Published in final edited form as:

Ann Hum Genet. 2011 September ; 75(5): 559–568. doi:10.1111/j.1469-1809.2011.00662.x.

Genetic variants and monoallelic expression of surfactant protein D in inflammatory bowel disease

Zhenwu Lin^{a,†}, Gerrit John^a, John P. Hegarty^a, Arthur Berg^b, Wei Yu^a, Yunhua Wang^{a,c}, Ashley A. Kelly^a, Blaise Z. Peterson^c, Lisa S. Poritz^{a,c}, Joanna Floros^{d,e}, and Walter A. Koltun^{a,†}

^aDepartments of Surgery, Pennsylvania State University College of Medicine, Hershey, PA, USA

^bPublic Health Sciences, Pennsylvania State University College of Medicine, Hershey, PA, USA

^cCellular and Molecular Physiology, Pennsylvania State University College of Medicine, Hershey, PA, USA

^dPediatrics, Pennsylvania State University College of Medicine, Hershey, PA, USA

^eObstetrics and Gynecology, Pennsylvania State University College of Medicine, Hershey, PA, USA

SUMMARY

Surfactant protein D (*SP-D*) is a mucosal collectin that functions in the innate immune response to microorganisms in the pulmonary tract, and possibly the gastrointestinal tract. We studied the genetic association of the two non-synonymous *SP-D* SNPs rs721917 (C/T Met11Thr) and rs2243639 (G/A Ala160Thr) in 256 IBD cases (123 CD and 133 UC) and 376 unrelated healthy individuals from an inflammatory bowel disease (IBD) population from Central Pennsylvania. Case-control analysis revealed a significant association of rs2243639 with susceptibility to Crohn's disease (CD) ($p=0.0036$), but not ulcerative colitis (UC) ($p=0.883$), and no association of rs721917 with CD ($p=0.328$) or UC ($p=0.218$). Using intestinal tissues from 19 individuals heterozygous for each SNP, we compared allele expression of these two SNPs between diseased and matched normal tissues. rs2243639 exhibited balanced biallele expression; while rs721917 exhibited differential allele expression (balanced biallelic 37%, imbalanced biallelic 45%, and dominant monoallelic 18%). Comparison of allelic expression pattern between diseased and matched normal tissues, 13 out 19 individuals (14 UC, 5 CD) showed a similar pattern. The six patients exhibiting a different pattern were all UC patients. The results suggest that differential allelic expression may affect penetrance of the SNP rs721917 disease-susceptibility allele in IBD. The potential impact of *SP-D* monoallelic expression on incomplete penetrance is discussed.

Keywords

surfactant protein-D; inflammatory bowel disease; allele specific expression; genetic association

[†]Corresponding authors: zlin@hmc.psu.edu OR wkoltun@hmc.psu.edu Department of Surgery, Division of Colon & Rectal Surgery Pennsylvania State University College of Medicine H137, 500 University Drive, Hershey, PA 17033, USA Tel.: +1-717-531-5164 Fax: +1-717-531-0646.

Conflict of interest statement The authors declare no conflict of interest.

INTRODUCTION

Surfactant protein-D (*SP-D*) belongs to the collagen-containing C-type (calcium dependent) lectins called collectins. *SP-D* is mainly produced by alveolar type II cells in the lung, where it is involved in pulmonary immunity and initiates a wide range of defense mechanisms against microorganisms, including direct opsonization, neutralization, agglutination, complement activation, and enhanced phagocytosis (Kishore *et al.*, 2006). *SP-D* has further been implicated in clearance of apoptotic and necrotic cells (Clark *et al.*, 2002) and has direct bactericidal effects (Wu *et al.*, 2003). Other functions include the control of pulmonary inflammation (Wright, 2005), T_H-cell polarization (Madan *et al.*, 2001), and suppression of lipid peroxidation (Bridges *et al.*, 2000). Thus, *SP-D* plays an important role in linking adaptive and innate immune cell functions in the first line of the host defense and is thereby important in human health and disease (Kishore *et al.*, 2005). However, *SP-D* is not restricted to the lung, but is instead widely distributed on mucosal surfaces of various tissues (Madsen *et al.*, 2000a), including the reproductive and gastrointestinal tract, and is most likely involved in extra-pulmonary host defenses. In the porcine intestine, *SP-D* has been localized to epithelial cells of the intestinal glands (crypts of Lieberkuhn) in the duodenum, jejunum and ileum (Soerensen *et al.*, 2005), where it exerts a scavenger function by promoting uptake of pathogenic bacteria by these cells (Hogenkamp *et al.*, 2007).

The pathogenesis of inflammatory bowel disease (IBD), consisting of two main subtypes, Crohn's disease (CD) and ulcerative colitis (UC), is believed to be a result of both genetic predisposition and environmental influences (Russell *et al.*, 2004, Schreiber *et al.*, 2005, Elson *et al.*, 2005). Currently over 50 IBD-associated genes/loci have been identified (Lees & Satsangi, 2009). In addition, *SP-D* has recently been associated with IBD in a Japanese IBD population. Of the five known SNPs of the human *SP-D* gene, the 2-allele haplotype GG of an intronic SNP rs911887 and the nonsynonymous SNP rs2243639 reached statistical significance for susceptibility to both IBD and UC (Tanaka *et al.*, 2009). Subset analysis in association with early onset UC further suggested that *SP-D* could be a disease-modifier gene as well as a disease-susceptibility gene. Interestingly, random monoallelic and/or heterogeneous allele expression occurring in genes involved in immunity (Ohlsson *et al.*, 1998) has also been described for *SP-D* in the large intestine and other tissues in the rat (Lin & Floros, 2002). Given the role of *SP-D* in innate immunity and the observation that variable or heterogeneous allele expression occurs, we further examined the genetic association of *SP-D* nonsynonymous variants with IBD in a population of central Pennsylvania, and studied the effect of allele expression on incomplete penetrance in disease phenotype. The two genetic variants analyzed were rs721917 (C/T Met11Thr) in exon 2 and rs2243639 (G/A Ala160Thr) in exon 5. The study population consisted of 256 IBD patients and 376 unrelated healthy controls from Central Pennsylvania.

MATERIALS AND METHODS

Study samples

IBD patients—A total of 256 IBD cases were studied, including: 1) 131 individuals with IBD (80 CD patients, 51 UC patients) from 72 families with familial IBD history and 2) 125 sporadic IBD patients (43 CD patients, 82 UC patients) from the Milton S. Hershey Medical Center from Central Pennsylvania. Samples were obtained from our IBD familial registry established in 1999. Peripheral blood was collected from study participants and used to derive B cell lines by Epstein Barr virus (EBV) transformation. Intestinal tissues were obtained at the time of surgery. IBD diagnosis was made using standard clinical, radiological, and endoscopic/histopathological procedures.

Controls—Peripheral blood samples from 376 unrelated healthy individuals from the Milton S. Hershey Medical Center were used as controls.

Ethical considerations

All human tissues described above were approved by the Human Subjects Protection Offices of The Pennsylvania State University College of Medicine, and were undertaken with the understanding and written consent of each subject.

Genomic DNA and RNA isolation

Genomic DNA from B cell lines and peripheral blood was isolated using a QIAamp DNA blood kit (Qiagen Inc. Valencia, CA), and DNA from intestinal tissues was isolated with a QIAamp DNA Mini Kit according to the manufacturer's instruction. DNA concentrations were measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technology, Wilmington, DE).

Total RNA was extracted from 38 diseased and non-diseased adjacent intestinal tissues from IBD patients using the RNeasy Mini Kit according to the manufacture's instruction (Qiagen Inc.). cDNA was synthesized from 1 µg of total RNA using the Superscript III First Strand Synthesis Kit (Invitrogen, Carlsbad, CA).

Reverse transcription (RT)-PCR

cDNA was used as template to amplify a 484 bp fragment of *SP-D* using primers SPD18s (5'-CTCCAGGCTGCTTTCTCTCAG-3') and SPD26r (5'-TGGCAGCATGAGGGTCTAAG-3'), as well as β-actin (266 bp fragment) with primers Actin7s (5'-TGTGGATCAGCAAGCAGGAG-3') and Actin8r (5'-GTGAACCTTTGGGGGATGCTC-3'). PCR reactions were prepared as previously described (Lin et al, 2009), and amplifications were performed with the following PCR profile: 95 °C 2 min, 35 cycles of 95 °C 30 sec, 58 °C 1 min, and 72 °C 1 min, and then 72 °C 4 min.

Northern analysis

Total mRNA (3.5 µg) was separated on 1% agarose gel containing formaldehyde (0.22 M) and transferred to positively charged nylon transfer membrane (GeneScreen Plus, Perkin Elmer Life Science, Boston, MA) and fixed under a UV Stratalinker 2400 (Stratagene, La Jolla, CA). The procedure for Northern analysis was described before (Lin *et al.*, 2001). Briefly, the blot was hybridized with ³²P-ATP-labeled *SP-D* antisense oligonucleotide probe. The probe was prepared with the KinaseMax kit (Ambion, Austin, TX) and hybridization was performed in Ultrahyb solution according to the manufacturer's instructions (Ambion, Austin, TX). The blot was then exposed to Kodak XAR5 film at −80°C for 5 min for lung RNA or 24 hours for intestine RNA.

Analysis of genotypes and allele expression with PCR-based RFLP

Genotypes and allele expression of the two genetic variants located in exon regions of the human *SP-D* gene, rs721917 (C/T Met11Thr) in exon 2 and rs2243639 (G/A Ala160Thr) in exon 5 (Tanaka *et al.*, 2009), were determined using PCR-based restriction fragment length polymorphism (RFLP) and converted RFLP methods (Lin et al *et al.*, 1996).

For genotype analysis, an 11 kb DNA fragment of the *SP-D* gene was amplified with primers SPD30s and SPD26r (SPD30s 5'-TCACCTCTAGAAGCTGAGCCAAGCC-3', SPD26r 5'-TGGCAGCATGAGGGTCTAAG-3') using Elongase Enzyme Mix (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions with the following PCR profile: 95 °C 2 min, 10 cycles of 95 °C 30 sec, 58 °C 1 min and 68 °C 10 min, followed by 20

cycles of 95 °C 30 sec, 62 °C 1 min and 68 °C 12 min, and then 68 °C 20 min. One µl of the PCR products was used as template to further amplify a 102 bp DNA fragment containing rs721917 with primers SPD31s (5'-CTCCTCTCTGCACTGGTCAT-3') and SPD20r (5'-ACCAGGGTGCAAGCACTGCG-3'), and a 164 bp DNA fragment containing rs2243639 with primers SPD34s (5'-AGCGTGGAGTCCCTGGAAGC-3') and SPD35r (5'-AGATTCTCTCCATGTTCCCAG-3'). The PCR profile consisted of 95 °C 2 min, 5 cycles of 95 °C 30 sec, 50 °C 1 min and 72 °C 1 min followed by 30 cycles of 95 °C 30 sec, 58 °C 1 min and 72 °C 1 min, and a final extension for 4 min at 72 °C.

For allele expression, 1 µl cDNA was used as template to amplify a 102 bp fragment containing rs721917 with primers SPD31s and SPD20r, and a 118 bp fragment containing rs2243639 with primers SPD34s and SPD23r (5'-TCTCCAGGAATGCCTTTG-3'). All PCR reactions were prepared as previously described (Lin et al, 2009).

During PCR amplification, SNPs rs721917 and rs2243639 were converted to Fsp I and Hha I restriction enzyme recognition sites, respectively. PCR products were digested with Fsp I (for rs721917) and Hha I (for rs2243639) according to the manufacturer's instructions (New England Biolabs, Ipswich, MA). The digested PCR products were separated on 8% PAGE and visualized by ethidium bromide staining. Genotypes were scored according to the gel pattern of allelic products following digestion (Figure 1).

Quantification of allelic products and classification of the allele-specific expression pattern

Allele expression was quantified by densitometry using ImageJ software (NIH). The ratio of the digested PCR products of allele T to allele C for rs721917 and of allele G to allele A for rs2243639 in all heterozygous individuals was used to classify the allele-specific expression pattern in each individual (Lin & Floros, 2002). If the difference between the two allelic PCR products was less than 20% (the ratio = 0.8-1.2), the allele expression was classified as balanced biallelic (BB), whereas a difference of more than 20% but within 5-fold (the ratio = 0.2 - <0.8 or 1.2 - ≤5.0) was classified as imbalanced biallelic (IB) expression. If the difference was more than 5-fold (the ratio = <0.2 or >5.0), the allele expression was classified as dominant monoallelic (DM).

Statistical analysis of genetic association

For the genetic association study, Pearson's χ^2 -test with one degree of freedom for allelic association was performed using Haploview (MIT/Harvard Broad Institute). We also calculated genotype-based Odds Ratio (OR) using Fisher's exact test on the contingency tables, and the effect of the association is measured by the resulting p-values. A genotype difference was considered significant when $p < 0.05$. The two SNPs were analyzed for linkage disequilibrium and found to be significantly correlated ($p < 0.001$, $D' = 0.944$, $\text{Corr} = 0.611$). As only two SNPs in linkage disequilibrium were considered in this study, the reported p-values are not adjusted for multiple comparisons.

RESULTS

Genetic association of *SP-D* rs2243639 (Ala160Thr) with IBD

For genotype analysis of *SP-D* variants rs721917 and rs2243639, 257 IBD cases including 132 familial IBD (81 CD patients, 51 UC patients) and 125 sporadic IBD patients (43 CD patients, 82 UC patients) and 376 unrelated healthy controls were studied using PCR-based RFLP and converted RFLP methods. After digestion with Fsp I (for rs721917) and Hha I (for rs2243639), PCR products were separated on 8% PAGE and visualized by ethidium bromide staining. Genotypes were scored according to the gel pattern of allelic products

after digestion: for rs721917 C allele is 102 bp, T allele is 82 and 20 bp; for rs2243639 A allele is 164 bp, G allele is 144 and 20 bp. The 20 bp fragments from T allele of rs721917 and G allele of rs2243639 in the restriction enzyme digested PCR products ran off of the gels. Figure 1 shows exemplary results for genotypes at rs721917 (Fig. 1A) and rs2243639 (Fig. 1B) using genomic DNA from 16 different and randomly chosen IBD cases.

For rs721917, the frequency of genotypes in the control samples was 13% CC, 53% CT, and 34% TT calculated from Table 1. The IBD case group exhibited a comparable genotype frequency of 15% CC, 50% CT, and 35% TT. The distribution of both alleles is very similar between cases and controls. For rs2243639, the frequency of genotypes in the control samples was 39% GG, 45% GA, and 16% AA. The IBD case group exhibited a genotype frequency of 28% GG, 55% GA, and 17% AA. The distribution of both alleles is different between cases and controls. In the CD subgroup, genotypes were presented as 23.6% GG, 53.6% GA, and 22.8% AA. The UC subgroup exhibited a genotype frequency of 32.3% GG, 56.4% GA, and 11.3% AA.

To further analyze a possible genetic association of these genetic variants with IBD, CD and UC, genotyping data were subjected to statistical analyses. The results indicate that rs2243639 was significantly associated with CD (OR=1.5531, 95% CI= 1.14992-2.0983, $p=0.0036$), but not with UC (OR=1.0300, 95% CI=0.7645-1.3845, $p=0.8836$). For the association with IBD the p value is 0.054. Although very close, it failed to reach the significance level ($p<0.05$). On the other hand, genetic variation at rs721917 did not reach statistical significance for susceptibility to either IBD or subgroups CD and UC (Table 1).

***SP-D* mRNA expression in the human intestine**

SP-D has been demonstrated to be expressed in rat colon and small intestine by Northern and RT-PCR analysis and in human ileum and colon by RT-PCR and immunohistochemistry analysis (Madsen *et al.*, 2000b, Bourbon & Chailley-Heu, 2001, van Rozendaal *et al.*, 2001). We performed Northern and RT-PCR analyses for the intestinal tissues from IBD patients. The result from Northern analysis showed that *SP-D* gene is expressed in both small intestine and colon. Compared to the human lung as shown in Fig 2A, the intestine *SP-D* mRNA band on Northern blot with 24 hours of exposure was lower than that in the lung with only 5 min exposure at -80°C , suggesting that the level of *SP-D* transcripts in intestinal tissues is much lower. The size of *SP-D* mRNA from intestinal tissues was the same as that in the lung. The results from RT-PCR also confirmed that *SP-D* is expressed in both small intestine and colon, and no difference was observed in the size of PCR product between *SP-D* mRNA from intestine and lung (Fig 2B).

Allele expression patterns of *SP-D* rs2243639 and rs721917 in human intestinal tissues

We used a PCR-based RFLP method to distinguish the PCR products amplified from different alleles. First we amplified cDNA fragments with *SP-D* SNPs rs721917 and rs2243639, and then we used the same restriction enzymes to digest the PCR products. The digested PCR products were separated by 8% PAGE. The allelic PCR products were quantitatively analyzed. For homozygous individuals, only one allelic product was shown; however, for the heterozygous individuals two allelic products should be shown. If the two alleles are not equally transcribed, the patterns of allelic product will vary. We quantified each allelic product and used the ratio of allelic products to classify the allele expression as: balanced biallelic (BB, 0.8-1.2), imbalanced biallelic (IB, 0.2-<0.8 and 1.2-<5.0) or dominant monoallelic (DM, <0.2 and >5.0). The results from 38 intestinal tissues revealed that *SP-D* variant rs721917 showed a differential allele expression (Fig. 3A and Table 2) with 14 BB samples, 17 IB samples, and 7 DM samples. For rs2243639, only balanced biallelic expression was observed in all the intestinal tissues examined (Fig. 3B).

The 38 intestinal tissues studied were 19 pairs of diseased and matched non-diseased tissues from the same organ, such as diseased colon/matched non-diseased colon, diseased terminal ileum/matched non-diseased terminal ileum. Thirteen individuals exhibited the same allelic expression pattern between diseased and normal tissues, and only 6 individuals exhibited different allelic expression patterns between diseased and normal tissues (Table 2). In these 6 individuals, 5 were from BB to IB, or from IB to DM, and only one exhibited a large difference from BB to DM. We have previously studied *SP-D* allelic expression in 8 rat pedigrees with 3-generations. We observed that the parents from a family with high frequency of *SP-D* monoallelic expression their family members exhibited high frequency of monoallelic expression, suggesting that inheritable factors may regulate allelic expression (Lin & Floros, 2002). The present results from human *SP-D* further support the notion that genetic factors play a role in allelic expression traits. It is very interesting to point out that all 6 individuals with a different allelic expression pattern were UC patients. Calculated from Table 2, 43% (6 out of 14) of UC patients exhibited *SP-D* allele differential expression pattern between disease and normal tissues, while all 6 CD patients exhibited the same pattern.

DISCUSSION

SP-D plays an important function in host defense and regulation of inflammation. The SNP rs721917 has been shown to associate with risk at tuberculosis (Floros *et al.*, 2000). Although genetic variants of *SP-D* alone failed to show a significant association of with lung diseases, such as ARDS (acute respiratory distress syndrome), COPD (chronic obstructive pulmonary disease), and IPF (idiopathic pulmonary fibrosis) (Guo *et al.*, 2001, Lin *et al.*, 2000, Selman *et al.*, 2003), haplotypes of *SP-D* with other surfactant proteins were associated with BPD (bronchopulmonary dysplasia) and ARDS (acute RDS) (Pavlovic *et al.*, 2006, Thomas *et al.*, 2007). While human *SP-D* has been intensively studied in lung diseases (Floros & Hoover, 1998, Floros & Kala, 1998, Lin *et al.*, 2000), it is also believed to play a role in extra-pulmonary systems (Madsen *et al.*, 2000b). Recently the genetic association of *SP-D* with IBD was studied in a Japanese population. Among 5 *SP-D* SNPs (2 exonic, 3 intronic) analyzed, the minor allele G of the intronic rs911887 was associated with UC, whereas both nonsynonymous SNPs, rs721917 in exon 2 and rs2243639 in exon 5 were not associated with IBD. However a 2-allele haplotype GG of the intronic rs911887 and exonic rs2243639 were associated with UC (Tanaka *et al.*, 2009).

In the present study, we investigated two exonic nonsynonymous SNPs of *SP-D* in 257 IBD cases (124 CD and 133 UC) and 376 unrelated healthy individuals in an IBD population from Central Pennsylvania. Different from the previous published study, we found that the exonic SNP rs2243639 is significantly associated IBD alone, and also the association is with IBD subtype CD ($p=0.0036$), and not UC. Our results are consistent with the published report in the sense that *SP-D* has been associated with IBD. However, we found that rs2243639 alone was associated with IBD, specifically CD; whereas their IBD association was with a 2-allele haplotype comprising rs2243639 with another minor allele (rs911887), specifically for UC. In different IBD populations differences in genetic background and environmental factors may collectively contribute to differential IBD susceptibilities and/or other complexities discussed below. Differences in the association of genetic variations with IBD in different populations have been observed for several IBD susceptibility genes. Some risk alleles are absent or fixed in a population, which implies that risk alleles identified in one population do not necessarily account for disease prevalence in all human populations (Myles *et al.*, 2008). Indeed, risk loci for one population may not be found for a different population (Nakagome *et al.*, Morgan *et al.*). Therefore disease-association studies in diverse populations are required in order to determine whether different alleles are responsible for disease prevalence in different populations. Recently, a correlation between

SP-D SNPs and SP-D serum protein concentrations has been demonstrated (Foreman MG *et al.* 2011). The reason that *SP-D* nonsynonymous SNP rs721917 shows no association with IBD may be that the amino acid change from Met>Thr at that position has no effect on protein function, or that there is incomplete penetrance of genetic information.

The penetrance of a disease-causing mutation is the proportion of individuals with the mutation exhibiting clinical symptoms. If the proportion is less than 100%, it is defined as an incompletely penetrant mutation. Incomplete penetrance likely results from genetic, epigenetic, and environmental factors, all which affect gene function from transcription to protein activation. Figure 4 illustrates the effect of allelic expression on penetrance of a disease allele. As a result of variation in allele expression, the final gene expression products (mRNA and protein) of a heterozygous individual (TC) may have be an equal amount of product from wild type allele T and mutated allele C; only the product of wild type allele T or mutated allele C; or only the dominant product of wild type T or mutated allele C. The complete loss of heterozygosity or quantitative changes in expression of a mutated allele may lead to a different disease phenotype outcome. From the point of view of gene function, a heterozygous individual will be the same as a homozygous wild type allele individual, or a homozygous mutated allele individual. Monoallelic expression is widely observed in human and other diploid eukaryotic organisms (Zakharova *et al.*, 2009, Chess *et al.*, 1994, Rajewsky, 1996). The genes exhibiting monoallelic expression include imprinted genes (Reik & Walter, 2001), X-inactivated genes (Lyon, 1986), and some autosomal genes (Chess *et al.*, 1994). Using a genome-wide assay, it was assessed that more than 5% of genes were subjected to random monoallelic expression (Gimelbrant *et al.*, 2007). Monoallelic expression can lead to differences in levels of gene expression. These include loss of imprinting in imprinted genes and loss of heterozygosity in autosomal genes. A role of monoallelic expression has been observed in disease susceptibility, such as for imprinted genes IGF-2 and MEST in lung adenocarcinoma (Kohda *et al.*, 2001), IGF-2 in Wilms' tumor (Ogawa *et al.*, 1993), and GABRA5, GABRB3, and WI-14946 in Angelman syndrome (Bittel *et al.*, 2005); imprinted and autosomal genes in cancers (Das *et al.*, 2009, Morison *et al.*, 2005, Barrois *et al.*, 2001); and APOE4 in Alzheimer disease (Xu *et al.*, 2007). Recently, a decrease in allele-specific expression of the TGFBR1 gene was shown to predispose to colorectal cancer (Tomsic J *et al.* 2010).

We examined allele expression analyses of the two *SP-D* SNPs in intestinal tissues and observed that rs2243639 exhibited balanced biallelic expression; while rs721917 exhibited a differential allele expression (balanced biallelic 37%, imbalanced biallelic 45%, and dominant monoallelic 18%). Interestingly, from our comparison of the allelic expression pattern between diseased and matched normal tissues, 13 out 19 individuals (14 UC, 5 CD) showed a similar expression pattern, while the other 6 patients exhibiting a different pattern were all UC patients. These results are consistent with our previous finding that the family background is involved in the regulation of *SP-D* allele expression in a rat pedigree, thereby suggesting that a genetic factor(s) may also be involved in the regulation of *SP-D* allele expression. Our data suggest that heterogeneous allele expression may be a factor affecting incomplete penetrance of mutations which thus may cause disease, and may help in understanding why *SP-D* rs2243639 is associated, while rs721917 is not associated with CD. A recent paper (Arai *et al.* 2011) showed allelic expression imbalance in intestinal tissue (n=10) of the risk haplotype of the IBD-associated gene, Nkx2-3. It is thought that the expression imbalance of the risk haplotype may confer susceptibility to UC in the colonic mucosa. We conclude that *SP-D* rs2243639 susceptibility to Crohn's disease (CD) (p=0.0036), but not ulcerative colitis (UC) (p=0.883), and no association of rs721917 with CD (p=0.328) or UC (p=0.218). The present observations need to be confirmed with a large number of samples, and the mechanism of *SP-D* monoallelic expression on disease outcome should be investigated.

Acknowledgments

The authors thank Tammy Lin for her aid in genotype analysis.

Support This work is supported by a grant from the Philadelphia Health Care Trust (WAK), a feasibility research grant from Department of Surgery (ZL), a research grant from American Lung Association (ZL), and NIH HL34788 (JF).

REFERENCES

- Arai T, Kakuta Y, Kinouchi Y, Kimura T, Negoro K, Aihara H, Endo K, Shiga H, Kanazawa Y, Kuroha M, Moroi R, Nagasawa H, Shimodaira Y, Takahashi S, Shimosegawa T. Increased expression of NKX2.3 mRNA transcribed from the risk haplotype for ulcerative colitis in the involved colonic mucosa. *Hum Immuno*. 2011 Epub Apr 8.
- Barrois M, Eychenne MK, Terrier-Lacombe MJ, Duarte N, Dubourg C, Douc-Rasy S, Chompret A, Khagad M, Hartmann O, Caput D, Benard J. Genomic and allelic expression status of the p73 gene in human neuroblastoma. *Med Pediatr Oncol*. 2001; 36:45–7. [PubMed: 11464903]
- Bittel DC, Kibiryeva N, Talebizadeh Z, Driscoll DJ, Butler MG. Microarray analysis of gene/transcript expression in Angelman syndrome: deletion versus UPD. *Genomics*. 2005; 85:85–91. [PubMed: 15607424]
- Bourbon JR, Chailley-Heu B. Surfactant proteins in the digestive tract, mesentery, and other organs: evolutionary significance. *Comp Biochem Physiol A Mol Integr Physiol*. 2001; 129:151–61.
- Bridges JP, Davis HW, Damodarasamy M, Kuroki Y, Howles G, Hui DY, McCormack FX. Pulmonary surfactant proteins A and D are potent endogenous inhibitors of lipid peroxidation and oxidative cellular injury. *J Biol Chem*. 2000; 275:38848–55. [PubMed: 10969075]
- Chess A, Simon I, Cedar H, Axel R. Allelic inactivation regulates olfactory receptor gene expression. *Cell*. 1994; 78:823–34. [PubMed: 8087849]
- Clark H, Palaniyar N, Strong P, Edmondson J, Hawgood S, Reid KB. Surfactant protein D reduces alveolar macrophage apoptosis in vivo. *J Immunol*. 2002; 169:2892–9. [PubMed: 12218102]
- Das R, Hampton DD, Jirtle RL. Imprinting evolution and human health. *Mamm Genome*. 2009; 20:563–72. [PubMed: 19830403]
- Elson CO, Cong Y, McCracken VJ, Dimmitt RA, Lorenz RG, Weaver CT. Experimental models of inflammatory bowel disease reveal innate, adaptive, and regulatory mechanisms of host dialogue with the microbiota. *Immunol Rev*. 2005; 206:260–76. [PubMed: 16048554]
- Floros J, Hoover RR. Genetics of the hydrophilic surfactant proteins A and D. *Biochim Biophys Acta*. 1998; 1408:312–22. [PubMed: 9813381]
- Floros J, Kala P. Surfactant proteins: molecular genetics of neonatal pulmonary diseases. *Annu Rev Physiol*. 1998; 60:365–84. [PubMed: 9558469]
- Floros J, Lin HM, Garcia A, Salazar MA, Guo X, Diangelo S, Montano M, Luo J, Pardo A, Selman M. Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis*. 2000; 182:1473–8. [PubMed: 11023470]
- Foreman MG, Kong X, Demeo DL, Pillai SG, Hersh CP, Bakke P, Gulsvik A, Lomas DA, Litonjua AA, Shapiro SD, Tal-Singer R, Silverman EK. Polymorphisms in surfactant protein-d are associated with chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol*. Mar; 2011 44(3):316–22. [PubMed: 20448057]
- Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. Widespread monoallelic expression on human autosomes. *Science*. 2007; 318:1136–40. [PubMed: 18006746]
- Guo X, Lin HM, Lin Z, Montano M, Sansores R, Wang G, Diangelo S, Pardo A, Selman M, Floros J. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J*. 2001; 18:482–90. [PubMed: 11589345]
- Hogenkamp A, Herias MV, Tooten PC, Veldhuizen EJ, Haagsman HP. Effects of surfactant protein D on growth, adhesion and epithelial invasion of intestinal Gram-negative bacteria. *Mol Immunol*. 2007; 44:3517–27. [PubMed: 17477970]

- Kishore U, Bernal AL, Kamran MF, Saxena S, Singh M, Sarma PU, Madan T, Chakraborty T. Surfactant proteins SP-A and *SP-D* in human health and disease. *Arch Immunol Ther Exp (Warsz)*. 2005; 53:399–417. [PubMed: 16314824]
- Kishore U, Greenhough TJ, Waters P, Shrive AK, Ghai R, Kamran MF, Bernal AL, Reid KB, Madan T, Chakraborty T. Surfactant proteins SP-A and *SP-D*: structure, function and receptors. *Mol Immunol*. 2006; 43:1293–315. [PubMed: 16213021]
- Kohda M, Hoshiya H, Katoh M, Tanaka I, Masuda R, Takemura T, Fujiwara M, Oshimura M. Frequent loss of imprinting of IGF2 and MEST in lung adenocarcinoma. *Mol Carcinog*. 2001; 31:184–91. [PubMed: 11536368]
- Lees CW, Satsangi J. Genetics of inflammatory bowel disease: implications for disease pathogenesis and natural history. *Expert Rev Gastroenterol Hepatol*. 2009; 3:513–34. [PubMed: 19817673]
- Lin Z, Cui X, Li H. Multiplex genotype determination at a large number of gene loci. *Proc Natl Acad Sci U S A*. 1996; 93:2582–7. [PubMed: 8637917]
- Lin Z, Demello D, Phelps DS, Koltun WA, Page M, Floros J. Both human SP-A1 and Sp-A2 genes are expressed in small and large intestine. *Pediatr Pathol Mol Med*. 2001; 20:367–86. [PubMed: 11552738]
- Lin Z, Floros J. Heterogeneous allele expression of pulmonary *SP-D* gene in rat large intestine and other tissues. *Physiol Genomics*. 2002; 11:235–43. [PubMed: 12464693]
- Lin Z, Pearson C, Chinchilli V, Pietschmann SM, Luo J, Pison U, Floros J. Polymorphisms of human SP-A, SP-B, and *SP-D* genes: association of SP-B Thr131Ile with ARDS. *Clin Genet*. 2000; 58:181–91. [PubMed: 11076040]
- Lin Z, Poritz L, Franke a, Li Tong-Yi, Ruether a, Byrnes K a, Wang Y, Gebhard a W, Macneill C, Thomas NJ, Schreiber S, Koltun WA. Genetic association of DLG5 R30Q with familial and sporadic inflammatory bowel disease in men. *Disease Markers*. 2009; 27(5):193–201. 2009. [PubMed: 20037206]
- Lyon MF. X chromosomes and dosage compensation. *Nature*. 1986; 320:313. [PubMed: 3960115]
- Madan T, Kishore U, Singh M, Strong P, Clark H, Hussain EM, Reid KB, Sarma PU. Surfactant proteins A and D protect mice against pulmonary hypersensitivity induced by *Aspergillus fumigatus* antigens and allergens. *J Clin Invest*. 2001; 107:467–75. [PubMed: 11181646]
- Madsen J, Kliem A, Tornøe I, Skjodt K, Koch C, Holmskov U. Localization of lung surfactant protein D on mucosal surfaces in human tissues. *J Immunol*. 2000a; 164:5866–70. [PubMed: 10820266]
- Madsen J, Kliem A, Tornøe I, Skjodt K, Koch C, Holmskov U. Localization of lung surfactant protein D on mucosal surfaces in human tissues. *J Immunol*. 2000b; 164:5866–70. [PubMed: 10820266]
- Morgan AR, Han DY, Huebner C, Lam WJ, Fraser AG, Ferguson LR. PTPN2 but not PTPN22 is associated with Crohn's disease in a New Zealand population. *Tissue Antigens*.
- Morison IM, Ramsay JP, Spencer HG. A census of mammalian imprinting. *Trends Genet*. 2005; 21:457–65. [PubMed: 15990197]
- Myles S, Davison D, Barrett J, Stoneking M, Timpson N. Worldwide population differentiation at disease-associated SNPs. *BMC Med Genomics*. 2008; 1:22. [PubMed: 18533027]
- Nakagome S, Takeyama Y, Mano S, Sakisaka S, Matsui T, Kawamura S, Oota H. Population-specific susceptibility to Crohn's disease and ulcerative colitis; dominant and recessive relative risks in the Japanese population. *Ann Hum Genet*. 74:126–36. [PubMed: 20367632]
- Ogawa O, Eccles MR, Szeto J, Mcnøe LA, Yun K, Maw MA, Smith PJ, Reeve AE. Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. *Nature*. 1993; 362:749–51. [PubMed: 8097018]
- Ohlsson R, Tycko B, Sapienza C. Monoallelic expression: 'there can only be one'. *Trends Genet*. 1998; 14:435–8. [PubMed: 9825668]
- Pavlovic J, Papagaroufalos C, Xanthou M, Liu W, Fan R, Thomas NJ, Apostolidou I, Papathoma E, Megaloyianni E, Diangelo S, Floros J. Genetic variants of surfactant proteins A, B, C, and D in bronchopulmonary dysplasia. *Dis Markers*. 2006; 22:277–91. [PubMed: 17264398]
- Rajewsky K. Clonal selection and learning in the antibody system. *Nature*. 1996; 381:751–8. [PubMed: 8657279]
- Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nature reviews*. 2001; 2:21–32.

- Russell RK, Nimmo ER, Satsangi J. Molecular genetics of Crohn's disease. *Curr Opin Genet Dev.* 2004; 14:264–70. [PubMed: 15172669]
- Schreiber S, Rosenstiel P, Albrecht M, Hampe J, Krawczak M. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nature Reviews Genetics.* 2005; 6:376–88.
- Selman M, Lin HM, Montano M, Jenkins AL, Estrada A, Lin Z, Wang G, Diangelo S, Guo X, Umstead TM, Lang CM, Pardo A, Phelps DS, Floros J. Surfactant protein A and B genetic variants predispose to idiopathic pulmonary fibrosis. *Hum Genet.* 2003; 113:542–550. [PubMed: 13680361]
- Soerensen CM, Nielsen OL, Willis A, Heegaard PM, Holmskov U. Purification, characterization and immunolocalization of porcine surfactant protein D. *Immunol.* 2005; 114:72–82.
- Tanaka M, Arimura Y, Goto A, Hosokawa M, Nagaishi K, Yamashita K, Yamamoto H, Sonoda T, Nomura M, Motoya S, Imai K, Shinomura Y. Genetic variants in surfactant, pulmonary-associated protein D (SFTPD) and Japanese susceptibility to ulcerative colitis. *Inflam Bowel Dis.* 2009a; 15:918–25.
- Thomas NJ, Fan R, Diangelo S, Hess JC, Floros J. Haplotypes of the surfactant protein genes A and D as susceptibility factors for the development of respiratory distress syndrome. *Acta Paediatr.* 2007; 96:985–9. [PubMed: 17524024]
- Tomsic J, Guda K, Liyanarachchi S, Hampel H, Natale L, Markowitz SD, Tanner SM, de la Chapelle A. Allele-specific expression of TGFBR1 in colon cancer patients. *Carcinogenesis.* Oct; 2010 31(10):1800–4. [PubMed: 20705955]
- Van Rozendaal BA, Van Golde LM, Haagsman HP. Localization and functions of SP-A and *SP-D* at mucosal surfaces. *Ped Pathol & Mol Med.* 2001; 20:319–39.
- Wright JR. Immunoregulatory functions of surfactant proteins. *Nature reviews.* 2005; 5:58–68.
- Wu H, Kuzmenko A, Wan S, Schaffer L, Weiss A, Fisher JH, Kim KS, McCormack FX. Surfactant proteins A and D inhibit the growth of Gram-negative bacteria by increasing membrane permeability. *J Clin Investig.* 2003; 111:1589–602. [PubMed: 12750409]
- Xu PT, Li YJ, Qin XJ, Kroner C, Green-Odlum A, Xu H, Wang TY, Schmechel DE, Hulette CM, Ervin J, Hauser M, Haines J, Pericak-Vance MA, Gilbert JR. A SAGE study of apolipoprotein E3/3, E3/4 and E4/4 allele-specific gene expression in hippocampus in Alzheimer disease. *Mol Cell Neurosci.* 2007; 36:313–31. [PubMed: 17822919]
- Zakharova IS, Shevchenko AI, Zakian SM. Monoallelic gene expression in mammals. *Chromosoma.* 2009; 118:279–90. [PubMed: 19242715]

Figure 1. Genotype analysis of *SP-D* SNPs rs721917 and rs2243639

PCR primers used for genotyping were SPD31s and SPD20r for rs721917, and SPD34s and SPD35r for rs2243639. The primer sequences and PCR amplification conditions were described in the Materials and Methods. The allelic products of C and T for rs721917 in Panel A, and the allelic products of A and G for rs2243639 in Panel B are indicated by arrows. Individual samples are indicated below the gels. The 20 bp fragments from the T allele of rs721917 and the G allele of rs2243639 in the restriction enzyme digested PCR products have run off of the gels. M is a 1 kb DNA ladder. C is a negative control of PCR using dH₂O instead of DNA. Number 1-16 is the number of the 16 individual samples.

Figure 2. *SP-D* expression in human intestine from Northern and RT-PCR analysis

Panel A. Northern analysis Total mRNA (3.5 μ g) was separated on 1% agarose gel containing formaldehyde (0.22 M) and blotted on GeneScreen Plus membrane. The blot was hybridized with a 32 P-ATP-labeled *SP-D* antisense oligonucleotide probe. The Northern blot procedure and hybridization was described in the Materials and Methods. **Panel B. RT-PCR** PCR Primers for PCR amplification were SPD18s and SPD26r for *SP-D*, and Actin7s and Actin8r for β -actin. The PCR products were 484 bp for *SP-D* and 266 bp for actin. Primer sequences and PCR conditions were described in the Materials and Methods. Samples are indicated below the gels. Sm int: small intestine. M is a 1 kb DNA ladder, C is a PCR negative control.

Figure 3. Differential allele expression of *SP-D* in human intestinal tissues

cDNA (1 μ l) was used as template to amplify a 102 bp fragment containing rs721917 with primers SPD31s and SPD20r, and a 118 bp fragment containing rs2243639 with primers SPD34s and SPD23r as described in the Methods. PCR products (5 μ l) were digested with Fsp I (for rs721917) and Hha I (for rs2243639) and then separated on 8% PAGE. The PCR products of allele C and T for rs721917 in Panel A, and of allele A and G for rs2243639 in Panel B are indicated by arrows. The 20 bp fragments from T allele of rs721917 and G allele of rs2243639 in the restriction enzyme digested PCR products have run off of the gels. All samples in Panel A were heterozygous for the rs721917, and in Panel B were heterozygous for the rs2243639. Numbers 1-16 represent the number of the individual samples. Individual samples A1-16 are for rs721917 and B1-16 for rs2243639. C is a negative PCR control.

Figure 4. Illustration of the effect allelic expression on incomplete penetrance of disease phenotype

As a result of variation in allele expression, the final gene expression products (mRNA and protein) of a heterozygous individual (TC) may have an equal (balanced biallelic) amount of product from wild type allele T and mutated allele C; only a single (monoallelic) product of wild type allele T or mutated allele C; or only the dominant (imbalanced biallelic) product of wild type T or mutated allele C. The complete loss of heterozygosity or quantitative changes in expression of a mutated allele may lead to a different disease phenotype outcome.

Table 1

Genotype analysis of SP-D rs2243639 and rs721917 in CD, UC, and IBD

Genotypes	Cases	Controls	OR	95%CI	P value
rs2243639					
<u>For CD:</u>	28	60	1.5531	1.149-2.098	0.0036
AA	66	170			
AG	29	144			
GG					
<u>For UC:</u>					
AA	15	60	1.0300	0.765-1.385	0.8846
AG	75	170			
GG	43	144			
<u>For IBD:</u>					
AA	43	60	1.2577	0.985-1.590	0.0543
AG	141	170			
GG	72	144			
rs721917					
<u>For CD:</u>	14	50	0.8629	0.625-1.160	0.3285
CC	61	198			
CT	49	128			
TT					
<u>For UC:</u>					
CC	25	50	1.1961	0.892-1.602	0.2180
CT	67	198			
TT	41	128			
<u>For IBD:</u>					
CC	39	50	1.0189	0.805-1.289	0.9070
CT	128	198			
TT	90	128			

Table 2

Allelic expression patterns between diseased and normal tissues

Sample	Disease	Tissue type	C/T ratio	Balanced 0.8-1.2	Imbalanced 0.2-<0.8 >1.2-5.0	Monoallelic <0.2 & >5.0
888-46-1	UC	ND	1.451		IB	
		D	5.886			M
888-113-1	UC	ND	1.552		IB	
		D	2.983		IB	
888-68-1	UC	ND	1.058	B		
		D	3.963		IB	
888-128-1	UC	ND	1.487		IB	
		D	1.971		IB	
777-35-1	CD	ND	1.572		IB	
		D	1.354		IB	
777-35-1	CD	ND	1.572		IB	
		D	1.354		IB	
888-44-1	UC	ND	1.508		IB	
		D	1.312		IB	
888-53-1	UC	ND	1.008	B		
		D	1.281		IB	
888-69-1	UC	ND	1.208	B		
		D	1.064	B		
888-50-1	UC	ND	1.201		IB	
		D	0.641		IB	
132-01-1	CD	ND	0.860	B		
		D	1.194	B		
888-83-1	UC	ND	1.156	B		
		D	0.935	B		
777-57-1	CD	ND	0.986	B		
		D	0.986	B		

Sample	Disease	Tissue type	C/T ratio	Balanced 0.8-1.2	Imbalanced 0.2-<0.8 >1.2-5.0	Monoallelic <0.2 & >5.0
888-90-1	UC	ND	0.145			M
777-60-1	UC	D	0.959	B		
		ND	0.936	B		
145-01-1	UC	D	0.399		IB	
		ND	0.935	B		
888-52-1	UC	D	0.936	B		
		ND	0.154			M
777-14-1	CD	D	0.487		IB	
		ND	0.095			M
888-49-1	UC	D	0.137			M
		ND	0.128			M
		D	0.061			M