RACIAL DIFFERENCES IN ALLELIC DISTRIBUTION AT THE HUMAN PULMONARY SURFACTANT PROTEIN B GENE LOCUS (SP-B)

Stavroula V. Veletza □ Department of Cellular and Molecular Physiology, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA

Peter K. Rogan □ Department of Pediatrics, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA

Tom TenHave □ Department of Biostatistics, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA

Samuel A. Olowe □ Department of Pediatrics, University of Lagos, Lagos, Nigeria

Joanna Floros □ Department of Cellular and Molecular Physiology, College of Medicine, The Pennsylvania State University, Hershey, Pennsylvania, USA

□ *Variable numbers of composite repetitive motifs are found in different individuals within intron 4 of the surfactant protein B (SP-B) gene (Biochem J. 1995;305:583). This study tests the hypothesis that the distribution of SP-B alleles differs among racial/ethnic groups. A total of 412 SP-B alleles were analyzed: 206 from Caucasian, 68 from African-American, and 138 from Nigerian individuals. Twelve groups of alleles (A–L) carrying 3 to 18 motifs were found. The distribution of the 12 alleles in the Caucasian group differs from that found in the Nigerian ($p < .001$) and African-American ($p < .001$) populations. The overall distribution of allels between the African-American and the Nigerian populations were not statistically different. Specific alleles were also present in different proportions among the groups studied. For example, the most common allele (allele E) in all three populations is present at a significantly higher frequency in Caucasians than in the other two populations, but its frequency does not differ from the Nigerian and African-American groups. A less frequent allele, H, also differs significantly when Caucasians are compared with each of the other two populations, but the frequency of this allele is comparable between the African-American and Nigerian populations. To assess the importance of having comparable racial composition between the control and the case groups, a group of African-Americans with respiratory distress syndrome (RDS) ($n = 40$) was compared with the African American and the Caucasian groups studied above. No significant*

Received 27 September 1995; accepted 29 February 1996.

We thank Susan DiAngelo for excellent technical assistance. This work was supported by National Institutes of Health grant NIH HL34788.

Present address for Peter K. Rogan is Allegheny-Singer Research Institute, Pittsburgh, PA, USA.

Address correspondence to Joanna Floros, PhD, Department of Cellular and Molecular Physiology, The Pennsylvania State University, College of Medicine, 500 University Drive, Hershey, PA 17033, USA.

0190-2148/96 $12.00 + .00

difference was observed between the racially matched groups but a significant difference (p = .006) was observed between the racially mixed groups. The results indicate that the distribution of SP-B alleles differs between the racial groups but not between the ethnic groups studied. Thus, racial composition of the groups under study is important when considering whether particular alleles at this locus predispose to inherited disorders.

Keywords *alleles, genetic, polymorphisms, pulmonary disease, racial differences, SP-B, surfactant*

Pulmonary surfactant is essential for normal lung function in mammals and consists of lipids and proteins (SP-A, B, C, and D). The surfactant proteins play important roles in surfactant biology/physiology including surfactant function [1]. The importance of SP-B for normal lung function has been affirmed by in vivo studies. When neonatal rabbits were treated with a monoclonal antibody to SP-B, they exhibited reduced compliance, inflammation, and hyaline membranes [2] as well as reduced tidal volumes [3]. In addition, surfactant replacement to preterm rabbits with reconstituted SP-B containing surfactant restored function to levels similar to those obtained with natural sheep surfactant [4]. Absence of SP-B is not compatible with life, as has been shown in patients with congential alveolar proteinosis [5, 6]. The major mutation observed in most cases of congenital alveolar proteinosis is codon 12lins2, but other mutations within the SP-B gene have also been identified [7], suggesting possible allelic heterogeneity.

Varying levels of genetic polymorphism have been observed for the human SP-A, B, and C genes: SP-A is very polymorphic [8–12], whereas SP-C is conserved [13] among humans. We have previously reported a polymorphism in intron 4 of the human SP-B gene [14] that is associated with respiratory distress syndrome (RDS) or with prematurity, which is one of the confounding factors for RDS. This polymorphism involves the gain or loss of a motif that consists of a 20-bp conserved sequence adjacent to a variable length of CA dinucleotides. A number of SP-B alleles, at the intron 4 polymorphic site, with varied copy number of these motifs were previously identified [14]. In the present study, we compare the allele distribution of this polymorphism in normal Caucasian, African-American, and Nigerian populations, as well as in African-Americans affected with RDS. Our goal is to determine whether the distribution of these SP-B alleles differs in racial/ethnic groups and to assess the importance of having a comparable composition between control and case groups as a prelude to studies designed to investigate the contribution of this locus to the etiology of pulmonary disease.

MATERIALS AND METHODS

Specimens and PCR

DNAs were prepared from (1) blood from newborn and adult Caucasian-Americans and African-Americans, (2) adult Nigerians, (3) African-American infants with RDS. DNA preparations and PCR amplification were performed as described previously [14]. The primers used were #172: CTGGTCATCGACTACTTCCA and #161: TGTGTGTGAGAGTGAGGGTGTAAG (positions 2591–2610 and 3174–3197, respectively; GenBank locus HUMSPBAA).

Statistical Analysis

To test the overall differences in the distributions of alleles between populations, we used the two-sided Fisher's Exact Test. Specifically, this test is used to assess population differences with respect to the ratios of frequencies for each pair of alleles. We also used the two-sided Fisher's Exact Test to test for a population difference in a single ratio between the frequencies of particular pairs of alleles. Polymorphism information content (PIC) values and heterozygosity indexes (h) were calculated according to Ott [15].

RESULTS AND DISCUSSION

Characterization of SP-B Alleles

The polymorphism in intron 4 of the human SP-B gene was examined in three populations. We analyzed 412 alleles: 206 from Caucasian, 68 from African-American, and 138 from Nigerian individuals. We found 12 groups of alleles of varying size, ranging from 321–829 bp, carrying 3–18 motifs (Table 1 and Figure 1). These motifs are composed of a series of 20-bp imperfect repeats, followed by a variable length of CA repeats, in a tandem arrangement [14, 16]. The number of motifs present in each of the four alleles (A, B, E, and K) have been determined by sequencing [14]. Alleles of similar sizes were grouped in bins according to the number of motifs present. The expected sizes of alleles other than these four were inferred from the number of motif units present in the PCR product. Each designed allele differs by a single motif unit (Table 1).

The polymorphism information content (PIC) value and heterozygosity index (h) vary for the three populations: for the Nigerian, PIC = 0.319, h = 0.329; for the Caucasian, PIC = 0.151, h = 0.156; for the

Table 1 Description of the 12 groups of alleles observed for the polymorphic site in intron 4 of the human SP-B gene

Allele	Number	Motifs inserted/deleted	Caucasian	African American	Nigerian	African-American RDS
A	3	−8	2	0	0	0
B	6	−5	12[a]	2	1[a]	2
C	9	−2	0	0	0	0
D	10	−1	0[a]	2	4[a]	0
E	11	0	187[a,b]	51[b]	115[a]	30
F	12	+1	1	3	2	0
G	13	+2	1	2	4	1
H	14	+3	0[a,b]	4[b]	6[a]	1
I	15	+4	0	1	0	1
J	16	+5	2	2	2	5
K	17	+6	1	1	3	0
L	18	+7	0	0	1	0
Total			206	68	138	40

Note. The numbers of motifs present in each one of the alleles are shown in the second column. The numbers of motifs considered either deleted or inserted in each allele, in relation to the most prevalent allele E, are shown in the third column. [a]Significant differences in the frequency of alleles between Caucasians and Nigerians; [b]Significant differences in the frequency of alleles between Caucasian and African-Americans.

African-American, PIC = 0.387, $h = 0.404$. These findings indicate that this SP-B variant is more polymorphic in African-Americans and Nigerians than in Caucasians.

Distribution of Alleles in Caucasian, African American, and Nigerian Populations

The distribution of the 12 alleles in the Caucasian population differs from that found in Nigerian ($p < .001$) and African-American ($p < .001$) populations. The distributions of alleles in the African-American and the Nigerian populations are not statistically different ($p = .999$).

Specific alleles are disproportionately represented in these three ethnic groups when compared to the distribution of the most common allele, E, among these ethnic groups. Relative to allele E, allele B is present in the Caucasian population at a higher frequency than in the Nigerian group ($p = 0.04$), whereas the frequency of this allele is similar in the Caucasian and African-American populations. Allele H is absent in this Caucasian cohort population but present in the African-American or the Nigerian groups (Caucasian vs. African-American; $p = .002$; Caucasian vs. Nigerian, $p = .003$). However, the frequency of this allele relative to the frequency of E in the African-American and Nigerian populations

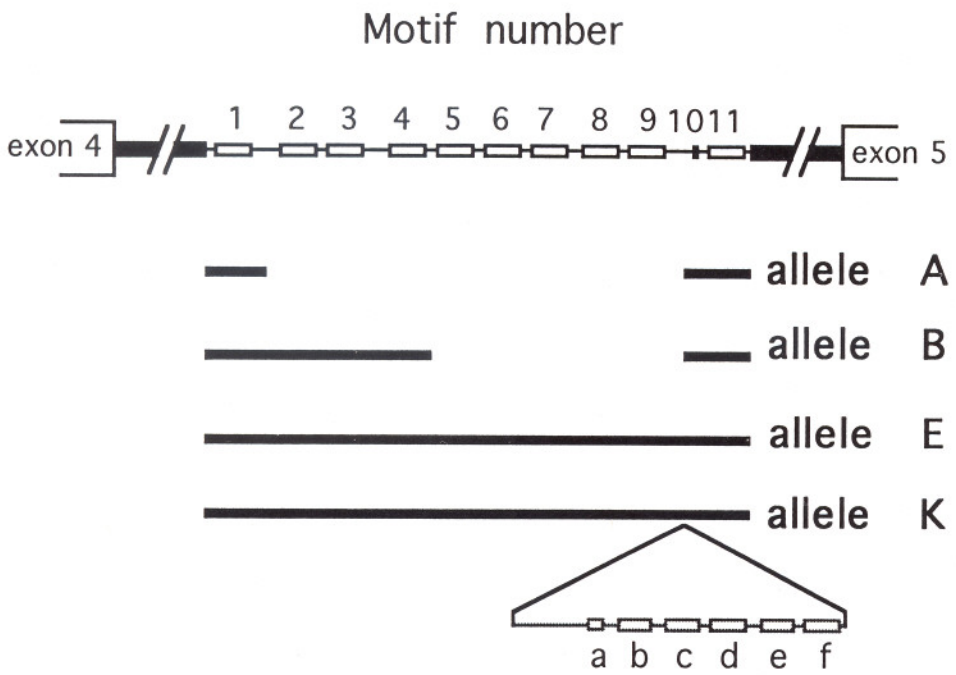

Figure 1. Organization of intron 4. From sequence data [14] allele A contains motifs 1, 10, and 11; allele B contains motifs 1–4, 10, and 11; allele E contains motifs 1–11; and allele K contains motifs 1–9, a–f, 10, and 11. Motifs a–f are inserted degenerate motifs [14]. Sequences in intron 4 adjacent to the repetitive motifs are marked with interrupted thick lines. There are approximately 68 bases in intron 4 before motif 1 and 398 bases after motif 11.

does not differ ($p = .3$). Thus, the distributions of the SP-B alleles differ relative to allele E between the Caucasian and Nigerian, and between the Caucasian and the African-American populations. Genetic differences among these groups have also been observed for the SP-A genotype [8, 9]. These observations together indicate that the overall profile of genetic differences in the surfactant protein genes A and B may be distinguishable in certain racial groups.

We next investigated the importance of having a comparable racial composition between the control and the case group (with RDS). We compared the distribution of SP-B alleles in a small group of African-Americans with RDS ($n = 40$) with either the distribution in African-Americans with no prior history of RDS (studied above), or with the Caucasian group (described above). We found that relative to the distribution of allele E, the distribution of SP-B alleles between the African-American RDS and Caucasian groups differed significantly ($p = .006$), whereas the distribution of SP-B alleles between the African-American RDS and non-RDS groups was not significantly different. This observation

points to the possibility that erroneous conclusions regarding the association of particular alleles with the disease state could be drawn, if the genetic composition of the groups studied is not considered. The presence, however, of race-dependent allelic heterogeneity at the SP-B locus cannot be determined from studies of patients with alveolar proteinosis since there was mention of the race of these patients [7].

In conclusion, the distribution of SP-B alleles at the intron 4 polymorphisms differs between Caucasians and Africans or African-Americans. Therefore, for this SP-B locus the genetic background of case and control groups must be taken into consideration to avoid misleading conclusions regarding the potential role that SP-B plays in the etiology of respiratory disease.

REFERENCES

1. Weaver TE, Whitsett JA. Function and regulation of expression of pulmonary surfactant-associated proteins. Biochem J. 1991;273:249–264.
2. Robertson B, Kobayashi T, Ganzuka M, Grossmann G, Li WZ, Suzuki Y. Experimental neonatal respiratory failure induced by a monoclonal antibody to the hydrophobic surfactant-associated protein SP-B. Pediatr Res. 1991;30:239–243.
3. Kobayashi T, Nitta K, Takahashi R, Kurashima K, Robertson B, Susuki Y. Activity of pulmonary surfactant after blocking the associated proteins SP-A and SP-B. J Appl Physiol. 1991; 71:530–536.
4. Rider ED, Ikegami M, Whitsett JA, Hill Q, Absolom D, Joke AH. Treatment responses to surfactants containing natural surfactant proteins in preterm rabbits. Am Rev Respir Dis. 1993;147:669–676.
5. Nogee LM, deMello DE, Dehner LP, Colten HR. Brief report: Deficiency of pulmonary surfactant protein B in congenital alveolar proteinosis. New Engl J Med. 1993;328:406–410.
6. Nogee LM, Garnier G, Dietz HC, et al. A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. J Clin Invest. 1994;93:1860–1863.
7. deMello DE, Nogee LM, Heyman, S, et al. Molecular and phenotype variability in the congenital alveolar proteinosis syndrome associated with inherited surfactant protein B deficiency. J Pediatr. 1994;125:43–50.
8. Rishi A, Hatzis D, McAlmon K, Floros J. An allelic variant of the 6A gene for human surfactant protein A. Am J Physiol. 1992;262 (Lung Cell. Mol. Physiol. 6):L566–L573.
9. Krizkova L, Sakthivel R, Olowe SA, Rogan P, Floros J. Human SP-A: genotype and single strand conformation polymorphism analysis. Am J Physiol. 1994;266 (Lung Cell. Mol. Physiol. 10): L519–L527.
10. McCormick SM, Boggaram V, Mendelson CR. Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2 genes. Am J Physiol. 266 (Lung Cell. Mol. Physiol. 10): L354–L366.
11. Karinch AM, Floros J. 5′ splicing and allelic variants in the human pulmonary surfactant protein A genes. Am J Respir. Cell Mol Biol. 1995;12:77–88.
12. Floros J, Karinch A. Human SP-A: then and now. Am J Physiol. 1995;268 (Lung Cell. Mol. Physiol. 12):L162–L165.
13. Hatzis D, Deiter G, deMello DE, Floros J. Human surfactant protein-C: genetic homogeneity and expression in RDS; comparison with other species. Exp. Lung Res. 1994;20:57–72.
14. 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.
15. Ott J. Strategies for characrterizing highly polymorphic markers in human gene mapping. Am J Hum Genet. 1992;51:283–290.
16. Pilot–Matias TJ, Kister SE, Fox JL, Kropp K, Glasser SW, Whitsett JA. Structure and organization of the gene encoding human pulmonary surfactant proteolipid SP-B. DNA. 1989;8:75–86.