

# Haplotypes of the surfactant protein genes A and D as susceptibility factors for the development of respiratory distress syndrome

Neal J. Thomas (nthomas@psu.edu)<sup>1,2</sup>, Ruzong Fan<sup>5</sup>, Susan DiAngelo<sup>3</sup>, Joseph C. Hess<sup>1</sup>, Joanna Floros<sup>1,3,4</sup>

Departments of 1.Pediatrics, 2.Health Evaluation Sciences, 3.Cellular and Molecular Physiology, 4.Obstetrics and Gynecology, Penn State Children's Hospital and The Pennsylvania State University College of Medicine, Hershey, Pennsylvania, USA

5.Department of Statistics, Texas A&M University, College Station TX

## Keywords

Genetics, Haplotype, Surfactant protein A, Surfactant protein D, Respiratory distress syndrome

## Correspondence

Neal J. Thomas, M.D., M.Sc, Pediatric Critical Care Medicine, Penn State Children's Hospital, 500 University Drive, MC, H085, Hershey, Pennsylvania 17033, USA.  
Tel: +(717)531-5337 | Fax: +(717)531-0809 | Email: nthomas@psu.edu

## Received

27 October 2006; revised 29 January 2007; accepted 15 March 2007.

DOI:10.1111/j.1651-2227.2007.00319.x

## Abstract

**Aims:** Polymorphisms of genes are transmitted together in haplotypes, which can be used in the study of the development of complex diseases such as respiratory distress syndrome (RDS). The surfactant proteins (SPs) play important roles in lung function, and genetic variants of these proteins have been linked with lung diseases, including RDS. To determine whether haplotypes of SP-A and SP-D are transmitted disproportionately from parents to offspring with RDS, we hypothesized that previously unstudied genetic haplotypes of these SP genes are associated with the development of RDS. **Methods:** DNA was collected from 132 families of neonates with RDS. Genotyping was performed, and haplotype transmission from parent to offspring was determined by transmission disequilibrium test.

**Results:** The two-marker SP-D/SP-A haplotype DA160\_A/SP-A2 1A<sup>1</sup> is protective against the development of RDS ( $p = 0.035$ ). Four three- and four-marker haplotypes containing one or both loci from the significant two-marker haplotype are also protective against the development of RDS.

**Conclusions:** These data identify protective haplotypes against RDS and support findings related to SP genetic differences in children who develop RDS. Study of haplotypes in complex diseases with both genetic and environmental risk factors may lead to better understanding of these types of diseases.

## INTRODUCTION

After many years of work by multiple investigative teams, it has been determined that the human genome contains 10 million polymorphisms, which are defined as allele differences that occur in at least 1% of the population. This immense number of potential 'targets' of candidate genes has made research aimed at determining genetic predisposition to multifactorial diseases such as neonatal respiratory distress syndrome (RDS) overwhelming. However, genetic researchers do have an advantage, in that many of these polymorphisms are often transmitted not independent from one another, and the presence of one gene variant can tag the presence of another polymorphism from the same chromosome. These groups of polymorphisms, or haplotypes, are transmitted together. Haplotype assessment can provide a higher level of specificity, sensitivity, and accuracy in 'true' associations with disease risk or severity. By focusing on haplotypes instead of SNPs, researchers are now able to more effectively, efficiently, and accurately study genetic predisposition to various diseases of interest. With the recent report of the International HapMap Consortium (1), and the identification and cataloging of haplotypes now available, the utility of this type of study is brought into focus as an important tool to guide genetic association studies on complex human diseases such as RDS.

RDS is a well-described and frequently studied disease. It is known that the environmental insult of premature birth predisposes newborns to develop this acute lung injury, and

that the host inflammatory response plays an important role in the development of RDS. However, there is wide variation in both the severity of RDS, the long-term sequelae that result from RDS, and also the development of this disease in the population of children who are at risk secondary to environmental insults. These wide variations of a single 'syndrome' likely are based in the genetic background of the individual at risk. It is likely that a combination of multifactorial environmental insults and multigenic-inherited predispositions interact and lead to the development of RDS (2,3). Genetic variation is obvious from the well-documented increased incidence of RDS in males (4,5) as well as the lower incidence of disease and milder disease severity in black populations when compared with that in Caucasians (4,6). While it is likely that multiple genes, and/or their interactions, predispose to the development of RDS, genetic association studies on this disease should focus on the genes that are involved in lung development and injury repair. Therefore, the SPs have become a target of study. Multiple genetic variants of the surfactant proteins A (SP-A) and B (SP-B) have been shown to be both 'at risk' for and protective against the development of RDS (7-12).

There are four SPs that comprise the protein portion of pulmonary surfactant and play important roles in lung function and host defense of the lung. These proteins collectively play important roles in surfactant function, such as surface tension regulation, as well as in the structure and metabolism of surfactant. Moreover, the SP-A and SP-D are particularly important in the innate immune response to pulmonary

insults and infections by regulation of the inflammatory processes of the lung (13,14). Not only are SP-A and SP-D coupled by function, but they also are linked by virtue of being mapped to a short region on chromosome 10 (15–17). A three-marker linkage disequilibrium (LD) analysis revealed significant LD in several ethnic groups between the SP-D and SP-A loci (18). Therefore, examination of the transmission of haplotypes of these two closely related proteins to the development of disease is of significant value.

Based on the above information, we undertook the present study to determine if haplotypes of SP-A and SP-D were transmitted disproportionately from parents to the offspring affected with RDS. We hypothesized that novel genetic haplotypes of these SP genes, in combination with the environmental insult of premature birth, are associated with the development of RDS in a heterogeneous population of infants.

## PATIENTS AND METHODS

Genomic DNA was extracted from umbilical cord blood samples or buccal swabs from subjects with RDS and one or both of 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, and informed consent was obtained from each participant or guardian. 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 number of affected and the total number of children per family, zygosity, race, sex, ethnicity and age.

### Genomic analysis

DNA was extracted from 200  $\mu$ L of blood or buccal swab using QIAamp DNA mini kit (Qiagen, Valencia, CA) according to manufacturer's instructions. PCR-cRFLP was performed as described for SP-A and SP-D markers (19) for samples collected prior to 2003. All other samples were genotyped for SP-A and SP-D using a pyrosequencing method that has been previously described in detail (20).

### Statistical analysis

The nuclear family-based transmission disequilibrium test (TDT) analysis was performed using GENEHUNTER (www.broad.mit.edu) (Whitehead Institute for Biomedical Research, MIT) to determine (a) transmission of individ-

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

No. of families	132	
No. of families (No. of children)	99 (1), 24 (2), 6 (3), 2 (4), 1 (6)	
No. of families (No. of affected children)	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

ual SP-A and SP-D markers from parents to affected offspring, and to test for (b) transmission of haplotypes of two-, three- and four-marker loci (21). Statistical significance was defined as  $p < 0.05$ .

## RESULTS AND DISCUSSION

The difficulties in finding causative genetic etiologies to complex disease processes such as RDS are likely based on these syndromes resulting from multifactorial environmental insults and multigenetic inherited predispositions. While many SNPs have been found to relate to complex diseases such as RDS, the search for a single-candidate gene to determine disease causality is likely to be extremely challenging. Therefore, the change in focus to groups of linked polymorphisms being transmitted from parents to the affected offspring together is emerging. In this study, we demonstrate, for the first time, haplotype transmission disequilibrium of the SP-A and SP-D loci on chromosome 10 to the development of RDS. Analyses were performed examining the transmission of two-, three-, and four-marker haplotypes. The TDT results (Table 2) demonstrated that one two-marker haplotype of the SP-D/SP-A2 locus, DA160\_A/SP-A2 1A<sup>1</sup>, is protective against the development of RDS ( $\chi^2 = 4.45$ ;  $p = 0.035$ ). Moreover, two three-marker haplotypes and two four-marker haplotypes were also found to be protective against the development of RDS (Table 2). These data further the findings of our group and others related to the SP genetic differences in children who develop RDS

**Table 2** Results of the haplotype analysis of the surfactant proteins on the development of respiratory distress syndrome

Haplotype	Protection/ risk	Transmitted	Untransmitted	p
Two marker analysis				
DA160_A/SP-A2 1A <sup>1</sup>	Protection	2	9	0.035
Three marker analysis				
DA11_T/ DA160_A/ SP-A2 1A <sup>1</sup>	Protection	1	9	0.011
DA160_A/SP-A2 1A <sup>2</sup> / SP-A1 6A <sup>4</sup>	Protection	0	4	0.046
Four marker analysis				
DA11_T/ DA160_A/SP-A2 1A <sup>1</sup> /SP-A1 6A <sup>3</sup>	Protection	2	9	0.035
DA11_T/ DA160_A/SP-A2 1A <sup>2</sup> /SP-A1 6A <sup>4</sup>	Protection	0	4	0.046

(7–12,22). Outside of the haplotype analyses presented, we found no significant differences in transmission of the SP-D markers [DA11(C/T) and DA160(A/G)] that were tested.

Genetic variants of SP-A have been associated in multiple studies to either an increased or decreased risk of the development of RDS in pre-term infants (7,10,23,24). However, this is likely only one piece of the puzzle in understanding the genetics of RDS. The SP-A and SP-D loci map on the long arm of chromosome 10 and have been identified with significant LD in several ethnic groups (18). Therefore, it is likely that certain haplotypes containing SNPs of both genes may have an impact on diseases that relate to pulmonary host defense functions. In fact, transmission of haplotypes from parents to affected children that contain SNPs of SP-A and SP-D identified a protective haplotype. The DA160\_A variant was found to be protective against RDS when haplotyped with SP-A2 variant 1A<sup>1</sup>. This was observed in a two- (p = 0.035), a three- (p = 0.011) and a four- (p = 0.035) marker analysis (Table 2). Moreover, the DA160\_A variant was found to be protective against RDS when haplotyped with a second SP-A2 variant (1A<sup>2</sup>) in three- (p = 0.046) and four- (p = 0.046) marker analyses. Interestingly, a second SP-D SNP (DA11\_T) was found in the protective DA160\_A-containing haplotypes of three- (p = 0.011) and four-marker analyses (p = 0.035 and p = 0.046, Table 2). These data indicate that both DA160\_A and DA11\_T, when present in SP-A-containing haplotypes, are likely to confer protection from RDS. In the DA160 (A/G) SNP, the nucleotide alteration to an A (160\_A) from a G (160\_G) leads to a change to threonine from alanine in amino acid 160, which is located in the carbohydrate recognition domain of SP-D. Whether this amino acid change affects properties of SP-D remains to be determined. Moreover, in the DA11 (C/T) SNP, the DA11\_T and DA11\_C encode methionine and threonine, respectively. The Thr/Thr<sup>11</sup> genotype has been shown to associate with lower SP-D levels when compared with the Met/Met<sup>11</sup> genotype (25). Since the two SP-D loci [DA11(C/T) and DA160(A/G)] are linked (18), it is un-

clear whether the protection is conferred by the changes at amino acid 11 or amino acid 160.

The significant protective haplotypes identified in the present study contain the SP-A2 variants 1A<sup>1</sup> and 1A<sup>2</sup>. SP-A2 has been shown to be more active than SP-A1 with regards to its ability to modulate proinflammatory cytokine production by a macrophage-like cell line (26–28) enhance phagocytosis of *Pseudomonas aeruginosa* (29), and inhibit surfactant secretion (30). Whether these SP-A2 variants provide protection against certain lung disease via enhanced activities in host defense or surfactant homeostasis remains to be determined. However, in other lung diseases, the 1A<sup>1</sup> variant and a 1A<sup>1</sup>-containing haplotype were shown to associate with worse pulmonary function in adults with cystic fibrosis (31), and the 1A<sup>1</sup>/1A<sup>1</sup> genotype was shown to associate with susceptibility to the development of respiratory syncytial virus (32), as was the DA11\_T variant discussed earlier (33).

Differences of SP-A variants or haplotypes inferring protection in RDS and risk in other lung diseases (or the converse) have been previously observed. For example, the 6A<sup>2</sup>/1A<sup>0</sup> haplotype, or the 6A<sup>2</sup> or 1A<sup>0</sup> variants, have been shown to associate with RDS risk (7,11), whereas, in recurrent acute otitis media, the 6A<sup>2</sup>/1A<sup>0</sup> haplotype was shown to associate with protection (34). Comparable observations were made for the SPA-1 locus variant present in the significant haplotypes shown in Table 2. One haplotype of the three-marker analysis and the two haplotypes of the four-marker analysis (Table 2) contained the SP-A1 6A<sup>4</sup> or 6A<sup>3</sup> variant marker locus. Interestingly, the 6A<sup>3</sup> variant, in the presence of the SP-B (1580 T/T) genotype, has been shown previously to associate with reduced risk for RDS in black subjects (12). Moreover, two of these three haplotypes contained, in addition of the 6A<sup>4</sup> locus, the 1A<sup>2</sup> locus. The protective effect of the two 6A<sup>4</sup>/1A<sup>2</sup>-containing haplotypes is consistent with our previous findings where the 6A<sup>4</sup>/1A<sup>2</sup> haplotype, in a pilot study, showed a trend as being a protective factor for RDS (7). However, the 6A<sup>4</sup> variant has been found to associate with risk in other diseases such as interstitial pulmonary fibrosis (35), tuberculosis (36) and cancer (37). Perhaps these types of differences in association studies where a given haplotype associates with risk or protection in neonates and the converse in children and/or adults point to differences in pathways involved in the pathogenesis of RDS and other lung diseases.

In summary, we found that a two-marker haplotype of the two SPs on chromosome 10, SP-A and SP-D, associates with protection from the development of RDS in prematurely born infants at risk for developing acute lung injury. We have also demonstrated that three- and four-marker haplotypes (each containing one or both markers found in the two-marker haplotype) associate with protection against the development of RDS. A limitation of the present study is the moderate sample size, which is likely to result in under-reporting of causative haplotypes, since our power to detect small differences is limited by our number of complete trios for each haplotype studied. The data presented in this

project should be considered 'pilot' in nature, and therefore these associations should be strengthened and validated by increased sample size, as well as by replication in other groups of premature children. Only by exploring and then validating, in alternative settings, the association of these haplotype with lung disease of prematurity, we will be able to understand the complex nature of the genetic impact on RDS.

#### ACKNOWLEDGEMENTS

This work was supported by NIH R37HL34788, by the General Clinical Research Center at The Pennsylvania State University College of Medicine (M01 RR010732 and C06 RR016499), and by Children's Miracle Network research grant from Penn State Children's Hospital. The authors wish to thank Yunhua Wang for her help with genotyping.

#### References

1. The International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; 437: 1299–320.
2. Floros J, Karinch AM. Genetics of neonatal disease as they related to the surfactant protein genes. In: Robertson B, Taesch HW, editors. *Surfactant therapy for lung diseases*. New York: Marcel Dekker, 1995; 95–106.
3. Floros J, Kala P. Surfactant proteins: molecular genetics of neonatal pulmonary diseases. *Annu Rev Physiol* 1998; 60: 365–84.
4. Farrell PM, Wood RE. Epidemiology of hyaline membrane disease in the United States: analysis of national mortality statistics. *Pediatrics* 1976; 58: 167–76.
5. Khoury MJ, Marks JS, McCarthy BJ, Zaro SM. Factors affecting the sex differential in neonatal mortality: the role of respiratory distress syndrome. *Am J Obstet Gynecol* 1985; 151: 777–82.
6. Hulsey TC, Alexander GR, Robillard PY, et al. Hyaline membrane disease: the role of ethnicity and maternal risk characteristics. *Am J Obstet Gynecol* 1993; 168: 572–6.
7. Floros J, Fan R, Matthews A, DiAngelo S, Luo J, Nielsen H, et al. 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* 2001; 60: 178–87.
8. Floros J, Veletzka SV, Kotikalapudi P, Krizkova L, Karinch AM, Friedman C, et al. Dinucleotide repeats in the human surfactant protein-B gene and respiratory-distress syndrome. *Biochem J* 1995; 305(Pt 2): 583–90.
9. Haataja R, Ramet M, Marttila R, Hallman M. Surfactant proteins A and B as interactive genetic determinants of neonatal respiratory distress syndrome. *Hum Mol Genet* 2000; 9: 2751–60.
10. Kala P, Ten HT, Nielsen H, Dunn M, Floros J. Association of pulmonary surfactant protein A (SP-A) gene and respiratory distress syndrome: interaction with SP-B. *Pediatr Res* 1998; 43: 169–77.
11. Ramet 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.
12. Floros J, Fan R, Diangelo S, Guo X, Wert J, Luo J. Surfactant protein (SP) B associations and interactions with SP-A in white and black subjects with respiratory distress syndrome. *Pediatr Int* 2001; 43: 567–76.
13. Floros J, Phelps DS. Pulmonary surfactant. In: Yaksch TL, Lynch C, Maze M, Biebuyck JF, Saidman LJ, editors. *Anesthesia: biological foundations*. New York: Lippincott-Raven, 1997; 1257–79.
14. Phelps DS. Pulmonary surfactant modulation of host-defense function. *Appl Cardiopulm Pathophysiol* 1995; 5: 221–9.
15. Hoover RR, Floros J. Organization of the human SP-A and SP-D loci at 10q22-q23. Physical and radiation hybrid mapping reveal gene order and orientation. *Am J Respir Cell Mol Biol* 1998; 18: 353–62.
16. Floros J, Hoover RR. Genetics of the hydrophilic surfactant proteins A and D. *Biochim Biophys Acta* 1998; 1408: 312–22.
17. Kolble K, Lu J, Mole SE, Kaluz S, Reid KB. Assignment of the human pulmonary surfactant protein D gene (SFTPD) to 10q22-q23 close to the surfactant protein A gene cluster. *Genomics* 1993; 17: 294–8.
18. Liu W, Bentley CM, Floros J. Study of human SP-A, SP-B and SP-D loci: allele frequencies, linkage disequilibrium and heterozygosity in different races and ethnic groups. *BMC Genet* 2003; 4: 13.
19. Diangelo S, Lin Z, Wang G, Phillips S, Ramet M, Luo J, et al. Novel, non-radioactive, simple and multiplex PCR-cRFLP methods for genotyping human SP-A and SP-D marker alleles. *Dis Markers* 1999; 15: 269–81.
20. Pavlovic J, Papagaroufalis C, Xanthou M, Liu W, Fan R, Thomas NJ, et al. Genetic variants of surfactant proteins A, B, and D in bronchopulmonary dysplasia. *Dis Markers* 2006; 22: 227–232.
21. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; 58: 1347–65.
22. Fan R, Floros J, Xiong M. Transmission disequilibrium test of two unlinked disease loci; Application to respiratory distress syndrome. *Adv Appl Stat* 2001; 1: 277–308.
23. Marttila R, Haataja R, Ramet M, Pokela ML, Tammela O, Hallman M. Surfactant protein A gene locus and respiratory distress syndrome in Finnish premature twin pairs. *Ann Med* 2003; 35: 344–52.
24. Floros J, Fan R. Surfactant protein A and B genetic variants and respiratory distress syndrome: allele interactions. *Biol Neonate* 2001; 80(Suppl 1): 22–5.
25. Leth-Larsen R, Garred P, Jensenius H, Meschi J, Hartshorn K, Madsen J, et al. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005; 174: 1532–8.
26. Wang G, Phelps DS, Umstead TM, Floros J. Human SP-A protein variants derived from one or both genes stimulate TNF-alpha production in the THP-1 cell line. *Am J Physiol Lung Cell Mol Physiol* 2000; 278(5): L946–54.
27. Wang G, Umstead TM, Phelps DS, Al-Mondhiry H, Floros J. The effect of ozone exposure on the ability of human surfactant protein a variants to stimulate cytokine production. *Environ Health Perspect* 2002; 110: 79–84.
28. Huang W, Wang G, Phelps DS, Al-Mondhiry H, Floros J. Human SP-A genetic variants and bleomycin-induced cytokine production by THP-1 cells: effect of ozone-induced SP-A oxidation. *Am J Physiol Lung Cell Mol Physiol* 2004; 286: L546–53.
29. Mikerov AN, Umstead TM, Huang W, Liu W, Phelps DS, Floros J. SP-A1 and SP-A2 variants differentially enhance association of *Pseudomonas aeruginosa* with rat alveolar macrophages. *Am J Physiol Lung Cell Mol Physiol* 2005; 288: L150–8.

30. Wang G, Bates-Kenney SR, Tao JQ, Phelps DS, Floros J. Differences in biochemical properties and in biological function between human SP-A1 and SP-A2 variants, and the impact of ozone-induced oxidation. *Biochemistry* 2004; 43: 4227–39.
31. Choi EH, Ehrmantraut M, Foster CB, Moss J, Chanock SJ. Association of common haplotypes of surfactant protein A1 and A2 (SFTPA1 and SFTPA2) genes with severity of lung disease in cystic fibrosis. *Pediatr Pulmonol* 2006; 41: 255–62.
32. Lofgren J, Ramet 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. Lahti M, Lofgren J, Marttila R, Renko M, Klaavuniemi T, Haataja R, et al. Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 2002; 51(6): 696–9.
34. Ramet M, Lofgren J, Alho OP, Hallman M. Surfactant protein-A gene locus associated with recurrent otitis media. *J Pediatr* 2001; 138: 266–8.
35. Selman M, Lin HM, Montano M, Jenkins AL, Estrada A, Lin Z, et al. Surfactant protein A and B genetic variants predispose to idiopathic pulmonary fibrosis. *Hum Genet* 2003; 113: 542–50.
36. Floros J, Lin HM, Garcia A, Salazar MA, DiAngelo S, Montano M, et al. Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis* 2000; 182: 1473–8.
37. Seifart C, Lin HM, Seifart U, Plagens A, DiAngelo S, von Wichert P, et al. Rare SP-A alleles and the SP-A1–6A(4) allele associate with risk for lung carcinoma. *Clin Genet* 2005; 68: 128–36.