Although cystic fibrosis (CF) is a monogenic disease, its clinical manifestations are influenced in a complex manner. Severity of lung disease, the main cause of mortality among CF patients, is likely modulated by several genes. The mannose-binding lectin 2 (MBL2) gene encodes an innate immune response protein and has been implicated as a pulmonary modifier in CF. However, reports have been conflicting, and interactions with other modifiers have not been investigated. We therefore evaluated the association of MBL2 with CF pulmonary phenotype in a cohort of 1,019 Canadian pediatric CF patients. MBL2 genotypes were combined into low-, intermediate-, and high-expression groups based on MBL2 levels in plasma. Analysis of age at first infection with *Pseudomonas aeruginosa* demonstrated that MBL2 deficiency was significantly associated with earlier onset of infection. This MBL2 effect was amplified in patients with high-producing genotypes of transforming growth factor beta 1 (TGFB1). Similarly, MBL2 deficiency was associated with more rapid decline of pulmonary function, most significantly in those carrying the high-producing TGFB1 genotype. These findings provide evidence of gene-gene interaction in the pathogenesis of CF lung disease, whereby high TGF-β1 production enhances the modulatory effect of MBL2 on the age of first bacterial infection and the rate of decline of pulmonary function.
Complex two-gene modulation of lung disease severity in children with cystic fibrosis

Ruslan Dorfman,1 Andrew Sandford,2 Chelsea Taylor,3 Baisong Huang,3 Daisy Frangiolis,2 Yongqian Wang,1 Richard Sang,1 Lilian Pereira,1 Lei Sun,4 Yves Berthiaume,5 Lap-Chee Tsui,1 Peter D. Paré,2 Peter Durie,6 Mary Corey,3 and Julian Zielenki1

1Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada. 2The James Hogg ICAPTURE Centre for Cardiovascular and Pulmonary Research, University of British Columbia, Vancouver, British Columbia, Canada. 3Child Health Evaluative Sciences Program, Hospital for Sick Children, Toronto, Ontario, Canada. 4Department of Public Health Sciences and Department of Statistics, University of Toronto, Toronto, Ontario, Canada. 5Department of Medicine, Centre de Recherche Centre Hospitalier de l’Université de Montréal, Montreal, Quebec, Canada. 6Program in Physiology and Experimental Medicine, Hospital for Sick Children, Toronto, Ontario, Canada.

Although cystic fibrosis (CF) is a monogenic disease, its clinical manifestations are influenced in a complex manner. Severity of lung disease, the main cause of mortality among CF patients, is likely modulated by several genes. The mannose-binding lectin 2 (MBL2) gene encodes an innate immune response protein and has been implicated as a pulmonary modifier in CF. However, reports have been conflicting, and interactions with other modifiers have not been investigated. We therefore evaluated the association of MBL2 with CF pulmonary phenotype in a cohort of 1,019 Canadian pediatric CF patients. MBL2 genotypes were combined into low-, intermediate-, and high-expression groups based on MBL2 levels in plasma. Analysis of age at first infection with Pseudomonas aeruginosa demonstrated that MBL2 deficiency was significantly associated with earlier onset of infection. This MBL2 effect was amplified in patients with high-producing genotypes of transforming growth factor beta 1 (TGFβ1). Similarly, MBL2 deficiency was associated with more rapid decline of pulmonary function, most significantly in those carrying the high-producing TGFβ1 genotype. These findings provide evidence of gene-gene interaction in the pathogenesis of CF lung disease, whereby high TGF-β1 production enhances the modulatory effect of MBL2 on the age of first bacterial infection and the rate of decline of pulmonary function.

Introduction
Cystic fibrosis (CF) is an autosomal-recessive disease typically manifesting with progressive obstructive lung disease, insufficiency of the exocrine pancreas, and elevated electrolyte concentration in sweat glands (OMIM 219700). Although CF is a monogenic disease caused by mutations in the CF transmembrane conductance regulator (CFTR) gene, genotype-phenotype correlation studies of some manifestations suggest more complex inheritance, involving other genetic and environmental factors (1, 2). In particular, the severity of pulmonary disease, which is the major cause of morbidity and mortality in CF, is poorly correlated with CFTR genotype. Patients carrying the same CFTR mutations show extremely variable lung phenotypes at all ages. It is postulated that the severity and progression of pulmonary disease in CF is modulated by secondary genetic factors called CF modifiers (1, 3).

The hallmark of CF pulmonary disease is chronic infection with characteristic pathogens such as Pseudomonas aeruginosa, Staphylococcus aureus, and Burkholderia cepacia, which leads to chronic inflammatory damage to lung tissue and progressive loss of pulmonary function. Genetic factors that influence susceptibility to infection are therefore of major interest. One of the first genes implicated as a pulmonary modifier in CF was the mannose-binding lectin 2 (MBL2) gene (4). MBL2 produces a protein of the collectin family, a key factor in the innate immune response (5, 6). MBL2 selectively binds D-mannose and N-acetyl-D-glucosamine (GlcNAc) on bacte-
tion between MBL2 variants and pulmonary function (25). The Drumm study was specifically designed to include age-specific extremes of pulmonary function, and patients under 8 years of age were excluded because of inability to define the severity of pulmonary disease. This precluded the ability to test for potential effects of the MBL2 polymorphisms at an early stage of CF lung disease. TGF-β1 is a growth factor that plays a critical role in inflammatory processes likely to have pleiotropic effects in the lung. It prevents spontaneous activation of alveolar macrophages and secretion of proinflammatory cytokines and also causes inflammation, apoptosis, remodeling, and fibrosis of lung tissue, but does not affect the efficiency of phagocytosis (26–30). Several transforming growth factor beta 1 (TGFB1) gene variants were previously shown to affect the level of secreted protein (31, 32) and were reported to affect severity of lung disease in CF patients (25). Excessive tissue fibrosis in CF lungs was proposed as the basis underlying the high lung CF-related diabetes 2.8% 2.9% 2.7% 0.85
CF liver disease 4.4% 3.8% 5.4% 0.22
Meconium ileus 21.2% 22.3% 19.6% 0.35
Median age of diagnosis (yr) 0.36 0.35 0.38 0.35
Mean FEV1 % predicted 84.1% 83.7% 84.8% 0.50
MBL2 gene variants and genotypes. Four MBL2 gene variants (Table 2) were typed in 1,393 CF patients and 600 randomly selected parents from the entire CF Modifier Study cohort (see Methods). The distributions of these variants and related genotypes and diplotype groups in patients and parents are shown in Supplemental Table 2 (supplemental material available online with this article; doi:10.1172/JCI33754). The overall allelic and genotypic frequencies were similar in the patient and parent cohorts and were not different from the previously published data (4, 11). We then established 6 major diplotype groups from family members’ genotypes at 4 variant sites (Table 3); their frequencies were similar in patients and parents (Supplemental Table 2).

Table 1
Demographic and clinical characteristics of CF patients (age <18.5 yr) with PI

<table>
<thead>
<tr>
<th></th>
<th>All patients</th>
<th>By CFTR mutation</th>
<th>P²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔF508/ΔF508</td>
<td>Other PI</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>1,019</td>
<td>611</td>
<td>408</td>
</tr>
<tr>
<td>Mean age (yr)</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>Median age of diagnosis (yr)</td>
<td>0.36</td>
<td>0.35</td>
<td>0.38</td>
</tr>
<tr>
<td>Meconium ileus</td>
<td>21.2%</td>
<td>22.3%</td>
<td>19.6%</td>
</tr>
<tr>
<td>CF liver disease</td>
<td>4.4%</td>
<td>3.8%</td>
<td>5.4%</td>
</tr>
<tr>
<td>CF-related diabetes</td>
<td>2.8%</td>
<td>2.9%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Mean FEV1 % predicted</td>
<td>84.1%</td>
<td>83.7%</td>
<td>84.8%</td>
</tr>
<tr>
<td>P. aeruginosa positive</td>
<td>69.0%</td>
<td>67.3%</td>
<td>71.6%</td>
</tr>
<tr>
<td>Median age of first infection (yr)</td>
<td>7.2</td>
<td>7.5</td>
<td>6.9</td>
</tr>
<tr>
<td>B. cepacia positive</td>
<td>5.5%</td>
<td>5.9%</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

⁻CFTR mutations for patients in this cohort are listed in Supplemental Table 4. ¹ΔF508/ΔF508 versus other PI; calculated by Student t test or Fisher exact test except where noted. ²Includes ΔF508/other and other/other severe genotypes. ³Except mean FEV1, % predicted (n = 801 total patients, 483 ΔF508/ΔF508, 318 other PI). ⁴Wilcoxon rank-sum test. ⁵Log-rank test of time to event.

Results

Demographic and clinical characteristics of the pediatric CF patient cohort. The cohort of 1,019 CF pediatric patients is summarized in Table 1. The largest group of CF patients sharing the same CFTR genotype consists of those homozygous for the ΔF508 mutation (35). This severely affected CF group serves as a clinical reference for patients with other CFTR genotypes (36, 37). We compared the demographic and clinical parameters in CF patients homozygous for ΔF508 with the rest of the pancreatic insufficiency (PI) group (Table 1). None of the demographic and clinical parameters were significantly different between the 2 groups.
Table 2

**MBL2** gene variants and notation in the current study

<table>
<thead>
<tr>
<th>Variant</th>
<th>Location</th>
<th>Nucleotide variant</th>
<th>Amino acid change</th>
<th>rs no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y⁵⁺</td>
<td>Promoter</td>
<td>–221G</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>X</td>
<td>Promoter</td>
<td>–221C</td>
<td>–</td>
<td>7096206</td>
</tr>
<tr>
<td>A⁴⁺</td>
<td>Exon 1</td>
<td>154C, 161G, 170G</td>
<td>Arg52, Gly54, Gly57</td>
<td>–</td>
</tr>
<tr>
<td>B⁺⁺</td>
<td>Exon 1</td>
<td>164T</td>
<td>Cys52</td>
<td>5030737</td>
</tr>
<tr>
<td>C⁺⁺</td>
<td>Exon 1</td>
<td>161A</td>
<td>Asp54</td>
<td>1800450</td>
</tr>
<tr>
<td>D⁺⁺</td>
<td>Exon 1</td>
<td>170A</td>
<td>Glu 57</td>
<td>1800451</td>
</tr>
</tbody>
</table>

⁵⁺Wild type. ⁺⁺Due to the same structural effect on the **MBL2** protein, the B, C, and D variants were combined as the 0 variant group.

**MBL2** expression in **CF** patients and parents. Nonparametric analysis of the median normalized **MBL2** expression between **CF** patients and the parent (control) group demonstrated that **CF** patients had significantly higher levels of **MBL2** expression (Figure 1C and Supplemental Table 3). The higher **MBL2** expression in **CF** patients was most likely associated with induction of **MBL2** due to the persistent bacterial challenge and mobilization of innate defense mechanisms. The lack of significant elevation in **MBL2** expression in the ****00 diplotype almost certainly resulted from the structural mutations that prevent oligomerization of the protein which in turn leads to an inability to detect the **MBL2** oligomer in the assay system.

**MBL2** expression in the pediatric and adult cohorts. To investigate a potential effect of age on **MBL2** expression, the **CF** patients were divided into pediatric (age <18.5 yr; n = 848) and adult (age ≥18.5 yr; n = 545) groups. The pediatric group had significantly higher plasma levels of **MBL2** than the adult group for all diplotype except ****00 (Figure 1D).

**MBL2** and age of first *P. aeruginosa* infection. Figure 2 shows the Kaplan-Meier curves for the 3 **MBL2** diplotypes in the pediatric cohort. The median age of first *P. aeruginosa* infection was significantly different for low, intermediate, and high **MBL2** groups (4.4, 7.0, and 8.0 yr, respectively; *P* = 0.0003), indicating that **MBL2** deficiency is associated with significantly earlier onset of *P. aeruginosa* infection.

Polymorphisms in the **TGFB1** gene influence age of onset of first *P. aeruginosa* infection in **MBL2** deficient **CF** patients. High-expressing variants of **TGFB1** have been shown to be associated with a more severe lung disease phenotype (25). To investigate whether this could be explained by early *P. aeruginosa* infection, we analyzed age of first infection for the **TGFB1** codon 10 (rs1982073) genotypes in the pediatric **CF** cohort. There was no difference in the age of first *P. aeruginosa* infection between the genotypes (*P* = 0.7; Supplemental Figure 1).

Analysis of the **MBL2** diplotype groups stratified by the **TGFB1** codon 10 genotypes demonstrated an influence of the high-expressing **TGFB1** allele C on age of first *P. aeruginosa* infection among the 3 **MBL2** diplotype groups. No effect of **MBL2** diplotype on age of first *P. aeruginosa* infection was observed in the subset of patients carrying the low-expressing **TGFB1** genotype TT (*P* = 0.16; Figure 3A). However, in the subjects homozygous for the high-expressing **TGFB1** C allele (CC genotype), the differences in median age of first *P. aeruginosa* infection by **MBL2** diplotype groups were magnified (low **MBL2**, 3.1 yr; intermediate **MBL2**, 6.7 yr; high **MBL2**, 8.7 yr; *P* = 0.01; Figure 3C). In patients with low levels of **MBL2**, the **CC** genotype was associated with the earliest infection age, while the **TT** genotype was associated with the oldest age at first infection. The **TGFB1** variant (rs2241715) exhibited a similar result as that observed for the codon 10 variant (data not shown). Although this observation suggested the possibility of gene-gene interaction between **MBL2** and **TGFB1**, testing by the Cox proportional hazards regression model failed to detect such interaction (*P* = 0.2 for the interaction term **MBL2**×**TGFB1**).

**MBL2** and **MBL2**/**TGFB1** modulation of pulmonary function in **CF** patients. Of the patients studied, 511 were old enough to have multiple pulmonary function measurements in the 3 years before study enrollment. In a mixed regression model of forced expiratory volume in 1 s as a percentage of the predicted value (FEV₁ % predicted; see Methods) versus age, **MBL2**, **TGFB1**, and all interaction terms, the 3-way interaction was significant (*P* = 0.002), indicating that the slope of FEV₁ versus age was affected by **MBL2** and **TGFB1** genotypes. Table 3 shows how the interaction of these modifiers affected lung function. Rate of decline of FEV₁ was significantly steeper in the **MBL2** diplotype group associated with low expression (*P* = 0.05). When stratified by **TGFB1** genotype, the significant effect of **MBL2** on FEV₁ was only seen in the high-expressing **CC** genotype group (*P* = 0.0002). The intercept at age 10 was relatively stable across all genotype combinations, except in the high **TGFB1**/low **MBL2** group, where it was significantly lower. Analysis of **TGFB1** codon 10 alone showed a significant effect on lung function, consistent with earlier reports. The rate of FEV₁ decline was –2.62% (95% CI, –3.29% to –1.95%), –1.22% (95% CI, –1.61% to –0.83%), and –2.04% (95% CI, –2.45% to –1.62%) per year in CC, CT, and TT genotypes, respectively (*P* = 0.0005), while the respective intercepts were 88.5%, 87.5%, and 86.3% (*P* = 0.7).

Taken together, these data suggest that the effect of **MBL2** expression is a critical determinant of the severity of lung disease at a young age in patients with **CF** disease and that high-expressing variants of **TGFB1** enhance this effect. Age at first infection was a significant predictor of FEV₁ intercept at age 10 (age <6 yr,

![Table 3](http://www.jci.org)
82.2%, n = 161; age ≥6 yr, 89.3%, n = 350; P = 0.0001) but not of rate of decline (−1.95 and −1.72, respectively; P = 0.4). When infection age was added to the models of FEV₁ versus MBL2 or TGFB1, both infection age and genotype remained significant predictors of lung function (data not shown).

Effect of gender on age at first infection and lung function. The age at first *P. aeruginosa* infection was significantly lower in females than in males (median age, 6.7 versus 7.9 years; P = 0.04, log-rank test). However there was no interaction between gender and MBL2 diplotype group (P = 0.62) or TGFB1 codon 10 (P = 0.98) in a Cox regression model.

Consistent results were seen for lung function. Females showed a more rapid mean rate of decline in FEV₁ than males (−1.45% versus −1.08% predicted per year; P < 0.0001), but there was a similar intercept at age 10 years (87.3% versus 86.7%; P = 0.73). Sex and the 3-way interaction of sex, age, and genotype were added to mixed models of FEV₁ based on MBL2 or TGFB1, and there was no suggestion that the effects of either were different for males and females. The 3-way interaction terms were not significant in the MBL2 model (P = 0.80) or in the codon 10 model (P = 0.41). The effect of gender on the interaction between TGFB1 codon 10 and MBL2 diplotype groups for lung function decline could not be tested. This regression model involves a 4-way interaction and all possible 2- and 3-way interactions, and the number of patients in the smallest categories of sex, MBL2 group, and TGFB1 group were insufficient.

Table 4

<table>
<thead>
<tr>
<th>Diplotype combination</th>
<th>Diplohyte group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y0/Y0, X0/Y0, and XA/Y0</td>
<td>Low MBL2</td>
</tr>
<tr>
<td>YA/Y0 and XA/XA</td>
<td>Intermediate MBL2</td>
</tr>
<tr>
<td>XA/YA and YA/YA</td>
<td>High MBL2</td>
</tr>
</tbody>
</table>

Figure 1

Normalized MBL2 levels corresponding to 6 MBL2 gene diplotypes (**00, XYA0, YYA0, XXAA, XYAA, and YYAA) in CF patient and parent cohorts. (A and B) Distribution of normalized MBL2 expression values and median MBL2 expression values for each diplotype is shown for 600 parents (A) and 1,393 PI CF patients (B). (C and D) Comparison of normalized MBL2 expression values between patients and parents (C) and between 848 pediatric and 545 adult patients (D). Statistical difference was tested using the Wilcoxon rank-sum test. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.
MBL2 and age at first infection in ΔF508 homozygotes. Table 1 demonstrates that lung function and age at first infection were similar in patients who were homozygous for the ΔF508 mutation and other PI patients. The effect of MBL2 on age at first infection was computed for this subset homozygous for ΔF508, and the results were consistent with the main analysis. The median age at first infection was 3.9 yr (95% CI, 2.9 to 9.7 yr), 7.4 yr (95% CI, 6.1 to 7.9 yr), and 8.1 yr (95% CI, 7.2 to 8.8 yr) for the high MBL2 group. The survival probability differences were tested using log-rank test (P = 0.0003).

Figure 2
Results of time-to-event analysis showing median age and 95% CI at first positive culture for P. aeruginosa. The proportion of patients not infected with P. aeruginosa versus age is plotted for low MBL2 (**00 and XYA0; n = 136), intermediate MBL2 (YYA0 and XXAA; n = 311), and high MBL2 (YYAA and XYAA; n = 572) diplotype groups. The graph shows data for 1,019 pediatric (<18.5 years of age) PI CF patients. Median age of infection was 4.4 yr (95% CI, 3.2 to 6.4 yr) for the low MBL2 group, 7.0 yr (95% CI, 6.1 to 7.9 yr) for the intermediate MBL2 group, and 8.0 yr (95% CI, 7.2 to 8.8 yr) for the high MBL2 group. The survival probability differences were tested using log-rank test (P = 0.01, log-rank test). No further analysis was conducted on the subset of ΔF508 homozygotes because of the small subgroup sizes.

Figure 3
Results of time-to-event analysis (as in Figure 2) plotted for the 3 MBL2 diplotype groups and stratified by the codon 10 (rs1982073, L10P) TGFB1 genotypes. Data are shown for 1,003 pediatric PI CF patients; no codon 10 results were available for 16 patients. (A) For the TT genotype (n = 389), median age at first positive P. aeruginosa culture was 5.7 yr (95% CI, 2.4 to 9.7 yr) for the low MBL2 group (n = 56), 7.1 yr (95% CI, 5.9 to 8.5 yr) for the intermediate MBL2 group (n = 115), and 7.4 yr (95% CI, 6.1 to 9.1 yr) for the high MBL2 group (n = 218). (B) For the CT genotype (n = 454), median age at first positive P. aeruginosa culture was 4.6 yr (95% CI, 3.5 to 9.7 yr) for the low MBL2 group (n = 53), 7.4 yr (95% CI, 5.9 to 8.9 yr) for the intermediate MBL2 group (n = 155), and 8.1 yr (95% CI, 7.1 to 9.4 yr) for the high MBL2 group (n = 246). (C) For the CC genotype (n = 160), median age at first positive P. aeruginosa culture was 3.1 yr (95% CI, 2.8 to 6.2 yr) for the low MBL2 group (n = 26), 6.7 yr (95% CI, 3.4 to 9.1 yr) for the intermediate MBL2 group (n = 34), and 8.7 yr (95% CI, 7.1 to 10.4 yr) for the high MBL2 group (n = 100). Differences between groups were determined by the log-rank test.
The patient cohort was recruited from a population-based sample representing nearly 75% of the current Canadian CF population (34). The resource consists of DNA and plasma samples from CF patients and parents as well as demographic and clinical data from patients. Comparison with the Canadian CF Patient Registry demonstrates that the demographic and clinical data in the recruited patient cohort were nationally representative (34). Consequently, the results obtained from these cohorts can be generalized to the entire Canadian CF population. Treatment and survival of CF patients has been steadily improving over the 2 decades in which the study patients were born. All analyses in this report were repeated, including a binary birth cohort variable. In no case did birth cohort affect the results of MBL2 and TGFB1 genotypes (data not shown).

Secondary modulation of the severity of CF lung disease can be confounded by CFTR genotype, the primary determinant of disease severity. Patients carrying at least 1 mild CFTR mutation tend to have a milder form of CF (i.e., almost exclusively pancreatic sufficiency). Pulmonary disease is somewhat milder, although highly variable, than in patients carrying severe mutations on both alleles (due to lack of CFTR protein synthesis, processing, or function). Therefore, to avoid a potential confounding effect of CFTR genotypes, only pancreatic insufficient patients who had verified severe CFTR mutations on both alleles were included in this study. The demographic and clinical characteristics of the study population (Table 1) demonstrate that the subset of patients homozygous for the AF508 mutation were remarkably similar to the cohort patients carrying other severe CFTR genotypes, justifying the combination of the 2 patient groups as the study population for our analyses.

To our knowledge, this is the largest and most comprehensive study of MBL2 as a CF pulmonary disease modifier. We verified previous findings on the relationship between genotypes of the key MBL2 variants and expression of the MBL2 protein in peripheral blood (4, 12, 19). We also classified MBL2 diplotypes on the basis of plasma protein expression of oligomerized MBL2 and genotyped for the variants affecting production or assembly of
MBL2 protein (Table 2). Our results (Figure 1, A and B) confirmed earlier reports (4, 12, 19) showing a correlation between the 6 MBL2 diplotype and corresponding MBL2 expression levels. It should be noted, however, that there was considerable variation of MBL2 expression between subjects in each diplotype cohort. Other unspecified factors, such as unrecognized MBL2 mutations and modulation of MBL2 expression by disease factors, could account for the variable MBL2 expression levels (41).

Carlsson et al. (23) reported an increased plasma level of MBL2 in CF patients. The results of the present study not only confirmed this observation, but also demonstrated that pediatric CF patients had significantly higher MBL2 plasma levels than adults. The higher MBL2 expression in pediatric CF patients could be attributed to an enhanced acute inflammatory response due to repeated bacterial challenge, which is not observed in older patients who have long-established chronic sinopulmonary infections.

In a previous study, Garred et al. (4) demonstrated an earlier age of onset of chronic P. aeruginosa infection in CF patients as a result of MBL2 deficiency. Because it is difficult to reliably establish the onset of chronic P. aeruginosa infection, we used age at the first positive respiratory culture as a surrogate. The time-to-event analysis showed that MBL2 deficiency was strongly associated with earlier acquisition of P. aeruginosa in young children with CF (P = 0.0003; Figure 2). This modulatory effect most likely reflects the critical role of MBL2 in the first defense against bacterial colonization (16) and supports the similar but less significant effect previously observed in a small Italian CF patient cohort (42). However, it should also be noted that although MBL2 deficiency has been implicated in many studies as contributing to higher susceptibility to infections, there are also data suggesting the opposite (43). This MBL2 duality may account for the relatively high frequency of MBL2 variants encoding low MBL serum concentrations in some populations and should be seen in the context of specific circumstances of bacterial challenge and infection (43).

Notably, the MBL2-deficient patients carrying the high-expressing TGFβ1 genotype CC had the most significant earlier onset of P. aeruginosa acquisition. Interestingly, although TGFβ1 affected the modulation of age of onset by MBL2 in the pediatric patients, no direct impact of its codon 10 genotypes was observed in the time-to-event analysis using the same cohort unstratified for MBL2 (Supplemental Figure 1). This modulation of the MBL2 effect by TGFβ1 could result from impaired bacterial eradication and clearance in airways due to macrophage inactivation by elevated levels of TGFβ1 (44, 45). This, in conjunction with the MBL2 deficiency, would lead to the enhanced vulnerability of young CF patients to infection, manifesting as significantly earlier P. aeruginosa acquisition. This is in agreement with the well-documented dual role of TGFβ1 as a pro- and antiinflammatory cytokine in the immune defense system (46, 47).

Because CF children with MBL2 deficiency acquire P. aeruginosa pulmonary infections at a significantly younger age than those with normal levels, one might expect that MBL2-deficient children would have faster decline of pulmonary function. Indeed, as shown in Table 5, we demonstrated that MBL2 deficiency correlated with a more rapid rate of decline of pulmonary function. The observed effect could be explained by the recognized role of MBL2 in immunity. Because MBL2 deficiency prevents the activation of the MBL2-associated complement pathway, there is failure of the primary antibacterial defense, namely macrophage phagocytosis of bacteria. In this regard, MBL2 is not only involved in innate antibacterial defense through the complement pathway, it is also a potent regulator of inflammatory pathways (9), which are highly dependent on the concentration of circulating MBL2. Low plasma MBL2 levels induce monocyte production of proinflammatory cytokines, whereas high MBL2 levels have a strong antiinflammatory effect (9). This observation provides a potential explanation for the magnified modulatory effect on the severity of lung disease in the MBL2-deficient children homozygous for the high-producing TGFβ1 codon 10 variant C (P = 0.0002). Specifically, MBL2 deficiency (<4 μg/ml) has been shown to stimulate release of proinflammatory cytokines (TNF-α, IL-1β, and IL-6) from monocytes, whereas cytokine production is greatly suppressed with MBL2 levels higher than 6 μg/ml (9). The context of MBL2-induced high levels of proinflammatory cytokines could support activation of the TGFβ1-stimulated commitment of Th17 (i.e., IL-17-producing CD4+ T cell lineage producing neutrophil-mobilizing cytokines instrumental in the host defense to infections (46–49). In this situation, enhanced production of TGFβ1 by the C allele could further inhibit the host defense by a variety of mechanisms. These include downregulation of macrophage receptors critical for phagocytosis (44, 50), downregulation of IgG receptors on mononuclear cells, and reduction of phagocytosis of IgG-opsonized bacteria (51) as well as other mechanisms critical for antibacterial defenses, such as attenuation of innate immune signaling pathways mediated by TLR4 on LPS-induced macrophages (52).

The 2 clinical variables used in this analysis, namely the onset of lung infection and rate of decline of pulmonary function, have different contributions to CF lung pathophysiology. The first is a causative event, because lung infection is presumed to be an initiating event leading to progressive lung damage. The second is an outcome measure. Although the rate of decline of pulmonary function is clearly influenced by infection status, many other environmental and genetic factors will have protective or deleterious consequences. In this regard, statistical modeling of gene-gene interactions between MBL2 and TGFβ1 revealed strong interaction for lung function (P < 0.002) but not for the age of first P. aeruginosa infection (P = 0.2). Failure to detect MBL2/TGFβ1 interaction for the infection variable may reflect the imprecision in assessment of first P. aeruginosa infection. Nevertheless, these results warrant investigation of the biochemical and physiological mechanisms of interaction between these 2 proteins.

Consideration should be given to the clinical implications of these findings. Genetic testing for MBL2 and TGFβ1 variants may permit early identification of at-risk patients either at diagnosis or through newborn screening. This would allow physicians to closely monitor these subjects and give consideration to more aggressive efforts to prevent early onset of bacterial infections. Knowledge of specific modifier genotypes would also permit more accurate patient stratification for entry into therapeutic clinical trials. However, incorporation of genotyping into clinical trials and practice awaits determination of sensitivity and specificity for Pseudomonas prediction, ideally in a population of patients identified by newborn screening. Finally, recognition of the clinical consequences of MBL2 deficiency may have therapeutic implications. For example, intravenous administration of human MBL2 derived from pooled plasma to an MBL2-deficient CF patient with end-stage lung disease was considered to have some beneficial effects (53). Nevertheless, our findings suggest that success with MBL2 replacement therapy would require that it be administered prior to the onset of chronic bacterial infection. Now that recombinant MBL2 protein is available, there is wide interest in MBL2 supplementation. Trials are in progress...
with recombinant MBL2 protein therapy in MBL2-deficient adults who are at increased risk of infection during chemotherapy or following liver or bone marrow transplantation (54, 55). A therapeutic trial with a recombinant MBL2 protein has also been proposed for MBL2-deficient CF patients (56). More research has to be done to refine the details of MBL2 modulation before seriously considering clinical trials of human recombinant MBL2 replacement therapy to delay or prevent early onset of chronic pulmonary infections in CF disease. However, MBL2-deficient patients may benefit from more frequent microbiological screening in order to detect earlier signs of infection that may respond to current antibiotic treatment.

In conclusion, based on its critical role in innate antibacterial defense, we postulate that MBL2 contributes to the variability of lung disease by modulating the age of acquisition of \textit{P. aeruginosa} in pediatric CF patients. Furthermore, we provide evidence of a complex interaction between MBL2 and TGF-\(\beta\)1 and rate of decline of pulmonary function in young patients with CF, possibly related to a synergistic effect on the onset of acquisition of \textit{P. aeruginosa} infection. Gaining insight into the mechanisms of MBL2 modulation of lung pathophysiology in CF patients and its interaction with other pulmonary modifiers (e.g., TGF-\(\beta\)1) will be critical to explain our findings at a molecular level and provide a basis for developing modifier-based therapies.

\section*{Methods}

\textbf{CF patients and parent cohorts.} CF patients and their family members were recruited from 37 specialized CF clinics across Canada (shown in Supplemental Table S), during a period from 2003 to 2006, as participants in the Canadian CF Modifier Study. The protocols for this study were approved by ethical review boards at the Hospital for Sick Children and all participating institutions. Informed consent was obtained from each individual or his/her guardian. The Canadian CF Modifier Study population consists of 2,441 patients diagnosed with CF between 1951 and 2006 and 3,092 parents. This report of first infection as a phenotype was restricted to pediatric patients (<18.5 years of age at the time of recruitment) in order to capture the most reliable information on first infection and to compute average lung function decline unbiased by mortality selection. The age of first infection was generally not available in the records of patients reported from clinics specializing in adult CF care. CF patients included in this study had PI and carried severe \textit{CFTR} mutations on both alleles. Patients reported as PI but carrying at least 1 \textit{CFTR} mutation typically associated with pancreatic sufficiency were excluded. The demographic and clinical characteristics of the pediatric patient cohort are shown in Table 1.

\textit{Pulmonary phenotypes.} To evaluate MBL2 (and MBL2 stratified by TGF-\(\beta\)1 variants) as a pulmonary CF modifier, 2 lung-related clinical outcome variables were used: age at first \textit{P. aeruginosa} infection and FEV\textsubscript{1}. Patients attending Canadian CF clinics have sputum cultures and lung function testing at every clinic visit, which are routinely scheduled at 3-month intervals. Age at the time of the first positive respiratory culture for \textit{P. aeruginosa} was used as an indicator of susceptibility to infection with this bacterium. FEV\textsubscript{1} (as percent of predicted for sex, height, and age) reflects the progression of CF lung disease and is the best predictor of survival (57). Predicted values were based on the formula of Knudson et al. (58) for those over the age of 10 years and Corey et al. (59) for those 6–10 years of age. FEV\textsubscript{1} measurements in the 3 years prior to enrollment were recorded in order to capture the current status and rate of decline for each patient.

\textit{Plasma and DNA samples.} Individual blood samples were collected in glass tubes containing anticoagulant (Acid Citrate Dextrose, ACD-Vacutainer tubes; BD) and maintained at room temperature. Plasma samples were isolated by centrifugation at 524 g for 10 minutes and immediately frozen and stored at –70°C. Genomic DNA was extracted using the phenol/chloroform procedure (60). The DNA stocks were stored at 4°C. Patients were recruited during the years 2003–2006, and the age in all analyses is based on the date the blood sample was taken.

\textit{Genotyping.} For MBL2, 4 functional MBL2 gene variants were genotyped in genomic DNA samples: the promoter variant at position –221 (\(Y\) or \(X\) variant) and 3 structural mutations in exon 1 (B, C, and D variants), as shown in Table 2. The genotyping was performed using Luminex Bead Array technology (61, 62) implemented for the MBL2 variants as described in Supplemental Methods and Supplemental Table 1. MBL2 diplotypes (Table 3) were inferred from promoter variants X and Y and exon 1 structural variants A, B, C, and D according to Madsen et al. (11).

For \textit{TGF\(\beta\)1}, the codon 10 expression variant (rs192073, L10P) in \textit{TGF\(\beta\)1} was genotyped using TaqMan genotyping assay (63). Genotyping of the promoter variant (–509; rs1800469) was performed using Luminex Bead Array technology (Luminex Corp.).

\textit{Measurement of MBL2 protein plasma levels.} The plasma level of oligomerized MBL2 protein was measured in blood samples from 1,393 CF patients and 600 parents as described in Supplemental Methods. The MBL2 protein levels were normalized to adjust for batch differences using the mean and standard deviation for 10 non-CF control samples included in every batch. Therefore the normalized MBL2 expression level reflects the variation above or below that expected for samples unrelated to CF.

\textit{Statistics.} Patients who had not acquired \textit{P. aeruginosa} infection at the time of observation will likely acquire the infection in the future; therefore, their age at first infection is unknown. To account for this, age at first infection was treated as a right-censored variable, and time-to-event analysis was used to compare age at first infection in different subgroups. Because infection is present for some unknown time before being cultured, the age at first positive culture is a left-censored estimate of true infection age. Left censoring is not considered in standard time-to-event analysis. Restriction of the study subjects to children and to those followed in specialized CF clinics minimizes the unknown time from infection to detection. We plotted Kaplan-Meier survival curves to demonstrate the distribution of age at first \textit{P. aeruginosa} infection, and curves were compared using the log-rank statistic. \textit{P. aeruginosa}–free survival curves were compared for MBL2 genotype groups, for \textit{TGF\(\beta\)1} genotypes, and for MBL2/\textit{TGF\(\beta\)1} combinations. Cox proportional hazards regression was used to test the effects of sex and of multiple genotype groups on age at first infection.

Pulmonary function testing involves volitional breathing maneuvers that are generally not possible in children until 5–7 years of age. Data on current lung function and decline over the past 3 years, represented by FEV\textsubscript{1} % predicted, were compiled for patients that had been tested. Longitudinal pulmonary function patterns (in patients with at least 2 measurements) were analyzed using a random effects mixed-model regression, as described by Schluchter et al. (64), and the mean intercept at 10 years of age and the mean slope were compared in potential modifier categories. The dependent variable in all models was FEV\textsubscript{1} % predicted, and the time variable was age 10 yr. Genotype groups for potential modifiers were included as categorical variables. Interaction terms for modifier categories and age provided a statistical test for category differences in the change of lung function over time. The effects of gender and age at first infection were also added to selected models. Because age at first infection was censored, it was dichotomized into <6 years and ≥6 years, the latter including those not yet infected. All statistical analyses were performed using SAS version 9.1 software. A \(P\) value of 0.05 or less was considered significant.

\section*{Acknowledgments}

This project was supported by Genome Canada through the Ontario Genomics Institute per research agreement 2004-
Ogi-3-05, by the Canadian Cystic Fibrosis Foundation, by the
Ontario Research Fund – Research Excellence Program, and by
donations from the Lloyd Carr-Harrison Foundation. R. Dorfman
is supported by the joint Fellowship of Canadian Institutes of
Health Research and Ontario Women’s Health Council.

The authors thank Nicole Anderson, Jennifer Breeton, Mary Crist-
tofi, Roxanne Rousseau, and Michael Van Spall for their excep-
tional effort in recruiting and ascertaining Canadian CF fami-
lies for the study as well as Gouri Mukiereje, Aditi Master, Fan
Lin, Danuta Markiewicz, Qiuju Huang, Xiao-Wei Yuan, Dorota
Stefanowicz, and Anluan O’Brien for their excellent technical
assistance. The authors express gratitude to all CF patients and
their families for participating in this study. The authors
are indebted to the following individuals from member institutions
of the Canadian Consortium for CF Genetic Studies (see Supple-
mental Table 5) for ascertaining patient data and blood samples
from CF patients and their families: S. Aaron, P. Barrett, B.
Beaurivage, Y. Berthiaume, P. Bigonesse, M. Boland, L. Boucher,
A. Cantin, L. Charette, G. Cote, A. Dale, G. Davidson, K.
Devesceri, R. Dicaire, V. Fauvel, A. Freitag, D.N. Garey, M. Gaul,
W. Gervais, J. Gjerv, F. Gosse, A. Gravelle, B. Habbick, R. Hen-
nesey, S.B. Holmes, J. Hopkins, D. Hughes, M. Jackson, J. Jacob,
A. Jeanneret, P. Kean, W. Kepron, T. Kovessi, V.J. Kumar, L. Lands,
M. LaPerriere, J. Leong, R. Levesque, D. Lougheed, M. Lowe, B.
Lyttle, K. Malhotra, J.E. Marcotte, S. Marsolais, C. Martinez, E.
Matouk, D. McCulloch, R.T. Michael, M. Montgomery, R. Mor-
ris, E.M. Nakielna, F. Paquett, H. Pasterkamp, N. Patterson, L.
Peddler, L. Peterson, N. Petit, C. Piche, M. Plante, H.R. Rabin,
K. Ramlall, F. Raymont, L. Rivard, G. Rivard, M. Rousson, M.
Ruel, J. Salgado, L. Semple, E. Sheppard, F. Simard, A. Smith,
M. Solomon, R. Stackhouse, J. Tabak, L. Taylor, A. Tsang, E. Tullis,
C. Turtle, K. Vandamheen, M. van Spall, R. van Wylick, I. Waters,
T. Wells, S. Wilte, and P. Zuberbuhler.

Received for publication August 28, 2007, and accepted in revised
form January 2, 2008.

Address correspondence to: Julian Zielenksi, Program in Genetics
and Genome Biology, Hospital for Sick Children, 555 University A-
venue, Toronto, Ontario MSG 1X8, Canada.

Lilian Pereira’s present address is: Universidade Federal do Parana,
Curitiba, Brazil.

Lap-Chee Tsui’s present address is: University of Hong Kong, Pok-
falum, People’s Republic of China.