMan assays for genotyping. When compared with the first method described above, a different conclusion was obtained: SNPs 1038 and 16442 of the HMOX1 gene showed a P-value of less than 0.05 indicating that there might be unidentified errors in the genotyping results of these two SNPs. However, the magnitude of the difference was small (<10% for both SNPs) and the discrepancy in allele frequency between the participating parents and the Caucasian samples listed by the Seattle SNPs website might be explained by two reasons. Firstly, only 23 Caucasian samples were included in these the Seattle SNPs website. This small sample size would reduce the accuracy of the calculated allele frequency using these data. It would be ideal if there were other research centers that had done the genotyping of these 2 SNPs so another comparison would be possible. However, only Seattle SNPs contains data for SNP HMOX_1038. For the other SNP HMOX1_16442, an additional 60 DNA samples from another European population were genotyped by Seattle SNPs and the allele frequency was found to be 0.92 and 0.08 for alleles A and T, respectively. These values are closer to the ones found for the samples in this study, which favored the accuracy of the genotyping data. Secondly, Caucasian is a general racial classification and people in this group could possess different genetic backgrounds. Therefore, there could be genuine allele frequency differences between our study group and the individuals genotyped for the Seattle SNPs project. For these reasons the results from both SNP 1038 and 16442 of the HMOX1 gene were considered acceptable.

4. ANOVA, Regression Analysis and FBAT

Two approaches were employed for testing the existence of any correlation between the 22 selected SNPs and the 3 phenotypic traits. Results from these two methods were then compared to reinforce the conclusions.

The first method was the one-way ANOVA. The dependent variables, i.e. the selected phenotypic characteristics, in an ANOVA must be
normally distributed before any statement can be deduced. A goodness of fit test was used to check for the normality of the data, and it was found that only the dataset of age of diagnosis was skewed. Therefore, age of diagnosis was transformed by taking the logarithm to the base 10 in order to achieve a normally distributed variable. ANOVA of 10 pairs of genotype-phenotype comparisons were found to have a P value of less than 0.05, which suggested the existence of an association. These 10 pairs were: BF_2557 with FEV1 predicted value and standard deviation, BF_7202 with age of diagnosis, FEV1 predicted value and standard deviation, TLR4_1859 with age of diagnosis, HMOX1_2790 with FEV1 predicted value and standard deviation; and lastly, HMOX1_9531 with FEV1 predicted value and standard deviation. However, as FEV1 predicted value and standard deviation were highly correlated ($r^2=0.64$) associations with both variables were to be expected and could not be considered as independent observations. Further investigation was done to confirm these significant associations.

Regression analysis was performed for those 10 SNP-phenotype pairs, with inclusion of age, sex and CFTR genotype as the confounding factors. Those factors were chosen since they might contribute to the observed phenotypes in addition to the tested SNP. For example, it is known that FEV1 percent predicted values decline with age in CF patients. Therefore, it was essential to take into consideration those confounding factors before a conclusion could be finalized. Sex had no confounding effect on the observed phenotypes as indicated in the tables in the Results section. CFTR mutation contributed to part of the association for BF_7202 when tested with age of diagnosis and for TLR4_1859 when tested with age of diagnosis. CFTR genotype was related to the age of diagnosis since it was observed that individuals were diagnosed at an older age when they had a milder form of CFTR mutation. Of course, this is one of the reasons that a CFTR mutation is classified as mild. Age had a significant effect on FEV1 percent predicted value, but not on the FEV1 standard deviation values. This could be explained by the fact that FEV1 is a measurement of lung function that is derived from a comparison to a random sample of the healthy population. Therefore, as CF progresses this is reflected by a decrease in the FEV1 percent predicted. However, such decrease would not be detected in the FEV1 standard deviation values because the FEV1
standard deviation value collected for this study was recorded as the difference from the average of CF patients in same age group. Among the 10 pairs of genotype-phenotype associations, only (1) BF_2557 and FEV1 standard deviation value, (2) BF_7202 and FEV1 standard deviation value, (3) HMOX1_2790 and FEV1 standard deviation value and (4) HMOX1_9531 and FEV1 standard deviation value demonstrated a significant association after adjustment for the confounding factors. Such multiple tests may generate significant results by chance. This is mitigated to some degree by the fact that the outcome variables were related instead of independent in the multiple tests. Nevertheless, the results of this study should be considered as hypothesis generating and need to be confirmed by repeating the experiment with a larger group of samples.

FBAT was another approach used in this study for determining the existence of a correlation between selected SNPs and phenotypes. Two models (additive and dominant) were included in the FBAT analyses. Under the additive model, 7 pairs of genotype-phenotype associations were found to be significant. They were: (1) BF_2557 with FEV1 predicted value, (2) BF_4022 with FEV1 standard deviation value, (3) BF_7202 with FEV1 standard deviation value, (4) C3_28795 with age of diagnosis, (5) C3_43118 with age of diagnosis, (6) TLR4_15884 with FEV1 predicted value and (7) HMOX1_1038 with FEV1 predicted value. On the other hand, there were 12 pairs of genotype-phenotype associations determined to be significant under the dominant model. They were (1) BF_7202 with age of diagnosis and FEV1 standard deviation, (2) BF_2557 with FEV1 predicted value and FEV1 standard deviation, (3) BF_4022 with FEV1 predicted value and standard deviation, (4) BF_6484 with FEV1 standard deviation, (5) C3_28795 with age of diagnosis and FEV1 predicted value, (6) TLR4_15884 with FEV1 predicted value, (7) HMOX1_1038 with FEV1 predicted value and (8) HMOX1_9531 with FEV1 standard deviation. Thus, there was a slight difference in the conclusions between the additive and dominant models. This was likely due to the different underlying assumptions of each model. The dominant model assumes that one of the alleles is dominant over the other. In other words, it was assumed that phenotypic trait of the patients with either the dominant homozygous genotype or the heterozygous genotype was significantly different from those with recessive homozygous genotype. Under the additive model it was assumed that the phenotypic
trait of heterozygous patients was intermediate between the two groups of homozygous patients. Both models were fitted to the data as there was no a priori information regarding the mode of inheritance for any of the selected polymorphisms.

FBAT analysis involved families with parents who had a heterozygous genotype since this test analyzed the inheritance pattern from parents to the affected individual. Although the theory, calculation and criteria for both ANOVA and FBAT were different, it was rational to expect that the results from FBAT should overlap with those obtained from ANOVA and lead to a consistent conclusion. However, only a few matches were observed when the results of the two methods were compared. For some SNPs, a significant association was detected by ANOVA but not by FBAT. There are several possible reasons that could explain this inconsistency. Firstly, a false positive might have resulted due to the recruitment of patients of different ethnicity. Although all patients were self-reported as Caucasians, they might have had subtly different genetic backgrounds. In another words, there might have been unrecognized subgroups in the study, for example, patients with ancestry in England might be genetically dissimilar to those from Quebec. Different ethnic subgroups could show different allele frequencies and therefore this population stratification might lead to a false positive conclusion in the ANOVA. It might be inappropriate to compare results from ANOVA with results from FBAT since they were different with regard to the groups of individuals being analyzed. ANOVA analyzed all individuals recruited in this study while FBAT excluded families with homozygous parents. Lastly, the significant relationship seen in ANOVA may have resulted by chance.

On the other hand, significant associations were detected by FBAT for some other SNPs, which were not significant in the ANOVA. FBAT was used to analyze the inheritance pattern of alleles of patients from their parents. There should not be a problem of false positives when considering a sample population with mixed ethnicity, because the genetic background does not affect the inheritance pattern. In addition, the mean values were compared in the ANOVA in order to determine the presence of any association between the SNPs and phenotypic characteristics. In this type of analysis, it is optimal if
the number of individuals in each group is identical. However, this was not the case for most of the SNPs. Therefore, the ANOVA might have been underpowered to detect an association. Furthermore, some important information might have been missed by only looking at the mean values. For example, two groups might have the same mean but one group might have had the patients clustered around the mean values and the other group might have included some patients who had extreme values.

5. Age of Onset Analysis

*Pseudomonas aeruginosa* is one of the most frequent pathogens found in lungs of CF patients. It is common to see that some patients get colonized in their early childhood while others are not infected with this bacterium until adulthood. However, testing for association between the selected SNPs and age of patients when they first get colonized with *Pseudomonas aeruginosa* requires a different type of analysis than those described above. Some of the recruited individuals had never been infected with *Pseudomonas aeruginosa*. ANOVA was inappropriate to use for this type of censored data and therefore this analysis was performed by age of onset analysis or so called survival analysis. In general, this test includes two phases: (1) The Kaplan-Meir technique is used to analyze the time to an endpoint (in this case first infection with *Pseudomonas aeruginosa*); and 2) The Log Rank test is then used to compare two or more Kaplan-Meier "curves" (in this case the curves for each genotype). By examining all the 22 SNPs by this survival analysis, only TLR4_9263 was revealed to be significantly associated with the age of first *Pseudomonas aeruginosa* infection. In this survival analysis the data were censored as some individuals had not experienced their first *Pseudomonas aeruginosa* infection at the time of the data collection. Therefore, the current age was entered for patients recorded as no *Pseudomonas aeruginosa* infection so that they would contribute to the analysis. However, inclusion of this group of patients might cause misinterpretation of the data since they might have had one or two colonizations many years ago and the documentation could not be traced. Additionally, this group of patients might get infected in the future so it would be inappropriate to group them as no *Pseudomonas aeruginosa* infection and make a conclusion that their
genotype was not linked to the *Pseudomonas aeruginosa* infection at the time of recruitment.

6. *Pseudomonas aeruginosa* Infection Status

As described in the Introduction section, colonization with *Pseudomonas aeruginosa* occurs in about 80% of CF patients and infection by this pathogen greatly deteriorates the health status of the patient. In this study, all the participating patients were classified into one of the four groups according to their exposure to *Pseudomonas aeruginosa*. A chi-square test was performed in order to determine if one or more SNPs contributed to susceptibility to *Pseudomonas aeruginosa* infection in CF patients. Among the 22 selected SNPs of the four candidate modifier genes, only C3_963 illustrated a borderline P-value which suggested that it might play a role in the patients' ability to eradicate *Pseudomonas aeruginosa* infection. Thus, no informative conclusions could be drawn in this section. Usually, microbial testing of patients' sputum is one of the routine tests in CF clinics and there should be minimal error in the testing protocol for the determination of *Pseudomonas aeruginosa* infection status for CF patients. However, errors might have arisen from the procedures in data collection. It was time-consuming and complicated to look for relevant data if the patients had only one or a few *Pseudomonas aeruginosa* infections in the past. This is especially true for patients who had been transferred from different clinics and those who were seriously ill and had voluminous clinic/hospital charts. For example, it would be easy to misclassify an adult patient as group 0 with no *Pseudomonas aeruginosa* infection if no related information could be found in his/her charts. However, the patient might have been colonized once many years ago in childhood and there was no way to trace back to the previous records. In addition, the number of patients in one of the four categories was small for most SNPs and such a small sample size might reduce the power of the analysis.
7. Haplotype Analysis Utilizing the RGui and FBAT programs

The selected SNPs in each candidate modifier genes were grouped together as a haplotype in the Haplotype analysis. There were two main reasons for performing this investigation: (1) to test for any interaction between SNPs which might result in the observed phenotype even though no association could be found when considering SNPs separately, and (2) to see if there was another unidentified SNP in the vicinity of the selected ones which might have contributed to the phenotypic trait and was in linkage disequilibrium with the tested SNPs. Haplotype analysis was done using two programs: RGui and FBAT, and the results were then compared for consistency in order to strengthen the conclusion of an association between the haplotype and the phenotypic trait. The major difference between these two programs in the haplotype analysis was similar to that between ANOVA and FBAT in the single locus analysis: the RGui program included all participants in the study while the FBAT program looked for the inheritance pattern and only considered those trios with heterozygous parents.

(a) Haplotype analysis utilizing the RGui program

In the haplotype analysis done utilizing the RGui program, only two haplotypes in Factor B, h11111 and h22112, were found to be significantly associated with the age of diagnosis. By visual inspection of the two haplotypes, it was apparent that they shared the same alleles at SNPs 6484 and 7202. By examining the results table it could be observed that only three haplotypes were analyzed individually while the rest (with lower frequencies) were grouped together as a pool. The third individual haplotype (with a P-value of greater than 0.05) did not have the same alleles at SNP 6484 and 7202 as the two significant haplotypes. It was possible that SNPs 6484 and 7202 were interacting and affecting the age of diagnosis. However, this result might be a false positive as two additional haplotypes in the pooled group also had the same alleles at SNP 6484 and 7202 but an association with age of diagnosis was not apparent for these two haplotypes. In addition, it was more likely that SNP 7202 was the cause for the significant relationship since it was found to be significant by both ANOVA and FBAT when the SNPs were considered independently.
An association was only found for one of the 16 haplotypes of the Complement factor 3 gene, i.e., h1212. Comparison of this haplotype with the others, suggested that the first two SNPs, SNP 963 and 28795, might be interacting and thus responsible for the significant result. C3_28795 was observed to be significant in the single locus FBAT analysis but none of the SNPs in Complement factor 3 were significant in the ANOVA. Thus, it may be that C3_28795 was the true driving force in the detected haplotype result.

Even though TLR4_1859 and TLR4_15884 were found to be significantly associated with age of diagnosis and FEV1 predicted value respectively, haplotypes consisting of these two SNPs showed no significant association in the analysis. Only the pooled group was shown to have a P value of less than 0.05 and thus no specific haplotype could be concluded to be responsible.

In the last candidate gene Heme oxygenase-1, one haplotype (h112221) and two haplotypes (h112221 and h211221) were discovered to be associated with the FEV1 predicted value and the FEV1 standard deviation, respectively. Therefore, the combination of the last three SNPs of Heme oxygenase-1, SNP 3303, 9531 and 16442, might have been responsible for the observed associations. However, it was interesting to note that another haplotype with the same alleles at those three positions did not have a P-value of less than 0.05. Therefore, it might be that another unobserved SNP in the HMOX1 gene found on the h112221 and h211221 backgrounds was responsible for the observed association.

(b) Haplotype analysis utilizing the FBAT program

More haplotypes were found to be significantly associated with the phenotypes by FBAT analysis. In Factor B, haplotype h11111 was found to be associated with both FEV1 predicted value and standard deviation, while another haplotype h22121 was observed to be strongly associated with the FEV1 standard deviation value. When comparing all the haplotypes seen in Factor B, no consistent pattern was observed in the haplotypes which had a P value of less than 0.05. In the single locus analysis of the SNPs in factor B by FBAT, almost all of them were
found to be related to FEV1 predicted value, FEV1 standard deviation or both. This might be a reason for the lack of a consistent pattern of Factor B haplotypes associated with these phenotypic traits. It is possible that several Factor B SNPs independently influence the level of FEV1.

In Complement factor 3, haplotypes h1111 and h1212 were discovered to be significantly associated with age of diagnosis and FEV1 standard deviation value, respectively. When comparing the results of individual SNPs and haplotypes, it was difficult to discern a consistent pattern. Sixteen haplotypes were detected and each had a sample size sufficient for individual haplotype analysis. Such a large number of haplotypes made it difficult to search for any interaction among SNPs and the relationship to the phenotypic traits.

Twenty-three haplotypes were seen in Toll-like receptor 4 in the patients and only 9 of them were used for the analysis since the frequency of the rest was too low. One of the haplotypes, h1122112, was shown to have a P value of less than 0.05 for its correlation with FEV1 predicted value. The only difference in this haplotype compared with the remaining 8 lay in the sixth SNP, which was SNP 15884. This conclusion was consistent with the single SNP FBAT analysis, and suggested that there was no interaction of SNPs or unidentified SNPs contributing to the phenotypic trait.

A similar observation was made in the heme oxygenase-1 gene analysis. Among the 21 haplotypes found in this candidate gene, only 7 of them were frequent enough for analysis. One haplotype, h221211, was determined to be associated with the FEV1 % predicted value. The only noticeable difference between this haplotype and the rest was in the second SNP, SNP 1038. This SNP was significantly associated with the FEV1 predicted value in the single locus FBAT analysis and thus was likely to be the driving factor for the haplotype association seen here.

There were some differences between the results obtained by the two programs. Firstly, the possible haplotypes determined in the patients were slightly different by the two programs due to the
selection procedure of participants eligible for analysis. Even if the same haplotype was found in both programs, the haplotype frequency was not the same due to the difference in the number of patients in each analysis. Secondly, the RGui program grouped the haplotypes with low frequency together while the FBAT program ignored the haplotypes with less than 10 families. Obviously, some information would be lost in both cases. On the other hand, a similar pattern was observed in the results obtained by both types of analysis. For most of the haplotypes that were found to have a significant association with an outcome variable, it appeared that the driving factor behind the detected association was due to a single SNP, not the interaction of SNPs. This was because the haplotype results were consistent with the single SNP analyses by both ANOVA and FBAT.

8. Haplotype Analysis of the Age of First *Pseudomonas aeruginosa* Infection utilizing Hapstat

Haplotype analysis was also performed for investigation of the age of first *Pseudomonas aeruginosa* infection, and this was done using the Hapstat program. The same concerns applied to this analysis, with respect to the difficulty of recording the age of first *Pseudomonas aeruginosa* infection accurately, as for the single SNP analysis.

Although TLR4_9263 was concluded to be associated with the age of first *Pseudomonas aeruginosa* infection, none of the haplotypes in the four candidate genes were found to have a significant relationship with this outcome. As illustrated in the Results section, the number of possible haplotypes analyzed by the program was less than the haplotype analysis done by both RGui and FBAT, due to the fact that the program was set to ignore the haplotypes with a frequency of less than 0.05. Inevitably this decreased the number of participants and therefore the power of this part of the study.
9. Position of SNPs with Significant Association and their Effect on Gene Function

Among the 22 selected SNPs, 12 of them were found to be significantly linked to one of the tested phenotypic traits and the results are summarized in Tables 55-57.
Table 55: Summary of all significant SNP-phenotype associations by ANOVA and FBAT analyses

<table>
<thead>
<tr>
<th>Analytical Test</th>
<th>Gene</th>
<th>Polymorphism</th>
<th>Phenotypic trait</th>
<th>Position in the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Factor B</td>
<td>BF_2557</td>
<td>FEV1 predicted value</td>
<td>Exon</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BF_7202</td>
<td>Age of Diagnosis</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 predicted value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heme oxygenase-1</td>
<td>HMOX1_2790</td>
<td>FEV1 predicted value</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HMOX1_9531</td>
<td>FEV1 predicted value</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Factor B</td>
<td>BF_2557</td>
<td>FEV1 predicted value</td>
<td>Exon</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BF_4022</td>
<td>FEV1 predicted value</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BF_6484</td>
<td>FEV1 standard deviation value</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BF_7202</td>
<td>Age of Diagnosis</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td>Complement Factor 3</td>
<td>C3_28795</td>
<td>Age of Diagnosis</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FEV1 predicted value</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C3_43118</td>
<td>Age of Diagnosis</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td>Toll-like Receptor 4</td>
<td>TLR4_15844</td>
<td>FEV1 predicted value</td>
<td>Exon</td>
</tr>
<tr>
<td></td>
<td>Heme oxygenase-1</td>
<td>HMOX1_1038</td>
<td>FEV1 predicted value</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HMOX1_9531</td>
<td>FEV1 standard deviation value</td>
<td></td>
</tr>
</tbody>
</table>
Table 56: Summary of all significant haplotype-phenotype associations when testing by RGui and FBAT analyses

<table>
<thead>
<tr>
<th>Analytical Test</th>
<th>Gene</th>
<th>Haplotype</th>
<th>Phenotypic trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGui</td>
<td>Factor B</td>
<td>h11111</td>
<td>Age of Diagnosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h22112</td>
<td>Age of Diagnosis</td>
</tr>
<tr>
<td></td>
<td>Complement Factor 3</td>
<td>h1212</td>
<td>Age of Diagnosis</td>
</tr>
<tr>
<td></td>
<td>Heme oxygenase-1</td>
<td>h221211</td>
<td>Age of Diagnosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h112221</td>
<td>FEV1 predicted value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h112221</td>
<td>FEV1 standard deviation value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h211221</td>
<td>FEV1 standard deviation value</td>
</tr>
<tr>
<td>FBAT</td>
<td>Factor B</td>
<td>h11111</td>
<td>FEV1 predicted value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h11111</td>
<td>FEV1 standard deviation value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h22121</td>
<td>FEV1 standard deviation value</td>
</tr>
<tr>
<td></td>
<td>Complement Factor 3</td>
<td>h22111</td>
<td>Age of Diagnosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h11112</td>
<td>FEV1 predicted value</td>
</tr>
<tr>
<td></td>
<td></td>
<td>h1212</td>
<td>FEV1 standard deviation value</td>
</tr>
<tr>
<td></td>
<td>Toll-like Receptor 4</td>
<td>h1122112</td>
<td>FEV1 predicted value</td>
</tr>
<tr>
<td></td>
<td>Heme oxygenase-1</td>
<td>h221211</td>
<td>FEV1 predicted value</td>
</tr>
</tbody>
</table>
As indicated in the above table, those 13 SNPs which demonstrated a significant relationship with one of the phenotypic traits could be divided into two main groups according to their position in the gene: intron or exon. Most of the SNPs are situated in the intronic regions. Although introns are spliced out and therefore not translated into amino acids, polymorphisms in this region could also affect the protein products. It is common to find the nucleotides C and T at the beginning of an intron, whereas nucleotides A and G are found at the end of it. ([44]) However, there are other sequences that control the exact position of where to cut and the process of splicing. ([45]) This signal is included in the intron, therefore splicing error might occur if there was a mutation in the intron and this might influence the function of the gene. In addition, two of the SNPs, C3_963 and HMOX1_1038 are in a promoter region, which might affect the initiation step of transcription. On the other hand, two SNPs are in an exonic region, i.e., BF_2557 and TLR4_15844. For example, BF_2557 is located in exon 3 of the Factor B gene, however, it changes the third base of the codon from G to A without altering the amino acid produced. The position of this SNP is near the end of exon 3 which may be an exonic splicing enhancer. That is, it is in the region of the exon which promotes the cutting of the adjacent intron and therefore might affect the structure and function of the protein products being formed. This is similar to the C3_43118 SNP, which is a tagSNP of C3_44692 in this study since

<table>
<thead>
<tr>
<th>Analytical Test</th>
<th>Gene</th>
<th>Polymorphism / Haplotype</th>
<th>Phenotypic trait</th>
<th>Position in the gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survival analysis</td>
<td>Complement Factor 3</td>
<td>C3_963</td>
<td>Age of first <em>Pseudomonas aeruginosa</em> infection</td>
<td>Intron</td>
</tr>
<tr>
<td></td>
<td>Toll-like Receptor 4</td>
<td>TLR4_9263</td>
<td>Age of first <em>Pseudomonas aeruginosa</em> infection</td>
<td>Intron</td>
</tr>
<tr>
<td>Chi squared test</td>
<td>Complement Factor 3</td>
<td>C3_963</td>
<td><em>Pseudomonas aeruginosa</em> infection status</td>
<td>Intron</td>
</tr>
</tbody>
</table>

Table 57: Summary table of SNPs which revealed a significant association with the *Pseudomonas aeruginosa* infection status and the age of first *Pseudomonas aeruginosa* infection.
C3_44692 is located in the beginning region of exon 41 of the Complement factor 3 gene.

The large number of statistical tests performed in this study posed a potential problem in the analysis. A P value of less than 0.05 was reported to be significant in this project. However, such an observation might have been due to chance rather than reflecting a true association. In order to have equal statistical power as the case when only one hypothesis is being investigated, the Bonferroni correction procedure can be used where the statistical significance level is multiplied by the number of tests done when independent hypotheses are being examined for the same dataset. [46] Therefore, Bonferroni correction might be a solution for solving such potential risk, however, it was not used in this project as the correction factor is overly conservative since the phenotypic outcomes were correlated with each other and some of the SNPs showed low levels of linkage disequilibrium i.e. the hypotheses being tested were not independent. [47]

One way to assess the potential functional consequences of a SNP is to determine the degree of conservation of the surrounding sequence between species. Among the 22 SNPs in the four candidate genes, a few of them were conserved among 17 vertebrate species [48] and these are summarized in table 58:

Table 58: Summary table of the conservation score of the selected SNPs in the four candidate genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Conservation Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td>2557</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>6484</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>11912</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>17050</td>
<td>0.120</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>3303</td>
<td>0.269</td>
</tr>
</tbody>
</table>

A conservation score from 0 (not conserved) to 1 (highly conserved) was assigned when comparing the sequence within the 17 vertebrate species. This conservation score can be interpreted as the probability that any given base is in a conserved element. The SNPs in the above table were found to be conserved among that 17 vertebrate
species, which meant that they remained the same through evolution. In another words, they might have an important function in the gene. It is important if the above listed SNPs were found to be significant with one or more of the tested phenotypes, since it further strengthened the conclusion that those SNPs affected the gene function. For example, BF_2557 was shown to be highly conserved among those 17 vertebrate species and therefore, it is likely to be functionally important. This conclusion is interesting because BF_2557 was one of the SNPs which was shown to have an association with the tested phenotypes.

10. Summary

In summary, 23 families were deleted from the study due to the presence of Mendelian errors in them. With the remaining 1605 individuals or 535 trios, only 13 of the selected 22 SNPs were found to have a significant association by either ANOVA or FBAT with one or more of the four selected phenotypic traits (age of diagnosis, FEV1 predicted and standard deviation value, age of first Pseudomonas aeruginosa infection). Also only one of the SNPs was determined to have a borderline P value when testing for the Pseudomonas aeruginosa infection status among the participants. In the haplotype analyses done by RGui or FBAT, some haplotypes were found to have a P value of less than 0.05. Most of these haplotypes contained a SNP which was found to have a significant relationship with the phenotype by either ANOVA or FBAT, and such inclusion of this SNP would probably be the driving force behind the observed result.

11. Future studies

This is a sub-study of a large, Canada-wide and international research project on CF modifier genes. There were a total of 1605 patients from 535 trios that participated in this part of the study. The recruitment of patients was ongoing at the time when this sub-study started. Therefore, the trends that we observed can be investigated in additional trio members who were consented later in the project and families with more than one affected child. This is particularly critical for some of the selected SNPs where borderline P values were obtained. If the same result is found in a larger dataset then such a
consistent pattern would be more convincing. Such a replication study is ongoing.

Replication of these results could also be sought in the two large US CF modifier gene consortia. The other two cohorts of patients are from the University of North Carolina/Case Western Reserve University and Johns Hopkins University.

For those SNPs that were found to be significantly associated with one or more of the tested phenotypes, it is possible that they affect the gene function and therefore lead to different severity of the disease. This hypothesis could be confirmed by performing further investigation. For example, if the SNP is in the promoter region of the DNA molecule, it is in a region that serves as a control point for regulating transcription. Therefore the amount of mRNA and protein produced would be affected, which could be measured by Real-time PCR and ELISA, respectively. If the SNP is in the intronic region, it may affect the splicing procedure and this could be tested by either sequencing or measuring the length of the transcript by amplifying the cDNA sequence.

Although the frequency of *Burkholderia cepacia* infection of CF patients is lower than that of *Pseudomonas aeruginosa*, it is even more difficult to eradicate the pathogen once infection is established. An experiment could be done utilizing the SNPs shown to be associated with disease status severity in this project in order to determine the relationship between these SNPs and *Burkholderia cepacia* infection. For example, the genes that are responsible for the attachment of the pathogen in lungs (TLR4 would be a possible choice since *Burkholderia cepacia* is a Gram-negative bacteria and its LPS also attaches to TLR4 as is the case for *Pseudomonas aeruginosa*) and other genes in the immune system for clearing the micro-organism.

Finally, it is possible to perform a genome-wide association analysis with 300,000 - 500,000 SNPs being analyzed in a single experiment for each study participant. This can be done by recently developed high-throughput genotyping technologies such as the Illumina or Affymetrix systems. With appropriate adjustments for multiple
comparisons and the use of several populations to provide replication; this approach provides a systematic assessment of the entire human genome.
References


27. Mueller-Ortiz, S.L., S.M. Drouin, and R.A. Wetsel, The alternative activation pathway and complement component C3 are critical for a protective immune response against Pseudomonas


Appendix

Table A1: Concentration of a subset of the DNA samples in the original source plates #1-8

<table>
<thead>
<tr>
<th>Individual ID</th>
<th>DNA concentration (ng/µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0003-01</td>
<td>4.43</td>
</tr>
<tr>
<td>0003-02</td>
<td>3.93</td>
</tr>
<tr>
<td>0003-03</td>
<td>1.13</td>
</tr>
<tr>
<td>0019-01</td>
<td>2.52</td>
</tr>
<tr>
<td>0019-02</td>
<td>1.16</td>
</tr>
<tr>
<td>0019-03</td>
<td>0.25</td>
</tr>
<tr>
<td>0052-01</td>
<td>1.95</td>
</tr>
<tr>
<td>0052-02</td>
<td>0.87</td>
</tr>
<tr>
<td>0052-03</td>
<td>0.72</td>
</tr>
<tr>
<td>0063-01</td>
<td>1.94</td>
</tr>
<tr>
<td>0063-02</td>
<td>0.93</td>
</tr>
<tr>
<td>0063-03</td>
<td>0.46</td>
</tr>
<tr>
<td>0069-01</td>
<td>2.54</td>
</tr>
<tr>
<td>0069-02</td>
<td>1.60</td>
</tr>
<tr>
<td>0069-03</td>
<td>-0.03</td>
</tr>
<tr>
<td>0083-01</td>
<td>2.63</td>
</tr>
<tr>
<td>0083-02</td>
<td>1.47</td>
</tr>
<tr>
<td>0083-03</td>
<td>0.55</td>
</tr>
<tr>
<td>0101-01</td>
<td>3.17</td>
</tr>
<tr>
<td>0101-02</td>
<td>2.47</td>
</tr>
<tr>
<td>0101-03</td>
<td>1.29</td>
</tr>
<tr>
<td>0150-01</td>
<td>3.40</td>
</tr>
<tr>
<td>0150-02</td>
<td>9.77</td>
</tr>
<tr>
<td>0150-03</td>
<td>2.75</td>
</tr>
</tbody>
</table>
Table A2: Families with non-Mendelian inheritance when genotyping a particular SNP

<table>
<thead>
<tr>
<th>Family ID</th>
<th>SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0167</td>
<td>HMOX1_2790</td>
</tr>
<tr>
<td>0313</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>0374</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>0445</td>
<td>BF_2557</td>
</tr>
<tr>
<td>0451</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>0453</td>
<td>TLR4_11912</td>
</tr>
<tr>
<td>0478</td>
<td>BF_2557</td>
</tr>
<tr>
<td>0482</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1251</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1300</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1324</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1350</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1383</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1431</td>
<td>C3_963</td>
</tr>
<tr>
<td>1444</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1878</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>1881</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>2543</td>
<td>C3_963</td>
</tr>
<tr>
<td>2800</td>
<td>C3_963</td>
</tr>
<tr>
<td>2809</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>2821</td>
<td>HMOX1_149</td>
</tr>
<tr>
<td>4120</td>
<td>HMOX1_149</td>
</tr>
</tbody>
</table>
Table A3: Genotype frequency of each of the SNPs examined in the gene of Factor B

<table>
<thead>
<tr>
<th>SNP</th>
<th>Group</th>
<th>Total # of Individuals</th>
<th># of Individuals</th>
<th>Genotype</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557</td>
<td>Parents</td>
<td>1066</td>
<td>36</td>
<td>AA</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>314</td>
<td>AG</td>
<td>29.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>705</td>
<td>GG</td>
<td>66.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>undetermined</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>533</td>
<td>20</td>
<td>AA</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>135</td>
<td>AG</td>
<td>25.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>367</td>
<td>GG</td>
<td>68.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>undetermined</td>
<td>2.1</td>
</tr>
<tr>
<td>BF_4022</td>
<td>Parents</td>
<td>1070</td>
<td>107</td>
<td>AA</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>458</td>
<td>AG</td>
<td>42.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>489</td>
<td>GG</td>
<td>45.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>undetermined</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>67</td>
<td>AA</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>201</td>
<td>AG</td>
<td>37.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>259</td>
<td>GG</td>
<td>48.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>undetermined</td>
<td>1.5</td>
</tr>
<tr>
<td>BF_6484</td>
<td>Parents</td>
<td>1070</td>
<td>858</td>
<td>AA</td>
<td>80.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>189</td>
<td>AG</td>
<td>17.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td>GG</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>undetermined</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>423</td>
<td>AA</td>
<td>79.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>AG</td>
<td>18.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>GG</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>undetermined</td>
<td>0.8</td>
</tr>
<tr>
<td>BF_7202</td>
<td>Parents</td>
<td>1070</td>
<td>197</td>
<td>AA</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>508</td>
<td>AG</td>
<td>47.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>344</td>
<td>GG</td>
<td>32.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>21</td>
<td>undetermined</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>105</td>
<td>AA</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>252</td>
<td>AG</td>
<td>47.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>172</td>
<td>GG</td>
<td>32.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>undetermined</td>
<td>1.1</td>
</tr>
<tr>
<td>BF_8311</td>
<td>Parents</td>
<td>1070</td>
<td>862</td>
<td>CC</td>
<td>80.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>182</td>
<td>CT</td>
<td>17.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>TT</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td>undetermined</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>428</td>
<td>CC</td>
<td>80.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90</td>
<td>CT</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>TT</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>undetermined</td>
<td>2.2</td>
</tr>
</tbody>
</table>
Table A4: Genotype frequency of each of the SNPs examined in the gene of Complement factor 3

<table>
<thead>
<tr>
<th>SNP</th>
<th>Group</th>
<th>Total # of Individuals</th>
<th># of Individuals</th>
<th>Genotype</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>Parents</td>
<td>1064</td>
<td>270</td>
<td>GG</td>
<td>25.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>527</td>
<td>GT</td>
<td>49.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>249</td>
<td>TT</td>
<td>23.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>undetermined</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>532</td>
<td>152</td>
<td>GG</td>
<td>28.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>255</td>
<td>GT</td>
<td>47.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>119</td>
<td>TT</td>
<td>22.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>undetermined</td>
<td>1.1</td>
</tr>
<tr>
<td>C3_28795</td>
<td>Parents</td>
<td>1068</td>
<td>355</td>
<td>AA</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>521</td>
<td>AG</td>
<td>48.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>177</td>
<td>GG</td>
<td>16.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>undetermined</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>534</td>
<td>202</td>
<td>AA</td>
<td>37.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>237</td>
<td>AG</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>92</td>
<td>GG</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>undetermined</td>
<td>0.6</td>
</tr>
<tr>
<td>C3_36735</td>
<td>Parents</td>
<td>1070</td>
<td>202</td>
<td>AA</td>
<td>18.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>519</td>
<td>AG</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>330</td>
<td>GG</td>
<td>30.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19</td>
<td>undetermined</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>96</td>
<td>AA</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>266</td>
<td>AG</td>
<td>49.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>163</td>
<td>GG</td>
<td>30.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>undetermined</td>
<td>1.9</td>
</tr>
<tr>
<td>C3_43118</td>
<td>Parents</td>
<td>1070</td>
<td>239</td>
<td>AA</td>
<td>22.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>529</td>
<td>AG</td>
<td>49.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>292</td>
<td>GG</td>
<td>27.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>undetermined</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>123</td>
<td>AA</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>251</td>
<td>AG</td>
<td>46.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>157</td>
<td>GG</td>
<td>29.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>undetermined</td>
<td>0.8</td>
</tr>
</tbody>
</table>
Table A5: Genotype frequency of each of the SNPs examined in the gene of Toll-like receptor 4

<table>
<thead>
<tr>
<th>SNP</th>
<th>Group</th>
<th>Total # of Individuals</th>
<th># of Individuals</th>
<th>Genotype</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR4_851</td>
<td>Parents</td>
<td>1070</td>
<td>551 AA</td>
<td>51.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>434 AG</td>
<td>40.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>81 GG</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 undetermined</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>285 AA</td>
<td>53.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>203 AG</td>
<td>37.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>38 GG</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9 undetermined</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>Parents</td>
<td>1070</td>
<td>143 AA</td>
<td>13.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>509 AG</td>
<td>47.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>410 GG</td>
<td>38.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 undetermined</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>77 AA</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>244 AG</td>
<td>45.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>209 GG</td>
<td>39.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 undetermined</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>Parents</td>
<td>1070</td>
<td>24 CC</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>266 CT</td>
<td>24.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>770 TT</td>
<td>72.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 undetermined</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>10 CC</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>140 CT</td>
<td>26.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>382 TT</td>
<td>71.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 undetermined</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>Parents</td>
<td>1070</td>
<td>16 AA</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>207 AC</td>
<td>19.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>841 CC</td>
<td>78.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 undetermined</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>8 AA</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>106 AC</td>
<td>19.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>418 CC</td>
<td>78.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 undetermined</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>Parents</td>
<td>1068</td>
<td>478 GG</td>
<td>44.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>467 GT</td>
<td>43.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>114 TT</td>
<td>10.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9 undetermined</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>534</td>
<td>234 GG</td>
<td>43.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>229 GT</td>
<td>42.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65 TT</td>
<td>12.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 undetermined</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>Parents</td>
<td>1070</td>
<td>23 CC</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>270 CG</td>
<td>25.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>764 GG</td>
<td>71.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13 undetermined</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>12 CC</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>120 CG</td>
<td>22.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>392 GG</td>
<td>73.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11 undetermined</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>
Table A5: Genotype frequency of each of the SNPs examined in the gene of Toll-like receptor 4

<table>
<thead>
<tr>
<th>TLR4_17050</th>
<th>Parents 1070</th>
<th></th>
<th></th>
<th>Patients 535</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>CC</td>
<td>1.9</td>
<td>15</td>
<td>CC</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>277</td>
<td>CT</td>
<td>25.9</td>
<td>132</td>
<td>CT</td>
<td>24.7</td>
</tr>
<tr>
<td></td>
<td>765</td>
<td>TT</td>
<td>71.5</td>
<td>384</td>
<td>TT</td>
<td>71.8</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>undetermined</td>
<td>0.8</td>
<td>4</td>
<td>undetermined</td>
<td>0.8</td>
</tr>
<tr>
<td>SNP</td>
<td>Group</td>
<td>Total # of Individuals</td>
<td># of Individuals</td>
<td>Genotype</td>
<td>Percentage</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>------------------------</td>
<td>------------------</td>
<td>----------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>Parents</td>
<td>1036</td>
<td>122</td>
<td>AA</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>404</td>
<td>AG</td>
<td>39.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>479</td>
<td>GG</td>
<td>46.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>31</td>
<td>undetermined</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>518</td>
<td>61</td>
<td>AA</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>208</td>
<td>AG</td>
<td>40.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>227</td>
<td>GG</td>
<td>43.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td>undetermined</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>Parents</td>
<td>1070</td>
<td>975</td>
<td>CC</td>
<td>91.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>83</td>
<td>CT</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>TT</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>undetermined</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>478</td>
<td>CC</td>
<td>89.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>51</td>
<td>CT</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>TT</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>undetermined</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>Parents</td>
<td>1068</td>
<td>341</td>
<td>AA</td>
<td>31.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>517</td>
<td>AT</td>
<td>48.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>200</td>
<td>TT</td>
<td>18.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>undetermined</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>534</td>
<td>166</td>
<td>AA</td>
<td>31.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>253</td>
<td>AT</td>
<td>47.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>108</td>
<td>TT</td>
<td>20.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td>undetermined</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>Parents</td>
<td>1070</td>
<td>2</td>
<td>CC</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>106</td>
<td>CG</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>954</td>
<td>GG</td>
<td>89.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>undetermined</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>0</td>
<td>CC</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>53</td>
<td>CG</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>476</td>
<td>GG</td>
<td>89.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>undetermined</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>Parents</td>
<td>1068</td>
<td>225</td>
<td>AA</td>
<td>21.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>528</td>
<td>AG</td>
<td>49.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>302</td>
<td>GG</td>
<td>28.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13</td>
<td>undetermined</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>534</td>
<td>108</td>
<td>AA</td>
<td>20.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>255</td>
<td>AG</td>
<td>47.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>163</td>
<td>GG</td>
<td>30.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>undetermined</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>Parents</td>
<td>1070</td>
<td>952</td>
<td>AA</td>
<td>89.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>109</td>
<td>AT</td>
<td>10.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>TT</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>undetermined</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patients</td>
<td>535</td>
<td>475</td>
<td>AA</td>
<td>88.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td>AT</td>
<td>10.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>TT</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>undetermined</td>
<td>0.8</td>
<td></td>
</tr>
</tbody>
</table>
Table A7: Phenotypic characteristics of the CF patients

<table>
<thead>
<tr>
<th>Characteristics</th>
<th># of Individuals</th>
<th>Average</th>
<th>Range</th>
<th># (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>488</td>
<td>16.2</td>
<td>1.0 - 61.1</td>
<td>na</td>
</tr>
<tr>
<td>Age at Diagnosis (years)</td>
<td>464</td>
<td>2.7</td>
<td>0 - 59.4</td>
<td>na</td>
</tr>
<tr>
<td>Age of First PA Infection (years)</td>
<td>284</td>
<td>8.5</td>
<td>0.2 - 39.3</td>
<td>na</td>
</tr>
<tr>
<td>FEV1 % Predicted</td>
<td>409</td>
<td>75.7</td>
<td>14.2 - 140.1</td>
<td>na</td>
</tr>
<tr>
<td>FEV1 SD</td>
<td>409</td>
<td>0.45</td>
<td>(-1.93) - 4.08</td>
<td>na</td>
</tr>
<tr>
<td>Sex</td>
<td>495</td>
<td>na</td>
<td>na</td>
<td>236 (47.7%) female, 259 (52.3%) male</td>
</tr>
<tr>
<td>Pancreatic Function</td>
<td>478</td>
<td>na</td>
<td>na</td>
<td>428 (89.5%) PI, 49 (10.3%) PS, 1 (0.2%) unknown</td>
</tr>
<tr>
<td>Meconium Ileus</td>
<td>473</td>
<td>na</td>
<td>na</td>
<td>380 (80.3%) no, 90 (19%) yes, 3 (0.7%) unknown</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa status</td>
<td>476</td>
<td>na</td>
<td>na</td>
<td>141 (29.6%) none, 73 (15.4%) once, 102 (21.4%) sporadic, 160 (33.6%) chronic</td>
</tr>
<tr>
<td>Genotype</td>
<td>463</td>
<td>na</td>
<td>na</td>
<td>272 (58.7%) ΔF508/ΔF508, 155 (33.4%) ΔF508/other, 36 (7.9%) other/other</td>
</tr>
</tbody>
</table>
Table A8: The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Complement factor 3. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3_963</td>
<td>GG</td>
<td>128</td>
<td>-0.28</td>
<td>0.08</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>208</td>
<td>-0.19</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>102</td>
<td>-0.18</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>AA</td>
<td>170</td>
<td>-0.14</td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>194</td>
<td>-0.25</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>80</td>
<td>-0.28</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>AA</td>
<td>79</td>
<td>-0.37</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>225</td>
<td>-0.15</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>134</td>
<td>-0.19</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>AA</td>
<td>107</td>
<td>-0.16</td>
<td>0.08</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>208</td>
<td>-0.23</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>128</td>
<td>-0.23</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

Table A9: The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Complement factor 3.

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3_963</td>
<td>GG</td>
<td>120</td>
<td>73.86</td>
<td>2.39</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>195</td>
<td>76.40</td>
<td>1.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>86</td>
<td>77.59</td>
<td>2.83</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>AA</td>
<td>161</td>
<td>77.33</td>
<td>2.07</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>175</td>
<td>74.74</td>
<td>1.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>70</td>
<td>74.72</td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>AA</td>
<td>73</td>
<td>73.00</td>
<td>69.67</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>202</td>
<td>77.89</td>
<td>1.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>126</td>
<td>75.65</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>AA</td>
<td>96</td>
<td>77.37</td>
<td>2.68</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>193</td>
<td>76.86</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>117</td>
<td>72.85</td>
<td>2.42</td>
<td></td>
</tr>
</tbody>
</table>
Table A10: The ANOVA result of examining FEV1 sd value among different genotypes of the selected SNPs in Complement factor 3

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3_963</td>
<td>GG</td>
<td>120</td>
<td>0.31</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>195</td>
<td>0.53</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>86</td>
<td>0.56</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>AA</td>
<td>161</td>
<td>0.54</td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>175</td>
<td>0.39</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>70</td>
<td>0.47</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>AA</td>
<td>73</td>
<td>0.24</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>202</td>
<td>0.52</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>126</td>
<td>0.49</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>AA</td>
<td>96</td>
<td>0.59</td>
<td>0.10</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>193</td>
<td>0.48</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>117</td>
<td>0.34</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

Table A11: The ANOVA result of examining FEV1 predicted value among different genotypes of the selected SNPs in Toll-like receptor 4

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR4_851</td>
<td>AA</td>
<td>210</td>
<td>76.04</td>
<td>1.82</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>163</td>
<td>75.96</td>
<td>2.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>30</td>
<td>71.76</td>
<td>4.81</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>AA</td>
<td>55</td>
<td>79.18</td>
<td>3.54</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>184</td>
<td>74.31</td>
<td>1.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>166</td>
<td>76.19</td>
<td>2.04</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>CC</td>
<td>7</td>
<td>66.41</td>
<td>9.92</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>106</td>
<td>77.02</td>
<td>2.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>294</td>
<td>75.54</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>AA</td>
<td>6</td>
<td>81.71</td>
<td>10.73</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>85</td>
<td>75.89</td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>316</td>
<td>75.66</td>
<td>1.48</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>GG</td>
<td>176</td>
<td>74.78</td>
<td>1.98</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>177</td>
<td>77.02</td>
<td>1.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>50</td>
<td>75.11</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>CC</td>
<td>8</td>
<td>71.57</td>
<td>9.35</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>90</td>
<td>76.16</td>
<td>2.79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>301</td>
<td>75.69</td>
<td>1.52</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>CC</td>
<td>10</td>
<td>91.25</td>
<td>8.28</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>105</td>
<td>74.90</td>
<td>2.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>291</td>
<td>75.52</td>
<td>1.54</td>
<td></td>
</tr>
</tbody>
</table>
Table A12: The ANOVA result of examining FEV1 sd value among different genotypes of the selected SNPs in Toll-like receptor 4

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLR4_851</td>
<td>AA</td>
<td>210</td>
<td>0.44</td>
<td>0.07</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>163</td>
<td>0.49</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>30</td>
<td>0.44</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>AA</td>
<td>55</td>
<td>0.57</td>
<td>0.13</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>184</td>
<td>0.43</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>166</td>
<td>0.46</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>CC</td>
<td>7</td>
<td>-0.02</td>
<td>0.36</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>106</td>
<td>0.47</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>294</td>
<td>0.47</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>AA</td>
<td>6</td>
<td>0.66</td>
<td>0.39</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>85</td>
<td>0.50</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>316</td>
<td>0.45</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>GG</td>
<td>176</td>
<td>0.47</td>
<td>0.07</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>177</td>
<td>0.51</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>50</td>
<td>0.31</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>CC</td>
<td>8</td>
<td>0.13</td>
<td>0.34</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>90</td>
<td>0.50</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>301</td>
<td>0.46</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>CC</td>
<td>10</td>
<td>1.02</td>
<td>0.30</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>105</td>
<td>0.43</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>291</td>
<td>0.45</td>
<td>0.06</td>
<td></td>
</tr>
</tbody>
</table>

Table A13: The ANOVA result of examining age of diagnosis among different genotypes of the selected SNPs in Heme oxygenase-1. The age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_149</td>
<td>AA</td>
<td>54</td>
<td>-0.20</td>
<td>0.12</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>167</td>
<td>-0.13</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>195</td>
<td>-0.30</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CC</td>
<td>398</td>
<td>-0.20</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>43</td>
<td>-0.30</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>2</td>
<td>-1.15</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>AA</td>
<td>142</td>
<td>-0.26</td>
<td>0.07</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>207</td>
<td>-0.17</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>91</td>
<td>-0.21</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CC</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>41</td>
<td>-0.36</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>401</td>
<td>-0.20</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>AA</td>
<td>96</td>
<td>-0.27</td>
<td>0.09</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>205</td>
<td>-0.15</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>138</td>
<td>-0.25</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>AA</td>
<td>403</td>
<td>-0.20</td>
<td>0.04</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>39</td>
<td>-0.34</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>1</td>
<td>0.57</td>
<td>0.86</td>
<td></td>
</tr>
</tbody>
</table>
Table A14: Age of onset analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Factor B

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Variable</th>
<th>Sub-group</th>
<th>Estimated Value</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF_2557</td>
<td>BF_2557</td>
<td>AA</td>
<td>0.098</td>
<td>0.284</td>
<td>0.322</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>-0.204</td>
<td>0.189</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.148</td>
<td>0.088</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.036</td>
<td>0.135</td>
<td>0.965</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.009</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>BF_4022</td>
<td>AA</td>
<td>0.0246</td>
<td>0.175</td>
<td>0.620</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>-0.105</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.146</td>
<td>0.087</td>
<td>0.0965</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.029</td>
<td>0.135</td>
<td>0.974</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.015</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>BF_6484</td>
<td>AA</td>
<td>-0.0445</td>
<td>0.253</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>-0.0465</td>
<td>0.273</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.156</td>
<td>0.086</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0275</td>
<td>0.133</td>
<td>0.976</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.0148</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>BF_7202</td>
<td>AA</td>
<td>-0.0684</td>
<td>0.146</td>
<td>0.859</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.058</td>
<td>0.118</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.152</td>
<td>0.087</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0465</td>
<td>0.135</td>
<td>0.942</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.00099</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>BF_8311</td>
<td>CC</td>
<td>-0.135</td>
<td>0.257</td>
<td>0.719</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CT</td>
<td>-0.24</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.149</td>
<td>0.087</td>
<td>0.0887</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0406</td>
<td>0.134</td>
<td>0.955</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ΔF508/other</td>
<td>-0.005</td>
<td>0.144</td>
<td></td>
</tr>
</tbody>
</table>
Table A15: Age of onset analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Complement factor 3

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Variable</th>
<th>Sub-group</th>
<th>Estimated Value</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3_963</td>
<td>C3_963</td>
<td>GG</td>
<td>0.26</td>
<td>0.125</td>
<td>0.0786</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GT</td>
<td>-0.177</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.16</td>
<td>0.087</td>
<td>0.0668</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>AF508/AF508</td>
<td>0.0014</td>
<td>0.137</td>
<td>0.9988</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>0.007</td>
<td>0.146</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>C3_28795</td>
<td>AA</td>
<td>-0.025</td>
<td>0.131</td>
<td>0.2087</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.216</td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.15</td>
<td>0.0869</td>
<td>0.0842</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>AF508/AF508</td>
<td>-0.0068</td>
<td>0.133</td>
<td>0.9981</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>0.0045</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>C3_36735</td>
<td>AA</td>
<td>0.0735</td>
<td>0.157</td>
<td>0.1213</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.168</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.173</td>
<td>0.0876</td>
<td>0.0495</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>AF508/AF508</td>
<td>-0.031</td>
<td>0.134</td>
<td>0.9605</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>-0.027</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>C3_43118</td>
<td>AA</td>
<td>0.112</td>
<td>0.136</td>
<td>0.2875</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AG</td>
<td>0.107</td>
<td>0.116</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.153</td>
<td>0.0866</td>
<td>0.0958</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>AF508/AF508</td>
<td>-0.03</td>
<td>0.133</td>
<td>0.4232</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AF508/other</td>
<td>-0.0128</td>
<td>0.142</td>
<td></td>
</tr>
</tbody>
</table>
Table A16: Age of onset analysis investigating association between age of first *Pseudomonas aeruginosa* infection and selected SNPs in Heme oxygenase-1

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Variable</th>
<th>Sub-group</th>
<th>Estimated Value</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMOX1_149</td>
<td>AA</td>
<td>0.0189</td>
<td>0.18</td>
<td>0.9514</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>-0.0377</td>
<td>0.131</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.132</td>
<td>0.089</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.032</td>
<td>0.135</td>
<td>0.9654</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>-0.02</td>
<td>0.146</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CC</td>
<td>3.283</td>
<td>61.03</td>
<td>0.5595</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.31</td>
<td>61.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.148</td>
<td>0.086</td>
<td>0.0861</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0262</td>
<td>0.133</td>
<td>0.9797</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>-0.0093</td>
<td>0.142</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>AA</td>
<td>0.077</td>
<td>0.131</td>
<td>0.7633</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>0.031</td>
<td>0.118</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.145</td>
<td>0.087</td>
<td>0.0966</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>0.0039</td>
<td>0.136</td>
<td>0.9996</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>0.00041</td>
<td>0.146</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CG</td>
<td>0.02</td>
<td>0.158</td>
<td>0.8987</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.151</td>
<td>0.0867</td>
<td>0.0819</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.0346</td>
<td>0.133</td>
<td>0.9663</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>-0.0095</td>
<td>0.142</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>AA</td>
<td>0.0048</td>
<td>0.143</td>
<td>0.9069</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>-0.047</td>
<td>0.118</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.162</td>
<td>0.088</td>
<td>0.0647</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.033</td>
<td>0.134</td>
<td>0.9628</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>-0.021</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>AA</td>
<td>-0.886</td>
<td>0.363</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>-0.938</td>
<td>0.396</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>Female</td>
<td>0.144</td>
<td>0.087</td>
<td>0.0986</td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/ΔF508</td>
<td>-0.03</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CFTR mutation</td>
<td>ΔF508/other</td>
<td>-0.0198</td>
<td>0.143</td>
<td>0.9688</td>
</tr>
</tbody>
</table>
Table A17: Chi squared test for investigating the relationship between different genotypes of the selected SNPs in Factor B and *Pseudomonas aeruginosa* infection status

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>PA status (in percentage)</th>
<th>Chi Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>BF 2557</td>
<td>AA</td>
<td>36.84</td>
<td>10.53</td>
<td>15.79</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>26.09</td>
<td>13.91</td>
<td>20.87</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>30.00</td>
<td>16.36</td>
<td>21.21</td>
</tr>
<tr>
<td>BF 4022</td>
<td>AA</td>
<td>30.51</td>
<td>10.71</td>
<td>25.42</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>28.25</td>
<td>12.99</td>
<td>20.34</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>29.61</td>
<td>17.60</td>
<td>21.89</td>
</tr>
<tr>
<td>BF 6484</td>
<td>AA</td>
<td>29.79</td>
<td>16.22</td>
<td>20.48</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>30.00</td>
<td>11.11</td>
<td>24.44</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>28.57</td>
<td>14.29</td>
<td>28.57</td>
</tr>
<tr>
<td>BF 7202</td>
<td>AA</td>
<td>30.43</td>
<td>11.96</td>
<td>22.83</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>29.15</td>
<td>14.35</td>
<td>19.28</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>29.03</td>
<td>18.71</td>
<td>23.87</td>
</tr>
<tr>
<td>BF 8311</td>
<td>CC</td>
<td>29.92</td>
<td>14.44</td>
<td>21.52</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>27.50</td>
<td>20.00</td>
<td>20.00</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>40.00</td>
<td>0.00</td>
<td>20.00</td>
</tr>
</tbody>
</table>
Table A18: Chi square test for investigating the relationship between different genotypes of the selected SNPs in Toll-like receptor 4 and *Pseudomonas aeruginosa* infection status

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>PA status (in percentage)</th>
<th>Chi Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>TLR4_851</td>
<td>AA</td>
<td>32.16</td>
<td>16.08</td>
<td>21.96</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>28.89</td>
<td>13.33</td>
<td>20.56</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>20.59</td>
<td>8.82</td>
<td>20.59</td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>AA</td>
<td>35.29</td>
<td>17.56</td>
<td>22.06</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>30.91</td>
<td>13.64</td>
<td>20.00</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>26.09</td>
<td>16.30</td>
<td>22.83</td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>CC</td>
<td>37.50</td>
<td>12.50</td>
<td>25.00</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>32.80</td>
<td>12.80</td>
<td>21.60</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>28.45</td>
<td>16.13</td>
<td>21.11</td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>AA</td>
<td>14.29</td>
<td>14.29</td>
<td>28.57</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>22.83</td>
<td>18.48</td>
<td>21.74</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>31.47</td>
<td>14.40</td>
<td>21.33</td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>GG</td>
<td>31.10</td>
<td>13.88</td>
<td>20.10</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>26.96</td>
<td>16.18</td>
<td>22.06</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>29.82</td>
<td>17.54</td>
<td>24.56</td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>CC</td>
<td>30.00</td>
<td>20.00</td>
<td>30.00</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>33.33</td>
<td>11.11</td>
<td>22.22</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>27.87</td>
<td>16.38</td>
<td>21.26</td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>CC</td>
<td>27.27</td>
<td>9.09</td>
<td>27.27</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>25.64</td>
<td>16.24</td>
<td>23.08</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>31.01</td>
<td>15.07</td>
<td>20.58</td>
</tr>
</tbody>
</table>
Table A19: Chi square test for investigating the relationship between different genotypes of the selected SNPs in Heme oxygenase-1 and Pseudomonas aeruginosa infection status

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Genotype</th>
<th>PA status (in percentage)</th>
<th>Chi Test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>AA</td>
<td>29.63</td>
<td>11.11</td>
<td>22.22</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>30.39</td>
<td>12.71</td>
<td>19.34</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>30.14</td>
<td>18.66</td>
<td>22.49</td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>CC</td>
<td>30.35</td>
<td>15.29</td>
<td>21.41</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>26.09</td>
<td>13.04</td>
<td>21.74</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0.00</td>
<td>50.00</td>
<td>0.00</td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>AA</td>
<td>30.07</td>
<td>18.95</td>
<td>19.61</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>31.82</td>
<td>14.09</td>
<td>20.00</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>23.71</td>
<td>11.34</td>
<td>27.84</td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>CC</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>CG</td>
<td>23.40</td>
<td>19.15</td>
<td>34.04</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>30.35</td>
<td>14.82</td>
<td>20.00</td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>AA</td>
<td>29.13</td>
<td>19.42</td>
<td>20.39</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>30.91</td>
<td>15.91</td>
<td>20.45</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>27.59</td>
<td>11.03</td>
<td>22.76</td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>AA</td>
<td>30.05</td>
<td>14.79</td>
<td>20.19</td>
</tr>
<tr>
<td></td>
<td>AT</td>
<td>23.91</td>
<td>19.57</td>
<td>32.61</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>100.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table A20: FBAT analysis of the age of diagnosis under the additive model. The age of diagnosis was first log-transformed before entering into the program.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPs</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td>BF_2557</td>
<td>A</td>
<td>224</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>BF_2557</td>
<td>G</td>
<td>224</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>BF_4022</td>
<td>A</td>
<td>298</td>
<td>0.755</td>
</tr>
<tr>
<td></td>
<td>BF_4022</td>
<td>G</td>
<td>298</td>
<td>0.755</td>
</tr>
<tr>
<td></td>
<td>BF_6484</td>
<td>A</td>
<td>148</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>BF_6484</td>
<td>G</td>
<td>148</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td>BF_7202</td>
<td>A</td>
<td>317</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td>BF_7202</td>
<td>G</td>
<td>317</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td>BF_8311</td>
<td>C</td>
<td>139</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>BF_8311</td>
<td>T</td>
<td>139</td>
<td>0.136</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_963</td>
<td>G</td>
<td>344</td>
<td>0.443</td>
</tr>
<tr>
<td></td>
<td>C3_963</td>
<td>T</td>
<td>344</td>
<td>0.443</td>
</tr>
<tr>
<td></td>
<td>C3_28795</td>
<td>A</td>
<td>332</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>C3_28795</td>
<td>G</td>
<td>332</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>C3_36735</td>
<td>A</td>
<td>320</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>C3_36735</td>
<td>G</td>
<td>320</td>
<td>0.250</td>
</tr>
<tr>
<td></td>
<td>C3_43118</td>
<td>A</td>
<td>342</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>C3_43118</td>
<td>G</td>
<td>342</td>
<td>0.048</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_851</td>
<td>A</td>
<td>300</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>TLR4_851</td>
<td>G</td>
<td>300</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>TLR4_1859</td>
<td>A</td>
<td>333</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td>TLR4_1859</td>
<td>G</td>
<td>333</td>
<td>0.730</td>
</tr>
<tr>
<td></td>
<td>TLR4_2856</td>
<td>C</td>
<td>207</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>TLR4_2856</td>
<td>T</td>
<td>207</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>TLR4_9263</td>
<td>A</td>
<td>153</td>
<td>0.855</td>
</tr>
<tr>
<td></td>
<td>TLR4_9263</td>
<td>C</td>
<td>153</td>
<td>0.855</td>
</tr>
<tr>
<td></td>
<td>TLR4_11912</td>
<td>G</td>
<td>317</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>TLR4_11912</td>
<td>T</td>
<td>317</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>TLR4_15884</td>
<td>C</td>
<td>194</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>TLR4_15884</td>
<td>G</td>
<td>194</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>TLR4_17050</td>
<td>C</td>
<td>196</td>
<td>0.344</td>
</tr>
<tr>
<td></td>
<td>TLR4_17050</td>
<td>T</td>
<td>196</td>
<td>0.344</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_149</td>
<td>A</td>
<td>250</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>HMOX1_149</td>
<td>G</td>
<td>250</td>
<td>0.289</td>
</tr>
<tr>
<td></td>
<td>HMOX1_1038</td>
<td>C</td>
<td>65</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>HMOX1_1038</td>
<td>T</td>
<td>65</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>HMOX1_2790</td>
<td>A</td>
<td>333</td>
<td>0.688</td>
</tr>
<tr>
<td></td>
<td>HMOX1_2790</td>
<td>T</td>
<td>333</td>
<td>0.688</td>
</tr>
<tr>
<td></td>
<td>HMOX1_3303</td>
<td>C</td>
<td>80</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>HMOX1_3303</td>
<td>G</td>
<td>80</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>HMOX1_9531</td>
<td>A</td>
<td>334</td>
<td>0.553</td>
</tr>
<tr>
<td></td>
<td>HMOX1_9531</td>
<td>G</td>
<td>334</td>
<td>0.553</td>
</tr>
<tr>
<td></td>
<td>HMOX1_16442</td>
<td>A</td>
<td>79</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>HMOX1_16442</td>
<td>T</td>
<td>79</td>
<td>0.194</td>
</tr>
</tbody>
</table>
Table A21: FBAT analysis of the age of diagnosis under the dominant model. The age of diagnosis was first log-transformed before entering into the program.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPs</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>A</td>
<td>214</td>
<td>0.370</td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>G</td>
<td>57</td>
<td>0.832</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>A</td>
<td>253</td>
<td>0.818</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>G</td>
<td>136</td>
<td>0.802</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>A</td>
<td>22</td>
<td>0.545</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>G</td>
<td>142</td>
<td>0.342</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>A</td>
<td>231</td>
<td>0.889</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>G</td>
<td>196</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>C</td>
<td>8</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>T</td>
<td>138</td>
<td>0.124</td>
<td></td>
</tr>
<tr>
<td><strong>Complement factor 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>G</td>
<td>222</td>
<td>0.942</td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>T</td>
<td>227</td>
<td>0.213</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>A</td>
<td>189</td>
<td>0.366</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>G</td>
<td>244</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>A</td>
<td>243</td>
<td>0.891</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>G</td>
<td>179</td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>A</td>
<td>239</td>
<td>0.177</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>G</td>
<td>223</td>
<td>0.096</td>
<td></td>
</tr>
<tr>
<td><strong>Toll-like receptor 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>A</td>
<td>94</td>
<td>0.401</td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>G</td>
<td>275</td>
<td>0.413</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>A</td>
<td>277</td>
<td>0.533</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>G</td>
<td>165</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>C</td>
<td>201</td>
<td>0.131</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>T</td>
<td>34</td>
<td>0.231</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>A</td>
<td>149</td>
<td>0.607</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>C</td>
<td>20</td>
<td>0.332</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>G</td>
<td>136</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>T</td>
<td>261</td>
<td>0.930</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>C</td>
<td>187</td>
<td>0.457</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>G</td>
<td>31</td>
<td>0.577</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>C</td>
<td>192</td>
<td>0.318</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>T</td>
<td>37</td>
<td>0.884</td>
<td></td>
</tr>
<tr>
<td><strong>Heme oxygenase-1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>A</td>
<td>205</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>G</td>
<td>117</td>
<td>0.495</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>C</td>
<td>4</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>T</td>
<td>65</td>
<td>0.198</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>A</td>
<td>185</td>
<td>0.304</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>T</td>
<td>246</td>
<td>0.762</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>C</td>
<td>80</td>
<td>0.064</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>G</td>
<td>7</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>A</td>
<td>226</td>
<td>0.195</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>G</td>
<td>208</td>
<td>0.648</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>A</td>
<td>7</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>T</td>
<td>79</td>
<td>0.080</td>
<td></td>
</tr>
</tbody>
</table>
Table A22: FBAT analysis of FEV1 predicted value under the additive model

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td>BF_2557</td>
<td>A</td>
<td>201</td>
<td>0.019</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_2557</td>
<td>G</td>
<td>201</td>
<td>0.019</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_4022</td>
<td>A</td>
<td>265</td>
<td>0.090</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_4022</td>
<td>G</td>
<td>265</td>
<td>0.090</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_6484</td>
<td>A</td>
<td>138</td>
<td>0.965</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_6484</td>
<td>G</td>
<td>138</td>
<td>0.965</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_7202</td>
<td>A</td>
<td>293</td>
<td>0.598</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_7202</td>
<td>G</td>
<td>293</td>
<td>0.598</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_8311</td>
<td>C</td>
<td>132</td>
<td>0.573</td>
</tr>
<tr>
<td>Factor B</td>
<td>BF_8311</td>
<td>T</td>
<td>132</td>
<td>0.573</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_963</td>
<td>G</td>
<td>321</td>
<td>0.124</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_963</td>
<td>T</td>
<td>321</td>
<td>0.124</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_28795</td>
<td>A</td>
<td>307</td>
<td>0.087</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_28795</td>
<td>G</td>
<td>307</td>
<td>0.087</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_36735</td>
<td>A</td>
<td>293</td>
<td>0.307</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_36735</td>
<td>G</td>
<td>293</td>
<td>0.307</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_43118</td>
<td>A</td>
<td>310</td>
<td>0.823</td>
</tr>
<tr>
<td>Complement factor 3</td>
<td>C3_43118</td>
<td>G</td>
<td>310</td>
<td>0.823</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_851</td>
<td>A</td>
<td>273</td>
<td>0.247</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_851</td>
<td>G</td>
<td>273</td>
<td>0.247</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_1859</td>
<td>A</td>
<td>297</td>
<td>0.610</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_1859</td>
<td>G</td>
<td>297</td>
<td>0.610</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_2856</td>
<td>C</td>
<td>187</td>
<td>0.876</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_2856</td>
<td>T</td>
<td>187</td>
<td>0.876</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_9263</td>
<td>A</td>
<td>146</td>
<td>0.721</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_9263</td>
<td>C</td>
<td>146</td>
<td>0.721</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_11912</td>
<td>G</td>
<td>285</td>
<td>0.092</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_11912</td>
<td>T</td>
<td>285</td>
<td>0.092</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_15884</td>
<td>C</td>
<td>170</td>
<td>0.036</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_15884</td>
<td>G</td>
<td>170</td>
<td>0.036</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_17050</td>
<td>C</td>
<td>185</td>
<td>0.940</td>
</tr>
<tr>
<td>Toll-like receptor 4</td>
<td>TLR4_17050</td>
<td>T</td>
<td>185</td>
<td>0.940</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_149</td>
<td>A</td>
<td>228</td>
<td>0.830</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_149</td>
<td>G</td>
<td>228</td>
<td>0.830</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_1038</td>
<td>C</td>
<td>59</td>
<td>0.015</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_1038</td>
<td>T</td>
<td>59</td>
<td>0.015</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_2790</td>
<td>A</td>
<td>308</td>
<td>0.469</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_2790</td>
<td>T</td>
<td>308</td>
<td>0.469</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_3303</td>
<td>C</td>
<td>77</td>
<td>0.333</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_3303</td>
<td>G</td>
<td>77</td>
<td>0.333</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_9531</td>
<td>A</td>
<td>316</td>
<td>0.886</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_9531</td>
<td>G</td>
<td>316</td>
<td>0.886</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_16442</td>
<td>A</td>
<td>73</td>
<td>0.693</td>
</tr>
<tr>
<td>Heme oxygenase-1</td>
<td>HMOX1_16442</td>
<td>T</td>
<td>73</td>
<td>0.693</td>
</tr>
</tbody>
</table>
Table A23: FBAT analysis of FEV1 predicted value under the dominant model

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>A</td>
<td>193</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>G</td>
<td>49</td>
<td>0.773</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>A</td>
<td>232</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>G</td>
<td>118</td>
<td>0.556</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>A</td>
<td>22</td>
<td>0.847</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>G</td>
<td>132</td>
<td>0.908</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>A</td>
<td>217</td>
<td>0.132</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>G</td>
<td>179</td>
<td>0.397</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>C</td>
<td>9</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>T</td>
<td>131</td>
<td>0.812</td>
<td></td>
</tr>
<tr>
<td><strong>Complement factor 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>G</td>
<td>207</td>
<td>0.267</td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>T</td>
<td>220</td>
<td>0.226</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>A</td>
<td>166</td>
<td>0.736</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>G</td>
<td>233</td>
<td>0.038</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>A</td>
<td>224</td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>G</td>
<td>166</td>
<td>0.305</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>A</td>
<td>216</td>
<td>0.616</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>G</td>
<td>200</td>
<td>0.865</td>
<td></td>
</tr>
<tr>
<td><strong>Toll-like receptor 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>A</td>
<td>93</td>
<td>0.487</td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>G</td>
<td>250</td>
<td>0.312</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>A</td>
<td>247</td>
<td>0.448</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>G</td>
<td>150</td>
<td>0.919</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>C</td>
<td>180</td>
<td>0.575</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>T</td>
<td>31</td>
<td>0.298</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>A</td>
<td>142</td>
<td>0.748</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>C</td>
<td>21</td>
<td>0.849</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>G</td>
<td>124</td>
<td>0.294</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>T</td>
<td>236</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>C</td>
<td>165</td>
<td>0.037</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>G</td>
<td>27</td>
<td>0.539</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>C</td>
<td>181</td>
<td>0.838</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>T</td>
<td>36</td>
<td>0.762</td>
<td></td>
</tr>
<tr>
<td><strong>Heme oxygenase-1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>A</td>
<td>190</td>
<td>0.984</td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>G</td>
<td>103</td>
<td>0.668</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>C</td>
<td>4</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>T</td>
<td>59</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>A</td>
<td>177</td>
<td>0.654</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>T</td>
<td>230</td>
<td>0.528</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>C</td>
<td>77</td>
<td>0.509</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>G</td>
<td>6</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>A</td>
<td>216</td>
<td>0.800</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>G</td>
<td>195</td>
<td>0.625</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>A</td>
<td>5</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>T</td>
<td>73</td>
<td>0.887</td>
<td></td>
</tr>
</tbody>
</table>
Table A24: FBAT analysis of FEV1 standard deviation value under the additive model

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td>BF_2557</td>
<td>A</td>
<td>201</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>BF_2557</td>
<td>G</td>
<td>201</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>BF_4022</td>
<td>A</td>
<td>264</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>BF_4022</td>
<td>G</td>
<td>264</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>BF_6484</td>
<td>A</td>
<td>137</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td>BF_6484</td>
<td>G</td>
<td>137</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td>BF_7202</td>
<td>A</td>
<td>293</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>BF_7202</td>
<td>G</td>
<td>293</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>BF_8311</td>
<td>C</td>
<td>131</td>
<td>0.859</td>
</tr>
<tr>
<td></td>
<td>BF_8311</td>
<td>T</td>
<td>131</td>
<td>0.859</td>
</tr>
<tr>
<td><strong>Complement factor 3</strong></td>
<td>C3_963</td>
<td>G</td>
<td>320</td>
<td>0.848</td>
</tr>
<tr>
<td></td>
<td>C3_963</td>
<td>T</td>
<td>320</td>
<td>0.848</td>
</tr>
<tr>
<td></td>
<td>C3_28795</td>
<td>A</td>
<td>306</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>C3_28795</td>
<td>G</td>
<td>306</td>
<td>0.316</td>
</tr>
<tr>
<td></td>
<td>C3_36735</td>
<td>A</td>
<td>293</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>C3_36735</td>
<td>G</td>
<td>293</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>C3_43118</td>
<td>A</td>
<td>309</td>
<td>0.224</td>
</tr>
<tr>
<td></td>
<td>C3_43118</td>
<td>G</td>
<td>309</td>
<td>0.224</td>
</tr>
<tr>
<td><strong>Toll-like receptor 4</strong></td>
<td>TLR4_851</td>
<td>A</td>
<td>272</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>TLR4_851</td>
<td>G</td>
<td>272</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td>TLR4_1859</td>
<td>A</td>
<td>296</td>
<td>0.473</td>
</tr>
<tr>
<td></td>
<td>TLR4_1859</td>
<td>G</td>
<td>296</td>
<td>0.473</td>
</tr>
<tr>
<td></td>
<td>TLR4_2856</td>
<td>C</td>
<td>187</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td>TLR4_2856</td>
<td>T</td>
<td>187</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td>TLR4_9263</td>
<td>A</td>
<td>146</td>
<td>0.356</td>
</tr>
<tr>
<td></td>
<td>TLR4_9263</td>
<td>C</td>
<td>146</td>
<td>0.356</td>
</tr>
<tr>
<td></td>
<td>TLR4_11912</td>
<td>G</td>
<td>285</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>TLR4_11912</td>
<td>T</td>
<td>285</td>
<td>0.377</td>
</tr>
<tr>
<td></td>
<td>TLR4_15884</td>
<td>C</td>
<td>170</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>TLR4_15884</td>
<td>G</td>
<td>170</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>TLR4_17050</td>
<td>C</td>
<td>185</td>
<td>0.643</td>
</tr>
<tr>
<td></td>
<td>TLR4_17050</td>
<td>T</td>
<td>185</td>
<td>0.643</td>
</tr>
<tr>
<td><strong>Heme oxygenase-1</strong></td>
<td>HMOX1_149</td>
<td>A</td>
<td>227</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>HMOX1_149</td>
<td>G</td>
<td>227</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>HMOX1_1038</td>
<td>C</td>
<td>59</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>HMOX1_1038</td>
<td>T</td>
<td>59</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td>HMOX1_2790</td>
<td>A</td>
<td>307</td>
<td>0.723</td>
</tr>
<tr>
<td></td>
<td>HMOX1_2790</td>
<td>T</td>
<td>307</td>
<td>0.723</td>
</tr>
<tr>
<td></td>
<td>HMOX1_3303</td>
<td>C</td>
<td>77</td>
<td>0.521</td>
</tr>
<tr>
<td></td>
<td>HMOX1_3303</td>
<td>G</td>
<td>77</td>
<td>0.521</td>
</tr>
<tr>
<td></td>
<td>HMOX1_9531</td>
<td>A</td>
<td>315</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>HMOX1_9531</td>
<td>G</td>
<td>315</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>HMOX1_16442</td>
<td>A</td>
<td>73</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td>HMOX1_16442</td>
<td>T</td>
<td>73</td>
<td>0.614</td>
</tr>
</tbody>
</table>
Table A25: FBAT analysis of FEV1 standard deviation value under the dominant model

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Allele</th>
<th># of Individuals</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor B</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>A</td>
<td>193</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td>BF_2557</td>
<td>G</td>
<td>49</td>
<td>0.290</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>A</td>
<td>231</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td>BF_4022</td>
<td>G</td>
<td>118</td>
<td>0.153</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>A</td>
<td>22</td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td>BF_6484</td>
<td>G</td>
<td>131</td>
<td>0.779</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>A</td>
<td>217</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>BF_7202</td>
<td>G</td>
<td>179</td>
<td>0.258</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>C</td>
<td>9</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>BF_8311</td>
<td>T</td>
<td>130</td>
<td>0.964</td>
<td></td>
</tr>
<tr>
<td><strong>Complement factor 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>G</td>
<td>207</td>
<td>0.602</td>
<td></td>
</tr>
<tr>
<td>C3_963</td>
<td>T</td>
<td>219</td>
<td>0.803</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>A</td>
<td>166</td>
<td>0.763</td>
<td></td>
</tr>
<tr>
<td>C3_28795</td>
<td>G</td>
<td>232</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>A</td>
<td>224</td>
<td>0.502</td>
<td></td>
</tr>
<tr>
<td>C3_36735</td>
<td>G</td>
<td>166</td>
<td>0.087</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>A</td>
<td>215</td>
<td>0.438</td>
<td></td>
</tr>
<tr>
<td>C3_43118</td>
<td>G</td>
<td>199</td>
<td>0.280</td>
<td></td>
</tr>
<tr>
<td><strong>Toll-like receptor 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>A</td>
<td>92</td>
<td>0.694</td>
<td></td>
</tr>
<tr>
<td>TLR4_851</td>
<td>G</td>
<td>249</td>
<td>0.826</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>A</td>
<td>246</td>
<td>0.195</td>
<td></td>
</tr>
<tr>
<td>TLR4_1859</td>
<td>G</td>
<td>149</td>
<td>0.648</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>C</td>
<td>180</td>
<td>0.911</td>
<td></td>
</tr>
<tr>
<td>TLR4_2856</td>
<td>T</td>
<td>31</td>
<td>0.320</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>A</td>
<td>142</td>
<td>0.509</td>
<td></td>
</tr>
<tr>
<td>TLR4_9263</td>
<td>C</td>
<td>21</td>
<td>0.236</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>G</td>
<td>124</td>
<td>0.827</td>
<td></td>
</tr>
<tr>
<td>TLR4_11912</td>
<td>T</td>
<td>236</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>C</td>
<td>165</td>
<td>0.220</td>
<td></td>
</tr>
<tr>
<td>TLR4_15884</td>
<td>G</td>
<td>27</td>
<td>0.404</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>C</td>
<td>181</td>
<td>0.955</td>
<td></td>
</tr>
<tr>
<td>TLR4_17050</td>
<td>T</td>
<td>36</td>
<td>0.176</td>
<td></td>
</tr>
<tr>
<td><strong>Heme oxygenase-1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>A</td>
<td>189</td>
<td>0.812</td>
<td></td>
</tr>
<tr>
<td>HMOX1_149</td>
<td>G</td>
<td>102</td>
<td>0.664</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>C</td>
<td>4</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_1038</td>
<td>T</td>
<td>59</td>
<td>0.069</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>A</td>
<td>176</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>HMOX1_2790</td>
<td>T</td>
<td>229</td>
<td>0.404</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>C</td>
<td>77</td>
<td>0.608</td>
<td></td>
</tr>
<tr>
<td>HMOX1_3303</td>
<td>G</td>
<td>6</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>A</td>
<td>215</td>
<td>0.0016</td>
<td></td>
</tr>
<tr>
<td>HMOX1_9531</td>
<td>G</td>
<td>194</td>
<td>0.558</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>A</td>
<td>5</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>HMOX1_16442</td>
<td>T</td>
<td>73</td>
<td>0.628</td>
<td></td>
</tr>
</tbody>
</table>
Table A26: Frequencies of possible haplotypes generated for the Factor B gene when determining the presence of any correlation between the haplotypes and age of diagnosis. 430 participants were included.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h11111</td>
<td>0.1850</td>
<td>0.0133</td>
</tr>
<tr>
<td>f.h12112</td>
<td>0.0005</td>
<td>0.0019</td>
</tr>
<tr>
<td>f.h12121</td>
<td>0.0116</td>
<td>0.0041</td>
</tr>
<tr>
<td>f.h21111</td>
<td>0.0385</td>
<td>0.0067</td>
</tr>
<tr>
<td>f.h21211</td>
<td>0.1040</td>
<td>0.0104</td>
</tr>
<tr>
<td>f.h22111</td>
<td>0.0180</td>
<td>0.0046</td>
</tr>
<tr>
<td>f.h22112</td>
<td>0.0947</td>
<td>0.0101</td>
</tr>
<tr>
<td>f.h22121</td>
<td>0.5477</td>
<td>0.0170</td>
</tr>
</tbody>
</table>

Table A27: Frequencies of possible haplotypes generated for the Factor B gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value. 380 participants were included.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h11111</td>
<td>0.1804</td>
<td>0.0140</td>
</tr>
<tr>
<td>f.h12112</td>
<td>0.0004</td>
<td>0.0022</td>
</tr>
<tr>
<td>f.h12121</td>
<td>0.0134</td>
<td>0.0047</td>
</tr>
<tr>
<td>f.h21111</td>
<td>0.0411</td>
<td>0.0074</td>
</tr>
<tr>
<td>f.h21211</td>
<td>0.0979</td>
<td>0.0108</td>
</tr>
<tr>
<td>f.h22111</td>
<td>0.0195</td>
<td>0.0051</td>
</tr>
<tr>
<td>f.h22112</td>
<td>0.1030</td>
<td>0.0112</td>
</tr>
<tr>
<td>f.h22121</td>
<td>0.5445</td>
<td>0.0181</td>
</tr>
</tbody>
</table>

Table A28: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Factor B and FEV1 predicted value, with no adjustment for the confounding factors.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h11111</td>
<td>-2.0478</td>
<td>3.7388</td>
<td>-0.5477</td>
<td>0.5839</td>
</tr>
<tr>
<td>h21211</td>
<td>1.4678</td>
<td>4.1147</td>
<td>0.3567</td>
<td>0.7213</td>
</tr>
<tr>
<td>h22121</td>
<td>2.8597</td>
<td>3.2461</td>
<td>0.8810</td>
<td>0.3783</td>
</tr>
<tr>
<td>pooled</td>
<td>-4.6232</td>
<td>4.6197</td>
<td>-1.0008</td>
<td>0.3169</td>
</tr>
</tbody>
</table>
Table A29: Haplotype analysis for investigation of possible correlation between combinations of selected SNPs in Factor B and FEV1 predicted value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h11111</td>
<td>-1.4185</td>
<td>3.7412</td>
<td>-0.3791</td>
<td>0.7046</td>
</tr>
<tr>
<td>h21211</td>
<td>1.9765</td>
<td>4.1065</td>
<td>0.4813</td>
<td>0.6303</td>
</tr>
<tr>
<td>h22121</td>
<td>3.4425</td>
<td>3.2503</td>
<td>1.0591</td>
<td>0.2895</td>
</tr>
<tr>
<td>pooled</td>
<td>-5.2740</td>
<td>4.6159</td>
<td>-1.1426</td>
<td>0.2532</td>
</tr>
<tr>
<td>SEXM</td>
<td>-2.8884</td>
<td>2.7585</td>
<td>-1.0471</td>
<td>0.2951</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>4.5233</td>
<td>3.1719</td>
<td>1.4261</td>
<td>0.1538</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>-0.4898</td>
<td>4.3212</td>
<td>-0.1134</td>
<td>0.9098</td>
</tr>
</tbody>
</table>

Table A30: Frequencies of possible haplotypes generated for the Factor B gene when determining the presence of any correlation between the haplotypes and FEV1 standard deviation value. 380 participants were included

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h11111</td>
<td>0.1804</td>
<td>0.0140</td>
</tr>
<tr>
<td>f.h12112</td>
<td>0.0003</td>
<td>0.0022</td>
</tr>
<tr>
<td>f.h12121</td>
<td>0.0135</td>
<td>0.0047</td>
</tr>
<tr>
<td>f.h21111</td>
<td>0.0410</td>
<td>0.0074</td>
</tr>
<tr>
<td>f.h21211</td>
<td>0.0979</td>
<td>0.0108</td>
</tr>
<tr>
<td>f.h22111</td>
<td>0.0194</td>
<td>0.0051</td>
</tr>
<tr>
<td>f.h22112</td>
<td>0.1031</td>
<td>0.0112</td>
</tr>
<tr>
<td>f.h22121</td>
<td>0.5444</td>
<td>0.0181</td>
</tr>
</tbody>
</table>
Table A31: Frequencies of possible haplotypes generated for the Complement Factor 3 gene when determining the presence of any correlation between the haplotypes and age of diagnosis. 434 participants were included

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h1111</td>
<td>0.0468</td>
<td>0.0112</td>
</tr>
<tr>
<td>f.h1112</td>
<td>0.0814</td>
<td>0.0122</td>
</tr>
<tr>
<td>f.h1121</td>
<td>0.1127</td>
<td>0.0143</td>
</tr>
<tr>
<td>f.h1122</td>
<td>0.1105</td>
<td>0.0142</td>
</tr>
<tr>
<td>f.h1211</td>
<td>0.0492</td>
<td>0.0116</td>
</tr>
<tr>
<td>f.h1212</td>
<td>0.0596</td>
<td>0.0115</td>
</tr>
<tr>
<td>f.h1221</td>
<td>0.0484</td>
<td>0.0117</td>
</tr>
<tr>
<td>f.h1222</td>
<td>0.0146</td>
<td>0.0076</td>
</tr>
<tr>
<td>f.h2111</td>
<td>0.0395</td>
<td>0.0099</td>
</tr>
<tr>
<td>f.h2112</td>
<td>0.0174</td>
<td>0.0077</td>
</tr>
<tr>
<td>f.h2121</td>
<td>0.0877</td>
<td>0.0141</td>
</tr>
<tr>
<td>f.h2122</td>
<td>0.1057</td>
<td>0.0141</td>
</tr>
<tr>
<td>f.h2211</td>
<td>0.0496</td>
<td>0.0135</td>
</tr>
<tr>
<td>f.h2212</td>
<td>0.0968</td>
<td>0.0134</td>
</tr>
<tr>
<td>f.h2221</td>
<td>0.0407</td>
<td>0.0125</td>
</tr>
<tr>
<td>f.h2222</td>
<td>0.0396</td>
<td>0.0107</td>
</tr>
</tbody>
</table>

Table A32: Frequencies of possible haplotypes generated for the Complement factor 3 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value. 384 participants were included

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h1111</td>
<td>0.0428</td>
<td>0.0120</td>
</tr>
<tr>
<td>f.h1112</td>
<td>0.0853</td>
<td>0.0133</td>
</tr>
<tr>
<td>f.h1121</td>
<td>0.1191</td>
<td>0.0153</td>
</tr>
<tr>
<td>f.h1122</td>
<td>0.1087</td>
<td>0.0147</td>
</tr>
<tr>
<td>f.h1211</td>
<td>0.0545</td>
<td>0.0125</td>
</tr>
<tr>
<td>f.h1212</td>
<td>0.0656</td>
<td>0.0122</td>
</tr>
<tr>
<td>f.h1221</td>
<td>0.0461</td>
<td>0.0119</td>
</tr>
<tr>
<td>f.h1222</td>
<td>0.0092</td>
<td>0.0063</td>
</tr>
<tr>
<td>f.h2111</td>
<td>0.0359</td>
<td>0.0099</td>
</tr>
<tr>
<td>f.h2112</td>
<td>0.0111</td>
<td>0.0063</td>
</tr>
<tr>
<td>f.h2121</td>
<td>0.0895</td>
<td>0.0147</td>
</tr>
<tr>
<td>f.h2122</td>
<td>0.1158</td>
<td>0.0149</td>
</tr>
<tr>
<td>f.h2211</td>
<td>0.0492</td>
<td>0.0134</td>
</tr>
<tr>
<td>f.h2212</td>
<td>0.0875</td>
<td>0.0131</td>
</tr>
<tr>
<td>f.h2221</td>
<td>0.0418</td>
<td>0.0132</td>
</tr>
<tr>
<td>f.h2222</td>
<td>0.0377</td>
<td>0.0102</td>
</tr>
</tbody>
</table>
Table A33: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Complement factor 3 and FEV1 predicted value, with no adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-3.1249</td>
<td>4.4955</td>
<td>-0.6951</td>
<td>0.4870</td>
</tr>
<tr>
<td>h1122</td>
<td>-5.7612</td>
<td>4.2983</td>
<td>-1.3403</td>
<td>0.1801</td>
</tr>
<tr>
<td>h1211</td>
<td>-3.2693</td>
<td>6.0141</td>
<td>-0.5436</td>
<td>0.5867</td>
</tr>
<tr>
<td>h1212</td>
<td>-7.1414</td>
<td>5.6725</td>
<td>-1.2590</td>
<td>0.2080</td>
</tr>
<tr>
<td>h2122</td>
<td>2.7698</td>
<td>4.1147</td>
<td>0.6731</td>
<td>0.5009</td>
</tr>
<tr>
<td>h2212</td>
<td>-5.0863</td>
<td>4.5827</td>
<td>-1.1099</td>
<td>0.2670</td>
</tr>
<tr>
<td>pooled</td>
<td>-1.8448</td>
<td>3.1835</td>
<td>-0.5795</td>
<td>0.5623</td>
</tr>
</tbody>
</table>

Table A34: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Complement factor 3 and FEV1 predicted value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-3.1173</td>
<td>4.5129</td>
<td>-0.6908</td>
<td>0.4897</td>
</tr>
<tr>
<td>h1122</td>
<td>-5.7416</td>
<td>4.3112</td>
<td>-1.3318</td>
<td>0.1829</td>
</tr>
<tr>
<td>h1211</td>
<td>-3.2399</td>
<td>6.1242</td>
<td>-0.5290</td>
<td>0.5968</td>
</tr>
<tr>
<td>h1212</td>
<td>-7.5332</td>
<td>5.8414</td>
<td>-1.2896</td>
<td>0.1972</td>
</tr>
<tr>
<td>h2122</td>
<td>2.3798</td>
<td>4.1276</td>
<td>0.5765</td>
<td>0.5642</td>
</tr>
<tr>
<td>h2212</td>
<td>-5.2010</td>
<td>4.6088</td>
<td>-1.285</td>
<td>0.2591</td>
</tr>
<tr>
<td>pooled</td>
<td>-1.7596</td>
<td>3.1975</td>
<td>-0.5503</td>
<td>0.5821</td>
</tr>
<tr>
<td>SEXM</td>
<td>-2.2927</td>
<td>2.7407</td>
<td>-0.8365</td>
<td>0.4029</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>3.4788</td>
<td>3.1592</td>
<td>1.1012</td>
<td>0.2708</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>1.8634</td>
<td>4.3248</td>
<td>0.4309</td>
<td>0.6666</td>
</tr>
</tbody>
</table>
Table A35: Frequencies of possible haplotypes generated for the Complement factor 3 gene when determining the presence of any correlation between the haplotypes and FEV1 standard deviation value. 384 participants were included.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h1111</td>
<td>0.0438</td>
<td>0.0121</td>
</tr>
<tr>
<td>f.h1112</td>
<td>0.0850</td>
<td>0.0133</td>
</tr>
<tr>
<td>f.h1121</td>
<td>0.1180</td>
<td>0.0153</td>
</tr>
<tr>
<td>f.h1122</td>
<td>0.1090</td>
<td>0.0147</td>
</tr>
<tr>
<td>f.h1211</td>
<td>0.0534</td>
<td>0.0123</td>
</tr>
<tr>
<td>f.h1212</td>
<td>0.0662</td>
<td>0.0122</td>
</tr>
<tr>
<td>f.h1221</td>
<td>0.0468</td>
<td>0.0117</td>
</tr>
<tr>
<td>f.h1222</td>
<td>0.0091</td>
<td>0.0061</td>
</tr>
<tr>
<td>f.h2111</td>
<td>0.0349</td>
<td>0.0098</td>
</tr>
<tr>
<td>f.h2112</td>
<td>0.0112</td>
<td>0.0064</td>
</tr>
<tr>
<td>f.h2121</td>
<td>0.0897</td>
<td>0.0147</td>
</tr>
<tr>
<td>f.h2122</td>
<td>0.1166</td>
<td>0.0151</td>
</tr>
<tr>
<td>f.h2211</td>
<td>0.0506</td>
<td>0.0134</td>
</tr>
<tr>
<td>f.h2212</td>
<td>0.0870</td>
<td>0.0131</td>
</tr>
<tr>
<td>f.h2221</td>
<td>0.0415</td>
<td>0.0130</td>
</tr>
<tr>
<td>f.h2222</td>
<td>0.0371</td>
<td>0.0101</td>
</tr>
</tbody>
</table>

Table A36: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Complement factor 3 and FEV1 standard deviation value, with no adjustment for the confounding factors.

<table>
<thead>
<tr>
<th>SNP Combination</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-0.1705</td>
<td>0.1811</td>
<td>-0.9411</td>
<td>0.3467</td>
</tr>
<tr>
<td>h1122</td>
<td>-0.2052</td>
<td>0.2000</td>
<td>-1.0263</td>
<td>0.3048</td>
</tr>
<tr>
<td>h1211</td>
<td>-0.1801</td>
<td>0.2443</td>
<td>-0.7372</td>
<td>0.4610</td>
</tr>
<tr>
<td>h1212</td>
<td>-0.1882</td>
<td>0.2235</td>
<td>-0.8421</td>
<td>0.3998</td>
</tr>
<tr>
<td>h2121</td>
<td>0.1773</td>
<td>0.2090</td>
<td>0.8486</td>
<td>0.3961</td>
</tr>
<tr>
<td>h2122</td>
<td>0.1110</td>
<td>0.1619</td>
<td>0.6858</td>
<td>0.4929</td>
</tr>
<tr>
<td>h2212</td>
<td>-0.1177</td>
<td>0.1899</td>
<td>-0.6196</td>
<td>0.5355</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0260</td>
<td>0.1428</td>
<td>-0.1818</td>
<td>0.8557</td>
</tr>
</tbody>
</table>
Table A37: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Complement factor 3 and FEV1 standard deviation value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1112</td>
<td>-0.1686</td>
<td>0.1817</td>
<td>-0.9278</td>
<td>0.3535</td>
</tr>
<tr>
<td>h1122</td>
<td>-0.1966</td>
<td>0.2001</td>
<td>-0.9826</td>
<td>0.3258</td>
</tr>
<tr>
<td>h1211</td>
<td>-0.1766</td>
<td>0.2469</td>
<td>-0.7152</td>
<td>0.4745</td>
</tr>
<tr>
<td>h1212</td>
<td>-0.2058</td>
<td>0.2274</td>
<td>-0.9051</td>
<td>0.3654</td>
</tr>
<tr>
<td>h2121</td>
<td>0.1790</td>
<td>0.2093</td>
<td>0.8550</td>
<td>0.3926</td>
</tr>
<tr>
<td>h2122</td>
<td>0.1032</td>
<td>0.1623</td>
<td>0.6362</td>
<td>0.5246</td>
</tr>
<tr>
<td>h2211</td>
<td>0.1142</td>
<td>0.1901</td>
<td>0.6009</td>
<td>0.5479</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0204</td>
<td>0.1430</td>
<td>-0.1427</td>
<td>0.8865</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0384</td>
<td>0.0975</td>
<td>0.3940</td>
<td>0.6936</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.0948</td>
<td>0.1120</td>
<td>0.8463</td>
<td>0.3974</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>-0.0210</td>
<td>0.1534</td>
<td>-0.1366</td>
<td>0.8914</td>
</tr>
</tbody>
</table>

Table A38: Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and age of diagnosis. 432 participants were included

<table>
<thead>
<tr>
<th>Haplotype Frequency</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h112212</td>
<td>0.1447</td>
<td>0.0121</td>
</tr>
<tr>
<td>f.h1122122</td>
<td>0.2303</td>
<td>0.0144</td>
</tr>
<tr>
<td>f.h121112</td>
<td>0.0012</td>
<td>0.0012</td>
</tr>
<tr>
<td>f.h121221</td>
<td>0.0428</td>
<td>0.0070</td>
</tr>
<tr>
<td>f.h121222</td>
<td>0.1042</td>
<td>0.0104</td>
</tr>
<tr>
<td>f.h122121</td>
<td>0.0016</td>
<td>0.0015</td>
</tr>
<tr>
<td>f.h1221221</td>
<td>0.1026</td>
<td>0.0104</td>
</tr>
<tr>
<td>f.h1221222</td>
<td>0.0057</td>
<td>0.0027</td>
</tr>
<tr>
<td>f.h1222112</td>
<td>0.0013</td>
<td>0.0014</td>
</tr>
<tr>
<td>f.h1222122</td>
<td>0.0097</td>
<td>0.0035</td>
</tr>
<tr>
<td>f.h122221</td>
<td>0.0010</td>
<td>0.0015</td>
</tr>
<tr>
<td>f.h1222222</td>
<td>0.0729</td>
<td>0.0089</td>
</tr>
<tr>
<td>f.h2222112</td>
<td>0.0015</td>
<td>0.0015</td>
</tr>
<tr>
<td>f.h2222121</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h22222122</td>
<td>0.2788</td>
<td>0.0153</td>
</tr>
</tbody>
</table>
Table A39: Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value. 382 participants were included.

<table>
<thead>
<tr>
<th>Haplotype Frequency</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h1122112</td>
<td>0.1350</td>
<td>0.0125</td>
</tr>
<tr>
<td>f.h1122122</td>
<td>0.2354</td>
<td>0.0154</td>
</tr>
<tr>
<td>f.h12112112</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
<tr>
<td>f.h12122211</td>
<td>0.0405</td>
<td>0.0072</td>
</tr>
<tr>
<td>f.h12122222</td>
<td>0.1022</td>
<td>0.0110</td>
</tr>
<tr>
<td>f.h12211212</td>
<td>0.0018</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h12212211</td>
<td>0.1043</td>
<td>0.0111</td>
</tr>
<tr>
<td>f.h12212222</td>
<td>0.0065</td>
<td>0.0031</td>
</tr>
<tr>
<td>f.h12221112</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
<tr>
<td>f.h12221222</td>
<td>0.0070</td>
<td>0.0031</td>
</tr>
<tr>
<td>f.h12222211</td>
<td>0.0011</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h12222222</td>
<td>0.0733</td>
<td>0.0095</td>
</tr>
<tr>
<td>f.h22221112</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h22221212</td>
<td>0.0019</td>
<td>0.0019</td>
</tr>
<tr>
<td>f.h22222222</td>
<td>0.2867</td>
<td>0.0164</td>
</tr>
</tbody>
</table>

Table A40: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 predicted value, with no adjustment for the confounding factors.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1122112</td>
<td>-0.6825</td>
<td>3.1893</td>
<td>-0.2140</td>
<td>0.8306</td>
</tr>
<tr>
<td>h1122122</td>
<td>1.7004</td>
<td>2.7327</td>
<td>0.6222</td>
<td>0.5338</td>
</tr>
<tr>
<td>h1212222</td>
<td>-1.1586</td>
<td>3.6077</td>
<td>-0.3212</td>
<td>0.7481</td>
</tr>
<tr>
<td>h1221221</td>
<td>-0.0822</td>
<td>3.6590</td>
<td>-0.0225</td>
<td>0.9821</td>
</tr>
<tr>
<td>h1222222</td>
<td>2.0593</td>
<td>4.0267</td>
<td>0.5114</td>
<td>0.6091</td>
</tr>
<tr>
<td>pooled</td>
<td>1.7087</td>
<td>4.4352</td>
<td>0.3853</td>
<td>0.7001</td>
</tr>
</tbody>
</table>
### Table A41: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 predicted value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1122112</td>
<td>-0.9215</td>
<td>3.2089</td>
<td>-0.2872</td>
<td>0.7740</td>
</tr>
<tr>
<td>h1122122</td>
<td>1.7599</td>
<td>2.7359</td>
<td>0.6433</td>
<td>0.5201</td>
</tr>
<tr>
<td>h1212222</td>
<td>-1.1055</td>
<td>3.6083</td>
<td>-0.3064</td>
<td>0.7593</td>
</tr>
<tr>
<td>h1221221</td>
<td>-0.1823</td>
<td>3.6909</td>
<td>-0.0494</td>
<td>0.9606</td>
</tr>
<tr>
<td>h1222222</td>
<td>2.1089</td>
<td>4.0556</td>
<td>0.5200</td>
<td>0.6031</td>
</tr>
<tr>
<td>pooled</td>
<td>1.7826</td>
<td>4.4310</td>
<td>0.4023</td>
<td>0.6875</td>
</tr>
<tr>
<td>SEXM</td>
<td>-3.0651</td>
<td>2.7789</td>
<td>-1.1030</td>
<td>0.2700</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>2.8110</td>
<td>3.1823</td>
<td>0.8833</td>
<td>0.3771</td>
</tr>
<tr>
<td>genotypeOO</td>
<td>0.5437</td>
<td>4.3542</td>
<td>0.1249</td>
<td>0.9006</td>
</tr>
</tbody>
</table>

### Table A42: Frequencies of possible haplotypes generated for the Toll-like receptor 4 gene when determining the presence of any correlation between the haplotypes and FEV1 standard deviation value. 382 participants were included

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h1122112</td>
<td>0.1350</td>
<td>0.0125</td>
</tr>
<tr>
<td>f.h1122122</td>
<td>0.2354</td>
<td>0.0154</td>
</tr>
<tr>
<td>f.h1212112</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
<tr>
<td>f.h1212221</td>
<td>0.0405</td>
<td>0.0072</td>
</tr>
<tr>
<td>f.h1212222</td>
<td>0.1022</td>
<td>0.0110</td>
</tr>
<tr>
<td>f.h1221212</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h1222121</td>
<td>0.1043</td>
<td>0.0111</td>
</tr>
<tr>
<td>f.h1222122</td>
<td>0.0065</td>
<td>0.0031</td>
</tr>
<tr>
<td>f.h1222112</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
<tr>
<td>f.h1222122</td>
<td>0.0070</td>
<td>0.0031</td>
</tr>
<tr>
<td>f.h1222221</td>
<td>0.0011</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h1222222</td>
<td>0.0733</td>
<td>0.0095</td>
</tr>
<tr>
<td>f.h2222112</td>
<td>0.0017</td>
<td>0.0017</td>
</tr>
<tr>
<td>f.h2222121</td>
<td>0.0020</td>
<td>0.0020</td>
</tr>
<tr>
<td>f.h2222122</td>
<td>0.2867</td>
<td>0.0164</td>
</tr>
</tbody>
</table>
Table A43: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Toll-like receptor 4 and FEV1 standard deviation value, with no adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h122112</td>
<td>-0.0400</td>
<td>0.1136</td>
<td>-0.3521</td>
<td>0.7248</td>
</tr>
<tr>
<td>h122122</td>
<td>-0.0148</td>
<td>0.0974</td>
<td>-0.1518</td>
<td>0.8793</td>
</tr>
<tr>
<td>h121222</td>
<td>-0.1856</td>
<td>0.1285</td>
<td>-1.4440</td>
<td>0.1487</td>
</tr>
<tr>
<td>h122121</td>
<td>-0.0675</td>
<td>0.1303</td>
<td>-0.5182</td>
<td>0.6043</td>
</tr>
<tr>
<td>h122222</td>
<td>-0.1347</td>
<td>0.1437</td>
<td>-0.9371</td>
<td>0.3487</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.1072</td>
<td>0.1580</td>
<td>-0.6782</td>
<td>0.4976</td>
</tr>
</tbody>
</table>

Table A44: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Toll-like receptor 4 and FEV1 standard deviation value, with adjustment for the confounding factors

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h122112</td>
<td>-0.0315</td>
<td>0.1144</td>
<td>-0.2756</td>
<td>0.7828</td>
</tr>
<tr>
<td>h122122</td>
<td>-0.0125</td>
<td>0.0977</td>
<td>-0.1280</td>
<td>0.8981</td>
</tr>
<tr>
<td>h121222</td>
<td>-0.1776</td>
<td>0.1287</td>
<td>-1.3798</td>
<td>0.1676</td>
</tr>
<tr>
<td>h122121</td>
<td>-0.0629</td>
<td>0.1316</td>
<td>-0.4782</td>
<td>0.6325</td>
</tr>
<tr>
<td>h122222</td>
<td>-0.1501</td>
<td>0.1449</td>
<td>-1.0354</td>
<td>0.3005</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0991</td>
<td>0.1581</td>
<td>-0.6268</td>
<td>0.5308</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0269</td>
<td>0.0991</td>
<td>0.2714</td>
<td>0.7861</td>
</tr>
<tr>
<td>genotypeF0</td>
<td>0.0901</td>
<td>0.1135</td>
<td>0.7938</td>
<td>0.4273</td>
</tr>
<tr>
<td>genotypeO0</td>
<td>-0.0533</td>
<td>0.1553</td>
<td>-0.3433</td>
<td>0.7314</td>
</tr>
</tbody>
</table>
Table A45: Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and age of diagnosis. 433 participants were included.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h111211</td>
<td>0.0074</td>
<td>0.0031</td>
</tr>
<tr>
<td>f.h111221</td>
<td>0.0015</td>
<td>0.0015</td>
</tr>
<tr>
<td>f.h112122</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
<tr>
<td>f.h112221</td>
<td>0.3273</td>
<td>0.0160</td>
</tr>
<tr>
<td>f.h211211</td>
<td>0.3730</td>
<td>0.0166</td>
</tr>
<tr>
<td>f.h211212</td>
<td>0.0100</td>
<td>0.0040</td>
</tr>
<tr>
<td>f.h211221</td>
<td>0.1085</td>
<td>0.0107</td>
</tr>
<tr>
<td>f.h211222</td>
<td>0.0036</td>
<td>0.0022</td>
</tr>
<tr>
<td>f.h212121</td>
<td>0.0135</td>
<td>0.0040</td>
</tr>
<tr>
<td>f.h212122</td>
<td>0.0363</td>
<td>0.0064</td>
</tr>
<tr>
<td>f.h212221</td>
<td>0.0612</td>
<td>0.0084</td>
</tr>
<tr>
<td>f.h212222</td>
<td>0.0019</td>
<td>0.0016</td>
</tr>
<tr>
<td>f.h221211</td>
<td>0.0522</td>
<td>0.0077</td>
</tr>
<tr>
<td>f.h221212</td>
<td>0.0010</td>
<td>0.0018</td>
</tr>
<tr>
<td>f.h221221</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

Table A46: Haplotype analysis for investigation of correlation relationship between combinations of selected SNPs in Heme oxygenase-1 and the age of diagnosis, with no adjustment for the confounding factors. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-0.1633</td>
<td>0.1525</td>
<td>-1.0710</td>
<td>0.2842</td>
</tr>
<tr>
<td>h211221</td>
<td>-0.1412</td>
<td>0.2224</td>
<td>-0.6348</td>
<td>0.5255</td>
</tr>
<tr>
<td>h212221</td>
<td>-0.2404</td>
<td>0.3017</td>
<td>-0.7967</td>
<td>0.4256</td>
</tr>
<tr>
<td>h21211</td>
<td>-0.5300</td>
<td>0.3007</td>
<td>-1.7626</td>
<td>0.0780</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0541</td>
<td>0.2641</td>
<td>-0.2049</td>
<td>0.8377</td>
</tr>
</tbody>
</table>
Table A47: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Heme oxygenase-1 and the age of diagnosis, with adjustment for the confounding factors. The age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Regression Coefficient</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h112221</td>
<td>-0.1535</td>
<td>0.1517</td>
<td>-1.0116</td>
<td>0.3117</td>
</tr>
<tr>
<td>h211221</td>
<td>-0.1162</td>
<td>0.2216</td>
<td>-0.5245</td>
<td>0.6000</td>
</tr>
<tr>
<td>h212221</td>
<td>-0.2301</td>
<td>0.3008</td>
<td>-0.7647</td>
<td>0.4445</td>
</tr>
<tr>
<td>h221211</td>
<td>-0.5149</td>
<td>0.3017</td>
<td>-1.7068</td>
<td>0.0879</td>
</tr>
<tr>
<td>pooled</td>
<td>-0.0804</td>
<td>0.2632</td>
<td>-0.3053</td>
<td>0.7602</td>
</tr>
<tr>
<td>SEXM</td>
<td>0.0633</td>
<td>0.1820</td>
<td>0.3477</td>
<td>0.7281</td>
</tr>
<tr>
<td>genotypeFO</td>
<td>0.4132</td>
<td>0.2078</td>
<td>1.9883</td>
<td>0.0468</td>
</tr>
<tr>
<td>genotype00</td>
<td>0.3573</td>
<td>0.2910</td>
<td>1.2278</td>
<td>0.2195</td>
</tr>
</tbody>
</table>

Table A48: Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and FEV1 predicted value. 383 participants were included.

<table>
<thead>
<tr>
<th>Haplotype Frequency</th>
<th>Estimate of Haplotype Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h111211</td>
<td>0.0085</td>
<td>0.0035</td>
</tr>
<tr>
<td>f.h111221</td>
<td>0.0018</td>
<td>0.0018</td>
</tr>
<tr>
<td>f.h112122</td>
<td>0.0014</td>
<td>0.0014</td>
</tr>
<tr>
<td>f.h112221</td>
<td>0.3308</td>
<td>0.0171</td>
</tr>
<tr>
<td>f.h211211</td>
<td>0.3744</td>
<td>0.0176</td>
</tr>
<tr>
<td>f.h211212</td>
<td>0.0100</td>
<td>0.0039</td>
</tr>
<tr>
<td>f.h211221</td>
<td>0.1072</td>
<td>0.0113</td>
</tr>
<tr>
<td>f.h211222</td>
<td>0.0026</td>
<td>0.0022</td>
</tr>
<tr>
<td>f.h212121</td>
<td>0.0097</td>
<td>0.0036</td>
</tr>
<tr>
<td>f.h212122</td>
<td>0.0413</td>
<td>0.0072</td>
</tr>
<tr>
<td>f.h212221</td>
<td>0.0592</td>
<td>0.0088</td>
</tr>
<tr>
<td>f.h212222</td>
<td>0.0020</td>
<td>0.0018</td>
</tr>
<tr>
<td>f.h221211</td>
<td>0.0499</td>
<td>0.0079</td>
</tr>
<tr>
<td>f.h221221</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
</tbody>
</table>
Table A49: Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene when determining the presence of any correlation between the haplotypes and FEV1 standard deviation value. 383 participants were included.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimate of Frequency</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>f.h111211</td>
<td>0.0085</td>
<td>0.0035</td>
</tr>
<tr>
<td>f.h111221</td>
<td>0.0018</td>
<td>0.0018</td>
</tr>
<tr>
<td>f.h112122</td>
<td>0.0014</td>
<td>0.0014</td>
</tr>
<tr>
<td>f.h112221</td>
<td>0.3307</td>
<td>0.0171</td>
</tr>
<tr>
<td>f.h211211</td>
<td>0.3745</td>
<td>0.0176</td>
</tr>
<tr>
<td>f.h211212</td>
<td>0.0098</td>
<td>0.0039</td>
</tr>
<tr>
<td>f.h211221</td>
<td>0.1073</td>
<td>0.0113</td>
</tr>
<tr>
<td>f.h211222</td>
<td>0.0026</td>
<td>0.0022</td>
</tr>
<tr>
<td>f.h212121</td>
<td>0.0097</td>
<td>0.0036</td>
</tr>
<tr>
<td>f.h212122</td>
<td>0.0413</td>
<td>0.0072</td>
</tr>
<tr>
<td>f.h212221</td>
<td>0.0591</td>
<td>0.0088</td>
</tr>
<tr>
<td>f.h212222</td>
<td>0.0021</td>
<td>0.0018</td>
</tr>
<tr>
<td>f.h221121</td>
<td>0.0499</td>
<td>0.0079</td>
</tr>
<tr>
<td>f.h221221</td>
<td>0.0013</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

Table A50: Frequencies of possible haplotypes generated for the Factor B gene by the FBAT program.

<table>
<thead>
<tr>
<th>haplotypes</th>
<th>Estimates of frequency:</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>22121</td>
</tr>
<tr>
<td>h2</td>
<td>11111</td>
</tr>
<tr>
<td>h3</td>
<td>21211</td>
</tr>
<tr>
<td>h4</td>
<td>22112</td>
</tr>
<tr>
<td>h5</td>
<td>21111</td>
</tr>
<tr>
<td>h6</td>
<td>22111</td>
</tr>
<tr>
<td>h7</td>
<td>12121</td>
</tr>
<tr>
<td>h8</td>
<td>22122</td>
</tr>
<tr>
<td>h9</td>
<td>22221</td>
</tr>
<tr>
<td>h10</td>
<td>21221</td>
</tr>
</tbody>
</table>
Table A51: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Factor B and age of diagnosis by the FBAT program. Age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of Family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>251</td>
<td>-104.948</td>
<td>-92.713</td>
<td>76.79</td>
<td>-1.396</td>
<td>0.163</td>
</tr>
<tr>
<td>h2</td>
<td>182</td>
<td>-27.32</td>
<td>-33.303</td>
<td>54.699</td>
<td>0.809</td>
<td>0.418</td>
</tr>
<tr>
<td>h3</td>
<td>117</td>
<td>-14.405</td>
<td>-17.557</td>
<td>27.811</td>
<td>0.598</td>
<td>0.550</td>
</tr>
<tr>
<td>h4</td>
<td>120</td>
<td>0.241</td>
<td>-7.648</td>
<td>21.898</td>
<td>1.686</td>
<td>0.092</td>
</tr>
<tr>
<td>h5</td>
<td>55</td>
<td>-15.496</td>
<td>-10.599</td>
<td>11.493</td>
<td>-1.444</td>
<td>0.149</td>
</tr>
<tr>
<td>h6</td>
<td>31</td>
<td>-1.428</td>
<td>-2.042</td>
<td>6.609</td>
<td>0.239</td>
<td>0.811</td>
</tr>
<tr>
<td>h7</td>
<td>19</td>
<td>-0.349</td>
<td>-0.24</td>
<td>2.519</td>
<td>-0.068</td>
<td>0.946</td>
</tr>
<tr>
<td>h8</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h9</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table A52: Frequencies of possible haplotypes generated for the Complement factor 3 gene by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimates of frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>2122</td>
</tr>
<tr>
<td>h2</td>
<td>1121</td>
</tr>
<tr>
<td>h3</td>
<td>1122</td>
</tr>
<tr>
<td>h4</td>
<td>2212</td>
</tr>
<tr>
<td>h5</td>
<td>2121</td>
</tr>
<tr>
<td>h6</td>
<td>1212</td>
</tr>
<tr>
<td>h7</td>
<td>1112</td>
</tr>
<tr>
<td>h8</td>
<td>2221</td>
</tr>
<tr>
<td>h9</td>
<td>1221</td>
</tr>
<tr>
<td>h10</td>
<td>2111</td>
</tr>
<tr>
<td>h11</td>
<td>1211</td>
</tr>
<tr>
<td>h12</td>
<td>2211</td>
</tr>
<tr>
<td>h13</td>
<td>1111</td>
</tr>
<tr>
<td>h14</td>
<td>2222</td>
</tr>
<tr>
<td>h15</td>
<td>1222</td>
</tr>
<tr>
<td>h16</td>
<td>2112</td>
</tr>
</tbody>
</table>
Table A53: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Complement factor 3 and FEV1 predicted value by FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>146</td>
<td>7487.676</td>
<td>7443.953</td>
<td>0.09</td>
<td>0.928</td>
<td></td>
</tr>
<tr>
<td>h2</td>
<td>133</td>
<td>6542.610</td>
<td>6142.430</td>
<td>0.847</td>
<td>0.397</td>
<td></td>
</tr>
<tr>
<td>h3</td>
<td>124</td>
<td>6316.309</td>
<td>5745.721</td>
<td>1.34</td>
<td>0.180</td>
<td></td>
</tr>
<tr>
<td>h4</td>
<td>108</td>
<td>4553.519</td>
<td>5070.933</td>
<td>-1.233</td>
<td>0.218</td>
<td></td>
</tr>
<tr>
<td>h5</td>
<td>93</td>
<td>4105.946</td>
<td>4166.365</td>
<td>-0.158</td>
<td>0.874</td>
<td></td>
</tr>
<tr>
<td>h6</td>
<td>89</td>
<td>3466.416</td>
<td>3999.288</td>
<td>-1.417</td>
<td>0.156</td>
<td></td>
</tr>
<tr>
<td>h7</td>
<td>97</td>
<td>4850.073</td>
<td>4148.990</td>
<td>1.893</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>h8</td>
<td>63</td>
<td>2794.117</td>
<td>3115.907</td>
<td>-0.956</td>
<td>0.339</td>
<td></td>
</tr>
<tr>
<td>h9</td>
<td>70</td>
<td>2983.738</td>
<td>2909.471</td>
<td>0.219</td>
<td>0.826</td>
<td></td>
</tr>
<tr>
<td>h10</td>
<td>66</td>
<td>2559.862</td>
<td>2640.701</td>
<td>-0.261</td>
<td>0.794</td>
<td></td>
</tr>
<tr>
<td>h11</td>
<td>64</td>
<td>2687.309</td>
<td>2723.217</td>
<td>-0.148</td>
<td>0.882</td>
<td></td>
</tr>
<tr>
<td>h12</td>
<td>61</td>
<td>2630.163</td>
<td>2708.422</td>
<td>-0.247</td>
<td>0.805</td>
<td></td>
</tr>
<tr>
<td>h13</td>
<td>62</td>
<td>2878.470</td>
<td>2702.198</td>
<td>0.563</td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td>h14</td>
<td>33</td>
<td>1396.275</td>
<td>1337.111</td>
<td>0.26</td>
<td>0.795</td>
<td></td>
</tr>
<tr>
<td>h15</td>
<td>30</td>
<td>998.844</td>
<td>1197.097</td>
<td>-1.017</td>
<td>0.309</td>
<td></td>
</tr>
<tr>
<td>h16</td>
<td>29</td>
<td>807.060</td>
<td>997.585</td>
<td>-1.135</td>
<td>0.256</td>
<td></td>
</tr>
</tbody>
</table>
Table A54: Frequencies of possible haplotypes generated for the gene of Toll-like receptor 4 by the FBAT program

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>Estimates of frequency:</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>2222122</td>
</tr>
<tr>
<td>h2</td>
<td>1122122</td>
</tr>
<tr>
<td>h3</td>
<td>1122112</td>
</tr>
<tr>
<td>h4</td>
<td>1212222</td>
</tr>
<tr>
<td>h5</td>
<td>1221221</td>
</tr>
<tr>
<td>h6</td>
<td>1222222</td>
</tr>
<tr>
<td>h7</td>
<td>1212221</td>
</tr>
<tr>
<td>h8</td>
<td>1222122</td>
</tr>
<tr>
<td>h9</td>
<td>1221222</td>
</tr>
<tr>
<td>h10</td>
<td>2222112</td>
</tr>
<tr>
<td>h11</td>
<td>1222221</td>
</tr>
<tr>
<td>h12</td>
<td>1122121</td>
</tr>
<tr>
<td>h13</td>
<td>1212122</td>
</tr>
<tr>
<td>h14</td>
<td>1212121</td>
</tr>
<tr>
<td>h15</td>
<td>1212212</td>
</tr>
<tr>
<td>h16</td>
<td>1222112</td>
</tr>
<tr>
<td>h17</td>
<td>1112122</td>
</tr>
<tr>
<td>h18</td>
<td>1222212</td>
</tr>
<tr>
<td>h19</td>
<td>1122222</td>
</tr>
<tr>
<td>h20</td>
<td>1221212</td>
</tr>
<tr>
<td>h21</td>
<td>2222121</td>
</tr>
<tr>
<td>h22</td>
<td>1221111</td>
</tr>
<tr>
<td>h23</td>
<td>1221121</td>
</tr>
</tbody>
</table>
Table A55: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Toll-like receptor 4 and age of diagnosis by the FBAT program. Age of diagnosis was logarithmically transformed for normality.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>238</td>
<td>-43.878</td>
<td>-53.555</td>
<td>61.458</td>
<td>1.234</td>
<td>0.217</td>
</tr>
<tr>
<td>h2</td>
<td>221</td>
<td>-31.153</td>
<td>-22.668</td>
<td>51.439</td>
<td>-1.183</td>
<td>0.237</td>
</tr>
<tr>
<td>h3</td>
<td>170</td>
<td>-24.737</td>
<td>-30.509</td>
<td>34.862</td>
<td>0.978</td>
<td>0.328</td>
</tr>
<tr>
<td>h4</td>
<td>133</td>
<td>-20.116</td>
<td>-16.067</td>
<td>26.067</td>
<td>-0.793</td>
<td>0.428</td>
</tr>
<tr>
<td>h5</td>
<td>117</td>
<td>-17.313</td>
<td>-12.126</td>
<td>25.422</td>
<td>-1.029</td>
<td>0.304</td>
</tr>
<tr>
<td>h6</td>
<td>90</td>
<td>-20.642</td>
<td>-19.783</td>
<td>21.797</td>
<td>-0.184</td>
<td>0.854</td>
</tr>
<tr>
<td>h7</td>
<td>61</td>
<td>-15.974</td>
<td>-14.514</td>
<td>15.556</td>
<td>-0.370</td>
<td>0.711</td>
</tr>
<tr>
<td>h8</td>
<td>21</td>
<td>-3.244</td>
<td>-4.737</td>
<td>4.760</td>
<td>0.684</td>
<td>0.494</td>
</tr>
<tr>
<td>h9</td>
<td>13</td>
<td>-2.272</td>
<td>-5.136</td>
<td>4.983</td>
<td>1.283</td>
<td>0.200</td>
</tr>
<tr>
<td>h10</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h22</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h23</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
### Table A56: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Toll-like receptor 4 and FEV1 standard deviation value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>215</td>
<td>89.665</td>
<td>87.686</td>
<td>83.367</td>
<td>0.217</td>
<td>0.828</td>
</tr>
<tr>
<td>h2</td>
<td>198</td>
<td>86.617</td>
<td>78.920</td>
<td>81.316</td>
<td>0.854</td>
<td>0.393</td>
</tr>
<tr>
<td>h3</td>
<td>149</td>
<td>38.735</td>
<td>50.966</td>
<td>56.578</td>
<td>-1.626</td>
<td>0.104</td>
</tr>
<tr>
<td>h4</td>
<td>114</td>
<td>28.836</td>
<td>28.579</td>
<td>33.093</td>
<td>0.045</td>
<td>0.964</td>
</tr>
<tr>
<td>h5</td>
<td>112</td>
<td>44.000</td>
<td>41.269</td>
<td>37.156</td>
<td>0.448</td>
<td>0.654</td>
</tr>
<tr>
<td>h6</td>
<td>76</td>
<td>20.103</td>
<td>16.635</td>
<td>22.980</td>
<td>0.724</td>
<td>0.469</td>
</tr>
<tr>
<td>h7</td>
<td>57</td>
<td>13.377</td>
<td>13.234</td>
<td>20.769</td>
<td>0.031</td>
<td>0.975</td>
</tr>
<tr>
<td>h8</td>
<td>18</td>
<td>0.364</td>
<td>5.476</td>
<td>7.912</td>
<td>-1.817</td>
<td>0.069</td>
</tr>
<tr>
<td>h9</td>
<td>10</td>
<td>-0.178</td>
<td>0.898</td>
<td>1.261</td>
<td>-0.959</td>
<td>0.338</td>
</tr>
<tr>
<td>h10</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h22</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h23</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Table A57: Frequencies of possible haplotypes generated for the Heme oxygenase-1 gene by the FBAT program

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>Estimates of frequency:</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>211211</td>
</tr>
<tr>
<td>h2</td>
<td>112221</td>
</tr>
<tr>
<td>h3</td>
<td>211221</td>
</tr>
<tr>
<td>h4</td>
<td>212221</td>
</tr>
<tr>
<td>h5</td>
<td>221211</td>
</tr>
<tr>
<td>h6</td>
<td>212122</td>
</tr>
<tr>
<td>h7</td>
<td>212121</td>
</tr>
<tr>
<td>h8</td>
<td>211212</td>
</tr>
<tr>
<td>h9</td>
<td>111211</td>
</tr>
<tr>
<td>h10</td>
<td>111221</td>
</tr>
<tr>
<td>h11</td>
<td>112211</td>
</tr>
<tr>
<td>h12</td>
<td>211222</td>
</tr>
<tr>
<td>h13</td>
<td>221221</td>
</tr>
<tr>
<td>h14</td>
<td>112122</td>
</tr>
<tr>
<td>h15</td>
<td>212222</td>
</tr>
<tr>
<td>h16</td>
<td>112222</td>
</tr>
<tr>
<td>h17</td>
<td>122221</td>
</tr>
<tr>
<td>h18</td>
<td>212211</td>
</tr>
<tr>
<td>h19</td>
<td>221212</td>
</tr>
<tr>
<td>h20</td>
<td>121211</td>
</tr>
<tr>
<td>h21</td>
<td>111222</td>
</tr>
</tbody>
</table>
Table A58: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Heme oxygenase-1 and age of diagnosis by the FBAT program. Age of diagnosis was logarithmically transformed for normality

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>S</th>
<th>E(S)</th>
<th>Var(S)</th>
<th>Z</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>255</td>
<td>-41.328</td>
<td>-49.460</td>
<td>71.522</td>
<td>0.962</td>
<td>0.336</td>
</tr>
<tr>
<td>h2</td>
<td>221</td>
<td>-22.466</td>
<td>-30.480</td>
<td>58.029</td>
<td>1.052</td>
<td>0.293</td>
</tr>
<tr>
<td>h3</td>
<td>124</td>
<td>-13.264</td>
<td>-14.073</td>
<td>23.826</td>
<td>0.166</td>
<td>0.868</td>
</tr>
<tr>
<td>h4</td>
<td>83</td>
<td>-11.206</td>
<td>-8.905</td>
<td>17.249</td>
<td>-0.554</td>
<td>0.580</td>
</tr>
<tr>
<td>h5</td>
<td>52</td>
<td>-16.957</td>
<td>-12.072</td>
<td>12.119</td>
<td>-1.403</td>
<td>0.161</td>
</tr>
<tr>
<td>h6</td>
<td>49</td>
<td>-5.699</td>
<td>-2.119</td>
<td>8.887</td>
<td>-1.201</td>
<td>0.230</td>
</tr>
<tr>
<td>h7</td>
<td>18</td>
<td>-8.772</td>
<td>-5.713</td>
<td>3.132</td>
<td>-1.728</td>
<td>0.084</td>
</tr>
<tr>
<td>h8</td>
<td>10</td>
<td>-1.623</td>
<td>-2.611</td>
<td>2.088</td>
<td>0.684</td>
<td>0.494</td>
</tr>
<tr>
<td>h9</td>
<td>9</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Table A59: Haplotype analysis for investigation of correlation between combinations of selected SNPs in Heme oxygenase-1 and FEV1 standard deviation value by the FBAT program

<table>
<thead>
<tr>
<th>Haplotype</th>
<th># of family</th>
<th>$S$</th>
<th>E($S$)</th>
<th>Var($S$)</th>
<th>Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>233</td>
<td>115.283</td>
<td>105.434</td>
<td>89.373</td>
<td>1.042</td>
<td>0.298</td>
</tr>
<tr>
<td>h2</td>
<td>198</td>
<td>75.282</td>
<td>76.587</td>
<td>69.100</td>
<td>-0.157</td>
<td>0.875</td>
</tr>
<tr>
<td>h3</td>
<td>114</td>
<td>17.461</td>
<td>26.854</td>
<td>33.189</td>
<td>-1.630</td>
<td>0.103</td>
</tr>
<tr>
<td>h4</td>
<td>69</td>
<td>14.362</td>
<td>14.004</td>
<td>14.980</td>
<td>0.092</td>
<td>0.926</td>
</tr>
<tr>
<td>h5</td>
<td>44</td>
<td>19.581</td>
<td>15.284</td>
<td>12.130</td>
<td>1.234</td>
<td>0.217</td>
</tr>
<tr>
<td>h6</td>
<td>47</td>
<td>7.874</td>
<td>10.227</td>
<td>12.511</td>
<td>-0.665</td>
<td>0.506</td>
</tr>
<tr>
<td>h7</td>
<td>15</td>
<td>1.953</td>
<td>2.011</td>
<td>1.592</td>
<td>-0.046</td>
<td>0.963</td>
</tr>
<tr>
<td>h8</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h9</td>
<td>8</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h10</td>
<td>7</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h11</td>
<td>3</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h12</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h13</td>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h14</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h15</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h16</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h17</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h18</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h19</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h20</td>
<td>1</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>h21</td>
<td>0</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table A60: Haplotype analysis between the haplotypes formed by the five selected SNPs in the Factor B gene and age of Pseudomonas aeruginosa infection by Hapstat, in order to determine any association between the haplotypes and the phenotype

<table>
<thead>
<tr>
<th></th>
<th>Estimated value</th>
<th>Standard Error</th>
<th>Z score</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000</td>
<td>-0.306</td>
<td>0.196</td>
<td>-1.549</td>
<td>0.121</td>
</tr>
<tr>
<td>11001</td>
<td>0.014</td>
<td>0.179</td>
<td>0.076</td>
<td>0.939</td>
</tr>
<tr>
<td>11000</td>
<td>-0.174</td>
<td>0.229</td>
<td>-0.759</td>
<td>0.448</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.117</td>
<td>0.171</td>
<td>-0.687</td>
<td>0.492</td>
</tr>
<tr>
<td>CFTR genotype</td>
<td>0.089</td>
<td>0.120</td>
<td>0.739</td>
<td>0.460</td>
</tr>
</tbody>
</table>
Table A61: Haplotype analysis between the haplotypes formed by the four selected SNPs in the Complement factor 3 gene and age of *Pseudomonas aeruginosa* infection by Hapstat, in order to determine any association between the haplotypes and the phenotype

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimated value</th>
<th>Standard Error</th>
<th>Z score</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>-0.341</td>
<td>0.317</td>
<td>-1.077</td>
<td>0.282</td>
</tr>
<tr>
<td>0010</td>
<td>-0.401</td>
<td>0.238</td>
<td>-1.687</td>
<td>0.092</td>
</tr>
<tr>
<td>0011</td>
<td>-0.264</td>
<td>0.268</td>
<td>-0.983</td>
<td>0.326</td>
</tr>
<tr>
<td>0100</td>
<td>-0.601</td>
<td>0.383</td>
<td>-1.571</td>
<td>0.116</td>
</tr>
<tr>
<td>0101</td>
<td>-0.506</td>
<td>0.385</td>
<td>-1.315</td>
<td>0.188</td>
</tr>
<tr>
<td>1010</td>
<td>-0.526</td>
<td>0.487</td>
<td>-1.078</td>
<td>0.281</td>
</tr>
<tr>
<td>1101</td>
<td>-0.070</td>
<td>0.323</td>
<td>-0.216</td>
<td>0.829</td>
</tr>
<tr>
<td>1110</td>
<td>-0.474</td>
<td>0.360</td>
<td>-1.317</td>
<td>0.188</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.090</td>
<td>0.184</td>
<td>-0.489</td>
<td>0.625</td>
</tr>
<tr>
<td>CFTR genotype</td>
<td>0.0589</td>
<td>0.1305</td>
<td>0.4512</td>
<td>0.6518</td>
</tr>
</tbody>
</table>

Table A62: Haplotype analysis between the haplotypes formed by the seven selected SNPs in the Toll-like receptor 4 gene and age of *Pseudomonas aeruginosa* infection by Hapstat, in order to determine any association between the haplotypes and the phenotype

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Estimated value</th>
<th>Standard Error</th>
<th>Z score</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0011001</td>
<td>0.133</td>
<td>0.198</td>
<td>0.670</td>
<td>0.503</td>
</tr>
<tr>
<td>0011011</td>
<td>-0.070</td>
<td>0.160</td>
<td>-0.435</td>
<td>0.664</td>
</tr>
<tr>
<td>0101111</td>
<td>-0.080</td>
<td>0.236</td>
<td>-0.340</td>
<td>0.734</td>
</tr>
<tr>
<td>0110110</td>
<td>0.089</td>
<td>0.249</td>
<td>0.358</td>
<td>0.720</td>
</tr>
<tr>
<td>0111111</td>
<td>0.096</td>
<td>0.209</td>
<td>0.457</td>
<td>0.648</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.188</td>
<td>0.178</td>
<td>-1.060</td>
<td>0.289</td>
</tr>
<tr>
<td>CFTR genotype</td>
<td>0.169</td>
<td>0.134</td>
<td>1.261</td>
<td>0.207</td>
</tr>
</tbody>
</table>
Table A63: Haplotype analysis between the haplotypes formed by the six selected SNPs in the Heme oxygenase-1 gene and age of Pseudomonas aeruginosa infection by Hapstat, in order to determine any association between the haplotypes and the phenotype

<table>
<thead>
<tr>
<th></th>
<th>Estimated value</th>
<th>Standard Error</th>
<th>Z score</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>001110</td>
<td>0.072</td>
<td>0.156</td>
<td>0.464</td>
<td>0.643</td>
</tr>
<tr>
<td>100110</td>
<td>0.034</td>
<td>0.220</td>
<td>0.156</td>
<td>0.876</td>
</tr>
<tr>
<td>101011</td>
<td>0.363</td>
<td>0.271</td>
<td>1.338</td>
<td>0.181</td>
</tr>
<tr>
<td>110100</td>
<td>0.173</td>
<td>0.265</td>
<td>0.653</td>
<td>0.514</td>
</tr>
<tr>
<td>Sex</td>
<td>0.171</td>
<td>0.182</td>
<td>0.943</td>
<td>0.346</td>
</tr>
<tr>
<td>CFTR genotype</td>
<td>0.107</td>
<td>0.129</td>
<td>0.827</td>
<td>0.409</td>
</tr>
</tbody>
</table>