

A study of collectin genes in spontaneous preterm birth reveals an association with a common surfactant protein D gene polymorphism

Minna K. Karjalainen¹, Johanna M. Huusko¹, Anu Tuohimaa¹, Aino Luukkonen¹, Ritva Haataja¹ and Mikko Hallman¹

INTRODUCTION: Preterm birth is the major cause of mortality and morbidity in neonates. Intrauterine infection and/or inflammatory response are evident in 60–70% of spontaneous preterm births (SPTBs). Genetic factors significantly increase this risk. However, the genetic background associated with SPTB is poorly understood. Surfactant protein (SP) A, SP-D, and mannose-binding lectin (MBL) are structurally and functionally related collectins that bind pathogen-associated molecular patterns, and mostly suppress innate immune responses.

RESULTS: We detected an overrepresentation of the methionine allele of the *SFTPD* gene (encoding SP-D) Met31Thr polymorphism in preterm infants as compared to term infants. This association was highly significant in infants of families with recurrent SPTBs ($P = 0.001$, odds ratio = 1.65, 95% confidence interval = 1.22–2.22); however, there was no such association with *SFTPD* in the mothers of these infants. Polymorphism of the genes encoding SP-A and MBL did not influence the risk of SPTB.

DISCUSSION: Our results suggest that the fetal *SFTPD* Met31Thr polymorphism plays a significant role in genetic predisposition to SPTB. We propose that fetal immune responses influence sensitivity to preterm labor-inducing signals.

METHODS: Genes encoding SP-A, SP-D, and MBL were investigated as potential candidates for association with SPTB in a population of preterm singleton infants ($n = 406$) and their mothers ($n = 308$), and in mothers with term deliveries ($n = 201$) and their infants ($n = 201$), all originating from northern Finland.

Preterm birth (a birth that occurs before 37 completed weeks of gestation) is a serious global health problem; ~10% of all infants are born preterm (1). Prematurity is the main cause of mortality and morbidity in newborn infants. Preterm infants are at an increased risk of developing serious chronic diseases, including bronchopulmonary dysplasia, retinopathy of prematurity, cerebral palsy, and cognitive problems (2). More than 60% of all preterm births are due to spontaneous onset of labor. In multiple pregnancies, the risk of spontaneous preterm birth (SPTB) is several-fold higher. Bacterial vaginosis, ascending infections of the genital tract, and systemic infections, including asymptomatic infections of the urinary tract, serve as major risk

factors. In 25% of all preterm births, preterm premature rupture of the fetal membranes precedes the spontaneous onset of labor (3). Taken together, most risk factors for SPTB are related to intrauterine infection or inflammatory reaction.

There is evidence to suggest that genetics plays an important role in the risk of preterm birth. There is a higher incidence of preterm deliveries among mothers with a history of previous such deliveries (4); it is therefore likely that specific maternal and/or fetal genotypes are involved (5). Several genes have been considered as candidates for an association with SPTB (6), and the first genome-wide linkage analysis of SPTB was recently published (7). Given that infection and inflammation are involved in the initiation of SPTB, genes encoding proteins involved in the regulation of inflammatory mediators are plausible candidate genes. Indeed, polymorphism in several genes associated with inflammation pathways, including the cytokine genes encoding tumor necrosis factor- α and interleukin-4, have been reported to associate with preterm birth in mothers with preterm labor and in fetuses born preterm (6). Despite the reported associations, the genetic background of this multifactorial phenotype is still poorly understood. To date, genome-wide association analyses have not been reported.

Collectins are a family of C-type lectins with collagenous domains. Members of this family, including surfactant protein (SP) A, SP-D, and mannose-binding lectin (MBL), participate in several immunoregulatory functions and bind the carbohydrate domains of invading microorganisms (8). SP-A and -D are produced mainly in the type II lung alveolar cells but are also expressed in other cell types. These collectins are present in several tissues and compartments, including the airways, female genital tract, amniotic fluid, extraembryonic tissues, and blood (9). They have partially overlapping roles in host defense, particularly in the lung. Both SP-A and SP-D participate in the killing and clearance of pathogens and as opsonins; they also modulate inflammatory responses by interacting directly with pattern recognition receptors (10). In addition, SP-A and SP-D have complex roles in pulmonary surfactant metabolism and homeostasis (11). MBL is produced mainly in the liver and secreted into the blood. It functions by opsonization and activation of the lectin

¹Department of Pediatrics, Institute of Clinical Medicine, University of Oulu, Oulu, Finland. Correspondence: Minna K. Karjalainen (minna.k.karjalainen@oulu.fi)

The last two authors contributed equally to this work.

Received 25 April 2011; accepted 12 July 2011. doi:10.1038/pr.2011.2

pathway of the complement system, thereby leading to lysis of MBL-bound pathogens (12).

We propose that polymorphisms in collectin genes have a role in susceptibility to SPTB. The aim of this study was to assess the genes encoding SP-A (*SFTPA1* and *SFTPA2*), SP-D (*SFTPD*), and MBL (*MBL2*) for their possible role in SPTB in a genetically homogeneous northern Finnish population. Both fetal and maternal polymorphisms were considered. We also hypothesized that families with recurrent SPTBs may have a stronger genetic predisposition to this outcome as compared to families in which SPTBs were only sporadic. We found that, a common polymorphism of the gene encoding SP-D was associated with SPTB in preterm infants of families with recurrent SPTBs. Interestingly, this variation, Met31Thr, has previously been shown to influence the concentration, oligomerization, and binding properties of SP-D (13,14,15).

RESULTS

SP-A and SP-D Gene Polymorphisms

The frequency distribution of the *SFTPD* amino acid 31 allele was significantly different in SPTB infants of families with recurrent SPTBs as compared to controls; the frequency of the Met allele was 0.717 in cases and 0.605 in controls (nominal $P = 0.001$, permutation-corrected $P = 0.008$) (Table 1). The association of Met31Thr with SPTB in these infants was further analyzed using logistic regression analysis under three alternative models (codominant, dominant, and recessive) with sex as a covariate (Table 2). Logistic regression revealed that this polymorphism significantly influenced the

SPTB phenotype under the codominant model (Met/Met vs. Met/Thr vs. Thr/Thr, $P = 0.002$) and the recessive model (Met/Met vs. Met/Thr and Thr/Thr, $P = 0.0005$, odds ratio for Met/Met 2.06). No significant differences were observed among the mothers as regards the *SFTPD* Met31Thr allele frequencies (Table 1). In families with recurrent SPTB, the allele distribution in term siblings ($n = 176$) was comparable (0.707 and 0.293 for the Met and Thr alleles, respectively) to that in preterm siblings.

The frequency distributions of the alleles of *SFTPA1* Val50Leu and *SFTPA2* Gln223Lys were similar in SPTB infants or mothers (Table 1), as well as in term siblings of SPTB infants (data not shown), compared to control groups; this continued to hold true even after the SPTB cases were subdivided according to the number of SPTB cases in the family. The frequency distributions of the SP polymorphisms did not differ significantly between SPTB infants with gestational ages of <32 weeks ($n = 139$) and those of 32–36 weeks ($n = 263$) (minor allele frequencies of 0.381 and 0.409 for *SFTPA1* Val50Leu, 0.194 and 0.182 for *SFTPA2* Gln223Lys, and 0.371 and 0.317 for *SFTPD* Met31Thr, respectively). Additional stratification by gender, maternal smoking, and presence/absence of preterm premature rupture of the fetal membranes did not reveal significant differences for the SP polymorphism frequencies across the subgroups (data not shown). The SP gene polymorphisms did not deviate from Hardy–Weinberg equilibrium. The *SFTPA1* and *SFTPA2* SNPs displayed moderate linkage disequilibrium (LD) ($D' = 0.8$, $r^2 = 0.24$; Figure 1) and were therefore also analyzed as haplotypes. However, this analysis did not reveal any significant differences

Table 1. *SFTPA1*, *SFTPA2*, and *SFTPD* allele frequencies in the case–control study population of spontaneous preterm birth

Gene and polymorphism	Allele and corresponding aa	All SPTB cases	Recurrent SPTB	Sporadic SPTB	Controls
<i>Infants</i>		$n = 406$	$n = 189$	$n = 217$	$n = 201$
<i>SFTPA1</i> , rs1136450	G (Val50)	0.601	0.597	0.605	0.571
	C (Leu50)	0.399	0.403	0.395	0.429
<i>SFTPA2</i> , rs1965708	C (Gln223)	0.814	0.826	0.804	0.787
	A (Lys223)	0.186	0.174	0.196	0.213
<i>SFTPD</i> , rs721917	T (Met31)	0.664^a	0.717^a	0.618	0.605^a
	C (Thr31)	0.336	0.283	0.382	0.395
<i>Mothers</i>		$n = 308$	$n = 94$	$n = 214$	$n = 201$
<i>SFTPA1</i> , rs1136450	G (Val50)	0.582	0.635	0.559	0.568
	C (Leu50)	0.418	0.365	0.441	0.432
<i>SFTPA2</i> , rs1965708	C (Gln223)	0.825	0.845	0.816	0.801
	A (Lys223)	0.175	0.155	0.184	0.199
<i>SFTPD</i> , rs721917	T (Met31)	0.675	0.681	0.673	0.614
	C (Thr31)	0.325	0.319	0.327	0.386

SPTB cases, infants born preterm (gestational age <37 weeks) or mothers who had preterm deliveries; recurrent SPTB, infants or mothers from families with multiple infants born preterm; sporadic SPTB, infants or mothers from families with a single preterm infant; controls, term infants (gestational age >37 weeks) or mothers with term deliveries. Significantly different allele frequency differences ($P < 0.05$) between cases and controls are shown by bold; all other frequency differences were NS ($P > 0.05$).

aa, amino acid; CI, confidence interval; NS, nonsignificant; OR, odds ratio; SPTB, spontaneous preterm birth; *SFTPA1*, *SFTPA2*, and *SFTPD*, surfactant protein A1, A2, and D genes.

^aThe *SFTPD* Met31 allele was overrepresented in SPTB infants compared to controls ($P = 0.04$, OR = 1.29, 95% interval, CI = 1.01–1.66, permutation-corrected P NS). This association was restricted to SPTB infants with recurrent SPTB in the family ($P = 0.001$, OR = 1.65, 95% CI = 1.22–2.22, permutation-corrected $P = 0.008$). If only a single preterm infant per family was allowed in the analysis, frequency of the Met allele was 0.715 in infants ($n = 93$) from families with recurrent SPTB (compared to controls: $P = 0.009$, OR = 1.64, 95% CI = 1.12–2.38, permutation-corrected $P = 0.048$).

Table 2. Logistic regression analysis of *SFTPD* Met31Thr genotypes in preterm infants (gestational age <37 weeks) from families with recurrent spontaneous preterm births (SPTBs) and term control infants (gestational age >37 weeks) with sex as a covariate

Genotype	Frequency (n)		Logistic regression analysis		
	SPTB infants	Control infants	Model	P	OR (95% CI)
Met/Met	0.534 (101)	0.357 (71)	Codominant ^a	0.002	2.16 (1.13–4.16) ^b
Met/Thr	0.365 (69)	0.497 (99)	Dominant	0.183	1.52 (0.82–2.82) ^c
Thr/Thr	0.101 (19)	0.146 (29)	Recessive	0.0005	2.06 (1.37–3.10) ^d

CI, confidence interval; OR, odds ratio.

^aMet/Met vs. Met/Thr vs. Thr/Thr. ^bOR for Met/Met compared to Thr/Thr; if only a single preterm infant per family is analyzed, OR = 2.25, 95% CI = 0.98–5.17, *P* = 0.02. ^cOR for Met/Met and Met/Thr compared to Thr/Thr; if only a single preterm infant per family is analyzed, OR = 1.60, 95% CI = 0.73–3.35, *P* = 0.25. ^dOR for Met/Met compared to Thr/Thr and Met/Thr; if only a single preterm infant per family is analyzed, OR = 2.02, 95% CI = 1.23–3.34, *P* = 0.006

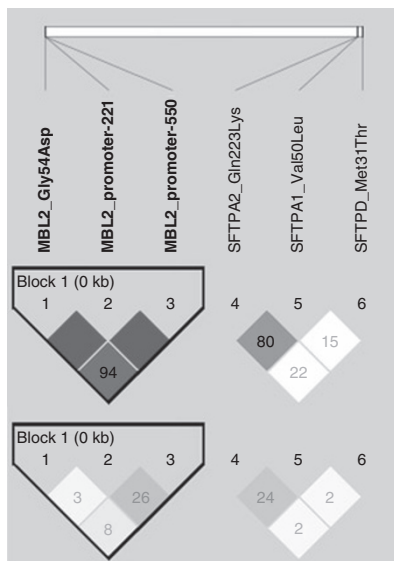


Figure 1. Linkage disequilibrium (LD) plots of the *SFTPA1*, *SFTPA2*, *SFTPD*, and *MBL2* polymorphisms for the control population consisting of term infants (gestational age >37 weeks) and their mothers. The names and relative positions of the polymorphisms are shown at the top. Pairwise D' (upper panel) and r^2 (lower panel) values are shown in the squares. Where no number appears, LD is complete ($D' = 1$). Darker colors indicate higher LD.

between SPTB infants or mothers, and controls (data not shown). Although the *SFTPD* locus lies in close physical proximity to the *SFTPA1* and *SFTPA2* loci on chromosome 10, the *SFTPD* Met31Thr polymorphism was not in LD with the two *SP-A* gene polymorphisms (Figure 1). Therefore, we did not analyze intergenic haplotypes of these three genes.

MBL2 Polymorphisms and Haplotypes

As expected, the *MBL2* promoter polymorphisms –550 (H/L) and –221 (X/Y), and exon 1 Gly54Asp (A/B) displayed high LD (Figure 1). Altogether, eight 3-SNP *MBL2* haplotypes were predicted. Of these, four (HYA, LXA, LYA, and LYB) were common (frequencies >0.05), whereas the other four (LXB, HYB, HXA, and HXB) were rare (frequencies <0.01). No significant differences were detected between the SPTB and control groups with respect to *MBL2* allele or haplotype frequencies (Table 3). Allele frequencies of term siblings of SPTB infants were comparable to those of their preterm siblings (data not shown).

Allele frequency distributions did not differ significantly between SPTB infants with gestational ages of <32 weeks ($n = 131$) and those of 32–36 weeks ($n = 261$) for any of the *MBL2* polymorphisms (minor allele frequencies of 0.357 and 0.420 for promoter –550; 0.229 and 0.252 for promoter –221; and 0.115 and 0.132 for Gly54Asp, respectively) or haplotypes (data not shown). When the study population was stratified by gender, maternal smoking, or presence/absence of preterm premature rupture of the fetal membranes, there were no significant differences in allele or haplotype frequencies among the subgroups (data not shown). The *MBL2* polymorphisms that were analyzed did not deviate from Hardy–Weinberg equilibrium.

MBL2 haplogenotypes are known to be associated with differences in serum MBL concentrations (16,17). Therefore, the haplogenotypes were ranked into classes according to their MBL scores (1–10), reflecting their varying potential to induce high-serum MBL concentrations as described previously (17), and MBL scores were calculated for each of the study groups. The frequencies of the high (MBL scores 6–10), intermediate (MBL scores 3–5), and low (MBL scores 1–2) MBL-conferring genotype groups were 70.8, 21.4, and 7.8%, respectively, in SPTB infants; 75.3, 19.7, and 5.1%, respectively, in control infants; and 77.8, 16.8, and 5.4%, respectively, in term siblings of SPTB infants. In mothers with SPTB and control mothers, the frequencies were 71.8 and 76% for high, 20.1 and 18.5% for intermediate, and 8.1 and 5.5% for low MBL-conferring genotype groups, respectively. No significant differences were detected in haplogenotype frequencies (data not shown) or MBL scores (Figure 2) between SPTB infants or mothers and controls; the median MBL score was 7 in all of the groups. However, the median MBL score was higher (median of 8) in term siblings of SPTB infants ($P = 0.03$ in comparison with term infants in the control group and $P = 0.005$ in comparison with SPTB infants) (Figure 2). When the SPTB infants or their mothers were divided into groups according to the nature of SPTBs in the family (recurrent or sporadic), or according to gestational age, no significant differences were observed across the subgroups (data not shown).

DISCUSSION

In SPTBs, the prominent feature is a mostly silent infection or inflammation involving the urogenital tract and the fetoplacental

Table 3. MBL2 allele and haplotype frequencies in the case–control study population of spontaneous preterm birth

			All SPTB cases	Recurrent SPTB	Sporadic SPTB	Controls	
Infants	Polymorphism	Allele	<i>n</i> = 395	<i>n</i> = 188	<i>n</i> = 207	<i>n</i> = 201	
		Promoter -550					
		rs11003125	C (L)	0.598	0.568	0.626	0.563
			G (H)	0.402	0.432	0.374	0.437
		Promoter -221					
		rs7096206	G (Y)	0.765	0.783	0.749	0.751
			C (X)	0.235	0.217	0.251	0.249
		aa 54					
		rs1800450	G, Gly (A)	0.875	0.888	0.862	0.897
			A, Asp (B)	0.125	0.112	0.138	0.103
		Haplotype					
			HYA	0.397	0.435	0.365	0.435
		LXA	0.225	0.209	0.241	0.246	
		LYA	0.247	0.244	0.249	0.216	
		LYB	0.119	0.106	0.129	0.099	
Mothers	Polymorphism	Allele	<i>n</i> = 300	<i>n</i> = 90	<i>n</i> = 210	<i>n</i> = 201	
		Promoter -550					
		rs11003125	C (L)	0.553	0.524	0.564	0.555
			G (H)	0.447	0.476	0.436	0.445
		Promoter -221					
		rs7096206	G (Y)	0.780	0.810	0.767	0.744
			C (X)	0.220	0.190	0.233	0.256
		aa 54					
		rs1800450	G, Gly (A)	0.868	0.872	0.867	0.898
			A, Asp (B)	0.132	0.128	0.133	0.102
		Haplotype					
			HYA	0.433	0.475	0.417	0.439
		LXA	0.207	0.186	0.218	0.254	
		LYA	0.216	0.205	0.220	0.203	
		LYB	0.127	0.127	0.123	0.099	

aa, amino acid; MBL2, mannose-binding lectin gene; SPTB, spontaneous preterm birth.

SPTB cases, infants born preterm (gestational age <37 weeks) or mothers who had preterm deliveries; recurrent SPTB, infants or mothers from families with multiple infants born preterm; sporadic SPTB, infants or mothers from families with a single preterm infant; controls, term infants (gestational age >37 weeks) or mothers with term deliveries. Allele and haplotype frequency differences between cases and controls were nonsignificant ($P > 0.05$).

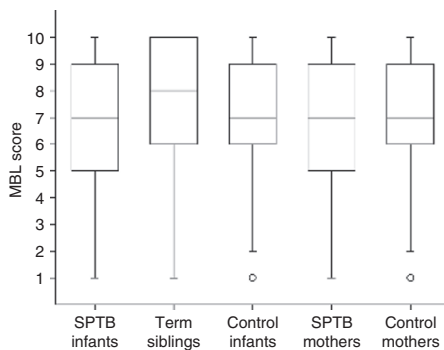


Figure 2. Mannose-binding lectin (MBL) genotype classes (MBL scores) in the case–control study population of spontaneous preterm birth (SPTB) and term siblings of preterm-born infants. Medians and ranges along with the 25 and 75% points are shown. Outliers are indicated by dots. Mean MBL scores were 6.81 (SPTB infants), 7.43 (term siblings), 6.98 (control infants), 6.85 (SPTB mothers), and 6.96 (control mothers). No differences in MBL scores were observed between SPTB infants or mothers and controls. The term siblings of SPTB infants displayed higher MBL scores than either the controls ($P = 0.03$) or SPTB infants ($P = 0.005$).

compartments (3). Although acquired factors play an important role, it is evident that susceptibility to SPTB is often an inherited feature, conferred by genetic factors of in the mother and the fetus (6). Given the important role of intrauterine infection and inflammation, we chose to explore variations of the

genes encoding SP-A, SP-D, and MBL, which have important roles in innate immunity. We found an association between a nonsynonymous variation at codon 31 of the *SFTPD* gene and the risk of SPTB when the outcome phenotype was being born preterm (affected fetus phenotype) (Table 1). To our knowledge, this is the first study to report an association of this gene with SPTB.

According to our results, the *SFTPD* Met31Thr polymorphism can be considered a potential factor to explain genetic susceptibility to SPTB. We observed a significant difference in the allele frequency of *SFTPD* methionine 31 ($P = 0.001$) between term infants of families with exclusively term deliveries and preterm infants of families with multiple preterm deliveries (Table 1). This association was further verified by logistic regression analysis, which suggested that the Met31Thr genotypes influence the SPTB phenotype under the codominant ($P = 0.002$) and recessive models ($P = 0.0005$; Table 2). Because the allele distribution of Met31Thr was similar in SPTB infants and their term siblings, this polymorphism may have a permissive role, possibly by influencing resistance to preterm labor–promoting signals. With the phenotype of giving birth preterm (affected mother phenotype), no significant associations were observed (Table 1). In an epidemiological study of large families with predisposition to preterm birth, the genome of the mother was shown to influence the duration of gestation of her preterm infant, whereas the

fetal genome influenced the premature birth *per se* but had no detectable influence on the degree of prematurity (18). This is consistent with the proposed permissive role for the fetal *SFTPD* Met31Thr polymorphism.

The biological significance of the associated *SFTPD* polymorphism has been studied previously. This polymorphism influences serum SP-D concentrations and oligomerization of the SP-D protein (13–15). The significance of the Met31Thr polymorphism is probably due to its influence on the binding of pathogen-associated molecular patterns. Indeed, highly multimerized and lower-molecular-weight forms of SP-D have been shown to exhibit differential binding to microbial ligands, with the former preferentially binding mannan and lipoteichoic acid and the latter, lipopolysaccharide and peptidoglycan (13,19). In the Finnish population, term-born young children carrying the *SFTPD* codon 31 methionine allele were shown to be predisposed to severe respiratory syncytial virus infection (20). In other studies, the alternative allele, threonine, was overrepresented in patients with tuberculosis in a Mexican population (21) and in those with allergic rhinitis in a Chinese population (22). This further suggests that genetic susceptibility depends on the specific ligand that causes the inflammatory challenge.

The potential role of *SFTPD* polymorphism in the process of preterm labor remains to be investigated. Although the main production site for SP-D is the lung, it is present in other tissues including the female reproductive tract, placenta, and fetal membranes (9). The concentration of this protein increases in the amniotic fluid as gestation proceeds (23). SP-D is currently thought to play a dual role in inflammation, with either proinflammatory or anti-inflammatory effects defined by the binding orientation of the protein (24), possibly depending on the degree of SP-D multimerization. Our study finding of an association between an *SFTPD* polymorphism and SPTB may be related to this dual function, which may be partly defined by the alternative variants of the Met31Thr polymorphism.

In this study, the genes encoding SP-A and MBL were not found to be associated with SPTB. According to the findings of an experimental study, SP-A in the amniotic fluid potentially regulates the duration of pregnancy, as implicated by the fact that its concentration levels increase strikingly as gestation proceeds toward term labor; moreover, a high concentration of SP-A was shown to induce proinflammatory cytokine expression in macrophages *in vitro* (25). However, in our study, the polymorphisms of *SFTPA1* and *SFTPA2* did not show any association with SPTB. In the case of the *MBL2* gene, our study finding of an absence of association was unexpected, because the association of distinct *MBL2* polymorphisms and haplotypes with preterm birth has been previously demonstrated (26–29). However, the earlier studies were performed in different populations and there are inconsistencies in phenotype definitions of preterm birth. In our study, the median MBL score, which indicates the potential to induce high serum MBL concentrations, tended to be higher in the term siblings (median of 8) of the SPTB infants as compared to all other

groups (median of 7) (Figure 2). This may actually mean that these individuals were “protected” from SPTB because of the presence of the higher serum MBL-conferring genotypes as compared to their preterm-born siblings. The roles of MBL and the pulmonary collectins in the etiology of SPTB require further investigation.

The strengths of the study include the use of a carefully selected, relatively homogeneous population (30), consideration of pregnancy history, and analysis of both maternal and fetal alleles. To avoid type I errors, we used a limited number of polymorphisms and applied corrections for multiple testing for the allele and haplotype frequency comparisons. However, this study had some limitations. According to the dbSNP database, 31, 30, 35, and 32 common SNPs exist within *SFTPA1*, *SFTPA2*, *SFTPD*, and *MBL2*, respectively, in Caucasians. Because our sample sizes did not provide adequate power to detect associations with a large number of polymorphisms, we decided to analyze only the SNPs that were most likely to have biological significance rather than aim at capturing complete variation of these genes. As a consequence, the analyzed polymorphisms captured only a portion of the variation within these genes, and our results do not exclude the possibility that other polymorphisms within these genes could affect the SPTB phenotype. For the *SFTPD* polymorphism, the observed association may be attributable to LD with other potential functional variations within the gene or with other genes in close proximity. Two other common nonsynonymous variations located on the collagen domain–encoding region are known to exist within the *SFTPD* gene. Three of the four common *MBL2* SNPs that correlate with MBL concentrations (16) were studied, with a careful analysis of haplogenotypes and associated MBL scores. However, P/Q variation, which is considered to be one of the potentially significant variations, was not analyzed. Therefore, we were not able to discriminate between the common *MBL2* LYPA and LYQA haplotypes.

In summary, a specific *SFTPD* polymorphism that may increase the concentration and aggregate size of SP-D showed association with SPTB when the phenotype was defined as being spontaneously born preterm. SP-D and other collectins may prove to be important proteins involved in activation of the labor process when the fetoplacental compartment is challenged with infection.

METHODS

Study Population and DNA Sample Preparation

The case study population consisted of mothers and their preterm infants (gestational age <37 weeks); all the women had delivered after spontaneous onset of labor. The participants were selected retrospectively from the 1973–2003 birth diaries of Oulu University Hospital, and prospectively during 2003–2005. The exclusion criteria were factors that are known to contribute to the risk of preterm birth (multiple gestation, preeclampsia, polyhydramnios, septic infection, and fetus with congenital disease), and chronic disease or alcohol/narcotics abuse in the mother. The control population consisted exclusively of mothers who had experienced at least three term births (gestational age >37 weeks; $n = 201$) and their singleton infants ($n = 201$), sampled prospectively from Oulu University Hospital between 2004 and 2007.

All of the families that were studied originated from northern Finland, where the population is known to exhibit genetic homogeneity (30). The two outcome phenotypes studied were (i) being delivered spontaneously preterm (affected fetus/infant phenotype), and (ii) delivering the infant spontaneously preterm (affected mother phenotype). Written informed consent was obtained for use of the samples, and the study was approved by the ethics committee of Oulu University Hospital.

The clinical characteristics of the case study population are presented in **Table 4**. The population comprised two groups: (i) mothers who had given birth preterm (gestational age <37 weeks) at least twice ($n = 94$), and their preterm infants ($n = 189$) ("recurrent SPTB"), and (ii) mothers who had given birth preterm once ($n = 214$), and their preterm infants ($n = 217$) ("sporadic SPTB"). The mean numbers of children in the families with recurrent and sporadic SPTBs were 4.7 and 2.9, respectively, consistent with a high fertility rate in this region. In 35% of all cases, preterm premature rupture of the fetal membranes preceded the SPTB (**Table 4**). We also analyzed data from term-born siblings ($n = 176$) of preterm infants belonging to families with recurrent SPTBs.

Genomic DNA was extracted from whole-blood specimens ($n = 706$) using the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN) and UltraClean DNA Blood Isolation Kit (MO BIO Laboratories, Carlsbad, CA). Chelex 100 (Bio-Rad, Hercules, CA) was used to extract genomic DNA from buccal cells ($n = 586$). Buccal samples were whole-genome amplified using an Illustra GenomePhi V2 DNA Amplification Kit (GE Healthcare, Buckinghamshire, UK). Ethidium bromide-stained agarose gels and UV absorbance measurements were used to control for the quality and quantity of the whole-genome amplified samples.

Table 4. Characteristics of the study population involved in spontaneous preterm birth

SPTB infants	
Total, n	406
Gestational age, weeks	32.2 ± 2.97 (22.7–36.1) ^a
Gestational age <32 weeks/≥32 weeks	140/263 ^b
Birth weight, g	1,938 ± 639 (538–3,375) ^a
PPROM ^c , yes/no	143/263
Male/female	232/174
Age of the mother at time of birth, years	29.1 ± 5.99 (16–46) ^a
Maternal smoking during pregnancy, yes/no/unknown	42/205/159
SPTB mothers	
Total n	308
Single spontaneous preterm delivery	214
Two spontaneous preterm deliveries	71
3–6 spontaneous preterm deliveries	23
Smoking during pregnancy, yes/no/unknown	37/189/82

PPROM, preterm premature rupture of fetal membranes; SPTB, spontaneous preterm birth.

^aMean ± s.d. (range). ^bPrecise gestational age of three preterm infants was unknown.

^cPregnancies with PPROM defined as preterm births with leakage of amniotic fluid as the presenting symptom before the onset of contractions.

SNP Selection and Genotyping

For each of the three SP genes, one nonsynonymous single-nucleotide polymorphism (SNP) was analyzed: *SFTPA1* SNP Val50Leu, *SFTPA2* SNP Gln223Lys, and *SFTPD* SNP Met31Thr, with accession numbers rs1136450, rs1965708, and rs721917, respectively, in the dbSNP database (<http://ncbi.nih.gov/SNP>). The *SFTPA1* and *SFTPA2* SNPs were selected because of their known utility as tagging SNPs, and also because they are located at the functional domains of the SP-A1 and SP-A2 polypeptides (collagenous and carbohydrate recognition regions, respectively) (31,32). The *SFTPD* Met31Thr was selected for study because this variation has previously been shown to affect SP-D oligomerization and concentration in the serum (13–15). Of the six *MBL2* SNPs known to correlate with serum MBL concentrations (16), three were included in this study: the promoter variations –550 (H/L) and –221 (X/Y) and the exon 1 variation Gly54Asp (A/B), with accession numbers rs11003125, rs7096206, and rs1800450, respectively, in the dbSNP database. Given their low minor allele frequency (<0.05) in Caucasians, the *MBL2* exon 1 variations Cys52Arg (A/D, rs5030737) and Gly57Glu (A/C, rs1800451) were not included in the analysis. The P/Q variation (rs7095891) at position +4 in the 5' untranslated region was excluded from analysis because of deviation from Hardy–Weinberg equilibrium.

The *SFTPA1*, *SFTPA2*, and *SFTPD* SNPs were genotyped using PCR-restriction fragment-length polymorphism analysis. Two rounds of PCR were first performed to amplify the region surrounding each SNP. One of the primers in the reamplification had either a one-nucleotide or a two-nucleotide mismatch, resulting in the creation of an arbitrary recognition site for the restriction endonucleases *DdeI*, *TaqI*, and *FspI* for the *SFTPA1*, *SFTPA2*, and *SFTPD* SNPs, respectively, for the purpose of genotype detection. Restriction endonucleases were purchased from New England Biolabs (Beverly, MA). The *MBL2* SNPs were genotyped by template-directed dye-terminator incorporation with fluorescence polarization detection (33) using AcycloPrime II SNP Detection Kits (Perkin Elmer Life Sciences, Boston, MA). Details of the PCR and SNP primers used are available on request.

Statistical Analysis, Haplotype Prediction, and Study Power

Haploview, v. 4.2 (34) was used for Hardy–Weinberg equilibrium testing and to obtain pairwise LD (D' and r^2) values, as well as for estimation of haplotype frequencies and comparisons of allele and haplotype frequencies. Haploview creates haplotype frequency estimates based on an accelerated expectation maximization algorithm, and uses the χ^2 test for case–control analysis of alleles and haplotypes. The P values obtained from allele and haplotype analyses were corrected for multiple testing using 10,000 permutations; a corrected P value <0.05 was considered significant. Genotype frequency comparisons (logistic regression analysis with fetal sex as a covariate under codominant, dominant, and recessive models) and calculations of common odds ratios were performed using Predictive Analytics SoftWare statistics, v. 17.0.3 (IBM SPSS). *MBL2* phased haplotypes were constructed using SNP-HAP, v. 1.3.1. (<http://www-gene.cimr.cam.ac.uk/clayton/software>). *MBL2* haplotypes were classified according to their potential to induce high serum MBL concentrations into classes ranging from 1 (lowest potential) to 10 (highest potential), as described previously (17), and subsequently analyzed by means of the nonparametric Mann–Whitney U -test using Predictive Analytics SoftWare statistics. The genetic power of the study was estimated using Genetic Power Calculator (<http://pngu.mgh.harvard.edu/~purcell/gpc>) (35). Using the allelic 1 degree of freedom test and an additive risk model, and assuming a causal SNP with a minor allele frequency ranging from 0.1 to 0.4 and a prevalence of 0.05 for SPTB, our sample size was estimated to provide 80% power ($\alpha = 0.008$, taking into account multiple testing of six SNPs) to detect genotypic relative risks of 1.6–2.0 for heterozygotes and 2.3–2.9 for homozygotes in the infants. Using the same parameters, the sample size in the mothers was estimated to provide 80% power to detect relative risks of 1.7–2.0 for the heterozygous and 2.4–3.0 for the homozygous risk genotypes.

ACKNOWLEDGMENTS

The authors thank Maarit Haarala, Outi Kajula, Mirrka Ovaska, and Riitta Vikeväinen for sample collection and technical assistance.

STATEMENT OF FINANCIAL SUPPORT

This work was supported by grants from the Finnish Academy and the Sigrid Jusélius Foundation (M.H.), and the Foundation of Pediatric Research in Finland (R.H.).

REFERENCES

1. Beck S, Wojdyla D, Say L, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. *Bull World Health Organ* 2010;88:31–8.
2. Ward RM, Beachy JC. Neonatal complications following preterm birth. *BJOG* 2003;110:Suppl 20:8–16.
3. Goldenberg RL, Culhane JE, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet* 2008;371:75–84.
4. Porter TF, Fraser AM, Hunter CY, Ward RH, Varner MW. The risk of preterm birth across generations. *Obstet Gynecol* 1997;90:63–7.
5. Wilcox AJ, Skjaerven R, Lie RT. Familial patterns of preterm delivery: maternal and fetal contributions. *Am J Epidemiol* 2008;167:474–9.
6. Plunkett J, Muglia LJ. Genetic contributions to preterm birth: implications from epidemiological and genetic association studies. *Ann Med* 2008;40:167–95.
7. Haataja R, Karjalainen MK, Luukkonen A, et al. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. *PLoS Genet* 2011;7:e1001293.
8. Wright JR. Immunoregulatory functions of surfactant proteins. *Nat Rev Immunol* 2005;5:58–68.
9. Kishore U, Greenhough TJ, Waters P, et al. Surfactant proteins SP-A and SP-D: structure, function and receptors. *Mol Immunol* 2006;43:1293–315.
10. Sano H, Kuroki Y. The lung collectins, SP-A and SP-D, modulate pulmonary innate immunity. *Mol Immunol* 2005;42:279–87.
11. Haagsman HP, Hogenkamp A, van Eijk M, Veldhuizen EJ. Surfactant collectins and innate immunity. *Neonatology* 2008;93:288–94.
12. Ip WK, Takahashi K, Ezekowitz RA, Stuart LM. Mannose-binding lectin and innate immunity. *Immunol Rev* 2009;230:9–21.
13. Leth-Larsen R, Garred P, Jensenius H, et al. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005;174:1532–8.
14. Heidinger K, König IR, Bohnert A, et al. Polymorphisms in the human surfactant protein-D (SFTPD) gene: strong evidence that serum levels of surfactant protein-D (SP-D) are genetically influenced. *Immunogenetics* 2005;57:1–7.
15. Sørensen GL, Hjelmberg JB, Kyvik KO, et al. Genetic and environmental influences of surfactant protein D serum levels. *Am J Physiol Lung Cell Mol Physiol* 2006;290:L1010–7.
16. Frakking FN, Brouwer N, Zweers D, et al. High prevalence of mannose-binding lectin (MBL) deficiency in premature neonates. *Clin Exp Immunol* 2006;145:5–12.
17. Pesonen E, Hallman M, Sarna S, et al. Mannose-binding lectin as a risk factor for acute coronary syndromes. *Ann Med* 2009;41:591–8.
18. Plunkett J, Feitosa MF, Trusgnich M, et al. Mother's genome or maternally-inherited genes acting in the fetus influence gestational age in familial preterm birth. *Hum Hered* 2009;68:209–19.
19. Sorensen GL, Hoegh SV, Leth-Larsen R, et al. Multimeric and trimeric subunit SP-D are interconvertible structures with distinct ligand interaction. *Mol Immunol* 2009;46:3060–9.
20. Lahti M, Lofgren J, Marttila R, et al. Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 2002;51:696–9.
21. Floros J, Lin HM, García A, et al. Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis* 2000;182:1473–8.
22. Deng YQ, Tao ZZ, Kong YG, et al. Association between single nucleotide polymorphisms of surfactant protein D and allergic rhinitis in Chinese patients. *Tissue Antigens* 2009;73:546–52.
23. Miyamura K, Malhotra R, Hoppe HJ, et al. Surfactant proteins A (SP-A) and D (SP-D): levels in human amniotic fluid and localization in the fetal membranes. *Biochim Biophys Acta* 1994;1210:303–7.
24. Gardai SJ, Xiao YQ, Dickinson M, et al. By binding SIRPalpha or calreticulin/CD91, lung collectins act as dual function surveillance molecules to suppress or enhance inflammation. *Cell* 2003;115:13–23.
25. Condon JC, Jeyasuria P, Faust JM, Mendelson CR. Surfactant protein secreted by the maturing mouse fetal lung acts as a hormone that signals the initiation of parturition. *Proc Natl Acad Sci USA* 2004;101:4978–83.
26. Annells MF, Hart PH, Mullighan CG, et al. Interleukins-1, -4, -6, -10, tumor necrosis factor, transforming growth factor-beta, FAS, and mannose-binding protein C gene polymorphisms in Australian women: Risk of preterm birth. *Am J Obstet Gynecol* 2004;191:2056–67.
27. Bodamer OA, Mitterer G, Maurer W, Pollak A, Mueller MW, Schmidt WM. Evidence for an association between mannose-binding lectin 2 (MBL2) gene polymorphisms and pre-term birth. *Genet Med* 2006;8:518–24.
28. van de Geijn FE, Dolhain RJ, van Rijs W, Willemsen SP, Hazes JM, de Groot CJ. Mannose-binding lectin genotypes are associated with shorter gestational age. An evolutionary advantage of low MBL production genotypes? *Mol Immunol* 2008;45:1514–8.
29. Swierczko AS, Atkinson AP, Cedzynski M, et al. Two factors of the lectin pathway of complement, I-ficolin and mannan-binding lectin, and their associations with prematurity, low birthweight and infections in a large cohort of Polish neonates. *Mol Immunol* 2009;46:551–8.
30. Jakkula E, Rehnström K, Varilo T, et al. The genome-wide patterns of variation expose significant substructure in a founder population. *Am J Hum Genet* 2008;83:787–94.
31. Rämetsä M, Haataja R, Marttila R, Floros J, Hallman M. Association between the surfactant protein A (SP-A) gene locus and respiratory-distress syndrome in the Finnish population. *Am J Hum Genet* 2000;66:1569–79.
32. Löfgren J, Rämetsä M, Renko M, Marttila R, Hallman M. Association between surfactant protein A gene locus and severe respiratory syncytial virus infection in infants. *J Infect Dis* 2002;185:283–9.
33. Chen X, Kwok PY. Homogeneous genotyping assays for single nucleotide polymorphisms with fluorescence resonance energy transfer detection. *Genet Anal* 1999;14:157–63.
34. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5.
35. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19:149–50.