ORIGINAL INVESTIGATION

Molecular cloning of two novel mucin-like genes in the disease-susceptibility locus for diffuse panbronchiolitis

**Minako Hijikata · Ikumi Matsushita · Goh Tanaka · Tomoko Tsuchiya · Hideyuki Ito ·
Katsushi Tokunaga · Jun Ohashi · Sakae Homma · Yoichiro Kobashi · Yoshio Taguchi ·
Arata Azuma · Shoji Kudoh · Naoto Keicho**

Abstract Diffuse panbronchiolitis (DPB) is a rare complex genetic disease affecting East Asians and is strongly associated with the class I human leukocyte antigens (HLA)-B54 in Japanese and HLA-A11 in Koreans. We recently showed that an HLA-associated major susceptibility gene for DPB is probably located within the 200 kb in the class I region 300 kb telomeric of the HLA-B locus on the chromosome 6p21.3. We cloned two novel mucin-like genes designated panbronchiolitis related mucin-like 1 and 2 (*PBMUCL1* and *PBMUCL2*) in the candidate region, which form a mucin-like gene cluster together with two adjacent genes, *MUC21* and *DPCR1*. PBMUCL1 gene expression was remarkably upregulated by polyinosine–polycytidylic acid [poly(I:C)] stimulation in normal human bronchial epithelial cells redifferentiated at the air–liquid interface. We found genetic polymorphisms in PBMUCL1 gene which were associated with DPB: the A-allele of the

M. Hijikata · I. Matsushita · G. Tanaka · T. Tsuchiya ·
N. Keicho (✉)
Department of Respiratory Diseases,
Research Institute, National Center for Global Health
and Medicine, Tokyo 162-8655, Japan
e-mail: nkeicho-tky@umin.ac.jp

H. Ito
Department of Thoracic Surgery, National Center for Global
Health and Medicine, Tokyo 162-8655, Japan

K. Tokunaga
Department of Human Genetics, Graduate School of Medicine,
The University of Tokyo, Tokyo 113-0033, Japan

J. Ohashi
Molecular and Genetic Epidemiology,
Department of Public Health Medicine,
Doctoral Program in Life System Medical Sciences,
Graduate School of Comprehensive Human Sciences,
University of Tsukuba, Tsukuba, Ibaraki 305-8077, Japan

S. Homma
Department of Respiratory Medicine,
Toho University School of Medicine, Tokyo 143-8541, Japan

Y. Kobashi
Department of Pathology, Tenri Hospital,
Nara 632-8552, Japan

Y. Taguchi
Department of Respiratory Medicine,
Tenri Hospital, Nara 632-8552, Japan

A. Azuma
Division of Pulmonary Medicine, Infectious Diseases
and Oncology, Department of Internal Medicine,
Nippon Medical School, Tokyo 113-8602, Japan

S. Kudoh
Fukujuji Hospital, Tokyo 204-8522, Japan

Present Address:
G. Tanaka
Department of Respiratory Medicine,
The University of Tokyo Hospital, Tokyo 113-0033, Japan

Present Address:
T. Tsuchiya
Central Laboratory, ECI, Inc., Tokyo 153-0042, Japan

PBMUCL1 intron 2 single nucleotide polymorphism (SNP) was positively associated and variable numbers of tandem repeats (VNTR) polymorphism in exon 3 (1,890-base pair deletion) was negatively associated. Despite a strong association with HLA-B in the Japanese, the mucin-like gene PBMUCL1 is also one of the candidate genes of DPB susceptibility.

Introduction

Diffuse panbronchiolitis (DPB) is characterized by sino-bronchial infection and diffuse bilateral micronodular pulmonary lesions (Fraser et al. 1990) and is considered to be a complex genetic disease mostly affecting East Asians, occurring with a frequency of 0.00011 in Japanese people. DPB was first described in Japanese patients (Homma et al. 1983) and subsequently in Koreans (Kim et al. 1992) and Chinese (Chu et al. 1992; Tsang et al. 1998). Sugiyama et al. (1990) reported that the frequency of human leukocyte antigen (HLA)-B54 was much higher in Japanese DPB patients than that in the normal population. HLA-B54 is predominantly seen in East Asians (Imanishi et al. 1991). We confirmed that HLA-B gene *B*5401* encoding the B54 antigen was most strongly associated with the disease in the Japanese (Keicho et al. 1998). On the other hand, it was reported that HLA-A11 was most significantly associated with DPB in Koreans, while HLA-B54 was not (Park et al. 1999). These inconsistent observations suggested that the disease-susceptibility gene(s) might be located between *HLA-B* and *HLA-A* loci in the class I region of the major histocompatibility complex (MHC) on chromosome 6p21.3. By analyzing polymorphic genetic markers between *HLA-B* and *HLA-A* loci in disease and control populations, it was demonstrated that the most likely region for the putative disease locus is a 200-kb segment in the class I region, 300 kb telomeric of the *HLA-B* locus (Keicho et al. 2000). Besides positive association of HLA-B54 in Japanese or HLA-A11 in Koreans with DPB, negative association of HLA-B44 with DPB was observed both in Japanese (Keicho et al. 1998) and in Koreans (Park et al. 1999). Because there was a strong linkage disequilibrium (LD) between *HLA-B*4403* and *HLA-A*3303* both in Japanese (Saito et al. 2000) and in Koreans (Lee et al. 2005), it is also likely that genetic determinant of disease-resistance is not attributable solely to *HLA-B* loci.

The candidate region of disease-susceptibility was roughly bordered by the *CDSN* and *GTF2H4* loci (Keicho et al. 2000); however, only a few genes had been recognized (Matsuzaka et al. 2002; Itoh et al. 2008). In the present study, we tried molecular cloning of new genes in the candidate region, and investigated the association of polymorphisms in the new mucin-like genes with the disease.

Materials and methods

Gene prediction

The nucleotide sequence of 53L9 clone (GenBank accession number: AB023048), that is 192,650 base pairs in length and covers the candidate region, was subjected to GENSCAN computer program (http://genes.mit.edu/GEN-SCAN.html) to predict gene-like structures. Repetitive elements were searched by RepeatMasker computer program (http://www.repeatmasker.org).

Human bronchial epithelial cells

Primary-cultured human bronchial epithelial (HBE) cells were obtained from cancer-free bronchi of the surgically resected lung. Written informed consent was obtained from each individual. The ethical committee approved the study protocol. We isolated and cultured HBE cells as described (Gray et al. 1996), and cells of passages 3–5 were used in this study. To obtain redifferentiated mucociliary cultures, we further supplemented bronchial epithelial growth medium (BEGM; Biowhittaker, Walkersville, MD, USA) with human epidermal growth factor (final 25 ng/ml) and retinoic acid (5×10^{-8} M) as described (Gray et al. 1996). Cells were transferred at a density of 5×10^5/well onto collagen-coated 6-well Transwell plates (Corning, Corning, NY, USA) and cultured with a 1:1 mixture of BEGM and Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with the same final concentration of media additives as above. Media was applied both apically and basally until cells reached confluence, and when the cells reached confluence (day 0), media was applied only basally to establish an air–liquid interface (ALI). The cells were maintained until a differentiated cell population with mucus secretion was present (day 14).

Reverse transcription/polymerase chain reaction (RT/PCR)

Total RNA was extracted from primary-cultured HBE cells and from NCI-H292 cells (ATCC Number: CRL-1848) with Trizol (Invitrogen). Total RNA from adult lung (Stratagene, La Jolla, CA, USA) was also used. Five micrograms of the total RNA were subjected to reverse transcription (RT) with oligo(dT)$_{12–18}$ primers using SuperScript III RNaseH$^-$ Reverse Transcriptase (Invitrogen) as recommended by the manufacturer. Two microliters of RT reaction was subjected to PCR amplification with primer sets listed in Suppl. Table 1 using AmpliTaq Gold DNA polymerase (Applied Biosystems, Foster City, CA, USA). Amplified products were purified and sequenced bidirectionally with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) using 3100 Genetic Analyzer

(Applied Biosystems). Human MTC panel (Clontech, Mountain View, CA, USA) was subjected to PCR amplification with primer sets shown in Suppl. Table 1 to determine the tissue expression of candidate genes.

Cloning of full-length cDNA

Longer cDNA fragments of G2 and G4 transcripts were amplified using TaKaRa LA Taq (TaKaRa, Shiga, Japan) with primer sets listed in Suppl. Table 2. PCR cycling condition was 40 cycles of 95°C for 15 s, 65°C for 15 s and 72°C for 4 min. RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) was carried out using First-Choice RLM-RACE Kit (Ambion, Austin, TX, USA) following the manufacturer's instructions. Gene-specific primers were listed in Suppl. Table 3. PCR products were purified and sequenced as described above. Full-length cDNA sequences were obtained by the overlapping three PCR fragments.

The coding region of G4 transcript [5,322 nucleotides (nt) from ATG codon in exon 2 to stop codon in exon 5] was amplified by RT/PCR using Platinum Taq DNA Polymerase High Fidelity (Invitrogen). PCR condition was 45 cycles of 94°C for 15 s, 60°C for 15 s, 72°C for 6 min. Total RNA was reverse transcribed with random nonamers, and 5′ alternate-exons of G4 gene were amplified by PCR using AmpliTaq Gold DNA polymerase (Applied Biosystems). Primers were listed in Suppl. Table 4.

Real-time RT/PCR

HBE cells were stimulated with polyinosine–polycytidylic acid [poly(I:C)]; Sigma–Aldrich, St. Louis, MO, USA) (100 µg/ml) or *Pseudomonas aeruginosa* lipopolysaccharide (LPS; Sigma–Aldrich) (20 µg/ml). Cells were harvested after 24 h, and the total RNA was extracted using RNeasy Mini Kit (QIAGEN, Hamburg, Germany). After reverse transcription with random nonamers, G4 mRNA expression was analyzed by real-time PCR using primers (Suppl. Table 5) and SYBR Premix Ex Taq (TaKaRa) by iCycler (BioRad, Hercules, CA, USA) following the manufacturer's instructions. The cycle threshold (Ct) of gene transcript was normalized to the Ct of β-actin, and the relative gene expression-fold was determined by the $\Delta\Delta$Ct method. The pure amplification of the target was determined from a single peak of a dissociation curve. Results were calculated as fold induction over control, and analyzed by means of the paired Student's *t* test. Data are expressed as the means ± SD.

HBE cells were cultured at ALI for 14 days, and relative expression of G4 gene was compared before and after redifferentiation. Redifferentiated HBE cells were also stimulated 24 h with the addition of poly(I:C) (100 µg/ml)

onto the apical surface and into the lower chamber, and fold induction of G4 gene expression was compared.

Immunohistochemistry

To detect G4 expression in the lung tissue, rabbit polyclonal anti-G4 peptide (68 N-TPTNVIKPSGYLQP-C) antibody was made and affinity purified (GENENET, Fukuoka, Japan). The rabbit preimmune serum was used as a negative antibody control. Autopsy samples of formalin-fixed paraffin-embedded lung tissues from patients with DPB and from those without apparent lung disease were stained as described (Kamio et al. 2005).

Cases and controls

In total, 108 unrelated Japanese patients with DPB participated in the present study, in which 92 were initially analyzed and 16 were newly added later. All fulfilled the diagnostic criteria described elsewhere (Keicho et al. 2000). As a control, 98 healthy unrelated Japanese were initially studied (control 1), and then 220 individuals were added as another set of controls (control 2). Because of the rarity of this disease, another independent case group could not be prepared. Control subjects and patients were from the same area. Distribution of polymorphisms analyzed here was not affected by age and sex of each individual. The ethical committee approved the study protocol, and written informed consent was obtained from each individual.

DNA typing

Luminex Multi-Analyte Profiling system (xMAP) with WAKFlow HLA typing kit (Wakunaga Pharmaceutical, Hiroshima, Japan) was used to determine HLA-A and -B genotypes. Microsatellite marker C2_4_4 which showed the strongest association with DPB in the previous study was determined as described (Keicho et al. 2000). Exon regions of GENSCAN-predicted genes (G1, G2, G3, G4, G7 and G10), *CDSN*, *C6orf15*, *MUC21*, *DPCR1* and *SFTA2* were amplified by PCR and directly sequenced. Variations were genotyped in 92 cases and 98 controls.

VNTR and promoter polymorphisms of G2 and G4

The VNTR polymorphism in exon 3 of G4 gene was analyzed. Genomic DNA was amplified with primers 5′-GGCGGATCCCCCACCACCTTATTTGTTCTCCAC-3′ and 5′-GGCGGATCCTTAAAATGTGTGGGCTCAAAGAAG-3′ and electrophoresed on 2% agarose gels with ethidium bromide. The PCR products were classified by their length. Amplified products of respective length were
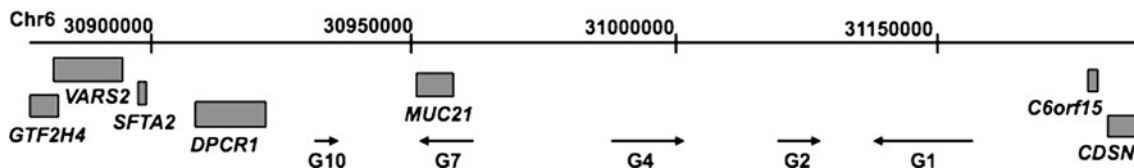
Fig. 1 Organization of genes and GENSCAN-predicted genes in the candidate region. The candidate region of an HLA-associated susceptibility gene for DPB is the 200-kb region on 6p21.3, bordered by the *CDSN* and *GTF2H4* loci (Keicho et al. 2000). Position of known genes (*boxes*) and GENSCAN-predicted genes (*arrows*) in the 200-kb region are shown

selected and digested by *Mbo*I and ligated into *Bam*HI digested pBlueScript II SK+ vector (Stratagene). Subcloned plasmids using DH5α cells were sequenced bidirectionally with T3 and T7 primers, and obtained sequences were reconnected to cover the whole VNTR region. Thus, the location and the length of insertion/deletion polymorphisms in the VNTR region were determined. Direct sequencing of PCR products without subcloning was also carried out. Primers used for sequencing are listed in Suppl. Table 6. VNTR regions of *G2* and *MUC21* were analyzed by direct sequencing of PCR products without subcloning.

Comprehensive SNP screening of the candidate region

Three disease-associated haplotypes containing *HLA-B*5401* or *B*5504* were identified in the previous study (Keicho et al. 2000). They shared 200-kb candidate segment containing disease-susceptible allele, C2_4_4-231. Twelve DPB samples and 12 control samples which were homozygotes for C2_4_4-231 were selected and the candidate 100-kb region with high LD obtained from the SNP screening described above and G4 gene region including intron sequences were amplified by overlapping PCR products. Direct sequencing of PCR products was effected using appropriate inner primers. When an allele which was carried by disease-associated haplotype(s) was found, the polymorphism was genotyped in 92 cases and 98 controls. A microsatellite marker D6S2694 which lies 500 bp upstream of G2 gene transcription start site was genotyped by GeneScan (version 3.7) and Genotyper (version 2.0) software (Applied Biosystems), using fluorescent-labeled primer 5′-(HEX)-GGCTAGTTTACCACATGTAACTG-3′ and 5′-ATGTTTGAAAATGTTGCTGG-3′.

Statistical analysis

Disease associations with markers were assessed by Fisher's exact test. In this study, P values less than 0.05 were generally considered significant. The odds ratio (OR) was defined as the cross-product ratio of the numbers shown in the 2×2 table. To examine whether genotype frequencies in the populations are compatible with Hardy–Weinberg equilibrium, Hardy–Weinberg exact tests were carried out using the program Arlequin version 3.11 (Excoffier et al. 2005). To assess the extent of pairwise LD between polymorphisms, we calculated Lewontin's D' and r^2 for polymorphisms by Haploview version 4.2 (Barrett et al. 2005). Frequencies of haplotypes containing multiple polymorphic sites were estimated by Arlequin (Excoffier et al. 2005).

Results

Molecular cloning and expression of cDNAs coding for mucin-like structures

By GENSCAN program, 12 genes were predicted in the 200-kb candidate region that showed association with the disease in the previous study (Keicho et al. 2000). When predicted genes were serially numbered G1 (centromeric) to G12 (telomeric), G12 was already known as valyl-tRNA synthetase 2, mitochondrial (*VARS2*, RefSeq NM_020442). Two exons of G11 coincided partially with those of diffuse panbronchiolitis critical region 1 (*DPCR1*, GenBank accession no. AB064272) (Matsuzaka et al. 2002). When 53L9 sequence was analyzed by RepeatMasker to find repetitive elements, exons of four predicted genes (G3, G5, G6 and G8) were located in the long terminal repeat (LTR) regions. G9 was not tested, being a single-exon gene. The putative amino acid sequences of G2 and G4 showed characteristic of mucin, that is, tandem repeat of serine and threonine-rich peptide. Human mucoepidermoid pulmonary carcinoma cell line NCI-H292 and primary-cultured HBE cells were investigated for the mRNA expression of predicted genes. Total RNA from NCI-H292 and HBE cells were subjected to RT/PCR with primer pairs that span at least one intron to detect five predicted genes' transcripts (G1, G2, G4, G7 and G10) (Fig. 1).

Among five predicted genes, G2 transcript was positive in primary-cultured HBE cells by RT/PCR and was negative in NCI-H292 cells. G4 transcript was positive both in NCI-H292 cells and in primary-cultured HBE cells. G1, G7 and G10 transcripts were negative in both cells. PCR products were sequenced to confirm the amplification of the predicted genes. By combination of RT/PCR and 5′- and 3′-RACE procedures, three overlapping cDNA fragments that covered full-length cDNAs of G2 and G4 were obtained. Full-length coding region of G4 consisting

Fig. 2 Deduced amino acid sequences of G4 and G2 genes. Deduced amino acid sequence of G4 gene (**a**) and G2 gene (**b**) are shown. Putative signal peptide sequences are *underlined*, and transmembrane domain is *double-underlined*

a

MRRGNISPAFWFLWLLLFGLLGPSSENTTAFTKGSDTTTASITGSETTMASTMASTTALT
TGSKITTDSTTGSETTSASTMASTAAFTTGSETNTASTTDSGTTIASTRTFTTGSDTTTG
STAGSETIVASTTVSGTTTTFTIASTTVPETT

MASSTTSTAGSE KTMASSIISE TTMASTTGSE TATVSTTGSE TTTTSTASSE ATKVSTTGSE
TTTASTAGSE TTTTSTSMAGSE ATTTSTADSK VITASSMSSE TTVAPAAGSN TTTASTTGSE
TTTILIKASE TTTASTAGSE TTTPSPTGSQ TTIVSISGSE ITTTSTAGSE NTTVSSAGSG
TTTASMAGSE TTVSTAGSE TTTVSITGTE TTMVSAMGSE TTTNSTTSSE TTVTSTAGSE
TTTVSTVGSE TTTAYTADSE TTAASTTGSE MTTVFTAGSE TITPSTAGSE TTTVSTAGSE
TTTVSTTGSE TTTASTAHSE TTAASTMGSE TTKVSTAGSE TTVSTAGSE TTAASTEDSE
TNTAFTEDSK TTTASTTGFE TTAASTTGSE PTMASTMGSE TTMASTIGPE TTKVSTASSE
VTTVFAAGSE TIRASTVGSE TTTVSTTGSE TTTASIMGSE TSTDSTTGSE TTTASTEGSE
TTTASTEGSE ATTVSTTGSE TTTVSITDSE TTTTCTEGSE MTAVSTTVFE TTTASTEGSE
ITIASTSDSE TTTASTEGSE TTTVTTAGSE TKTAYTTGSE TTTASNTGLE TTTVFTIGSD
TTTASTEGSE TTAVSATGSE MTTVSTEGSE NTTVSTTGSE TTTVSTTGLE TTTTSTEGSE
MTTVSTTGAE TTTDSTEGSG TTAASTAGSE TTTVSTADSE NTTASTADSE TTSASTTGSE
TTTASTTSSE TTTASTEGSE TTTVSTTDSE TTMVSTTGSE RTITSTEGSE TTTVSATGSE
TTVSTEGSG TTTVSITGSE TTKVSTTGSE TTTTSTEGSE ITTASITGSE TTTASTEGSE
TTTASTEGSE TTSASTTGSE TTTASTTSSE TTMASIMGSE TTMASTIGSE TTKVSTASSK
MTTVFTENSE TTIASTTASE TTTVSTAGSE TIPASTAGSE TTTTTSTEGSE TTTASTEGSE
TTTASTESSE TTTATTIGSE TTTASTEGSE TTTTSTEGSE TTTASTEGSE ITTVSTTGSE
TTTASTEGSE TTTASTEGSE LTTVSTTGSE TITVSAEGSE TTTVTTMGSE TTTASTAGSE
TTTVSTAGSE TTTASIEGSE TTTVSSTGSE TTTVSTTGTE TTITSTEGSE TTTVTTAGSE
TTAVYTTGSE TTTTSTEGSE TTTVSTTGSE TTTASTADLE TTTVSTSGSG TTTASTAGSE
TTTVYITGSK TTTASTEGSE ATTVSTTSSE TTTASTTGSE MTTVFTTVSE TTTVSTIGSE
ATTSSAAGSE ATTTSTEGSE TTTASTAGSE TTTASTAGSE TTTASTSGSE TNTACTTGSE
TSTPSSAGSE TNTAFIIGSE STIASTASLE PTATSLTGSE

TTTVSITASGATAASTTVSSTTFVLTKATDVSIQPITNTPMSGTRTTGTRLTASSSVTMAPGMDFT
ASAASHTVPGIVLNTSGLGTSTMGASSTTSAHGVRTTTGSTREPTSSTFQETGPVSMGTNTVSMSH
TPTNVIKPSGYLQPWAIILISLAAVVAAVGLSVGLSFCLRNLFFPLRYCGIYYPHGHSHSLGLDLN
LGLGSGTFHSLGNALVHGGELEMGHGGTHGFGYGVGHGLSHIHGDGYGVNHGGHYGHGGGH

b

MPRYVPLLLLLLLLRCSERGGGVNFGEKDAKVPGTWRDGVRVPGEGASWDSDRASPERRY
GIVGLSQSISTKHPETSPKDSRIRENDV

TADGRTTEDHI TADPGTTEDSV TADPGTTEDNV TVDPGTTEGSV TADPATTKDYV
SADPGTTKDSV TADPGTTENFV TADPGTTKDSI TADPRTTEDSV TADPGTTKHSI
TVDPGTTEDSV TADPGTTKHSI TADPGTTEDSV TADPGTTEDET TKHGDTHLL

of 5,322 nt was amplified by single RT/PCR from NCI-H292.

G4 gene consisted of five exons and spanned approximately 29 kb. The cDNA sequence obtained consisted of 6,005 nt (AB560770), and the deduced amino acid sequence was 1,773 amino acids in length (Fig. 2a). Signal peptide of 24 amino acids was predicted in its N-terminus, and one transmembrane domain was predicted near its C-terminus, thus G4 was predicted to be a type 1 membrane protein according to SOSUI ver1.11 (http://bp.nuap.nagoya-u.ac.jp/sosui/) (Hirokawa et al. 1998). Large exon 3 (4,599 nt) encoded 124 tandem repeats of 10 amino acids rich in serine and threonine residues which could be *O*-glycosylated, but proline residues was poor. In total, serine was 18.4% and threonine was 33.6%. Almost all tandem repeats consisted of 10 amino acids, with a few exceptions of variable length (9–14 amino acids). Repeating sequences were not identical, and 85 repeat sequences out of 124 had a unique sequence. The most frequent sequence was TTTASTEGSE, which appeared 14 times. Seven N-linked glycosylation motifs were predicted (Genetyx-Mac, Software Development, Tokyo, Japan). Taken these characteristics together, G4 gene was predicted to encode a mucin-like glycoprotein. By 5′-RACE method, transcription start sites were different between HBE cells and NCI-H292 cells (Fig. 3). First exon obtained from NCI-H292 cells overlapped exon 2 from HBE cells. Transcription start site of HBE cells was 4,523 bp upstream of that of NCI-H292. The same 5′ exon as that in HBE cells was obtained from the total RNA from adult lung tissue (Stratagene).
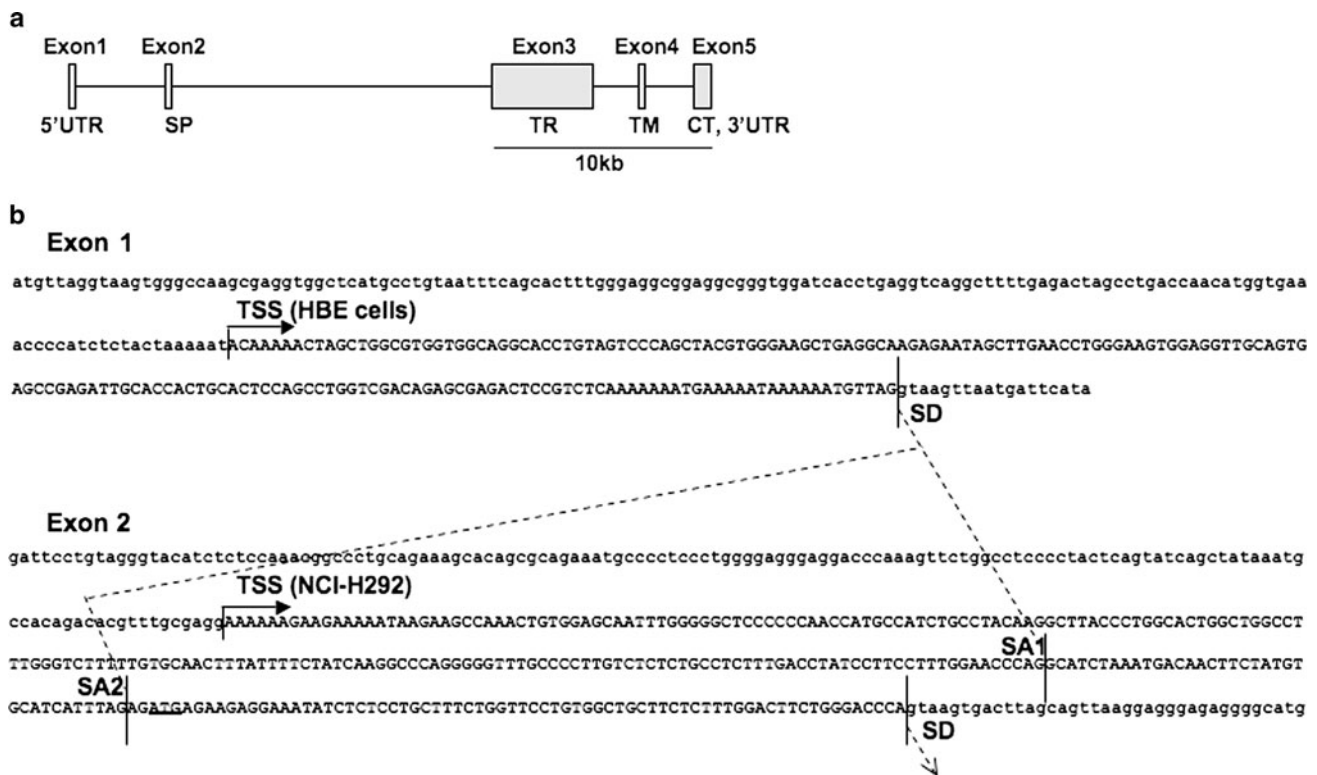
Fig. 3 Alternate splicing of 5′exon in G4 gene. Genomic structure of G4 gene (**a**) and nucleotide sequences of exon 1 and exon 2 (**b**) are shown. Exon sequences are indicated by *uppercase letters* and upstream region and intron sequences by *lowercase letters*. *UTR* untranslated region, *SP* signal peptide, *TR* tandem repeats, *TM* trans-membrane, *CT* cytoplasmic tail, *TSS* transcription start site, *SD* splice donor, *SA* splice acceptor. In NCI-H292 cells, transcription start site was upstream of exon 2 splice-acceptor sites of HBE cells. Translation start site (ATG) in exon 2 is *underlined*
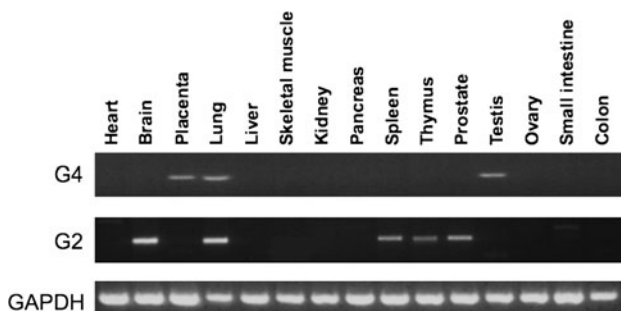

Fig. 4 Expression of G4 and G2 genes in normal tissues. Transcripts from G4 and G2 genes in human organs were detected by RT/PCR. Human MTC Panel (CLONTECH) was used as template for PCR as described by the manufacturer

G2 gene consisted of four exons and cDNA sequence was 1,671 nt in length (AB560771). It coincided partially with AK094333 cDNA sequence of *HCG22*. The deduced amino acid sequence consisted of 251 amino acids (Fig. 2b). Signal peptide of 22 amino acids was predicted in its N-terminus, and no transmembrane domain was found. Therefore, G2 protein was predicted to be secreted (SOSUI). Exon 3 encoded 15 tandem repeats of 11 amino acids. The putative protein was rich in proline (8.0%), serine (7.6%) and threonine (18.7%). One N-linked glycosylation motif was predicted.

The mRNA expression of G2 and G4 genes was examined by PCR screening in commercial human multiple tissue cDNA panel, human MTC panel (CLONTECH). G4 mRNA was detected in the placenta, lung and testis. G2 mRNA was detected in the brain, lung, spleen, thymus and prostate (Fig. 4).

Expression of G4 mRNA in redifferentiated HBE cells

G4 mRNA expression in passage-3 HBE cells under submerged and unstimulated culture conditions from 10 individuals were analyzed by real-time RT/PCR, however, the expression levels were relatively low compared to the total RNA of lung tissue (Stratagene). HBE cells were stimulated with poly(I:C) or LPS, and the G4 expression was upregulated after stimulation with poly(I:C) in HBE cells (Fig. 5a). Redifferentiation of HBE cells from three individuals at ALI for 14 days was accompanied with eightfold induction of expression. Redifferentiated cells were also stimulated with poly(I:C), and more than 100-fold
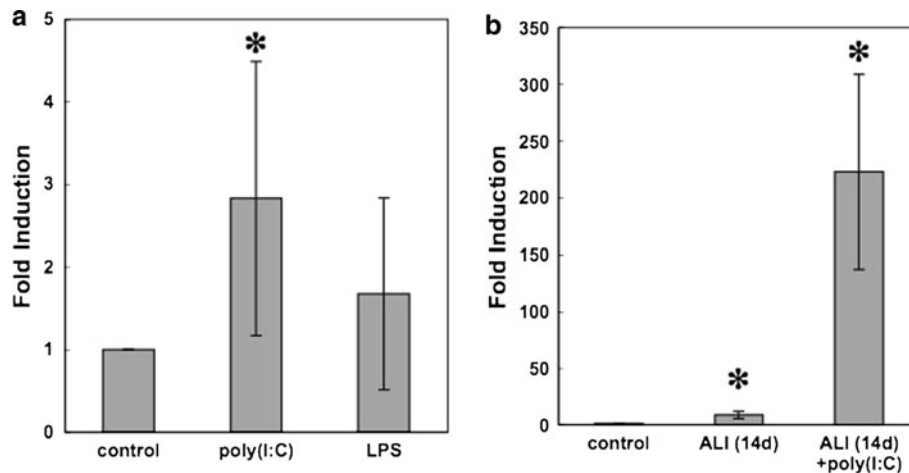
Fig. 5 Induction of G4 transcripts by redifferentiation and by poly(I:C) stimulation. **a** HBE cells were stimulated with poly(I:C) ($n = 9$) or with LPS ($n = 3$). After 24 h, the cells were harvested and G4 mRNA expression level was compared to unstimulated cells by real-time RT/PCR as described in "Materials and methods". Fold inductions of G4 mRNA with poly(I:C) or with LPS are shown as means ± SD. *$P < 0.05$ compared with nonstimulated control cells ($n = 9$). **b** HBE cells ($n = 3$) were redifferentiated at ALI for 14 days, and redifferentiated cells were stimulated with poly(I:C). G4 mRNA expression level was compared to undifferentiated and unstimulated control cells. Data are shown as means ± SD. *$P < 0.05$ compared with control ($n = 3$)

Fig. 6 Immunohistochemistry of **a** lung section from a DPB patient showed strong staining of serous cells in hypertrophic submucosal gland with rabbit anti-G4 peptide antibody; **b** lung section from patients free of any apparent lung disease showed moderate staining of serous cells in submucosal gland. Lung section from DPB (**c**) and normal (**d**) showed no reactivity with the preimmune rabbit sera
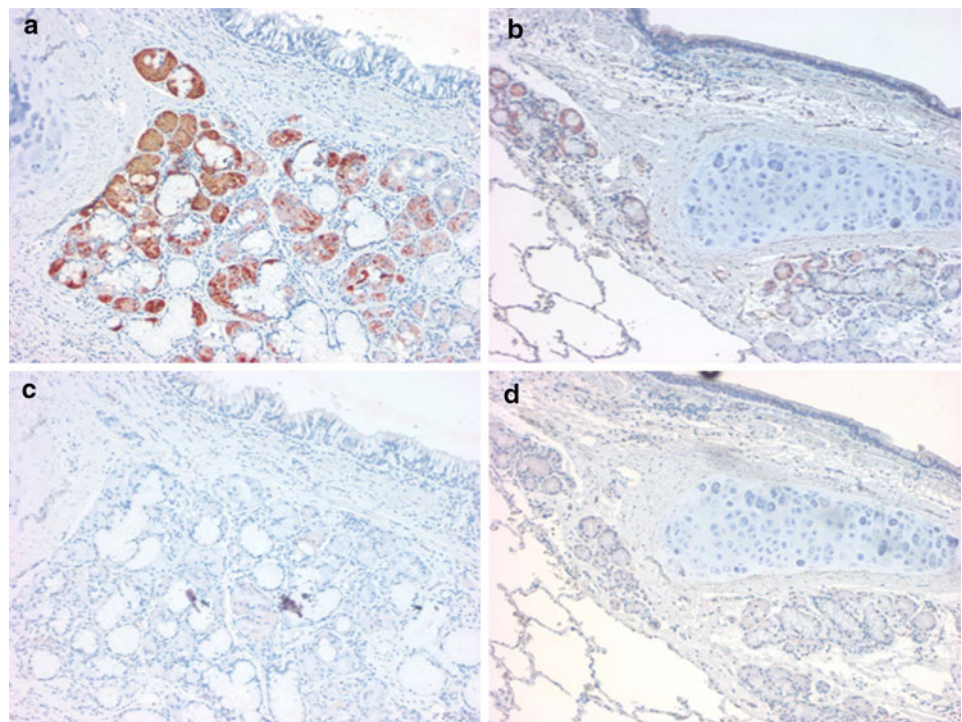
induction was observed (Fig. 5b). To identify whether the redifferentiation or poly(I:C) stimulation changed the transcription start site, RT/PCR with 5′ exon-specific primers were performed. In HBE cells, 5′ exon was always the same, and the transcription start site of NCI-H292 cells was not used.

Immunohistochemistry

Immunohistochemistry for G4 was performed on formalin-fixed paraffin-embedded lung tissues from autopsied cases of DPB and cases with no airway disease. In the intact lung tissue without chronic airway disease, serous cells of the

submucosal glands were moderately stained (Fig. 6). The epithelial cells of bronchi and goblet cells were not stained. In lung tissues with DPB, serous cells in the hypertrophic submucosal gland were strongly stained. They were not stained with the preimmune rabbit sera.

Association of DPB with SNPs in exon regions of the genes and predicted genes located in the 200-kb region

Genomic DNA containing predicted exons and exon–intron boundaries of the candidate 200-kb was amplified by PCR and sequenced. One hundred and sixty single nucleotide polymorphisms (SNPs) were identified and genotyped in 92 cases and 98 controls. Hardy–Weinberg exact tests were performed, and SNPs that significantly deviated ($P < 0.05$) from Hardy–Weinberg equilibrium in the control population were excluded from further study. Consequently, 157 SNPs were analyzed. Twenty-three alleles were significantly associated with the disease by 2×3 Fisher's exact test ($P < 0.05$). Twenty-two SNPs among them were located in the centromeric 100-kb region bordered by *C6orf15* and G4 gene, and one was in *CDSN*. Many of the SNPs that were found to be significantly associated with the disease were in high LD with each other ($r^2 > 0.9$). Fifty-seven SNPs located in the telomeric 80-kb region including *MUC21, DPCR1* and *SFTA2* were not associated with the disease. Supplementary Figure 1 shows the LD pattern of the 200-kb candidate region in the Japanese population from the HapMap project (The International HapMap Consortium 2003). A nonsynonymous SNP (N1712D) in exon 5 of G4 gene (rs4248153) showed the strongest positive association with DPB among these SNPs [$P = 0.010$, OR = 2.3 (95% confidence interval (CI) 1.24–4.18)] assuming a dominant model from the findings in the previous study (Keicho et al. 2000).

VNTR polymorphism

VNTR sequence in the large exon 3 of G4 gene was analyzed. Variations in the length of exon 3 observed in DNA samples from patients and controls were 4,599, 4,598, 4,596, 3,735, 3,006 and 2,709 bp (Fig. 7). One-bp deletion caused frame-shift and premature stop codon. In the VNTR of G2 gene, there was 4-bp deletion (rs71984633) which caused frame-shift and premature stop codon.

We also investigated association of VNTR polymorphism with the disease. The shortest VNTR in G4 gene (1,890-bp deletion) was negatively associated with DPB [$P = 0.012$, OR = 0.29 (95% CI 0.10–0.84)]. It was in strong LD with HLA-B*4403 ($D' = 0.98$, $r^2 = 0.94$). Four-bp insertion in G2 gene was positively associated with the disease [$P = 0.034$, OR = 2.5 (95% CI 1.1–5.6)]. The
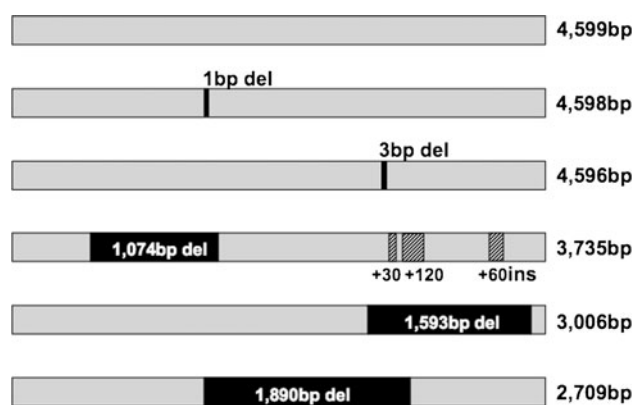

Fig. 7 Schematic diagram of VNTR polymorphisms in exon 3 of G4 gene. VNTR of G4 gene was highly polymorphic and varied in length and position of insertion/deletion. The locations of insertion/deletion polymorphisms in exon 3 are shown. Exon 3 consisted of 4,599 bp in 53L9 sequence. The exon 3 nucleotide length of each variation is shown on the right hand side

MUC21 VNTR polymorphisms did not show any significant association with the disease.

Comprehensive SNP screening of the 100-kb candidate region

Since 22 SNPs, significantly associated with the disease, were located in the centromeric 100-kb region bordered by *C6orf15* and G4 gene, we tried comprehensive SNP screening of the 100-kb region. We selected 12 DPB samples and 12 control samples which were homozygotes for C2_4_4-231. Among the 24 samples, 9 had *HLA-B*5401* and one had *HLA-B*5504*. In the G4 gene region, approximately 10-kb sequence in intron 2 could not be analyzed due to repeat sequence. In total, 247 polymorphisms were observed in the 100-kb region from *C6orf15* to G4. Two polymorphisms were related to the disease-associated haplotype and showed strong association: One of these two polymorphisms, A/G SNP in intron 2 of G4 gene (ss252451736) was genotyped in 92 cases and 98 controls and showed association with DPB [$P = 0.019$, OR = 2.8 (95% CI 1.2–6.2)]. Intron 2 of the G4 gene was long and consisted of 14,726 bp and the A/G SNP was located 1,962 bp upstream of the exon 3 splice-acceptor site. In the 5′ near gene region of G2, D6S2694-428 allele [(TTCC)$_4$ (TTTC)$_{13}$(14nt)$_{ins}$ or (TTCC)$_5$(TTTC)$_{12}$(14nt)$_{ins}$] of microsatellite polymorphism [(D6S2694 (TTCC)n (TTTC)n(14nt)ins/del)] also showed positive association with the disease [$P = 0.0049$, OR = 2.3 (95% CI 1.4–4.8)]. In the previous association study where microsatellite markers were used, microsatellite C2_4_4 showed the strongest association with DPB (Keicho et al. 2000), whereas the new marker polymorphism D6S2694 showed stronger association with DPB than C2_4_4, though marker

Table 1 Genotype frequencies of HLA-B and G4 gene polymorphisms in 108 DPB patients and controls

Polymorphism	Patients with DPB (n = 108)	Control1 (n = 98)	Control 2 (n = 220)	OR		P value	
				Versus control 1	Versus control 2	Versus control 1	Versus control 2
HLA-B							
5401/5401	4 (3.7)	0 (0.0)	0 (0.0)				
5401/others	36 (33.3)	11 (11.2)	28 (12.7)	4.7 (2.2–9.7)	4.0 (2.3–7.0)	P < 0.0001	P < 0.0001
Others/others	68 (63.0)	87 (88.8)	192 (87.3)				
HLA-B							
4403/4403	0 (0.0)	1 (1.0)	0 (0.0)				
4403/others	5 (4.6)	15 (15.3)	33 (15.0)	0.25 (0.09–0.71)	0.28 (0.10–0.73)	P = 0.0097	P = 0.0054
Others/others	103 (95.4)	82 (83.7)	187 (85.0)				
G4 exon 5 (rs4248153)							
A/A	24 (22.2)	15 (15.3)	30 (13.6)				
A/G	52 (48.1)	38 (38.8)	99 (45.0)	2.0 (1.1–3.6)	1.7 (1.0–2.7)	P = 0.021	P = 0.040
G/G	32 (29.6)	45 (45.9)	91 (41.4)				
G4 exon 3 (1,890-bp deletion)							
del/del	1 (0.9)	1 (1.0)	0 (0.0)				
del/ins	5 (4.6)	15 (15.3)	32 (14.5)	0.30 (0.11–0.81)	0.35 (0.14–0.85)	P = 0.014	P = 0.017
ins/ins	102 (94.4)	82 (83.7)	188 (85.5)				
G4 intron 2 (ss252451736)							
A/A	1 (0.9)	0 (0.0)	0 (0.0)				
A/G	24 (22.2)	10 (10.2)	24 (10.9)	2.7 (1.2–5.9)	2.5 (1.3–4.6)	P = 0.016	P = 0.0049
G/G	83 (76.9)	88 (89.8)	196 (89.1)				

Odds ratios (OR) and P values assuming a dominant model of the minor alleles are shown. Data are presented as n (%) or mean (95% confidence interval)

ins insertion, *del* deletion

allele C2_4_4-231 was also associated with the disease [$P = 0.022$, OR = 2.7 (95% CI 1.2–6.0)] in 92 patients and 98 controls of the present study.

Representative disease-associated marker alleles were finally genotyped in 108 cases, control 1 ($n = 98$) and control 2 ($n = 220$), and the result is summarized in Table 1. Consistent with the previous studies (Keicho et al. 1998, 2000), the frequency of *HLA-B*5401* was significantly increased in the patient population than in the controls, and the frequency of *HLA-B*4403* was significantly decreased in the patient population. Except for these association with HLA alleles, the A-allele of the SNP in G4 intron 2 (ss252451736) showed the strongest positive association, and VNTR polymorphism in G4 gene (1,890-bp deletion) showed the strongest negative association in the candidate region, assuming a dominant model. As shown in Table 2, haplotype frequencies were estimated among data from 318 controls by EM algorithm with the Arlequin software. The A-allele of the SNP in G4 intron 2 was found in HLA-B*5401 and A*2402 containing haplotype, and VNTR polymorphism in G4 gene (1,890-bp deletion) was found in HLA-B*4403 and A*3303 containing haplotype. The frequency of HLA-B*5401, HLA-B*5504 and A*1101

containing haplotype was low, and the haplotype structure was not clearly estimated.

Discussion

We raised the hypothesis that a major disease-susceptibility gene for DPB is located between the *HLA-B* and *HLA-A* loci and that its founder mutation occurred on an ancestral chromosome bearing HLA-B54 and HLA-A11 alleles in East Asia. In Japan, it has been confirmed that HLA-B54 is associated with DPB (Sugiyama et al. 1990; Keicho et al. 1998, 2000). However, Park et al. (1999) reported that HLA-B54 was not associated with DPB in Koreans, though the allele frequency of HLA-B54 is as high in the Korean population as in Japanese. Instead, HLA-A11 showed significant association with DPB in Koreans. In recent report from China, HLA-A11 was significantly increased, and HLA-B54 was not associated with the disease (She et al. 2007).

The first screening demonstrated that a disease-susceptibility gene is most likely located within the 200 kb in the class I region, 300 kb telomeric of the *HLA-B* locus

Table 2 Estimated haplotypes and their frequencies in the control population (n = 318)

Frequencies	HLA-B	G4 exon 5 (rs4248153)	G4 exon 3 (1,890 bp ins/del)	G4 intron 2 (ss252451736)	HLA-A
0.1066	B*5201	G	ins	G	A*2402
0.0660	B*4403	G	del	G	A*3303
0.0494	B*0702	A	ins	G	A*2402
0.0330	B*4601	A	ins	G	A*0207
0.0263	B*5401	A	ins	A	A*2402
0.0257	B*1501	G	ins	G	A*1101
0.0253	B*5101	G	ins	G	A*3101

ins insertion, *del* deletion

(Keicho et al. 2000). In the present study, we first predicted exon-like structures within the candidate region by GEN-SCAN computer program, and subsequently cloned two novel mucin-like genes. In the candidate region, *DPCR1* which encoded mucin-like domain (Matsuzaka et al. 2002) and a novel membrane-tethered mucin gene *MUC21* (Itoh et al. 2008) were cloned. *DPCR1*, *MUC21*, *G4* and *G2* are all located in cluster in 6p21.3. Mucins may be defined as having one long mucin domain rich in proline, threonine, and serine (PTS domain), which is heavily glycosylated through GalNAc *O*-linkage to the serine and threonine residues (Lang et al. 2007). To date, at least 21 different human mucin genes have been cloned (Rose and Voynow 2006; Itoh et al. 2008); however, with this definition of mucins, 18 mucin genes have been assigned to the MUC gene family according to the HUGO nomenclature (http://www.genenames.org). In human genome, two mucin gene clusters are already known (Rose and Voynow 2006). Oligomeric mucus/gel-forming mucin genes *MUC6*, *MUC2*, *MUC5AC* and *MUC5B* are located in a cluster of genes in 11p15.5, and cell-surface associated mucin genes *MUC3A*, *MUC3B*, *MUC12*, *MUC17* are located in a cluster in 7q22. In addition, *MUC20* and *MUC4* are adjacent genes in 3q29.

DPCR1 and MUC21 transcripts were expressed in lung tissue (Matsuzaka et al. 2002; Itoh et al. 2008). We could detect mRNA expression of both G4 and G2 genes by RT/PCR in the lung tissue and HBE cells. Tissue distribution of G4 by immunohistochemistry resembled that of MUC7, a secreted mucin, which is localized to the serous cells of submucosal gland (Sharma et al. 1998). G4 mRNA expression levels in undifferentiated HBE cells were relatively low, but they increased when HBE cells were redifferentiated by air–liquid interface culture like some other mucin genes (Ross et al. 2007). G4 mRNA was upregulated by poly(I:C), especially when HBE cells were redifferentiated at air–liquid interface culture. Therefore, under certain inflammatory conditions related to dsRNA stimulation, not only serous cells in submucosal gland but also airway epithelial cells in vivo may express G4. Increased mucus production is associated with bacterial or viral infections (Rose and Voynow 2006; Voynow et al. 2006). As regards

dsRNA analog poly(I:C) stimulation, it was initially reported that poly(I:C) activated MUC2 transcription activity (Londhe et al. 2003). Recently, Tadaki et al. (2009) reported that poly(I:C) induced a small but significant increase in MUC5AC mRNA expression in normal human bronchial epithelial cells and in NCI-H292 cells. It was also reported that synthetic dsRNA upregulated expression of MUC5AC in differentiated primary human tracheobronchial epithelial cells and in NCI-H292 cells and that viral infection induced mucin production through TLR-3 pathway (Zhu et al. 2009). One of the major symptoms of DPB is mucus hypersecretion, therefore it is intriguing if new mucin-like genes in the candidate region had any role in mucus hypersecretion in airway of DPB. G4 were membrane-tethered glycoproteins, and clarification is needed as to whether they could be secreted by alternate splicing like MUC4 (Moniaux et al. 2000), or by protein cleavage (Davies et al. 2002). Recently, we found an insertion/deletion polymorphism in the promoter region of MUC5B gene, which was significantly associated with DPB (Kamio et al. 2005). Mucus hypersecretion in DPB may be explained partly by MUC5B, but other mucin genes may also play a role. G4 gene expression is remarkably upregulated by poly(I:C) and otherwise its expression level is low; inflammatory condition mediated by TLR3 may be needed for the induction of G4 expression. It is not known if viral infection has a role in DPB. More investigations are needed to reveal the role of mucin genes including G4 in DPB.

Many genetic polymorphisms including allelic length variation of tandem repeat region are known in mucin genes (Fowler et al. 2001). The new mucin-like genes were also highly polymorphic, and some of the polymorphisms in G4 and G2 gene regions were associated with DPB. In G2 gene, four-nucleotide insertion which causes frameshift and premature stop codon was associated with DPB, but the significance was marginal. The putative protein from the insertion allele will be shorter than that from the deletion allele by 89 amino acids. In G4 gene, A-allele of the SNP in intron 2 was positively associated with the disease, and the shortest VNTR variant was negatively associated. A-allele of intron 2 SNP was borne by the major

disease-associated haplotypes in Japan (Keicho et al. 2000), HLA-B54-A24. However, its relation with the other two disease-associated haplotypes, HLA-B54-A11 and HLA-B*5504-A11, was not clear in the present study because these two haplotypes were much less frequent among the Japanese. The A-allele in intron 2 might not explain the disease association of all three haplotypes, but our present study was restricted in finding a solution to the problem: The study population was small, mainly because DPB is categorized as a rare disease even among Asians. Because of this small sample size and necessity of multiple comparisons, all associations in this study were not robust enough. The disease association of the intron SNP and VNTR should be tested in other populations such as Koreans and Chinese. It is possible that the polymorphisms in G4 gene cause some functional differences in the mucin-like glycoprotein or in its expression; however, further investigations are needed. We designated G4 gene as panbronchiolitis related mucin-like 1 (*PBMUCL1*) and G2 gene as *PBMUCL2*.

The polymorphisms in the new gene showed both the *P* values and odds ratios which could not reach those of *HLA-B*. An unidentified polymorphism which is in linkage disequilibirium with HLA-B54 in the Japanese and with HLA-A11 in Koreans may be more strongly associated with DPB. Search for such polymorphism should be continued, and more studies are needed to reveal the function of these mucin-like genes clustered in 6p21.3 and its difference due to genetic polymorphisms. Further collaborative studies in Asian countries are needed to investigate the genetic polymorphisms and a potential role of the new mucin-like genes in pathogenesis of DPB.

We declare that these experiments comply with the current laws of Japan.

Conflict of interest None declared.

References

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Chu YC, Yeh SZ, Chen CL, Chen CY, Chang CY, Chiang CD (1992) Diffuse panbronchiolitis: report of a case. J Formos Med Assoc 91:912–915

Davies JR, Herrmann A, Russell W, Svitacheva N, Wickström C, Carlstedt I (2002) Respiratory tract mucins: structure and expression patterns. Novartis Found Symp 248:76–88

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Fowler J, Vinall L, Swallow D (2001) Polymorphism of the human muc genes. Front Biosci 6:D1207–D1215

Fraser RG, Pare PD, Fraser RS, Genereux GP (1990) Diagnosis of diseases of the chest, vol 3. W.B. Saunders, Philadelphia, pp 2224–2229

Gray TE, Guzman K, Davis CW, Abdullah LH, Nettesheim P (1996) Mucociliary differentiation of serially passaged normal human tracheobronchial epithelial cells. Am J Respir Cell Mol Biol 14:104–112

Hirokawa T, Boon-Chieng S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378–379

Homma H, Yamanaka A, Tanimoto S, Tamura M, Chijimatsu Y, Kira S, Izumi T (1983) Diffuse panbronchiolitis. A disease of the transitional zone of the lung. Chest 83:63–69

Imanishi T, Akaza T, Kimura A, Tokunaga K, Gojobori T (1991) Estimation of allele and haplotype frequencies for HLA and complement loci. In: Tsuji K, Aizawa M, Sasazuki T (eds) HLA, vol 1. Oxford University Press, New York, pp 76–79

Itoh Y, Kamata-Sakurai M, Denda-Nagai K, Nagai S, Tsuiji M, Ishii-Schrade K, Okada K, Goto A, Fukuyama M, Irimura T (2008) Identification and expression of human epiglycanin/MUC21:a novel transmembrane mucin. Glycobiology 18:74–83

Kamio K, Matsushita I, Hijikata M, Kobashi Y, Tanaka G, Nakata K, Ishida T, Tokunaga K, Taguchi Y, Homma S, Nakata K, Azuma A, Kudoh S, Keicho N (2005) Promoter analysis and aberrant expression of the MUC5B gene in diffuse panbronchiolitis. Am J Respir Crit Care Med 171:949–957

Keicho N, Tokunaga K, Nakata K, Taguchi Y, Azuma A, Bannai M, Emi M, Ohishi N, Yazaki Y, Kudoh S (1998) Contribution of HLA genes to genetic predisposition in diffuse panbronchiolitis. Am J Respir Crit Care Med 158:846–850

Keicho N, Ohashi J, Tamiya G, Nakata K, Taguchi Y, Azuma A, Ohishi N, Emi M, Park MH, Inoko H, Tokunaga K, Kudoh S (2000) Fine localization of a major disease-susceptibility locus for diffuse panbronchiolitis. Am J Hum Genet 66:501–507

Kim YW, Han SK, Shim YS, Kim KY, Han YC, Seo JW, Im JG (1992) The first report of diffuse panbronchiolitis in Korea: five case reports. Intern Med 31:695–701

Lang T, Hansson GC, Samuelsson T (2007) Gel-forming mucins appeared early in metazoan evolution. Proc Natl Acad Sci USA 104:16209–16214

Lee KW, Oh DH, Lee C, Yang SY (2005) Allelic and haplotypic diversity of HLA-A, -B, -C, and–DQB1 genes in the Korean population. Tissue Antigens 65:437–447

Londhe V, McNamara N, Lemjabbar H, Basbaum C (2003) Viral dsRNA activates mucin transcription in airway epithelial cells. FEBS Lett 553:33–38

Matsuzaka Y, Tounai K, Denda A, Tomizawa M, Makino S, Okamoto K, Keicho N, Oka A, Kulski JK, Tamiya G, Inoko H (2002) Identification of novel candidate genes in the diffuse panbronchiolitis critical region of the class I human MHC. Immunogenetics 54:301–309

Moniaux N, Escande F, Batra SK, Porchet N, Laine A, Aubert JP (2000) Alternative splicing generates a family of putative secreted and membrane-associated MUC4 mucins. Eur J Biochem 267:4536–4544

Park MH, Kim YW, Yoon HI, Yoo CG, Han SK, Shim YS, Kim WD (1999) Association of HLA class I antigens with diffuse panbronchiolitis in Korean patients. Am J Respir Crit Care Med 159:526–529

Rose MC, Voynow JA (2006) Respiratory tract mucin genes and mucin glycoproteins in health and disease. Physiol Rev 86:245–278

Ross AJ, Dailey LA, Brighton LE, Devlin RB (2007) Transcriptional profiling of mucociliary differentiation in human airway epithelial cells. Am J Respir Cell Mol Biol 37:169–185

Saito S, Ota S, Yamada E, Inoko H, Ota M (2000) Allele frequencies and haplotypic associations defined by allelic DNA typing at HLA class I and class II loci in the Japanese population. Tissue Antigens 56:522–529

Sharma P, Dudus L, Nielsen PA, Clausen H, Yankaskas JR, Hollingsworth MA, Engelhardt JF (1998) MUC5B and MUC7 are differentially expressed in mucous and serous cells of submucosal glands in human bronchial airways. Am J Respir Cell Mol Biol 19:30–37

She J, Sun Q, Fan L, Qin H, Bai C, Shen C (2007) Association of HLA genes with diffuse panbronchiolitis in Chinese patients. Respir Physiol Neurobiol 157:366–373

Sugiyama Y, Kudoh S, Maeda H, Suzaki H, Takaku F (1990) Analysis of HLA antigens in patients with diffuse panbronchiolitis. Am Rev Respir Dis 141:1459–1462

Tadaki H, Arakawa H, Mizuno T, Suzuki T, Takeyama K, Mochizuki H, Tokuyama K, Yokota S, Morikawa A (2009) Double-stranded RNA and TGF-alpha promote MUC5AC induction in respiratory cells. J Immunol 182:293–300

The International HapMap Consortium (2003) The International HapMap Project. Nature 426:789–796

Tsang KW, Ooi CG, Ip MS, Lam WK, Ngan H, Chan EY, Hawkins B, Ho CS, Amitani R, Tanaka E, Itoh H (1998) Clinical profiles of Chinese patients with diffuse panbronchiolitis. Thorax 53:274–280

Voynow JA, Gendler SJ, Rose MC (2006) Regulation of mucin genes in chronic inflammatory airway diseases. Am J Respir Cell Mol Biol 34:661–665

Zhu L, Lee PK, Lee WM, Zhao Y, Yu D, Chen Y (2009) Rhinovirus-induced major airway mucin production involves a novel TLR3-EGFR-dependent pathway. Am J Respir Cell Mol Biol 40:610–619