Identification of *trans* Protein QTL for Secreted Airway Mucins in Mice and a Causal Role for *Bpifb1*

Lauren J. Donoghue,*,†,1 Alessandra Livraghi-Butrico,‡,1 Kathryn M. McFadden,* Joseph M. Thomas,* Gang Chen,‡ Barbara R. Grubb,‡ Wanda K. O'Neal,‡ Richard C. Boucher,‡ and Samir N. P. Kelada*,†,‡,2

*Department of Genetics, †Curriculum in Genetics and Molecular Biology, and ‡Marsico Lung Institute, University of North Carolina, Chapel Hill, North Carolina 27599

ORCID ID: 0000-0003-2676-9232 (S.N.P.K.)

ABSTRACT Mucus hyper-secretion is a hallmark feature of asthma and other muco-obstructive airway diseases. The mucin proteins MUC5AC and MUC5B are the major glycoprotein components of mucus and have critical roles in airway defense. Despite the biomedical importance of these two proteins, the loci that regulate them in the context of natural genetic variation have not been studied. To identify genes that underlie variation in airway mucin levels, we performed genetic analyses in founder strains and incipient lines of the Collaborative Cross (CC) in a house dust mite mouse model of asthma. CC founder strains exhibited significant differences in MUC5AC and MUC5B, providing evidence of heritability. Analysis of gene and protein expression of *Muc5ac* and *Muc5b* in incipient CC lines (n = 154) suggested that post-transcriptional events were important regulators of mucin protein content in the airways. Quantitative trait locus (QTL) mapping identified distinct, *trans* protein QTL for MUC5AC (chromosome 13) and MUC5B (chromosome 2). These two QTL explained 18 and 20% of phenotypic variance, respectively. Examination of the MUC5B QTL allele effects and subsequent phylogenetic analysis allowed us to narrow the MUC5B QTL and identify *Bpifb1* as a candidate gene. *Bpifb1* mRNA and protein expression were upregulated in parallel to MUC5B after allergen challenge, and *Bpifb1* knockout mice exhibited higher MUC5B expression. Thus, BPIFB1 is a novel regulator of MUC5B.

KEYWORDS asthma; mucins; mucus; MUC5AC; MUC5B; quantitative trait locus (QTL)

EXCESSIVE mucus in the airways is a hallmark feature of muco-obstructive airway diseases including asthma, chronic obstructive pulmonary disease (COPD), and cystic fibrosis (CF) (Fahy and Dickey 2010). MUC5AC and MUC5B are the primary secreted mucins in the airways (Evans *et al.* 2009) and alterations in the levels and/or post-translational modifications of each mucin, as well as the ratio between the two, are associated with different airway diseases (Kirkham *et al.* 2002; Fahy and Dickey 2010). MUC5AC is increased at the mRNA and protein levels in the airways of asthmatics (Ordoñez *et al.* 2001) and plays an important role in airway narrowing and hyper-responsiveness (Evans *et al.* 2015). MUC5B is significantly elevated in COPD and CF (Fahy and Dickey 2010), as well as idiopathic pulmonary fibrosis (IPF) (Seibold *et al.* 2011), and is the primary mucin in plugs removed from the lungs after fatal status asthmaticus (Sheehan *et al.* 1995; Groneberg *et al.* 2002). In addition to associations with disease, MUC5B regulates features of lung homeostasis, such as mucociliary clearance, suggesting it has a critical role in lung defense to pathogens (Roy *et al.* 2013).

While there has been a great deal of research on the biochemical pathways that regulate mucin genes and corresponding proteins (Rose and Voynow 2006; Davis and Dickey 2008; Thai *et al.* 2008), the genetic architecture of mucus phenotypes in both humans and animal models is not well characterized. Previous studies provided evidence that the chronic mucus hyper-secretion phenotype in COPD (also known as chronic bronchitis) has a familial component (Viegi *et al.* 1994; Silverman *et al.* 1998), and results from a twin study indicated that ~40% of variation in the chronic mucus hyper-secretion phenotype is attributable to genetic variation (Hallberg *et al.* 2008). Genome-wide association

doi: https://doi.org/10.1534/genetics.117.300211
Manuscript received May 19, 2017; accepted for publication August 22, 2017; published Early Online August 28, 2017.
Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.117.300211/-/DC1.

[1]These authors contributed equally to this work.

[2]Corresponding author: University of North Carolina, Department of Genetics, 120 Mason Farm Rd., Genetic Medicine Bldg., Room 5072, Chapel Hill, NC 27599. E-mail: samir_kelada@med.unc.edu

studies have identified two SNPs (rs6577641 in *SATB1* and rs34391416 near the genes *EFCAB4A*, *CHID1*, and *AP2A2*) associated with chronic mucus hyper-secretion in the context of COPD (Dijkstra *et al.* 2014; Lee *et al.* 2014), but these loci do not explain a large fraction of the estimated heritability. This could in part be due to the need to rely on survey data to ascertain phenotype status in these studies, which dichotomizes the quantitative nature of muco-obstructive phenotypes and is subject to recall bias, resulting in reduced power to detect significant associations.

Given the importance of MUC5AC and MUC5B in muco-obstructive airway diseases, we sought to identify key regulators of these secreted mucins using an unbiased, genome-wide search. To accomplish this goal, we utilized incipient lines of the Collaborative Cross (CC). The CC is composed of a panel of recombinant inbred mouse lines derived from eight-way crosses using five classical inbred strains (C57BL/6J, 129S1/SvImJ, A/J, NOD/ShiLtJ, and NZO/H1LtJ) and three wild-derived inbred strains (WSB/EiJ, PWK/PhJ, and CAST/EiJ) (Collaborative Cross Consortium 2012). Studies with incipient and established CC lines have identified quantitative trait loci for numerous traits of biomedical interest (Aylor *et al.* 2011; Kelada *et al.* 2012, 2014; Rutledge *et al.* 2014; Gralinski *et al.* 2015; Mosedale *et al.* 2017; Venkatratnam *et al.* 2017). Within this framework, we applied a house dust mite (HDM) model (Kelada *et al.* 2011) to elicit MUC5AC and MUC5B expression and secretion in mouse airways. We used quantitative protein measurements of secreted mucins, QTL mapping, gene expression, and genome sequence data to identify candidate regulators of these secreted mucins. Further, we interrogated the relationship between MUC5B and the candidate regulator *Bpifb1* with gene and protein expression data, followed by the use of a knockout strain to validate BPIFB1 as a novel regulator of MUC5B in the airways.

Materials and Methods

Mice

We obtained 154 male pre-CC mice (ages 10–14 weeks) from Oak Ridge National Laboratory (Chesler *et al.* 2008). Each mouse was from an independent CC line that had undergone 5–14 generations of inbreeding. We also obtained male mice of each of the eight CC founder strains from The Jackson Laboratory (n = 3–4/strain were phenotyped, plus additional C57BL/6J mice). All mice were singly housed, with alpha-dri bedding, under normal 12-hr light/dark cycles. All experiments conducted with mice in this study were compliant with an Institutional Animal Care and Use Committee protocol at an animal facility approved and accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International.

To generate *Bpifb1* KO mice, sperm from hemizygous male *Bpifb1* mice was obtained from the Knockout Mouse Project (KOMP) (www.komp.org). These mice were created from embryonic stem cell clone EPD0704_3_A09 obtained from the KOMP Repository and generated by the Wellcome Trust Sanger Institute. Methods used on the targeted alleles have been published (Skarnes *et al.* 2011). In this work, the *Tm1a* allele was used without subsequent modification or removal of cassette elements. Sperm was used to fertilize eggs from C57BL/6NJ mice, and heterozygous offspring were mated to generate a colony of mice used for the experiment.

Mice were genotyped using a single PCR reaction containing two different forward primers [wild-type (WT) allele primer: 5′-GAAGAGGGAAGGTCAGGAGAAAGGG-3′ and knockout (CSD-neoF) allele primer 5′-GGGATCTCATGCTGGAGTTCT TCG-3′] and one common reverse primer (5′-TAAGGTCC TTTGGTCAACGGTTTGC-3′). PCR products were analyzed by gel electrophoresis, with the WT allele producing a 535-bp product and the knockout allele producing a 646-bp product.

By and large, matings of heterozygous mice produced Mendelian ratios of offspring. However, in the third generation we observed transmission ratio distortion with a depletion of homozygous mutant pups (four born, nine expected), suggesting that either loss of *Bpifb1* has some effect on genotype ratios that is incompletely penetrant or that an additional mutation on the strain background contributed to non-Mendelian ratios. This effect was not observed in three subsequent generations.

Homozygous WT and knockout littermates (males and females) were used in experiments in which mice were either used at baseline (*i.e.*, naïve, no allergen exposure) or sensitized twice (i.p.) with 10 μg of Der p 1 and challenged with PBS or 50 μg of Der p 1. Within each genotype, mice were randomized to treatment groups.

Phenotyping protocol

We employed an HDM model of allergic airway disease that produces hallmark asthma phenotypes including T_h2-biased airway inflammation, elevated serum IgE, mucous cell metaplasia, and airway hyper-responsiveness in a strain-dependent fashion (Kelada *et al.* 2011). Results for inflammatory cell recruitment [eosinophils (Kelada *et al.* 2014) and neutrophils (Rutledge *et al.* 2014)], responsiveness to methacholine (Kelada 2016), and lung gene expression (eQTL) (Kelada *et al.* 2014) have been previously published. At the start of our study, mice were phenotyped by whole body plethysmography for baseline responsiveness to methacholine (on days 1 and 8). The mice were then sensitized with 10 μg (in 100 μl) of the immunodominant allergen from the *Dermatophagoides pteronyssinus* species of HDM, Der p 1 (Cat. No. LTN-DP1-1; Indoor Biotechnology), administered by intraperitoneal injection on days 11 and 18 of the protocol. On day 25, mice were challenged with 50 μg of Der p 1 (in 100 μl), administered by oro-pharyngeal aspiration. Mice were phenotyped for responsiveness to methacholine again by whole body plethysmography on day 28. Immediately following this measurement, mice were killed by an overdose of pentobarbitol, followed by collection of blood, whole lung lavage fluid, and lung tissue. Lung lavage was performed using two successive volumes of 0.5 and 1.0 ml PBS. The first

fraction of lavage fluid was spun down (500 g for 10 min) to separate out cells and the supernatant was retained for analysis of mucins by agarose gel electrophoresis followed by immunoblotting.

Agarose gel electrophoresis and immunoblotting

We used a previously established method (Kirkham *et al.* 2002) involving agarose gel electrophoresis following by immunoblotting to detect MUC5AC and MUC5B in bronchoalveolar lavage fluid (BALF). Briefly, the addition of urea powder (40 mg/0.1 ml) to BALF produced a 6 M urea BALF, which was reduced with 10 mM dithiothreitol for 90 min and alkylated with 25 mM iodoacetamide to obtain constituent mucin monomer subunits. Equal volumes of reduced samples (40 μl) were run on 1% agarose gel. Gels were vacuum-blotted onto nitrocellulose membranes, blocked with Odyssey blocking buffer (LI-COR Biosciences), and probed with a rabbit polyclonal antibody (Ab) raised against murine MUC5AC or MUC5B (Ehre *et al.* 2012) [UNC294 and UNC222, respectively, generously provided by Dr. Camille Ehre (UNC)]. Using sequence data (Keane *et al.* 2011), we determined that there were no missense variants encoded in *Muc5ac* or *Muc5b* at the epitopes recognized by these antibodies, thus each antibody-based measurement should have been unbiased with respect to strain. The secondary Ab was Alexa Fluor 680 goat anti-rabbit IgG, diluted 1/15,000 in Odyssey blocking buffer. Detection and analysis of specific signals were performed using the Odyssey CLx infrared imaging system (LI-COR Biosciences). For quantification, the integrated intensity of the major high-molecular-weight bands in each lane was measured by defining a rectangular region of interest (ROI) encompassing these bands in the sample with the strongest signal and using this ROI for all lanes (as shown in Supplemental Material, Figure S1, Odyssey software version 3.0 or ImageStudio version 1.0.11). The median intensity of three pixels on the right and the left of the defined ROI was used to define the background, which was subtracted from the integrated intensity value. Owing to the large number of pre-CC samples, they had to be run on five different gels. To normalize the integrated intensity values for transfer efficiency and antibody binding, the same volume of two different positive control samples (BALF from C57BL/6N mice treated with IL-13) was loaded on each row of every gel (as illustrated in Figure S1), and their integrated intensity values were used to normalize the values for the pre-CC samples. BALF from *Muc5ac* KO and *Muc5b* KO mice were used as controls to confirm antibody specificity as reported in Livraghi-Butrico *et al.* (2017). Phenotype data for CC founder strains and pre-CC mice are given in Table S1 and Table S2, respectively.

For statistical analyses of pre-CC mice, MUC5AC and MUC5B intensity values were natural log-transformed. For MUC5AC, we observed a spike in the distribution at the low end of the range, suggesting that these values were near the limit of detection, and we removed these data points (n = 10 mice) from further analysis.

Heritability estimates

Analysis of variance was used to test whether phenotypes varied among CC founder strains. From these ANOVA models, we estimated broad-sense heritability (H^2) by calculating the interclass correlation (r_1) and the coefficient of genetic determination (g^2) (Festing 1979):

$$r_1 = (\text{MSB-MSW})/(\text{MSB} + (n\text{-}1)\text{MSW}),$$

$$g^2 = (\text{MSB-MSW})/(\text{MSB} + (2n\text{-}1)\text{MSW}),$$

where MSB and MSW are the mean squares between and within, respectively, from the eight-way ANOVA model described above and n is the number of mice per strain (n = 3–4/strain).

Genotyping, quantitative trait locus (QTL) mapping, and phylogeny

We genotyped each mouse at the University of North Carolina at Chapel Hill using one of two Affymetrix SNP arrays (A or B) that were produced during the development of the Mouse Diversity array (Yang *et al.* 2009), resulting in 27,059 mapping intervals, as described in a previous study (Rutledge *et al.* 2014). These genotyping arrays were annotated to NCBI Build 36 of the mouse genome, but we mapped QTL boundaries to Build 37 positions and we report Build 37 positions in our results. We used BAGPIPE (Valdar *et al.* 2009) to fit a regression model and report LOD scores for QTL mapping of mucin protein levels. For MUC5AC, QTL mapping was carried out using the data set after removal of values below the apparent threshold of detection. Inclusion of values below the threshold resulted in no significant QTL. Significance thresholds were determined by permutation (n = 10,000). We used the 1.5 LOD drop method to approximate confidence interval for QTL (Dupuis and Siegmund 1999). The proportion of variance explained for each QTL was estimated by regression of phenotypes on genotype for the peak locus for each QTL. Allele effects plots were used to identify putative functional haplotype groupings using a previously established method (Kelada *et al.* 2012). Post hoc statistical tests of these haplotypes were conducted using Tukey's honestly significant difference test. To build a phylogenetic tree relating the eight CC founder strains at the Chr 2 QTL, we generated sequence alignments using sequence data (http://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1211) (Keane *et al.* 2011) and then generated a maximum likelihood tree using RAxML version 7.2.9 (Stamatakis 2006). One hundred bootstrap samples were used to assess the support for each tree branch. Please see File S1 for code used to run RAxML.

Candidate gene identification

To identify candidate genes for the MUC5B QTL on Chr 2, we applied a weight of evidence approach based on a series of criteria. This included the following criteria:

1. Evidence that the gene is expressed in airway epithelia, based on expression data from a previous study (GSE34764) (Doherty *et al.* 2012). Specifically, we used samples from the

control group (PBS, $n = 3$) of GSE34764 to characterize the distribution of gene expression and then categorized genes in the bottom fifth percentile as not expressed while genes above or equal to the fifth percentile were considered expressed.
2. Evidence that the gene is differentially expressed in whole lung RNA samples as a function of allergen treatment, based on a previous study we conducted (GSE19223) (Kelada *et al.* 2011). For this criterion, we conducted an analysis of differential expression using Partek software (version 6.6) and used nominal *P*-values <0.05 and fold change >1.5.
3. Evidence that the gene is differentially expressed in normal lung tissue as a function of Chr 2 founder haplotype, which was assessed using two independent gene expression data sets involving strains from different subspecies [A/J (*musculus* haplotype) and C57BL/6J (*domesticus* haplotype) in GSE30015 and WSB/EiJ (*domesticus* haplotype) and PWK/PhJ (*musculus* haplotype) in GSE44555] (Crowley *et al.* 2015). This analysis was conducted using R (version 3.1.2) using the MAANOVA package after masking for probesets that contain variants from *musculus* haplotypes. For this criterion, we did not specify a fold change cutoff, only a nominal statistical significance threshold of <0.05. However, we did specify that the direction of the gene expression difference (either higher or lower) be consistent across the two studies.
4. Unique sharing of nonsynonymous variants among the Chr 2 *musculus* haplotypes. We used sequence data (Keane *et al.* 2011) and a corresponding web query tool (http://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1211) to identify nonsynonymous variants that are shared by A/J, NOD/ShiLtJ, and PWK/PhJ strains and not present in other CC founder strains.
5. Prior evidence of biological relevance. We conducted literature searches in PubMed and used the Mouse Genome Informatics website to find previously published evidence that a particular gene may be relevant to allergic airway inflammation and/or mucus production/secretion.

Gene expression by qRT-PCR

RNA from the right lung middle lobe was isolated using the RNEasy Kit (Qiagen, Valencia, CA) and reverse transcribed using the High Capacity RNA to cDNA kit (Applied Biosystems, Foster City, CA) according to manufacturers' protocols. TaqMan assays for *Muc5b* and *Bpifb1* were custom designed to avoid genetic variants at these loci in the CC founder strain genomes. Assays were validated for efficient amplification using serial dilutions of template. Primer and probe pairs are as follows: *Muc5b* F: 5′-AAGGTGTATGTACCAACTGG-3′, R: 5′-TGTACTATCCTGGGTCCTC-3′, Probe: 5′-TGTGACCTCTCCTGTCCACCCAC-3′; *Bpifb1* F: 5′-CACAAGCTCTCCTTCGTG-3′, R: 5′-TACATGTCATCAAAGGCCTG-3′, Probe: 5′-TGAATCTCCTGGTGCCAGCCCTG-3′. ΔCt values were calculated by normalizing to *Rps20* (Mm02342828_g1, Applied Biosystems). A two-sided or one-sided Student's *t*-test was used to analyze the data as noted, and a significant change was defined as $\alpha = 0.05$.

Histology and immunohistochemistry

Formalin-fixed left lung lobes were cut in cross-section starting at the hilum and 2 mm apart along the main stem bronchus, yielding three slices/lung (level 1, 2, and 3 from rostral to caudal). All three slices were embedded in paraffin, and 5-μm-thick sections were cut and stained with AB-PAS. Immunohistochemistry for MUC5B and BPIFB1 was performed as previously described (Livraghi-Butrico *et al.* 2017). Briefly, sections were incubated for 1 hr with FC Block (BD Pharmigen) at 1:200 in 5% normal goat serum (NGS) (JacksonImmuno) in phosphate buffered saline with 0.1% Tween-20 and for 3 hr with 1:1000 MUC5B (H-300, Santa Cruz Biotech) or 1:200 BPIFB1 (HP8044, Hycult Biotech) primary antibodies in 5% NGS. Rabbit IgG (JacksonImmuno) was used as isotype control for both antibodies. Goat anti-rabbit biotinylated secondary antibody (JacksonImmuno) was applied on all sections at 1:200 in 5% NGS for 30 min. The VECTSTAIN Elite ABC R.T.U and Immpact NovaRed horseradish peroxidase detection systems (Vector Laboratories, Burlingame, CA) were used according to the manufacturer's protocols. Slides were counterstained with Mayer's hematoxylin before dehydration and mounting with DPX (Sigma Aldrich). Slides were imaged on an Olympus BX60F5 microscope with CellSens Standard software (Olympus Life Sciences version 1.13).

Morphometric analysis

Total airway or luminal histological staining was quantified morphometrically on level 1 (the level most affected by mucous cell metaplasia and mucus accumulation) by investigators blind to treatment and genotype. Total, intraepithelial (stored), and luminal AB-PAS positive and MUC5B or BPIFB1-positive volume densities (Vs) were quantified using 10× or 20× magnification micrographs as previously described (Harkema *et al.* 1987).

Meta-analysis of gene expression data

Eight gene expression data from mouse models of asthma were identified in NCBI's GEO database (GSE1301, GSE13032, GSE13382, GSE1438, GSE19223, GSE27066, GSE34764, and GSE6858). In all but one case (GSE1438), data were downloaded then normalized (using the default robust multi-array average with quantile normalization method) using Partek software (version 6.6). Statistical analysis of differential gene expression was also conducted in Partek, accounting for strain if >1 strain was included in the study. For GSE1438, we used NCBI's GEO2R web-based software to conduct data analysis because of the custom arrays used in that experiment. Meta-analysis of gene expression data from NCBI's GEO repository was conducted using Fisher's combined probability test.

Data availability

Supplemental Material is provided in File S1, a zip file containing perl scripts used in QTL mapping and code use for phylogenetic analysis. Phenotype data are provided in Table S1 and Table S2. Genotype calls for pre-CC mice are available at http://www.genetics.org/content/198/2/735.supplemental.

Results

To identify regulators of secreted MUC5AC and MUC5B in the airway, we applied an HDM model of asthma (Kelada *et al.* 2011) involving the immunodominant antigen Der p 1. In preliminary experiments using C57BL/6J mice, we quantified secreted MUC5AC and MUC5B levels in BALF after Der p 1 allergen sensitization and challenge using agarose gel electrophoresis followed by immunoblotting. These experiments revealed that the vast majority of MUC5AC and MUC5B were contained in the soluble fraction of BALF and that allergen treatment increased MUC5AC 10-fold and MUC5B 6-fold (Figure S2). We then phenotyped CC founder strains for MUC5AC and MUC5B BALF levels after allergen sensitization and challenge, and we included a subset of A/J mice that were sensitized but challenged only with PBS (vehicle). As shown in Figure 1, levels of MUC5AC and MUC5B in BALF varied significantly by strain. For both mucins, the effect of strain was greater than the effect of treatment alone in the A/J strain. Estimates of broad-sense heritability for MUC5AC among allergen sensitized and challenged mice based on interclass correlations (r_1) and the coefficient of genetic determination (g^2) were 0.77 and 0.63, respectively. For MUC5B, the corresponding heritability estimates were 0.69 and 0.53. These results indicate that variation in the levels of both mucins is significantly associated with variation in the genome.

We then phenotyped 154 incipient lines of the CC, which we refer to as "pre-CC" mice. Each pre-CC mouse was derived from a different CC breeding funnel and contained allelic contributions from all eight founder strains (Aylor *et al.* 2011). Pre-CC mice were treated with Der p 1 allergen and exhibited a wide range of MUC5AC and MUC5B levels in BALF (Figure S3). We examined pairwise correlations between MUC5AC and MUC5B and found a weak but statistically significant and positive correlation between the two mucin proteins (Figure S4; $r = 0.25$, $P = 0.005$), where the correlation was strongest toward the higher ranges of MUC5AC and MUC5B expression. In contrast, pairwise correlations of *Muc5ac* and *Muc5b* mRNA [assessed using previously generated data (Kelada *et al.* 2014)] were quite strong (Figure S4; $r = 0.83$, $P < 1 \times 10^{-4}$). Interestingly, pairwise correlations between each mucin mRNA and protein were weak or absent, suggesting post-transcriptional events that uncouple intracellular mRNA and secreted protein levels.

To identify the specific genetic loci that control mucin levels in response to allergen sensitization and challenge, we used the pre-CC phenotype and genotype data to map QTL for each mucin protein (Figure 2). For MUC5AC, we identified a QTL on Chr 13 (*Muc5acq1*, LOD = 7.2, Figure 2A), spanning 71.6–81.0 Mb, and for MUC5B, we identified a QTL on Chr 2 (*Muc5bq1*, LOD = 7.9, Figure 2B), spanning 152.5–159.8 Mb. These two QTL explained 18 and 20% of phenotype variation, respectively. The genes encoding both mucins are located on Chr 7, indicating these two pQTL act in *trans*, and the fact that the two *trans*-pQTL are different indicates differential regulation of each mucin.

We then sought to identify candidate genes for each QTL, initially using the QTL allele effects pattern as a way to identify potentially functionally distinct alleles at each locus and prioritize genes (Kelada *et al.* 2012). For *Muc5acq1*, there were no obvious differences in phenotype among pre-CC mice categorized by most probable CC founder haplotype (Figure S5). As such, we turned to other data to identify potential candidate genes for *Muc5acq1*. This locus contains 47 protein-coding genes and 50 noncoding RNAs. Expression data were available for 45 of the 47 protein-coding genes (based on GSE51768), and we assessed pairwise correlation of each gene with BALF MUC5AC. Eight genes were associated with MUC5AC at a nominal *P*-value threshold of 0.05 (Table S3). We also asked whether any of these genes had local eQTL using a previously generated eQTL data set from the same set of mice (Kelada *et al.* 2014). Fifteen genes of the 47 had eQTL (Table S3). Finally, we conducted a literature search to see if any of the 47 genes had potential links to MUC5AC. In aggregate, no single gene stood out in terms of pairwise correlation and eQTL; however, evidence for biological plausibility for four genes (*Slc9a3*, *Nr2f1*, *Glrx1*, and *Spata9*) was obtained from the literature search. As such, these genes may be worthy of additional investigation in relation to MUC5AC.

In contrast to the allele effects for *Muc5acq1*, the allele effects for *Muc5bq1* were striking. We found that alleles from A/J, NOD/ShiLtJ, and PWK/PhJ were associated with significantly lower levels of MUC5B (Figure 2C), indicating that these three alleles share functional similarities that result in low MUC5B BALF levels. Visualization of founder haplotypes based on high-density genotyping (Wang *et al.* 2012) in this region revealed that these three strains all share a *Mus musculus musculus*-derived haplotype spanning ~152–154 Mb (Figure S6), indicating intersubspecific introgression (Yang *et al.* 2011) of *musculus* haplotypes into the genomes of the A/J and NOD/ShiLtJ strains (classical inbred strains of *domesticus* origin). Using unbiased sequence data (Keane *et al.* 2011), we constructed a phylogenetic tree for this region and confirmed that these three strains have *musculus*-derived alleles that are distinct from all other CC founder strains (Figure 2D). This result indicates (based on parsimony) that the *musculus* haplotypes contain a functional variant (or variants) that cause lower MUC5B.

To prioritize candidate genes underlying the *Muc5bq1*, we could not rely on gene expression data for genes in this region because the microarray platform we used (Kelada *et al.* 2014) was based on C57BL/6J sequence and therefore was biased against *M. m. musculus* and *castaneus* transcripts emanating from *Muc5bq1*. We therefore chose to leverage other data sets to prioritize candidate genes. As described in the methods, we applied a weight of evidence approach to the gene list to prioritize genes based on expression in a relevant cell type (airway epithelia), change in expression in the lung due to allergen treatment, differences in expression in the lung due to subspecies origin, the presence of predicted amino acid variants shared by A/J, NOD/ShiLtJ, and PWK/PhJ founder

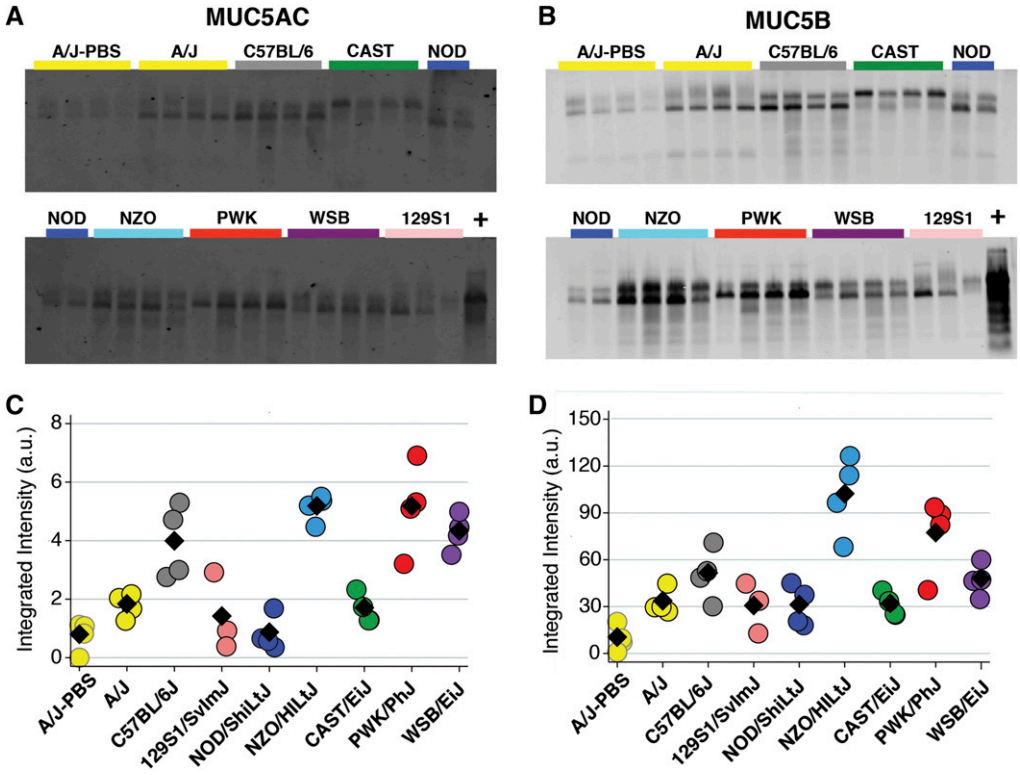

Figure 1 MUC5AC and MUC5B expression in CC founder strains. Immunoblots for (A) MUC5AC and (B) MUC5B in BALF of CC founder strains sensitized and challenged with Der p 1 (A/J challenged with PBS or Der p 1), with n = 3–4 mice/strain. + denotes BALF sample from a C57BL/6J mouse treated with allergen in a separate experiment and used as positive control for antibody reactivity. Quantification of protein signal (integrated intensity a.u.) for (C) MUC5AC and (D) MUC5B show significant effect of strain as calculated by ANOVA for MUC5AC (P strain = 4.0×10^{-7}) and MUC5B (P strain = 1.1×10^{-5}). Each circle represents one mouse and the diamonds indicate strain means.

strains, and previously published papers documenting a role for a gene in relevant biological pathways (Figure S7). This approach identified *Bpifb1* as a priority candidate gene because it met all of the aforementioned criteria.

Bpifb1 is a member of the BPI (bactericidal/permeability-increasing protein)/LBP (lipopolysaccharide-binding protein) family of lipid-transfer proteins, and was previously known as *Lplunc1* for long palate lung and nasal epithelia clone (Bingle *et al.* 2011). Previous studies in humans documented expression of BPIFB1 in goblet cells of the respiratory epithelium (Bingle *et al.* 2010), secretion into the airways (Bingle *et al.* 2010; Gao *et al.* 2015), and upregulation in CF (Bingle *et al.* 2012) and COPD (Gao *et al.* 2015; Titz *et al.* 2015). In contrast, BPIFA1 is down-regulated in asthma models (Chu *et al.* 2007; Wu *et al.* 2017).

To begin evaluating a potential role of *Bpifb1* in MUC5B regulation, we examined lung gene expression in pre-CC mice followed by gene and protein expression in control and allergen-treated inbred strains. For the pre-CC mice, we performed qRT-PCR with gene expression probes designed to avoid genetic variation in the CC founder genomes. We found that *Muc5b* and *Bpifb1* mRNA were strongly and positively correlated among pre-CC mice ($r = 0.75$, $P < 1 \times 10^{-4}$, Figure S8). This result prompted us to ask whether *Muc5b* and *Bpifb1* were jointly upregulated in other mouse models of asthma. As shown in Table 1, we found that expression of both *Muc5b* and *Bpifb1* were nearly always increased in response to allergens such as ovalbumin, *Alternaria*, and HDM across three different inbred strains. Meta-analysis confirmed significant differential expression, with meta-analysis P-values of 1.5×10^{-13} and 1.8×10^{-10} for *Muc5b* and *Bpifb1*, respectively. Thus, this correlation appears to be generalizable to other mouse models of asthma. Lastly, we evaluated the correlation between MUC5B and BPIFB1 in C57BL/6J mice after sensitization and challenge with allergen or PBS (Figure 3). We found that the expression of both proteins increased significantly as a function of allergen (Figure 3B) in a highly correlated fashion ($r = 0.98$, $P < 0.001$, Figure 3C).

The strong correlations between *Bpifb1* and *Muc5b* at the mRNA and protein level were potentially suggestive of a model in which the Chr 2 QTL leads to variation in *Bpifp1* mRNA (*i.e.*, is an eQTL), which in turn leads to variation in MUC5B protein levels. However, neither *Bpifb1* nor *Muc5b* mRNAs were correlated with MUC5B protein levels in BALF in pre-CC mice (Figure S9), ruling out the proposed eQTL-based model. Rather, our data suggest an alternative mechanism by which genetic variation in *Bpifb1* affects MUC5B, potentially through amino acid variants that affect BPIFB1 protein structure/function. *Musculus*-derived haplotypes (*i.e.*, A/J, NOD/ShiLtJ, and PWK/PhJ) associated with low MUC5B levels share the alternative allele of rs48926494, which encodes a valine to glycine substitution at amino acid 23 of BPIFB1. As such, this SNP is a potential causal variant.

Based on these results, we sought to further evaluate the role of *Bpifb1* using knockout (KO) mice. Using qRT-PCR, we found that *Bpifb1* mRNA was undetectable in KO mice (*i.e.*, Ct values $\geq$ 38). To confirm the loss of expression of BPIFB1 protein in KO mice, we sensitized and challenged WT and KO mice with allergen to induce BPIFB1 protein expression and

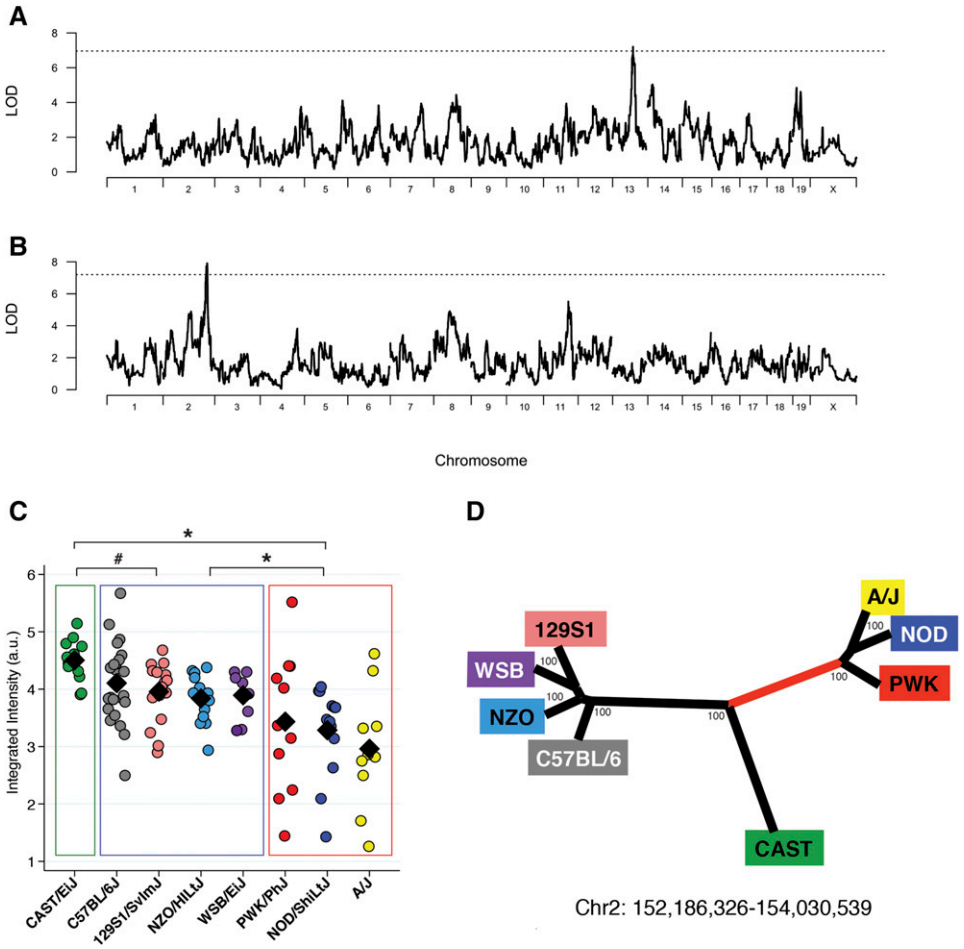

Figure 2 QTL mapping for MUC5AC and MUC5B expression and genetic analysis of MUC5B QTL on Chr 2. Genome scans for (A) MUC5AC and (B) MUC5B reveal distinct *trans*-pQTL. Note that loci encoding Muc5ac and Muc5b are located on chromosome 7. Dashed line represents 95% significance threshold determined by permutation. (C) Allele effects plot of MUC5B. Each pre-CC mouse is plotted by most probable CC founder strain haplotype at Chr 2:154 Mb. Differences between haplotype groupings were determined using Tukey's HSD test indicated as # $P = 0.06$, * $P < 1.0 \times 10^{-4}$. (D) Phylogenetic tree of CC founder strains for Chr 2:152,186,326–154,030,539 bp reconstructed from sequence data shows that there are three distinct alleles derived from *domesticus*, *castaneus*, or *musculus* (red branch) subspecies. Numbers indicate bootstrap support estimates for each branch (out of 100).

did not detect any BPIFB1 protein by immunohistochemistry in KO mice (Figure S10).

We then measured secreted MUC5B in *Bpifb1* WT and KO mice at baseline and after Der p 1 allergen sensitization and challenge with PBS or Der p 1 allergen (Figure 4A). *Bpifb1* KO mice had a threefold increase in MUC5B in BALF under baseline conditions ("naïve" mice), which persisted after sensitization with allergen and challenge with PBS. After sensitization and allergen challenge, MUC5B in BALF increased proportionally in both genotypes. Thus our results indicate independent and additive effects for genotype and allergen treatment (*i.e.*, no genotype by treatment interaction). MUC5AC was not affected by genotype in naïve or PBS-exposed mice and did increase after allergen challenge as expected, though the effect of genotype was not statistically significant (Figure 4B). Results from histological analysis of airway sections for luminal MUC5B content were consistent with BALF MUC5B data (Figure 5). By quantitative morphometry, there was a significant increase in the fraction of intraluminal AB-PAS positive material in *Bpifb1* KO mice and a matching trend for MUC5B (Figure 5C). Conversely, there was a trend toward decreased MUC5B intracellular staining in KO mice (Figure S11). These differences were present despite roughly equal levels of airway inflammation as assessed by differential bronchoalveolar lavage cell counts (Figure S12).

Discussion

We leveraged a series of genetic approaches, including an inbred strain survey, QTL mapping, and functional analysis of a candidate gene, to identify novel regulators of secreted mucin levels in the airways. Our initial studies involving the eight CC founder strains demonstrated that a large fraction of the variation in secreted MUC5AC and MUC5B was attributable to genetic differences between strains. Having established a genetic effect broadly, we turned to QTL mapping in the pre-CC population to identify the specific genetic loci that regulated these two secreted mucins. For the MUC5AC QTL (*Muc5acq1*), we identified potential candidate genes that merit exploration in future studies. *Bpifb1* emerged as a priority candidate gene for the MUC5B QTL (*Muc5bq1*) based on a series of criteria, and we subsequently used *Bpifb1* knockout mice to validate that BPIFB1 is causally related to levels of MUC5B levels in the airways.

Our ability to identify QTL and candidate genes was greatly facilitated by the use of the CC. In particular, the presence of *musculus*-derived alleles in the CC was beneficial in at least two ways. First, *musculus*-derived alleles were associated with lower MUC5B compared to *domesticus*-derived alleles, and because of the intersubspecific introgression (Yang *et al.* 2011) of *musculus* alleles into the genomes of the A/J and

Table 1 Meta-analysis of *Muc5b* and *Bpifb1* gene expression in previous studies

						Differential expression: fold change *vs.* control			
GEO ID	Array platform	Mouse strain	Allergen	Tissue	$n_{treat}/n_{control}$	*Muc5b*	*P*-value	*Bpifb1*	*P*-value
GSE1301[a]	Affymetrix mouse expression 430A array	Not reported	HDM	Whole lung	3/3	2.68	6.73E−05	2.67	8.56E−03
GSE13032[b]	Affymetrix mouse genome 430 2.0 array	A/J and C57BL/6J	OVA	Whole lung	3/3 A/J, 3/3 C57BL/6J	2.66	1.62E−06	1.86	1.28E−06
GSE13382	Agilent-014868 whole mouse genome microarray 4x44K	Not reported	OVA	Whole lung	3/3	1.48	6.87E−02	1.61	3.61E−02
GSE1438	Custom arrays	BALB/c	OVA	Whole lung	5/5	1.55	2.50E−01	1.60	1.10E−01
GSE19223[b]	Affymetrix mouse exon 1.0 ST array	C57BL/6J and BALB/cJ	HDM	Whole lung	11/12 C57BL/6J, 11/14 BALB/cJ	2.21	7.45E−03	3.77	2.11E−03
GSE27066[c]	Affymetrix mouse genome 430 2.0 array	C57BL/6	OVA	Whole lung	4/4	−1.14	4.27E−03	−1.43	5.21E−01
GSE34764	Affymetrix mouse gene 1.0 ST array	C57BL/6	*Alternaria*	Airway epithelia	3/3	1.38	3.05E−02	1.21	1.02E−02
GSE6858	Affymetrix mouse genome 430 2.0 array	BALB/cJ	OVA	Whole lung	4/4	3.89	5.74E−04	2.67	9.64E−03
					Meta-analysis *P*-value		1.48E−13		1.81E−10

[a] Some studies included additional mice with engineered mutations which were not included in this meta-analysis.
[b] Analysis adjusted for strain.
[c] No upregulation of *Muc5ac* detected in this study.

NOD/ShiLtJ founder strains, *musculus*-derived alleles were present in >30% of pre-CC mice. Second, the intersubspecific introgression of *musculus* alleles allowed us to narrow the MUC5B QTL from ∼7.5 to <2 Mb. In terms of quantitative trait gene identification, these features complemented the already enhanced mapping resolution (∼7.5–10 Mb) in the CC compared to typical F2 intercrosses.

While we detected QTL for both MUC5AC and MUC5B that explain a relatively large fraction of phenotypic variance, our data indicate that there are additional QTL to identify. Indeed, the broad-sense heritability estimates based on CC founder strains were high, and the mucin phenotype distributions in the pre-CC population were consistent with a polygenic basis of inheritance. Future studies involving inbred CC strains, with biological replicates, will facilitate the identification of these additional loci. Study designs involving CC line intercrosses (recombinant inbred intercrosses) will also help explore the possibility of epistasis between loci, which was potentially suggested by the relatively large variance of MUC5B levels among pre-CC mice with the *M. m. musculus*-derived haplotype at the Chr 2 QTL.

The existence of distinct QTL for MUC5AC (Chr 13) and MUC5B (Chr 2) indicates that these mucins are under the control of distinct regulatory mechanisms in this model of muco-obstructive airway disease. Additionally, we found that while *Muc5ac* and *Muc5b* mRNA levels in the lung were strongly correlated, ostensibly due to shared transcriptional regulation, neither mRNA was correlated with the levels of secreted MUC5AC or MUC5B. Thus, our results highlight the existence of multiple regulatory pathways controlling the levels of mucins in the airway that could be affected by genetic variation. Given such complexity, it is perhaps not surprising that we identified *trans*-pQTL for both mucins. Consequently, we conclude that distal regulatory features separately control the levels of these two mucins in the airway. These results contrast with what is typically observed for eQTL detected under baseline conditions, which are often located close to genes, but are consistent with a recent study showing that *trans*-pQTL are abundant and often result from protein–protein interactions (Chick *et al.* 2016). As such, we hypothesize that there may be protein–protein interactions underlying the observed relationship between MUC5B and BPIFB1. In addition to the observed strong correlation of MUC5B and BPIFB1 in the airways of allergen-challenged mice, which is suggestive of protein colocalization, BPIFB1 has been shown to have strong interactions with mucins in human bronchial epithelial cell secretions (Radicioni *et al.* 2016), supporting this hypothesis. Furthermore, we can rule out the possibility that variation in *Bpifb1* gene expression underlies the QTL since *Bpifb1* mRNA was not correlated with secreted MUC5B.

Previous studies have provided evidence that BPIFB1 is involved in human airway diseases in which MUC5B is upregulated, including CF (Scheetz *et al.* 2004; Bingle *et al.* 2012) and COPD (Gao *et al.* 2015; Titz *et al.* 2015), as well as IPF (Yang *et al.* 2013; Gao *et al.* 2015). Data for the role of BPIFB1 in asthma are more sparse, but upregulation has been observed in a study involving segmental allergen challenge (Wu *et al.* 2005). There are conflicting reports of how MUC5B levels change in asthma, with reports of increased MUC5B in spontaneous sputum and asthmatic plugs (Groneberg *et al.* 2002; Kirkham *et al.* 2002) contrasting with a report of decreased MUC5B in induced sputum (Lachowicz-Scroggins *et al.* 2016). A very recent report linked differences in the glycosylation status of MUC5B to asthma exacerbations

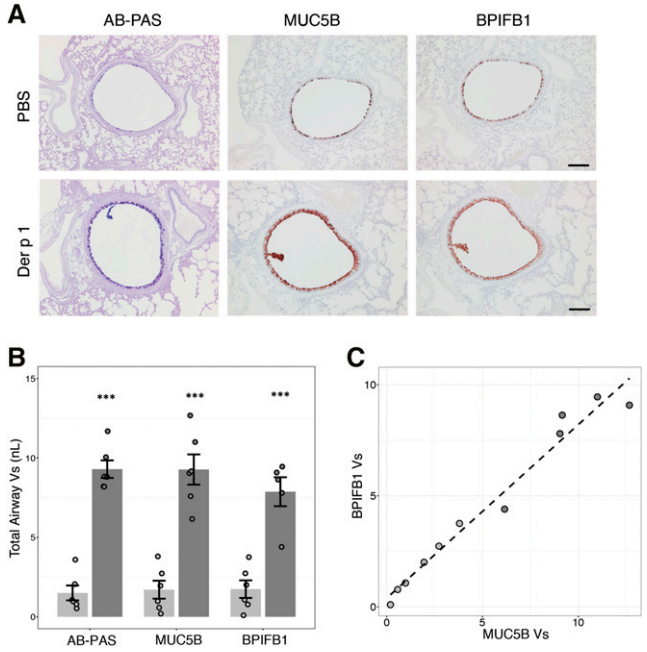

Figure 3 Correlated expression of MUC5B and BPIFB1 in airway epithelium. Der p 1 challenge elicits a correlated increase in MUC5B and BPIFB1 in allergen-sensitized C57BL/6J mice. (A) Large airway staining shows increase in AB-PAS, MUC5B, and BPIFB1 in airway epithelium. Images representative of group based on average quantified signal. Bar, 100 μM. (B) Quantitative morphometry of total airway signal (Vs, mean ± SEM) for PBS control (light gray) or Der p 1 treated (dark gray) mice. n = 5–6 mice/treatment group. Fold change for treatment: 6.2 (AB-PAS), 5.4 (MUC5B), 4.5 (BPIFB1). Significance of two-sided t-tests of treatment indicated as *** $P < 0.001$. (C) Correlation of BPIFB1 and MUC5B total airway signal. Pearson's correlation $r = 0.98$ ($P < 0.001$). Mice challenged with PBS (light gray) or Der p 1 (dark gray).

(Welsh *et al.* 2017). We note that while some studies have reported decreased MUC5B in asthma, antibody-based approaches may be biased due to protein degradation that can occur in the context of inflamed or infected airways (Henderson *et al.* 2014). This does not appear to have biased our results, though, as we detected differences in MUC5B in *Bpifb1* KO mice compared to WT mice under baseline conditions. However, quantification by mass spectrometry will be necessary to measure the levels of MUC5B in a completely unbiased manner. Nevertheless, the body of evidence that MUC5B and BPIFB1 are upregulated in different airway diseases suggests that the association between BPIFB1 and MUC5B may be a common feature of obstructive lung diseases in which MUC5B is elevated. Additionally, we showed that this correlation is present at the level of gene expression in multiple mouse models of asthma and at the protein level upon allergen challenge.

Our data provide clear evidence that BPIFB1 is causally related to MUC5B BALF levels. A rather simple hypothesis to put forth is that BPIFB1 regulates some aspect (s) of MUC5B accumulation in the airways. Indeed, we found evidence of an increase in the fraction of AB-PAS positive luminal substance in *Bpifb1* KO mice whereas the total histological signal in the airways was not affected by genotype. There are at least two possible mechanisms that can explain our observations of higher MUC5B levels in the *Bpifb1* KO strain, namely enhanced secretion into the airways or reduced clearance from the airways. Given that there is already evidence that BPIFB1 and MUC5B interact (Radicioni *et al.* 2016), it will be important to investigate how these interactions affect secretion and/or clearance in future studies. It is interesting to note that in contrast to *Bpifb1* KO mice, which had high MUC5B levels, pre-CC mice with the *musculus* haplotype of *Bpifb1* had lower levels of MUC5B. Thus, the putative causal variant we identified (rs48926494) could impair secretion or facilitate clearance. The use of models that can produce either high or low MUC5B in the airways as a function of *Bpifb1* genotype will be an asset to the investigation of mucus accumulation driven by the interactions between these two proteins.

In summary, through our unbiased, genome-wide search, we have identified distinct loci that regulate MUC5AC and MUC5B levels in the airway after allergen sensitization and challenge. We identified *Bpifb1* as a candidate gene for the MUC5B QTL, and confirmed a causal role for this gene in regulating the levels of secreted MUC5B in the airways, a phenotype that contributes to airway defense and multiple airways diseases.

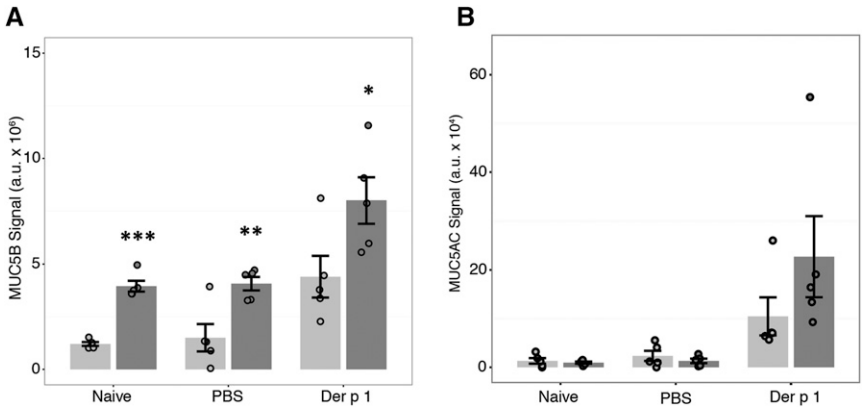

Figure 4 MUC5B expression increased in BALF of *Bpifb1* KO mice independent of treatment. (A) MUC5B and (B) MUC5AC protein in BALF of naïve mice and allergen-sensitized mice challenged with PBS or Der p 1 for WT (light gray) or *Bpifb1* KO (dark gray) mice. Fold change of genotype for MUC5B: 3.2 (naïve), 2.7 (PBS), 1.8 (Der p 1); MUC5AC: 0.93 (naïve), 0.95 (PBS), 2.2 (Der p 1). n = 5 mice/genotype/treatment. Mean ± SEM depicted. Significance of two-sided t-tests of genotype indicated as * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

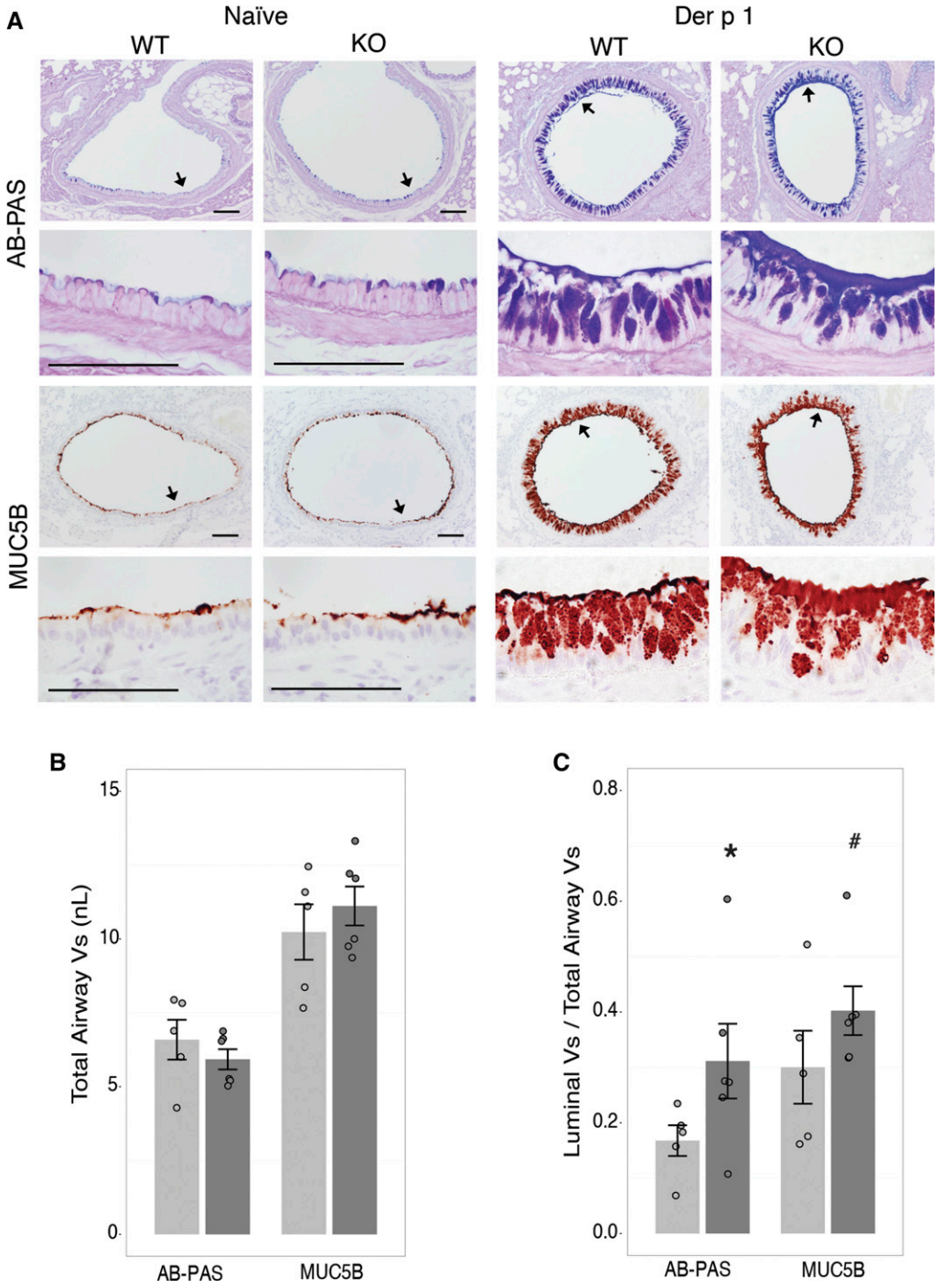

Figure 5 Increased luminal MUC5B in large airways of *Bpifb1* KO mice. (A) AB-PAS staining and MUC5B immunohistochemistry of large airways of naïve or Der p 1 sensitized and challenged WT and *Bpifb1* KO mice. Arrows indicate regions of higher magnification panels. Images representative of group based on average quantified total and luminal signal. n = 5 WT naïve, 3 KO naïve, 5 WT Der p 1, and 6 KO Der p 1. Bar, 50 μM. (B and C) Quantitative morphometry of large airway AB-PAS and MUC5B sections for (B) total airway Vs and (C) luminal Vs expressed as a fraction of total airway Vs, for WT (light gray) and *Bpifb1* KO (dark gray) mice. Mean ± SEM depicted. Significance of one-sided t-tests of genotype indicated as $^{\#}$ $P \leq 0.1$, * $P \leq 0.05$.

Acknowledgments

We are grateful to Kim Burns for her assistance with histology, Dr. Yunjung Kim for assistance with parts of the gene expression analyses, and Holly Rutledge for technical assistance with the early phases of this work. We also thank Colin Bingle for helpful discussions and input. This work was supported in part by the intramural program of the National Human Genome Research Institute, National Institutes of Health (NIH) (ZIA-HG200361), by the US Department of Energy (Office of Biological and Environmental Research), and by NIH grants U01CA134240, U01CA105417, 1R01ES024965, 1R01HL122711, R01HL116228, P01HL108808, P01HL110873, UH2HL123645, P50HL107168, P30ES10126, HL-089708, and P30DK065988 and grants from the Cystic Fibrosis Foundation (BOUCHE15R0 and R026-CR11). L.J.D. was supported in part by a grant from the National Institute of General Medical Sciences under award 5T32 GM007092. We also acknowledge use of a knockout mouse from the Knockout Mouse Project, which was funded by NIH grant 3U01HG004080.

Author contributions: L.J.D., A.L.-B., B.R.G., W.K.O., and S.N.P.K. designed the experiments. L.J.D., A.L.-B., K.M.M., J.M.T., G.C., and S.N.P.K. performed experiments and analyzed data. L.J.D., A.L.-B., B.R.G., W.K.O., R.C.B., and S.N.P.K. wrote the manuscript. All authors edited and reviewed the manuscript.

Literature Cited

Aylor, D. L., W. Valdar, W. Foulds-mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic analysis of complex traits in the emerging collaborative cross. Genome Res. 21: 1213–1222.

Bingle, C. D., K. Wilson, H. Lunn, F. A. Barnes, A. S. High *et al.*, 2010 Human LPLUNC1 is a secreted product of goblet cells and minor glands of the respiratory and upper aerodigestive tracts. Histochem. Cell Biol. 133: 505–515.

Bingle, C. D., L. Bingle, and C. J. Craven, 2011 Distant cousins: genomic and sequence diversity within the BPI fold-containing (BPIF)/PLUNC protein family. Biochem. Soc. Trans. 39: 961–965.

Bingle, L., K. Wilson, M. Musa, B. Araujo, D. Rassl *et al.*, 2012 BPIFB1 (LPLUNC1) is upregulated in cystic fibrosis lung disease. Histochem. Cell Biol. 138: 749–758.

Chesler, E. J., D. R. Miller, L. R. Branstetter, L. D. Galloway, B. L. Jackson *et al.*, 2008 The collaborative cross at Oak Ridge National Laboratory: developing a powerful resource for systems genetics. Mamm. Genome 19: 382–389.

Chick, J. M., S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi *et al.*, 2016 Defining the consequences of genetic variation on a proteome-wide scale. Nature 2801: 1–39.

Chu, H. W., J. Thaikoottathil, J. G. Rino, G. Zhang, Q. Wu *et al.*, 2007 Function and regulation of SPLUNC1 protein in Mycoplasma infection and allergic inflammation. J. Immunol. 179: 3995–4002.

Collaborative Cross Consortium; F. A. Iraqi, M. Mahajne, Y. Salaymah, H. Sandovski, H. Tayem *et al.*, 2012 The genome architecture of the collaborative cross mouse genetic reference population. Genetics 190: 389–401.

Crowley, J. J., V. Zhabotynsky, W. Sun, S. Huang, I. K. Pakatci *et al.*, 2015 Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. Nat. Genet. 47: 353–360.

Davis, C. W., and B. F. Dickey, 2008 Regulated airway goblet cell mucin secretion. Annu. Rev. Physiol. 70: 487–512.

Dijkstra, A. E., J. Smolonska, M. van den Berge, C. Wijmenga, P. Zanen *et al.*, 2014 Susceptibility to chronic mucus hypersecretion, a genome wide association study. PLoS One 9: e91621.

Doherty, T. A., N. Khorram, K. Sugimoto, D. Sheppard, P. Rosenthal *et al.*, 2012 Alternaria induces STAT6-dependent acute airway eosinophilia and epithelial FIZZ1 expression that promotes airway fibrosis and epithelial thickness. J. Immunol. 88: 2622–2629.

Dupuis, J., and D. Siegmund, 1999 Statistical methods for mapping quantitative trait loci from a dense set of markers. Genetics 151: 373–386.

Ehre, C., E. N. Worthington, R. M. Liesman, B. R. Grubb, D. Barbier *et al.*, 2012 Overexpressing mouse model demonstrates the protective role of Muc5ac in the lungs. Proc. Natl. Acad. Sci. USA 109: 16528–16533.

Evans, C. M., K. Kim, M. J. Tuvim, and B. F. Dickey, 2009 Mucus hypersecretion in asthma: causes and effects. Curr. Opin. Pulm. Med. 15: 4–11.

Evans, C. M., D. S. Raclawska, F. Ttofali, D. R. Liptzin, A. A. Fletcher *et al.*, 2015 The polymeric mucin Muc5ac is required for allergic airway hyperreactivity. Nat. Commun. 6: 6281.

Fahy, J. V., and B. F. Dickey, 2010 Airway mucus function and dysfunction. N. Engl. J. Med. 363: 2233–2247.

Festing, M., 1979 *Inbred Strains in Biomedical Research*. Oxford University Press, New York.

Gao, J., S. Ohlmeier, P. Nieminen, T. Toljamo, S. Tiitinen *et al.*, 2015 Elevated sputum BPIFB1 levels in smokers with chronic obstructive pulmonary disease: a longitudinal study. Am. J. Physiol. Lung Cell. Mol. Physiol. 309: L17–L26.

Gralinski, L. E., M. T. Ferris, D. L. Aylor, A. C. Whitmore, R. Green *et al.*, 2015 Genome wide identification of SARS-CoV susceptibility loci using the collaborative cross. PLoS Genet. 11: e1005504.

Groneberg, D. A., P. R. Eynott, S. Lim, T. Oates, R. Wu *et al.*, 2002 Expression of respiratory mucins in fatal status asthmaticus and mild asthma. Histopathology 40: 367–373.

Hallberg, J., A. Dominicus, U. K. Eriksson, M. G. De Verdier, N. L. Pedersen *et al.*, 2008 Interaction between smoking and genetic factors in the development of chronic bronchitis. Am. J. Respir. Crit. Care Med. 177: 486–490.

Harkema, J. R., C. G. Plopper, D. M. Hyde, J. A. St George *et al.*, 1987 Regional differences in quantities of histochemically detectable mucosubstances in nasal, paranasal, and nasopharyngeal epithelium of the bonnet monkey. J. Histochem. Cytochem. 35: 279–286.

Henderson, A. G., C. Ehre, B. Button, L. H. Abdullah, L. H. Cai *et al.*, 2014 Cystic fibrosis airway secretions exhibit mucin hyperconcentration and increased osmotic pressure. J. Clin. Invest. 124: 3047–3060.

Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. Nature 477: 289–294.

Kelada, S. N. P., 2016 Plethysmography phenotype QTL in mice before and after allergen sensitization and challenge. G3 6: 2857–2865.

Kelada, S. N. P., M. S. Wilson, U. Tavarez, K. Kubalanza, B. Borate *et al.*, 2011 Strain-dependent genomic factors affect allergen-induced airway hyperresponsiveness in mice. Am. J. Respir. Cell Mol. Biol. 45: 817–824.

Kelada, S. N. P., D. L. Aylor, B. C. E. Peck, J. F. Ryan, U. Tavarez *et al.*, 2012 Genetic analysis of hematological parameters in incipient lines of the collaborative cross. G3 2: 157–165.

Kelada, S. N. P., D. E. Carpenter, D. L. Aylor, P. Chines, H. Rutledge *et al.*, 2014 Integrative genetic analysis of allergic inflammation in the murine lung. Am. J. Respir. Cell Mol. Biol. 51: 436–445.

Kirkham, S., J. K. Sheehan, D. Knight, P. S. Richardson, and D. J. Thornton, 2002 Heterogeneity of airways mucus: variations in the amounts and glycoforms of the major oligomeric mucins MUC5AC and MUC5B. Biochem. J. 361: 537–546.

Lachowicz-Scroggins, M., S. Yuan, S. Kerr, E. Dunican, M. Yu *et al.*, 2016 Abnormalities in MUC5AC and MUC5B protein in airway mucus in asthma. Am. J. Respir. Crit. Care Med. 194: 1296–1300.

Lee, J., M. H. Cho, C. P. Hersh, M. L. McDonald, J. D. Crapo *et al.*, 2014 Genetic susceptibility for chronic bronchitis in chronic obstructive pulmonary disease. Respir. Res. 15: 113.

Livraghi-Butrico, A., B. R. Grubb, K. J. Wilkinson, A. S. Volmer, K. A. Burns *et al.*, 2017 Contribution of mucus concentration and secreted mucins Muc5ac and Muc5b to the pathogenesis of muco-obstructive lung disease. Mucosal Immunol. 10: 395–407.

Mosedale, M., Y. Kim, W. J. Brock, S. E. Roth, T. Wiltshire *et al.*, 2017 Candidate risk factors and mechanisms for tolvaptan-induced liver injury are identified using a collaborative cross approach. Toxicol. Sci. 156: 438–454.

Ordoñez, C. L., R. Khashayar, H. H. Wong, R. Ferrando, R. Wu *et al.*, 2001 Mild and moderate asthma is associated with airway goblet cell hyperplasia and abnormalities in mucin gene expression. Am. J. Respir. Crit. Care Med. 163: 517–523.

Radicioni, G., R. Cao, J. Carpenter, A. A. Ford, T. T. Wang *et al.*, 2016 The innate immune properties of airway mucosal surfaces are regulated by dynamic interactions between mucins and interacting proteins: the mucin interactome. Mucosal Immunol. 9: 1442–1454.

Rose, M. C., and J. A. Voynow, 2006 Respiratory tract mucin genes and mucin glycoproteins in health and disease. Physiol. Rev. 86: 245–278.

Roy, M. G., A. Livraghi-Butrico, A. A. Fletcher, M. M. McElwee, S. E. Evans *et al.*, 2013 Muc5b is required for airway defence. Nature 505: 412–416.

Rutledge, H., D. L. L. Aylor, D. E. E. Carpenter, B. C. C. Peck, P. Chines *et al.*, 2014 Genetic regulation of Zfp30, CXCL1, and neutrophilic inflammation in murine lung. Genetics 198: 735–745.

Scheetz, T. E., J. Zabner, M. J. Welsh, J. Coco, M. D. F. Eyestone *et al.*, 2004 Large-scale gene discovery in human airway epithelia reveals novel transcripts. Physiol. Genomics 17: 69–77.

Seibold, M. A., A. L. Wise, M. C. Speer, M. P. Steele, K. K. Brown *et al.*, 2011 A common MUC5B promoter polymorphism and pulmonary fibrosis. N. Engl. J. Med. 364: 1503–1512.

Sheehan, J. K., P. S. Richardson, D. C. Fung, M. Howard, and D. J. Thornton, 1995 Analysis of respiratory mucus glycoproteins in asthma: a detailed study from a patient who died in status asthmaticus. Am. J. Respir. Cell Mol. Biol. 13: 748–756.

Silverman, E. K., H. A. Chapman, J. M. Drazen, S. T. Weiss, B. Rosner *et al.*, 1998 Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Am. J. Respir. Crit. Care Med. 157: 1770–1778.

Skarnes, W. C., B. Rosen, A. P. West, M. Koutsourakis, W. Bushell *et al.*, 2011 A conditional knockout resource for the genome-wide study of mouse gene function. Nature 474: 337–342.

Stamatakis, A., 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.

Thai, P., A. Loukoianov, S. Wachi, and R. Wu, 2008 Regulation of airway mucin gene expression. Annu. Rev. Physiol. 70: 405–429.

Titz, B., A. Sewer, T. Schneider, A. Elamin, F. Martin *et al.*, 2015 Alterations in the sputum proteome and transcriptome in smokers and early-stage COPD subjects. J. Proteomics 128: 306–320.

Valdar, W., C. C. Holmes, R. Mott, and J. Flint, 2009 Mapping in structured populations by resample model averaging. Genetics 182: 1263–1277.

Venkatratnam, A., S. Furuya, O. Kosyk, A. Gold, W. Bodnar *et al.*, 2017 Collaborative cross mouse population enables refinements to characterization of the variability in toxicokinetics of trichloroethylene and provides genetic evidence for the role of PPAR pathway in its oxidative metabolism. Toxicol. Sci. DOI: 10.1093/toxsci/kfx065

Viegi, G., L. Carrozzi, F. Di Pede, S. Baldacci, M. Pedreschi *et al.*, 1994 Risk factors for chronic obstructive pulmonary disease in a north Italian rural area. Eur. J. Epidemiol. 10: 725–731.

Wang, J., F. de Villena, and L. McMillan, 2012 Comparative analysis and visualization of multiple collinear genomes. BMC Bioinformatics 13: S13.

Welsh, K. G., K. Rousseau, G. Fisher, L. R. Bonser, P. Bradding *et al.*, 2017 MUC5AC and a glycosylated variant of MUC5B alter mucin composition in children with acute asthma. Chest DOI: 10.1016/j.chest.2017.07.001.

Wu, J., M. Kobayashi, E. A. Sousa, W. Lui, J. Cai *et al.*, 2005 Differential proteomic analysis of bronchoalveolar lavage fluid in asthmatics following segmental antigen challenge. Mol. Cell. Proteomics 4: 1251–1264.

Wu, T., J. Huang, P. J. Moore, M. S. Little, W. G. Walton *et al.*, 2017 Identification of BPIFA1/SPLUNC1 as an epithelium-derived smooth muscle relaxing factor. Nat. Commun. 8: 14118.

Yang, H., Y. Ding, L. N. Hutchins, J. Szatkiewicz, T. A. Bell *et al.*, 2009 A customized and versatile high-density genotyping array for the mouse. Nat. Methods 6: 663–666.

Yang, H., J. R. Wang, J. P. Didion, R. J. Buus, T. A. Bell *et al.*, 2011 Subspecific origin and haplotype diversity in the laboratory mouse. Nat. Genet. 43: 648–655.

Yang, I. V, C. D. Coldren, S. M. Leach, M. A. Seibold, E. Murphy *et al.*, 2013 Expression of cilium-associated genes defines novel molecular subtypes of idiopathic pulmonary fibrosis. Thorax 68: 1114–1121.

Communicating editor: D. Threadgill