

Original Article

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

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, Rämetsä M, Merrill J. 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: 178–187. © Munksgaard 2001

A key cause of respiratory distress syndrome (RDS) in the prematurely born infant is deficiency of pulmonary surfactant, a lipoprotein complex. Both low levels of surfactant protein A (SP-A) and SP-A alleles have been associated with RDS. Using the candidate gene approach, we performed family-based linkage studies to discern linkage of SP-A to RDS and identify SP-A susceptibility or protective alleles. Moreover, we performed case–control studies of whites and blacks to detect association between RDS and SP-A alleles. Transmission disequilibrium test (TDT) analysis revealed that the frequency of transmission (from parent to the offspring with RDS) of alleles 6A² and 1A⁰ and of 1A⁰/6A² haplotype in RDS was increased, whereas transmission of alleles 1A⁵ and 6A⁴ and of haplotype 1A⁵/6A⁴ was decreased. Extended TDT analysis further strengthened the observations made. The case–control studies showed that in whites or blacks with RDS the frequencies of specific genotypes, 1A⁰ and 6A² or 1A⁰, were increased, respectively, but the frequency of specific 6A³ genotypes was increased in certain white subgroups and decreased in blacks. Regression analysis revealed gestational age (GA) and 6A³ genotypes are significant factors in blacks with RDS. In whites with RDS, GA and antenatal steroids are important factors. The data together indicate linkage between SP-A and RDS; certain SP-A alleles/haplotypes are susceptibility (1A⁰, 6A², 1A⁰/6A²) or protective (1A⁵, 6A⁴, 1A⁵/6A⁴) factors for RDS. Some differences between blacks and whites with regard to SP-A alleles may exist.

Abbreviations: DPPC, disaturated phosphatidylcholine – ETDT, extended transmission disequilibrium test – GA, gestational age – RDS, respiratory distress syndrome – SP-A, surfactant protein A – SP-A1, surfactant protein A1 gene – SP-A2, surfactant protein A2 gene – SP-B, surfactant protein B – SP-C, surfactant protein C – SP-D, surfactant protein D – TDT, transmission disequilibrium test – TGF- α , transforming growth factor α

¹ Work performed while at Pennsylvania State University.

**J Floros^{a,b}, R Fan^c,
A Matthews^c, S DiAngelo^a,
J Luo^a, H Nielsen^d, M Dunn^e, I
H Gewolb^f, J Koppe^g, L van
Sonderen^g,
L Farri-Kostopoulos^h, M Tzakiⁱ,
M Rämetsä^j and J Merrill^k**

^a Departments of Cellular and Molecular Physiology, ^b Pediatrics, and ^c Health Evaluation Sciences, The Pennsylvania State University College of Medicine, Hershey, PA, USA, ^d Department of Pediatrics, New England Medical Center, Boston, MA, ^e Department of Pediatrics, University of Toronto, Toronto, Canada, ^f Department of Pediatrics, University of Maryland, Baltimore, MD, USA, ^g Department of Neonatology, Emma Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, ^h Department of Pediatrics, University of Patras, Patras, Greece, ⁱ Department of Neonatology, Elena Hospital, Athens, Greece, ^j Department of Pediatrics and Biocenter Oulu, University of Oulu, Oulu, Finland, ^k Department of Pediatrics, University of Pennsylvania, Philadelphia, PA, USA

Key words: ETDT, extended TDT – race differences – RDS, respiratory distress syndrome – SP-A variants – TDT, transmission disequilibrium test

Corresponding author: Joanna Floros, Ph.D., Professor, Department of Cellular and Molecular Physiology, M.C. #H166, The Milton S. Hershey Medical Center, The Pennsylvania State University College of Medicine, 500 University Drive, P.O. Box 850, Hershey, PA 17033, USA. Tel: +1 717 531 6972; fax: +1 717 531 7667; e-mail: jfloros@psu.edu

Received 20 April 2001, revised and accepted for publication 28 June 2001

A number of factors may predispose the prematurely born infant to neonatal respiratory distress syndrome (RDS), which can result in significant morbidity and mortality of the neonate. Pulmonary surfactant, which is a lipoprotein complex and the biochemical hallmark of lung maturity, is essential for normal lung function (1). Deficiency of surfactant leads to RDS in the prematurely born infant. Measurements of surfactant components in amniotic fluid have been used to predict RDS and/or have been associated with the severity of RDS (2–4). Moreover, steroid therapy of mothers who threaten to deliver prematurely is used clinically, in an attempt to accelerate fetal lung maturity and surfactant production. Although the general belief is that maternal steroid therapy is beneficial, it has not led to the elimination of RDS (5).

Prematurity, race, sex, and genetic components are among the etiologic contributors to RDS (6, 7). Black infants, although they have a higher rate for prematurity, appear to have lower incidence of RDS compared to their white counterparts (8). These racial differences are not due to quantitative differences in surfactant lipids (3, 9), but may be due to surfactant protein differences (10–12) or other race-related factors (13). Genetics may play a role in the pathogenesis of RDS (6).

The surfactant proteins (SP)-A, SP-B, SP-C, and SP-D play key roles both in surfactant functions, such as alveolar integrity, by preventing alveolar collapse at low lung volumes, and in the innate host defense and the regulation of inflammatory processes in the lung (1, 14, 15). Infection is an important complication of the prematurely born infant. Therefore, aberration in surfactant or its components may affect both alveolar integrity and host defense in the prematurely born infant.

The human SP-A locus consists of one pseudogene and two functional genes, *SP-A1* and *SP-A2*, in opposite transcriptional orientation (16). A number of alleles that occur in > 1% in the population (17) have been identified and characterized for the *SP-A1* (6A, 6A², 6A³, 6A⁴) and the *SP-A2* (1A, 1A⁰, 1A¹, 1A², 1A³, 1A⁵) genes (15, 18, 19). Variability in the levels of SP-A mRNA has been observed in lungs of adult subjects (20, 21) and in fetal lung explant cultures (22). An association between low levels of SP-A mRNA and the SP-A 6A²6A²/1A⁰1A⁰ genotype has been observed in human lungs of unrelated individuals (21, 22). Although low levels of SP-A protein have been associated with the severity of RDS, genotype analysis was not performed in that study (4). Very low amounts of SP-A have also been shown by immunohistochemistry and electron microscopy in lung tissues of babies who died from RDS (23, 24). Candidate gene

approaches have revealed associations between SP-A and SP-B alleles and RDS in case-control association studies (10, 12, 25). The frequencies of SP-A alleles or genotypes shown previously to associate with low levels of SP-A mRNA (21, 22) were found to be increased in white infants with RDS of a mixed (12), or a homogeneous (Finnish) (25) ethnicity.

In the present study, we wished to not only build on previous findings, but also extend them by performing a family-based linkage study using transmission disequilibrium test (TDT) analysis and by carrying out a larger (n = 584) population-based case-control association study of control and RDS groups, including both white and black subjects (since race is thought to be a factor in RDS). TDT measures the transmission of marker alleles of heterozygous parents to the affected offspring. The non-transmitted parental marker alleles serve as an internal control to the transmitted alleles. Hence, TDT can avoid the problem of 'spurious' associations arising from conventional case-control studies due to admixture, heterogeneity, or stratification of the population (26). Moreover, the extended TDT (ETDT) analysis (27) allows for composite test of all the alleles of a multi-allele marker locus such as the SP-A1 and SP-A2 loci and can further strengthen observations. Both TDT and ETDT analyses together may determine whether the SP-A locus is linked to RDS and confirm or refute susceptibility or protective SP-A allele associations observed in case-control studies.

Subjects and methods

Specimens

Blood samples, cheek swab, or discarded tissues were collected from subjects with or without RDS and parents of affected offspring following institutional guidelines for use of human tissues. Diagnosis of RDS was made by clinical criteria (flaring, grunting, and retraction) and verified by the reticulogranular pattern on X-ray. Individuals treated with surfactant therapy prophylactically were excluded from the study. Genomic DNA was extracted and served as template in PCR for genotype analysis. SP-A gene-specific amplifications and gene-specific allele determinations were performed using a non-radioactive, converted polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) approach, described previously (17).

Family-based TDT and extended TDT studies

The study group consisted of 32 families, each of which had genotype data for both parents and for at least 1 affected offspring. Table 1 shows the demographic data of the population studied.

For the 32 families, 10 of them had 1 offspring, 18 had 2, and 4 had 3. The total number of subjects studied was 122 (58 offspring and 64 parents). The race, sex, ethnicity, and age category of the subjects are shown in Table 1. There were a total of 47 affected offspring. Nineteen of the families had only 1, 11 had 2, and 2 had 3 affected offspring. There were 5 affected monozygous twin sets and 1 affected monozygous triplet set. There were 12 dizygous twin sets and 2 dizygous triplet sets in which at least 1 individual was affected. For TDT analysis to be valid, we analyzed only 1 affected offspring from each of the 5 affected monozygous twin sets and 1 affected monozygous from the triplet set, (i.e. we removed 7 affected offspring and used the remaining 40 affected offspring in the study).

For each SP-A1 or SP-A2 allele, the number of alleles transmitted indicates the times the allele was transmitted to the offspring by the heterozygous parent(s). For this analysis, only parents heterozygous for the SP-A alleles were used. Each heterozygous parent was used in turn to create a new bi-allelic system for each SP-A1 or SP-A2 allele under study, by collapsing the alleles, other than the one under examination, to a new allele category X. For example, if the transmission of allele 6A/X was being studied, the remaining SP-A1 alleles 6A², 6A³, and 6A⁴ were collapsed to form a new allele category X. Therefore, the data in Figs 1 and 2 indicate the frequency with which heterozygous parents with, for example, genotype

Table 1. Characteristics of population used in the TDT study

| | | |
|--|------------------------|-----|
| No. of families | 32 | |
| No. of families (No. of children) | 10 (1), 18 (2), 4 (3) | |
| No. of families (No. of affected children) | 19 (1), 11 (2), 2 (3) | |
| Zygote | Monozygotic twins | 5 |
| | Monozygotic triplets | 1 |
| | Dizygotic twins | 12 |
| | Dizygotic triplets | 2 |
| Race | Black, no Hispanic | 12 |
| | White, no Hispanic | 105 |
| | Other or Mixed parents | 3 |
| | Missing | 2 |
| Sex | Female | 60 |
| | Male | 62 |
| Ethnicity | American | 72 |
| | The Netherlands | 47 |
| | Missing | 3 |
| Age category | Adult | 64 |
| | Child | 9 |
| | Newborn | 49 |

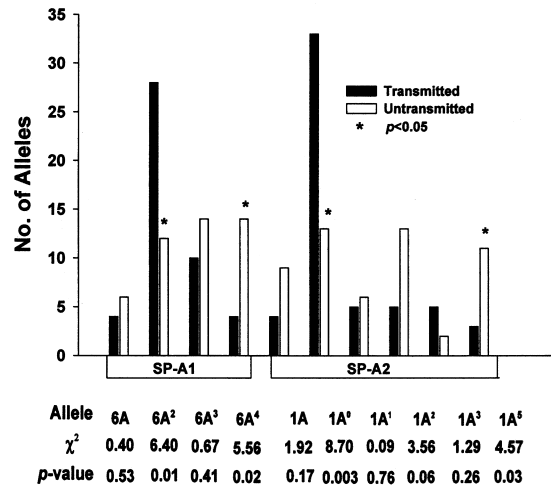


Fig. 1. TDT analysis of SP-A1 and SP-A2 alleles. Thirty-two families with both parents and at least one affected child were used in this analysis. The number of transmitted and untransmitted alleles for 6A, 6A², etc., are shown along with the corresponding χ^2 and p-value.

6A/X or haplotype 1A⁰6A²/X transmitted allele 6A or haplotype 1A⁰6A² to their affected offspring. Thus, from the 10 heterozygous parents with genotype 6A/X (Fig. 1), 4 transmitted allele 6A, and 6 transmitted allele X to their affected offspring.

ETDT, which allows for a composite test analysis of all alleles of a multi-allele locus, such as SP-A1 and SP-A2 loci, was performed (27). In this analysis, we tested the null hypothesis of no linkage between RDS and SP-A loci using parsimonious and saturated models. Moreover, we tested the goodness-of-fit of the parsimonious model by comparing it to the saturated model.

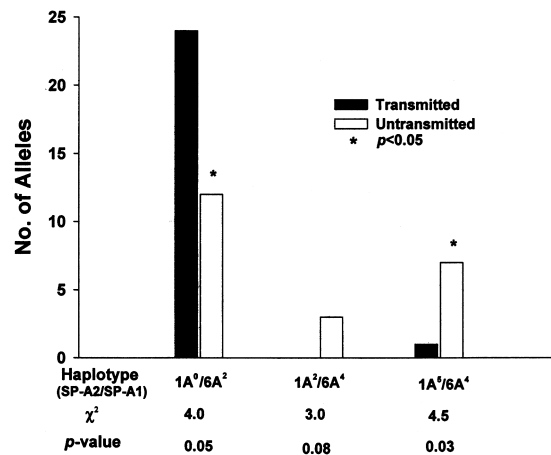


Fig. 2. TDT analysis of SP-A haplotypes. Thirty-two families with both parents and at least one affected child were used in the analysis. The transmission of haplotypes (p < 0.1) is shown along with the χ^2 .

Case-control associations: characteristics of the study group

The study group included a total of 584 prematurely born white or black babies of both sexes, with and without RDS, born to mothers who either did or did not receive prenatal glucocorticoid therapy. The characteristics of the study group are shown in Table 2.

We compared the frequency of alleles, shown previously to be significant in control and RDS groups of white American and Finnish (12, 25). For our analysis we took into consideration race, sex, gestational age (GA), and maternal steroid therapy (Table 2). The GAs between RDS and control subjects in white and blacks are not well matched (Fig. 3). To better match GAs between RDS and control groups, we kept for the analysis of whites, subjects with GA > 28 weeks and for blacks, subjects with 28 weeks < GA < 35 weeks. In addition, we carried out analyses with GA sub-

groups that further matched GA for control and RDS groups. For whites, we performed analyses with individuals of 28 weeks < GA < 33 weeks and GA ≥ 33 weeks and for blacks, with individuals of 28 weeks < GA < 31 weeks and 31 weeks ≤ GA ≤ 35 weeks. Similarly, we conducted analyses with individuals regardless of maternal steroid therapy and then again by separating them into those who received maternal steroid therapy and those who did not.

Statistical analysis

The family-based TDT analysis utilized the GENEHUNTER program from the Whitehead Institute for Biomedical Research, MIT. The TDT statistic is the McNemar's test. The ETDT program is from Dr Curtis, <http://www.gene.ucl.ac.uk/~dcurtis/>.

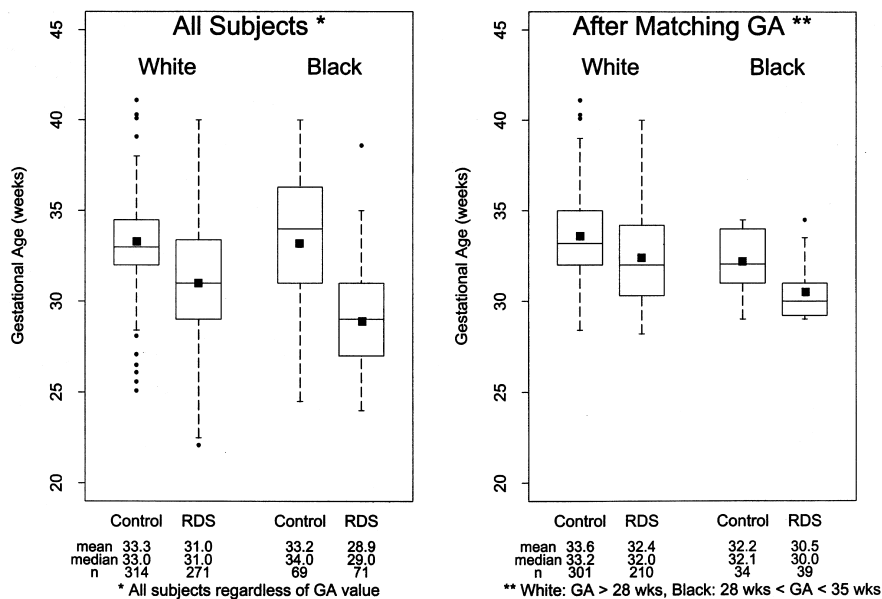
Analysis of the case-control associations used SAS Version 7 to perform Fisher's exact tests to

Table 2. Characteristics of case-control study population

| Race | Group | No. of subject | No. in sex | No. in steroid treatment† | GA # | Mean of GA | SD of GA |
|-------|---------|----------------|------------|---------------------------|--------------|------------|----------|
| White | Control | 301 | 129F:172M | 181S:120NS | 28 < GA | 33.6 | 2.5 |
| | RDS | 210 | 84F:124M* | 102S:107NS** | 28 < GA | 32.4 | 2.6 |
| Black | Control | 34 | 21F:12M*** | 26S:8NS | 28 < GA < 35 | 32.2 | 1.6 |
| | RDS | 39 | 17F:22M | 24S:15NS | 28 < GA < 35 | 30.5 | 1.5 |

† Maternal steroid therapy received 24–168 h before birth.
 F: female; M: male; S: steroid treatment; NS: no steroid treatment.
 GA #: Gestational age in weeks.
 * The sex of 2 subjects is missing.
 ** One subject's steroid status is missing.
 *** The sex of one subject is missing.

Fig. 3. GA distribution in whites and blacks. The boxplot depicts the distribution of gestational ages in whites, in control and RDS groups, before (left panel) and after (right panel) GA adjustment. The horizontal line within each box indicates the median GA, and the solid black box represents the mean GA. The number (n) of subjects in each category is indicated. The lower and upper broken horizontal lines of the box represent the quartiles of the data or the 25th and 75th percentile of the GA dataset, respectively. The size of the entire box represents the interquartile range and the solid circles represent outliers.



assess association of alleles or genotypes with RDS. A p-value of ≤ 0.05 is considered to be significant. Logistic regression analysis was also conducted, using GA, sex, steroid treatment, and genotypes ($1A^0/1A^0$ or $1A^0/*$, $6A^2/6A^2$ or $6A^2/*$, $6A^3/6A^3$ or $6A^3/*$ separately), as possible regressors.

Results

TDT analysis of SP-A alleles and haplotypes

Based on nucleotide differences, four SP-A1 alleles ($6A$, $6A^2$, $6A^3$, $6A^4$) and eight SP-A2 alleles ($1A$, $1A^0$, $1A^1$, $1A^2$, $1A^3$, $1A^5$, $1A^6$, $1A^9$) were observed in the study population. These 12 alleles were used to test whether there is linkage between RDS and the SP-A1 and SP-A2 loci. Of these alleles, all, except $1A^6$ and $1A^9$, occur in $> 1\%$ of the general population (17). Fig. 1 shows the results of the TDT analysis.

Of the marker loci, SP-A1 and SP-A2, alleles $6A^2$, $6A^4$, $1A^0$, and $1A^5$ show linkage with a potential RDS allele (p-values 0.01, 0.02, 0.003, 0.03, respectively). Based on the high frequency of transmission of alleles $6A^2$ and $1A^0$ to the affected child, these alleles may be viewed as linked susceptibility factors for RDS. The low transmission of alleles $1A^5$ and $6A^4$ indicates that these alleles may be viewed as protective factors for RDS.

The SP-A1 and SP-A2 loci are linked and based on both our published (19) and unpublished data of CEPH (Centre d'Etude du Polymorphisme Humaine) family pedigree analysis, the $6A^2$ and $1A^0$ alleles represent a common genomic haplotype. DNAs for CEPH pedigrees are commercially available and, because these have multiple members, are suited for haplotype analysis. Fig. 2 presents the TDT analysis of haplotypes that were transmitted with higher or lower frequency to the offspring ($p < 0.1$). Haplotype $1A^0/6A^2$ is over-represented in RDS ($p = 0.05$), but haplotype $1A^5/6A^4$ is under-represented in RDS ($p = 0.03$), indicating that the former is a susceptibility factor and the latter a protective factor for RDS. Other haplotypes present in the families studied, but did not show a differential transmission ($p > 0.1$) to the affected offspring, include $1A/6A$, $1A^0/6A^3$, $1A^0/6A^4$, $1A^1/6A^2$, $1A^1/6A^3$, $1A^2/6A^2$, $1A^2/6A^3$, $1A^3/6A^2$.

ETDT analysis of SP-A1 and SP-A2 alleles

We performed ETDT analysis in order to further assess linkage of the SP-A1 and SP-A2 multi-allele loci with the RDS locus. For ETDT, we tested the hypothesis of no linkage between RDS and SP-A loci by examining two models, a parsimonious

(allele-wise) and a saturated (genotype-wise). The log likelihoods and the respective chi-square p-values are shown. The goodness-of-fit of the parsimonious model when compared to the saturated model is also shown for each SP-A1 and SP-A2 locus.

For SP-A1 the results are:

Log likelihood under null hypothesis:

$$L0 = -31.89$$

Log likelihood using parsimonious (allele-wise) model: $L1 = -27.83$

Log likelihood using saturated (genotype-wise) model: $L2 = -23.01$

Chi-squared for allele-wise TDT
 $= 2 * (L1 - L0) = 8.10$, 3 df, $p = 0.04$

Chi-squared for genotype-wise TDT
 $= 2 * (L2 - L0) = 17.74$, 5 df, $p = 0.003$

Chi-squared for goodness-of-fit of allele-wise model = $2 * (L2 - L1) = 9.64$, 2 df, $p = 0.01$.

For SP-A2, the results are:

Log likelihood under null hypothesis:

$$L0 = -38.82$$

Log likelihood using parsimonious (allele-wise) model: $L1 = -30.35$

Log likelihood using saturated (genotype-wise) model: $L2 = -21.73$

Chi-squared for allele-wise TDT
 $= 2 * (L1 - L0) = 16.93$, 7 df, $p = 0.02$

Chi-squared for genotype-wise TDT
 $= 2 * (L2 - L0) = 34.18$, 11 df, $p = 0.0003$

Chi-squared for goodness-of-fit of allele-wise model = $2 * (L2 - L1) = 17.25$, 4 df, $p = 0.002$

These results strengthen considerably our findings in Figs 1 and 2, namely, the SP-A1 and SP-A2 loci are linked to the RDS locus. The goodness-of-fit indicated that the saturated (genotype-wise) model has a better explanatory power.

Frequency of SP-A genotypes

Comparisons between control and RDS genotypes were made by considering race, GA, sex, and maternal steroid therapy.

White subjects. The significant findings for white subjects are shown in Tables 3 and 4. The data in Table 3 indicate that the frequency of $1A^0$ allele versus all other alleles is higher in babies with RDS and GA > 28 weeks, regardless of whether maternal steroid therapy was considered as a factor. The frequency of $6A^3/6A^3$ genotype versus $6A^3/*$ was higher in subjects with RDS (GA > 28 weeks) who received maternal steroid therapy. For the $1A^0/1A^0$

Table 3. Frequency of allele 1A⁰ and 6A³/6A³ or 1A⁰/1A⁰ genotypes in whites with and without RDS

| Genotype comparison | Steroid† | Sex | GA # | RDS | | Control | | p-value | Odds ratio | 95% confidence interval |
|--|----------|------|------|------|-----|---------|-----|---------|------------|-------------------------|
| | | | | f | n | f | n | | | |
| 1A ⁰ allele vs. all others | +, - | F, M | >28 | 0.57 | 420 | 0.52 | 602 | 0.04 | 1.3 | 1.0, 1.6 |
| 1A ⁰ /1A ⁰ vs. all other genotypes | + | F, M | >28 | 0.34 | 102 | 0.24 | 181 | 0.05 | 1.6 | 1.0, 2.8 |
| 6A ³ /6A ³ vs. 6A ³ /* | + | F, M | >28 | 0.26 | 38 | 0.11 | 80 | 0.04 | 2.8 | 1.0, 7.7 |

† Maternal steroid therapy received 24–168 h before birth; the symbol (+, -) indicates that some of the subjects under study had (+) steroid therapy and others did not (-).

GA # = gestational age in weeks; F = female; M = male.

f = frequency, allele 1A⁰ vs. all other alleles (first row); Accordingly, in other rows, n = number of individuals or alleles.

The frequency of 1A⁰/1A⁰ genotype was compared to that of the combined frequencies of 'all other' genotypes in control and RDS.

The frequency of 6A³/6A³ genotype was compared to that of 6A³/* in control and RDS.

Table 4. Frequency of allele 6A² or 6A²/6A² genotype in whites with and without RDS

| Genotype comparison | Steroid† | Sex | GA # | RDS | | Control | | p-value | Odds ratio | 95% confidence interval |
|--|----------|------|--------------|------|-----|---------|-----|---------|------------|-------------------------|
| | | | | f | n | f | n | | | |
| 6A ² allele vs. all others | +, - | F | >28 | 0.61 | 168 | 0.52 | 258 | 0.05 | 1.4 | 1.0, 2.1 |
| 6A ² allele vs. all others | +, - | F, M | 28 < GA < 33 | 0.60 | 254 | 0.52 | 222 | 0.05 | 1.4 | 1.0, 2.0 |
| 6A ² /6A ² vs. */* | +, - | F, M | 28 < GA < 33 | 0.72 | 60 | 0.55 | 53 | 0.05 | 2.1 | 1.0, 4.6 |
| 6A ² /6A ² vs. all other genotypes | + | F, M | >28 | 0.35 | 102 | 0.24 | 181 | 0.03 | 1.8 | 1.0, 3.0 |
| 6A ² /6A ² vs. all other genotypes | +, - | F | >28 | 0.38 | 84 | 0.25 | 129 | 0.03 | 1.9 | 1.0, 3.4 |
| 6A ² /6A ² vs. 6A ² /* | + | F, M | >28 | 0.43 | 83 | 0.29 | 147 | 0.02 | 1.9 | 1.1, 3.3 |
| 6A ² /6A ² vs. 6A ² /* | +, - | F | >28 | 0.46 | 70 | 0.31 | 103 | 0.04 | 1.9 | 1.0, 3.5 |

† Maternal steroid therapy received 24–168 h before birth; the symbol (+, -) indicates that some of the subjects under study had (+) steroid therapy and others did not (-).

GA # = gestational age in weeks; F = female; M = male.

f = frequency; n = number of individuals or alleles.

The frequency of allele 6A² was compared to that of all other alleles in control and RDS. Similar comparisons were made for the 6A²/6A² genotype, etc.

genotype, comparable observations were made for the subgroup that received steroid therapy. The frequency of the 6A² allele or 6A²/6A² genotype was higher in female babies with RDS and GA > 28 weeks, and in babies with RDS and 28 weeks < GA < 33 weeks, regardless of steroid therapy (Table 4). The frequency of the 6A²/6A² genotype was higher in babies with RDS and GA > 28 weeks and of mothers who received or did not receive steroid therapy (Table 4).

In an attempt to distinguish whether alleles 1A⁰, 6A², and 6A³ associate with prematurity rather than RDS, we determined the frequency of these alleles in control and RDS for two GA subgroups (< 32 weeks, ≥ 32 weeks). No significant difference was observed between these two groups either in control or RDS subjects (Fig. 4). Similar results were obtained when GA of 33 or 34 weeks was used to separate the two subgroups (not shown).

Black subjects. The significant findings for black subjects are shown in Table 5. The 1A⁰ allele and the (1A⁰/1A⁰ or 1A⁰/*) genotype are more frequent in RDS babies with younger GAs (28 weeks <

GA < 31 weeks), whereas the (6A³/6A³ or 6A³/*) genotype or the 6A³/* genotype alone was higher in control babies (31 ≤ GA < 35).

Regression analysis. The logistic regression analysis (Table 6) for whites revealed that both GA and maternal steroid therapy were important variables (p < 0.0001). In blacks, GA was similarly found to be important (p = 0.0003), but in contrast to

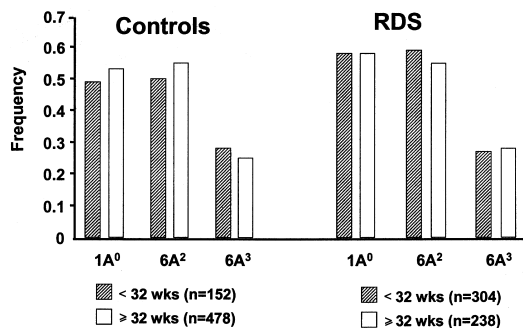


Fig. 4. Allele distribution in age-dependent subgroups (< 32 weeks, ≥ 32 weeks) in control and RDS groups. n denotes the number of subjects used.

Table 5. Frequency of SP-A alleles or genotypes in blacks with and without RDS

| Genotype comparison | Steroid† | GA # | Sex | RDS | | Control | | p-value | Odds ratio | 95% confidence interval |
|--|----------|----------|------|------|----|---------|----|---------|------------|-------------------------|
| | | | | f | n | f | n | | | |
| 1A ⁰ allele vs. all others | +, - | 28<GA<31 | F, M | 0.30 | 46 | 0.0 | 10 | 0.04 | N/A | N/A |
| (1A ⁰ /1A ⁰ or 1A ⁰ /*) vs. */* | +, - | 28<GA<31 | F, M | 0.52 | 23 | 0.0 | 5 | 0.04 | N/A | N/A |
| (6A ³ /6A ³ or 6A ³ /*) vs. */* | +, - | 31≤GA<35 | F, M | 0.38 | 16 | 0.69 | 29 | 0.04 | 0.3 | 0.1, 1.0 |
| 6A ³ /* vs. */* | +, - | 31≤GA<35 | F, M | 0.33 | 15 | 0.67 | 27 | 0.04 | 0.3 | 0.1, 1.0 |

† Maternal steroid therapy received 24–168 h before birth. The symbol (+, -) indicates that some of the subjects under study had (+) steroid therapy and others did not (-).

F: female; M: male; f: frequency; n: number of individuals, GA #: gestational age in weeks.

Table 6. Logistic regression analysis of 6A³ genotype in blacks and whites with and without RDS

| Variables | Est. | SD | Odds ratio | 95% confidence interval | p-value | Est. | SD | Odds ratio | 95% confidence interval | p-value |
|---|-------|------|------------|-------------------------|---------|-------|------|------------|-------------------------|---------|
| | | | | | | | | | | |
| Ind_GA | 2.34 | 0.65 | 10.38 | 2.90, 37.09 | 0.0003 | 1.32 | 0.21 | 3.75 | 2.47, 5.70 | <0.0001 |
| Ind_male | 0.88 | 0.60 | 2.40 | 0.75, 7.71 | 0.14 | 0.07 | 0.19 | 1.07 | 0.73, 1.56 | 0.73 |
| Ind_steroids | -1.12 | 0.63 | 0.33 | 0.10, 1.13 | 0.08 | -0.98 | 0.22 | 0.38 | 0.25, 0.58 | <0.0001 |
| Ind_(6A ³ /* or 6A ³ /6A ³) | -1.40 | 0.63 | 0.25 | 0.07, 0.84 | 0.03 | -0.14 | 0.19 | 0.87 | 0.60, 1.27 | 0.48 |

The logistic regression model is given by $\log(P(\text{RDS})/P(\text{Control})) = \alpha_0 + \alpha_1 \text{Ind_GA} + \alpha_2 \text{Ind_male} + \alpha_3 \text{Ind_steroids} + \alpha_4 \text{Ind_}(6A^3/* \text{ or } 6A^3/6A^3)$, where Ind_GA is 1 if 28<GA<31 and 0 if 31≤GA<35 for blacks, 1 if 28<GA<33 and 0 if 33≤GA for whites, Ind_male is 1 if male and 0 if female, Ind_steroids is 1 if steroid treatment is Yes and 0 if steroid treatment is No, Ind_(6A³/* or 6A³/6A³) is 1 if the genotype is 6A³/* or 6A³/6A³ and 0 if the genotype is */*.

* represent alleles other than 6A³.

Est. = estimates of α ; SD = standard deviation.

whites the (6A³/6A³ or 6A³/*) genotype was also important (p = 0.03). The maternal steroid therapy on the other hand did not appear to be as important (p = 0.08) in blacks. Given the smaller number of samples in blacks, further study with a larger sample size is warranted.

Discussion

The etiology of neonatal RDS is considered to be multigenic and/or multifactorial, with a major etiologic cause being pulmonary surfactant deficiency (6). Both low levels of SP-A (4, 23) and certain SP-A alleles have been associated with RDS in white subjects (12, 25). In the present study, we wished to build on existing findings and extend them by performing TDT and ETDT analyses and by carrying out a larger size case-control study of 584 individuals including both white and black subjects. The TDT analysis results indicate a significantly high transmission of 6A² or 1A⁰ allele or 1A⁰/6A² haplotype and low transmission of alleles 1A⁵, 6A⁴, or 1A⁵/6A⁴ haplotype to the affected child. The ETDT analysis supports and strengthens these observations by showing that *SP-A1* and *SP-A2* are linked to the RDS locus. The case-control studies indicate that the 1A⁰ and 6A² alleles (or specific 6A² genotypes) are over-represented in subgroups of white babies with RDS regardless of

maternal steroid therapy, whereas genotypes 1A⁰/1A⁰ and 6A³/6A³ are over-represented in RDS babies of mothers who received steroid therapy. The 6A² and 6A²/6A² appear also to be over-represented in females with RDS, with or without steroid therapy. Similar findings for males were not observed. In blacks with RDS, allele 1A⁰ or specific 1A⁰ genotypes are over-represented but 6A³ genotypes are under-represented, suggesting risk and protective factors, respectively. The latter is consistent with the regression analysis data, where the genotype (6A³/* or 6A³/6A³), (p = 0.03) and GA (p = 0.0003) were found to be protective and risk factors, respectively, in blacks with RDS whereas in whites maternal steroid therapy and GA were major risk factors (p < 0.0001). Because none of these alleles (1A⁰, 6A², 6A³) were associated with prematurity, it is likely that certain SP-A alleles, haplotypes, or genotypes are susceptibility or protective factors for RDS in whites or blacks. Although the data presented indicate some differences between whites and blacks, this warrants further confirmation given the small number of black subjects, and caution should be exercised with conclusions regarding race differences.

The data from TDT and ETDT analyses clearly link the SP-A locus to RDS, and indicate that the associations of 1A⁰ and 6A² alleles observed in case-control studies are not due to population

admixture, heterogeneity, or stratification (26, 28). The power of TDT stems from the fact that the non-transmitted allele serves as an internal control. Because genetically the parent and child are far more homogeneous than any other two unrelated individuals, 'spurious' associations can be avoided. TDT is also useful in the study of factors contributing to the pathogenesis of disease with multi-genetic etiology because it can detect small gene effects (29, 30).

Apparent discrepancies regarding the role of SP-A in normal lung function exist, as assessed by studies of SP-A (–/–) deficient mice born under normal conditions at term (31), and by *in vitro* studies (32). These may point to a developmental context-dependent role of SP-A (33) and/or genetic differences between mouse and human SP-A sequences. With regards to the former possibility several points may be considered. For example, the macromolecular context of the premature lung may differ from that at term. The surfactant components or other molecules may be at suboptimal concentrations (i.e. at a quantitative imbalance (33)). Reduced levels of surfactant lipids (2, 34) and surfactant proteins (23, 24, 35) have been observed in babies with or at risk for RDS. Thus, SP-A may have an impact and/or modify the severity of RDS depending on the presence and/or the network of interactions of other factors in the cellular or tissue environment (6, 28).

With regards to genetic differences between mouse and human SP-A sequences, we performed amino acid and nucleotide sequence comparisons between SP-A alleles and mouse sequence (GenBank accession # AK004788) to assess potential qualitative and quantitative differences, respectively. We compared a) the human 'core' (i.e. gene-specific) amino acids that distinguish all SP-A1 alleles from the SP-A2 alleles (36) and the corresponding mouse amino acids; and b) all amino acids that differ among the human alleles and the corresponding mouse amino acids. The results showed that the mouse SP-A is more similar to the SP-A1 gene and of the SP-A1 alleles, is more similar to the 6A, 6A² (a risk factor), and 6A³ (a protective factor in certain race/ethnic groups). Moreover, when all SP-A amino acids of the human SP-A alleles and the mouse SP-A were compared, the mouse SP-A was still more similar to 6A as well as to 6A² allele. To identify potential quantitative differences among the significant alleles, we compared nucleotide differences of the 3'UTR (untranslated region), because the 3'UTR is shown to mediate differences in basal levels and in response to dexamethasone among SP-A alleles (37, 38). The results showed that the mouse SP-A 3'UTR

sequence has the least nucleotide similarity with 6A² compared to all other alleles and the most with 6A and (a very close second) with RDS protective 6A³ and 6A⁴ alleles. Thus, the comparison data indicate that the mouse gene is not similar to 1A⁰ (risk factor), as determined by amino acid or 3'UTR nucleotide comparison. Although the mouse gene by amino acid sequence comparison is similar to 6A² (a risk factor), 6A, and 6A³ (a protective factor in certain race/ethnic groups) the mouse gene, by nucleotide 3'UTR comparison is most dissimilar with 6A² and most similar with 6A and with 6A³ and 6A⁴ alleles. Therefore, based on sequence comparison data, it is possible that genetically based SP-A quantitative differences along with other environmental differences, may explain some of the apparent discrepancies between the SP-A knockout mouse and humans.

Differences in SP-A mRNA levels (20, 21) and SP-A protein levels (39–41) have been observed among humans, and correlations between SP-A, 6A²6A²:1A⁰1A⁰, and low levels of SP-A mRNA have been made in both human lungs (21) and fetal lung explants (22). Moreover, the 1A⁰/6A² haplotype has been associated with RDS. These together allow one to speculate that infants born prematurely with 6A²6A²:1A⁰1A⁰ SP-A genotype have a higher risk of developing RDS compared to their counterparts with a different SP-A genotype. This genotype-dependent low concentration of SP-A may in turn compromise important lung functions. This speculation is supported in part by previous studies. SP-A content has been both correlated with the severity in RDS (4) and used to predict risk for RDS (34) and survival (42). Tissues of babies who died from RDS (23, 24) had low SP-A content. It is currently unknown whether developmental differences in the regulation of SP-A alleles exist. However, when SV40 promoter driven constructs containing the 3'UTR of 1A⁰, 6A², or 6A³ alleles were transiently transfected into lung adenocarcinoma NCI-H441 cells, the results showed that the 3'UTR mediates differences among alleles with regard to basal levels and in response to dexamethasone (37, 38).

Prematurity, race, and sex may be important contributing factors to RDS (6) and based on regression analysis, GA is a major variable for both blacks and whites. We minimized GA differences by matching control and RDS subjects within one and less than 2 weeks difference for whites and blacks, respectively. Comparable results were obtained for 1A⁰1A⁰ and 6A²6A² genotypes from a Finnish study group where the control and RDS subjects were also matched for GA within 1 week (25). Differences between races may exist, as

assessed by regression analysis, with regard to maternal steroid therapy and the contribution of SP-A alleles in the clinical outcome. Steroid therapy was found to be a major variable for whites but not for blacks under study, while the genotype (6A³/6A³ or 6A³/*) appears to be a protective factor for blacks but not for whites. The 6A² allele and genotype are over-represented in white females (but not in black females or males of both races) with RDS and with or without steroid therapy. The underlying mechanisms for these putative sex differences in whites are unknown. It is expected that multiple genetic and non-genetic factors contribute to and/or modify the severity of RDS, acting perhaps through networks of additive and/or epistatic interactions (33, 43, 44) at the cellular and/or tissue level.

Findings from previous case-control studies have implicated allele 6A³ as a protective factor in whites with RDS (25) and in blacks (but not in whites) with RDS (12). In the present study, comparable findings to those reported by Kala et al. (12) were observed. The frequency of the 6A³ genotype in blacks was increased in controls and in whites was increased in babies with RDS of mothers who had received steroid therapy. We speculate that the apparent discrepancy in whites of the 6A³/6A³ genotype among the case-control studies (present study, (12, 25)) is due to differences in the study group stratification such as the presence or absence of steroid therapy, or ethnicity.

Although no significant associations were made with either allele 1A⁵, or 6A⁴ and RDS in any of our case-control studies (12, 25), the transmission of 1A⁵, or 6A⁴ and of 1A⁵/6A⁴ haplotype allele to the affected infant was found to be low by TDT analysis, indicating that these alleles and haplotype are protective factors for RDS. The reasons for these differences are unknown. It is possible that the protective effect of the 1A⁵, or 6A⁴ allele, observed by TDT analysis, is diluted in the case-control studies due to genetic background differences, and/or the protective effect of these alleles is pronounced in a subgroup of RDS, yet to be determined.

In summary, we have demonstrated by TDT, ETDT, and case-control analyses that the SP-A locus is linked to RDS and that certain SP-A alleles or haplotypes within a developmental and perhaps a hormonal-dependent context are susceptibility or protective factors for RDS. Differences between whites and blacks and between white females and males may exist with regard to the role of SP-A alleles and possibly the role of steroids in RDS.

Acknowledgements

The authors thank Mr Scott Phillips for technical assistance and Dr Hung-Mo Lin for statistical help at the early stages of this project, Drs Xiaoxuan Guo for help with genotyping, and Guirong Wang for sequence comparisons at The Pennsylvania State University, College of Medicine and Dr Carolyn Levine at the Winthrop University Hospital SUNY at Stony Brook for specimen contributions, and Drs Yannis Sofagis at the Elena Hospital, Greece, and Dimitris Kouretas at the University Thessaly, Greece, for initial discussions of aspects of the project, and Sue Myers for typing the manuscript. This work was supported by NIH R37 HL34788, and the GCRC at Pennsylvania State University.

References

1. Floros J, Phelps D. Pulmonary surfactant. In: Yaksch TL, Lynch C III, Maze M, Biebuyck JF, Saidman LJ, eds. *Anesthesia: Biological Foundations*. New York: Lippincott-Raven, 1997: 1257–1279.
2. Torday J, Carson L, Lawson EE. Saturated phosphatidylcholine in amniotic fluid and prediction of the respiratory distress syndrome. *N Engl J Med* 1979; 301: 1013–1018.
3. Richardson DK, Torday JS. Racial differences in predictive value of the lecithin/sphingomyelin ratio. *Am J Obstet Gynecol* 1994; 170: 1273–1278.
4. Moya FR, Montes HF, Thomas VL, Mouzinho AM, Smith JF, Rosenfeld CR. Surfactant protein A and saturated phosphatidylcholine in respiratory distress syndrome. *Am J Respir Crit Care Med* 1994; 150: 1672–1677.
5. NIH Consensus Conference. Effect of corticosteroids for fetal maturation on perinatal outcomes. *J Am Med Assoc* 1995; 273: 413–418.
6. Floros J, Kala P. Surfactant proteins: molecular genetics of neonatal pulmonary disease. *Annu Rev Physiol* 1998; 60: 365–384.
7. Floros J, Karinch AM. Genetics of neonatal disease as they relate to the surfactant protein genes. In: Robertson B, Taeusch HW, eds. *Surfactant Therapy for Lung Disease*. New York: Marcel Dekker, 1995: 95–106.
8. Hulsey TC, Alexander GR, Robillard PY, Annibale DJ, Keenan A. Hyaline membrane disease: the role of ethnicity and maternal risk characteristics. *Am J Obstet Gynecol* 1993; 168: 572–576.
9. Berman S, Tanasijevic MJ, Alvarez JF, Ludmir J. Racial differences in the predictive value of the TDx fetal lung maturity assay. *Am J Obstet Gynecol* 1996; 56: 1018–1022.
10. Floros J, Veletza SV, Kotikalapudi P et al. Dinucleotide repeats in the human surfactant protein B gene and respiratory distress syndrome. *Biochem J* 1995; 305: 583–590.
11. Veletza SV, Rogan PK, TenHave T, Olowe SA, Floros J. Racial differences in allelic distribution at the human pulmonary surfactant protein B gene locus (SP-B). *Exp Lung Res* 1996; 22: 489–494.
12. Kala P, TenHave T, Nielson H, Dunn M, Floros J. Association of pulmonary surfactant protein A (SP-A) and RDS: interaction with SP-B. *Pediatr Res* 1998; 43: 169–177.
13. Olowe SA, Akinkugbe A. Amniotic fluid lecithin/sphingomyelin ratio: comparison between an African and a North American community. *Pediatrics* 1978; 62: 38–41.
14. Phelps DS. Pulmonary surfactant modulation of host-defense function. *Appl Cardiopulm Pathophysiol* 1995; 5: 221–229.

15. Floros J, Hoover RR. Genetics of the hydrophilic surfactant proteins A and D. *Biochim Biophys Acta* 1998; 1408: 312–322.
16. 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: 353–362.
17. DiAngelo S, Lin Z, Wang G, Phillips S, Ramet M, Floros J. Novel, non-radioactive, simple and multiplex PCR-cR-FLP methods for genotyping human SP-A and SP-D gene alleles. *Dis Markers* 1999; 15: 269–281.
18. Floros J, Karinch AM. Human SP-A: then and now. *Am J Physiol* 1995; 268 (12): L162–L165.
19. Floros J, DiAngelo S, Koptides M et al. Human SP-A locus: allele frequencies and linkage disequilibrium between the two *surfactant protein A* genes. *Am J Respir Cell Mol Biol* 1996; 15: 489–498.
20. Floros J, Phelps DS, deMello DE et al. The utility of postmortem lung for RNA studies; variability and correlation of the expression of surfactant proteins in human lung. *Exper Lung Res* 1991; 17: 91–104.
21. Karinch AM, deMello DE, Floros J. Effect of genotype on the levels of surfactant protein-A mRNA and on the SP-A2 splice variants in adult humans. *Biochem J* 1997; 321: 39–47.
22. Karinch AM, Deiter G, Ballard PL, Floros J. Regulation of expression of human *SP-A1* and *SP-A2* genes in fetal lung explant culture. *Biochim Biophys Acta* 1998; 1398: 192–202.
23. deMello DE, Phelps DS, Patel G, Floros J, Lagunoff D. Expression of the 35 kDa and low molecular surfactant-associated proteins in the lungs of infants dying with respiratory distress syndrome. *Am J Pathol* 1989; 134: 1285–1293.
24. deMello DE, Hayman S, Phelps D, Floros J. Immunolocalization of SP-A in lungs of infants dying from respiratory distress syndrome. *Am J Pathol* 1993; 142: 1631–1640.
25. Rämetsä M, Haataja R, Marttila R, Floros J, Hallman M. Surfactant protein A (SP-A) gene locus associated with respiratory distress syndrome in Finnish population. *Am J Hum Genet* 2000; 66: 1569–1579.
26. Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 1996; 59: 983–989.
27. Sham PC, Curtis D. An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet* 1995; 59 (Pt 3): 323–336.
28. Schork NJ. Genetics of complex disease: approaches, problems and solutions. *Am J Respir Crit Care Med* 1997; 156: S103–S109.
29. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; 52: 506–516.
30. Risch N, Merikangas K. The future of genetic studies of complex human disease. *Science* 1996; 273: 1516–1517.
31. Korfhagen TR, Bruno MD, Ross GF et al. Altered surfactant function and structure in *SP-A* gene targeted mice. *Proc Natl Acad Sci USA* 1996; 93: 9594–9599.
32. Schürch S, Possmayer F, Cheng S, Cockshutt AM. Pulmonary SP-A enhances adsorption and appears to induce surface sorting of lipid extract surfactant. *Am J Physiol* 1992; 263 (7): L210–L218.
33. Floros J, Wang G. A point of view: quantitative and qualitative imbalance in disease pathogenesis; pulmonary surfactant protein A genetic variants as a model. *Comp Biochem Physiol A Mol Integr Physiol* 2001; 129 (1): 295–303.
34. Hallman M, Merritt TA, Akino T, Bry K. Surfactant protein A, phosphatidylcholine, and surfactant inhibitors in epithelial lining fluid. *Am Rev Respir Dis* 1991; 144: 1376–1384.
35. Pryhuber GS, Hull WM, Fink I, McMahan MJ, Whittsett JA. Ontogeny of surfactant proteins A and B in human amniotic fluid as indices of fetal lung maturity. *Pediatr Res* 1991; 30 (6): 597–605.
36. Karinch AM, Floros J. 5' splicing and allelic variants of the human pulmonary *surfactant protein A* genes. *Am J Respir Cell Mol Biol* 1995; 12 (1): 77–88.
37. Oh MH, Wang G, DiAngelo S, Floros J. Differential response of surfactant protein A genetic variants to dexamethasone treatment. *Pediatr Res* 2001; 49: A1708.
38. Hoover RR, Floros J. SP-A 3'UTR is involved in the glucocorticoid inhibition of human *SP-A* gene expression. *Am J Physiol* 1999; 276 (20): L917–L924.
39. Doyle IR, Jones ME, Barr HA et al. Composition of human pulmonary surfactant varies with exercise and level of fitness. *Am J Respir Crit Care Med* 1994; 149: 1619–1627.
40. Honda Y, Takahashi H, Kuroki Y, Akino T, Abe S. Decreased contents of surfactant proteins A and D in BAL fluids of healthy smokers. *Chest* 1996; 109: 1006–1009.
41. Takahashi H, Honda Y, Kuroki Y, Imai K, Abe S. Pulmonary surfactant protein A: a serum marker of pulmonary fibrosis in patients with collagen vascular diseases. *Clin Chim Acta* 1995; 239: 213–215.
42. Stevens PA, Schadow B, Bartholain S, Segerer H, Obladen M. Surfactant protein A in the course of respiratory distress syndrome. *Eur J Pediatr* 1992; 151 (8): 596–600.
43. Frankel WN, Schork NJ. Who's afraid of epistasis? *Nat Genet* 1996; 14: 371–373.
44. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994; 265: 2037–2048.