

# Family-Based Association Tests Suggest Linkage Between Surfactant Protein B (SP-B) (and Flanking Region) and Respiratory Distress Syndrome (RDS): SP-B Haplotypes and Alleles From SP-B–Linked Loci Are Risk Factors for RDS

JOANNA FLOROS, NEAL J. THOMAS, WENLEI LIU, COSTAS PAPAGAROUFALIS, MARIETTA XANTHOU, SUNITA PEREIRA, RUZONG FAN, XIAOXUAN GUO, SUSAN DIANGELO, AND JELENA PAVLOVIC

*Departments of Cellular and Molecular Physiology [J.F., X.G., S.D., J.P.], Pediatrics [J.F., N.J.T.], Obstetrics and Gynecology [J.F.], and Health Evaluation Sciences [N.J.T., W.L.], The Pennsylvania State University College of Medicine, Hershey, PA 17033; Aghia Sophia Children's Hospital [C.P., M.X.], Athens 11527, Greece; Division of Newborn Medicine [S.P.], Tufts New England Medical Center, Boston, MA 02111; Department of Statistics [R.F.], TX A&M University, College Station, TX 77843*

**ABSTRACT:** Genetic variants of surfactant protein B (SP-B) have been associated with respiratory distress syndrome (RDS) in the prematurely born infant. We wished to determine linkage between RDS and SP-B single nucleotide polymorphisms (SNPs) [–18 (A/C), 1013 (A/C), 1580 (C/T), and 9306 (A/G)] or SP-B-linked microsatellite [(D2S388, D2S2232, (AAGG)<sub>n</sub>, and GATA41E01 (or D2S1331)] loci and identify susceptibility or protective alleles and haplotypes. We genotyped 132 families consisting of one or two parents and at least one child affected with RDS and performed biallelic and multiallelic family-based association test (FBAT) analysis, and extended transmission disequilibrium test (ETDT). ETDT analysis identified the microsatellite SP-B–linked loci (except D2S2232) to be linked to RDS. One allele from each of these three marker loci contributes to the risk of RDS. Multiallelic FBAT analysis detected a signal of linkage for the region of the four SNP loci. Three haplotypes within this region contribute to RDS risk. Although no other region showed significant linkage as judged by multiallelic FBAT, biallelic FBAT analysis revealed three potential susceptibility haplotypes formed by two to four loci within the SP-B and SP-B–linked microsatellite region. Each haplotype included GATA41E01, which was identified by ETDT analysis to be linked to RDS. We conclude that SP-B or SP-B–linked loci are linked to RDS and certain alleles or haplotypes are susceptibility or protective factors for the development of RDS in infants born prematurely.

(*Pediatr Res* 59: 616–621, 2006)

The cause of RDS of the prematurely born infant is likely a combination of multifactorial environmental insults and multigenic inherited predispositions (1,2). Deficiency of the lipoprotein complex called pulmonary surfactant can lead to RDS in the prematurely born infant. Surfactant protein deficiency (3–5) as well as genetic variants of the surfactant

protein A (SP-A) and SP-B have been associated with RDS (6–12).

SP-B is essential for normal lung function as assessed by animal (13) and human (14,15) studies. Lower than normal amounts of SP-B protein in SP-B heterozygous (+/–) mice, under normal conditions, have been associated with physiologic lung abnormalities such as decreased compliance and air trapping (16). The human SP-B gene has been localized on the short arm of chromosome 2 (17). It consists of 11 exons, and the precursor molecule of 42 kD (18) undergoes 5' and 3' posttranslational cleavages to give rise to the mature SP-B of 6 kD (19).

A number of SP-B SNPs (15) and four SP-B–linked microsatellite markers (20) have been characterized. Three of the microsatellite markers [D2S388, D2S2232, (AAGG)<sub>n</sub>] are located at the 5' end of the SP-B gene at approximately 130, 64, and 27 kb, respectively, according to the high-resolution TNG3 radiation hybrid panel analysis (20). However, according to the current version of the human genome map of chromosome 2 ([http://www.ensembl.org/Homo\\_sapiens](http://www.ensembl.org/Homo_sapiens)), D2S388 and D2S2232 markers are at approximately 191 and 105 kb from the centromeric end of SP-B, respectively. The microsatellite marker GATA41E01 (or D2S1331) is located at the 3' end of the SP-B at approximately 1074 kb by the G3 medium-resolution radiation hybrid panel (20) and at the 5' end of SP-B at approximately 697 kb by the current human genome map. Microsatellite marker loci are highly polymorphic as assessed by their high values of polymorphism information content (PIC) or heterozygosity index. For example, many polymorphic dinucleotide repeats have PIC values >0.70 (21), with an average heterozygosity index of about

Received May 13, 2005; accepted December 2, 2005.

Correspondence: Joanna Floros, Ph.D., Department of Cellular and Molecular Physiology, The Pennsylvania State University College of Medicine, 500 University Drive, P.O. Box 850, Hershey, PA 17033; e-mail: jfloros@psu.edu

This work was supported by NIH R37HL34788 and by the General Clinical Research Center at The Pennsylvania State University College of Medicine (M01 RR010732 and C06 RR016499).

DOI: 10.1203/01.pdr.0000203145.48585.2c

**Abbreviations:** ETDT, extended transmission disequilibrium test; FBAT, family-based association test; PIC, polymorphism information content; RDS, respiratory distress syndrome; SNP, single nucleotide polymorphism; SP-A, surfactant protein A; SP-B, surfactant protein B

0.75 (22). As a result, microsatellite markers are considered powerful informative markers in genetic studies of disease etiology. The higher level of resolution afforded by these multiallelic microsatellite markers can lead to a better detection of disease subgroups (21,22). In contrast, the average PIC value of a SNP is about 0.3 (22). In general, SNPs are not as informative as microsatellite markers in linkage studies. However, haplotype analysis of multiple linked SNPs can be more informative.

In this report, using biallelic and multiallelic FBAT (23–25) and ETDT/transmission disequilibrium test (TDT) analyses (for single marker analysis FBAT and TDT are essentially the same test), we aimed to a) determine whether SP-B or SP-B-linked loci are linked to the RDS locus and b) identify susceptibility and/or protective alleles and haplotypes for RDS.

## METHODS

Genomic DNA was extracted from blood samples, buccal swabs, or discarded tissues from subjects with and without RDS and their parents according to institutional guidelines. The protocol for the use of human samples in this study was approved by The Human Subjects Protection Office of The Pennsylvania State University College of Medicine. The diagnosis of RDS was made by clinical criteria (grunting, flaring, retraction, need for oxygen and continuous positive airway pressure or ventilatory support) and/or verified by the reticulogranular pattern on x-ray. The genomic DNA served as template for PCR in the genotype analysis.

The study group consisted of 132 families. These families consisted of one ( $n = 43$ ) or two ( $n = 89$ ) parents and had at least one affected child. For these families, genotype data of the loci under study (see below) for the parents and for at least one affected offspring were used for analyses in the present study. Table 1 shows the characteristics of the study group. These include the

**Table 1.** Characteristics of population used in the study ( $N = 399$ )

No. of families	132
No. of families*	99 (1), 24 (2), 6 (3), 2 (4), 1 (6)
No. of families†	113 (1), 13 (2), 4 (3), 2 (4)
Zygote	
Monozygotic twins	8
Monozygotic triplets	1
Dizygotic twins	43
Dizygotic triplets	18
Dizygotic quadruple	4
Race	
Black, no Hispanic	27
Hispanic	15
White, no Hispanic	343
Other or mixed parents	11
Unknown	3
Sex	
Female	204
Male	195
Ethnicity	
American	102
Greek	108
Gypsy	2
The Netherlands	53
Unknown	134
Age category	
Adult	224
Child	10
Newborn	165

\* Number of children in parentheses.

† Number of affected children in parentheses.

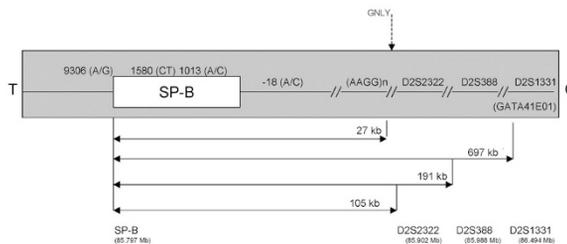
number of affected and the total number of children per family, zygosity, race, sex, ethnicity, and age.

For each microsatellite or SP-B SNP marker loci, the number of alleles or haplotypes transmitted indicates the frequency that the particular allele or haplotype was transmitted to the affected offspring by the heterozygous parent(s). Therefore, in the study, only parents heterozygous for the particular marker alleles or haplotypes were used. For each allele at loci with multiple alleles (or haplotypes), a new biallelic system was created by collapsing the alleles other than the one studied into a new allele category X. For example, to study the transmission of allele 2 of marker D2S388, the other alleles (1, 3, 4, 5, etc.) of marker D2S388 were collapsed into a new allele category X.

**Genotype analysis.** Genomic DNA was used as template for genotype analysis of four SNP loci, (-18(A/C), 1013 (A/C), 1580 (C/T), and 9306 (A/G) in human SP-B (15,26) and of four SP-B-linked microsatellite marker loci (20). The location of the microsatellite markers in relation to centromere (C) and telomere (T) as well as the SP-B gene SNPs are shown in Figure 1. Originally, the location of the microsatellite markers was determined by radiation hybrid panels (20). The G3 medium resolution hybrid panel positioned the D2S2232 more distally from the SP-B locus than the D2S388, whereas the TNG3 high-resolution hybrid panel positioned the D2S388 as the most distal marker (compared with D2S2232) from the SP-B locus (20). The GATA41E01 or D2S1331 microsatellite marker was placed at 3' end of SP-B by the G3 medium resolution radiation hybrid panel analysis (20) but at the 5' end of SP-B by the human genome map. The location of (AAGG) $_n$  microsatellite markers is as described by Kala *et al.* (20). In Figure 1, we show the relative location of the four microsatellite markers according to the current map of the human chromosome 2 ([http://www.ensembl.org/Homo\\_sapiens](http://www.ensembl.org/Homo_sapiens)). The genotyping and scoring of alleles for the microsatellites was done as described previously (15,20). A fraction of the SNPs were genotyped using pyrosequencing (Pavlovic *et al.*, in preparation). The PCR-based RFLP genotype method (15) provided the basis for the pyrosequencing protocol, which is a primer-based DNA sequencing method. Briefly, single-stranded templates containing the SNP under study are produced by polymerase chain reaction (PCR). Then a sequencing primer hybridizes to the single-stranded template and initiates nucleotide incorporation by DNA polymerase. Following incorporation of each nucleotide, a pyrophosphate group is released, which is converted to adenosine triphosphate (ATP). The ATP in turn drives the conversion of luciferin to oxyluciferin, producing a light that is detected by a charge-coupled device camera. Each light signal generates a pyrogram and is proportional to the number of nucleotides incorporated onto the DNA strand. Pyrograms were scored by a pattern-recognition software that compare the predicted SNP pattern (histogram) to the observed pattern (pyrogram) (Pyrosequencing AB, Uppsala, Sweden).

**Statistical analysis.** The PedCheck program was used to determine the compatibility of genotypes at each marker locus within families, before analysis (27). Marker loci with incompatible parental and offspring genotypes were treated as missing data in the families of question. Association between alleles of surfactant protein genes and RDS was tested using the FBAT (23,25,28). ETDT analysis was performed to assess linkage of a multiallelic locus to the disease locus (29,30).

FBAT analysis was performed using the online program <http://www.biostat.harvard.edu/fbat/fbat.htm>. Both multiallelic FBAT analysis for marker loci with multiple alleles or haplotypes (23) and biallelic FBAT analysis for SP-B SNP markers and selected microsatellites (23,25) were



**Figure 1.** Schematic presentation of SP-B SNPs and SP-B centromeric flanking microsatellite marker loci. The relative location of the marker loci used in the present analysis is shown from the telomere (T) to the centromere (C) according to the current map of human chromosome 2 (see “Methods”). The distance of the flanking microsatellite marker loci is noted in kilobase from the 3' end of the SP-B gene (or in megabase from the telomere). SP-B has been located from 5' to 3' within the 85.80697–85.79705 region from the telomere. The relative location of the granulysin precursor (GNLY) is noted by an arrow.

performed. Haplotypes formed by all contiguous marker loci (two to eight loci) were tested. All FBAT analyses assumed an additive model, with a minimum size (minisize) set to four, indicating that the test statistic was not computed when the number of informative families available was fewer than four. A significant  $p$  value and a positive  $Z$  statistic were indicative of a susceptibility marker allele for disease, whereas a significant  $p$  value and a negative  $Z$  statistic were indicative of a protective marker allele for disease. Results were not corrected for multiple comparisons because the eight marker loci under study may not be independent given the short distances of their location on chromosome 2 (15,20).

Missing parental haplotypes were reconstructed by the FBAT program based on observed parental and offspring genotypes assuming no recombination among marker loci. For individuals with multiple possible haplotypes, all the possible haplotypes were incorporated into the test. Different haplotype pairs were given different weights according to the conditional probability of observing the haplotype pairs given that they were compatible with the observed unphased genotype. FBAT evaluates the distribution of test statistics using the conditional offspring genotype distribution under the null hypothesis (24).

## RESULTS

**ETDT and TDT analyses.** ETDT analysis identified three microsatellite marker loci [D2S388, (AAGG)<sub>n</sub>, GATA41E01] to be linked to RDS. As indicated by the partial output for these marker loci (Table 2), the allele-wise TDT model is more powerful for marker loci D2S388 ( $p = 0.04$ ) and (AAGG)<sub>n</sub> ( $p = 0.045$ ) because the  $\chi^2$  for goodness of fit of the allele-wise model is not significant (Table 3). For marker locus GATA41E01 (or D2S1331), the genotype-wise TDT model is more powerful ( $p = 0.048$ ; goodness of fit:  $p = 0.03$ ). Table 2 depicts log likelihoods under the null and the parsimonious (allele-wise) hypotheses and for the saturated (genotype-wise) model. L0 of ETDT output represents the log base e likelihood of observing the transmission pattern, assuming there is no linkage and no linkage disequilibrium (LD). L1 of ETDT output represents the log-likelihood of the observed data maximized with respect to all  $m-1$  parameters;  $m$  is the number of alleles at the locus. Each one of them has a transmission probability that can be estimated by maximizing the log-likelihood. L2 of ETDT output represents the log-likelihood of the observed data maximized with respect to the  $m(m-1)/2$  parameters. Each heterozygote parent has a transmission probability for transmitting a certain allele that can be estimated by maximizing the log-likelihood. Table 3 depicts the  $\chi^2$  and the likelihood ratio for allele-wise and genotype-wise TDT and for the goodness of fit of the allele-wise model.

The likelihood ratio depicts the likelihood of the observed data under the alternative hypothesis relative to the likelihood of the observed data under the null hypothesis. For example, the first row for each marker in Table 3 represents the likeli-

hood of the observed data under the allele-wise model, assuming there is linkage and/or LD relative to the likelihood of the observed data, assuming there is no linkage and no LD. The likelihood ratio was computed using the unlogged likelihood under the alternative hypothesis divided by the unlogged likelihood under the null hypothesis. For example, the first row for each marker in Table 3 was computed by unlogged L1/unlogged L0, where L1 represented the maximized log base e likelihood of the observed data under the allele-wise model, and L0 represented the log base e likelihood of the observed data, assuming there is no linkage and no LD. One allele at each of these marker loci contributes to RDS risk (Fig. 2). These comprise two susceptibility alleles, D2S388\_4 ( $p = 0.04$ ) and GATA41E01\_1 ( $p = 0.01$ ) and one protective allele (AAGG)<sub>n\_8</sub> ( $p = 0.03$ ).

Based on the number of families listed in Table 1, we performed power calculation using the computer program TDT power calculator, version 1.2.1 (31). Assuming the disease allele has a frequency of 0.15 (with penetrance of 0.2, 0.1, and 0.05 for DD, Dd, and dd genotypes, respectively), the  $D'$  (a measure of LD between the marker and disease allele) is 0.9, and the recombination fraction (between the marker and disease locus) is 0, we have 72.7% power to detect a significant linkage and/or association at 0.05 significance level if the marker has a minor allele frequency (MAF) of 0.1. Keeping all parameters unchanged except for the marker MAF, we have, respectively, 82.5%, 64.1%, 49.2%, and 35.4% power to detect a significant linkage and/or association at 0.05 significance level if the markers have MAF of 0.2, 0.3, 0.4, and 0.5.

**FBAT analysis.** Multiallelic FBAT analysis of the four SP-B SNP regions (-18(A/C)/1013(A/C)/1580(C/T)/9306(A/G) indicated that the SP-B locus is linked to RDS (the overall  $p$  value for this region is 0.007). One susceptibility and two protective haplotypes of this region appear to contribute to this linkage (Fig. 2). The protective haplotypes are -18\_A/1013\_C/1580\_T/9306\_G ( $p = 0.04$ ) and -18\_A/1013\_A/1580\_T/9306\_A ( $p = 0.02$ ), and the susceptibility haplotype is -18\_A/1013\_C/1580\_T/9306\_A ( $p = 0.04$ ). Although no other region showed significant linkage as judged by multiallelic FBAT, biallelic FBAT analysis identified three susceptibility haplotypes with significant  $p$  values ( $p < 0.01$ ) formed by 2-4 SNP and/or microsatellite marker loci that can potentially contribute to RDS risk (Fig. 2). These are 9306\_A/GATA4E01\_1 ( $p = 0.003$ ), 9306\_A/1580\_C/GATA4E01\_1 ( $p = 0.004$ ), and 9306\_A/1580\_C/1013\_A/GATA41E01\_1 ( $p = 0.005$ ). Interestingly, the GATA41E01 (or D2S1331) locus shown by ETDT to be linked to RDS is present in all three susceptibility haplotypes identified by the biallelic FBAT.

## DISCUSSION

The cause of RDS of the prematurely born infant is likely a combination of multifactorial environmental insults and multigenic inherited predispositions (1,2,32). SP-B is essential for normal lung function (13-15), and SP-B polymorphisms from case-control studies of unrelated individuals

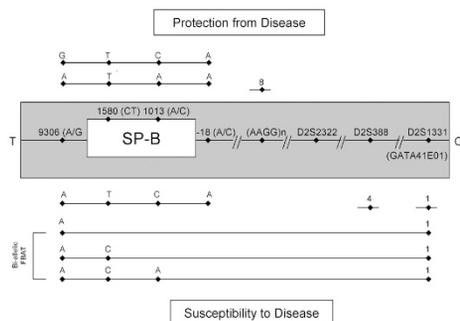
**Table 2.** Partial output of the ETDT analysis for the microsatellite markers: log likelihood under various hypotheses or \*model

Marker	Null	*Parsimonious (allele-wise)	*Saturated (genotype-wise)
D2S388	L0 = -105.358360	L1 = -98.109764	L2 = -92.822243
(AAGG) <sub>n</sub>	L0 = -85.950241	L1 = -75.932137	L2 = -73.817444
GATA41E01	L0 = -115.755577	L1 = -112.329918	L2 = -103.165428

\* Parsimonious or saturated model.

**Table 3.** Partial output of the ETDT analysis for the microsatellite markers:  $\chi^2$  for various TDTs or model

Marker	$\chi^2$	df	Likelihood ratio	p value
D25388	14.497192 for allele-wise TDT or 2* (L1-L0)	7	1406.1292562	0.043221
	25.072235 for genotype-wise TDT or 2* (L2-L0)	18	278205.9651	0.123557
	10.575042 for goodness of fit of allele-wise model or 2* (L2-L1)	11	197.85234101	0.479664
(AAGG)n	20.036209 for allele-wise TDT or 2* (L1-L0)	11	22428.864454	0.045147
	24.265594 for genotype-wise TDT or 2* (L2-L0)	17	185868.925	0.0112977
	4.229385 for goodness of fit of allele-wise model or 2* (L2-L1)	6	8.2870412539	0.645679
GATA41E01	6.851318 for allele-wise TDT or 2* (L1-L0)	6	30.742897747	0.335003
	25.180298 for genotype-wise TDT or 2* (L2-L0)	15	293651.50909	0.047968
	18.328979 for goodness of fit of allele-wise model or 2* (L2-L1)	9	9551.8487392	0.031761



**Figure 2.** Protective and susceptibility alleles and haplotypes. In the center (shaded area), the relative position of the SP-B SNP marker loci and the microsatellite markers present in the SP-B centromeric flanking region are shown (not drawn to scale). The protective and susceptibility haplotypes and alleles determined by multi-FBAT (haplotypes) and ETDT/TDT (alleles), are shown, respectively, above and below the SP-B diagram of the marker loci (shaded). For example, two protective haplotypes have been identified by FBAT ( $-18\_A/1013\_C/1580\_T/9306\_G$  and  $-18\_A/1013\_A/1580\_T/9306\_A$ ), and by ETDT/TDT analysis, one microsatellite allele [(AAGG) $n_8$ ]. FBAT and ETDT/TDT also identified a susceptibility haplotype and two microsatellite alleles. Three susceptibility haplotypes identified by biallelic FBAT analysis are also shown.

have been shown to be associated with RDS (7–9,11,12). In the present study, we used FBAT and ETDT analyses to avoid problems of false associations that can occur with case-control studies and to determine linkage and association of SP-B SNP or SP-B-linked microsatellite marker loci with RDS. The findings indicated that the SP-B four-SNP region (multiallele FBAT) and three microsatellite marker loci (ETDT) are linked to RDS. Two protective haplotypes and one susceptibility haplotype were identified for the SP-B four-SNP region, and of the significant microsatellite marker loci, two susceptibility alleles (D25388\_4, GATA41E01\_1) and one protective allele [(AAGG) $n_8$ ] were identified. These findings indicate that the SP-B genetic locus and the SP-B centromeric flanking region are linked to RDS and that certain haplotypes and alleles contribute to susceptibility or protection from RDS.

For single-marker analysis, FBAT and TDT are essentially the same. However, for haplotype analysis, FBAT may be more appropriate for nearby loci that are likely to be in LD. The haplotype FBAT test uses a weighted conditional approach and assumes no recombination (24), whereas the GeneHunter haplotype TDT uses the maximum likelihood approach based on an earlier test, the Lander-Green hidden Markov model, and assumes no LD (33). FBAT and TDT

measure the transmission frequency of alleles from a heterozygous parent to the affected offspring, whereas the non-transmitted allele serves as control. In this way, problems of false associations that may arise in case-control association studies due to unrecognized bias of stratification of the study group, admixture, or heterogeneity (34) may be minimized. Multiallelic FBAT and ETDT analyses can be used for a composite test analysis of all the alleles of a multiallelic locus, such as the microsatellite marker loci (29) or for the haplotypes of several SNP loci, and can determine linkage. Together, the data from the FBAT and ETDT analyses can provide evidence of linkage and association of the loci under study with the disease locus in question.

It is currently unknown whether the SP-B-linked microsatellite marker alleles, shown in the present study to be associated with increased or decreased risk of RDS, identify SP-B alleles or alleles of other (yet unknown) genes associated with RDS. A search of the human genome map specifically in the region between the microsatellite markers determined in the present study to be associated with and linked to RDS identified several known and unknown genes. One of the known genes is the granulysin gene, a T-cell product, which is located on the centromeric side of (AAGG) $n$  at a distance of about 36 kb from SP-B or at 85.833 Mb from the telomere. Granulysin is expressed in human cytolytic T lymphocytes and in natural killer cells and contributes to host defense *via* its antimicrobial properties (35–37). Interestingly, granulysin shares sequence similarity with a family of proteins known as saposins, of which SP-B is a family member. Granulysin has been shown to exhibit lytic activity against a wide range of microbes, inhibit growth of both Gram-positive and Gram-negative bacteria (38) and induce apoptosis in Jurkat cells (39). Recently, SP-B has been shown, in addition to its surfactant-related activity, to play a role in host defense (40–42). Therefore, the sequence similarity between SP-B and granulysin and their role in host defense raise the possibility that these molecules have similar or overlapping host defense activities. However, if and how granulysin contributes to the development of RDS remain to be determined. Several other genes were identified that were located between D2S388 and D2S1331 (DNA directed polymerase 1, mitochondrial inner membrane protein, ribosomal protein L35, and receptor expression enhancing protein 1) as well as between D2S388 and D2S2322 (sialyl transferase 9, atonal homologue 8). None of these genes appear to be expressed in the lung, and thus their role (if any) in the pathogenesis of RDS is unknown.

However, the significant microsatellite alleles identified in the present study may play a role in the regulation of RDS-associated alleles. It has been shown that microsatellites allow formation of Z-DNA (left-handed) conformation, which in turn may promote interaction with *trans*-acting factors. Such an interaction may be influenced by the length of the microsatellites (43–45). A role of microsatellites in the regulation of the human fetal globin gene has been suggested in a preliminary study (46).

Previous case-control studies have associated SP-B marker alleles and RDS. The frequency of an SP-B intron 4 size variant has been shown to be higher in RDS (7,9), and the combined frequency of the SP-B intron 4 variant and an SP-A variant was significantly higher in RDS than either marker allele alone, indicating a synergistic effect (9,11). Moreover, another case-control study identified SP-B as an interactive genetic determinant with SP-A for RDS (8). Although the intron 4 variant was not investigated in the present study, association of SP-B variants with RDS is consistent with the finding that the SP-B genetic locus is linked to RDS.

In summary, the present findings indicate linkage and association of the SP-B gene as well as of SP-B-linked loci upstream (on the centromeric side) of the SP-B locus to the RDS locus. Several alleles and/or haplotypes are identified to be associated with increased or decreased risk of RDS. The SP-B gene has been shown to be essential for life, and genetic variants may modulate risk of disease under certain environmental insults or stressors. The significant SP-B-linked marker loci may identify loci of unknown genes associated with RDS or SP-B variants linked to the microsatellite markers. Moreover, the marker loci studied here and shown to be linked to RDS may be useful in identifying RDS subgroups for which different therapies may be considered, and their underlying mechanisms may be investigated.

**Acknowledgments.** The authors thank Dr. Junming Luo for his help with genotyping and Caitlin Custer for typing this manuscript.

## REFERENCES

1. Floros J, Karinch AM 1995 Genetics of neonatal disease as they relate to the surfactant protein genes. In: Robertson B, Tausch HW (eds) *Surfactant Therapy for Lung Disease*. Marcel Dekker, New York, pp 95–106
2. Floros J, Kala P 1998 Surfactant proteins: molecular genetics of neonatal pulmonary diseases. *Annu Rev Physiol* 60:365–384
3. deMello DE, Phelps DS, Patel G, Floros J, Lagunoff D 1989 Expression of the 35kDa and low molecular weight surfactant-associated proteins in the lungs of infants dying with respiratory distress syndrome. *Am J Pathol* 134:1285–1293
4. Moya FR, Montes HF, Thomas VL, Mouzinho AM, Smith JF, Rosenfeld CR 1994 Surfactant protein A and saturated phosphatidylcholine in respiratory distress syndrome. *Am J Respir Crit Care Med* 150:1672–1677
5. Floros J, Fan R, Matthews A, DiAngelo S, Luo J, Nielsen H, Dunn M, Gewolb IH, Koppe J, van Sonderen L, Farri-Kostopoulos L, Tzaki M, Ramet M, Merrill J 2001 Family-based transmission disequilibrium test (TDT) and case-control association studies reveal surfactant protein A (SP-A) susceptibility alleles for respiratory distress syndrome (RDS) and possible race differences. *Clin Genet* 60:178–187
6. Floros J, Veletzka SV, Kotikalapudi P, Krizkova L, Karinch AM, Friedman C, Buchter S, Marks K 1995 Dinucleotide repeats in the human surfactant protein-B gene and respiratory-distress syndrome. *Biochem J* 305:583–590
7. Haataja R, Ramet M, Marttila R, Hallman M 2000 Surfactant proteins A and B as interactive genetic determinants of neonatal respiratory distress syndrome. *Hum Mol Genet* 9:2751–2760
8. Kala P, Ten Have T, Nielsen H, Dunn M, Floros J 1998 Association of pulmonary surfactant protein A (SP-A) gene and respiratory distress syndrome: interaction with SP-B. *Pediatr Res* 43:169–177

10. Rämetsä M, Haataja R, Marttila R, Floros J, Hallman M 2000 Association between the surfactant protein A (SP-A) gene locus and respiratory-distress syndrome in the Finnish population. *Am J Hum Genet* 66:1569–1579
11. Floros J, Fan R, DiAngelo S, Guo X, Wert J, Luo J 2001 Surfactant protein (SP) B associations and interactions with SP-A in white and black subjects with respiratory distress syndrome. *Pediatr Int* 43:567–576
12. Fan R, Floros J, Xiong M 2002 Transmission disequilibrium test of two unlinked disease loci: application to respiratory distress syndrome. *Adv Appl in Stat* 1:277–308
13. Clark JC, Wert SE, Bachurski CJ, Stahlman MT, Stripp BR, Weaver TE, Whitsett JA 1995 Targeted disruption of the surfactant protein B gene disrupts surfactant homeostasis, causing respiratory failure in newborn mice. *Proc Natl Acad Sci U S A* 92:7794–7798
14. Nogueira LM, Garnier G, Dietz HC, Singer L, Murphy AM, deMello DE, Colten HR 1994 A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. *J Clin Invest* 93:1860–1863
15. Lin Z, deMello DE, Wallot M, Floros J 1998 An SP-B gene mutation responsible for SP-B deficiency in fatal congenital alveolar proteinosis: evidence for a mutation hotspot in exon 4. *Mol Genet Metab* 64:25–35
16. Clark JC, Weaver TE, Iwamoto HS, Ikegami M, Jobe AH, Hull WM, Whitsett JA 1997 Decreased lung compliance and air trapping in heterozygous SP-B-deficient mice. *Am J Respir Cell Mol Biol* 16:46–52
17. Vamvakopoulos NC, Modi WS, Floros J 1995 Mapping the human pulmonary surfactant-associated protein B gene (SFTPB) to chromosome 2p12–>p11.2. *Cytogenet Cell Genet* 68:8–10
18. Jacobs KA, Phelps DS, Steinbrink R, Fisch J, Kriz R, Mitscock L, Dougherty JP, Tausch HW, Floros J 1987 Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa pulmonary surfactant-associated protein. *J Biol Chem* 262:9808–9811
19. Floros J, Phelps DS 1997 Pulmonary surfactant. In: Yaksh TL, Lynch C III, Zapol WM, Maze M, Bieback J, Saidman LJ (eds) *Anesthesia: Biologic Foundations*. Lippincott-Raven, New York, pp 1259–1279
20. Kala P, Koptides M, DiAngelo S, Hoover RR, Lin Z, Veletzka V, Kouretas D, Floros J 1997 Characterization of markers flanking the human SP-B locus. *Dis Markers* 13:153–167
21. Weber JL 1990 Human DNA polymorphisms based on length variations in simple-sequence tandem repeats. In: Davies E, Tilghman SM (eds) *Genetic and Physical Mapping*. Cold Spring Harbor Laboratory Press, New York, pp 159–181
22. Weber JL, May PE 1989 Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388–396
23. Horvath S, Xu X, Laird NM 2001 The family based association test method: strategies for studying general genotype-phenotype associations. *Eur J Hum Genet* 9:301–306
24. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM 2004 Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26:61–69
25. Laird NM, Horvath S, Xu X 2000 Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 19:S36–S42
26. Lin Z, Pearson C, Chinchilli V, Pietschmann SM, Luo J, Pison U, Floros J 2000 Polymorphisms of human SP-A, SP-B, and SP-D genes: association of SP-B Thr131Ile with ARDS. *Clin Genet* 58:181–191
27. O'Connell J, Weeks DE 1998 PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266
28. Rabinowitz D, Laird N 2000 A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered* 50:211–223
29. Sham PC, Curtis D 1995 An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet* 59:323–336
30. Curtis D 1997 Use of siblings as controls in case-control association studies. *Ann Hum Genet* 61:319–333
31. Chen WM, Deng HW 2001 A general and accurate approach for computing the statistical power of the transmission disequilibrium test for complex disease genes. *Genet Epidemiol* 21:53–67
32. Floros J, Hoover RR 1998 Genetics of the hydrophilic surfactant proteins A and D. *Biochim Biophys Acta* 1408:312–322
33. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES 1996 Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363
34. Spielman RS, Ewens WJ 1996 The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 59:983–989
35. Ernst WA, Thoma-Uszynski S, Teitelbaum R, Ko C, Hanson DA, Clayberger C, Krensky AM, Leippe M, Bloom BR, Ganz T, Modlin RL 2000 Granulysin, a T cell product, kills bacteria by altering membrane permeability. *J Immunol* 165:7102–7108
36. Ochoa MT, Stenger S, Sieling PA, Thoma-Uszynski S, Sabet S, Cho S, Krensky AM, Rollinghoff M, Nunes Sarno E, Burdick AE, Rea TH, Modlin RL 2001 T-cell release of granulysin contributes to host defense in leprosy. *Nat Med* 7:174–179
37. Kumar J, Okada S, Clayberger C, Krensky AM 2001 Granulysin: a novel antimicrobial. *Expert Opin Investig Drugs* 10:321–329
38. Krensky AM 2000 Granulysin: a novel antimicrobial peptide of cytolytic T lymphocytes and natural killer cells. *Biochem Pharmacol* 59:317–320
39. Gamen S, Hanson DA, Kaspar A, Naval J, Krensky AM, Anel A 1998 Granulysin-induced apoptosis. I. Involvement of at least two distinct pathways. *J Immunol* 161:1758–1764

40. Miles PR, Bowman L, Rao KM, Baatz JE, Huffman L 1999 Pulmonary surfactant inhibits LPS-induced nitric oxide production by alveolar macrophages. *Am J Physiol* 276:L186–L196
41. van Iwaarden JF, Claassen E, Jeurissen SH, Haagsman HP, Kraal G 2001 Alveolar macrophages, surfactant lipids, and surfactant protein B regulate the induction of immune responses via the airways. *Am J Respir Cell Mol Biol* 24:452–458
42. Epaud R, Ikegami M, Whitsett JA, Jobe AH, Weaver TE, Akinbi HT 2003 Surfactant protein B inhibits endotoxin-induced lung inflammation. *Am J Respir Cell Mol Biol* 28:373–378
43. Rich A, Nordheim A, Wang AH 1984 The chemistry and biology of left-handed Z-DNA. *Annu Rev Biochem* 53:791–846
44. McLean MJ, Wells RD 1988 The role of sequence in the stabilization of left-handed DNA helices in vitro and in vivo. *Biochim Biophys Acta* 950:243–254
45. Muller V, Takeya M, Brendel S, Wittig B, Rich A 1996 Z-DNA-forming sites within the human beta-globin gene cluster. *Proc Natl Acad Sci U S A* 93:780–784
46. Lapoumeroulie C, Castiglia L, Ruberto C, Fichera M, Amata S, Labie D, Ragusa A 1999 Genetic variations in human fetal globin gene microsatellites and their functional relevance. *Hum Genet* 104:307–314